

# Extremely Correlated Fermi Liquids

Compilation of Reprints \*

With Comments

B. Sriram Shastry

Physics Department  
University of California Santa Cruz, CA 95064

November 2, 2021

## Preface

This volume contains a compilation of papers on the *Extremely Strongly Correlated Fermi Liquid* (ECFL) theory. The theory provides a systematic framework for calculating the properties of a fermions on a lattice, when the interaction strength is so large that standard perturbation techniques based on Feynman diagrams, are rendered invalid. The models solved in the ECFL theory allow the local interaction to diverge, thereby eliminating or projecting out a finite fraction of states in the Hilbert space. This elimination turns the usual fermions into non-canonical *projected* fermions. These satisfy an anti-commuting non-canonical algebra, whereby Wick's theorem is irretrievably lost. The theory now uses the Schwinger-Tomonaga method of source fields to set up exact non-linear functional differential equations for Greens functions. Solving these equations is challenging, and several innovations are further necessary before one can find practical solutions that are testable by experiments. This package of essential steps forms the core of the ECFL formalism.

After the [initial paper](#) in 2011, we‡ have worked on finding explicit solutions and comparing with a few known, but rare exact solutions, found by other means. These include the  $d = 0$  Anderson impurity model (using the numerical renormalization group), the  $d = \infty$ ,  $U = \infty$  Hubbard model (using the dynamical mean field theory), the 1 dimensional  $t$ - $J$  model (using the density matrix renormalization group), and also with high temperature series expansions. After this extensive testing, we have recently applied the theory to the 2-dimensional  $t$ - $t'$ - $J$  model, and calculated several measurables in the density range relevant to cuprate superconductors. Recent papers in 2020-21 address two central issues of cuprate superconductivity, the [quasi-linear resistivity](#) and the [mechanism of superconductivity](#). At this stage it appears that a compilation of the (several) published papers from the ECFL theory providing some additional context could be helpful. Towards that goal, this compilation contains a topic-wise breakup of the papers, and also brief comments on each paper. Extensive hyperlinks are provided to enable easy paper-hopping.

### Co-authors‡

My grateful thanks to colleagues whose works are included here.

- *Graduate Students and Postdocs:* Michael Arciniaga, Wenxin Ding, Daniel Hansen, Peizhi Mai, Kazue Matsuyama, Edward Perepelitsky.
- *Senior Collaborators:* Professors Antoine Georges (Collège de France, Paris), G-H Gweon (UC Santa Cruz), Alex Hewson (Imperial College, London), Rok Žitko and Jernej Mravlje (Jožef Stefan Institute, Ljubljana), Ehsan Khatami (SJSU), Marcos Rigol (Penn State), Steven R White (UC Irvine) and H R Krishnamurthy (IISc, Bangalore).

# Contents

I	Topics (With Comments) . . . . .	4
I.1	Basic Formalism . . . . .	4
I.2	ECFL- Comparisons: Numerical Renormalization Group - Anderson Impurity Model and Dynamical Mean Field theory . . . . .	5
I.3	Resistivity in $d = \infty$ . . . . .	6
I.4	ECFL Solutions in $d = 1$ at $\mathcal{O}(\lambda^2)$ and preliminary solution in $d = 2$	7
I.5	ECFL Solutions in $d = 2$ at $\mathcal{O}(\lambda^2)$ and comparison with experiments on Raman Scattering and electrical conductivities . . . . .	8
I.6	Angle Resolved Photoemission Spectroscopy (ARPES) related . . . . .	9
I.7	High T expansion and ECFL comparison . . . . .	9
I.8	Quasi-linear resistivity and d-wave superconductivity in cuprates . . . . .	10
II	Included papers (most recent first) . . . . .	11

# I Topics (With Comments)

## I.1 Basic Formalism

### 1. Extremely Correlated Fermi Liquids

B. S. Shastry, arXiv:1102.2858 (2011), Phys. Rev. Letts. **107**, 056403 (2011). ...Page(14)

§This paper lays the foundations of the ECFL theory. A core issue in formulating a non canonical theory is to generalize Dyson's self energy of the canonical theory. This requires one to identify an object  $\Phi$  that can be tucked into the denominator of the Greens function. A subtle and important requirement is that  $\Phi$  must spawn a sensible (i.e. non-growing) vertex function (upon differentiating by a bosonic source). The vertex also defines the self energy through a closure relation. Thus a self-consistent loop of equations is set up, to be solved by a suitable iterative procedure. An earlier paper ([Extremely Correlated Quantum Liquids](#), B Sriram Shastry, Phys. Rev. **B 81**, 045121 (2010)) discusses this issue and proposes an alternate solution. However that solution leads inevitably to a fermi surface differing in volume from that of the free fermi gas, thereby violating the Luttinger-Ward theorem. While this may yet be relevant vey close to the Mott insulating state, one would like to retain the volume at lower particle densities.

This work finds another possible route for identifying the self energy, while maintaining continuity with the fermi gas, and thus avoiding the volume violation. Here the basic factorization of the physical Greens function is introduced, into an auxiliary Greens function  $g$  and a [comparison](#) function  $\mu$  through  $\mathcal{G}(k, \omega) = g(k, \omega)\mu(k, \omega)$ . The two Greens functions  $\mathcal{G}$  and  $g$  describe the fully interacting and a related canonical theory, with identical particle number sum-rules. This decomposition leads to a sensible self energy for the auxiliary Greens function  $g(k, \omega)$ , and  $\mu(k, \omega)$  is a second self energy, but now in the *numerator* of the physical  $\mathcal{G}$ . An iterative scheme is proposed here by introducing a parameter  $\lambda \in [0, 1]$  connecting the free fermi gas with the fully interacting problem. Exact formal equations for the two self energies are given and systematically expanded in  $\lambda$  to  $\mathcal{O}(\lambda^2)$ .

The paper also contains a phenomenological model for the two self energies which captures many features of the spectral line shapes.

### 2. Anatomy of the Self Energy

B. S. Shastry, arXiv:1104.2633; Phys. Rev. B 84, 165112 (2011); Phys. Rev. B 86, 079911(E) (2012). ...Page(24)

§The non-standard Greens function of the extremely correlated electrons proposed in the [paper\(1\)](#), is structurally very different from the Dyson construction for canonical fermions. This paper explores the difference, its consequences for the line shapes, and further refines the phenomenological spectral function proposed above.

This paper also contains a discussion of the Mori formalism. It can be used to define a Dyson-Mori self energy, if the moments of the spectral function are available from an independent calculation.

### 3. Extremely Correlated Fermi Liquids: The Formalism

B. S. Shastry, arXiv:1207.6826 (2012); Phys. Rev. **B 87**, 125124 (2013). ...Page(43)

§This paper introduces two *shift identities* for the  $t$ - $J$  model, and the use of a second Lagrange multiplier, required in addition to the thermodynamic chemical potential, to fulfill the number sum-rule on the physical as well as the auxiliary Greens functions. The paper also details a systematic procedure for computing the self energies and vertices to high order using functional derivatives on the lower order ones.

4. [Theory of extreme correlations using canonical Fermions and path integrals](#)  
B. S. Shastry, arXiv:1312.1892 (2013), Ann. Phys. **343**, 164 (2014). (Erratum) **373**, 717 (2016). ...Page(132)

§This paper shows the analogy between the  $\lambda$  expansion of the correlated electrons, and the large spin expansion using the Dyson-Maleev representation for spins. It also has a formulation of the theory in terms of novel  $\lambda$  fermions satisfying a  $\lambda$  dependent (deformed) fermi algebra, and also a path integral representation in terms of canonical fermions. This paper also simplifies the treatment using the *minimal equations of motion*, discarding extra symmetrizing terms employed earlier (in the [paper\(1\)](#) and in [paper \(3\)](#)). Later work has used only the minimal equations.

5. [Diagrammatic  \$\lambda\$  series for extremely correlated Fermi liquids](#)  
E. Perepelitsky and B. S. Shastry, arXiv:1410.5174, Ann. Phys. **357**, 1 (2015).  
...Page(183)

§This paper contains a detailed diagrammatic expansion of the Schwinger equations of motion in terms of the  $\lambda$  parameter. The rules for diagrams enable one to write down a complete set of distinct diagrams to any order, without recourse to functional differentiation as in the formalism [paper \(3\)](#). The diagrams contain a Feynman diagram subset, extended by adding certain additional (Schwinger) diagrams.

6. [Fermi Surface Volume of Interacting Systems](#)  
B. S. Shastry, arXiv:1808.00405v4; Annals of Physics **405**, 155 (2019). ...Page(315)

§This work provides an alternate proof of the theorem asserting the invariance of the Fermi surface volume at  $T=0$  upon turning on interactions, using the novel concept of *isothermal continuity*. For Fermi liquids this theorem is an important ingredient of Landau's argument, who used adiabatic continuity. It was demonstrated in the well known work of Luttinger and Ward, who assumed that perturbation theory is convergent. The new proof, motivated by [the results of ECFL\(6\)](#), extends the theorem to arbitrarily strong interactions, 1-d non-Fermi liquid systems, and to singlet superconductors. The sum-rule is expressed in terms of photoemission accessible spectral functions, and its rich  $T$  dependence is illustrated in [paper\(18\)](#).

## I.2 ECFL- Comparisons: Numerical Renormalization Group - Anderson Impurity Model and Dynamical Mean Field theory

7. [Extremely Correlated Fermi Liquid study of the  \$U = \infty\$  Anderson Impurity Model](#)  
B. S. Shastry, E. Perepelitsky and A. C. Hewson, arXiv:1307.3492 [cond-mat.str-el], Phys. Rev. **B 88**, 205108 (2013). ...Page(86)

§In this work the single impurity Anderson model is solved by the ECFL method, using the  $\lambda$  expansion to  $\mathcal{O}(\lambda^2)$ . For impurity density  $n_d \lesssim 0.7$  and the spectral properties

are shown to compare well with the numerically exact results from the numerical renormalization group method of Wilson. For higher densities  $n_d \gtrsim 0.7$  good agreement is also found provided the frequencies are scaled by a corrected quasiparticle weight  $Z$ .

8. **Reversal of particle-hole scattering-rate asymmetry in Anderson impurity model**

Rok Žitko, H. R. Krishnamurthy and B. Sriram Shastry, arXiv:1807.11343. Phys. Rev. B **98**, 161121 (R) (2018). ...Page(310)

§In [this work](#), motivated by a suggestion from the strong coupling solution in [paper \(7\)](#), the single impurity Anderson model is shown to have a novel separatrix in the density- $U$  plane, demarcating regions of weak and strong correlations. These regions are distinguished by opposite signs of the  $\omega^3$  term correction to  $\Im m \Sigma(\omega)$ , beyond the leading  $\omega^2$  Fermi liquid behavior.

9. **Extremely correlated Fermi liquid theory meets Dynamical mean-field theory: Analytical insights into the doping-driven Mott transition**

R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges and B. S. Shastry, arXiv:1309.5284 (2013), Phys. Rev. B **88**, 235132 (2013). ...Page(95)

§The ECFL spectral results for  $d = \infty$  at  $\mathcal{O}(\lambda^2)$ , are compared with the exact solution of the Hubbard model in the same limit, with varying values of density and the interaction  $U$ . For  $U \gtrsim 2D$  -the bandwidth, the two theories give a similar behavior at low frequencies, with a strong particle hole-asymmetry in the imaginary part of self energy. This benchmarking shows that the  $\mathcal{O}(\lambda^2)$  ECFL theory gives spectral functions that extend to higher energies compared to the exact results. The value of the quasiparticle weight  $Z$  is also over-estimated for densities  $n \gtrsim 0.75$ , similar to the impurity model results in [paper \(7\)](#). Once again, scaling the frequency by a corrected  $Z$  improves the agreement at high densities considerably.

10. **ECFL in the limit of infinite dimensions**

E. Perepelitsky and B. S. Shastry, arXiv:1309.5373 (2013), Annals of Physics **338**, 283-301 (2013). ...Page(113)

The ECFL equations are analyzed in the limit of  $d \rightarrow \infty$ . The resulting Dysonian self energy, found by inverting the physical Greens function, is shown to be momentum independent, as in the exact result. An expansion in  $\lambda$  is developed, giving the equations used in the DMFT comparison [paper\(9\)](#).

### I.3 Resistivity in $d = \infty$

11. **Low energy physics of the  $t$ - $J$  model in  $d = \infty$  using Extremely Correlated Fermi Liquid theory: Cutoff Second Order Equation**

B. S. Shastry and E. Perepelitsky, arXiv:1605.08213. Phys. Rev. B **94**, 045138 (2016). ...Page(222)

§Explores the origin and provides one convenient prescription for managing the high energy tails that the  $\mathcal{O}(\lambda^2)$  equations develop in the two (distinct limits) of high T at any density, and at high density  $n \rightarrow 1$  at low T. A cutoff procedure using a Tukey window confines the spectrum to roughly twice the (bare) bandwidth. Further the physical Greens function for the lower Hubbard band is required to satisfy the

exact normalization  $1 - \frac{n}{2}$  (without  $\lambda$  corrections as in earlier work), and argued to be essential for getting the right quasiparticle weight. With these ingredients, the resulting quasiparticle weight  $Z$  is close to the exact value from DMFT to fairly high  $n \lesssim 0.9$ .

12. [A Strange Metal from Gutzwiller correlations in infinite dimensions](#)

Wenxin Ding, Rok Žitko, Peizhi Mai, Edward Perepelitsky and B. Sriram Shastry, arXiv:1703.02206v2; Phys. Rev. **B 96**, 054114 (2017). DOI: 10.1103/PhysRevB.96.054114

...Page(256)

§The cutoff introduced in the above [paper \(11\)](#) is used to compute the resistivity, heat capacity and entropy at low T, and compared against exact results from DMFT. The different regimes of resistivity are reviewed, these include the fermi liquid, the strange metal, the bad metal and the high T metal, with three distinct crossovers between them. All these regimes follow from the Gutzwiller (i.e. large U) limit.

13. [A Strange Metal from Gutzwiller correlations in infinite dimensions II: Transverse Transport, Optical Response and Rise of Two Relaxation Rates](#)

Wenxin Ding, Rok Žitko, and B. Sriram Shastry, arXiv:1705.01914; Phys. Rev **B 96**, 115153 (2017).

...Page(284)

§This examines the T dependence of the Hall constant and the cotangent Hall angle, as well as the optical conductivity, using the same methods as in the [paper \(12\)](#) above. A noticeable *kink*, i.e. change in slope of the cotangent Hall angle versus  $T^2$  is noted at the effective fermi liquid scale, and the transverse time scale is extracted from the frequency dependent conductivities. Such a kink seems present in many experiments.

#### I.4 ECFL Solutions in $d = 1$ at $\mathcal{O}(\lambda^2)$ and preliminary solution in $d = 2$

14. [Extremely Correlated Fermi Liquids: Self consistent solution of the second order theory](#)

D. Hansen and B. S. Shastry, arXiv:1211.0594, (2012), Phys. Rev. **B 87**, 245101 (2013).

...Page(66)

§The first set of numerical results from the theory for the  $t$ - $J$  model in 2-d. The equations for the self energies are truncated to  $\mathcal{O}(\lambda^2)$ , and explicit results for electronic densities in the highly overdoped regime of cuprates  $n \leq 0.75$ , are found. The quasiparticle weight is larger than expected for  $n \gtrsim 0.75$ , since the normalization condition for the Greens function is different from the one later emphasized in [paper\(11\)](#), and employed in recent work.

15. [The  \$t\$ - \$t'\$ - \$J\$  model in one dimension using extremely correlated Fermi liquid theory and time dependent density matrix renormalization group](#)

Peizhi Mai, Steven R. White and B. Sriram Shastry, arXiv:1712.05396. ...Page(296)

§An application of the equations reported in the [paper \(11\)](#) and in [paper \(16\)](#) to the case of the 1-d  $t$ - $t'$ - $J$  model are made and the results compared with those from the density matrix renormalization group (DMRG). The physics of anomalous exponents, spin-charge separation, and the (non-fermi) Tomonaga-Luttinger liquid state brought about by quantum fluctuations in 1-d are captured by the ECFL equations, and also the DMRG solution. The DMRG and ECFL theories yield a self energy with strong k dependence of a similar nature.

## I.5 ECFL Solutions in $d = 2$ at $\mathcal{O}(\lambda^2)$ and comparison with experiments on Raman Scattering and electrical conductivities

16. **Extremely Correlated Fermi Liquid theory of the  $t$ - $J$  model in 2 dimensions: Low energy properties**  
 B Sriram Shastry and Peizhi Mai, arXiv:1703.08142, N. Jour. Phys. **20** 013027 (2018).  
 ...Page(269)

§An application of the equations reported in the [paper \(11\)](#), to the case of the 2-d  $t$ - $J$  model is made, at experimentally interesting densities  $.8 \leq n \leq .9$ , after adding the  $J$  terms. The model contains the second neighbor hopping  $t'$  as well. The quasiparticle weight is very sensitive to the sign of  $t'/t$ , and typically very small. The effective fermi temperature is calculated to be 2 or 3 orders of magnitude smaller than the bandwidth. The resistivity changes from convex  $T^2$  to linear- $T$ , and further to concave upwards, upon changing the sign of  $t'$ . This variation has the possibility of explaining the difference between electron doped and hole doped cuprate systems. The Hall number, cotangent Hall angle are calculated.

17. **Non-resonant Raman Scattering in Extremely Correlated Fermi Liquids**  
 Peizhi Mai and B. Sriram Shastry, arXiv:1805.09935, Phys. Rev. **B 98**, 115101 (2018).  
 ...Page(301)

§A detailed study of the non-resonant Raman scattering in a 2-d  $t$ - $J$  model is presented, using the ECFL equations developed in [paper \(16\)](#). A quartet of closely related variables are presented, three from standard polarization geometries and the optical conductivity. Together they reveal the role played by the model parameters, and the importance of using correct vertices in narrow band metals.

18. **Extremely correlated fermi liquid of  $t$ - $J$  model in two dimensions**  
 Peizhi Mai and B. Sriram Shastry, arXiv:1808.09788, Phys. Rev. B **98**, 205106 (2018).  
 ...Page(336)

§This work is a continuation of [paper \(16\)](#), and reports detailed calculations of the spectral functions, EDC and MDC dispersion relations, Hall number and a Fermi surface sum-rule reported in [paper \(6\)](#).

19. **Theory of anisotropic elasto-resistivity of 2-D extremely strongly correlated metals**  
 Michael Arciniaga, Peizhi Mai and B Sriram Shastry, arXiv:1909.06471; Phys. Rev. **B101**,245149 (2020).  
 ...Page(356)

§The effect of anisotropic strain on resistivities is calculated from equations generalizing those in [paper \(16\)](#) to include the effect of anisotropic strain. The elasto-resistivity tensor is calculated and predictions for the magnitude of this effect are made for cuprate superconductors. The anisotropic response of a few other measurables (the local density of states and optical sum-rule) are calculated.

20. **Aspects of the Normal State Resistivity of Cuprate Superconductors**  
 B. Sriram Shastry and Peizhi Mai, arXiv:1911.09119; Phys. Rev. **B101**,115121(2020).  
 ...Page(373)

§The electrical resistivity of three families of single layer cuprate superconductors are compared with the ECFL theory, as a function of temperature within a broad range of metallic densities. An almost perfect matching with data is possible.

## I.6 Angle Resolved Photoemission Spectroscopy (ARPES) related

21. [Extremely Correlated Fermi Liquid Description of Normal State ARPES in Cuprates](#)  
G.-H. Gweon, B. S. Shastry and G. D. Gu, arXiv:1104.2631 (2011), Phys. Rev. Letts. 107, 056404 (2011). ...Page(19)

§The ECFL spectral lines are compared with experiments on several cuprate materials. The ARPES experiments done at different incident photon energies give quite different looking spectra, this work shows that varying an elastic scattering parameter together with the line shapes from ECFL theory fit experiments very closely. This paper observes that a low energy kink  $\omega \sim 60\text{meV}$  present in the ECFL spectrum, is in close accord with experimental kinks.

22. [Dynamical Particle Hole Asymmetry in Cuprate Superconductors](#)  
B. S. Shastry, arXiv:1110.1032 (2011), Phys. Rev. Letts. **109**, 067004 (2012). ...Page(38)

§This work emphasizes the occurrence of a distinct skew or asymmetry of ECFL spectral lines, also seen in experiments. This implies the absence of particle hole symmetry in the self energy at even modestly frequencies  $\hbar\omega \sim 40\text{meV}$ . Further tests for establishing this asymmetry in experiments is proposed.

23. [Origin of Kinks in Energy Dispersion of Strongly Correlated Matter](#)  
Kazue Matsuyama, Edward Perepelitsky and B Sriram Shastry, arXiv:1610.08079, Phys. Rev. **B 95**, 165435 (2017). ...Page(236)

§This work provides a review of the kinks in ARPES, and contrasts two competing view points for explaining it, along with precise tests for each of them. One view of the kinks locates some bosonic degree of freedom at the kink energy, and leads to specific predictions for the energy dispersion of the electrons. Another view, initiated in [paper\(21\)](#) and developed further in [paper\(2\)](#), and [paper\(4\)](#) ascribes these to extremely strong correlations, and comes with another set of predictions. Considerable data is analyzed using the theory and several testable predictions made.

## I.7 High T expansion and ECFL comparison

24. [Electronic spectral properties of the two-dimensional infinite-U Hubbard model](#)  
E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, B. S. Shastry, arXiv:1303.2657 [cond-mat.str-el] (2013), Phys. Rev. **B 87**, 161120 (R) (2013). ...Page(81)

§The 2-d  $U = \infty$  Hubbard model electron Greens function is studied here. Several types of (low order) moments of the Greens functions are calculated here using high Temperature cumulant expansion. In this formalism non canonical fermions can be treated in a straightforward but laborious fashion. The comparison with ECFL spectra is found to be possible at low densities  $n \lesssim 0.75$ , and not too low T, and several useful results are compiled.

25. [Linked-Cluster Expansion for the Green's function of the Infinite-U Hubbard Model](#)  
E. Khatami, E. Perepelitsky, M. Rigol and B. S. Shastry, arXiv: 1310.8029 (2013), Phys. Rev. **E 89**, 063301 (2014). ...Page(170)

§The high T series for the  $U = \infty$  Hubbard model Greens function is calculated to eighth order in  $t$  using a linked cluster expansion. The electronic spectral function,

the thermopower and quasiparticle weight are calculated and compared to the ECFL results. The electron Greens function is tabulated up to  $\mathcal{O}(t^4)$  in k-space, and (the long) higher order terms up to  $\mathcal{O}(t^8)$  are available on request.

## I.8 Quasi-linear resistivity and d-wave superconductivity in cuprates

### 26. Aspects of the normal state resistivity of cuprate superconductors,

B. Sriram Shastry and Peizhi Mai, arXiv:1911.09119; Phys. Rev. B **101**, 115121 (2020).  
...Page(373)

§Currently available data on density and temperature dependence of resistivity of single layer cuprates, is compared with the results of ECFL theory. Calculations are performed to  $\mathcal{O}(\lambda^2)$ , with only one unknown parameter for each family of compounds, determined by the fits. Experiments on electron doped as well as hole doped materials are contrasted successfully with this theory.

### 27. Extremely correlated superconductors,

B. Sriram Shastry, arXiv:2102.08395; Annals of Physics **434**, 168614 (2021)....Page(384)

§Superconductivity in the t-J model is studied by generalizing the Gor'kov equations to include very strong correlations. Systematic equations for the superconducting instability are formulated within the  $\lambda$  expansion. The leading order equation is studied using a phenomenological spectral function for the metallic state. A superconducting phase appears away from half filling, and persists for a range of densities. A simple formula for the  $T_c$  is obtained, where the interplay between Mott-Hubbard localization, Cooper pairing due to the exchange energy J, and its dependence on the band parameters are made explicit.

## II Included papers (most recent first)

- **Extremely correlated superconductors**,  
B. Sriram Shastry , arXiv:2102.08395; Annals of Physics **434**, 168614 (2021).  
<https://doi.org/10.1016/j.aop.2021.168614> ...Page(384)
- **Aspects of the normal state resistivity of cuprate superconductors**,  
B. Sriram Shastry and Peizhi Mai, arXiv:1911.09119; Phys. Rev. B**101**, 115121 (2020).  
DOI: 10.1103/PhysRevB.101.115121 ...Page(373)
- **Theory of anisotropic elastoresistivity of two-dimensional extremely correlated metals**,  
Michael Arciniaga, Peizhi Mai and B. Sriram Shastry, arXiv:1909.06471; Phys. Rev. B**98**, 205106 (2018).  
DOI: 10.1103/PhysRevB.101.245149 ...Page(356)
- **Extremely correlated fermi liquid of  $t$ - $J$  model in two dimensions**,  
Peizhi Mai and B. Sriram Shastry, arXiv:1808.09788; Phys. Rev. B**98**, 205106 (2018).  
DOI: 10.1103/PhysRevB.98.205106 ...Page(336)
- **Fermi Surface Volume of Interacting Systems**,  
B. Sriram Shastry, arXiv:1808.00405v4; Annals of Physics **405** 155 (2019).  
DOI:<https://doi.org/10.1016/j.aop.2019.03.016> ...Page(315)
- **Reversal of particle-hole scattering-rate asymmetry in Anderson impurity model**,  
Rok Žitko, H. R. Krishnamurthy and B. Sriram Shastry, arXiv:1807.11343, Phys. Rev. B**98**, 161121(R) (2018).  
DOI: 10.1103/PhysRevB.98.161121  
...Page(310)
- **Non-resonant Raman Scattering in Extremely Correlated Fermi Liquids**,  
Peizhi Mai and B. Sriram Shastry, arXiv:1805.09935, Phys. Rev. B**98**, 115101 (2018).  
DOI: 10.1103/PhysRevB.98.115101 ...Page(301)
- **The  $t$ - $t'$ - $J$  model in one dimension using extremely correlated Fermi liquid theory and time dependent density matrix renormalization group**,  
Peizhi Mai, Steven R. White and B. Sriram Shastry, arXiv:1712.05396; Phys. Rev. B **98**, 035108 (2018).  
DOI: 10.1103/PhysRevB.98.035108 ...Page(296)
- **A Strange Metal from Gutzwiller correlations in infinite dimensions II: Transverse Transport, Optical Response and Rise of Two Relaxation Rates**,  
Wenxin Ding, Rok Žitko, and B. Sriram Shastry, arXiv:1705.01914; Phys. Rev B **96**, 115153 (2017).  
DOI: 10.1103/PhysRevB.96.115153  
...Page(284)
- **Extremely Correlated Fermi Liquid theory of the  $t$ - $J$  model in 2 dimensions: Low energy properties**,  
B Sriram Shastry and Peizhi Mai, arXiv:1703.08142, New Jour. Phys. **20** 013027 (2018).  
DOI:<https://doi.org/10.1088/1367-2630/aa9b74>  
...Page(269)

- **A Strange Metal from Gutzwiller correlations in infinite dimensions**,  
 Wenxin Ding, Rok Žitko, Peizhi Mai, Edward Perepelitsky and B. Sriram Shastry,  
 arXiv:1703.02206v2; Phys. Rev. **B 96**, 054114 (2017).  
 DOI: 10.1103/PhysRevB.96.054114 ...Page(256)
- **Origin of Kinks in Energy Dispersion of Strongly Correlated Matter**,  
 Kazue Matsuyama, Edward Perepelitsky and B Sriram Shastry, arXiv:1610.08079,  
 Phys. Rev. **B 95**, 165435 (2017).  
 DOI: http://link.aps.org/DOI: 10.1103/PhysRevB.95.165435  
 ...Page(236)
- **Low energy physics of the  $t$ - $J$  model in  $d = \infty$  using Extremely Correlated Fermi Liquid theory: Cutoff Second Order Equation**,  
 B. S. Shastry and E. Perepelitsky”, arXiv:1605.08213. Phys. Rev. **B 94**, 045138  
 (2016).  
 DOI: http://link.aps.org/doi/10.1103/PhysRevB.94.045138  
 ...Page(222)
- **Diagrammatic  $\lambda$  series for extremely correlated Fermi liquids**,  
 E. Perepelitsky and B. S. Shastry, arXiv: 1410.5174, Ann. Phys. **357**, 1 (2015).  
 DOI: doi:10.1016/j.aop.2015.03.010  
 ...Page(183)
- **Linked-Cluster Expansion for the Green’s function of the Infinite- $U$  Hubbard Model**,  
 E. Khatami, E. Perepelitsky, M. Rigol and B. S. Shastry, arXiv: 1310.8029 (2013),  
 Phys. Rev. **E 89**, 063301 (2014).  
 DOI:http://dx.doi.org/10.1103/PhysRevE.89.063301  
 ...Page(170)
- **Theory of extreme correlations using canonical Fermions and path integrals**,  
 B. S. Shastry, arXiv:1312.1892 (2013), Ann. Phys. **343**, 164-199 (2014).  
 DOI:http://dx.doi.org/10.1016/j.aop.2014.02.005.  
 (Erratum) Ann. Phys. Vol. 373, 717-718 (2016).  
 DOI:http://dx.doi.org/10.1016/j.aop.2016.08.015.  
 ...Page(132)
- **ECFL in the limit of infinite dimensions**,  
 E. Perepelitsky and B. S. Shastry, arXiv:1309.5373 (2013), Annals of Physics **338**, 283-  
 301 (2013).  
 DOI:http://dx.doi.org/10.1016/j.aop.2013.09.010  
 ...Page(113)
- **Extremely correlated Fermi liquid theory meets Dynamical mean-field theory: Analytical insights into the doping-driven Mott transition**,  
 R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges and B. S. Shastry,  
 arXiv:1309.5284 (2013), Phys. Rev. **B 88**, 235132 (2013).  
 DOI:http://dx.doi.org/10.1103/PhysRevB.88.235132  
 ...Page(95)
- **Extremely Correlated Fermi Liquid study of the  $U = \infty$  Anderson Impurity Model**,  
 B. S. Shastry, E. Perepelitsky and A. C. Hewson, arXiv:1307.3492 [cond-mat.str-el],

Phys. Rev. **B 88**, 205108 (2013).

DOI:<http://dx.doi.org/10.1103/PhysRevB.88.205108>

...Page(86)

- **Electronic spectral properties of the two-dimensional infinite-U Hubbard model**,  
 E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, B. S. Shastry, arXiv:1303.2657  
 [cond-mat.str-el] (2013), Phys. Rev. **B 87**, 161120 (R) (2013).  
 DOI: 10.1103/PhysRevB.87.161120

...Page(81)
- **Extremely Correlated Fermi Liquids: Self consistent solution of the second order theory**,  
 D. Hansen and B. S. Shastry, arXiv:1211.0594, (2012), Phys. Rev. **B 87**, 245101 (2013).  
 DOI: 10.1103/PhysRevB.87.245101

...Page(66)
- **Extremely Correlated Fermi Liquids: The Formalism**,  
 B. S. Shastry, arXiv:1207.6826 (2012); Phys. Rev. **B 87**, 125124 (2013).  
 DOI: 10.1103/PhysRevB.87.125124

...Page(43)
- **Dynamical Particle Hole Asymmetry in Cuprate Superconductors**,  
 B. S. Shastry, arXiv:1110.1032 (2011), Phys. Rev. Letts. **109**, 067004 (2012).  
 DOI: 10.1103/PhysRevLett.109.067004

...Page(38)
- **Anatomy of the Self Energy**,  
 B. S. Shastry, arXiv:1104.2633; Phys. Rev. **B 84**, 165112 (2011);  
 DOI: 10.1103/PhysRevB.84.165112  
 Phys. Rev. **B 86**, 079911(E) (2012).  
 DOI: 10.1103/PhysRevB.86.079911

...Page(24)
- **Extremely Correlated Fermi Liquid Description of Normal State ARPES in Cuprates**,  
 G.-H. Gweon, B. S. Shastry and G. D. Gu, arXiv:1104.2631 (2011), Phys. Rev. Letts.  
**107**, 056404 (2011).  
 DOI: 10.1103/PhysRevLett.107.056404

...Page(19)
- **Extremely Correlated Fermi Liquids**  
 B. S. Shastry, arXiv:1102.2858 (2011), Phys. Rev. Letts. **107**, 056403 (2011).  
 DOI: 10.1103/PhysRevLett.107.056403

...Page(14)

## Extremely Correlated Fermi Liquids

B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*

(Received 14 February 2011; published 29 July 2011)

We present the theory of an extremely correlated Fermi liquid with  $U \rightarrow \infty$ . This liquid has an underlying auxiliary Fermi liquid Green's function that is further caparisoned by extreme correlations. The theory leads to two parallel hierarchies of equations that permit iterative approximations in a certain parameter. Preliminary results for the spectral functions display a broad background and a distinct  $T$  dependent left skew. An important energy scale  $\Delta(\vec{k}, x)$  emerges as the average inelasticity of the FL Green's function, and influences the photoemission spectra profoundly. A duality is identified wherein a loss of coherence of the ECFL results from an excessively sharp FL.

DOI: 10.1103/PhysRevLett.107.056403

PACS numbers: 71.10.Fd

*Introduction.*—Correlated electron systems attract two distinct approaches. An intermediate to strong coupling approach is used when the interaction  $U$  is comparable to the band width  $2W$ , and has seen some success in recent times [1]. On the other hand, Anderson [2] has argued that myriad experiments on high  $T_c$  superconductors require a better understanding of the  $t$ - $J$  model physics. This model sets  $U \rightarrow \infty$  right away, i.e., leads to *extreme correlations* and involves Gutzwiller projected Fermi operators that are non canonical. Thus Wick's theorem is immediately lost, and perturbative schemes encoding the Feynman Dyson approach become useless. Since this approach is at the root of most current many body physics text books, the task of understanding the  $t$ - $J$  model is not lightly undertaken.

The Schwinger approach to interacting field theories is a powerful and attractive alternative. It is fundamentally non perturbative, where Wick's theorem is bypassed by dealing with suitable inverse Greens functions. Conventional many body theory for canonical Fermions can also be cast into this approach, and leads to the standard results. In Ref. [3] (henceforth referred to as paper I), the author has recently applied the Schwinger method to the  $t$ - $J$  model, and found a class of solutions that are termed as extremely correlated quantum liquids. That state is presumably realized under suitable conditions. However it gives a Fermi surface (FS) volume that is always distinct from that of the Fermi gas. This is contrast to the case of Fermi liquids (FL), where the important theorem of Luttinger and Ward (LW) [4,5] mandates the invariance of the FS volume under interactions.

In this Letter we propose a state of matter termed as an *extremely correlated Fermi liquid* (ECFL). The ECFL found here, represents an alternate class of solutions for the  $t$ - $J$  model, where the Fermi surface satisfies the Fermi gas (i.e., LW) volume. In this work we present the essentials of the formalism, and display preliminary results on spectral functions that are suggestive of the relevance of the ECFL state to cuprate materials. An inherent flexibility of the Schwinger approach permits the construction of an alternate class of solutions from the one found in paper I.

The excitations of the ECFL state may be thought of as bare electrons undergoing a double layer of renormalization: the FL dressing into quasiparticles that are further caparisoned (i.e., decorated) by extreme correlations.

*Formalism.*—The physical projected electronic Green's function  $\mathcal{G}$  satisfies an equation of motion (EOM) (I-29) written compactly in matrix form as

$$(\partial_{\tau_i} - \boldsymbol{\mu})\mathcal{G}(i, f) = -\delta(i, f)\{1 - \gamma(i)\} - \mathcal{V}_i \cdot \mathcal{G}(i, f) - X(i, \bar{\mathbf{j}}) \cdot \mathcal{G}(\bar{\mathbf{j}}, f) - Y(i, \bar{\mathbf{j}}) \cdot \mathcal{G}(\bar{\mathbf{j}}, f), \quad (1)$$

where  $\boldsymbol{\mu}$  is the chemical potential and an implicit integration over space-time variables such as  $\bar{\mathbf{j}}$ , written with bold overlined letters, is implied,

$$\begin{aligned} X(i, j) &= -t(i, j)[D(i) + D(j)] \\ &\quad + \frac{1}{2}J(i, \bar{\mathbf{k}})[D(i) + D(\bar{\mathbf{k}})]\delta(i, j) \\ Y(i, j) &= -t(i, j)[\mathbb{1} - \gamma(i) - \gamma(j)] \\ &\quad + \frac{1}{2}J(i, \bar{\mathbf{k}})[\mathbb{1} - \gamma(i) - \gamma(\bar{\mathbf{k}})]\delta(i, j). \end{aligned} \quad (2)$$

In the above expression [6], we used  $\gamma(i) = \mathcal{G}^k(i, i)$  with the  $k$  conjugation defined by  $(M^k)_{\sigma_1\sigma_2} = M_{\bar{\sigma}_2\bar{\sigma}_1}\sigma_1\sigma_2$ , and  $D_{\sigma_1\sigma_2}(i) = \sigma_1\sigma_2 \frac{\delta}{\delta \mathcal{V}_i^{\sigma_1\sigma_2}}$ . The added (bosonic) source term  $\mathcal{V}_i^{\sigma_1\sigma_2}(\tau_i)$  is central to this approach; it is a space-time dependent field that couples to the charge and spin densities through a term in the action:  $\sum_{i\sigma} \int_0^\beta d\tau \mathcal{V}_i^{\sigma_1\sigma_2}(\tau) X_i^{\sigma_1\sigma_2}(\tau)$ , where  $X_i^{\sigma_1\sigma_2}$  is the spin and density operator at site  $i$  that acts as  $|\sigma_1\rangle\langle\sigma_2|$ .

An important technical problem highlighted in I is to deal with the time dependence of the  $\gamma(i)$  term in Eq. (1) which makes the theory noncanonical. Here we use the decomposition into two factors [7]:

$$\mathcal{G}(a, b) = \mathbf{g}(a, \bar{\mathbf{b}}) \cdot \boldsymbol{\mu}(\bar{\mathbf{b}}, b), \quad (3)$$

and express  $\gamma(i) = [\mathbf{g}(i, \bar{\mathbf{j}}) \cdot \boldsymbol{\mu}(\bar{\mathbf{j}}, i)]^k$ . The object  $\mathbf{g}$  is an auxiliary FL Green's function and  $\boldsymbol{\mu}(\bar{\mathbf{b}}, b)$  is an appurtenant (or supplementary) factor that is determined below.

Antiperiodic boundary conditions  $\mathcal{G}(0, \tau_f) = -\mathcal{G}(\beta, \tau_f)$  and  $\mathcal{G}(\tau_i, 0) = -\mathcal{G}(\tau_i, \beta)$  imply that both factors  $\mathbf{g}$  and  $\mu$  are Fourier transformed using fermionic Matsubara frequencies. We define the inverse Green's function  $\mathbf{g}^{-1}(a, \bar{\mathbf{b}}) \cdot \mathbf{g}(\bar{\mathbf{b}}, b) = \mathbb{1} \delta(a, b)$ , and then a vertex function  $\Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p, q; r) = -\frac{\delta}{\delta \mathcal{V}_r^{\sigma_3 \sigma_4}} \{ \mathbf{g}_{\sigma_1 \sigma_2}^{-1}(p, q) \}$ . Thus  $\mathbf{g}$ ,  $\mu$  and  $\mathbf{g}^{-1}$  are matrices in the spin space, and the vertex  $\Lambda$  has four indices. We also define a linear operator

$$\mathbf{L}(i, f) = \left( t(i, \bar{\mathbf{j}}) \xi^* \cdot \mathbf{g}(\bar{\mathbf{j}}, f) - \frac{1}{2} J(i, \bar{\mathbf{j}}) \xi^* \cdot \mathbf{g}(i, f) \right) \times \left( \frac{\delta}{\delta \mathcal{V}_i^*} + \frac{\delta}{\delta \mathcal{V}_{\bar{\mathbf{j}}}^*} \right), \quad (4)$$

where the matrix  $\xi_{\sigma_1 \sigma_2}^* = \sigma_1 \sigma_2$ . The asterisk is used as a place holder that transmits the spin indices (after conjugation) of the  $\xi$  matrix to the source matrix  $\mathcal{V}$  in the functional derivative. This notation used is illustrated in component form by  $\cdots \xi_{\sigma_a \sigma_b}^* \cdots \delta / \delta \mathcal{V}_{\bar{\mathbf{j}}}^* = \cdots \sigma_a \sigma_b \cdots \delta / \delta \mathcal{V}_{\bar{\mathbf{j}}}^{\sigma_a \sigma_b}$ .

A useful chain rule for the functional derivative is noted

$$D(r) \mathcal{G}(a, b) = \xi^* \cdot \mathbf{g}(a, \bar{\mathbf{c}}) \cdot \Lambda_*(\bar{\mathbf{c}}, \bar{\mathbf{d}}; r) \cdot \mathcal{G}(\bar{\mathbf{d}}, b) + \xi^* \cdot \mathbf{g}(a, \bar{\mathbf{b}}) \cdot \left( \frac{\delta}{\delta \mathcal{V}_r^*} \mu(\bar{\mathbf{b}}, b) \right). \quad (5)$$

Using this chain rule, we see that

$$X(i, \bar{\mathbf{j}}) \cdot \mathcal{G}(\bar{\mathbf{j}}, f) \equiv \Phi(i, \bar{\mathbf{b}}) \cdot \mathcal{G}(\bar{\mathbf{b}}, f) + \Psi(i, f), \quad (6)$$

where

$$\begin{aligned} \Phi(i, m) &= \mathbf{L}(i, \bar{\mathbf{i}}) \cdot \mathbf{g}^{-1}(\bar{\mathbf{i}}, m) \\ \Psi(i, m) &= -\mathbf{L}(i, \bar{\mathbf{i}}) \cdot \mu(\bar{\mathbf{i}}, m). \end{aligned} \quad (7)$$

Thus the two fundamental functions of this formalism  $\Phi$ ,  $\Psi$  are closely connected as they arise from applying the same operator to the two factors of  $\mathcal{G}$ . Defining  $Y_0(i, j) = [-t(i, j) + \frac{1}{2} J(i, \bar{\mathbf{k}}) \delta(i, j)] \mathbb{1}$ , and  $Y_1(i, j) = t(i, j) [\gamma(i) + \gamma(j)] - \frac{1}{2} \delta(i, j) J(i, \bar{\mathbf{k}}) [\gamma(i) + \gamma(\bar{\mathbf{k}})]$ , also denote the Fermi gas Green's function

$$\mathbf{g}_0^{-1}(i, f) = \{ -(\partial_{\tau_i} - \boldsymbol{\mu}) \mathbb{1} - \mathcal{V}_i \} \delta(i, f) - Y_0(i, f). \quad (8)$$

Collecting everything, the exact EOM can now be written neatly as

$$\begin{aligned} \{ \mathbf{g}_0^{-1}(i, \bar{\mathbf{j}}) - \lambda Y_1(i, \bar{\mathbf{j}}) - \lambda \Phi(i, \bar{\mathbf{j}}) \} \cdot \mathbf{g}(\bar{\mathbf{j}}, \bar{\mathbf{f}}) \cdot \mu(\bar{\mathbf{f}}, f) \\ = \delta(i, f) [\mathbb{1} - \lambda \gamma(i)] + \lambda \Psi(i, f). \end{aligned} \quad (9)$$

We have introduced the parameter  $\lambda$  above, with  $0 \leq \lambda \leq 1$ , in order to provide an adiabatic path between the Fermi gas at  $\lambda = 0$  and the ECFL at  $\lambda = 1$ , and also an iterative scheme in powers of  $\lambda$  connecting the two endpoints.

We now choose the hitherto undetermined function  $\mu$  as

$$\mu(i, f) = \delta(i, f) [\mathbb{1} - \lambda \gamma(i)] + \lambda \Psi(i, f), \quad (10)$$

so that Eq. (9) reduces to a canonical FL type equation:

$$\{ \mathbf{g}_0^{-1}(i, \bar{\mathbf{j}}) - \lambda Y_1(i, \bar{\mathbf{j}}) - \lambda \Phi(i, \bar{\mathbf{j}}) \} \cdot \mathbf{g}(\bar{\mathbf{j}}, f) = \delta(i, f). \quad (11)$$

Notice that the right-hand side has a pure  $\delta$  function as in a canonical Fermi liquid type theory. To summarize, the EOM Eq. (1) under the decomposition Eq. (3) leads to Eq. (9). In turn this splits exactly into two coupled sets of equations Eq. (7), (10), and (11) for the two factors  $\mathbf{g}$  and  $\mu$ . Note that the entire procedure is exact, we write explicit forms of these equations below and then introduce approximate methods to solve them.

Inverting we find Dyson's equation for the auxiliary FL Green's function:

$$\mathbf{g}^{-1}(i, m) = \{ \mathbf{g}_0^{-1}(i, m) - \lambda Y_1(i, m) - \lambda \Phi(i, m) \}. \quad (12)$$

Taking functional derivatives of Eq. (10) and (12) with respect to  $\mathcal{V}$ , and comparing with Eq. (4) and (7) we generate *two parallel hierarchies of equations* for  $\mathbf{g}$  and  $\mu$  that form the core of this formalism. The hierarchy for  $\mathbf{g}$  is essentially autonomous and drives that for  $\mu$ . Starting with the Fermi gas at  $O(\lambda^0)$ , an iterative process similar to the skeleton graph expansion of LW [4] can be built up, such that terms of  $O(\lambda^n)$  arise from differentiating lower order terms of  $O(\lambda^{n-1})$ . Systematic approximations may thus be arranged to include all terms of  $O(\lambda^n)$  for various  $n$  [8]. The number of particles is given by  $\frac{1}{2} n(i) = \mathbf{g}(i, \bar{\mathbf{i}}) \mu(\bar{\mathbf{i}}, i)$ , and with

$$\mathcal{U}_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(a, b; c) \equiv \frac{\delta \mu_{\sigma_1 \sigma_2}(a, b)}{\delta \mathcal{V}_c^{\sigma_3 \sigma_4}}, \quad (13)$$

the equations to solve simultaneously are Eq. (7), (12), and (10). The density and spin density response functions (I-F1), (I-F7) can be found from differentiating  $\mathcal{G}$ , i.e.,  $\mathcal{Y}_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p, q; r) = \frac{\delta}{\delta \mathcal{V}_r^{\sigma_3 \sigma_4}} \{ \mathcal{G}_{\sigma_1 \sigma_2}(p, q) \}$ .

*Zero source limit in Fourier space.*—When we turn off the source  $\mathcal{V}$ , the various matrix function  $\mathcal{G}$ ,  $\mathbf{g}$ ,  $\mu$  become spin diagonal and translation invariant so we can Fourier transform these conveniently. We note the basic result expressing  $\mathcal{G}$  as a simple product of two functions in  $k$  space:

$$\begin{aligned} \mathcal{G}(k) &= \mathbf{g}(k) \mu(k), \quad \mu(k) = 1 - \lambda \frac{n}{2} + \lambda \Psi(k), \\ \mathbf{g}^{-1}(k) &= i\omega_k + \boldsymbol{\mu} - \varepsilon_k (1 - \lambda n) - \lambda \Phi(k), \end{aligned} \quad (14)$$

where  $\varepsilon_k$  is the Fourier transform of the hopping matrix  $-t(i, j)$ , and an uninteresting constant term is absorbed in  $\boldsymbol{\mu}$  here and below.

Here,  $\mathbf{g}$  plays the role of an underlying auxiliary FL with a self energy  $\Phi$ , and  $\Psi$  acts as an extra spectral weight that vanishes at high frequency, leaving the exact weight  $1 - \frac{n}{2}$  valid for a projected electron (as in paper I) for  $\lambda = 1$ . Denoting  $\sum_k \rightarrow \frac{1}{N_s \beta} \sum_{i\omega_k, \bar{k}}$  with  $N_s$  sites, the particle number sum rule is  $\sum_k \mu(k) \mathbf{g}(k) = \frac{n}{2}$ , i.e.,

$$\frac{n}{2} = \sum_k \mathbf{g}(k) + \lambda \sum_k \left( \Psi(k) - \frac{n}{2} \right) \mathbf{g}(k). \quad (15)$$

In this formalism, at  $k \sim k_F$ ,  $x = 0$  that is relevant to the LW sum rule, the  $\text{Re } \mathbf{g}(k)$  dominates  $\text{Re } \mathcal{G}(\vec{k}, 0)$  (since  $\text{Re } \Psi(\vec{k}, 0)$  is smooth through the FS). Requiring consistency with the LW theorem forces us to pin any sign change of  $\text{Re } \mathbf{g}(\vec{k}, 0)$  to the free case, whereby we impose a *second level sum rule*

$$\sum_k \Psi(k) \mathbf{g}(k) = \frac{n^2}{4}, \quad \text{and} \quad \sum_k \mathbf{g}(k) = \frac{n}{2}. \quad (16)$$

This can be viewed as a splitting of the usual number sum rule Eq. (15) [9]. With  $E(p_1, p_2) = [\varepsilon_{p_1} + \varepsilon_{p_2} + \frac{1}{2} \hat{J}(0) + \frac{1}{2} \hat{J}(p_1 - p_2)]$  we find

$$\begin{aligned} \Phi(k) &= \sum_p E(k, p) \mathbf{g}(p) \Lambda^{(a)}(p, k), \\ \Psi(k) &= \sum_p E(k, p) \mathbf{g}(p) \mathcal{U}^{(a)}(p, k), \end{aligned} \quad (17)$$

and the spin labels are from paper I with the usual significance  $\Lambda^{(a)} = \Lambda^{(2)} - \Lambda^{(3)} = \frac{1}{2} \Lambda^{(s)} - \frac{3}{2} \Lambda^{(t)}$ .

Next we introduce the spectral representation of various functions  $Q$  that vanish at infinity:  $Q(i\omega_Q) = \int_{-\infty}^{\infty} dx \frac{\rho_Q(x)}{i\omega_Q - x}$  and  $\rho_Q(x) = -\frac{1}{\pi} \text{Im } Q(x + i0^+)$ , with  $x^+ \equiv x + i0^+$ . The Matsubara frequency  $\omega_Q$  is fermionic (bosonic) if  $Q$  is fermionic (bosonic). Proceeding further, at any order in  $\lambda$ , the two hierarchies give us coupled equations for the spectral densities of the physical particles  $\rho_{\mathcal{G}}(\vec{k}, x)$  as well as the underlying Fermi liquid  $\rho_{\mathbf{g}}(\vec{k}, x)$ , in terms of the two objects  $\rho_{\bar{\Phi}}(\vec{k}, x)$  and  $\rho_{\Psi}(\vec{k}, x)$  and their Hilbert transforms. The Lehmann representation implies that  $\rho_{\mathcal{G}}(\vec{k}, x)$  is positive at all  $\vec{k}, x$ . In making approximations, this important and challenging constraint must be kept in mind.

*Solution of  $\mathbf{g}^{-1}$  and  $\mu$  to order  $O(\lambda)^2$ .*—We next discuss a systematic expansion in powers of  $\lambda$  [8], obtained by taking functional derivatives of Eq. (10) and (12) to generate expressions for the vertices given the Green's functions via  $\Lambda \sim -\frac{\delta}{\delta \mathcal{V}} \mathbf{g}^{-1}$  and  $\mathcal{U} \sim \frac{\delta}{\delta \mathcal{V}} \mu$ . To lowest order in  $\lambda$ , the bare vertex  $\Lambda^{(a)} = -1$ , this term is absorbed in a renormalization of the band dispersion to  $\bar{\varepsilon}_k$  in Eq. (14) [10], and the remaining term denoted by  $\bar{\Phi}(k)$ . To this order  $\mathcal{U}^{(a)} = 0$ . Proceeding to the next non trivial order in  $\lambda$ , by taking the functional derivative of Eq. (10) and (12) we find after a brief calculation:

$$\begin{aligned} \Psi(k) &= -2\lambda \sum_{p,q} E(k, p) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(q + p - k), \\ \bar{\Phi}(k) &= -2\lambda \sum_{p,q} E(k, p) [E(p, k) + E(q + p - k, p)] \\ &\quad \times \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(q + p - k). \end{aligned} \quad (18)$$

From Eq. (14) we note that these expressions Eq. (18) lead to a calculation of  $\mathbf{g}^{-1}$  and  $\mu$  correct up to  $O(\lambda^2)$ . Frequency dependent corrections arise only to second order in  $\lambda$ , which is analogous to the structure of the canonical many body theory within the skeleton graph expansion. We may now set  $\lambda = 1$  and study the resulting theory as the first step in exploring this formalism.

Denote  $f(x) = \frac{1}{(\exp \beta x + 1)}$  as the Fermi distribution functions and  $\bar{f}(x) = 1 - f(x)$ , and denote the usual Fermi factors from second order theory

$$\mathcal{W} = \{f(u)f(w)\bar{f}(v) + f(v)\bar{f}(u)\bar{f}(w)\} \delta(u + w - v - x),$$

a function of the frequencies  $u, v, w, x$ , and

$$\mathcal{Y} = \int_{u,v,w} \mathcal{W} \rho_{\mathbf{g}}(\vec{q}, w) \rho_{\mathbf{g}}(\vec{p}, u) \rho_{\mathbf{g}}(\vec{q} + \vec{p} - \vec{k}, v), \quad (19)$$

a function of  $\vec{k}, \vec{p}, \vec{q}$ , and  $x$ . We may then write the spectral functions corresponding to Eq. (18)

$$\begin{aligned} \rho_{\bar{\Phi}}(\vec{k}, x) &= 2 \sum_{\vec{p}, \vec{q}} E(\vec{k}, \vec{p}) [E(\vec{p}, \vec{k}) + E(\vec{q} + \vec{p} - \vec{k}, \vec{p})] \mathcal{Y}, \\ \rho_{\Psi}(\vec{k}, x) &= 2 \sum_{\vec{p}, \vec{q}} E(\vec{k}, \vec{p}) \mathcal{Y}. \end{aligned} \quad (20)$$

The functions appearing in Eq. (20) are familiar from Fermi liquids [4,5], and encode the usual phase space constraints of that theory. This leads to the low temperatures behavior  $\sim \max\{x^2, (\pi k_B T)^2\}$ , for both objects  $\text{Im } \Psi(k, x, T)$  and  $\text{Im } \bar{\Phi}(k, x, T)$ . The real parts of these objects are smooth through the Fermi surface, as one expects from the real part of the self energy in a FL, and hence motivates the second level sum rule Eq. (16).

From Eq. (14) we write the exact expression for the physical spectral function  $\rho_{\mathcal{G}}$ :

$$\rho_{\mathcal{G}}(\vec{k}, x) = \rho_{\mathbf{g}}(\vec{k}, x) \left( \left\{ 1 - \frac{n}{2} \right\} + \frac{\xi_k - x}{\Delta(\vec{k}, x)} + \eta(\vec{k}, x) \right), \quad (21)$$

where  $\xi_k = \bar{\varepsilon}_k - \mu$ , and the important energy scale  $\Delta(\vec{k}, x)$  and the term  $\eta$  is defined as

$$\Delta(\vec{k}, x) = -\frac{\rho_{\bar{\Phi}}(\vec{k}, x)}{\rho_{\Psi}(\vec{k}, x)}, \quad (22)$$

$$\eta(\vec{k}, x) = \text{Re} \Psi(\vec{k}, x^+) + \frac{1}{\Delta(\vec{k}, x)} \text{Re} \Phi(\vec{k}, x^+). \quad (23)$$

The sign of the energy scale  $\Delta$  in Eq. (22) is expected to be positive from Eq. (20). The dimensionless term  $\eta$  augments the spectral weight at the Fermi level. The equations necessary to solve the theory to  $O(\lambda^2)$  may be summarized as Eq. (14), (16), and (18) and Ref. [10] giving rise to the spectral function Eq. (21). These require further numerical work that is underway, it leads to spectral functions in 2 and 3 dimensions that will be published separately. However it also provides a very interesting insight about

the theory in high dimensions that is pursued analytically next.

*Solution in high dimensions.*—In sufficiently high dimensions, we show next that the dimensionless term  $\eta$  vanishes identically leading to a great simplification. For sufficiently high dimensions we can ignore the momentum dependence of  $\mathcal{Y}$  in Eq. (19) and assume  $\rho_{\Phi}(\vec{k}, x) \sim C_{\Phi}\sigma(x)$ , and  $\rho_{\Psi}(\vec{k}, x) \sim C_{\Psi}\sigma(x)$ , as functions of frequency only. Here  $\sigma(x)$  extends over energy range  $\omega_c \sim O(2W)$ , and  $C_{\Phi}$  has dimensions of inverse energy and is positive due to  $\rho_{\Phi}$ . Its Hilbert transform is called  $h(x) \equiv \mathcal{P} \int dy \frac{\sigma(y)}{x-y}$ . We use an analytically tractable Fermi liquid model [11] with  $\tau = \pi k_B T$ , where we set

$$\sigma(x) = \{x^2 + \tau^2\} e^{-C_{\Phi}\{x^2 + \tau^2\}/\omega_c}. \quad (24)$$

The peak value of  $C_{\Phi}\sigma(x)$  is of  $O(1)$  and independent of  $C_{\Phi}$  [12]. The other constant  $C_{\Psi}$  is dimensionless and negative. To complete the model, we note that the real parts are given in terms of  $h(x)$  as  $\text{Re } \bar{\Phi}(x^+) = C_{\Phi}h(x)$  and  $\text{Re } \Psi(x^+) = C_{\Psi}h(x)$ . With this choice the auxiliary spectral weight  $\eta(k, x)$  vanishes identically in Eq. (23). With  $\Gamma(x) \equiv \pi C_{\Phi}\sigma(x)$  and  $\epsilon(\xi, x) \equiv [x - \xi - C_{\Phi}h(x)]$  we may write  $\rho_{\mathcal{G}}(\xi, x) = \frac{1}{\pi} \frac{\Gamma(x)}{\Gamma^2(x) + \epsilon^2(\xi, x)}$  and  $\text{Re } \mathcal{g}(\xi, x) = \frac{\epsilon(\xi, x)}{\Gamma^2(x) + \epsilon^2(\xi, x)}$ . Denoting  $\langle Q(\xi) \rangle_{\xi} = \int d\xi N_B(\xi) Q(\xi)$ , where  $N_B(\xi)$  is the band density of states per spin, the chemical potential is fixed using  $\frac{n}{2} = \int_{-\infty}^{\infty} dx f(x) \langle \rho_{\mathcal{G}}(\xi, x) \rangle_{\xi}$ .

The energy parameter  $\Delta(\vec{k}, x)$  in Eq. (22) is a constant. We scale out a factor to define

$$\Delta_o = \frac{n^2}{4} \Delta(\vec{k}, x) = -\frac{n^2}{4} \frac{C_{\Phi}}{C_{\Psi}}. \quad (25)$$

The physically observable electronic spectral function reads

$$\rho_{\mathcal{G}}(\xi, x) = \frac{\Gamma(x)}{\pi} \frac{\{1 - \frac{n}{2}\} + \frac{(n^2)}{4} \{\frac{\xi - x}{\Delta_o}\}_+}{\Gamma^2(x) + \epsilon^2(\xi, x)}. \quad (26)$$

Here, the condition  $(f)_+ \equiv \max(0, f)$ , is inserted in the ECFL factor to guarantee the positivity of the spectral function for  $x \gg \xi$  [13]. We can determine  $\Delta_o$  directly from the second level sum rule Eq. (16):

$$\Delta_o = \int_{-\infty}^{\infty} dx f(x) \langle \rho_{\mathcal{G}}(\xi, x) \{ \xi - x \} \rangle_{\xi}. \quad (27)$$

Thus  $(2/n)\Delta_o$  is the average inelasticity  $\|(\xi - x)\|$  of the FL Green's function over the entire occupied band. It vanishes if  $\rho_{\mathcal{G}}$  were a pure delta function, as in a Fermi gas, but is non zero in a Fermi liquid. The linear energy term in Eq. (26) thus fundamentally arises to provide the extra density to  $\rho_{\mathcal{G}}$ , compensating the spectral depletion due to the first factor  $1 - \frac{n}{2}$  [originating in the non canonical nature of the projected electrons (paper I)].

In the numerical solution of the model, we can vary the shapes of the spectra from sharp to broad by controlling the energy scale  $\Delta_o$  via the parameters  $C_{\Phi}$  and  $\omega_0$  in the FL function  $\sigma(x)$ . For illustration we neglect the distinction between the band energy and the renormalized  $\bar{\epsilon}_k$ , choose a flat band density of states per spin  $\rho_0(\epsilon) = \frac{1}{2W} \Theta(W^2 - \epsilon^2)$  hence the band width is  $2W$ . Choose  $C_{\Phi} = 1$   $W = 10^4$  K [14], this gives  $\Delta_o \sim 600$  K in the cases studied. The spectral shapes from Eq. (26) have a characteristic left skew that is visible in Fig. 1, and also in many experimental spectra in high  $T_c$  systems. The marginal Fermi liquid hypothesis [15] assumes a linear correction to the spectral function, but is *symmetric* about the Fermi energy, i.e., of the form  $|\xi - x|$  instead of the term in Eq. (26).

From Eq. (27) a fascinating *duality* emerges between the FL and the ECFL [16]. When the FL is overall sharp such that  $\Delta_o$  is small, the ECFL is significantly broadened. This happens since in the ECFL factor in Eq. (26), the coefficient of  $\xi - x$  becomes large and dominates the  $1 - \frac{n}{2}$  contribution. The function  $\Delta(k)$  in Eq. (22) could vanish at points in  $k$  space in the full theory (without the assumption of  $k$  independence). At those points the ECFL spectra would lose all coherence by this duality. A loss of coherence would inevitably suggest a (false) pseudogap, if our current viewpoint were unavailable. The linear term also leads to a sloping term in the local density of states of the ECFL that the STM technique would probe, although its magnitude and sign are less reliably computed—depending as they do on the high energy scales  $W$  and  $\omega_0$ . In conclusion, we have presented essential ideas underlying the theory of extremely correlated Fermi liquids. We have

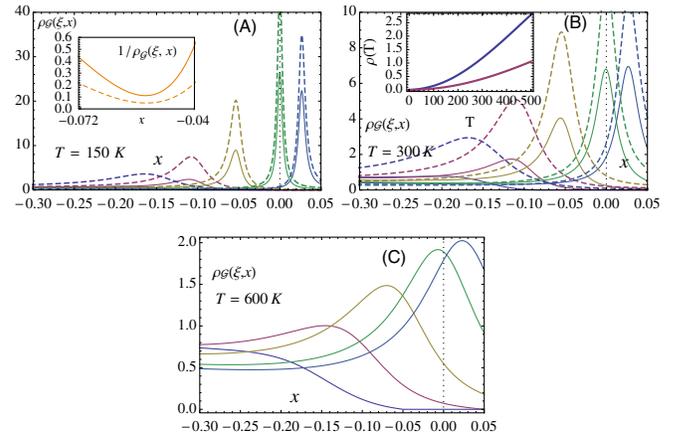


FIG. 1 (color online). The density  $n = .85$  and  $\omega_0 = 0.25$ . From left to right  $\rho_{\mathcal{G}}(x)$  for energies (in units of  $W$ )  $\xi = -0.3, -0.2, -0.1, 0, 0.05$  for both the FL (dashed) and the ECFL (solid) theories. Inset in (A): provides an enlarged view of the  $\xi = -0.1$  plots after *inversion*, and displays the left-skew asymmetry of the ECFL spectrum relative to the FL. Inset in (B) shows the DC resistivity  $\rho(T)$  within a bubble approximation as a function of  $T$  for the FL (blue) and the ECFL (red). Because of spectral redistribution, the ECFL reaches linear  $T$  behavior at a lower  $T$  than the FL.

shown that an explicit low order solution is very promising in the context of explaining the photoemission spectra of the cuprate materials.

Detailed numerics and comparison with experiments are currently underway.

This work was supported by DOE under Grant No. FG02-06ER46319.

- 
- [1] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989); A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [2] P.W. Anderson, *Science* **235**, 1196 (1987); *Phys. Rev. B* **78**, 174505 (2008).
- [3] B. S. Shastry, *Phys. Rev. B* **81**, 045121 (2010).
- [4] J.M. Luttinger and J.C. Ward, *Phys. Rev.* **118**, 1417 (1960); J.M. Luttinger, *Phys. Rev.* **119**, 1153 (1960); *Phys. Rev.* **121**, 942 (1961).
- [5] A.A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).
- [6] This expression differs slightly from (I-29) in Ref. [3]. The added terms vanish exactly due to the Pauli principle and Gutzwiller projection that eliminate states with double occupancy. They are added since they serve to usefully symmetrize the expressions for  $X(i, j)$  and  $Y(i, j)$  and thereby simplify the subsequent treatment.
- [7] This choice is the essential difference from a decomposition in paper I Eq. (I-31). In the present case, we are able to establish adiabatic continuity with the Fermi gas as indicated in Eq. (9) below.
- [8] It is important to realize that the nature of the  $\lambda$  expansion is different from that of a typical perturbative expansion, e.g., the  $U$  expansion in the Hubbard model. In the present case the scales of all parameters  $t, J, \lambda, \dots$  are similar  $\sim O(1)$ , since the large parameter  $U$  of the Hubbard model has been set at  $\infty$  at the outset. Thus  $\lambda$  should be viewed as a parameter that organizes the equations in a systematic fashion so that, e.g., keeping *all terms* of  $O(\lambda^2)$  gives a consistent theory that is structurally analogous to the skeleton expansion at  $O(U^2)$ . We may then examine the nature of the solution to this order, with the expectation that the  $O(\lambda^3)$  term would retain the qualitative features found at lower order. Further since  $\lambda$  couples to  $(\mathbf{g} \cdot \boldsymbol{\mu})^k \sim n$  in Eq. (10), an expansion in  $\lambda$  should be viewed as a low density expansion.
- [9] We have thus imposed the LW theorem in Eq. (16) rather than proved it in this theory.
- [10] In Eq. (14) we lump all dispersion type terms in the expression into an effective dispersion  $\bar{\epsilon}_k = (1 - \lambda n)\epsilon_k - \lambda \sum_p E(k, p) \mathbf{g}(p) e^{i\omega_p 0^+}$ , thus writing  $\mathbf{g}^{-1}(k) = i\omega_k + \boldsymbol{\mu} - \bar{\epsilon}_k - \lambda \bar{\Phi}(k)$ .
- [11] This functional form is based on the  $O(U^2)$  result for of the Hubbard model, after one sets the parameter  $U \rightarrow 1$  in view of the scales of variables of the  $t$ - $J$  model noted in [8]. For small energies the behavior of the Fermi liquid noted in [5] is captured in this form, and a reasonable extrapolation to large frequencies is provided.
- [12] The Hilbert transform of  $\sigma(x)$  is denoted by  $h(x)$ . Here  $h(x) = h_1(x) - h_2(x)$  with  $h_1(x) = \pi(x^2 + \tau^2) e^{-C_\Phi(x^2 + \tau^2)/\omega_c} \text{Erfi}(x\sqrt{C_\Phi/\omega_c})$  and  $h_2(x) = x\sqrt{\pi\omega_c/C_\Phi} e^{-\tau^2 C_\Phi/\omega_c}$ , where  $\text{Erfi}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2} dt$  is the imaginary error function.
- [13] In computing the various parameters self consistently, one finds that the positivity enforcing  $\Theta$  function in Eq. (26) can be dropped with very little ( $\sim 5\%$ ) error.
- [14] This gives a bandwidth of  $\sim 2$  eV that is typical of cuprates.
- [15] C.M. Varma, P.B. Littlewood, S. Schmitt-Rink, E. Abrahams, and A.E. Ruckenstein, *Phys. Rev. Lett.* **63**, 1996 (1989).
- [16] This physical duality is to be understood in the sense that a highly elastic auxiliary FL over all energies, with very sharp features would lead to a small  $\Delta_0$  through Eq. (27), so that the ECFL spectrum would have a very large coefficient of the linear  $x$  term in Eq. (26) and hence appear incoherent.

## Extremely Correlated Fermi-Liquid Description of Normal-State ARPES in Cuprates

G.-H. Gweon,<sup>1,\*</sup> B. S. Shastry,<sup>1,†</sup> and G. D. Gu<sup>2</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, California 95064, USA*

<sup>2</sup>*Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, New York 11973, USA*  
(Received 13 April 2011; published 29 July 2011)

The normal-state single particle spectral function of the high temperature superconducting cuprates, measured by the angle-resolved photoelectron spectroscopy (ARPES), has been considered both anomalous and crucial to understand. Here, we report an unprecedented success of the new extremely correlated Fermi liquid theory by one of us [B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011)] to describe both laser and conventional synchrotron ARPES data (nodal cut at optimal doping) on  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$  and synchrotron data on  $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ . It fits all data sets with the same physical parameter values, satisfies the particle sum rule and successfully addresses two widely discussed kink anomalies in the dispersion.

DOI: 10.1103/PhysRevLett.107.056404

PACS numbers: 71.10.Ay, 74.25.Jb, 74.72.Gh, 79.60.-i

Angle resolved photoelectron spectroscopy (ARPES) was the first probe to provide a detailed view of the anomalous nature of high temperature cuprate superconductors, discovering unexpectedly broad spectra with intense and asymmetric tails that have remained an enduring mystery for the last two decades. Conventional data taken with high energy ( $\geq 15$  eV) photons from synchrotron light sources have recently been supplemented with laser ARPES data [1,2] from lower energy (6 or 7 eV) sources. The latter show considerably sharper features near the Fermi energy. A drastic possibility to account for this distinction is that the sudden approximation could break down for the smaller photon energies used in laser ARPES [3].

An important unanswered question is whether the results of the two spectroscopies could be reconciled in a single theoretical framework that does not abandon the sudden approximation. More broadly, can we understand the wide variety of observed lines shapes in a theoretical framework with a sound microscopic basis and a single set of parameters?

In this Letter, we confront a recent theory of extremely correlated Fermi liquids (ECFL) proposed by Shastry [4] with the above challenge. The new formalism is complex and requires considerable further effort to yield numerical results in low dimensions. In the limit of high enough dimensions, however, a remarkably simple expression for the Green's function emerges; it is significantly different from the standard Fermi-liquid Dyson form, while satisfying the usual sum rules. We use this simple version of ECFL Green's function in this Letter, motivated by the attractive spectral shapes produced with very few parameters [4]. In this Letter we show that already the simplest version of the ECFL theory, with very few parameters, is very successful in detailed fitting of a wide variety of normal-state cuprate ARPES line shapes. Interesting predictions are made for the higher temperature spectral line skew.

Our focus in this Letter is on the data of optimally doped  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$  (Bi2212) and  $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$  (LSCO) superconductors in the normal state, taken with  $\vec{k}$  along the nodal direction connecting (0, 0) to  $(\pi/a, \pi/a)$ . Most of the data is taken from the published literature, while some original data are also presented (Bi2212 data in Figs. 4 and 5). Our sample is an optimally doped Bi2212 ( $T_c = 91$  K), grown by the floating zone method at the Brookhaven National Laboratory (BNL), and was measured at the Stanford Synchrotron Radiation Lightsource (SSRL) beam line 5-4 using 25 eV photons. The resolutions are 15 meV (energy) and  $0.3^\circ$  (angle).

*Line shape model.*—The ECFL spectral function is given as a product of an auxiliary Fermi-liquid (AFL) spectral function  $A_{\text{FL}}(\vec{k}, \omega)$  and a second frequency dependent “caparison” factor [4,5]:

$$A(\vec{k}, \omega) = A_{\text{FL}}(\vec{k}, \omega) \left( 1 - \frac{n}{2} + \frac{n^2}{4} \cdot \frac{\xi_{\vec{k}} - \omega}{\Delta_0} \right)_+, \quad (1)$$

where  $n$  is the number of electrons per  $\text{CuO}_2$  unit cell,  $(X)_+ \equiv \max(X, 0)$ ,  $\xi_{\vec{k}} = (1 - \frac{n}{2})\varepsilon(\vec{k})$ , where  $\varepsilon(\vec{k})$  is the bare one-electron band dispersion (see later). Here,  $A_{\text{FL}}(\vec{k}, \omega) = \frac{1}{\pi} \text{Im} \frac{1}{\omega - \xi_{\vec{k}} - \Phi(\omega)}$  with

$$\text{Im} \Phi(\omega) = \frac{\omega^2 + \tau^2}{\Omega_0} \exp\left(-\frac{\omega^2 + \tau^2}{\omega_0^2}\right) + \eta, \quad (2)$$

where  $\tau = \pi k_B T$ ,  $T$  is the temperature, and  $\omega$  is to be understood as  $\omega - i0^+$ . Here,  $\omega_0$  is the AFL energy scale (i.e., high  $\omega$  cutoff), and  $\Omega_0$  governs the lifetime, and, by causality, the quasiparticle weight (i.e., the wave function renormalization) of the AFL,  $Z_{\text{FL}} = (1 + \frac{\omega_0}{\sqrt{\pi}\Omega_0})^{-1}$ , as identified from  $\text{Re}\Phi$  [6].

The ECFL energy scale  $\Delta_0$  measures the “average intrinsic inelasticity” of the AFL. It is given [4] as

$$\Delta_0 = \int_{-\infty}^{\infty} d\omega f(\omega) \langle A_{FL}(\vec{k}, \omega) (\xi_{\vec{k}} - \omega) \rangle_{BZ}, \quad (3)$$

where  $\langle \cdot \rangle_{BZ}$  denotes averaging over the first Brillouin zone.

The parameters that enter this description are now listed. The “primary parameters” defining the ECFL fit consist of the dispersion  $\xi_{\vec{k}}$  taken from band theory, the density  $n$ , temperature  $T$ , and the AFL parameters  $\Delta_0$ ,  $\omega_0$ ,  $Z_{FL}$ ,  $\Omega_0$ . Of the last four parameters, only two are free parameters. For instance,  $\omega_0$  and  $Z_{FL}$  can be taken as free parameters, and  $\Omega_0$  and  $\Delta_0$  can be calculated using the equation for  $Z_{FL}$  and Eq. (3), respectively.

The parameter  $\eta$  in Eq. (2) is an additional “secondary parameter” [7] with respect to the ECFL theory [4]. Its origin is in impurity scattering as argued in [8], and additionally, in scattering with surface imperfections. Our fits determine  $\eta \approx 0.03$  eV for laser ARPES and  $\eta \approx 0.15$  eV for conventional ARPES. Greater penetration depth in laser ARPES suggests that it should be less sensitive to surface imperfections, thereby yielding a smaller  $\eta$ . We therefore propose that this parameter summarizes the effective sample quality in different experiments. The difference in line shapes arising from these values of  $\eta$  is demonstrated in Figs. 1(b) and 1(c).

Our strategy is to fix a common set of intrinsic parameters for all the materials, and allow  $\eta$  to be determined separately for each class of data. The most time consuming part is the calculation of  $\Delta_0$ , the results of which are summarized in Fig. 1(a).

In our line shape analysis (i) we first set  $n = 0.85$ , corresponding to the optimal doping. (ii) Here  $\xi_{\vec{k}}$  is taken to be the unrenormalized band dispersion, taken from the literature [9], and then scaled to fit the observed occupied band width, 1.5 eV, of the Bi2212 ARPES result [10,11]. (iii) We choose  $Z_{FL} = 1/3$ , to account for the dispersion renormalization due to the high energy kink [10,12], which in this theory is caused by the energy scale  $\omega_0$  (cf. Fig. 5). (iv) Finally, in all simulations, we include the finite energy

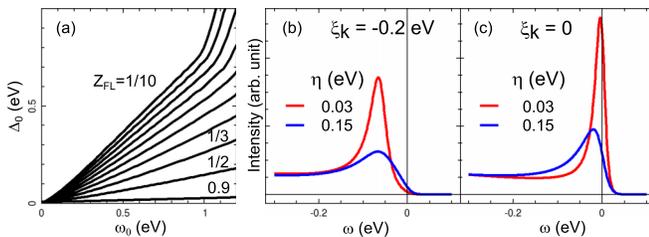


FIG. 1 (color online). (a)  $\Delta_0$  as a function of  $\omega_0$  for various  $Z_{FL}$ . Other primary ECFL parameter values are  $n = 0.85$ ,  $T = 100$  K, and  $\xi_{\vec{k}}$  as described in the text. A small  $\eta$  value, 0.010 eV, was used for this plot, which is used as a “lookup table” during the fit. (b),(c) Examples of the spectral function calculated with different values of the effective sample quality parameter  $\eta$ . See the caption of the next figure for parameter values used. The instrumental energy broadening of 10 meV (FWHM) is included.

resolution effect and the finite angle resolution effect as a combined Gaussian broadening (10 meV FWHM for laser ARPES and 25 meV FWHM for conventional ARPES) in energy [13].

*Line shape fit for laser ARPES.*—Figure 2 shows the fit of the laser ARPES data with the ECFL line shape. These fits were made using a procedure that is somewhat more restrictive than that in the recent work of Casey and Anderson [14,15] invoking the x-ray edge singularity ideas of Doniach-Sunjic, Anderson-Yuval, and Nozieres-de Dominicis [16] (CADS): we are using global, rather than perspectrum, fit parameters. However, our fit is somewhat less restricted than other fits shown in this Letter: here we allow a small variation of  $\xi_{\vec{k}}$  as in Ref. [14]. We find an excellent fit quality, at least comparable to CADS [14]. The gray line in panel (a) shows our calculation for  $k > k_F$ . Our expectation is that, were the data for  $k > k_F$  available, we would find a reasonable fit in this  $k$  region as well [17], as for other data sets below.

*Line shape fit for conventional ARPES.*—We find that the magnitude of the parameter  $\omega_0$  (0.5 eV) determined from the fit of the sharp laser data works very well also for the conventional ARPES data [18]. Thus, all parameters other than  $\eta$  are fixed, with one small exception in Fig. 4(d), where a slight change in  $\omega_0$  produces a much better fit over a larger energy range for LSCO.

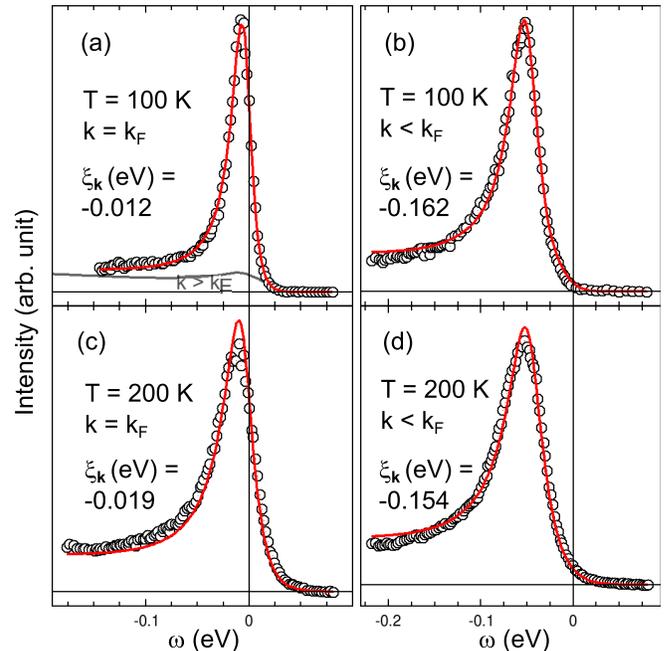


FIG. 2 (color online). Laser ARPES data (symbols, Bi2212) from Ref. [14] fit with the ECFL line shape (red lines). The free parameters of the fit were  $\omega_0$  (0.5 eV),  $\eta$  (0.032 eV), and  $\xi_{\vec{k}}$  (shown). Fixed parameters:  $n$  (0.85),  $Z_{FL}$  (1/3). Derived parameters:  $\Delta_0$  (0.12 eV),  $\Omega_0$  (0.14 eV). Other than  $\eta$  and  $\xi_{\vec{k}}$ , the same parameters are used elsewhere in the Letter. In (a), the gray line corresponds to the theoretical curve with  $\xi_{\vec{k}} = 0.15$  eV.

Figure 3 shows our fit of the data in Ref. [19] with a single free parameter  $\eta$ . The amount of the “extrinsic background” in ARPES is an issue of importance [20–22], especially when analyzing the conventional ARPES data. Here we fit the bg subtracted data, as well as the raw data (panel *d*). For subtracting the background, we use an often-used procedure [22,23] of equating the background to a fraction (“background scaling factor”) of the data far beyond the Fermi surface crossing ( $k = k_{10}$  for this data set). The background scaling factor, 1/2 for this figure, is determined to be the maximum value for which the resulting intensity is not negative. As shown in the panel (d), the ECFL fit remains good by adjusting  $\eta$ , whether or not the extrinsic background is subtracted. In contrast, we find that the CADS theory, notwithstanding its notable successes [14,15], cannot cope with even the background subtracted data [Fig. 3(e)], giving too steep a fall off towards the left. Likewise, the MFL fits [8,24] have been shown to compare well with the data only after substantial background subtraction [23,25].

Our own data on Bi2212 data, taken at  $T_c$  and well above  $T_c$ , covering a similar temperature range as the laser data of Fig. 2, can be fit equally well with the same background subtraction procedure, i.e., with the background scaling factor (1/2), as shown in Fig. 4.

We also find that the data for a lower- $T_c$  cuprate LSCO can be fit very well with the same intrinsic parameters. Here, we shall discuss only the  $k = k_F$  data for brevity.

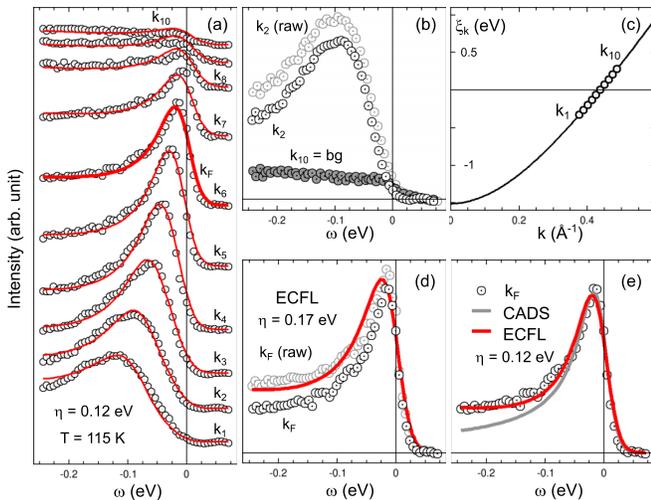


FIG. 3 (color online). Conventional ARPES data (Bi2212) fit with the ECFL line shape. The data are from Ref. [19] ( $T_c = 90$  K). (a) The data (symbols) and the fit (red lines) are shifted vertically by the same amount for ease of view. (b) An example of the raw data and the fit data is shown for  $k_2$ . The background (bg) spectrum (see text) was subtracted from each raw data, and the resulting data, shown in (a), are then fit. (c) The fixed  $\xi_k$  parameters used for the fit. Thus, in this figure,  $\eta$  is the only fit parameter (cf. Fig. 2 caption). (d) Raw data at  $k = k_F$  fit with a somewhat greater  $\eta$  value. (e) The current fit compared with a fit using the CADS line shape.

In this case, we determine that the background scaling factor be 1. The subtracted background data [26,27] is shown as the gray curve in Fig. 4(c). Given their weak superconductivity features [26,27], these LSCO data are taken to represent the normal-state property even if the temperature is slightly lower than  $T_c$ . As for the Bi2212 case, the data can be fit well even without the background subtraction, if a somewhat greater  $\eta$  value ( $\approx 0.17$  eV) is used. It is clear, from Fig. 4(c), that the data at a temperature as low as 25 K can be fit very well with the ECFL line shape. In addition, in working with LSCO line shapes, we noticed a steady and rapid rise in intensity beyond  $-\omega = 0.25$  eV, a behavior different from that of Bi2212. We leave the full discussion of this nonuniversal behavior for future work. However, we find it exceptional that the current theory is able to describe the line shape of LSCO up to very high energy, as shown in Fig. 4(d).

*Kinks in the spectra.*—The two independent energy scales  $\omega_0$  and  $\Delta_0$  are determined from our fit as  $\sim 0.5$  eV and  $\sim 0.1$  eV. These are natural candidates for the two main dispersion anomalies in the cuprates [12,28] as in Fig. 5. Well-defined energy distribution curve (EDC: intensity curve at a fixed  $\vec{k}$  value) peaks disappear in a wide energy from  $\sim 0.3$  eV to  $\sim 1$  eV, as observed experimentally for the high energy kink [10,12]. As this feature

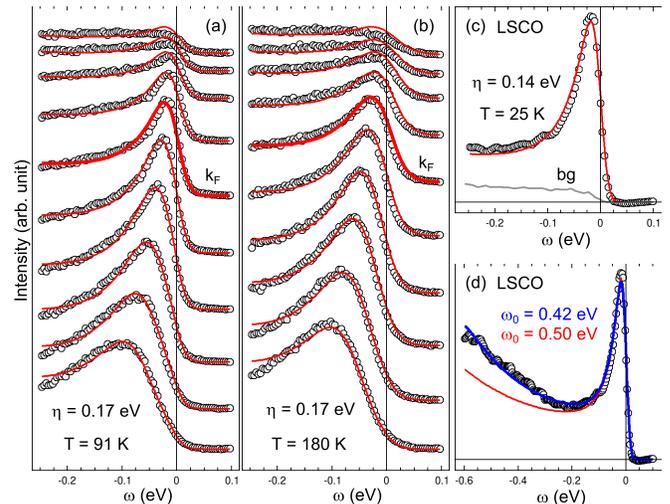


FIG. 4 (color online). Conventional ARPES data, including our own (a),(b), fit with the ECFL line shape. The procedure used to fit these data are identical with those of the previous figure, i.e., a fit with a single free parameter  $\eta$ , with (d) being a single exception. (a),(b), Optimally doped Bi2212 ( $T_c = 91$  K). (c) Optimally doped LSCO data [26,27]. (d) A test fit up to 0.6 eV for the LSCO data with a small change to  $\omega_0$  for the same data as in (c) but over a wider energy range. By changing  $\omega_0$  slightly from 0.50 to 0.42 eV, we see that an excellent fit up to 0.6 eV is found. The LSCO data, as far as we are aware, is fit only by the ECFL theory, since an energy dependence rising linearly for occupied states occurs naturally and uniquely in the ECFL spectral function.

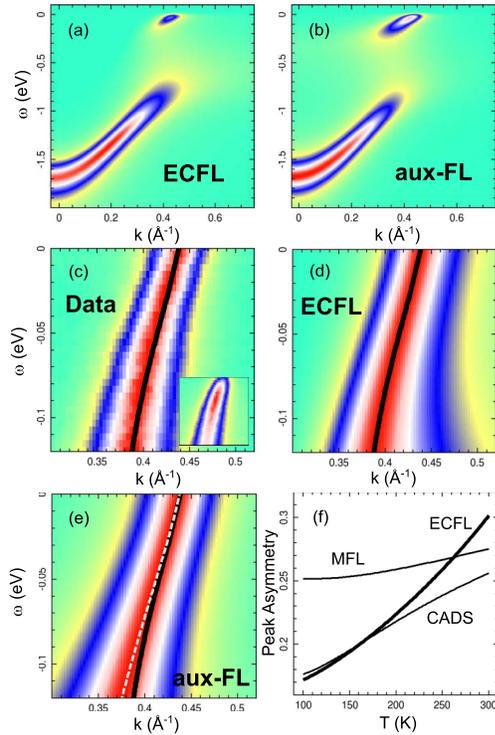


FIG. 5 (color online). Image plots of the spectral function for (a) the ECFL theory ( $\eta = 0.17$  eV), and (b) that of the auxiliary FL. (c) The data of Fig. 4(a) before (inset) and after (main) MDC normalization, by which each MDC is scaled and shifted to have minimum 0 (green) and maximum 1 (red). Blue corresponds to  $1/2$ . (d) The near- $E_F$  part of the ECFL spectral function of (a), after MDC-normalization with MDC peak positions traced by black line. The black line in (c) is from (d). (e) The near- $E_F$  part of the AFL spectral function of (b) after MDC-normalization. The MDC peak positions are traced by gray dashed line, while the black line is from (d). The bending such as shown by the black line here is commonly referred to as the kink. (f) The temperature dependence of the peak asymmetry compared for three different theories, rising as  $T^2$  for ECFL. Theory parameters for the calculation, apart from  $T$ , are taken from the fit of Fig. 2(a) for the ECFL, from the equivalent fit of Ref. [14] for the CADs, and from Ref. [25] for the MFL.

already exists in AFL, it cannot be associated with  $\Delta_0$  but rather with  $\omega_0$ . The (numerical) dynamical mean field theory [29] can already account for this feature as can the present ECFL (analytical) theory.

Turning to the low energy ARPES kink at  $\sim 70$  meV, Figs. 5(c)–5(e) illustrate the observed weak dispersion anomaly in the normal-state data (c), reproduced in the ECFL theory (d) but not in the AFL theory (e). Here we use a visualization method for momentum distribution curve (MDC: intensity curve at a fixed  $\omega$  value), an object discussed primarily for low energy kinks. Thus this feature originates from the scale  $\Delta_0$ , it causes an increased asymmetry and the (blue) shift of the peak to high hole energy, when the third term in the caparison factor ( $\frac{\mu^2}{4} \frac{\xi_{\vec{k}} - \omega}{\Delta_0}$ ) of

Eq. (1) becomes important. To our knowledge, the ECFL theory is a unique analytical theory that has both these kink features arising from purely electronic (extreme) correlations.

In Fig. 5(f), we show the temperature dependence of the dimensionless peak skew or asymmetry, defined as  $(HL - HR)/(HL + HR)$ , where HR (HL) is the half-width at half maximum on the right (left) side of the peak. The predicted  $T$ -dependent asymmetry, predicted even greater for  $\eta \approx 0.15$  eV (synchrotron data; not shown), would be interesting to explore in the future.

Further work is necessary to refine the picture suggested in this Letter. For example, as  $-\xi_{\vec{k}}$  increases, the line shape becomes somewhat too asymmetric. Work is also in progress to apply the theory to two particle response as seen, e.g., in optical conductivity. We have checked that the bubble approximation (conductivity as a product of two  $G$ 's) shows an agreement in the order of magnitude of the frequency scale and the conductivity.

*Conclusions.*—We have shown that it is possible to understand both ARPES data sets (laser or conventional) comprehensively, with identical physical parameters. Work going beyond the nodal cut and the optimal doping value is in progress. The theory is very tolerant of the uncertainty in the background subtraction for the conventional ARPES data. Additionally, the theory satisfies the global particle sum rule, and contains two interdependent energy scales ( $\omega_0$  and  $\Delta_0$ ) that correspond well to the energy scales of the two kinks. Thus the simplest version of the ECFL theory using a small number of parameters, provides a framework to understand the ARPES line shape data for the normal state of the cuprates: it works extremely well across techniques, samples, and temperatures.

Portions of this research were carried out at the SSRL, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the U.S. DOE Office of Science by Stanford University, and at the BNL, supported by the U.S. DOE under Grant No. DE-AC02-98CH10886. The work by GHG was supported partially by COR-FRG at UC Santa Cruz. BSS was supported by U.S. DOE under Grant No. FG02-06ER46319.

\*gweon@ucsc.edu

†sriram@physics.ucsc.edu

- [1] J.D. Koralek *et al.*, *Phys. Rev. Lett.* **96**, 017005 (2006).
- [2] W. Zhang *et al.*, *Phys. Rev. Lett.* **100**, 107002 (2008).
- [3] In the sudden approximation, the ARPES intensity  $I(\vec{k}, \omega) = M(\vec{k})A(\vec{k}, \omega)f(\omega)$ . Here,  $\vec{k}$  is the momentum and  $-\omega$  is the energy of the final  $N - 1$  state (as we set  $\hbar = 1$  in this Letter), where  $N$  is the number of electrons in the initial state.  $A(\vec{k}, \omega) = \text{Im} G(\vec{k}, \omega - i0^+)/\pi$  is the single particle spectral function,  $M(\vec{k})$  is the dipole matrix element, constant for an EDC, and  $f(\omega)$  is the Fermi-Dirac function. At the chemical potential  $\omega \equiv 0$ .

- [4] B. S. Shastry, this issue, Phys. Rev. Lett. **107**, 056403 (2011).
- [5] B. S. Shastry, arXiv:1104.2633.
- [6]  $\text{Re}\Phi(\omega) = -1/(\sqrt{\pi}\Omega_0)\exp(-\frac{\tau^2}{\omega_0^2})[\omega_0\omega - 2(\omega^2 + \tau^2)D(\frac{\omega}{\omega_0})]$ , where  $D(x) = (\sqrt{\pi}/2)e^{-x^2}\text{erfi}(x)$  is the Dawson function.  $C_\Phi = 1/(\pi\Omega_0)$  and  $\omega_c = C_\Phi\omega_0^2$ , where  $C_\Phi$  and  $\omega_c$  are parameters defined in [4].
- [7] By secondary, we mean “intrinsic [i.e., part of  $A(\vec{k}, \omega)$ ] but dependent on the effective sample quality.”
- [8] E. Abrahams and C. M. Varma, Proc. Natl. Acad. Sci. U.S.A. **97**, 5714 (2000).
- [9] R. S. Markiewicz *et al.*, Phys. Rev. B **72**, 054519 (2005).
- [10] W. Meevasana *et al.*, Phys. Rev. B **75**, 174506 (2007).
- [11] Also, a minor  $k$  shift of the theory was necessary ( $\sim 0.02 \text{ \AA}$ ) to match the slightly different  $k_F$  values reported in different experiments.
- [12] J. Graf *et al.*, Phys. Rev. Lett. **98**, 067004 (2007).
- [13] The minor effect of the finite angular resolution, significant only for the conventional ARPES data, can be modeled well by an energy broadening.
- [14] P. A. Casey *et al.*, Nature Phys. **4**, 210 (2008).
- [15] P. W. Anderson, Nature Phys. **2**, 626 (2006).
- [16] S. Doniach and M. Sunjic, J. Phys. C **3**, 285 (1970); P. Nozieres and C. T. De Dominicis, Phys. Rev. **178**, 1097 (1969); G. Yuval and P. W. Anderson, Phys. Rev. B **1**, 1522 (1970).
- [17] J. Koralek, Ph.D. thesis, University of Colorado, 2006.
- [18] When it is allowed to vary but  $k$  independent, it is 0.5 eV within  $\pm \sim 10\%$ .
- [19] A. Kaminski *et al.*, Phys. Rev. Lett. **86**, 1070 (2001).
- [20] C. G. Olson *et al.*, Phys. Rev. B **42**, 381 (1990).
- [21] G. Gweon, J. W. Allen, and J. D. Denlinger, Phys. Rev. B **68**, 195117 (2003).
- [22] A. Kaminski *et al.*, Phys. Rev. B **69**, 212509 (2004).
- [23] T. Valla *et al.*, Science **285**, 2110 (1999).
- [24] C. M. Varma *et al.*, Phys. Rev. Lett. **63**, 1996 (1989).
- [25] A. Kaminski *et al.*, Phys. Rev. B **71**, 014517 (2005).
- [26] T. Yoshida *et al.*, J. Phys. Condens. Matter **19**, 125209 (2007).
- [27] T. Yoshida, Ph.D. thesis, University of Tokyo, 2001.
- [28] A. Lanzara *et al.*, Nature (London) **412**, 510 (2001).
- [29] K. Byczuk *et al.*, Nature Phys. **3**, 168 (2007).



## Anatomy of the self-energy

B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*

(Received 13 April 2011; published 11 October 2011)

The general problem of representing the Greens function  $G(k, z)$  in terms of self-energy in field theories lacking Wick's theorem is considered. A simple construction shows that a Dyson-like representation with a self-energy  $\Sigma(k, z)$  is always possible, provided we start with a spectral representation for  $G(k, z)$  for finite-sized systems and take the thermodynamic limit. The self-energy itself can then be iteratively expressed in terms of another higher order self-energy, capturing the spirit of Mori's formulation. We further discuss alternative and more general forms of  $G(k, z)$  that are possible. In particular, a recent theory, by the author, of extremely correlated Fermi liquids at density  $n$ , for Gutzwiller projected noncanonical Fermi operators, obtains a new form of the Greens function:  $G(k, z) = [(1 - \frac{n}{2}) + \Psi(k, z)]/[z - \hat{E}_k - \Phi(k, z)]$ , with a pair of self-energies  $\Phi(z)$  and  $\Psi(z)$ . Its relationship with the Dyson form is explored. A simple version of the two-self-energies model was shown recently to successfully fit several data sets of photoemission line shapes in cuprates. We provide details of the unusual spectral line shapes that arise in this model, with the characteristic skewed shape depending upon a single parameter. The energy distribution curve (EDC) and momentum distribution curve (MDC) line shapes are shown to be skewed in opposite directions, and provide a testable prediction of the theory.

DOI: [10.1103/PhysRevB.84.165112](https://doi.org/10.1103/PhysRevB.84.165112)

PACS number(s): 79.60.-i

### I. INTRODUCTION

Our work explores the representation of the Greens function  $G(k, z)$  of a particle in field theories without Wick's theorem. While Wick's theorem for bosons and fermions automatically gives rise to the Dyson form of self-energy, its absence for noncanonical, i.e., general operators (other than bosons of fermions), leads to a conundrum that is poorly understood. This work addresses a particular type of noncanonical theory originating from Gutzwiller projection of electrons on a lattice.

The representation of the propagator or the Greens function in terms of its Dyson self-energy is a fundamental paradigm of standard interacting relativistic and nonrelativistic field theories. The structure of this representation and the generation of approximations for the self-energy in terms of Feynman diagrams, the vertex function, or higher order Greens functions form the dominant part of existing literature of many-particle physics.

In the context of extending these studies to extremely large and singular repulsive interactions, termed extreme correlations,<sup>1</sup> one needs to deal with noncanonical electrons. A standard noncanonical problem involves the Hubbard operators<sup>2</sup>  $X_j^{a,b}$  located at sites  $j$  of a lattice. These are "graded" projection operators with label  $a$  representing the three allowed local configurations  $0, \uparrow$ , and  $\downarrow$ . Among these operators,  $X_j^{0\sigma}$  and  $X_j^{\sigma 0}$  are fermionic destruction- and creation-type objects. Their Greens function is measured directly in angle-resolved photoemission experiments (ARPES) on certain experimental systems embodying extreme correlations, including the high-temperature superconductors.<sup>3</sup>

Quite recently, the author has formulated in Ref. 1, the theory of an extremely correlated Fermi liquid (ECFL) state of the  $t$ - $J$  model, where he has found another type of representation for the Greens function with a pair of self-energies [see Eq. (4) below] by using the nonperturbative Schwinger approach of source fields to depict the equations of motion. The use of more general forms of Greens functions is not completely new, there are examples in literature of multiple self-energies in

Refs. 4 and 5. The physics of extreme correlations treated here is based on nonperturbative considerations without obvious parallels in weak or intermediate coupling problems. It leads to the two-self-energies form, Eq. (4), whose distinctive signatures are strikingly different from those of Fermi liquids.

The technical details of the construction in Ref. 1 are intricate and require the processing of two parallel hierarchies for the two self-energies. A separate paper is in preparation detailing the involved technical details and the calculation in Ref. 1. Further background details of the notation, definitions, and sum rules satisfied by the Greens functions for extreme correlations, and its analyticity can be found in the earlier publication.<sup>2</sup> A suggestive functional form of the Greens function, Eq. (37), emerges from Ref. 1, by making the assumption of momentum independence of the two self-energies, valid in high dimensions. It satisfies the number sum rule and the total particle weight integrates to unity in each state. In Ref. 3, Gweon, Gu, and the author have shown that several experimental data sets on ARPES by different groups using both the traditional synchrotron light source and a laser light source can be reconciled very well with the line shape in Eqs. (37) and (60). This is the first satisfactory functional form that has been found to fit both laser and synchrotron data and to work very well with few adjustable parameters.

Therefore a major objective of this paper is to elucidate the detailed form of the spectral lines that emerged from the above simple version of the ECFL theory in Ref. 1 and successfully employed to understand experiments in Ref. 3. Our hope is that this detailed analysis will familiarize readers with the nuances of the new spectral function, and thereby facilitate ARPES line-shape analysis of further experiments on high-temperature superconductors and other materials, in a manner analogous to that in Ref. 3.

For the above purpose, we recall that in a lattice of finite (say small) number of sites, the state space is finite dimensional and hence the Greens function for arbitrarily complicated objects can be computed by numerical means, leading to rational functions of the complex frequency  $z$  as in Eq. (5) below.

We begin by studying this representation and see how the Dyson representation arises; we find that the two-self-energies representation, Eq. (4), is also quite natural from this view point. We further study the infinite size limits where the poles coalesce to give cuts in the complex  $z$  plane. We adopt a phenomenological model for an underlying auxiliary Fermi liquid (aux-FL) self-energy, enabling us to display analytic expressions for the Greens function. We provide a detailed perspective on the representation in Eq. (37), namely, the location of the poles and the subtle differences from a standard Fermi liquid.

Another result in this paper is to show that a Dyson-like representation with a self-energy  $\Sigma(k, z)$  is always possible, provided we start with a spectral representation. The self-energy itself can then be iteratively expressed in terms of another higher order self-energy. This hierarchical result is cast in the same form as the Mori formalism. While the Mori formalism is very abstract, and expressed in terms of projection operators, we can go beyond it in a certain sense. By working with standard spectral representation, we show that it is possible to express the higher self-energy spectral functions in terms of the lower ones, leading to an explicit hierarchy. Our construction completely bypasses the Mori projection operators, and should be useful in throwing light on the latter.

The plan of the paper is as follows. In Secs. II and III, we note the spectral representation and study the Greens function as a rational function of complex frequency  $z$  for a finite system. In Sec. IV, we note the representation in the limit of infinite size and introduce the high-dimensional expression with two self-energies. The detailed structure of the characteristic line shape as in Eq. (60) is discussed, and its dependence on physical parameters displayed with the help of a numerical example. An explicit example of the auxiliary Fermi-liquid Greens function is provided, and typical values of the parameters are argued for. In Sec. V, the line shapes in EDC and MDC are displayed in detail, in order to bring out the specific signatures of the theory, namely, a skew in the spectrum arising from the caparison factor in Ref. 1 and Eq. (37). In Sec. VI, we discuss the amusing connection with higher order self-energies of the type that Mori's formalism yields, but at a much more explicit level than what is available in the literature.

## II. SPECTRAL REPRESENTATION OF THE GREENS FUNCTION

Let us begin with the spectral representation<sup>10</sup> of the Matsubara Greens function at finite temperatures given by

$$G(k, z) = \int dx \frac{\rho_G(k, x)}{z - x}, \quad (1)$$

where  $G$  is the Greens function at a fixed wave vector  $k$ ,  $\rho_G$  is its spectral density, and the integration range is  $-\infty \leq x \leq \infty$ . To simplify notation, we call the Greens function as  $G(k, z)$ , the same object was denoted by  $\mathcal{G}(k, z)$  in Ref. 1. The index  $k$  can be also replaced by a spatial index when dealing with a local Greens function. The spectral function  $\rho_G(k, x)$  in most problems of interest in condensed matter physics has a compact support, so that  $G(k, z)$  has "reasonable" behavior in the complex  $z$  plane, with an asymptotic  $1/z$  fall off, and, apart

from a branch cut on a portion of the real line, it is analytic. The frequency  $z$  is either fermionic or bosonic depending on the statistics of the underlying particles. The spectral function is given by the standard formula<sup>2,10</sup>

$$\rho_G(k, x) = \sum_{\alpha, \beta} |\langle \alpha | A(k) | \beta \rangle|^2 (p_\alpha + p_\beta) \delta(x + \varepsilon_\alpha - \varepsilon_\beta), \quad (2)$$

where  $A(k)$  is the destruction operator,  $p_\alpha$  is the Boltzmann probability of the state  $\alpha$  given by  $e^{-\beta \varepsilon_\alpha} / Z$ , and  $\varepsilon_\alpha$  is the eigenvalue of the grand Hamiltonian of the system  $K = H - \mu \hat{N}$ . In the case of canonical particles,  $A(k)$  is the usual Fermi or Bose destruction operator. In Ref. 1, noncanonical Hubbard 'X' operators are considered; we will not require any detailed information about them here except that the anticommutator  $\{A, A^\dagger\}$  is not unity, but rather an object with a known expectation value  $(1 - n/2)$ , in terms of the dimensionless particle density  $n$ .

We consider two alternate representations of the Greens function in terms of the complex frequency  $z$  that are available in many-body physics: (a) for canonical bosons or fermions, the Dyson representation in terms of a single self-energy  $\Sigma(z)$  and (b) for noncanonical particles, a novel form proposed recently by the author with two self-energy type objects  $\Phi(z)$  and  $\Psi(z)$ :

$$G(k, z) = \frac{a_G}{z - \hat{E}_k - \Sigma(k, z)} \quad (\text{Dyson}) \quad (3)$$

$$= \frac{a_G + \Psi(k, z)}{z - \hat{E}_k - \Phi(k, z)} \quad (\text{ECFL}). \quad (4)$$

For canonical objects,  $a_G = 1$  and for Hubbard operators in the ECFL we write  $a_G = 1 - n/2$ . We start below from a finite-size system, where the Greens function is a meromorphic function expressible as the sum over isolated poles in the complex frequency plane with given residues. In fact, it is a rational function as well, expressible as the ratio of two polynomials. Using simple arguments, we will see that the above two representations in Eqs. (3) and (4) are both natural ways of proceeding with the self-energy concept. In the limit of a large system, the poles coalesce to give us cuts in the complex frequency plane with specific spectral densities. In this limit, we display the equations relating the different spectral functions.

## III. FINITE-SYSTEM GREENS FUNCTION

We drop the explicit mention of the wave vector  $k$ , and start with the case of a finite-sized system, where we may diagonalize the system exactly and assemble the Greens function from the matrix elements of the operators  $A$  and the eigenenergies as in Eq. (2). We see that  $\rho_G$  is a sum over say  $m$  delta functions located at the eigenenergies  $E_j$  (assumed distinct), so we can write the meromorphic representation

$$\tilde{G}(z) = \sum_{j=1}^m \frac{a_j}{z - E_j}. \quad (5)$$

The overbar in  $\bar{G}(z)$  is to emphasize that we are dealing with the finite-sized version of the Greens function  $G(z)$ . Here,  $a_j$  and  $E_j$  constitute  $2m$  known real parameters. The sum

$$\sum_{j=1}^m a_j = a_G, \quad (6)$$

where  $a_G = 1$  for canonical objects and we denote  $a_G = 1 - n/2$  for the noncanonical case of ECFL. In the infinite-size limit, we set  $\bar{G}(z) \rightarrow G(z)$ . It is clear that for  $z \gg \{E_j\}_{\max}$ , we get the asymptotic behavior  $\bar{G} \rightarrow \frac{a_G}{z}$ , and therefore  $\bar{G}$  is a rational function that may be expressed as the ratio of two polynomials in  $z$  of degrees  $m - 1$  and  $m$ :

$$\bar{G}(z) = a_G \frac{P_{m-1}(z)}{Q_m(z)}, \quad Q(z) = \prod_{j=1}^m (z - E_j),$$

$$P(z) = \prod_{r=1}^{m-1} (z - \gamma_r), \quad (7)$$

where the roots  $\gamma_r$  are expressible in terms of  $a_j$  and  $E_j$ . We use the convention that all polynomials  $Q_m$  have the coefficient of the leading power of  $z$  as unity, and the degree is indicated explicitly.

We now proceed to find the self-energy type expansion for  $\bar{G}$ , and for this purpose, multiplying Eq. (5) by  $z$  and rearranging we get the ‘‘equation of motion:’’

$$(z - \hat{E})\bar{G}(z) = a_G + \bar{I}(z), \quad (8)$$

where we introduced a mean energy  $\hat{E}$ :

$$\hat{E} = \frac{1}{a_G} \sum a_j E_j,$$

$$\bar{I}(z) = \sum_{j=1}^m \frac{a_j (E_j - \hat{E})}{z - E_j}, \quad (9)$$

so that asymptotically at large  $z$  we get  $\bar{I}(z) \sim O(1/z^2)$ . In standard theory,  $\hat{E}$  plays the role of the Hartree-Fock self-energy so that the remaining self-energy vanishes at high frequencies.<sup>12</sup> Motivated by the structure of the theory of extremely correlated Fermi systems,<sup>1</sup> we next introduce the basic decomposition

$$\bar{I}(z) = \bar{G}(z)\Phi(z) + \Psi(z), \quad (10)$$

where we have introduced two self-energy type functions  $\Phi(z)$  and  $\Psi(z)$  that will be determined next. Clearly, Eq. (10) leads immediately to the Greens function (4) (or Eq. (3), if we set  $\Psi \rightarrow 0$ ). The rationale for Eq. (10) lies in the fact that the function  $\bar{I}$  has the same poles as  $\bar{G}(z)$ . Thus it has a representation as a ratio of two polynomials:

$$\bar{I}(z) = i_0 \frac{R_{m-2}(z)}{Q_m(z)}, \quad (11)$$

with  $R_{m-2}$  a polynomial of degree  $m - 2$ ,  $i_0$  a suitable constant, and the same polynomial  $Q$  from Eq. (7), thereby it is natural to seek a proportionality with  $\bar{G}$  itself. If we drop  $\Psi$  and rename  $\Phi \rightarrow \Sigma$ , then this gives the usual Dyson self-energy  $\Sigma(z)$  determined uniquely using Eqs. (7) and (11) as

$$\Sigma(z) = \frac{i_0}{a_G} \frac{R_{m-2}(z)}{P_{m-1}(z)}. \quad (12)$$

Expression (10) offers a more general possibility, where  $\Phi(z)$  and  $\Psi(z)$  may be viewed as the quotient and remainder obtained by dividing  $\bar{I}(z)$  by  $\bar{G}(z)$ . It is straightforward to see that  $\Psi(z)$  and  $\Phi(z)$  are also rational functions expressible as ratios of two polynomials:

$$\Psi(z) = \psi_o \frac{K_{m-3}(z)}{D_{m-1}(z)}, \quad \Phi(z) = \phi_o \frac{L_{m-2}(z)}{D_{m-1}(z)}, \quad (13)$$

where  $K$ ,  $L$ , and  $D$  are polynomials of the displayed degree. Comparing the poles and the zeros of  $\bar{G}$  in Eq. (4) with Eqs. (12) and (7), we write down two equations:

$$a_G P_{m-1} = a_G D_{m-1} + \psi_o K_{m-3},$$

$$Q_m = (z - \bar{E})D_{m-1} - \phi_o L_{m-2}, \quad (14)$$

so that we may eliminate  $D$  and write an identity,

$$(z - \bar{E})P_{m-1} - Q_m = \frac{\psi_o}{a_G} (z - \bar{E})K_{m-3} + \phi_o L_{m-2}. \quad (15)$$

Here, the left-hand side is assumed known and we have two polynomials to determine from this equation. Therefore there are multiple solutions of this problem, and indeed setting  $K \rightarrow 0$  gives the Dyson form as a special case.

#### A. A simple example with two sites

The Greens function of the  $t$ - $J$  model at density  $n$  with  $J = 0$  and only two sites is a trivial problem that illustrates the two possibilities discussed above. The two quantum numbers  $k = 0, \pi$  correspond to the bonding and antibonding states with energies  $e_k = \mp t$ , and a simple calculation at a given  $k$  gives Eq. (5) as

$$\bar{G}(k, z) = \frac{a_1}{z - e_k} + \frac{a_2}{z + e_k}, \quad (16)$$

where  $z = i\omega_n + \mu$ ,  $a_2 = e^{\beta\mu} [1 + e^{\beta(\mu - e_k)}] / (2Z)$ ,  $a_1 = 1 - n/2 - a_2$ , and the grand partition function  $Z = 1 + 4e^{2\beta\mu} + 4e^{\beta\mu} \cosh(\beta t)$ . This can be readily expressed as

$$\bar{G}(k, z) = \frac{(1 - \frac{n}{2}) + \Psi(k, z)}{z - E_k - \Phi(k, z)}, \quad \Psi(k, z) = \frac{B_k}{z + E_k},$$

$$\Phi(k, z) = \frac{A_k}{z + E_k}, \quad (17)$$

where  $E_k$  is arbitrary,  $A_k = (E_k^2 - e_k^2)$ , and  $B_k = (1 - n/2)(\bar{E}_k - E_k)$  and with the first moment of energy  $\bar{E}_k = e_k(a_1 - a_2)/(1 - n/2)$ . As we expected, the functions  $\Psi, \Phi$  thus have a single pole, as opposed to  $\bar{G}$  with two poles. In this case the dynamics is rather trivial, so that the choice of  $E_k$  is free. If we set  $E_k = \bar{E}_k$ , the residue  $B_k$  vanishes and so the second form collapses.

#### B. Summary of analysis

In summary, guided by analyticity and the pole structure of  $\bar{G}(k, z)$ , we find it possible to go beyond the standard Dyson representation. However, we end up getting more freedom than we might have naively expected. This excess freedom is not unnatural, since we haven't yet discussed the microscopic origin of these two self-energies. The theory in Ref. 1 provides

an explicit expression for the two objects  $\Psi$  and  $\Phi$ , where a common linear functional differential operator  $\mathbf{L}$  generates these self-energies by acting upon different “seed” functions as in Eq. (7) of Ref. 1. The above discussion therefore provides some intuitive understanding of the novel form of the Greens function in Eq. (4), without actually providing an alternative derivation to that in Ref. 1.

#### IV. INFINITE-SYSTEM SPECTRAL DENSITIES AND RELATIONSHIPS

In the infinite-size limit, the various functions will be represented in terms of spectral densities obtained from the coalescing of the poles. Following Eq. (1), we will denote a general function

$$Q(z) = \int dx \frac{\rho_Q(x)}{z-x}, \quad (18)$$

where  $Q = \Sigma, \Phi, \Psi$  in terms of its density  $\rho_Q(x)$ . The density is given by  $\rho_Q(x) = (-\frac{1}{\pi})\Im m Q(x+i0^+)$ , as usual. In parallel to the discussion of Eq. (1), the assumption of a compact support of  $\rho_Q$  gives us well-behaved functions. We now turn to the objective of relating the spectral functions in the two representations discussed above.

##### A. Spectral representation for the Dyson self-energy

Let us start with Eq. (1) and the standard Dyson form (3) where we drop the overbar and study the infinite system function  $G(z)$ . We use the symbolic identity:

$$\frac{1}{x+i0^+} = \mathcal{P}\frac{1}{x} - i\pi\delta(x), \quad (19)$$

with real  $x$ ,  $\mathcal{P}$  denoting the principal value, and the Hilbert transform of a function  $f(u)$  is defined by

$$\mathcal{H}[f](x) = \mathcal{P} \int_{-\infty}^{\infty} dy \frac{f(y)}{x-y}. \quad (20)$$

We note the following standard result for completeness:

$$\rho_G(x) = a_G \frac{\rho_\Sigma(x)}{[\pi\rho_\Sigma(x)]^2 + [x - \hat{E} - \mathcal{H}[\rho_\Sigma](x)]^2}. \quad (21)$$

A more interesting inverse problem is to solve for  $\rho_\Sigma(x)$  given  $G(z)$ . Toward this end, we rewrite the Dyson equation as

$$\Sigma(z) = z - \hat{E} - \frac{a_G}{G(z)}, \quad (22)$$

where the self-energy vanishes asymptotically as  $1/z$ , provided the constant part, if any, is absorbed in  $\hat{E}$ . Therefore this object can be decomposed in the fashion of Eq. (18). We compare Eq. (22) with Eq. (18) with  $Q \rightarrow \Sigma$  and conclude that

$$\rho_\Sigma(x) = \frac{1}{\pi} \Im m \frac{a_G}{G(x+i0^+)} = \frac{a_G \rho_G(x)}{[\pi\rho_G(x)]^2 + [\Im e G(x)]^2}. \quad (23)$$

The real part can be found either by taking the Hilbert transform,

$$\Re e \Sigma(x) = \mathcal{H}[\rho_\Sigma](x), \quad (24)$$

or more directly as

$$\begin{aligned} \Re e \Sigma(x) &= x - \hat{E} - \Re e \frac{a_G}{G(x+i0^+)} \\ &= x - \hat{E} - \frac{a_G \Re e G(x)}{[\pi\rho_G(x)]^2 + [\Im e G(x)]^2}. \end{aligned} \quad (25)$$

##### B. Spectral representation for the ECFL self-energies

For the ECFL Greens function in Eq. (4), we set  $a_G = (1 - \frac{n}{2})$  and write  $\hat{E} \rightarrow \xi$  representing the single-particle energy measured from the chemical potential. We start with the expression:

$$G(\xi, z) = \frac{1}{z - \xi - \Phi(z)} \times \left[ \left(1 - \frac{n}{2}\right) + \Psi(z) \right], \quad (26)$$

and express it in terms of the two spectral functions  $\rho_\Psi$  and  $\rho_\Phi$ .<sup>13</sup> We can write spectral function  $\rho_G$ :

$$\begin{aligned} \rho_G(\xi, x) &= \frac{\rho_\Phi(x)}{[\pi\rho_\Phi(x)]^2 + [x - \xi - \mathcal{H}[\rho_\Phi](x)]^2} \\ &\times \left[ \left(1 - \frac{n}{2}\right) + \frac{\xi - x}{\Delta(\xi, x)} + \eta(\xi, x) \right], \end{aligned} \quad (27)$$

where  $\Delta(\xi, x)$  and the term  $\eta$  are defined as

$$\Delta(\xi, x) = -\frac{\rho_\Phi(\xi, x)}{\rho_\Psi(\xi, x)}, \quad (28)$$

$$\eta(\xi, x) = \mathcal{H}[\rho_\Psi](\xi, x) + \frac{1}{\Delta(\xi, x)} \mathcal{H}[\rho_\Phi](\xi, x). \quad (29)$$

The real part of  $G$  is also easily found as

$$\Re e G(\xi, x) = \frac{\left[ \left(1 - \frac{n}{2}\right) + \mathcal{H}[\rho_\Psi](\xi, x) \right] [x - \xi - \mathcal{H}[\rho_\Phi](\xi, x)] - \pi^2 \rho_\Psi(\xi, x) \rho_\Phi(\xi, x)}{[\pi\rho_\Phi(\xi, x)]^2 + [x - \xi - \mathcal{H}[\rho_\Phi](\xi, x)]^2}. \quad (30)$$

Thus given the ECFL form of the Greens function, we can calculate the Dyson Schwinger form of self-energy in a straightforward way using the inversion formula, Eqs. (23) and (24). The inverse problem of finding  $\Phi$  and  $\Psi$  from a given  $\Sigma$  or  $G$  is expected to be ill defined, as discussed above for finite systems.

The first Fermi liquid (FL) factor in Eq. (27) has a peak at the Fermi-liquid quasiparticle frequency  $E_k^{\text{FL}}$  for a given  $\xi_k$  given as the root of

$$E_k^{\text{FL}} - \xi_k - \mathcal{H}[\rho_\Phi](\xi_k, E_k^{\text{FL}}) = 0, \quad (31)$$

however,  $\rho_G$  itself has a slight shift in the peak due to the linear- $x$  dependence in the numerator. This is analyzed in detail in the next section for a model self-energy. At this solution  $x(\xi)$ , Eq. (30) gives a relation:

$$\rho_\Psi(\xi_k, E_k^{\text{FL}}) = -\rho_\Phi(\xi_k, E_k^{\text{FL}}) \times \Im e G(\xi_k, E_k^{\text{FL}}). \quad (32)$$

### C. High-dimensional ECFL model with $\vec{k}$ -independent self-energies and its Dyson representation

In this section, we illustrate the two self-energies and their relationships in the context of the recent work on the ECFL of Ref. 1, and in Ref. 3. Here, we study a model Greens function, proposed in Ref. 1 for the  $t$ - $J$  model, that should be suitable in high enough dimensions. It is sufficiently simple so that most calculations can be done analytically. The model Greens function satisfies the Luttinger-Ward sum rule<sup>6</sup> and thereby maintains the Fermi surface of the Fermi gas, but yields spectral functions that are qualitatively different from the Fermi liquid. This dichotomy is possible since it corresponds to a simple approximation within a formalism that is very far from the standard Dyson theory, as explained in the previous sections. Our aim in this section is to take this model Greens function of the ECFL and to express it in terms of the Dyson self-energy so as to provide a greater feel for the model.

Here, the two self-energies are taken to be frequency dependent but momentum independent, and by using the formalism of Ref. 1, they become related through  $\Delta_0$ , an important physical parameter of the theory:

$$\Psi(z) = -\frac{n^2}{4\Delta_0} \Phi(z). \quad (33)$$

The physical meaning of  $\Delta_0$  as the mean inelasticity of the auxiliary Fermi liquid (aux-FL) is emphasized in Ref. 1, and follows from Eq. (56). Thus  $\rho_\Psi = -\frac{n^2}{4\Delta_0} \rho_\Phi$ , and hence we get the simple result:<sup>11</sup>

$$G(\xi_k, z) = g(\xi_k, z) \left[ \left(1 - \frac{n}{2}\right) - \frac{n^2}{4\Delta_0} \Phi(z) \right]. \quad (34)$$

The auxiliary Fermi liquid has a Greens function  $g^{-1}(\xi_k, z) = z - \xi_k - \Phi(z)$ , where  $\xi_k$  is the electronic energy at wave vector  $k$  measured from the chemical potential  $\mu$ , and therefore we may write the model Greens function as

$$G(\xi_k, z) = \frac{n^2}{4\Delta_0} + \left( \frac{n^2}{4\Delta_0} \right) \frac{\varepsilon_0 + \xi_k - z}{z - \xi_k - \Phi(z)}, \quad (35)$$

where

$$\varepsilon_0 = \Delta_0 \frac{4}{n^2} \left(1 - \frac{n}{2}\right). \quad (36)$$

With  $\Gamma(x) = \pi \rho_\Phi(x)$ ,  $\Im e \Phi(x + i0^+) = \mathcal{H}[\rho_\Phi](x)$  and  $\varepsilon(\xi_k, x) \equiv [x - \xi_k - \mathcal{H}[\rho_\Phi](x)]$ , we can express the spectral function and the real part of the Greens function as

$$\rho_G(\xi_k, x) = \left( \frac{n^2}{4\pi \Delta_0} \right) \frac{\Gamma(x)}{\Gamma^2(x) + \varepsilon^2(\xi_k, x)} (\varepsilon_0 + \xi_k - x), \quad (37)$$

$$\Re e G(\xi_k, x) = \left( \frac{n^2}{4\Delta_0} \right) \left[ 1 + \frac{\varepsilon(\xi_k, x)(\varepsilon_0 + \xi_k - x)}{\Gamma^2(x) + \varepsilon^2(\xi_k, x)} \right]. \quad (38)$$

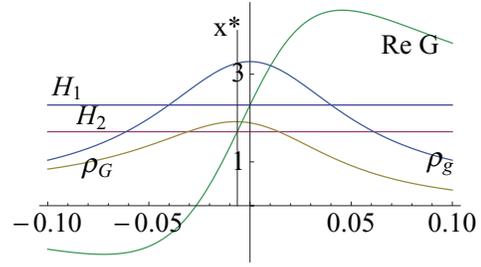


FIG. 1. (Color online) The density  $n = 0.85$ , temperature  $T = 600$  K,  $\Delta_0 = 0.0786$ , and parameters are set I of Eq. (54). At this rather high temperature, we can see the details of the spectral shape clearly. The vertical line is at  $x^* = E_{k_F}^*$ , this energy is the location of the peak of the physical spectral function  $\rho_G$ . Its leftward (i.e., red) shift relative to the Fermi-liquid peak at the chemical potential is clearly seen. The two horizontal lines specify the magnitude of the  $\Re e G(0, x)$  at  $x = 0$  (H1) and  $x = E_{k_F}^*$  (H2). The line H1 is at height  $n^2/(4\Delta_0)$  and H2 is at height  $n^2/(4\Delta_0)(1 - Z_k/2)$ .

The linear-frequency term in braces in Eq. (37) is termed the caparison factor in Ref. 1 and leads to significant features of the spectrum as discussed below. For completeness, we note the auxiliary Fermi-liquid part of the problem as

$$\rho_g(\xi_k, x) = \frac{1}{\pi} \frac{\Gamma(x)}{\Gamma^2(x) + \varepsilon^2(\xi_k, x)}. \quad (39)$$

In Fig. 1, we plot the above three functions for a model system described more fully in Sec. IV F.

### D. EDC or constant-wave-vector scans and energy dispersion

We first study the peak structure corresponding to fixing  $\vec{k}$  the wave vector and hence  $\xi_k$ , and sweeping the energy  $x$ . These give rise to the energy distribution curves, i.e., the EDC's. The aux-FL part has a peak at  $x = E_k^{\text{FL}}$  for a given  $\xi_k$ , as in standard FL theory from solving for the roots of Eq. (31). For  $k \sim k_F$ , we find  $E_k^{\text{FL}} = \xi_k Z_k$  with the momentum-independent self-energy  $\Phi(z)$ , where

$$Z_k = \lim_{x \rightarrow E_k^{\text{FL}}} [1 - \partial \Im e \Phi(x) / \partial x]^{-1}. \quad (40)$$

Expanding around this solution, we write

$$\varepsilon(\xi_k, x) \sim \frac{1}{Z_k} (x - Z_k \xi_k). \quad (41)$$

We will also write  $\Gamma_k \equiv \Gamma(x)/x \rightarrow E_k^{\text{FL}}$  at the FL quasiparticle location, where we expect for the Fermi liquid  $\Gamma_k \sim c_1(k - k_F)^2 + c_2 T^2$ , with suitable values as described more fully in Sec. IV F. At this value, we have the identity  $\Re e G(\xi_k, E_k^{\text{FL}}) = n^2/(4\Delta_0)$  as remarked above. As a consequence, in Fig. 1 the intersection of the line H1 and the vertical  $y$  axis also coincides with the value of  $\Re e G$  at the chemical potential. To elucidate the line shape of the ECFL, we start with the FL solution and perturb around it to find the corrected location of the peaks in the full spectral function.

$$\rho_G^{\text{Peak}}(\xi_k, x) = \frac{1}{\pi} \frac{Z_k^2 \Gamma_k}{Z_k^2 \Gamma_k^2 + (x - E_k^{\text{FL}})^2} \frac{n^2}{4\Delta_0} (\varepsilon_0 + \xi_k - x). \quad (42)$$

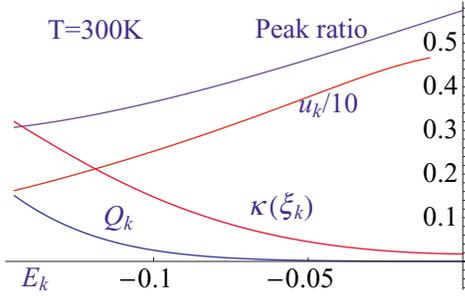


FIG. 2. (Color online) The density  $n = 0.85$ ,  $T = 300$  K,  $\Delta_0 = 0.0678$ , and parameters are from set I in Eq. (54). The various dimensionless variables shown against  $E_k^{\text{FL}}$  are the peak ratio from Eq. (50), the variable  $u_k$  (scaled by ten) from Eq. (44), the skew asymmetry factor  $\kappa(\xi_k)$  from Eq. (65), and the variable  $Q_k$  from Eq. (61).

Similarly, the real part is found:

$$\Re G^{\text{Peak}}(\xi_k, x) = \frac{n^2}{4\Delta_0} \left[ 1 + Z_k \frac{(x - E_k^{\text{FL}})(\varepsilon_0 + \xi_k - x)}{Z_k^2 \Gamma_k^2 + (x - E_k^{\text{FL}})^2} \right]. \quad (43)$$

We introduce the following convenient positive variable  $u_k$ :

$$\sinh u_k \equiv \frac{\varepsilon_0 + \xi_k - E_k^{\text{FL}}}{Z_k \Gamma_k}, \quad (44)$$

so that near the Fermi energy and at low  $T$  the small  $\Gamma_k \sim T^2$  drives it to a large and positive value, i.e.,  $\exp u_k \rightarrow 1/T^2$ . At higher binding energies,  $u_k$  decreases toward zero, as discussed below. We will also define a dimensionless variable  $Q_k$  below in Eq. (61) that depends on  $u_k$  only and determines the shape of the peak. To analyze the shape at a given  $\xi_k$ , we introduce a dimensionless energy variable  $\epsilon$  through the relation

$$x = E_k^{\text{FL}} + Z_k \Gamma_k \epsilon, \quad (45)$$

where we must require that  $|\epsilon| \sim 1$  for the expansion around the FL peak to be valid. The spectral function is expressible as

$$\rho_G^{\text{Peak}}(\xi_k, \epsilon) = \rho_G^*(k) \left[ \frac{\sinh(u_k) - \epsilon}{1 + \epsilon^2} \right] 2e^{-u_k}, \quad (46)$$

$$\Re G^{\text{Peak}}(\xi_k, x) = \frac{n^2}{4\Delta_0} \left\{ 1 + Z_k \epsilon \left[ \frac{\sinh(u_k) - \epsilon}{1 + \epsilon^2} \right] \right\}. \quad (47)$$

From Eq. (46) we see that at any  $k$ , the spectral function peaks at  $\epsilon^* \equiv -e^{-u_k}$  with the true quasiparticle peak  $E_k^*$  corrected from the Fermi-liquid value  $E_k^{\text{FL}} (= Z_k \xi_k)$  as  $E_k^* \equiv E_k^{\text{FL}} - e^{-u_k} Z_k \Gamma_k$  at  $\epsilon^* = -e^{-u_k}$ . Simplifying, we find the EDC energy dispersion or spectrum

$$E_k^* = \xi_k + \varepsilon_0 - \sqrt{[\varepsilon_0 + (1 - Z_k) \xi_k]^2 + Z_k^2 \Gamma_k^2}. \quad (48)$$

We provide examples of this dispersion later in Fig. 5.

For a given  $\xi_k$ , the magnitude of the spectral function at this peak is given by

$$\Re G^*(k) = \frac{n^2}{4\Delta_0} \left( 1 - \frac{1}{2} Z_k \right),$$

$$\rho_G^*(k) = \frac{n^2 Z_k}{8\pi \Delta_0} e^{u_k}. \quad (49)$$

The magnitude of  $\Re G^*(k)$  is a little smaller than the value  $n^2/(4\Delta_0)$  arising at the FL solution  $\epsilon = 0$ . In Fig. 1, this is reflected in the line H2 that lies a little below H1.<sup>14</sup>

The peak value  $\rho_G^*(k)$  falls off with  $\xi_k \ll 0$ , and is always smaller relative to the peak of the aux-FL peak value  $\rho_g^*(k)$ . The ratio of the two peak values is given by

$$\frac{\rho_G^*(k)}{\rho_g^*(k)} = \frac{n^2 Z_k \Gamma_k}{8\Delta_0} e^{u_k}. \quad (50)$$

We see below numerical examples of these functions. Figure 2 illustrates the peak ratio and other features for a typical set of parameters.

### E. MDC or constant-energy scans and energy dispersion.

It is also useful to study the momentum distribution curves obtained by fixing the energy  $x$  and scanning the energy  $\xi_k$ .<sup>15</sup> In the model of a  $k$ -independent self-energy, this is a particularly convenient strategy, and hence maximizing Eq. (37) at a fixed  $x$ , we find the MDC energy dispersion or spectrum:

$$\xi^*(x) = x - \varepsilon_0 + \sqrt{\Gamma^2(x) + [\varepsilon_0 - \Re \Phi(x)]^2}. \quad (51)$$

Thus  $\xi^*(x)$  is the peak position of  $\xi_k$  in constant-energy scans, whereas  $E_k^*$  in Eq. (48) represents peak position of energy at a fixed  $\xi_k$ . It is amusing to compare this with Eq. (48). Unlike Eq. (48), this formula is valid at all energies, not just near the chemical potential where the two agree closely. We will see below in Fig. 5 that this function is multivalued in a range of values of energy  $x$  leading to characteristic features of the spectrum.

### F. Numerical example of highdimensional ECFL model

In this section, we use a rectangular band with height  $1/(2W)$  and width  $2W$ , and take  $W = 0.86$  eV (i.e.,  $10^4$  K) as a typical value. In Ref. 3, a more realistic band structure is used as described in detail there. The model for the Fermi liquid introduced in Ref. 1 [see Eq. (24)] is given by the expression

$$\Gamma(x) = \pi \rho_\Phi(x) = \pi C_\Phi (x^2 + \tau^2) e^{-C_\Phi (x^2 + \tau^2)/\omega_c} + \eta, \quad (52)$$

with  $\tau = \pi k_B T$ . We have added a scattering width  $\eta$  as in Ref. 3, in order to account for scattering by off planar impurities. The real part of the self-energy is found from the Hilbert transform of  $\rho_\Phi(x)$ , and is given by

$$\Re \Phi(x) = C_\Phi \pi (x^2 + \tau^2) e^{-C_\Phi (x^2 + \tau^2)/\omega_c} \times \text{Erfi}(x/\sqrt{\omega_c}) - C_\Phi x \sqrt{\pi \omega_c} e^{-C_\Phi \tau^2/\omega_c}, \quad (53)$$

where  $\text{Erfi}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2} dt$  is the imaginary error function. A numerically small correction arising from  $\eta$  is dropped for brevity.

### G. Typical parameters

The same model is also used in the fit to experiments in Ref. 3 with a slight change of notation given by writing  $C_\Phi \rightarrow \frac{1}{\pi \Omega_0}$  and  $\omega_c \rightarrow \frac{\omega_0^2}{\pi \Omega_0}$ , in terms of the high- and low-frequency cutoff frequencies  $\omega_0$  and  $\Omega_0$ . We use two sets of standard

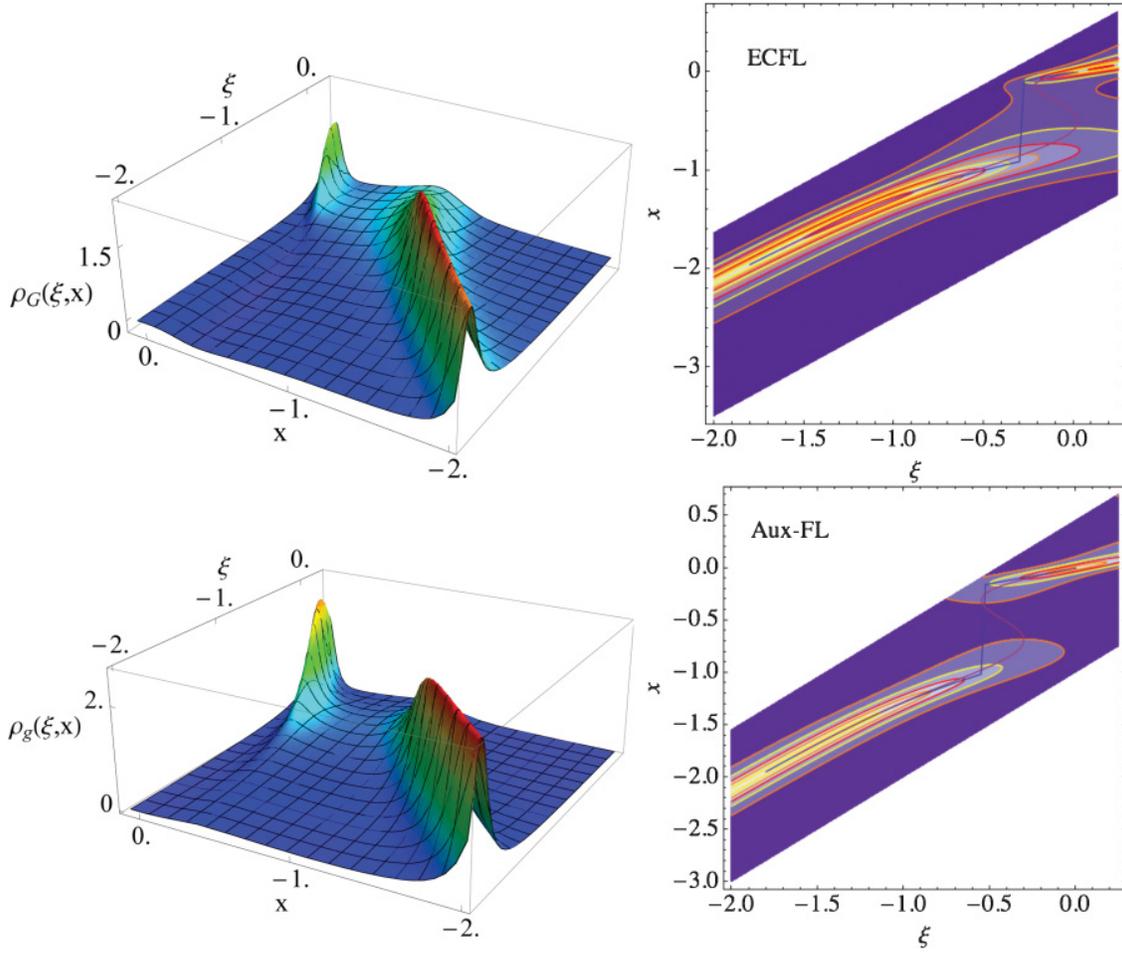


FIG. 3. (Color online) The ECFL (top left) and the auxiliary FL spectral functions (bottom left) at density  $n = 0.85$ ,  $T = 180$  K,  $\Delta_0 = 0.12$  eV,  $\eta = 0.12$  eV, and the other parameters are from Set II in Eq. (54). Here,  $\xi$  and  $x$  are in units of eV. In the ECFL curve on left, it is seen that the excitations near the Fermi energy become broad and dissolve into the continuum at an energy  $\sim -0.2$  eV, and reappear as sharp modes at a deeper binding energy. In the auxiliary FL, the excitations near the Fermi energy remain sharp and extend to lower energies than in the ECFL curves. The contour plots of the same functions in the right panel (top ECFL and bottom auxiliary FL) give a complementary perspective of the spectrum. The two superimposed solid lines at top right are from curves I and II of Fig. 5 and at bottom right curves III and IV of Fig. 5.

parameters;

$$\begin{aligned}
 \text{Set I: } & C_\Phi = 1 \text{ eV}^{-1}, \quad \omega_c = 0.25 \text{ eV} \quad \text{or} \\
 & \omega_0 = 0.5 \text{ eV}, \quad \Omega_0 = 0.318 \text{ eV} \\
 \text{Set II: } & C_\Phi = 2.274 \text{ eV}^{-1}, \quad \omega_c = 0.568 \text{ eV} \quad \text{or} \\
 & \omega_0 = 0.5 \text{ eV}, \quad \Omega_0 = 0.14 \text{ eV}. \quad (54)
 \end{aligned}$$

Set I was used in Ref. 1 for schematic plots employing a simple band density of states  $g_B(\epsilon) = \frac{1}{2W} \Theta(W^2 - \epsilon^2)$ . Set II was used in Ref. 3 employing a more elaborate dispersion described therein to successfully fit data on various high-temperature superconductors at optimal doping. The value of  $\eta$  is displayed in different plots. In Eq. (39), the spectral function  $\rho_g$  of the aux-FL is defined. The chemical potential is fixed by the number sum rule with  $\xi = \epsilon - \mu$

$$\frac{n}{2} = \int_{-\infty}^{\infty} dx f(x) \int d\epsilon g_B(\epsilon) \rho_g(\epsilon - \mu, x), \quad (55)$$

where  $f(x) = (1 + e^{\beta x})^{-1}$  is the Fermi function. We now write the contributions from extreme correlations that are described

in Ref. 1. The inelastic energy scale  $\Delta_0$  is found from the sum rule:

$$\Delta_0 = \int_{-\infty}^{\infty} dx f(x) \int d\epsilon g_B(\epsilon) \rho_g(\epsilon - \mu, x) (\epsilon - \mu - x). \quad (56)$$

Thus at a given density and temperature  $n, T$ , the model has only two parameters  $\omega_c$  and  $C_\Phi$  so that  $\Delta_0$  is fixed from Eq. (56). We study the details of the spectra next.

## V. THE SPECTRAL CHARACTERISTICS OF THE HIGH-DIMENSIONAL ECFL MODEL (37)

### A. Global view of the spectral function

We display in Fig. 3 the spectral function for the ECFL model Eq. (37) in three-dimensional (3D) plots and contour plots. Two distinct perspectives of the spectrum are found in the figure from the 3D and the contour plots. In both of these plots, we see that the excitations are sharply defined only for

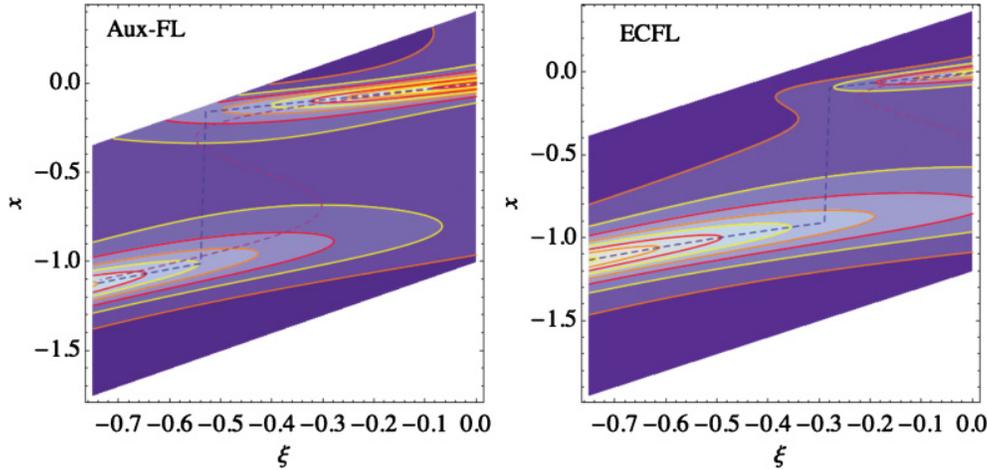


FIG. 4. (Color online) The contour plots of the aux-FL (left) and the ECFL model from Fig. 3 (right) with the same parameters as in Fig. 3 but over a smaller energy window. We superimpose the constant-wave-vector dispersion and MDC dispersion, with a value of  $\eta = 0.12$  common to the contour plots. The energy scale of the feature near the chemical potential is considerably reduced in the ECFL, and the “jump” in the EDC dispersion occurs at roughly half the corresponding energy in the aux-FL.

a certain range near the Fermi energy, and then merge into the continuum. At higher binding energies, the spectrum again looks quite sharp. For comparison, in Fig. 3, we also display the aux-FL spectral function. We note that the aux-FL spectra also become sharp at higher binding energies. This sharpening is modeled by the Gaussian in Eq. (39), its basic origin is the decrease in the weight of physical processes capable of quasiparticle damping as we move toward the band bottom. In order to look more closely at the low-energy part of the spectrum of the aux-FL and the ECFL, we show in Fig. 4 the contour plots of both over a smaller energy range.

We see that viewed in this rather broad sense, dispersions of the aux-FL and the model ECFL spectra share many characteristics, with somewhat different energy scales. However, there are crucial differences that emerge when we look at the distribution of spectral weight that arises in the ECFL, where the caparison factor in Eq. (37) pushes weight to higher binding energies. This is reflected most significantly in the line shapes that we study below. Since we use the momentum-independent self-energy for the aux-FL in this model calculation, we obtain very detailed EDC and MDC plots below. However, it must be borne in mind that refined calculations within the ECFL framework must necessarily introduce some momentum dependence, and hence several details are likely to change, in particular the structure far from the chemical potential would change somewhat more. Our view is that, this caveat apart, it is very useful to take the Eq. (37) seriously since it gives a simple framework to correlate different data.

### B. Dispersion relations in EDC and MDC

In Fig. 5, the EDC dispersion relation (i.e., locus of peaks of the spectral function at fixed  $\xi$ , found by numerical maximization), is plotted versus  $\xi$  along with the MDC spectrum Eq. (51). We recall that the latter expression is exact at all  $\xi$  and  $x$ , whereas Eq. (48) is not quite exact for the EDC dispersion. For comparison, we also show the corresponding figures for the aux-FL in Eq. (39), with the

same parameters. The dispersion relations Eq. (48) is displayed in the inset of Fig. 5, where it is compared with the result of numerically maximizing the spectral function at a fixed  $\xi$ . We see that Eq. (48) is only good for a range of energies near the Fermi energy.

We see that both sets of spectra for the aux-FL as well as the ECFL model exhibit similar global features, but with different scales of energy. In both cases, the constant energy scans show a jump discontinuity, whereas the MDC spectra show an “S” type or re-entrant type behavior. The origin of the latter is easy to see in the aux-FL, here a peak in  $-\Re\Phi(x)$  occurs at an energy approximately  $2\omega_0$ , so that as  $x$  decreases from zero,  $\xi_{\text{aux-FL}}^* = x - \Re\Phi(x)$  goes back up for a certain range. In the case of the ECFL, Eq. (51) shows that the energy scale  $\varepsilon_0$  enters the expression when  $\Gamma(x)$  becomes comparable to  $\varepsilon_0 - \Re\Phi(x)$ , and the net result is that the re-entrant behavior is pushed to lower binding energies.

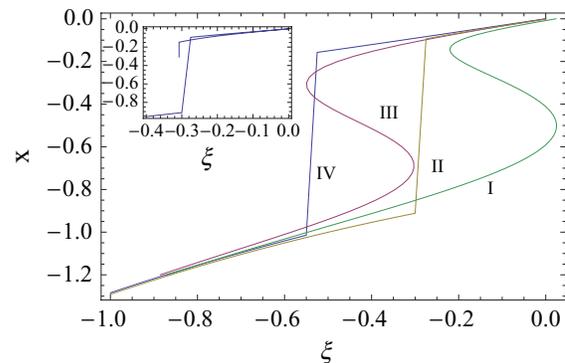


FIG. 5. (Color online) Energy dispersion curves in the ECFL and the aux-FL models. Here, the parameters are from set II in Eq. (54), with  $n = 0.85$  and  $T = 180$  K. With  $\eta = 0.12$ , curves I and II have the peaks in constant-wave-vector and constant-energy scans of the spectral function (37), and curves III and IV are corresponding figures for the aux-FL in Eq. (39). The inset compares Eq. (48) (the truncated curve) with the exact locus found by numerical maximization.

### C. The energy shift

The dispersion (48) corresponds to the ridge near the Fermi energy in Fig. 3. At low temperature, since  $\Gamma_{k_F} \sim O(T^2)$ , the corrected quasiparticle energy is always less than  $E_k^{\text{FL}}$ , so that there is always a leftward (i.e., red) shift of the dispersion, or from the hole (binding) energy point of view, we may say there is a blue shift. The peak shift is given by

$$\begin{aligned} \Delta E_k &= E_k^* - E_k^{\text{FL}} \\ &= (1 - Z_k)\xi_k + \varepsilon_0 - \sqrt{[\varepsilon_0 + (1 - Z_k)\xi_k]^2 + Z_k^2\Gamma_k^2}, \end{aligned} \quad (57)$$

which is a function of both  $T$  and  $k$ . Close to the Fermi energy, this can be written as

$$\Delta E_k = -\frac{Z_{k_F}^2 \Gamma_k^2}{2\varepsilon_0}. \quad (58)$$

At the Fermi momentum, this small shift is seen in Fig. 1 where the vertical line through  $x^*$  is displaced to the left from the  $y$  axis. As long as  $k \sim k_F$  this shift is very small  $\Delta E_k \sim O(T^4)$ , but as  $k$  moves away from  $k_F$  the shift (57) grows with  $\xi_k$ . This departure makes the dispersion in Eq. (48) depart significantly from the bare dispersion  $\xi_k$  as we move away from  $k_F$ . We see from Fig. 5 that the departure of the EDC peaks from the Fermi liquid is somewhat less pronounced than those of the MDC's, the latter is operationally called the low-energy kink. Our calculations therefore predict the magnitude of the shift (57) in terms of the energy scale  $\varepsilon_0$  and the Fermi-liquid parameter  $\Gamma_k$ . This energy shift is therefore also a useful method for extracting the fundamental parameter  $\Delta_0$  on using Eq. (36).

### D. Constant-energy cuts or MDC line shapes

We display the MDC line shapes in Fig. 6. Panel (a) shows the effect of the caparison factor  $[1 - \frac{\eta}{2} + \frac{\eta^2}{4\Delta_0}(\xi - x)]$ , whereby the curves are skewed to the right, in contrast to the EDC curves that are skewed to the left. The latter important feature is also seen below in Fig. 7(a) and noted in Ref. 1. Panel (b) shows the shallow peaks in the ‘‘S-like’’ region of the energy dispersion seen in Fig. 5, and panel (c) shows the deep interior region where the peaks are more symmetric.

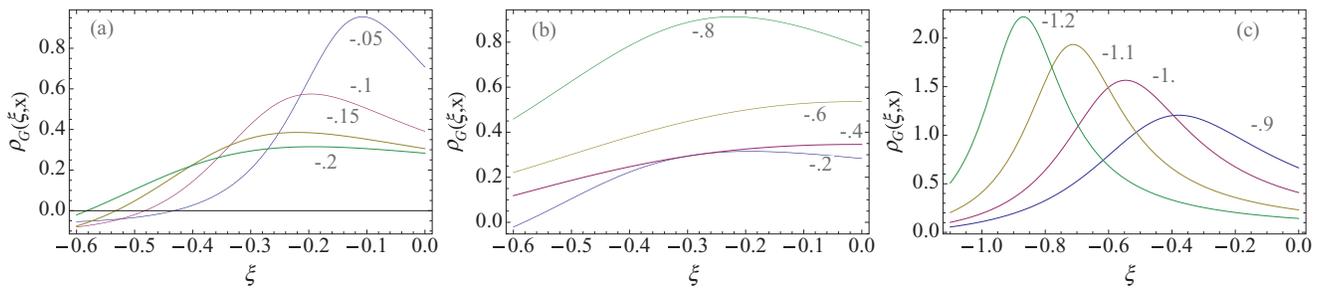


FIG. 6. (Color online) MDC line shapes at different values of energy  $x$  displayed in each curve. Here, the parameters are from set II in Eq. (54), with  $n = 0.85$ ,  $T = 180$  K, and  $\eta = 0.12$ . Panel (a) corresponds to  $x$  close to the chemical potential. It is interesting to note that the curves are skewed to the right, thus mirror imaging the leftward skew seen in the constant- $\xi$  (EDC) scans below Fig. 7(a), in a comparable range of energies and wave vectors. Panel (b) corresponds to the midenergy range, within the reentrant range of  $x$  from Figs. 5 or 3, with the counterintuitive movement of the shallow peak to the right with increasing  $x$ . Panel (c) corresponds to the second set of maxima in Fig. 3 far from the chemical potential, where the curves are quite symmetric.

### E. Constant wave-vector cuts or EDC line shapes

The spectral function and the real part of the Greens function are calculated from Eqs. (38) and (37). We display the EDC line shapes in Fig. 7. Panel (a) gives an overview of the spectral shapes for wave vectors near the Fermi surface, displaying a left skewed peak that falls rapidly in intensity as it broadens. This behavior is of great interest since it captures the experimental features in high- $T_c$  systems, as elaborated in Ref. 3. Panel (b) shows the spectra at higher binding energies, where a feature at lower energies begins to disperse significantly with  $\xi$ . It is evident that these two sets of dispersing features correspond to the two branches that are seen in the 3D plots and contour plots of Fig. 3. The inset in Fig. 7(b) shows the behavior of the aux-FL, where the two features are again seen but with different rates of intensity change.

We now turn to the task of understanding the reconstructed Dyson self-energy that leads to the above electron spectral functions. In Fig. 8, we show the spectral function  $\rho_G$  at various values of the energy  $\xi$  at  $T = 300$  K. The Fermi-liquid spectra at the same values of parameters are also shown for comparison. The Dyson self-energy  $\rho_\Sigma(\xi, x)$  necessary to produce these spectral functions is found using Eq. (23) and is displayed in Fig. 8 at two temperatures. The object  $\rho_\Sigma(\xi, x)$  has a distinctive minimum for each  $\xi$  that shifts to the left along with the energy  $\xi$ , which tracks the peaks in the physical spectral function  $\rho_G(\xi, x)$  from Eq. (23). It also shows the asymmetry between energies above and below the chemical potential that we noted at  $\xi = 0$  in Fig. 9. At the Fermi energy,  $\rho_\Sigma(\xi, x)$  is displayed in Fig. 9 over a large scale.

### F. The reduced line-shape function

An interesting aspect of the ECFL model Greens function (37) is the change in shape of the peaks as we leave the Fermi surface, so that the quasiparticles become hard to define at some point. This change in shape can be formulated neatly in terms of a single dimensionless parameter  $Q_k$  that we now define and explore. We examine Eq. (46) around its peak by writing

$$\varepsilon = \varepsilon^* + \cosh(u_k)\bar{\varepsilon}, \quad (59)$$

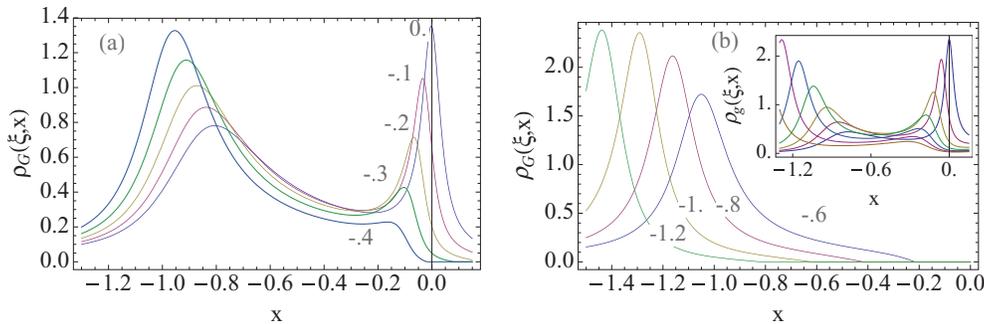


FIG. 7. (Color online) EDC line shapes at different values of energy  $\xi$  displayed in each curve. Here, the parameters are from set II in Eq. (54), with  $n = 0.85$ ,  $T = 180$  K, and  $\eta = 0.12$ . Panel (a) corresponds to  $\xi$  close to the chemical potential. Note that the curves are skewed to the left, i.e., a mirror image of the rightward skew seen in the constant- $x$  MDC scans above Fig. 6, in a comparable range of energies and wave vectors. Panel (b) corresponds to the higher energy range, and we see that only one broad maximum is found at a given  $\xi$ . The inset in (b) shows the aux-FL constant- $\xi$  scans for the same range; here, each  $\xi$  results in a pair of maxima, originating from the functional form of the self-energy in Eq. (37).

so that  $\rho_G^{\text{Peak}}(\xi_k, \bar{\epsilon}) = \rho_G^*(k)\gamma(Q_k, \bar{\epsilon})$ , with a characteristic line-shape function  $\gamma$  given by

$$\gamma(Q_k, \bar{\epsilon}) = \left[ \frac{Q_k(1 - \bar{\epsilon})}{Q_k(1 - \bar{\epsilon}) + \bar{\epsilon}^2} \right], \quad (60)$$

with

$$Q_k = 2 \frac{e^{-u_k}}{\cosh(u_k)}. \quad (61)$$

The parameter  $Q_k$  goes to zero near the Fermi surface at low  $T$  since  $u_k \rightarrow \infty$ , but at higher binding energies increases:  $Q_k \rightarrow 2$ .

As we get deeper into the occupied states  $\xi \ll 0$ , we find a remarkable change in shape of the spectral functions. This is illustrated in Fig. 10 where we plot  $\gamma$  of Eq. (60) after normalizing to unit area. In order to have a well-defined quasiparticle-type peak in  $\rho_G$  for  $\epsilon \sim O(1)$ ,  $Q_k$  must be small enough. This translates to the requirement of  $\xi_k$  being close to the Fermi surface. By setting  $Q \sim 1$  as the condition for losing a peak in the spectrum, we obtain the condition

$$\epsilon_0 + (1 - Z_k)\xi_k = \frac{1}{\sqrt{3}}Z_k\Gamma_k, \quad (62)$$

beyond which it is meaningless to talk of quasiparticles. This gives  $\epsilon_0$  as a rough characteristic scale for the disappearance of the quasiparticle peaks.

Figure 11 illustrates the change in shape somewhat differently by normalizing all curves to unity at the peak as in Eq. (60). The peak at  $\bar{\epsilon} = 0$  is sharp and quite symmetric for  $Q \ll 1$ , and becomes broader and more left skewed as  $Q_k$  increases toward its maximum value of two. Attaining the maximum value is possible, in principle, requirement being  $u_k = 0$  or from Eq. (44):

$$\Delta_0 = \frac{n^2}{4 - 2n} \eta e \Phi(E_k^{\text{FL}}). \quad (63)$$

Unless  $\Delta_0$  is very small, this condition is hard to satisfy. If this possibility is achieved, then there are several interesting consequences. Firstly, we note that from Eq. (49), the magnitude of the spectral function at  $u_k = 0$  becomes insensitive to disorder and temperature, etc. Its magnitude,  $n^2 Z_k / (8\pi \Delta_0)$ , should be

useful for finding  $\Delta_0$ . If this is approximately satisfied, then the peak structure loses meaning and the spectrum is essentially flat. Taking  $\bar{\epsilon} = -1$ , the fall off from the peak value of unity is 80%, and the spectrum becomes essentially featureless.

### G. Skewness parameter of the spectrum

We now estimate the skewness of the spectrum. The function (60) drops to half its peak value at two values of  $\bar{\epsilon}_{L,R}$  to the left and right of the peak given by

$$\bar{\epsilon}_{L,R} = -\frac{1}{2}Q_k \mp \sqrt{Q_k + \frac{1}{4}Q_k^2}. \quad (64)$$

For small  $Q \ll 1$ , the (dimensionless) width of the peak is small, it increases with  $Q$  as discussed further below. We define a dimensionless skew parameter in terms of the energies  $\epsilon^*, \epsilon_{R,L}$  [rather than the  $\bar{\epsilon}$ 's that are related via Eq. (59)]:

$$\begin{aligned} \kappa(\xi_k) &= \frac{2\epsilon^* - \epsilon_R - \epsilon_L}{\epsilon^* - \epsilon_L} \\ &= \tanh(u_k) - 1 + \sqrt{[2 - \tanh(u_k)]^2 - 1}, \end{aligned} \quad (65)$$

with the property that near the Fermi level when  $u_k \rightarrow +\infty$  the variable  $\kappa \rightarrow 0$ , and we get a symmetric curve about the maximum. On the other hand, for deeper occupation,  $u_k$  decreases in magnitude toward zero, driving  $\kappa \rightarrow 0.732$  and gives a curve that is increasingly biased to the left. The asymmetry  $\kappa$  grows as  $O(T^2)$  at low temperatures, and it is rather large at room temperature. As a rough estimate, the quasiparticle peak is lost when  $Q_k \sim 0.5$  where  $u_k \sim 0.98$  and  $\kappa(\xi_k) \sim 0.5$ . This loss of quasiparticle peak structure, skew factor, and its experimental signature is studied in greater detail in Ref. 3. See Fig. 2 for typical plot of skew parameter  $\kappa$  and  $Q$  versus the hole binding energy  $E_k$ .

## VI. SELF-ENERGY OF THE SELF-ENERGY AND A MORI-TYPE PROCEDURE.

Since the construction given above generates  $\Sigma(z)$  from  $G(z)$  given only the representation (1), we can as well repeat the trick. Since  $\Sigma(z)$  satisfies Eq. (18) with a density  $\rho_\Sigma(x)$  that is assumed known, and is analytic in the complex  $z$  plane

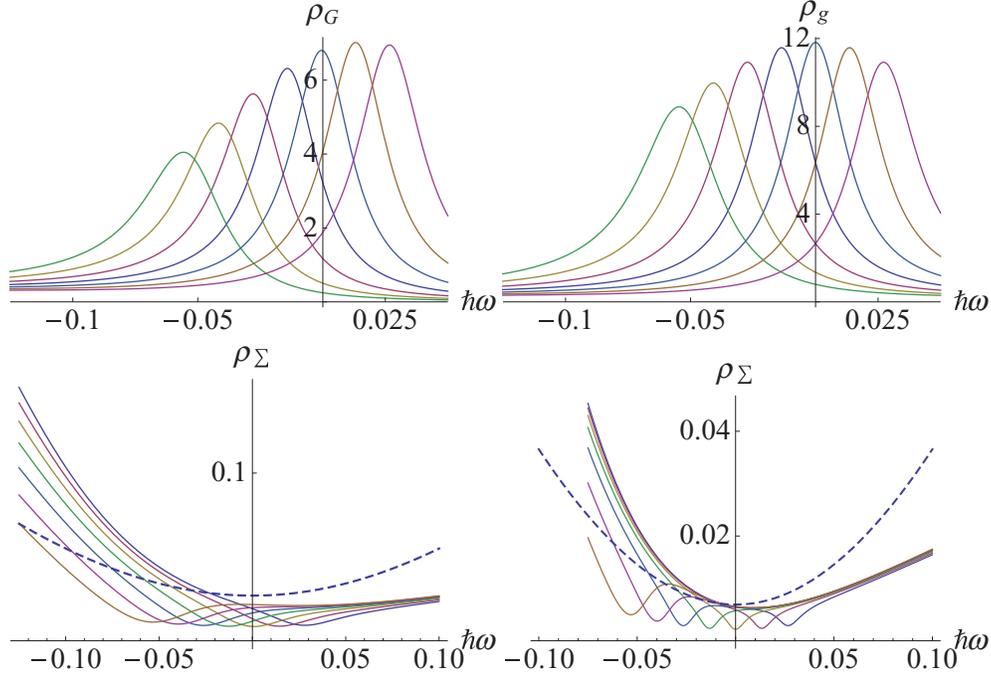


FIG. 8. (Color online) Top left panel: density  $n = 0.85$ , temperature  $T = 300$  K,  $\Delta_0 = 0.0678$ , and other parameters are from set I in Eq. (54). From left to right  $\rho_G(\xi, x)$  for energies in units of eV:  $\xi = -0.1, -0.075, -0.05, -0.025, 0, 0.025, 0.05$ . Top right panel: spectral function  $\rho_g(\xi, x)$  from Eq. (39) corresponding to the same  $\xi$  as in the left panel. The difference in the line shapes becomes clear when we examine the Dyson self-energy that produces these curves. Bottom left panel: the panel shows the spectral function at  $T = 300$  K for the inferred Dyson self-energy  $\rho_\Sigma(x)$  from Eqs. (23), (38), and (37) for the same energies. The dashed line is the input Fermi-liquid spectral function  $\rho_\Phi(\omega)$  at the same temperature. Bottom right panel: temperature  $T = 150$  K,  $\Delta_0 = 0.0642$ , and the identical data as in the bottom left panel.

with a  $1/z$  fall off, it satisfies the necessary conditions for a further decomposition. Consulting Eq. (23), we write down by inspection:

$$\begin{aligned}\Sigma(z) &= \frac{a_\Sigma}{z - \bar{E}_\Sigma - \Sigma^{(1)}(z)}, \\ \Sigma^{(1)}(z) &= \int dx \frac{\rho_{\Sigma^{(1)}}(x)}{z - x}.\end{aligned}\quad (66)$$

The constants  $a_\Sigma = \int dx \rho_\Sigma(x)$  and  $\bar{E}_\Sigma = \int dx x \rho_\Sigma(x) / a_\Sigma$  are known through  $\rho_\Sigma(x)$ . They may in practice be conve-

niently determined in terms of the moments of the Greens function<sup>7-9</sup> in applications. The spectral function is given by

$$\rho_{\Sigma^{(1)}}(x) = \frac{a_\Sigma \rho_\Sigma(x)}{[\pi^2 \rho_\Sigma(x)]^2 + [\text{Im} \Sigma(x)]^2}.\quad (67)$$

Comparing this representation with Eq. (21), we note the formal similarity between  $\rho_G(x)$  and  $\rho_{\Sigma^{(1)}}(x)$ . Thus for a Fermi liquid with momentum-independent self-energy, its

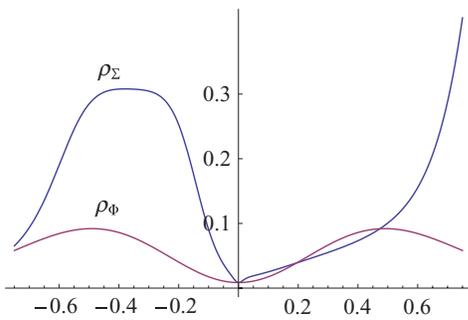


FIG. 9. (Color online) With  $n = 0.85$  and  $T = 300$  K, and the other parameters from set I in Eq. (54). The spectral function for the inferred Dyson self-energy  $\rho_\Sigma(\xi = 0, x)$  using Eq. (23) and the the input Fermi-liquid spectral function  $\rho_\Phi(x)$  over a larger energy range. Note the distinctive asymmetry in shape of  $\rho_\Sigma$  below and above the Fermi energy.

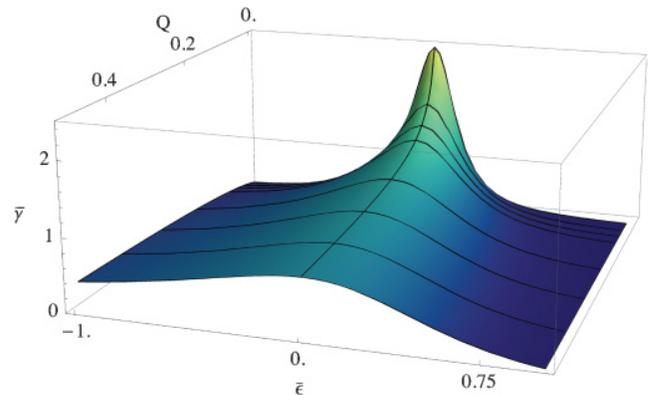


FIG. 10. (Color online) The spectral shapes possible are seen by plotting the shape function at different values of the parameter  $Q$ . In this curve,  $\bar{\gamma}$  is the  $\gamma(Q, \bar{\epsilon})$  of Eq. (60) normalized to unit area in the natural interval  $[-1, 1]$  for the variable  $\bar{\epsilon}$ .

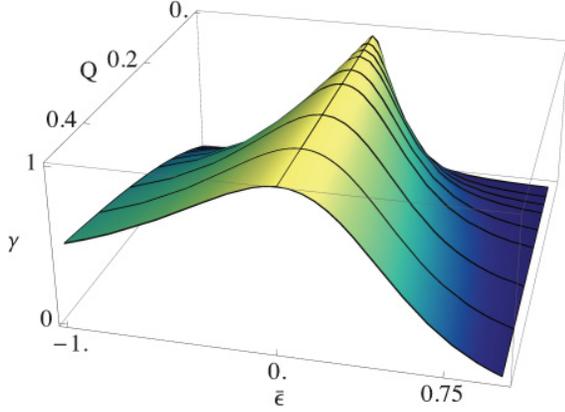


FIG. 11. (Color online) The same shapes as in Fig. 10 but now  $\gamma$  is normalized to unity at the peak as in Eq. (60). The sharp peaks for small  $Q \leq 0.25$  flatten out as  $Q$  increases with a left skew asymmetry that is characteristic of this functional dependence.

next self-energy resembles closely the spectral function  $\rho_G$ , especially at the Fermi energy.

We follow up briefly on the above amusing observation, and obtain a hierarchy of self-energies starting from an initial self-energy given by the spectral representation, Eqs. (18) and (66). This process parallels the continued fraction representation of analytic functions and seems intimately related to the formalism developed by Mori.<sup>7,8</sup> The latter is expressed in the language of projection operators for Liouville operators that is less straightforward than our simple treatment.

In order to conform to the notation popular in the Mori formalism, we will express the variables in Laplace representation rather than the one used above with complex frequencies. Let us consider the thermal and temporal correlation function for two operators  $A$  and  $B$  in Schrödinger time and its Laplace transform:

$$C_{AB}(t) = \int_0^\beta d\tau \langle A(t - i\tau)B(0) \rangle,$$

$$\hat{C}_{AB}(s) = \int_0^\infty e^{-st} C_{AB}(t) dt. \quad (68)$$

In the standard case, we find  $A = B^\dagger$ , where the product is real and also positive.<sup>16</sup> We see that the Laplace-transform function satisfies an integral representation:

$$\hat{C}_{AB}(s) = \int_{-\infty}^\infty dv \frac{\rho_{AB}(v)}{s - iv}, \quad \text{with a real density given by}$$

$$\rho_{AB}(v) = \sum_{nm} \frac{P_m - P_n}{\varepsilon_n - \varepsilon_m} \langle n|A|m \rangle \langle m|B|n \rangle \delta(\varepsilon_n - \varepsilon_m - v). \quad (69)$$

This object is closely connected with the correlation functions used in Eqs. (1) and (2) by using the fluctuation-dissipation theorem. Following Mori, we write down a relaxation function with the normalization property  $Z_0(0) = 1$ , and its Laplace transform

$$Z_0(t) \equiv \frac{C_{AB}(t)}{C_{AB}(0)}, \quad \hat{Z}_0(s) = \int_{-\infty}^\infty dv \frac{\rho_0(v)}{s - iv}. \quad (70)$$

Here, the real density  $\rho_0(v) = \frac{1}{C(0)}\rho_{AB}(v)$ , satisfies the normalization condition  $\int_{-\infty}^\infty dv \rho_0(v) = 1$ . Using the identity

$$\frac{1}{0^+ + i(u - v)} = \pi \delta(u - v) - i\mathcal{P} \frac{1}{u - v},$$

with  $\mathcal{P}$  denoting the principle value, an inverse relation expressing  $\rho_0(v) = \frac{1}{\pi} \Im e \hat{Z}_0(0^+ + iv)$  follows. In order to find a Dyson-type representation for  $\rho_0$ , following Eqs. (8) and (9), we take the “equation of motion” by multiplying Eq. (70) by  $s$  and write

$$s\hat{Z}_0(s) = 1 + i \int_{-\infty}^\infty dv \frac{v \rho_0(v)}{s - iv} \equiv Y_0(s) \hat{Z}_0(s). \quad (71)$$

The Dyson form of self-energy now emerges and we obtain

$$\hat{Z}_0(s) = \frac{1}{s - iY_0(s)}, \quad \text{with } Y_0(s) = \frac{\int_{-\infty}^\infty dv \frac{v \rho_0(v)}{s - iv}}{\int_{-\infty}^\infty dv \frac{\rho_0(v)}{s - iv}}. \quad (72)$$

As  $s \rightarrow \infty$ , the function  $Y_0(s)$  tends to  $\omega_1$ , with a real frequency  $\omega_1$  given by

$$\omega_1 = \int_{-\infty}^\infty dv v \rho_0(v). \quad (73)$$

Hence the function  $Y_0(s) - \omega_1$  falls off as  $1/s$  as  $s \rightarrow \infty$ . It is analytic everywhere except on the imaginary  $s$  axis. It therefore has a representation

$$Y_0(s) - \omega_1 = i \alpha_1 \int_{-\infty}^\infty dv \frac{\rho_1(v)}{s - iv}, \quad (74)$$

with a real density  $\alpha_1 \rho_1(v) = \frac{1}{\pi} \Im m [Y_0(0^+ + iv)]$ . With this, we may write

$$Z_0(s) = \frac{1}{s - i\omega_1 + \alpha_1 \int dv \frac{\rho_1(v)}{s - iv}}. \quad (75)$$

The real number  $\alpha_1$  is found using the convention that  $\rho_1(v)$  is normalized to unity. We may express  $\rho_1$  solely in terms of the lower density  $\rho_0(v)$  by using Eq. (70) as

$$\alpha_1 \rho_1(u) = \frac{\rho_0(u)}{\pi^2 \rho_0^2(u) + [\mathcal{H}[\rho_0](u)]^2}. \quad (76)$$

We determine  $\alpha_1$  from Eq. (76) by integrating over  $v$  and using the unit normalization of  $\rho_1(u)$ . It is evident from Eq. (76) that for the physically important case of a real and positive initial density  $\rho_0(v)$ , the derived density  $\rho_1(v)$  is also real positive.

This scheme is clearly generalizable to higher orders, and we simply iterate the above process. The answers may be written down by inspection as follows:

$$\hat{Z}_j(s) = \int_{-\infty}^\infty dv \frac{\rho_j(v)}{s - iv}, \quad \text{with normalization:}$$

$$\int_{-\infty}^\infty dv \rho_j(v) = 1. \quad (77)$$

These satisfy the recursion relation

$$\hat{Z}_j(s) = \frac{1}{s - i\omega_{j+1} + \alpha_{j+1} \hat{Z}_{j+1}(s)}, \quad (78)$$

where

$$\omega_{j+1} = \int_{-\infty}^\infty dv v \rho_j(v), \quad (79)$$

and  $\alpha_{j+1}$  as well as  $\rho_{j+1}(v)$  are defined through

$$\alpha_{j+1} \rho_{j+1}(u) = \frac{\rho_j(u)}{\pi^2 \rho_j^2(u) + [\mathcal{H}[\rho_j](u)]^2}. \quad (80)$$

Note that the numbers  $\alpha_j$  as well as  $\omega_j$  are real, and for all  $j$ , the densities  $\rho_j(v)$  are positive provided the the initial density  $\rho_0(v)$  is positive. This situation arises when the initial operators  $B = A^\dagger$ , as mentioned above.

It is clear that Eq. (76) is the precise analog of the relation (23) for the Greens function. The hierarchy of equations consisting of Eqs. (45)–(49) constitutes an iteration scheme that starts with the  $j = 0$  correlation function in Eq. (70). This is a forward hierarchy in the sense that successive densities at level  $j + 1$  are expressed explicitly in terms of the earlier ones at level  $j$ . In the reverse direction, it is rather simpler since level  $j$  is explicitly given in terms of level  $j + 1$  by Eq. (78). The use of this set of equations requires some *a priori* knowledge of the behavior of higher order self-energies to deduce the lower ones. Standard approximations<sup>7</sup> consist of either truncation of the series or making a physical assumption such as a Gaussian behavior at some level and then working out the lower level objects. Our object in presenting the above procedure is merely to point out that this iterative scheme is in

essence a rather simple application of the self-energy concept described above, with the repeated use of Eq. (23).

## VII. SUMMARY AND CONCLUSIONS

A new form of the electronic Greens function, departing widely from the Dyson form arises in the extreme correlation theory of the  $t$ - $J$  model. Motivated by its considerable success in explaining ARPES data of optimally doped cuprate superconductors,<sup>3</sup> we have presented in this paper results on the detailed structure of this Greens function and its spectral function. An illustrative example is provided, complete with numerical results, so that the novel line shape and its dependence on parameters is revealed. We have also presented a set of explicit results on the Mori form of the self-energy that holds promise in several contexts.

## ACKNOWLEDGMENTS

This work was supported by DOE under Grant No. FG02-06ER46319. I thank G-H. Gweon for helpful comments and valuable discussions. I thank A. Dhar and P. Wolfle for useful comments regarding the Mori formalism.

- 
- <sup>1</sup>B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).  
<sup>2</sup>B. S. Shastry, *Phys. Rev. B* **81**, 045121 (2010).  
<sup>3</sup>G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).  
<sup>4</sup>D. E. Logan, M. P. Eastwood and M. A. Tusch, *J. Phys. Condens. Matter* **10**, 2673 (1998).  
<sup>5</sup>Z. Wang, Y. Bang, and G. Kotliar, *Phys. Rev. Lett.* **67**, 2733 (1991).  
<sup>6</sup>J. M. Luttinger and J. C. Ward, *Phys. Rev.* **118**, 1417 (1960); J. M. Luttinger, *ibid.* **119**, 1153 (1960); **121**, 942 (1961).  
<sup>7</sup>H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965); **34**, 399 (1965).  
<sup>8</sup>M. Dupuis, *Prog. Theor. Phys.* **37**, 502 (1967).  
<sup>9</sup>J. J. Deisz, D. W. Hess, and J. W. Serene, *Phys. Rev.* **55**, 2089 (1997).  
<sup>10</sup>A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963). Our  $\rho_G(\xi_k, x)$  corresponds to the combination of spectral functions  $A(\vec{k}, x)$  and  $B(\vec{k}, x)$  used here.  
<sup>11</sup>The positive part constraint on the right of the following equation can be often omitted, we found that it is violated very slightly  $\sim 3\%$  in many cases. We omit it for simplicity in the following.  
<sup>12</sup>Shifting  $\hat{E}$  by a constant is also possible but the optimal choice is the one made here.  
<sup>13</sup>The notation is simplified from that in Ref. 1 by calling the extremely correlated Greens function as  $G$  rather than  $\mathcal{G}$  and the overbar in the self-energy  $\bar{\Phi}$  is omitted.

- <sup>14</sup>At  $T = 0$ , this finite value persists and  $\Im \text{leg}^{\text{Peak}}[\xi_k, x^*(\xi_k)]$  does not diverge, so that one might be concerned that the Luttinger-Ward volume theorem is being disobeyed. In comparison, note that the standard FL Greens function behaves in a slightly different way, at any finite  $T$  and  $\xi_k$ , at the energy  $x = E_k^{\text{FL}}$  we find both a peak in the spectral function  $\rho_g(\xi_k, E_k^{\text{FL}})$  and a zero of the  $\Re \text{leg}(\xi_k, E_k^{\text{FL}})$ , whereas at  $T = 0$ , we find a delta peak in the spectral function  $\rho_g(\xi_k, E_k^{\text{FL}})$  and a pole of the  $\Re \text{leg}(\xi_k, E_k^{\text{FL}})$ . However we see that the FL divergence of the real part of  $G$  of a typical Fermi liquid does occur, but displaced by a very small energy scale  $O(T^4)$ . The peak positions are displayed in Fig. 1 at a high enough temperature so that the features are distinguishable.  
<sup>15</sup>The process of scanning  $\xi_k$  used here differs slightly from the true MDC's, where one scans the wave vector  $\vec{k}$  rather than the energy  $\xi_k$ , but is more convenient here.  
<sup>16</sup>The spectral function  $\rho_{AA^\dagger}(x)$  defined below in Eq. (69) is positive as well. However, that condition can be relaxed and we can do with less, provided that the spectral function in Eq. (69) is real (rather than positive). For this to happen, we may allow for  $A \neq B^\dagger$ , but in this case, assume that both the matrix elements  $\langle nAm \rangle$  and  $\langle mBn \rangle$  are real numbers or imaginary numbers so that the product is real. These conditions correspond to both operators  $A$  and  $B$  being Hermitian (or anti Hermitian), and the absence of magnetic fields so that the wave functions may be chosen to be real. Thus we will assume the reality of the product of the matrix elements.

**Erratum: Anatomy of the self-energy [Phys. Rev. B 84, 165112 (2011)]**

B. Sriram Shastry

(Received 2 August 2012; published 13 August 2012)

DOI: [10.1103/PhysRevB.86.079911](https://doi.org/10.1103/PhysRevB.86.079911)

PACS number(s): 79.60.-i, 99.10.Cd

The spectral function  $\rho_{\Sigma}(\omega)$  of the Dyson self-energy, inferred from Eq. (23), is incorrectly shown in the lower two graphs of Fig. 8 and in Fig. 9, due to an erroneous code. The corrected figures are provided; the captions and text remain unchanged.

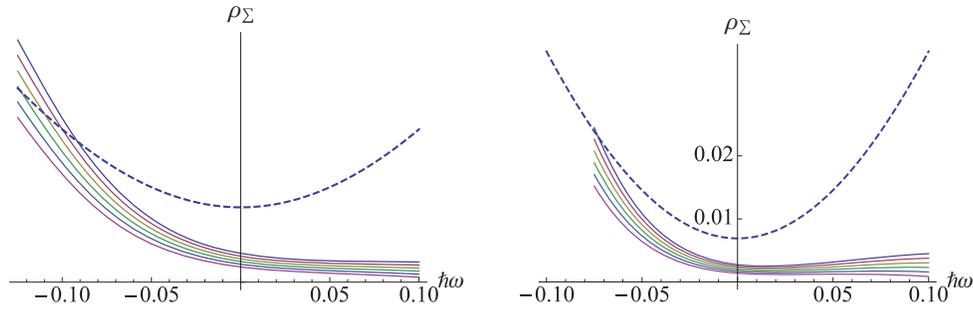


FIG. 1. (Color online) Replacement of lower panel in Fig. 8.

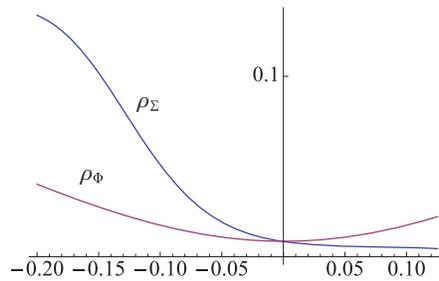


FIG. 2. (Color online) Replacement of Fig. 9.

## Dynamical Particle-Hole Asymmetry in High-Temperature Cuprate Superconductors

B. Sriram Shastry

*Department of Physics, University of California, Santa Cruz, California 95064, USA*

(Received 5 October 2011; published 10 August 2012)

Motivated by the form of recent theoretical results, a quantitative test for an important dynamical particle-hole asymmetry of the electron spectral function at low energies and long wavelengths is proposed. The test requires the decomposition of the angle resolved photo emission intensity, after a specific Fermi symmetrization, into odd and even parts to obtain its ratio  $\mathcal{R}$ . A large magnitude  $\mathcal{R}$  is implied in recent theoretical fits at optimal doping around the chemical potential, and I propose that this large asymmetry needs to be checked more directly and thoroughly. This processing requires a slightly higher precision determination of the Fermi momentum relative to current availability.

DOI: [10.1103/PhysRevLett.109.067004](https://doi.org/10.1103/PhysRevLett.109.067004)

PACS numbers: 74.72.-h, 74.20.Mn, 74.25.Jb

*Introduction.*—The search for a microscopic theory of the normal state of the cuprates is one of the main themes in condensed matter physics for the last two decades. The recent suggestions of describing the normal state in terms of theories with a quantum critical point [1] have also created wide interest in other branches of physics such as string theory and quantum gravity [2]. An initial theoretical objective is the derivation of the normal state low energy long wavelength single electron spectral function  $\rho_{\mathcal{G}}(\vec{k}, \omega)$  [or equivalently  $A(\vec{k}, \omega)$ ], encoding the complete set of symmetries.

In this Letter, I discuss the behavior of  $\rho_{\mathcal{G}}(\vec{k}, \omega)$  under a dynamical particle-hole transformation simultaneously inverting the wave vector and energy relative to the chemical potential  $\mu$  as

$$(\vec{k}, \omega) \rightarrow -(\vec{k}, \omega), \quad \text{with} \quad \vec{k} = \vec{k} - \vec{k}_{\text{F}}. \quad (1)$$

Invariance under this transformation has often been invoked in analyzing angle resolved photoemission (ARPES) data [3]. It is an emergent symmetry of the Fermi-liquid in the sense of Ref. [4], arising when correction terms of  $O(\omega/\varepsilon_{\text{F}})^3$  are neglected [5]. Fermi-liquids without disorder at intermediate coupling are invariant [6] under Eq. (1), as are most other contemporary theories of cuprates that I am aware of.

On the other hand two recent theories, the extremely correlated Fermi-liquid theory (ECFL) proposed by the author in Ref. [7], and the hidden Fermi-liquid theory proposed by Casey and Anderson (CA) in Ref. [8], yield a spectral function that lacks invariance under Eq. (1). In Ref. [9], a comparison between the ECFL spectral function and a large set of data at optimal doping shows excellent agreement and provides a useful parametrization of the data. To quantify the asymmetry: for optimally doped cuprates, in an energy range of  $\pm 25$  meV around  $\mu$ , the theories and the fits of Ref. [9] (extrapolated to lower  $\omega$ ) yield an asymmetry ratio  $\mathcal{R}$  [defined below Eq. (3)] between  $\sim 7\%$  and  $10\%$ . Because a large asymmetry makes a

decisive ruling on the allowed theories, we propose the direct experimental measurement of this effect and indicate a procedure for the same.

I first discuss a Fermi symmetrization procedure quite distinct from the symmetrization in Refs. [3,10]. I construct an object  $\mathcal{S}_{\mathcal{G}}(\vec{k}, \omega)$  [Eq. (2)] from the observed ARPES intensity and find expressions for this in the Fermi-liquid and the ECFL model. I further show how the momentum dependence of the dipole transition probability and the Fermi-liquid parameter  $Z_k$  can be absorbed into the constants.

The  $\mathcal{S}_{\mathcal{G}}(\vec{k}, \omega)$  function is detailed for a simplified version of ECFL (SECFL), providing an idealized picture of the predicted asymmetry effect in cuprates. I further discuss a related asymmetry of the tunneling conductance in the normal state, and also the expected angle integrated spectrum. Within the SECFL model, where the quasiparticle peaks are sharp over a large fraction of the zone, these exhibit unusual and possibly measurable features.

*Fermi symmetrization.*—Our first goal is to formulate a procedure for isolating terms in the spectral function near the Fermi energy that are linear in wave vector and frequency  $\sim \xi_k - \omega$  (with  $\xi_k = \vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}}$ ) found in the recent work [7]. The ARPES intensity is given in terms of the spectral function within the sudden approximation by the expression  $I(\vec{k}, \omega) = M(\vec{k})f_{\omega}\rho_{\mathcal{G}}(\vec{k}, \omega)$ , where  $M(\vec{k})$  is the dipole transition probability which is expected to be a smooth function of  $\vec{k}$  and independent of  $\omega$ . It also contains the Fermi function for occupied states  $f_{\omega} = \{1 + \exp(\beta\omega)\}^{-1}$ , a nonsymmetric function of  $\omega$ . Therefore, we first formulate a Fermi symmetrized object:

$$\mathcal{S}_{\mathcal{G}}(\vec{k}, \omega) \equiv f_{\omega}\bar{f}_{\omega}\rho_{\mathcal{G}}(\vec{k}, \omega) = \frac{1}{M(\vec{k})}\bar{f}_{\omega}I(\vec{k}, \omega), \quad (2)$$

where  $\bar{f}_{\omega} = 1 - f_{\omega} = f_{-\omega}$ . We may now decompose  $\mathcal{S}_{\mathcal{G}}(\vec{k}, \omega)$  under Eq. (1) into its antisymmetric  $\mathcal{S}_{\mathcal{G}}^{\text{a-s}}(\vec{k}_{\text{F}}|\vec{k}, \omega)$  and symmetric  $\mathcal{S}_{\mathcal{G}}^{\text{s}}(\vec{k}_{\text{F}}|\vec{k}, \omega)$  combinations

respectively  $\frac{1}{2}[\mathcal{S}_{\mathcal{G}}(\vec{k}_{\text{F}} + \vec{k}, \omega) \mp \mathcal{S}_{\mathcal{G}}(\vec{k}_{\text{F}} - \vec{k}, -\omega)]$ . We will also define the important asymmetry ratio:

$$\mathcal{R}_{\mathcal{G}}(\vec{k}_{\text{F}}|\vec{k}, \omega) = \mathcal{S}_{\mathcal{G}}^{\text{a-s}}(\vec{k}_{\text{F}}|\vec{k}, \omega) / \mathcal{S}_{\mathcal{G}}^{\text{s}}(\vec{k}_{\text{F}}|\vec{k}, \omega), \quad (3)$$

where normalization factors cancel out, giving a dimensionless function of order unity. Its magnitude can therefore be compared across different systems. We will quote  $\mathcal{R}_{\mathcal{G}}$  and  $\mathcal{S}_{\mathcal{G}}^{\text{s}}$  below for various theoretical models;  $\mathcal{S}_{\mathcal{G}}^{\text{a-s}}$  can be reconstructed from Eq. (3).

*Dynamical particle-hole symmetry of the Fermi-liquid theory.*—We begin by considering  $\mathcal{S}_{\mathcal{G}}$  for the Fermi-liquid theory. The spectral function of a Fermi-liquid  $\rho_{\mathcal{G}}^{\text{FL}}(\vec{k}, \omega)$  is given in terms of a smooth background plus a quasiparticle peak as in Eq. (4). Near the Fermi surface, we can linearize various objects in  $\vec{k}$  and  $\omega$ . With  $\vec{v}_{\vec{k}_{\text{F}}}$ , the Fermi velocity vector at  $\vec{k}_{\text{F}}$ , the quasiparticle piece is specified by three parameters (i) renormalization factor  $Z_{\vec{k}}$ , with a linear dependence  $Z_{\vec{k}} = Z_{\vec{k}_{\text{F}}}[1 + c_1(\vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}})]$ , (ii) the quasiparticle energy  $E_{\vec{k}}$  vanishing linearly at the Fermi surface  $E_{\vec{k}} = \frac{m}{m^*}(\vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}})$  with an effective mass renormalization  $\frac{m}{m^*}$ , and (iii) the line width  $\gamma_{\vec{k}} \propto [E_{\vec{k}}^2 + (\pi k_{\text{B}}T)^2]$  vanishes symmetrically at the Fermi surface. Thus near the Fermi surface:

$$\rho_{\mathcal{G}}^{\text{FL}}(\vec{k}, \omega) \sim \rho_{\mathcal{G}}^{(\text{bg})}(\vec{k}, \omega) + \frac{Z_{\vec{k}}}{\pi} \frac{\gamma_{\vec{k}}}{\gamma_{\vec{k}}^2 + (\omega - E_{\vec{k}})^2}. \quad (4)$$

For  $\vec{k}$  close to the Fermi surface, the background part is neglected compared to the large quasiparticle part. Defining the quasiparticle peak part

$$\mathcal{Q}(\vec{k}, \omega) = \frac{Z_{\vec{k}_{\text{F}}}}{4\pi \cosh^2(\beta\omega/2)} \frac{\gamma_{\vec{k}_{\text{F}}}}{\gamma_{\vec{k}_{\text{F}}}^2 + [\omega - \frac{m}{m^*}(\vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}})]^2}, \quad (5)$$

we write the Fermi symmetrized functions of  $(\vec{k}, \omega)$ :

$$\{\mathcal{S}_{\mathcal{G}_{\text{FL}}}^{\text{s}}, \mathcal{R}_{\mathcal{G}_{\text{FL}}}\} = \{\mathcal{Q}(\vec{k}, \omega), c_1(\vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}})\}, \quad (6)$$

where we retained only terms linear in  $\vec{k}$ ,  $\omega$  beyond the quasiparticle peak term  $\mathcal{Q}(\vec{k}, \omega)$ . Observe that to  $O(\omega^2)$  the asymmetry ratio  $\mathcal{R}$  is independent of  $\omega$ . The requirement of neglecting the background is necessary, because it is hard to make a general statement about the  $(k, \omega)$  dependence of the background part. Therefore, the discussion becomes sharp only in situations where the peak term overwhelms the background part—thus, forcing us to low temperatures. The same issue also impacts the synchrotron data adversely compared to the laser ARPES data, if we interpret the former to have more substantial elastic scattering correction as argued in Ref. [9].

We make a few remarks next. (1) The coefficient  $c_1$  vanishes in theories where the self-energy is  $\omega$  dependent but  $\vec{k}$  independent. To the extent that we can experimentally identify a  $\omega$  independent but  $k$  dependent term as in Eq. (6), one can say that the Fermi-liquid spectrum possesses the dynamical particle-hole invariance. (2) The momentum dependence of the dipole transition probability  $M(\vec{k})$ , if any, can be absorbed into  $c_1$  in Eq. (6) by Taylor expansion. This implies that the expression [Eq. (6)] is valid for the  $\mathcal{S}$ ,  $\mathcal{R}$  constructed from the ARPES intensities directly [i.e., omitting the  $1/M$  term in Eq. (2)]. The important asymmetry ratio  $\mathcal{R}$  gets rid of the overall scale factors. Therefore, its magnitude is a meaningful quantitative measure of the asymmetry. (3) It follows that the frequency independence of  $\mathcal{R}$  is also true for any theory where the Dyson self-energy  $\Im m \Sigma(k, \omega)$  is even (i.e., not necessarily quadratic) in  $\omega$ , such as the marginal Fermi-liquid [11] and also various refinements of the RPA. Subleading corrections of the type  $\omega \times T^2$  or  $\omega^3$  in  $\Im m \Sigma(k, \omega)$  [5], as well as intrinsic particle-hole asymmetric density of states (DOS) terms can lead to a nontrivial  $\mathcal{R}$ . However, these are estimated [5,6] to be an order of magnitude smaller than the predicted asymmetry of the theories discussed next.

*The asymmetry ratio in ECFL.*—In the recent work on the ECFL [7]  $\rho_{\mathcal{G}}(\vec{k}, \omega)$  is the product of a Fermi-liquid spectral function  $\rho_{\mathcal{G}}^{\text{g}}(\vec{k}, \omega)$  and a caparison factor  $[[1 - \frac{n}{2}] + \frac{\xi_{\vec{k}} - \omega}{\Delta(\vec{k}, \omega)} + \eta(\vec{k}, \omega)]$ , explicitly containing a linear dependence on the energy  $\omega$ . This important term redistributes the dynamical spectral weight within the lower Hubbard band, in such a way as to preserve the Fermi volume. In a further approximation of the formalism, a SECFL theory emerges where we obtain explicit analytical results. In this version,  $\eta(\vec{k}, \omega)$  is negligible and the coefficient  $\Delta$  is a constant determined by the number sum rule. In Refs. [9,12], the SECFL was tested against data on the high  $T_{\text{c}}$  cuprate  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ . The test spans a substantial range of occupied energies  $\sim 1$  eV, with quantitative fits in the 0.25 eV energy range. The remarkably close agreement between data and theory over the broad range of data sets appears to vindicate the form of the spectral function. The test proposed in this Letter is somewhat complementary, it is over a smaller energy range  $\sim 2k_{\text{B}}T$ , probing the asymptotic low energy region centered around the Fermi energy.

With the assumption of a smooth  $k$  dependence of  $\eta(\vec{k}, 0)$  and  $\Delta(\vec{k}, 0)$  in the expression for the spectral function [13] and  $p = d_0 + (1 - \frac{n}{2})$ , we obtain

$$\mathcal{S}_{\mathcal{G}_{\text{ECFL}}} \sim \mathcal{Q} \left[ p + d_1 \vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}} + d_2 \omega + \frac{(\vec{k} \cdot \vec{v}_{\vec{k}_{\text{F}}} - \omega)}{\Delta(\vec{k}_{\text{F}})} \right].$$

Here the term  $d_0$  arises from Taylor expanding  $\eta(\vec{k}_{\text{F}}, 0)$  and also from the shift of the chemical potential from the

free value,  $d_1$  from the momentum dependence of  $Z_k$  and this term can also absorb the momentum dependence of  $M(k)$ , and  $d_2$  from the frequency dependence of  $\eta(k, \omega)$ . We can thus compute the symmetric and antisymmetric parts  $\{\mathcal{S}_{\mathcal{G}_{\text{ECFL}}}^s, \mathcal{R}_{\mathcal{G}_{\text{ECFL}}}\}$  as

$$\sim \left\{ p\mathcal{Q}, \frac{d_1}{p} \vec{k} \cdot \vec{v}_{\vec{k}_F} + \frac{d_2}{p} \omega + \frac{(\vec{k} \cdot \vec{v}_{\vec{k}_F} - \omega)}{p\Delta(\vec{k}_F)} \right\}. \quad (7)$$

The asymmetry ratio  $\mathcal{R}$  therefore has a linear  $\omega$  and  $\vec{k}$  dependence. Using the frequency dependence as the signature, one should be able to distinguish between the results of Eqs. (6) and (7).

The SECFL model is described in detail in Ref. [12], where we write the spectral function near the Fermi energy  $\rho_{\mathcal{G}_{\text{SECFL}}}^{\text{Peak}}(\vec{k}_F + \vec{k}, \omega)$  as

$$\frac{1}{\pi} \frac{Z_k^2 \Gamma_k}{Z_k^2 \Gamma_k^2 + (\omega - E_k^{\text{FL}})^2} \frac{n^2}{4\Delta_0} \{\varepsilon_0 + \xi_k - \omega\}, \quad (8)$$

where  $\varepsilon_0 = \Delta_0 \frac{4}{n^2} (1 - \frac{n}{2})$ . Here  $E_k^{\text{FL}} = Z_k \xi_k$ , in view of the form of the self-energy  $\Phi$ . To leading order, we can set  $Z_k \rightarrow Z_F$  independent of  $k$ , and  $\xi_k = \vec{k} \cdot \vec{v}_{\vec{k}_F}$ ,  $E_k^{\text{FL}} = Z_F \vec{k} \cdot \vec{v}_{\vec{k}_F}$ , and set  $\Gamma_k = \eta + \pi C_\Phi [(\pi k_B T)^2 + (E_k^{\text{FL}})^2]$ , where  $\eta$  is the elastic broadening introduced in Ref. [9] [distinct from  $\eta(\vec{k}, \omega)$ ]. For the model Eq. (8), we can set  $\Gamma_k \rightarrow \Gamma_{k_F}$  and thus obtain the leading behavior near the Fermi energy of  $\{\mathcal{S}_{\mathcal{G}_{\text{SECFL}}}^s, \mathcal{R}_{\mathcal{G}_{\text{SECFL}}}\}$  as

$$\sim \left[ \left(1 - \frac{n}{2}\right) \mathcal{Q}(\vec{k}, \omega), \frac{\{\vec{k} \cdot \vec{v}_{\vec{k}_F} - \omega\}}{\varepsilon_0} \right], \quad (9)$$

where  $\mathcal{Q}(\vec{k}, \omega)$  is obtained from Eq. (5) by replacing  $m/m^* \rightarrow Z_F$  and  $\gamma_k \rightarrow \Gamma_k Z_k$ . Note that, e.g., at  $\vec{k} = \vec{0}$  and any convenient  $\omega_0$ ,  $|\mathcal{R}(0, \omega_0)| = \omega_0/\varepsilon_0$ , and thus its magnitude yields the important energy scale  $\Delta_0$ . We emphasize that Eq. (7) is more generally true within the ECFL approach. We display  $\mathcal{S}^{\text{a-s}}$  in Fig. 1 for a model calculation based in the SECFL model with a flat DOS [see Ref. [12], Sec. (IV.F)]. The values of the basic parameters in all figures are as follows:  $T = 180$  K,  $\omega_c = 0.25$  eV,  $C_\Phi = 1$  (eV) $^{-1}$ . Notice the distinctive increasing linear behavior with  $\vec{k}$  and a decreasing linear one with  $\omega$ , as in Eqs. (7) and (9).

*Single particle tunneling into the extremely correlated state.*—In the simplest model of tunneling in the  $t$ - $J$  model, the conductance is given in terms of the local DOS  $\rho_{\mathcal{G}}^{(\text{local})}(\omega) = \sum_{\vec{k}} \rho_{\mathcal{G}}(\vec{k}, \omega)$ . Its convolution with  $f_\omega$  and  $\bar{f}_\omega$  gives half the occupied  $\frac{n}{2}$ , and the unoccupied  $(1 - n)$  densities, thus providing useful sum rules for tunneling [14]. The sum rule implies asymmetry between adding particles and holes and thus a downward sloping conductance [15, 16]. Recent experiments in the overdoped regime

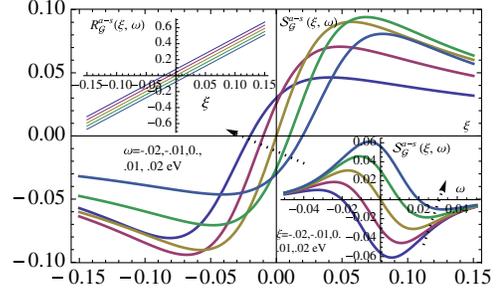


FIG. 1 (color online). Top inset shows the large predicted asymmetry  $\mathcal{R}_{\mathcal{G}_{\text{SECFL}}}^{\text{a-s}}$  versus  $\xi$  in the small energy range of 150 meV. Similar magnitudes are found as functions of  $\omega$  at various  $\xi$ . The figure shows  $\mathcal{S}_{\mathcal{G}_{\text{SECFL}}}^{\text{a-s}}$  from Eq. (9) versus  $\xi$  (main),  $\omega$  (inset) in electron volt at various  $\omega$  (main),  $\xi$  (inset). Arrows indicate the direction of increasing energies. We used  $n = 0.85$ ,  $\eta = 0.05$  eV, and  $\Delta_0 = 0.0796$  eV here.

[17, 18] display the same asymmetry, providing strong confirmation that  $t$ - $J$  model type extreme correlations are operative at high hole doping levels as well, and not just near half filling. More detailed information on the frequency dependence is clearly of experimental interest. We note that the angle integrated photo emission (AIP) technique obtains the local DOS  $\times f_\omega$ , and provides a complementary view to tunneling. Figure 2 presents the results from the SECFL model for both the (local) DOS and DOS  $\times f_\omega$  at various densities and elastic scattering parameter  $\eta$ . It shows an overall decrease of the local DOS with energy. Interestingly, the tunneling curve in the inset (III) shows an upturn followed by a rising piece near  $\omega \sim 0$ , and the AIP curve shows a related shallow minimum at  $\omega \sim -0.2$  eV.

To understand the unusual result, consider integrating the spectral function in Eq. (8) over  $\xi_k$ . As discussed in Refs. [9, 12], when the energy is less than  $\sim 1$  eV, the quasiparticles become sharp and this integral can be

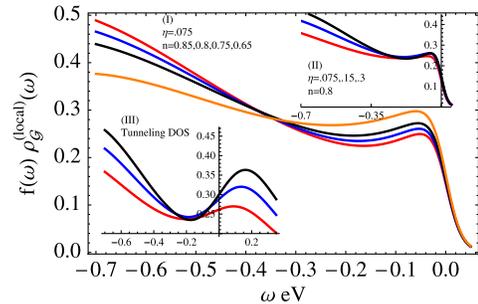


FIG. 2 (color online). (I) The predicted AIP spectrum showing a shallow minimum at  $\omega \sim -0.2$  eV, and a rise as the binding energy  $|\omega|$  increases. The rise is greater as the particle density  $n$  increases (bottom to top). Inset (II) reveals the role of elastic scattering width  $\eta$  (top to bottom). Inset (III) shows the local DOS relevant to the tunneling conductance, for the same parameters as in (II) with a remarkable rising piece near zero bias.

estimated by replacing the Fermi-liquid Lorentzian by  $\delta(\vec{k} \cdot \vec{v}_{k_F} - \frac{m^*}{m}\omega)$ . This yields the quasiparticle peak contribution:

$$\lim_{\omega \rightarrow \varepsilon_0} \rho_{G, \text{Peak}}^{(\text{local})}(\omega) \sim \text{const} \times \left\{ \varepsilon_0 + \left( \frac{m^*}{m} - 1 \right) \omega \right\}. \quad (10)$$

Because  $m \leq m^*$ , it follows that the slope is positive and hence the rising conductance! In the general version of ECFL, different parts of the Fermi surface contribute according to the weight of  $1/\Delta(\vec{k}_F)$ . We expect the resulting average to be less favorable to a rising term than in the SECFL model.

*Other theories.*—CA in Ref. [8] provide a spectral function that may be Taylor expanded at finite  $T$  and low enough energies as follows. With  $q = 1 - \frac{1}{4}n^2$  depending on the filling  $n$ , and  $\Gamma_{\hat{k}} = A(k_B T) + C v_{k_F}^2 \hat{k}^2$ , their expressions yield:

$$\{\mathcal{S}_{CA}^s, \mathcal{R}_{CA}^{a-s}\} = \left\{ \mathcal{Q}', \cot(q\pi/2) \frac{(v_F \hat{k} - \omega)}{\Gamma_0} \right\}, \quad (11)$$

with  $\mathcal{Q}' = \text{const} \times \frac{\sin(q\pi/2)}{4\pi \cosh^2(\beta\omega/2)} / [\Gamma_0^2 + (\omega - v_F \hat{k})^2]^{q/2}$ . Therefore, this work also implies a nontrivial  $\mathcal{R}$  with a linear  $\omega$ ,  $\hat{k}$  dependence, similar in form to that in ECFL, although with a non-Lorentzian peak factor replacing the  $\mathcal{Q}$  factor in Eq. (7). It is seen that the asymmetry of this theory as well as that of the ECFL theory vanishes continuously at low particle density  $n \rightarrow 0$ . An important characteristic energy  $\Delta^*(x, T)$ , say the inverse of the slope of the linear in  $\omega$  term in  $\mathcal{R}$  contains much physics. In the CA theory  $\Delta^*(x, T) \propto \Gamma_0$  vanishes at all densities  $x$  as  $T \rightarrow 0$ , thereby defining a line of quantum critical points. On the other hand in the ECFL calculations, the energy  $\Delta^*(x, T \rightarrow 0)$  is nonzero but much smaller than the (bare) Fermi energy. However, it could vanish at a specific filling  $x_c$ : as  $\Delta^*(x_c, T \rightarrow 0) \rightarrow 0$ , thereby locating an isolated quantum critical point.

Other contemporary theories have a different prediction from the ECFL and CA. The popular marginal Fermi-liquid model [11] for the spectral function has a Dyson self-energy that is symmetric under the transformation Eq. (1). Therefore, it leads to an  $\omega$  independent asymmetry ratio at small energies, as in the usual Fermi-liquid [6]. A similar  $\omega$  independent  $\mathcal{R}$  occurs for the RPA and its many variants emphasizing fluctuation contributions.

*Conclusions.*—The program of extraction of the asymmetry ratio from the “ideal” spectral weight is summarized in Fig. 3. A window of size  $\sim 2k_B T$  in  $\omega$  and  $v_F \hat{k}$  are highlighted in this construction. It is proposed that a careful examination of the ARPES intensity along these lines would determine the existence of dynamical particle-hole asymmetry. This asymmetry also relates to the difference in velocities (and amplitudes) of quasiparticles and quasiholes, of the type that are invoked in explaining the

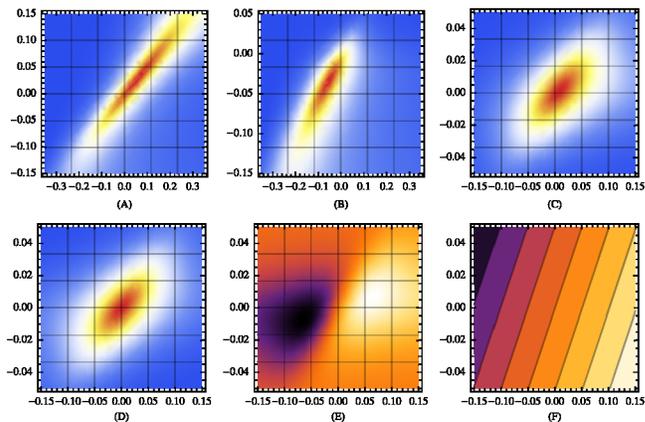


FIG. 3 (color online). Symmetry extraction illustrated for the simplified ECFL model. Here  $n = 0.85$ ,  $\eta = 0.05$ , and  $\Delta_0 = 0.0796$  with  $\omega$  (ordinate) and  $\xi$  (abscts antisymmetricissa) in electron volt. (A) Shows the spectral function  $\rho_G$ , (B)  $\rho_G f_\omega$ , (C)  $\rho_G f_\omega \tilde{f}_\omega$ , (D) the symmetrized object  $\mathcal{S}_G^s$ , (E) the antisymmetrized object  $\mathcal{S}_G^{a-s}$  showing a peak and a trough, and (F) the asymmetry ratio  $\mathcal{R}_G$  from Eq. (3).

peculiar sign of the Hall effect in the mixed state [19]. We thus expect it to be important in Hall and analogous transport contexts such as thermopower. This search is complementary, as well as a prerequisite, to the detailed characterization of the symmetric part  $\mathcal{S}^s$ . Specifically I propose that the search for a nontrivial (i.e.,  $\omega$  linear) asymmetry ratio  $\mathcal{R}$  is important for identifying the correct underlying theoretical description of the cuprates.

In order to implement the transformation Eq. (1) on the experimental data, we need a high resolution in frequency as well as momentum. Because the bare Fermi velocities are high  $\hbar v_F \sim 5 \text{ eV \AA}$ , the momentum resolution becomes critical. An error  $\Delta \xi \sim 15\text{--}20 \text{ meV}$  can lead to quite incorrect conclusions. Thus, in order to draw unambiguous conclusions we require  $\Delta k \sim 0.001 (\text{\AA})^{-1}$ , i.e.,  $\Delta \xi \sim 5 \text{ meV}$  or better, thereby posing an interesting challenge to the experimental ARPES community.

This work was supported by DOE under Grant No. FG02-06ER46319. I thank A. Chubukov, A. Georges, G. H. Gweon, D. Huse, H. R. Krishnamurthy, and T. V. Ramakrishnan for stimulating comments.

- 
- [1] C. P. Herzog, P. Kovtun, S. Sachdev, and D. T. Son, *Phys. Rev. D* **75**, 085020 (2007); T. Senthil, *Phys. Rev. B* **78**, 035103 (2008).
  - [2] S. S. Lee, *Phys. Rev. D* **79**, 086006 (2009).
  - [3] M. Norman *et al.*, *Nature (London)* **392**, 157 (1998); M. R. Norman, M. Randeria, H. Ding, and J. C. Campuzano, *Phys. Rev. B* **57**, R11093 (1998); M. R. Norman, H. Ding, H. Fretwell, M. Randeria, and J. C. Campuzano, *Phys. Rev. B* **60**, 7585 (1999); U. Chatterjee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9346 (2011).

- [4] J. Schmalian and C. D. Batista, *Phys. Rev. B* **77**, 094406 (2008).
- [5] C. Hodges, H. Smith, and J. W. Wilkins, *Phys. Rev. B* **4**, 302 (1971) find a small (odd in  $\omega$ ) correction to the usual  $\omega^2$  behavior  $\Sigma'' \propto \omega^2(1 - \frac{2}{3}\omega/\varepsilon_f)$ .
- [6] A small asymmetry is inevitably induced by DOS variation leading to an estimated asymmetry ratio  $\mathcal{R}$  [Eq. (3)] that is an order of magnitude smaller than that found in the ECFL and CA theories. Also the unavoidable disorder contributes, and thus one may operationally refer to cases as essentially invariant if  $|\mathcal{R}| \leq 1\%$  in an energy window of  $\pm 25$  meV around  $\mu$ .
- [7] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).
- [8] P. A. Casey and P. W. Anderson, *Phys. Rev. Lett.* **106**, 097002 (2011); P. A. Casey, J. D. Koralek, N. C. Plumb, D. S. Dessau, and P. W. Anderson, *Nature Phys.* **4**, 210 (2008); P. W. Anderson, *Nature Phys.* **2**, 626 (2006); S. Doniach and M. Sunjic, *J. Phys. C* **3**, 285 (1970).
- [9] G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).
- [10] The Fermi symmetrization defined in Eq. (2) corrects for the asymmetry of  $I(\omega)$  arising from  $f_\omega$ , by multiplying with  $\bar{f}_\omega$ . The term symmetrization of Ref. [3] refers to a different procedure. At the Fermi momentum, Ref. [3] assumes that  $\rho_G(\omega) = \rho_G(-\omega)$ . Under this assumption, the authors argue that  $I(\omega) + I(-\omega) \rightarrow \rho_G(\omega)M(\vec{k}_F)$ . The asymmetry ratio  $\mathcal{R}$  of Eq. (3) in this case is  $([\bar{f}_\omega I(\omega) - f_\omega I(-\omega)]/[\bar{f}_\omega I(\omega) + f_\omega I(-\omega)])$  and vanishes under their assumption.
- [11] C. M. Varma, P. B. Littlewood, S. Schmitt-Rink, E. Abrahams, and A. E. Ruckenstein, *Phys. Rev. Lett.* **63**, 1996 (1989).
- [12] B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011).
- [13] The energy variables  $\Delta(\vec{k}, \omega)$  and  $\eta(\vec{k}, \omega)$  are defined in Eqs. (21–23) of Ref. [7].
- [14] M. Randeria, R. Sensarma, N. Trivedi, and F. C. Zhang, *Phys. Rev. Lett.* **95**, 137001 (2005).
- [15] P. W. Anderson and N. P. Ong, *J. Phys. Chem. Solids* **67**, 1 (2006).
- [16] T. Hanaguri, C. Lupien, Y. Kohsaka, D.-H. Lee, M. Azuma, M. Takano, H. Takagi, and J. C. Davis, *Nature (London)* **430**, 1001 (2004).
- [17] A. Pasupathy, A. Pushp, K. Gomes, C. Parker, J. Wen, Z. Xu, G. Gu, S. Ono, Y. Ando, and A. Yazdani, *Science* **320**, 196 (2008), Figs. 4A, 1A, 1B.
- [18] C. Renner, B. Revaz, J. Y. Genoud, K. Kadowaki, and O. Fischer, *Phys. Rev. Lett.* **80**, 149 (1998).
- [19] J. M. Harris, N. P. Ong, and Y. F. Yan, *Phys. Rev. Lett.* **71**, 1455 (1993); A. T. Dorsey, *Phys. Rev. B* **46**, 8376 (1992).

## Extremely correlated Fermi liquids: The formalism

B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*

(Received 27 July 2012; published 15 March 2013)

We present the detailed formalism of the extremely correlated Fermi liquid theory, developed for treating the physics of the  $t$ - $J$  model. We start from the exact Schwinger equation of motion for the Green's function for projected electrons, and develop a systematic expansion in a parameter  $\lambda$ , relating to the double occupancy. The resulting Green's function has a canonical part arising from an effective Hamiltonian of the auxiliary electrons, and a caparison part playing the role of a frequency-dependent adaptive spectral weight. This adaptive weight balances the requirement at low  $\omega$  of the invariance of the Fermi volume, and at high  $\omega$  of decaying as  $\frac{c_0}{i\omega}$ , with a correlation-depleted  $c_0 < 1$ . The effective Hamiltonian  $H_{\text{eff}}$  describing the auxiliary fermions is given a natural interpretation with an effective interaction  $V_{\text{eff}}$  containing both the exchange  $J_{ij}$  and the hopping parameters  $t_{ij}$ . It is made Hermitian by adding suitable terms that ultimately vanish, in the *symmetrized theory* developed in this paper. Simple but important *shift invariances* of the  $t$ - $J$  model are noted with respect to translating its parameters uniformly. These play a crucial role in constraining the form of  $V_{\text{eff}}$  and also provide checks for further approximations. The auxiliary and physical Green's function satisfy two sum rules, and the Lagrange multipliers for these are identified. A complete set of expressions for the Green's functions to second order in  $\lambda$  is given, satisfying various invariances. A systematic iterative procedure for higher order approximations is detailed. A superconducting instability of the theory is noted at the simplest level with a high transition temperature.

DOI: [10.1103/PhysRevB.87.125124](https://doi.org/10.1103/PhysRevB.87.125124)

PACS number(s): 71.10.Fd

### I. INTRODUCTION

This work presents the detailed formalism of a newly developed framework for systematic calculation of the dynamical properties of the  $t$ - $J$  model, starting from the basic parameters  $t$  and  $J$  of the model. A subsequent paper Ref. 1 presents self-consistent numerical results from the initial application of this theory, for the case of a two-dimensional square lattice relevant to cuprate superconductors. We will refer to *extreme correlations* as the limit  $U \gg t$ , so that the single-occupancy constraint is enforced. The  $t$ - $J$  model Eq. (12) is the prime example of such a situation. In practice this theory applies already when  $U \gtrsim Zt$ , where  $Z$  is the coordination number of the lattice. The theory and calculations presented are in the *extremely correlated Fermi liquid* (ECFL) phase discussed in Ref. 2. This phase is liquid like, and connects continuously to the Fermi liquid phase of weak-coupling models such as the Hubbard model, while accommodating the extreme local interaction  $U \rightarrow \infty$ .

The  $t$ - $J$  model described by Eq. (12) is one of the standard models of condensed matter physics. It has been the focus of intense effort for the last few decades, as reviewed in Ref. 3. Interest in the model grew particularly after its identification by Anderson in Ref. 4, as governing many of the rich and complex set of phenomena in high- $T_c$  cuprate superconductors. The origin of the exchange part of the  $t$ - $J$  model in an inverse expansion in the interaction  $U$  is familiar from superexchange theory. The relation  $J = \frac{4t^2}{U}$  is found starting from the Hubbard model as in Ref. 5, so that large  $U$  leads to a small  $J$ . An early account of the model and the various sum rules can be found in the Ref. 5. More recently Zhang and Rice<sup>6</sup> gave an argument for reducing the three-band copper oxygen model to an effective single-band model, with a  $t$ - $J$  form. Their method, apart from being more realistic, gives independent magnitudes for  $t$  and  $J$  unconstrained by relations of the type inherent in superexchange within a single-band model.

Controlled calculations within this model are beset by two fundamental difficulties: (a) the noncanonical nature of the single-occupancy (Gutzwiller<sup>7</sup>) projection of the electrons that changes the canonical anticommutation relations to a more nontrivial Hubbard (Lie) algebra and (b) the absence of any obvious small parameter for generating a systematic theory. The present author has recently formulated a method in Ref. 2 and Ref. 8 that overcomes these difficulties to a large extent. The basic idea is to approach the system starting from the limit of low particle density  $n = N_e/N_s$  (ratio of electron number to the number of sites), i.e., a generalized virial expansion. The density can be increased towards half filling systematically, as described below. Early applications to angle-resolved photoemission (ARPES) experiments in Ref. 9 are promising, and the general structure of the solution already leads to nontrivial and experimentally testable predictions in Ref. 10. The present work gives the details of the method introduced in Ref. 2, and carries out a calculation to the lowest nontrivial order in a parameter  $\lambda$  described below. The main elements involved in this framework can be summarized as follows:

(1) *The Schwinger method.* Reference 2 utilizes the key observation that the Schwinger method dispenses with Wicks theorem, and replaces that step of canonical theory by a formal matrix (operator) inversion. The Schwinger equation for the Green's function typically involves a time derivative and a functional derivative with respect to a source potential  $\mathcal{V}$  (defined more fully below). It has the great advantage over standard equations of motion in that the functional derivative generates all required higher order Green's functions. This is unlike, say, the BBGKY hierarchy of quantum statistical mechanics, where one needs to import higher order correlations from elsewhere. For the  $t$ - $J$  model, Ref. 2 obtains an exact Schwinger equation described below in Eq. (42) and Eq. (43). For our purpose, that equation may be illustrated schematically

by the symbolic equation

$$[\hat{G}_0^{-1}(\boldsymbol{\mu}) - Y_1 - X] \cdot \mathcal{G} = (\mathbb{1} - \gamma), \quad (1)$$

where  $\hat{G}_0^{-1}$  [Eq. (64)] is a noninteracting Green's function and  $\gamma$  [Eq. (35)] is essentially the spatially localized but time-dependent Green's function itself  $\sim \mathcal{G}_{\text{local}}$ . Further,  $Y_1$  is a Hartree-type energy and  $X \sim (\text{something}) \times \frac{\delta}{\delta \mathcal{V}}$  contains the all-important functional derivative with respect to  $\mathcal{V}$  [both  $X, Y$  are defined in Eq. (43)]. The undefined ‘‘something’’ lumps together constants and the interaction potential, but is independent of  $\mathcal{G}$ . This is a convenient launching pad provided by Schwinger's method, since it is exact. However, it is also intractable as it stands. There is no obvious small parameter, and the presence of the time-dependent  $\gamma$  on the right-hand side represents the removal of states (and double occupancy) from the canonical theory and creates a new set of problems. We must understand and overcome these in order to create a practical and controlled scheme for calculations. We therefore push forward to the next set of steps.

(2) *Noncanonical nature of the problem and its consequences.* The noncanonical nature of the problem is reflected in the  $\gamma$  term on the right of Eq. (1); it is a time-dependent Green's function obtainable from  $\mathcal{G}$  itself [Eq. (45)]. This  $\gamma$  term contains an essential difficulty of the problem; it has a technical origin that we first discuss, and also an important physical aspect that we describe below.

(a) Consider first the canonical theories, such as the Hubbard model [see Eq. (4) below], where one only has the  $\mathbb{1}$  term on the right-hand side of Eq. (1). In order to get rid of the functional derivative operator  $X$  in favor of a (multiplicative) self-energy, one uses  $X \sim (\text{something}) \times \frac{\delta}{\delta \mathcal{V}}$  to write

$$\begin{aligned} X \cdot \mathcal{G} &\rightarrow \Sigma \mathcal{G} \equiv (\text{something}) \times \mathcal{G} \Gamma \mathcal{G}, \\ \text{using } \frac{\delta}{\delta \mathcal{V}} \mathcal{G} &= \mathcal{G} \Gamma \mathcal{G}, \end{aligned} \quad (2)$$

$$\text{following from } \Gamma \equiv -\frac{\delta}{\delta \mathcal{V}} \mathcal{G}^{-1},$$

wherein the vertex  $\Gamma$  is introduced. This gives the Schwinger-Dyson relationship between the self-energy  $\Sigma$  and vertex:

$$\Sigma = (\text{something}) \times \mathcal{G} \Gamma, \quad \text{so that } (\hat{G}_0^{-1}(\boldsymbol{\mu}) - Y_1 - \Sigma) \mathcal{G} = \mathbb{1}. \quad (3)$$

This Schwinger-Dyson construction necessarily requires that the vertex  $\Gamma$  reduce to unity at high frequencies, i.e., should be ‘‘asymptotically free.’’ In the case of the noncanonical theory Eq. (1), a similar procedure fails. It is easily verified that the required good behavior is lost because of the time-dependent term  $\gamma$  on the right-hand side of Eq. (1), as shown in Ref. 11. The so defined vertex grows linearly with frequency, and invalidates the Dysonian self-energy scheme.

(b) The physical problem that is related to the noncanonical  $\gamma$  term has to do with the spectral weight of the projected electrons in a  $t$ - $J$  model. Here basic sum rules give us insight into the origin, as well as a resolution of this fundamental problem. For noncanonical electrons, the high-frequency behavior of the Green's function is  $\mathcal{G} \sim \frac{c_0}{i\omega}$  with  $c_0 = 1 - \frac{n}{2}$ , rather than the familiar result for canonical electrons  $c_0 = 1$ . The depletion of  $c_0$  from unity arises from the physics of

single-occupancy projection of the (noncanonical) electrons  $\hat{c}_{i\sigma}$  (denoted by the Hubbard operator  $X_i^{\sigma 0}$  below). Consider the relation  $c_0 = \langle \hat{c}_{i\sigma} \hat{c}_{i\sigma}^\dagger + \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma} \rangle$ ; the process  $\langle \hat{c}_{i\sigma} \hat{c}_{i\sigma}^\dagger \rangle$  suffers from the inhibiting requirement that in order to create an electron with spin  $\sigma$ , the spin state  $\bar{\sigma}$  at site  $i$  must also be unoccupied (so that a double occupancy is not created by this process), resulting in  $c_0 < 1$ . On the other hand, if the numerator of  $\mathcal{G}(i\omega)$  remains as  $c_0$  at all frequencies, then the Fermi surface must enlarge in volume, and thereby violate the Luttinger-Ward theorem of invariance of this volume.<sup>12</sup> We thus arrive at an appreciation of the fundamental tension between the conflicting requirements: at high frequency of fixing a known coefficient  $c_0 < 1$ , and at low frequency of a numerator almost unshifted from unity, for preserving the Fermi surface volume. A resolution is provided by the possibility of an adaptive (or smart) spectral weight, i.e., the numerator of the Green's function. If a frequency-dependent spectral weight can be found, so as to interpolate smoothly between the high- and low-frequency requirements, then both could be satisfied.

(c) *The product ansatz:* The above points suggest that the Green's function of the  $t$ - $J$  model is usefully thought of as a product of two terms in frequency space i.e.,  $\mathcal{G} \sim \mathbf{g} \times \mu$  [Eq. (44)], where  $\mathbf{g}$  is a canonical Green's function and  $\mu$  the caparison factor playing the role of an adaptive (or smart) spectral weight factor. The  $\mathbf{g}$  term (i.e., the denominator) is required to be a canonical object with its poles and cuts as usual in a Fermi liquid, and defines the auxiliary Fermi liquid in this theory. The frequency-dependent  $\mu$  term (in the numerator) plays the role of the smart spectral weight; it reduces to the correct coefficient  $c_0$  at high frequencies while recovering weight at lower frequencies. Thus a convolution in time domain into two suitable time-dependent pieces could resolve this conundrum, and motivates the *product ansatz* in Eq. (44). This *product ansatz* is at the heart of the procedure described here and is seen to lead to a pair of *exact* equations for the two parts  $\mathbf{g}$  and  $\mu$  below in Eq. (67) and Eq. (68).

(d) The  $\mu$  term is also termed the caparison factor in Ref. 2, keeping in mind that it provides a second layer of dressing, over and above the dressing provided by the usual Fermi-liquid-type processes in  $\mathbf{g}$  itself.

(3) *Small parameter in theory.* The  $t$ - $J$  model is the sum of two highly nontrivial terms, the kinetic energy projected to the space of single occupancy, and the exchange energy. It has no obvious small parameters making it especially difficult to deal with. Some inspiration is gained by examining the form of the analogous Schwinger equation for canonical theories, such as the Hubbard model. Again omitting details, the relevant equation can be written symbolically as

$$\left[ \hat{G}_0^{-1}(\boldsymbol{\mu}) - U G - U \frac{\delta}{\delta \mathcal{V}} \right] \cdot G = \mathbb{1}, \quad (4)$$

where  $U$  is the Coulomb repulsion in the Hubbard model. Comparing with Eq. (1) suggests a simple approach to introduce a new parameter  $\lambda$ . In its simplest form, we propose to study the modified problem symbolically expressed as

$$[\hat{G}_0^{-1}(\boldsymbol{\mu}) - \lambda Y_1 - \lambda X] \cdot \mathcal{G} = (\mathbb{1} - \lambda \gamma), \quad (5)$$

with  $0 \leq \lambda \leq 1$ , so that this equation Eq. (5) interpolates smoothly between the Fermi gas and the  $t$ - $J$  model. This appearance of the parameter parallels the way the Hubbard parameter  $U$  enters Eq. (4). The complication of the noncanonical  $\gamma$  term on the right is handled analogously to the Hartree term  $Y_1$ . Unlike the repulsive Hubbard case, with an infinite interval  $[0, \infty]$  for  $U$ , the parameter  $\lambda$  lives in a small and finite interval  $[0, 1]$ . The expectation is that low-order perturbation expansion in  $\lambda$  has a reasonable chance of capturing the physics of extreme correlations at  $\lambda = 1$ . We show in Appendix A that in the atomic limit, the role of  $\lambda$  can be explicitly related to that of the fraction of double occupancy (and thus also density), so that tuning  $\lambda$  smoothly adjusts this fraction between its two limits. Further in Eq. (11) below, a suggestive expression for the fermionic operators is noted that relates  $\lambda < 1$  to a soft version of Gutzwiller projection.

(4) *Effective Hamiltonian for the auxiliary fermions with a pseudopotential.* Setting aside the caparison factor  $\mu$  for a moment, we examine further the equations of motion [Eq. (22) and Eq. (26)] for the auxiliary fermion  $\mathbf{g}$  following from Eq. (5) together with the product ansatz  $\mathcal{G} = \mathbf{g} \times \mu$ . We would like to interpret these as the actual (canonical) equations of a suitable Fermi liquid, obtainable from a Hermitian Hamiltonian. However, we find that the equations [Eq. (22) and Eq. (26)] as they stand do not immediately cooperate with this task. They require a process of symmetrization described next, where one adds extra terms that vanish when treated exactly, and after this lead to a Hermitean theory for  $\mathbf{g}$ . We term the resulting equations as the *symmetrized theory*, as outlined in this paper.

The theory based on Eq. (22) and Eq. (26) without symmetrization is of course also exact, and is potentially useful in its own right. We develop such a *minimal theory* elsewhere, with the expectation that this *minimal theory* would not admit a Hermitian Hamiltonian to describe the auxiliary  $\mathbf{g}$ . Also in an approximate treatment, e.g., through an expansion in the parameter  $\lambda$  to any fixed but finite order, we would expect the symmetrized and minimal versions of the theory to be different, converging only when *all orders* are taken into account.

Returning to the symmetrization procedure, we construct an effective Hamiltonian  $H_{\text{eff}}$  for canonical electrons ( $f_{i\sigma}, f_{i\sigma}^\dagger$ ), with the property that the (imaginary time) Heisenberg equation of motion for canonical electrons  $\dot{f}_{i\sigma} = -[f_{i\sigma}, H_{\text{eff}}]$  matches exactly the Heisenberg equation of motion for projected electrons  $\dot{\hat{c}}_{i\sigma} = -[\hat{c}_{i\sigma}, H_{t-J}]$ , except for terms that vanish on enforcing the single-occupancy constraint on the auxiliary  $f_{i\sigma}$  electrons. Thus we require

$$\begin{aligned} [f_{i\sigma}, H_{\text{eff}}] &= ([\hat{c}_{i\sigma}, H_{t-J}]_{(\hat{c}, \hat{c}^\dagger) \rightarrow (f, f^\dagger)} \\ &+ (\text{expressions involving } f, f^\dagger \text{ that vanish} \\ &\text{at single occupancy}). \end{aligned} \quad (6)$$

We can then add these missing terms with  $(f, f^\dagger) \rightarrow (\hat{c}, \hat{c}^\dagger)$  to the Heisenberg equation of motion (EOM) for  $\hat{c}$  and thereby obtain an auxiliary Fermi liquid that would be also “natural,” i.e., have all the standard properties of a Fermi liquid.<sup>13,14</sup> One should therefore be able to use standard Feynman diagrams

(Ref. 13) to compute the properties of this auxiliary theory in powers of  $\lambda$ , if one were so inclined.

We find it straightforward to find such an effective Hamiltonian  $H_{\text{eff}}$  [Eq. (27)] as described below in Sec. II C. The physical meaning of  $H_{\text{eff}}$  becomes clearer with the following remarks. The kinetic energy of the projected electrons could also be written differently. An alternate representation, occasionally used in literature, relates

$$\hat{c}_{i\sigma}^\dagger = X_i^{\sigma 0} \rightarrow f_{i\sigma}^\dagger (1 - n_{i\bar{\sigma}}), \quad \hat{c}_{i\sigma} = X_i^{0\sigma} \rightarrow f_{i\sigma} (1 - n_{i\bar{\sigma}}) \quad (7)$$

with  $\bar{\sigma} = -\sigma$  and  $n_{i\sigma} = f_{i\sigma}^\dagger f_{i\sigma}$ . Within this representation, the Hilbert space continues to allow for double occupancy, i.e., is canonical, but the various operators representing the physical processes act only upon the singly occupied subspace, *and produce states that are likewise singly occupied*. Thus we may write the kinetic energy part as

$$KE = - \sum_{ij} t_{ij} (1 - n_{i\bar{\sigma}}) f_{i\sigma}^\dagger f_{j\sigma} (1 - n_{j\bar{\sigma}}). \quad (8)$$

Since the exchange energy  $\sum_{ij} J_{ij} \vec{S}_i \cdot \vec{S}_j$  automatically conserves single occupancy, we will not write it out. The kinetic energy is thus a multi-Fermi operator and represents both the propagation and interaction between particles. To separate these functionalities, we introduce a parameter  $\lambda$  here—it will turn out to be the same parameter as in Eq. (5)—and write

$$\begin{aligned} KE(\lambda) &= - \sum_{ij} t_{ij} (1 - \lambda n_{i\bar{\sigma}}) f_{i\sigma}^\dagger f_{j\sigma} (1 - \lambda n_{j\bar{\sigma}}) \\ &= - \sum_{ij} t_{ij} f_{i\sigma}^\dagger f_{j\sigma} + \lambda \sum_{ij} t_{ij} f_{i\sigma}^\dagger f_{j\sigma} (n_{i\bar{\sigma}} + n_{j\bar{\sigma}}) \\ &\quad + \lambda^2 H_d, \end{aligned} \quad (9)$$

$$H_d = - \sum_{ij} t_{ij} f_{i\sigma}^\dagger f_{j\sigma} (n_{i\bar{\sigma}} n_{j\bar{\sigma}}), \quad H_d \rightarrow \text{dropped}. \quad (10)$$

The term  $H_d$  acts on the doubly occupied subspace and is null in the singly occupied space, and hence it may be dropped altogether. The remaining part of the kinetic energy term  $KE(\lambda)$  has the structure of a four-Fermi interaction between the canonical fermions, and turns out to be a large part of  $H_{\text{eff}}$  in Eq. (27). The introduction of the parameter  $\lambda$  can thus be viewed as replacing Eq. (7) by a “softer” representation of the Gutzwiller projection:

$$\hat{c}_{i\sigma}^\dagger \rightarrow f_{i\sigma}^\dagger (1 - \lambda n_{i\bar{\sigma}}), \quad \hat{c}_{i\sigma} \rightarrow f_{i\sigma} (1 - \lambda n_{i\bar{\sigma}}). \quad (11)$$

This  $\lambda$  representation discourages but does not completely eliminate double occupancy. However as  $\lambda \rightarrow 1$ , it does become the exact projected operators Eq. (7), and further provides a simple interpolation between standard (canonical) fermions and the projected electrons by varying  $\lambda$  in the range  $0 \leq \lambda \leq 1$ . Thus Eq. (11) suggests the interpretation of the parameter  $\lambda$  as the controller of the (partial) Gutzwiller projection.

In this representation (with  $\lambda = 1$ ), the physical electron Green’s function  $\mathcal{G}_{ij}$  corresponds to the correlator  $-\langle\langle (1 - n_{i\bar{\sigma}_i}) f_{i\sigma_i}, f_{j\sigma_j}^\dagger (1 - n_{i\bar{\sigma}_j}) \rangle\rangle$ , while  $-\langle\langle f_{i\sigma_i}, f_{j\sigma_j}^\dagger \rangle\rangle$  would represent

the auxiliary Green's function  $\mathbf{g}(i, j)$ . The caparison factor  $\mu$  seems hard to interpret in this language, though. The ECFL formalism developed here presents a procedure to splice together  $\mathbf{g}$  and  $\mu$  precisely, to yield the physical  $\mathcal{G}$ . Its otherwise formal structure becomes clearer upon making the above connection; in particular Eq. (11) helps in developing some intuition for  $\mathbf{g}$ . For instance a physical interpretation of the auxiliary fermions is provided by the  $f_{i\sigma}$  themselves, and thereby requiring the same number of auxiliary fermions as the physical ones, as done below, is perfectly natural.

(5) *Invariances of the effective Hamiltonian  $H_{\text{eff}}$  and the emergence of the second chemical potential  $u_0$ .* In  $H_{\text{eff}}$  [Eq. (27)], the hopping parameter  $t_{ij}$  is elevated to the role of an interaction coupling, in addition to its role as a band-hopping parameter. This feature needs attention, since we know that a constant ( $k$  independent) shift of the band energies  $\varepsilon_k \rightarrow \varepsilon_k + u_t$ , or adding an on-site interaction through  $J_{ij} \rightarrow J_{ij} + \delta_{ij}u_J$ , is inconsequential for the  $t$ - $J$  model, but makes a difference in Eq. (9), and in various approximations for the  $t$ - $J$  model. This “pure” gauge invariance is of primary importance in this kind of a theory, and must be addressed at the very outset to obtain a consistent and meaningful description of the  $t$ - $J$  model. Such shifts could potentially lead to a change of the interaction strengths in  $H_{\text{eff}}$ , unless they can be explicitly eliminated in the theory. This issue is addressed by first listing these *shift symmetries* of the model in Sec. II, and then requiring the approximation scheme to be shift invariant, at each order of  $\lambda$ .

Imposing the shift symmetries on  $H_{\text{eff}}$  Eq. (27) causes it to have a term with a Hubbard-Coulomb-like interaction with strength  $u_0$ , such that arbitrary shifts of  $t$  and  $J$  can be absorbed into the parameter  $u_0$ . Analogous to the standard chemical potential  $\mu$ , this  $u_0$  is a Lagrange multiplier of a term in the Hamiltonian  $H_{\text{eff}}$ . However it multiplies an interaction term that is *quartic* in the canonical fermions, unlike  $\mu$  that multiplies the usual (quadratic) number operator. The chemical potential  $\mu$  and the second chemical potential  $u_0$  are jointly determined by *two sum rules* Eq. (90) and Eq. (91), one for the number of physical electrons and the other for the (identical) number of auxiliary canonical electrons.

In this work, we obtain a set of equations for the Green's function. These are essentially of the same form as in our recent earlier Letter Ref. 2, but differ in a few details due to the usage of the idea of the effective Hamiltonian and its shift invariances. An iterative framework is carefully established, and calculations of the Green's function to second order in  $\lambda$  are carried out explicitly.

The outline of the paper is as follows. In Sec. II, we list the *shift symmetries* of the  $t$ - $J$  model and obtain the exact equation satisfied by the Green's function. We also determine the form of the effective Hamiltonian  $H_{\text{eff}}$  for the auxiliary fermions, such that the Heisenberg equations for the field operators are satisfied in a Hermitian framework. In Secs. III and IV, we use the *product ansatz* for the Green's function to introduce and find the exact equations for the auxiliary fermions and the caparison factor  $\mu$ . In Sec. V we turn off the time-dependent sources and write the exact momentum-space relations between the self-energy, the caparison factor, and the physical Green's functions—these are the analogs of the Schwinger-Dyson equations for this problem. Section VI

summarizes in tabular form the necessary equations needed for the next step in the iterative process that is analogous to the skeleton graph expansion. Section VII describes the  $\lambda$  expansion of various objects and the precise nature of the iterative expansion. Several detailed calculations are needed to obtain the second-order equations, and are detailed in Appendix B. Section VIII details the Ward identities of this theory, which splits into two parts following the splitting of the Green's functions. Section IX gives the set of vertices defining the random phase approximation for this theory and Sec. X gives the formal results for the charge and spin susceptibilities within RPA and its low-order expansion. Section XI concludes with some comments including a calculation of the superconducting transition temperature in this theory.

Appendix A gives a detailed calculation in the atomic limit. The simple calculation here may be useful in providing the reader some insight into the interpretation of the  $\lambda$  expansion in terms of the number of doubly occupied sites. Appendix B contains the detailed calculations of the various objects need to compile the second-order Green's function.

## II. THE $t$ - $J$ MODEL AND ITS SHIFT INVARIANCE

We write the projected Fermi operators in terms of the Hubbard  $X$  operators as usual  $\hat{c}_{i\sigma} \rightarrow X_i^{0\sigma}$ ,  $\hat{c}_{i\sigma}^\dagger \rightarrow X_i^{\sigma 0}$ , and  $\hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma} \rightarrow X_i^{\sigma'\sigma}$ . We study the  $t$ - $J$  model given by

$$\begin{aligned} H &= - \sum_{i,j,\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i,\sigma} X_i^{\sigma\sigma} \\ &\quad + \frac{1}{2} \sum_{i,j} J_{ij} \left\{ \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j \right\}, \\ &= - \sum_{i,j,\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i,\sigma} X_i^{\sigma\sigma} \\ &\quad + \frac{1}{4} \sum_{ij,\sigma} J_{ij} (X_i^{\sigma\bar{\sigma}} X_j^{\bar{\sigma}\sigma} - X_i^{\sigma\sigma} X_j^{\bar{\sigma}\bar{\sigma}}). \end{aligned} \quad (12)$$

We will treat the two terms on an equal footing as far as possible, and allow terms with  $i = j$ . The statement of the model is invariant under a particular “pure gauge” transformation that we next discuss. Let us first note the *shift invariance* of the two parameters in  $H$ . Consider the uniform (i.e., space independent) shifts of the basic parameters:

$$t_{ij} \rightarrow t_{ij} - u_t \delta_{ij}, \quad J_{ij} \rightarrow J_{ij} + u_J \delta_{ij}, \quad (13)$$

with independent parameters  $u_t, u_J$ . Under this transformation the Hamiltonian shifts as

$$H \rightarrow H + (u_t + \frac{1}{4}u_J) \hat{N}, \quad (14)$$

where  $\hat{N} = \sum_{i\sigma} X_i^{\sigma\sigma}$  is the number operator for the electrons. Let us note two simple theorems encoding this invariance:

*Shift theorem I.* A shift of either  $t$  or  $J$  can be absorbed into suitable parameters, leaving the physics unchanged.

*Shift theorem II.* The two shifts of  $t$  and  $J$  cancel each other when  $u_J = -4u_t$ .

The first theorem is illustrated in the initial Hamiltonian Eq. (12), where the shift in Eq. (14) can be absorbed in the chemical potential  $\mu \rightarrow \mu + u_t + \frac{1}{4}u_J$ . Later it serves to identify a second generalized chemical potential  $u_0$  encountered

in the following. The second theorem is subtle as it leaves the chemical potential  $\mu$  unchanged (see Ref. 15). It provides a measure of the equal-handed treatment of  $t$  and  $J$ . We will find these almost trivial theorems of great use in devising and validating various approximation schemes later.

In further work we need to add a source term via the operator  $\mathcal{A}$ :

$$\begin{aligned} \mathcal{A} &= \int_0^\beta \mathcal{A}(\tau) d\tau = \sum_{j,\sigma_1,\sigma_2} \int_0^\beta d\tau \mathcal{V}_j^{\sigma_1\sigma_2}(\tau) X_j^{\sigma_1\sigma_2}(\tau) \\ &+ \sum_{ij,\sigma_1\sigma_2} \int_0^\beta d\tau \mathcal{V}_{ij}^{\sigma_1\sigma_2}(\tau) X_i^{\sigma_1\sigma_2}(\tau) X_j^{0\sigma_2}(\tau), \end{aligned} \quad (15)$$

with the usual imaginary-time Heisenberg picture  $\tau$  dependence of the operators  $Q(\tau) = e^{\tau H} Q e^{-\tau H}$ , and the bosonic sources  $\mathcal{V}_j^{\sigma_1\sigma_2}(\tau)$  at every site and also  $\mathcal{V}_{ij}^{\sigma_1\sigma_2}(\tau)$  for every pair of sites, as arbitrary functions of time. We will denote these sources in a compact notation where the site index also carries the time argument as  $\mathcal{V}_i^{\sigma_1\sigma_2} \equiv \mathcal{V}_i^{\sigma_1\sigma_2}(\tau_i)$  and  $\mathcal{V}_{ij}^{\sigma_1\sigma_2} \equiv \mathcal{V}_{ij}^{\sigma_1\sigma_2}(\tau_i)\delta(\tau_i - \tau_j)$ . For any variable we define a modified expectation

$$\langle\langle Q(\tau_1, \tau_2, \dots) \rangle\rangle = \frac{\text{Tr}[e^{-\beta H} T_\tau e^{-\mathcal{A}} Q(\tau_1, \tau_2, \dots)]}{\text{Tr}[e^{-\beta H} T_\tau (e^{-\mathcal{A}})]}, \quad (16)$$

with a compact notation that includes the (imaginary) time ordering symbol  $T_\tau$  and the exponential factor automatically. With the abbreviation  $i \equiv (R_i, \tau_i)$  for spatial  $\vec{R}_i$  and imaginary-time ( $\tau$ ) coordinates, the physical electron is described by a Green's function:

$$\mathcal{G}_{\sigma_i\sigma_f}[i, f] = -\langle\langle X_i^{0\sigma_i} X_f^{\sigma_f 0} \rangle\rangle. \quad (17)$$

From this, the variation can be found from functional differentiation as

$$\begin{aligned} \frac{\delta}{\delta \mathcal{V}_j^{\sigma_1\sigma_2}(\tau_1)} \langle\langle Q(\tau_2) \rangle\rangle \\ = \langle\langle Q(\tau_2) \rangle\rangle \langle\langle X_j^{\sigma_1\sigma_2}(\tau_1) \rangle\rangle - \langle\langle X_j^{\sigma_1\sigma_2}(\tau_1) Q(\tau_2) \rangle\rangle. \end{aligned} \quad (18)$$

We note the fundamental anticommutator between the destruction and creation operators:

$$\{X_i^{0\sigma_1}, X_j^{\sigma_2 0}\} = \delta_{ij} (\delta_{\sigma_1\sigma_2} - (\sigma_1\sigma_2) X_i^{\bar{\sigma}_1\bar{\sigma}_2}). \quad (19)$$

### A. The Heisenberg equation of motion

Let us now study the time evolution of the destruction operator through its important commutator:

$$\begin{aligned} [X_i^{0\sigma_i}, H] &= - \sum_j t_{ij} [\delta_{\sigma_i\sigma_j} - (\sigma_i\sigma_j) X_i^{\bar{\sigma}_i\bar{\sigma}_j}] X_j^{0\sigma_j} + \underbrace{\frac{1}{4} J_0 X_i^{0\sigma_i}} \\ &- \mu X_i^{0\sigma_i} - \frac{1}{2} \sum_{j \neq i} J_{ij} (\sigma_i\sigma_j) X_j^{\bar{\sigma}_i\bar{\sigma}_j} X_i^{0\sigma_j}. \end{aligned} \quad (20)$$

Here  $J_0$  is the zero wave vector (i.e.,  $J_{ii}$  the on-site) exchange constant. The term in underbraces here and in the next equation ensures that the commutator reproduces the term with  $J_{ij} \rightarrow J_{ij} + u_J \delta_{ij}$  correctly. We note that under the transformation Eq. (14), the last term in Eq. (20) adds nothing, in view of the ordering of the operators as written, while the term with underbraces provides the correct transformation factor. Let us call this commutator

$$[X_i^{0\sigma_i}, H] = - \sum_j t_{ij} X_j^{0\sigma_j} + \underbrace{\frac{1}{4} J_0 X_i^{0\sigma_i}} - \mu X_i^{0\sigma_i} + A_{i\sigma_i}, \quad (21)$$

$$A_{i\sigma_i} = \sum_{j\sigma_j} t_{ij} (\sigma_i\sigma_j) X_i^{\bar{\sigma}_i\bar{\sigma}_j} X_j^{0\sigma_j} - \frac{1}{2} \sum_{j \neq i} J_{ij} (\sigma_i\sigma_j) X_j^{\bar{\sigma}_i\bar{\sigma}_j} X_i^{0\sigma_j}. \quad (22)$$

We next express the EOM for the Green's function in terms of  $A$ .

### B. Equation of motion for $\mathcal{G}$

Let us compute the time derivative of  $\mathcal{G}$ . For this we need the derivative

$$\begin{aligned} \partial_{\tau_i} T_\tau (e^{-\mathcal{A}} X_i^{0\sigma_i}(\tau_i)) &= -T_\tau (e^{-\mathcal{A}} [X_i^{0\sigma_i}(\tau_i), H]) + T_\tau (e^{-\mathcal{A}} [\mathcal{A}(\tau_i), X_i^{0\sigma_i}(\tau_i)]), \\ [\mathcal{A}(\tau_i), X_i^{0\sigma_i}(\tau_i)] &= \mathcal{V}_i^{\sigma_1\sigma_2}(\tau_i) [X_i^{\sigma_1\sigma_2}(\tau_i), X_i^{0\sigma_i}(\tau_i)] - \sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2}(\tau_i) \{X_i^{\sigma_1\sigma_2}(\tau_i), X_i^{0\sigma_i}(\tau_i)\} X_j^{0\sigma_2}(\tau_i) \\ &= -\mathcal{V}_i^{\sigma_1\sigma_2} X_i^{0\sigma_2} - \sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2} X_j^{0\sigma_2} + \sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2} (\sigma_1\sigma_i) X_i^{\bar{\sigma}_i\bar{\sigma}_1} X_j^{0\sigma_2}. \end{aligned} \quad (23)$$

This follows from the definition of the time ordering and Eq. (15) for  $\mathcal{A}$ . Using this we find

$$\begin{aligned} \partial_{\tau_i} \mathcal{G}_{\sigma_i\sigma_f}[i, f] &= -\delta(\tau_i - \tau_f) \delta_{i,f} \langle\langle (\delta_{\sigma_i\sigma_f} - \sigma_i\sigma_f X_i^{\bar{\sigma}_i\bar{\sigma}_f}) \rangle\rangle + \langle\langle [X_i^{0\sigma_i}(\tau_i), H] X_f^{\sigma_f 0}(\tau_f) \rangle\rangle \\ &- \mathcal{V}_i^{\sigma_1\sigma_2}(\tau_i) \mathcal{G}_{\sigma_2\sigma_f}[i, f] - \sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2} \mathcal{G}_{\sigma_2\sigma_f}[j, f] - \sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2} (\sigma_1\sigma_i) \langle\langle X_i^{\bar{\sigma}_i\bar{\sigma}_1}(\tau_i) X_j^{0\sigma_2}(\tau_j) X_f^{\sigma_f 0}(\tau_f) \rangle\rangle. \end{aligned} \quad (24)$$

To simplify notation, in such expressions for the Green's functions [or Eq. (26) below], the sum over an index implies a sum over the corresponding site and also an integration over the corresponding time; e.g.,  $\sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2} f(\dots, \tau_j, \dots) \rightarrow \sum_{R_j} \int_0^\beta d\tau_j \mathcal{V}_{ij}^{\sigma_1\sigma_2}(\tau_j) \delta(\tau_i - \tau_j) f(\dots, \tau_j, \dots)$ . A further bold letter summation convention is used after Eq. (41). However, note that in expressions for operators such as Eq. (21) or Eq. (22), the sum only refers to the site index summation. We further

use the abbreviations

$$\begin{aligned} \delta[i, j] &= \delta_{i, j} \delta(\tau_i - \tau_j), \quad t[i, j] = t_{ij} \delta(\tau_i - \tau_j), \\ J[i, j] &= J_{ij} \delta(\tau_i - \tau_j) \quad \mathcal{V}_r^{\sigma_a \sigma_b} = \mathcal{V}_r^{\sigma_a \sigma_b}[\tau_r]. \end{aligned} \quad (25)$$

In terms of these, and using Eq. (21), we find the equation of motion in terms of  $A$ :

$$\begin{aligned} (\partial_{\tau_i} - \mu) \mathcal{G}_{\sigma_i \sigma_f}[i, f] &= -\delta[i, f] \langle \langle \delta_{\sigma_i, \sigma_f} - \sigma_i \sigma_f X_i^{\bar{\sigma}_i \bar{\sigma}_f} \rangle \rangle + t[i, j] \mathcal{G}_{\sigma_i \sigma_f}[j, f] - \frac{1}{4} J_0 \mathcal{G}_{\sigma_i \sigma_f}[i, f] + \langle \langle A_{i\sigma_i}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \rangle \rangle \\ &\quad - \mathcal{V}_i^{\sigma_1 \sigma_j}(\tau_i) \mathcal{G}_{\sigma_j \sigma_f}[i, f] - \sum_j \mathcal{V}_{ij}^{\sigma_i \sigma_2} \mathcal{G}_{\sigma_2 \sigma_f}[j, f] - \sum_j \mathcal{V}_{ij}^{\sigma_1 \sigma_2}(\sigma_1 \sigma_i) \langle \langle X_i^{\bar{\sigma}_i \bar{\sigma}_1}(\tau_i) X_j^{0 \sigma_2}(\tau_j) X_f^{\sigma_f 0}(\tau_f) \rangle \rangle. \end{aligned} \quad (26)$$

We recall from the introduction the discussion regarding suitably generalizing  $A$  of Eq. (22), in order to make connection with a Hermitian  $H_{\text{eff}}$ , and therefore turn to this task next.

### C. Effective Hamiltonian

We now construct an effective Hamiltonian of canonical fermions that will turn out to govern the auxiliary Fermi liquid theory. The motivation for this construction is to cast the auxiliary fermionic part of the ECFL theory into a natural and canonical framework, so that the equation for the  $\mathbf{g}$ , i.e., the auxiliary piece of the full  $\mathcal{G}$ , is obtainable from a Hamiltonian that is Hermitian and respects the usual Fermi symmetry of interactions under exchange.

After some inspections we find that a suitable Hamiltonian is provided by the expression

$$\begin{aligned} H_{\text{eff}} &= - \sum_{ij} t_{ij} f_{i\sigma}^\dagger f_{j\sigma} + \sum_i \left( \frac{1}{4} J_0 - \mu \right) f_{i\sigma}^\dagger f_{i\sigma} + \lambda V_{\text{eff}}, \\ V_{\text{eff}} &= \frac{1}{4} \sum_{ij} t_{ij} (\sigma_1 \sigma_2) [(f_{i\sigma_1}^\dagger f_{i\bar{\sigma}_1}^\dagger + f_{j\sigma_1}^\dagger f_{j\bar{\sigma}_1}^\dagger) f_{i\bar{\sigma}_2} f_{j\sigma_2} + (\text{H.c.})] \\ &\quad - \frac{1}{4} \sum_{ij} J_{ij} (\sigma_1 \sigma_2) f_{i\sigma_1}^\dagger f_{j\bar{\sigma}_1}^\dagger f_{j\bar{\sigma}_2} f_{i\sigma_2} \\ &\quad + \frac{1}{4} \sum_i u_0 (\sigma_1 \sigma_2) f_{i\sigma_1}^\dagger f_{i\bar{\sigma}_1}^\dagger f_{i\bar{\sigma}_2} f_{i\sigma_2}, \end{aligned} \quad (27)$$

with a Hermitian effective potential  $V_{\text{eff}}^\dagger = V_{\text{eff}}$  (Fig. 1), and assume no constraint on double occupancy for these auxiliary (canonical) fermions  $f_{i\sigma}$ . The  $t$  and  $J$  parts reproduce the exact equations of motion as shown below with certain additional

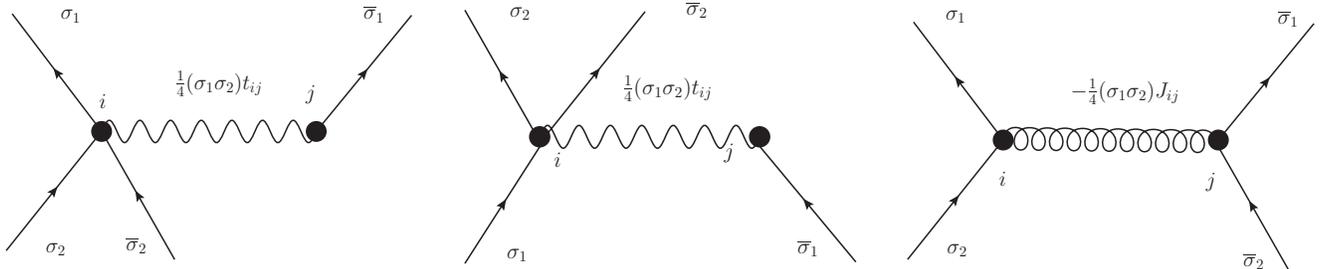


FIG. 1. The pseudopotential  $V_{\text{eff}}$  in the real-space representation, where the wavy line represents  $t_{ij}$  and the coiled line represents  $J_{ij}$ . The first two interaction vertices have two undisplayed symmetric partners with the exchange  $i \leftrightarrow j$ .

terms that vanish under the constraint of single occupancy. The parameter  $\lambda$  is set to unity at the end and provides an interpolation to the Fermi gas. The parameter  $u_0$  represent an effective Hubbard-type interaction for these fermions, giving a contribution  $u_0 \sum_i f_{i\uparrow}^\dagger f_{i\uparrow} f_{i\downarrow}^\dagger f_{i\downarrow}$ . Its magnitude is arbitrary at the moment, since it disappears under exclusion of double occupancy. Here it enables us to enforce the invariance in shift theorem I, where the shift of  $t$  and  $J$  can be absorbed in  $u_0$ . It will turn out to play the role of a second chemical potential or Lagrange multiplier, in fixing the second sum rule Eq. (91). To illustrate this remark, note that adding a constant to  $t$  or  $J$  as in Eq. (14) adds an on-site four-Fermi interaction term. In order to satisfy the shift theorem I, we must compensate for this suitably, leading to the extra on-site term with coefficient  $u_0$ , which can absorb this shift. It is also verified that the shift theorem II is satisfied without the  $u_0$  term. We emphasize that the  $u_0$  term is both natural and essential for the purpose of satisfying the shift theorem I. Since the structure of the  $u_0$  term is almost identical to that of  $J_{ij}$  we will most often “hide it” inside  $J_{ij}$ , and explicitly display it at the end. Thus unless explicitly displayed, we should read  $J_{ij} \rightarrow J_{ij} - u_0 \delta_{ij}$  below. For analogous terms involving the  $X_i^{\sigma \sigma'}$  operators as in Eq. (22), we can include  $u_0$  in  $J_{ij}$  without any errors, since the  $u_0$  term always vanishes due to the properties of these operators.

Defining symmetric Cooper pair singlet operators

$$\begin{aligned} \mathcal{P}^\dagger(i, j) &= \sum \sigma f_{i\sigma}^\dagger f_{j\bar{\sigma}}^\dagger = (f_{i\uparrow}^\dagger f_{j\downarrow}^\dagger - f_{i\downarrow}^\dagger f_{j\uparrow}^\dagger), \\ \mathcal{P}^\dagger(i, i) &= \sum \sigma f_{i\sigma}^\dagger f_{i\bar{\sigma}}^\dagger = 2 f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger, \end{aligned} \quad (28)$$

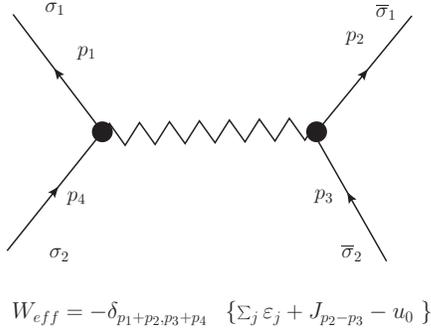


FIG. 2. The pseudopotential  $W_{\text{eff}}$  in the momentum-space representation. The zigzag line represents  $W_{\text{eff}}$ . Note that the momentum transfer in the argument of  $J$  is also expressible as  $J_{p_1-p_4}$ .

with  $\mathcal{P}^\dagger(j, i) = \mathcal{P}^\dagger(i, j)$ , we write

$$V_{\text{eff}} = \frac{1}{4} \sum_{ij} t_{ij} [(\mathcal{P}^\dagger(i, i) + \mathcal{P}^\dagger(j, j))\mathcal{P}(i, j) + (\text{H.c.})] - \frac{1}{4} \sum_{ij} J_{ij} \mathcal{P}_{ij}^\dagger \mathcal{P}_{ij}. \quad (29)$$

In momentum representation the effective Hamiltonian Eq. (27) reads

$$H_{\text{eff}} = \sum_k \left( \varepsilon_k + \frac{1}{4} J_0 - \mu \right) f_{k\sigma}^\dagger f_{k\sigma} + \frac{\lambda}{4N_s} \sum_p (\sigma_1 \sigma_2) W_{\text{eff}}(p_1, p_2; p_3, p_4) \times f_{p_1 \sigma_1}^\dagger f_{p_2 \bar{\sigma}_1}^\dagger f_{p_3 \bar{\sigma}_2} f_{p_4 \sigma_2},$$

$$W_{\text{eff}}(p_1, p_2; p_3, p_4) = -\delta_{p_1+p_2, p_3+p_4} \{ \varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + J_{p_2-p_3} - u_0 \} \quad (30)$$

(see Fig. 2), where the momentum-independent term  $u_0$  has been explicitly written out. In this effective Hamiltonian, the band energies  $\varepsilon_{p_j}$  of the original model are present, both in the band energy of the  $f$ 's and the interaction term. Therefore the shift Eq. (13) cannot be absorbed in the  $\mu$  alone, and  $u_0$  must also transform suitably to ensure that the effective Hamiltonian satisfies the shift theorem I. Thus in using the effective Hamiltonian we refine this theorem to

*Shift theorem I.1.* An arbitrary shift Eq. (14) of  $t$  and  $J$  can be absorbed by shifting the chemical potential  $\mu \rightarrow \mu + u_t + \frac{1}{4}u_J$  and  $u_0$  as

$$u_0 \rightarrow u_0 + 4u_t + u_J. \quad (31)$$

Note that the shift theorem II is manifestly satisfied: The combination of the band energies  $\varepsilon_{p_j}$  and the exchange term  $J_p$  in Eq. (30) guarantees that their shift adds up to  $u_J + 4u_t \rightarrow 0$ , which vanishes under the conditions of this theorem.

Since the standard notation for interaction reads  $\sum \langle ab|V|a'b' \rangle f_a^\dagger f_b^\dagger f_{b'} f_{a'}$  for a conventional two-body interaction, our notation corresponds to writing  $W_{\text{eff}}(p_1, p_2; p_3, p_4) = \langle p_1 p_2 | W | p_4 p_3 \rangle$ . Fermi symmetry implies the invariance  $W_{\text{eff}}(p_1, p_2; p_3, p_4) = W_{\text{eff}}(p_2, p_1; p_4, p_3)$ , and Hermiticity im-

plies the invariance  $W_{\text{eff}}(p_1, p_2; p_3, p_4) = W_{\text{eff}}(p_3, p_4; p_1, p_2)$ . For this canonical theory, we calculate the commutator:

$$[f_{i\sigma_i}, H_{\text{eff}}] = - \sum_j t_{ij} f_j \sigma_j + \left( \frac{1}{4} J_0 - \mu \right) f_{i\sigma_i} + \hat{A}_{i\sigma_i},$$

$$\hat{A}_{i\sigma_i} = [f_{i\sigma_i}, V_{\text{eff}}], \quad (32)$$

with

$$\hat{A}_{i\sigma_i} = \sum_{j\sigma_j} t_{ij} (\sigma_i \sigma_j) \times \left[ f_{i\bar{\sigma}_i}^\dagger f_{i\bar{\sigma}_j} f_{j\sigma_j} + \frac{1}{2} f_{j\bar{\sigma}_i}^\dagger f_{j\bar{\sigma}_j} f_{j\sigma_j} + \frac{1}{2} f_{j\bar{\sigma}_i}^\dagger f_{i\bar{\sigma}_j} f_{i\sigma_j} \right] - \frac{1}{2} \sum_{j \neq i} J_{ij} (\sigma_i \sigma_j) f_{j\bar{\sigma}_i}^\dagger f_{j\bar{\sigma}_j} f_{i\sigma_j}. \quad (33)$$

Let us note that  $\hat{A}_{i\sigma}$  Eq. (33) differs from  $A_{i\sigma}$  in Eq. (22), through terms (in underbraces) that vanish identically if we impose the single-occupancy constraint on the auxiliary electrons.

#### D. Equation of motion for $\mathcal{G}$ continued

We now return to the study of the equation of motion for  $\mathcal{G}$  in Eq. (26), expressed in terms of  $A_{i\sigma}$  of Eq. (22), the commutator of the destruction operator with  $H$ . This object yields the crucial Heisenberg equation of motion; therefore as discussed in Eq. (6), we next look for terms that can be added to it to make it identical to Eq. (33). Comparing Eq. (22) and Eq. (33) we see that these differ by terms [the second and third terms of the square bracket in Eq. (33)] that are automatically vanishing for the  $X_i^{ab}$  operators on using their standard rules. Thus we can add such vanishing terms to Eq. (22) that remain exact and also importantly preserve the Hermitian nature of the auxiliary fermionic theory in approximate schemes. We thus rewrite also an exact but more useful result:

$$A_{i,\sigma_i} = \sum_{ij\sigma_j} t_{ij} (\sigma_i \sigma_j) \times \left[ X_i^{\bar{\sigma}_i \bar{\sigma}_j} X_j^{0\sigma_j} + \frac{1}{2} X_j^{\bar{\sigma}_i \bar{\sigma}_j} X_j^{0\sigma_j} + \frac{1}{2} X_j^{\bar{\sigma}_i 0} X_i^{0\bar{\sigma}_j} X_i^{0\sigma_j} \right] - \frac{1}{2} \sum_{j \neq i} J_{ij} (\sigma_i \sigma_j) X_j^{\bar{\sigma}_i \bar{\sigma}_j} X_i^{0\sigma_j}, \quad (34)$$

so that  $A_{i,\sigma_i}$  and  $\hat{A}_{i,\sigma_i}$  contain terms that are in one to one correspondence. We will use Eq. (34) in place of Eq. (22) in Eq. (26) next.

The notation simplifies if we use the matrix notation for the spin indices introduced in Ref. 11 and Ref. 2, e.g.,  $\mathcal{G}_{\sigma_i \sigma_f}[i, f] \rightarrow [\mathcal{G}[i, f]]_{\sigma_i \sigma_f}$ , so that we may regard  $\mathcal{G}$  as a  $2 \times 2$  matrix. In short, the space-time indices are displayed but the spin indices are hidden in the above matrix structure. We next define  $\gamma$  through

$$\gamma_{\sigma_a \sigma_b}[i] = \sigma_a \sigma_b \mathcal{G}_{\bar{\sigma}_b \bar{\sigma}_a}[i^-, i], \quad \text{or} \quad \gamma[i] = \mathcal{G}^{(k)}[i^-, i], \quad (35)$$

where we denote the  $k$  conjugation of any matrix  $M$  by  $(M^{(k)})_{\sigma_1 \sigma_2} = M_{\bar{\sigma}_2 \bar{\sigma}_1} \sigma_1 \sigma_2$ . This conjugation corresponds to time reversal in the spin space. Let  $\mathbb{1}$  be the identity matrix in the  $2 \times 2$  dimensional spin space.

We employ a useful relation with an arbitrary operator  $\mathcal{Q}$  that follows from Eq. (16). We write

$$\begin{aligned} \langle\langle \sigma_a \sigma_b X_i^{\bar{\sigma}_a \bar{\sigma}_b}(\tau_i) \mathcal{Q} \rangle\rangle &= (\gamma_{\sigma_a \sigma_b}[i] - D_{\sigma_a \sigma_b}[i]) \langle\langle \mathcal{Q} \rangle\rangle, \\ \langle\langle \sigma_a \sigma_b X_i^{\bar{\sigma}_a 0}(\tau_i^+) X_j^{0 \bar{\sigma}_b}(\tau_j) \mathcal{Q} \rangle\rangle &= (\gamma_{\sigma_a \sigma_b}[i, j] - D_{\sigma_a \sigma_b}[i, j]) \langle\langle \mathcal{Q} \rangle\rangle, \end{aligned} \quad (36)$$

where we set  $\tau_j \rightarrow \tau_j^-$  and define

$$\gamma_{\sigma_a \sigma_b}[i, j] = (\sigma_a \sigma_b) \mathcal{G}_{\bar{\sigma}_b \bar{\sigma}_a}[j \tau_j^-, i \tau_i] = \langle\langle \sigma_a \sigma_b X_i^{\bar{\sigma}_a 0} X_j^{0 \bar{\sigma}_b} \rangle\rangle, \quad (37)$$

$$\gamma[i, i] = \gamma[i], \quad (38)$$

and

$$\begin{aligned} D_{\sigma_i \sigma_j}[i] &= \sigma_i \sigma_j \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i)}, \\ D_{\sigma_i \sigma_j}[i, j] &= \sigma_i \sigma_j \frac{\delta}{\delta \mathcal{V}_{i, j}^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i)}, \end{aligned} \quad (39)$$

and  $D[i, i] = D[i]$ .

In  $\gamma[i, i]$  and  $\gamma[i]$  we have equal-time objects with creation operators to the left of destruction operators. Let us note the rewriting of the last term in Eq. (26):

$$\begin{aligned} - \sum_j \mathcal{V}_{ij}^{\sigma_1 \sigma_2}(\sigma_i \sigma_1) \langle\langle X_i^{\bar{\sigma}_1 \bar{\sigma}_1}(\tau_i) X_j^{0 \sigma_2}(\tau_j) X_f^{\sigma_f 0}(\tau_f) \rangle\rangle \\ = + \sum_j \mathcal{V}_{ij}^{\sigma_1 \sigma_2} (\gamma_{\sigma_i \sigma_1}[i] - D_{\sigma_i \sigma_1}[i]) \mathcal{G}_{\sigma_2 \sigma_f}[j, f]. \end{aligned} \quad (40)$$

With this preparation, using Eq. (34) we rewrite Eq. (26) as

$$\begin{aligned} (\partial_{\tau_i} - \boldsymbol{\mu} + \frac{1}{4} J_0) \mathcal{G}_{\sigma_i \sigma_f}[i, f] \\ = -\delta[i, f] (\delta_{\sigma_i \sigma_f} - \gamma_{\sigma_i \sigma_f}[i]) - \mathcal{V}_i^{\sigma_i \sigma_i} \\ \times \mathcal{G}_{\sigma_3 \sigma_f}[i, f] - \mathcal{V}_{i, \mathbf{j}}^{\sigma_i \sigma_1}(\tau_i) \mathcal{G}_{\sigma_3 \sigma_f}[\mathbf{j}, f] \\ + \mathcal{V}_{ij}^{\sigma_1 \sigma_2} [\gamma_{\sigma_i \sigma_1}(i) - D_{\sigma_i \sigma_1}(i)] \mathcal{G}_{\sigma_2 \sigma_f}(\mathbf{j}, f) + t[i, \mathbf{j}] \\ \times \left\{ (\mathbb{1} - \gamma[i] + D[i] - \frac{1}{2} \gamma[\mathbf{j}] + \frac{1}{2} D[\mathbf{j}]) \cdot \mathcal{G}[\mathbf{j}, f] \right\}_{\sigma_i \sigma_f} \\ + t[i, \mathbf{j}] \left\{ (-\frac{1}{2} \gamma[\mathbf{j}, i] + \frac{1}{2} D[\mathbf{j}, i]) \cdot \mathcal{G}[i, f] \right\}_{\sigma_i \sigma_f} \\ + \frac{1}{2} J[i, \mathbf{j}] \{ (\gamma[\mathbf{j}] - D[\mathbf{j}]) \cdot \mathcal{G}[i, f] \}_{\sigma_i \sigma_f}, \end{aligned} \quad (41)$$

where the fixed variables are in normal letters and the repeated variables in bold letters are summed in space and integrated in time. This may be written compactly in matrix form as

$$\begin{aligned} (\partial_{\tau_i} - \boldsymbol{\mu}) \mathcal{G}[i, f] \\ = -\delta[i, f] (\mathbb{1} - \gamma[i]) - \mathcal{V}_i \cdot \mathcal{G}[i, f] \\ - \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G}[\mathbf{j}, f] + (\gamma(i) - D[i]) \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G}[\mathbf{j}, f] \\ - X[i, \mathbf{j}] \cdot \mathcal{G}[\mathbf{j}, f] - Y[i, \mathbf{j}] \cdot \mathcal{G}[\mathbf{j}, f], \end{aligned} \quad (42)$$

where we used the definitions (with fixed  $j$  and summed  $\mathbf{k}$ )

$$\begin{aligned} X[i, j] &= -t[i, j] (D[i] + \frac{1}{2} D[j]) \\ &\quad + \delta[i, j] \frac{1}{2} (J[i, \mathbf{k}] D[\mathbf{k}] - t[i, \mathbf{k}] D[\mathbf{k}, i]), \\ Y[i, j] &= -t[i, j] (\mathbb{1} - \gamma[i] - \frac{1}{2} \gamma[j]) + \frac{1}{4} J_0 \mathbb{1} \\ &\quad - \delta[i, j] \frac{1}{2} (J[i, \mathbf{k}] \gamma[\mathbf{k}] - t[i, \mathbf{k}] \gamma[\mathbf{k}, i]). \end{aligned} \quad (43)$$

These exact equations Eq. (42) and Eq. (43) form the basis for the remaining discussion. The coefficients in  $X$  and  $Y$

differ slightly from the ones in Ref. 2, in view of the usage of the effective Hermitean Hamiltonian idea in this paper. The extra terms arise from the form of Eq. (33), and actually vanish if we could treat either of these exactly. We will show that this formulation leads to approximations obeying the shift theorems I and II discussed earlier; note however that Eq. (42) and the forms of  $X, Y$  in Eq. (43) are manifestly invariant under these theorems.

### III. DECOMPOSITION OF $\mathcal{G}$ INTO THE AUXILIARY FERMION GREEN'S FUNCTION $\mathbf{g}$ AND THE CAPARISON FACTOR $\boldsymbol{\mu}$

As discussed in the introduction, we next write the *product ansatz* for  $\mathcal{G}$

$$\mathcal{G}[a, b] = \mathbf{g}[a, \mathbf{r}] \cdot \boldsymbol{\mu}[\mathbf{r}, b], \quad (44)$$

where  $\mathbf{g}$  is the canonical auxiliary Green's function and  $\boldsymbol{\mu}$  is the caparison factor, or the adaptive spectral weight. Since  $\mathcal{G}$  satisfies antiperiodic boundary conditions under  $\tau_a \rightarrow \tau_a + \beta$  and  $\tau_b \rightarrow \tau_b + \beta$  separately, we must Fourier-transform both factors  $\mathbf{g}$  and  $\boldsymbol{\mu}$  with fermionic frequencies  $\omega_n = (2n + 1)\pi k_B T$ . At this point  $\boldsymbol{\mu}$  and  $\mathbf{g}$  are undetermined. Let us first note in matrix notation the equal-time objects:

$$\begin{aligned} \gamma[i] &= \mathcal{G}[i^-, i] \rightarrow (\mathbf{g}[i, \mathbf{a}] \cdot \boldsymbol{\mu}[\mathbf{a}, i])^{(k)} \\ &= (\boldsymbol{\mu}[\mathbf{a}, i])^{(k)} \cdot (\mathbf{g}[i, \mathbf{a}])^{(k)}, \\ \gamma[i, j] &= \mathcal{G}[j^-, i] \rightarrow (\mathbf{g}[j, \mathbf{a}] \cdot \boldsymbol{\mu}[\mathbf{a}, i])^{(k)} \\ &= (\boldsymbol{\mu}[\mathbf{a}, i])^{(k)} \cdot (\mathbf{g}[j, \mathbf{a}])^{(k)}. \end{aligned} \quad (45)$$

We define three-point vertex functions

$$\begin{aligned} \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p, q; r) &\equiv -\frac{\delta}{\delta \mathcal{V}_r^{\sigma_3 \sigma_4}(\tau_r)} \{ \mathbf{g}_{\sigma_1 \sigma_2}^{-1}[p, q] \}, \\ \mathcal{U}_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}[a, b; c] &\equiv \frac{\delta \mu_{\sigma_1 \sigma_2}[a, b]}{\delta \mathcal{V}_c^{\sigma_3 \sigma_4}(\tau_c)}, \end{aligned} \quad (46)$$

or as an implicit matrix in the upper indices (but explicit in the lower ones):

$$\begin{aligned} \Lambda_{\sigma_3 \sigma_4}(p, q; r) &= -\frac{\delta}{\delta \mathcal{V}_r^{\sigma_3 \sigma_4}} \{ \mathbf{g}^{-1}[p, q] \}, \\ \mathcal{U}_{\sigma_3 \sigma_4}[a, b; c] &\equiv \frac{\delta \mu[a, b]}{\delta \mathcal{V}_c^{\sigma_3 \sigma_4}}. \end{aligned} \quad (47)$$

In a similar vein, to obtain the four-point vertex functions corresponding to the source  $\mathcal{V}_{r, s}$  with a pair of points  $r, s$  with  $\tau_r = \tau_s$ , we define

$$\begin{aligned} \Lambda_*(p, q; r, s) &= -\frac{\delta}{\delta \mathcal{V}_{r, s}^*(\tau_r)} \{ \mathbf{g}^{-1}[p, q] \}, \\ \mathcal{U}_*[a, b; c, d] &\equiv \frac{\delta \mu[a, b]}{\delta \mathcal{V}_{c, d}^*(\tau_c)}. \end{aligned} \quad (48)$$

In some expressions involving summations, it is convenient to think of the vertices  $\Lambda_*(p, q; r, s), \mathcal{U}_*[p, q; r, s]$  with independent times  $\tau_r, \tau_s$ , with the constraint of equal times imposed by multiplying by a delta function  $\delta(\tau_r - \tau_s)$ , as illustrated in Eq. (B4).

This set of vertices  $\Lambda$  and  $\mathcal{U}$  replaces the single vertex  $\Gamma$  of a canonical many-body system, and we will also find equations determining these below. Clearly in any exact treatment, the

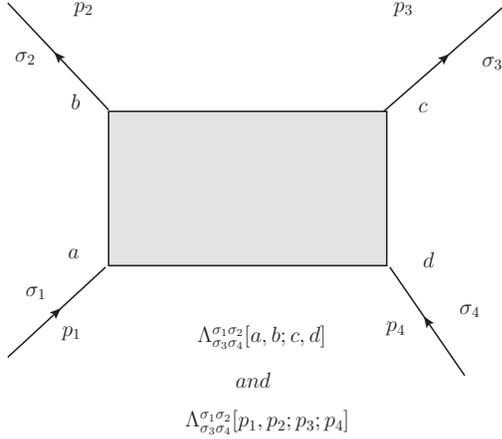


FIG. 3. The four-site vertex can be visualized from its definition for canonical theory:  $-\langle\langle f_{a\sigma_1} f_{b\sigma_2}^\dagger f_{c\sigma_3} f_{d\sigma_4} \rangle\rangle + \langle\langle f_{a\sigma_1} f_{b\sigma_2}^\dagger \rangle\rangle \langle\langle f_{c\sigma_3} f_{d\sigma_4} \rangle\rangle = \frac{\delta}{\delta \mathcal{V}_{cd}^{*3}} \mathbf{g}_{\sigma_1 \sigma_2}(ab) = \mathbf{g}_{\sigma_1 \sigma_1'}[a, \mathbf{a}] \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2'}[\mathbf{a}, \mathbf{b}; c, d] \mathbf{g}_{\sigma_2' \sigma_2}[\mathbf{b}, \mathbf{b}]$ . Therefore we may visualize that apart from the external legs,  $\Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}[a, b; c, d] \sim \langle\langle f_{a\sigma_1} f_{b\sigma_2}^\dagger f_{c\sigma_3} f_{d\sigma_4} \rangle\rangle$ . Note that in this convention, the labels differ by a cyclic permutation from those in Fig. 2.

four-point vertex contains the three-point vertex by collapsing the points:

$$\Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p, q; r) = \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p, q; r, s \rightarrow r), \quad (49)$$

and similarly for  $\mathcal{U}$ . However in any approximation scheme, this identity would follow only if the single-occupancy constraint at a given site  $i$ , namely  $\langle\langle X_i^{\sigma_1 \sigma_2} X_i^{0 \sigma_3} \dots \rangle\rangle = 0$ , is satisfied exactly, for all choices of the spin indices. Since typical approximations relax this constraint, if only slightly, it is useful to keep both sets of vertices in the theory as separate entities. Another attractive possibility is to require the identity Eq. (49), by making a different set of (controlled) approximations, and is also discussed below. Figure 3 illustrates the conventions used for the four-point vertex; the three-point vertex is obtained by the indicated contraction.

We now use a notation where  $*$  is used as a place holder, as illustrated in component form by

$$\dots \xi_{\sigma_a \sigma_b}^* \dots \frac{\delta}{\delta \mathcal{V}_j^*} = \dots \sigma_a \sigma_b \dots \frac{\delta}{\delta \mathcal{V}_j^{\sigma_a, \sigma_b}}, \quad (50)$$

with  $\xi_{\sigma_a \sigma_b} = \sigma_a \sigma_b$ , and an implicit spin flip in the indices of the attached derivative operator  $\delta / \delta \mathcal{V}_j^{\sigma_a, \sigma_b}$ .

We would like to rewrite Eq. (42) in terms of the vertex functions. We need to express

$$\begin{aligned} X[i, \mathbf{j}] \cdot \mathcal{G}[\mathbf{j}, f] &= -t[i, \mathbf{j}] (D[i] + \frac{1}{2} D[\mathbf{j}]) \cdot \mathcal{G}[\mathbf{j}, f] \\ &+ \frac{1}{2} (J[i, \mathbf{k}] D[\mathbf{k}] - t[i, \mathbf{k}] D[\mathbf{k}, i]) \cdot \mathcal{G}[i, f] \end{aligned} \quad (51)$$

in terms of the vertex functions. Differentiating Eq. (44) we find

$$\begin{aligned} \frac{\delta}{\delta \mathcal{V}_r^{\sigma, \sigma'}} \mathcal{G}[a, b] &= \mathbf{g}[a, \mathbf{c}] \cdot \Lambda_{\sigma, \sigma'}(\mathbf{c}, \mathbf{d}; r) \cdot \mathcal{G}[\mathbf{d}, b] \\ &+ \mathbf{g}[a, \mathbf{c}] \cdot \mathcal{U}_{\sigma, \sigma'}[\mathbf{c}, b; r]. \end{aligned} \quad (52)$$

Consulting Eq. (39) for the definition of  $D_{\sigma_1 \sigma_2}[i] = \xi_{\sigma_1 \sigma_2} \frac{\delta}{\delta \mathcal{V}_i^{\sigma_1 \sigma_2}}$ , where  $\xi_{\sigma_1 \sigma_2} = \sigma_1 \sigma_2$ , we rewrite this as

$$\begin{aligned} D[r] \cdot \mathcal{G}[a, b] &= \xi^* \cdot \mathbf{g}[a, \mathbf{c}] \cdot \Lambda_*(\mathbf{c}, \mathbf{d}; r) \cdot \mathcal{G}[\mathbf{d}, b] \\ &+ \xi^* \cdot \mathbf{g}[a, \mathbf{c}] \cdot \mathcal{U}_*[\mathbf{c}, b; r], \end{aligned} \quad (53)$$

where the spin flip in the derivatives is implied as stressed above.

Combining Eq. (51) and Eq. (53) we define the useful linear operator

$$\begin{aligned} \mathbf{L}[i, j] &= t[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[\mathbf{k}, j] \cdot \left( \frac{\delta}{\delta \mathcal{V}_i^*} + \frac{1}{2} \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}}^*} \right) \\ &+ \frac{1}{2} t[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, j] \cdot \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}, i}^*} \\ &- \frac{1}{2} J[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, j] \cdot \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}}^*}. \end{aligned} \quad (54)$$

Hence we may write Eq. (51) compactly as

$$X[i, j] \cdot \mathcal{G}[\mathbf{j}, f] \equiv \Phi[i, \mathbf{b}] \cdot \mathcal{G}[\mathbf{b}, f] + \Psi[i, f], \quad (55)$$

where the two central objects of this theory arise from the action of a common operator Eq. (54) on two seed objects  $\mathbf{g}^{-1}$  and  $\mu$  as follows:

$$\begin{aligned} \Phi[i, m] &\equiv \mathbf{L}[i, \mathbf{c}] \cdot \mathbf{g}^{-1}[\mathbf{c}, m] \\ &= -t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{c}] \cdot (\Lambda_*[\mathbf{c}, m; i] + \frac{1}{2} \Lambda_*[\mathbf{c}, m; \mathbf{j}]) \\ &- \frac{1}{2} t[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \Lambda_*[\mathbf{c}, m; \mathbf{k}, i] \\ &+ \frac{1}{2} J[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \Lambda_*[\mathbf{c}, m; \mathbf{k}] \end{aligned} \quad (56)$$

and

$$\begin{aligned} \Psi[i, m] &\equiv -\mathbf{L}[i, \mathbf{c}] \cdot \mu[\mathbf{c}, m] \\ &= -t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{c}] \cdot (\mathcal{U}_*[\mathbf{c}, m; i] + \frac{1}{2} \mathcal{U}_*[\mathbf{c}, m; \mathbf{j}]) \\ &- \frac{1}{2} t[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \mathcal{U}_*[\mathbf{c}, m; \mathbf{k}, i] \\ &+ \frac{1}{2} J[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \mathcal{U}_*[\mathbf{c}, m; \mathbf{k}]. \end{aligned} \quad (57)$$

We write Eq. (43) as

$$\begin{aligned} Y[i, j] &= -t[i, j] + Y_1[i, j], \\ Y_1[i, j] &= t[i, j] (\gamma[i] + \frac{1}{2} \gamma[j]) \\ &- \delta[i, j] \frac{1}{2} (J[i, \mathbf{k}] \gamma[\mathbf{k}] - t[i, \mathbf{k}] \gamma[\mathbf{k}, i]). \end{aligned} \quad (58)$$

We also need to process the object:

$$\begin{aligned} &(\gamma(i) - D[i]) \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G}[\mathbf{j}, f] \\ &= \gamma(i) \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G}[\mathbf{j}, f] - \xi^* \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \frac{\delta}{\delta \mathcal{V}_i^*} \mathcal{G}[\mathbf{j}, f] \\ &= \gamma(i) \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G}[\mathbf{j}, f] - \xi^* \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathbf{g}[\mathbf{j}, \mathbf{c}] \cdot \Lambda_*[\mathbf{c}, \mathbf{r}; i] \\ &\cdot \mathcal{G}[\mathbf{r}, f] - \xi^* \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathbf{g}[\mathbf{j}, \mathbf{c}] \cdot \mathcal{U}_*[\mathbf{c}, f; i]. \end{aligned} \quad (59)$$

#### IV. ASSEMBLING THE EQUATIONS

Let us rewrite the three relevant equations symbolically:

(1) Eq. (42) for  $\mathcal{G}$ :

$$\begin{aligned} (\partial_{\tau_i} - \mu) \mathcal{G} &= -\delta(\mathbb{1} - \gamma) - \mathcal{V}_i \cdot \mathcal{G} - \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G} \\ &+ (\gamma - D_i) \cdot \mathcal{V}_{i, \mathbf{j}} \cdot \mathcal{G} - X \cdot \mathcal{G} - Y \cdot \mathcal{G}. \end{aligned} \quad (60)$$

(2) Eq. (59) for the two site source  $\mathcal{V}_{ij}$ :

$$(\gamma - D_i) \cdot \mathcal{V}_{i,j} \cdot \mathcal{G} = \gamma(i) \cdot \mathcal{V}_{i,j} \cdot \mathcal{G}[\mathbf{j}, f] - \xi^* \cdot \mathcal{V}_{i,j} \cdot \mathbf{g} \cdot \Lambda_* \cdot \mathcal{G} - \xi^* \cdot \mathcal{V}_{i,j} \cdot \mathbf{g} \cdot \mathcal{U}_*. \quad (61)$$

(3) Eq. (55) the product rule:

$$X \cdot \mathcal{G} = \Phi \cdot \mathcal{G} + \Psi. \quad (62)$$

Combining these we rewrite Eq. (60) symbolically as

$$[\partial_{\tau_i} - \boldsymbol{\mu} + Y + \mathcal{V}_i + (\mathbb{1} - \gamma) \cdot \mathcal{V}_{i,j} + \xi^* \cdot \mathcal{V}_{i,j} \cdot \mathbf{g} \cdot \Lambda_* + \Phi] \mathcal{G} = -\delta(\mathbb{1} - \gamma) - \Psi - \xi^* \cdot \mathcal{V}_{i,j} \cdot \mathbf{g} \cdot \mathcal{U}_*. \quad (63)$$

Defining

$$\mathbf{g}_0^{-1}[i, f] = \left\{ \left[ (\boldsymbol{\mu} - \partial_{\tau_i} - \frac{1}{4} J_0) \mathbb{1} - \mathcal{V}_i \right] \delta[i, f] + t[i, f] - \mathcal{V}_{i,f}(\tau_i) \delta(\tau_i - \tau_f) \right\}, \quad (64)$$

the exact EOM Eq. (42) can be written in matrix form:

$$\begin{aligned} & \left\{ \mathbf{g}_0^{-1}[i, \mathbf{j}] + \gamma_i \cdot \mathcal{V}_{i,j} - \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \Lambda_*(\mathbf{b}, \mathbf{j}; i) - Y_1[i, \mathbf{j}] \right. \\ & \quad \left. - \Phi[i, \mathbf{j}] \cdot \mathbf{g}[\mathbf{j}, f] \cdot \mu[f, f] \right\} \\ & = \delta[i, f] (\mathbb{1} - \gamma[i]) + \Psi[i, f] \\ & \quad + \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \mathcal{U}_*(\mathbf{b}, f; i). \end{aligned} \quad (65)$$

At this point, a convenient parameter  $\lambda$  (finally set  $\lambda \rightarrow 1$ ) is now inserted into this equation as follows:

$$\begin{aligned} & \left\{ \mathbf{g}_0^{-1}[i, \mathbf{j}] + \underbrace{\lambda \gamma_i \cdot \mathcal{V}_{i,j} - \lambda \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \Lambda_*(\mathbf{b}, \mathbf{j}; i)}_{\text{}} \right. \\ & \quad \left. - \lambda Y_1[i, \mathbf{j}] - \lambda \Phi[i, \mathbf{j}] \cdot \mathbf{g}[\mathbf{j}, f] \cdot \mu[f, f] \right\} \\ & = \delta[i, f] (\mathbb{1} - \lambda \gamma[i]) + \lambda \Psi[i, f] \\ & \quad + \underbrace{\lambda \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \mathcal{U}_*(\mathbf{b}, f; i)}_{\text{}}. \end{aligned} \quad (66)$$

Clearly this becomes the exact equation Eq. (65) at  $\lambda = 1$ , and reduces to the Fermi gas Green's function Eq. (64) at  $\lambda = 0$ . We may now split Eq. (66) exactly into a pair of equations that are fundamental to the theory:

$$\begin{aligned} & \left\{ \mathbf{g}_0^{-1}[i, \mathbf{j}] + \lambda \gamma_i \cdot \mathcal{V}_{i,j} - \lambda \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \Lambda_*(\mathbf{b}, \mathbf{j}; i) \right. \\ & \quad \left. - \lambda Y_1[i, \mathbf{j}] - \lambda \Phi[i, \mathbf{j}] \cdot \mathbf{g}[\mathbf{j}, f] \right\} \cdot \mathbf{g}[\mathbf{j}, f] = \delta[i, f], \end{aligned} \quad (67)$$

$$\begin{aligned} & \mu[i, f] = \delta[i, f] (\mathbb{1} - \lambda \gamma[i]) + \lambda \Psi[i, f] \\ & \quad + \lambda \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \mathcal{U}_*(\mathbf{b}, f; i). \end{aligned} \quad (68)$$

We can usefully invert Eq. (67) and write

$$\begin{aligned} \mathbf{g}^{-1}[i, m] & = \left\{ \mathbf{g}_0^{-1}[i, m] + \lambda \gamma_i \cdot \mathcal{V}_{i,m} - \lambda \xi^* \cdot \mathcal{V}_{i,a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \right. \\ & \quad \left. \cdot \Lambda_*(\mathbf{b}, m; i) - \lambda Y_1[i, m] - \lambda \Phi[i, m] \right\}. \end{aligned} \quad (69)$$

We see that  $\mathbf{g}$  satisfies a canonical equation, with a delta function of weight unity on the right, and  $\mu$  soaks up the remaining factors on the right-hand side of Eq. (66). This decomposition is not unique; one has the obvious freedom of respectively post-multiplying  $\mathbf{g}$  and pre-multiplying  $\mu$  by a common function and its inverse. However, requiring  $\mathbf{g}$  to be canonical fixes the function to be unity. The motivation of introducing  $\lambda$  in the above equations is to establish adiabatic, or more properly parametric, continuity with the Fermi gas.<sup>16</sup> At this stage some remarks are necessary.

(1) At  $\lambda = 1$  Eq. (67) and Eq. (68) become exact equations for the EC phase, while it has the virtue that as  $\lambda = 0$  it gives a canonical equation for  $\mathbf{g}$ , with  $\mu[i, j] = \mathbb{1} \delta[i, j]$ . Procedurally, we can calculate objects to a given order in  $\lambda$  iteratively, and set  $\lambda = 1$  at the end of the calculation. We thus establish and maintain continuity with the Fermi gas in the equations of motion.

(2) The process of introducing  $\lambda$  into the EOM is not unique. For example the terms of Eq. (66) in the underbraces cancel at  $i = \mathbf{j}$  from the vanishing of Eq. (40). However this cancellation is exact only at  $\lambda = 1$ , so we will find below that an expansion in  $\lambda$  has the annoying feature of a slight violation of the contraction of indices result Eq. (49). We will show below that this is inconsequential to the orders in  $\lambda$  considered here. With hindsight, a better strategy would be to impose the constraint Eq. (49) to the order of the calculation. This can be achieved if we multiply the terms in underbraces by a sufficiently high power of  $\lambda^r$ , say with  $r \geq r_0$ , and thereby avoid dealing with this problem at low orders  $r < r_0$ . Below we will analyze the minimal choice  $r = 1$ , record the issues that crop up, and make suitable approximations later. The impatient may simply ignore the terms with underbraces.

(3) Another type of freedom is available at this stage: If necessary, we could add an arbitrary term that varies smoothly with  $\lambda$  and vanishes at both end points, e.g.,  $\propto \lambda(1 - \lambda)$ , to either side of Eq. (67) and Eq. (68). It will turn out that the first-order term  $[\mathbf{g}^{-1}]_1$  calculated below does need a simple term of this type to fulfill the Fermi surface sum rule. In general, however, the natural and minimal choice made in Eq. (66), without such a term, seems adequate for higher terms.

(4) We note that the shift theorems I and II are preserved by  $X, Y$  above in Eq. (43), and this invariance survives the introduction of  $\lambda$  in Eq. (66). As a result the various objects  $\Phi, \Psi, \mathbf{g}^{-1}, \mu$  satisfy these theorems individually. This property leads to a powerful consistency check on the approximations to each order in  $\lambda$ .

(5) Note that a  $\lambda$  expansion of  $\gamma[i]$  implies that the high-frequency fall-off of  $\mathcal{G} \sim \frac{c_0}{i\omega}$  now occurs with a coefficient  $c_0 = 1 - \lambda\gamma$  that is different from  $1 - \frac{n}{2}$  at finite orders of  $\lambda$ . While it is tempting to freeze this coefficient at the exact value, it would be inconsistent since we take its derivatives to find  $\Psi$ , etc. The departure of this coefficient from the exact value becomes increasingly significant near  $n \sim 1$ , and provides a criterion for the validity of a given order of approximation.

## V. EXPLICIT EQUATIONS AND THE ZERO-SOURCE LIMIT IN FOURIER SPACE

When we turn off the sources, the various matrix function  $\mathcal{G}, \mathbf{g}, \mu$  become spin diagonal. We will also take Fourier transforms (only) in this limit, since translation invariance in space and time is regained when the sources vanish.

We next express  $\Phi$  and  $\Psi$  explicitly in terms of the vertex functions. We need to take the Fourier transform of Eq. (56) and Eq. (57). In the ECFL theory, a rotationally invariant liquid phase is obtained by turning off the sources. We can use the standard spin rotational symmetry analysis illustrated here with  $\Lambda$  as in Ref. 11. We define the three nonvanishing matrix elements as  $\Lambda^{(1)} = \Lambda_{\sigma\sigma}^{\sigma\sigma}$ ,  $\Lambda^{(2)} = \Lambda_{\bar{\sigma}\bar{\sigma}}^{\sigma\sigma}$ , and  $\Lambda^{(3)} = \Lambda_{\sigma\bar{\sigma}}^{\sigma\bar{\sigma}}$ . We also record the Nozières identity for the

two expressions of a particle-hole singlet,  $\Lambda^{(1)} - \Lambda^{(2)} = \Lambda^{(3)}$ , which provides an important check on the theory. We further use a notation for the frequently occurring antisymmetric combination  $\Lambda^{(a)} = \Lambda^{(2)} - \Lambda^{(3)}$ . Armed with these, we next drop the matrix structure by utilizing an identity arising with a fixed  $\sigma$  (such as in the expression for  $\Phi_{\sigma\sigma}$  above):

$$\begin{aligned} \langle \sigma | \xi^* \cdot \mathbf{g} \cdot \Lambda_* | \sigma \rangle &= \sum_{\sigma_a \sigma_b} \sigma \sigma_a \mathbf{g}_{\sigma_a \sigma_b} \Lambda_{\bar{\sigma} \bar{\sigma}_a}^{\sigma_b \sigma} \\ &= \sum_{\sigma_b} (\mathbf{g}_{\sigma \sigma_b} \Lambda_{\bar{\sigma} \bar{\sigma}}^{\sigma_b \sigma} - \mathbf{g}_{\bar{\sigma} \sigma_b} \Lambda_{\bar{\sigma} \bar{\sigma}}^{\sigma_b \sigma}) \\ &= (\mathbf{g}_{\sigma \sigma} \Lambda_{\bar{\sigma} \bar{\sigma}}^{\sigma \sigma} - \mathbf{g}_{\bar{\sigma} \bar{\sigma}} \Lambda_{\bar{\sigma} \bar{\sigma}}^{\bar{\sigma} \bar{\sigma}}) \\ &= \mathbf{g}(\Lambda^{(2)} - \Lambda^{(3)}) \equiv \mathbf{g}\Lambda^{(a)}. \end{aligned} \quad (70)$$

Note that we dropped the spin index on  $\mathbf{g}$  due to the isotropy of the state.

We use the FT convention for the two-, three-, and four-site objects illustrated with the examples

$$\begin{aligned} \mathcal{G}[a, b] &= \sum_k e^{ik(a-b)} \mathcal{G}(k), \\ \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}[a, b; c] &= \sum_{p_1, p_2} e^{ip_1(a-c) + ip_2(c-b)} \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p_1, p_2), \\ \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}[a, b; c, d] &= \sum_{p_1 + p_4 = p_2 + p_3} e^{i(p_1 a - p_2 b - p_3 c + p_4 d)} \\ &\quad \times \Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p_1, p_2; p_3, p_4). \end{aligned} \quad (71)$$

The identity Eq. (49) in momentum space implies

$$\begin{aligned} \Lambda(p_1, p_2) &= \sum_{p_3, p_4} \Lambda(p_1, p_2; p_3, p_4), \\ \mathcal{U}(p_1, p_2) &= \sum_{p_3, p_4} \mathcal{U}(p_1, p_2; p_3, p_4). \end{aligned} \quad (72)$$

At zero source we get the exact relations between self-energies and vertices by Fourier-transforming Eq. (56) and Eq. (57):

$$\begin{aligned} \Phi(k) &= \sum_p \left( \varepsilon_p + \frac{1}{2} \varepsilon_k + \frac{1}{2} J_{k-p} \right) \mathbf{g}[p] \Lambda^{(a)}(p, k) \\ &\quad + \sum_{pq} \frac{1}{2} \varepsilon_{q+p-k} \mathbf{g}[p] \Lambda^{(a)}(p, k; q + p - k, q) \end{aligned}$$

TABLE I. Summary of defining equations. The computation of the Green's function  $\mathcal{G} = \mathbf{g} \cdot \mu$  in Eq. (44) requires several intermediate variables. The complete set of variables in this theory (first column) and their mutual and  $\lambda$  dependence (second column) are collected here for convenience. The corresponding equation number in the paper is given in the last column.

Object	Defining Equation	Eq. No.
$\mathbf{g}^{-1}[i, m]$	$\{\mathbf{g}_0^{-1}[i, m] + \lambda \gamma_i \cdot \mathcal{V}_{i, m} - \lambda \xi^* \cdot \mathcal{V}_{i, a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \Lambda_*(\mathbf{b}, m; i) - \lambda Y_1[i, m] - \lambda \Phi[i, m]\}$	Eq. (69)
$\mu[i, m]$	$\delta[i, m](\mathbb{1} - \lambda \gamma[i]) + \lambda \Psi[i, m] + \lambda \xi^* \cdot \mathcal{V}_{i, a} \cdot \mathbf{g}[\mathbf{a}, \mathbf{b}] \cdot \mathcal{U}_*(\mathbf{b}, m; i)$	Eq. (68)
$Y_1[i, m]$	$t[i, m](\gamma[i] + \frac{1}{2} \gamma[m]) - \delta[i, m] \frac{1}{2} (J[i, \mathbf{k}] \gamma[\mathbf{k}] - t[i, \mathbf{k}] \gamma[\mathbf{k}, i])$	Eq. (58)
$\gamma[i]$	$\mu^{(k)}[\mathbf{a}, i] \cdot \mathbf{g}^{(k)}[i, \mathbf{a}]$	Eq. (45)
$\gamma[i, m]$	$\mu^{(k)}[\mathbf{a}, i] \cdot \mathbf{g}^{(k)}[m, \mathbf{a}]$	Eq. (45)
$\Phi[i, m]$	$-t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{c}] \cdot (\Lambda_*[\mathbf{c}, m; i] + \frac{1}{2} \Lambda_*[\mathbf{c}, m; \mathbf{j}]) - \frac{1}{2} t[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \Lambda_*[\mathbf{c}, m; \mathbf{k}, i] + \frac{1}{2} J[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \Lambda_*[\mathbf{c}, m; \mathbf{k}]$	Eq. (56)
$\Psi[i, m]$	$-t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{c}] \cdot (\mathcal{U}_*[\mathbf{c}, m; i] + \frac{1}{2} \mathcal{U}_*[\mathbf{c}, m; \mathbf{j}]) - \frac{1}{2} t[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \mathcal{U}_*[\mathbf{c}, m; \mathbf{k}, i] + \frac{1}{2} J[i, \mathbf{k}] \xi^* \cdot \mathbf{g}[i, \mathbf{c}] \cdot \mathcal{U}_*[\mathbf{c}, m; \mathbf{k}]$	Eq. (57)

$$\begin{aligned} &= \sum_{pq} \left( \varepsilon_p + \frac{1}{2} \varepsilon_k + \frac{1}{2} \varepsilon_{q+p-k} + \frac{1}{2} J_{k-p} \right) \mathbf{g}[p] \\ &\quad \times \Lambda^{(a)}(p, k; q + p - k, q), \\ \Psi(k) &= \sum_p \left( \varepsilon_p + \frac{1}{2} \varepsilon_k + \frac{1}{2} J_{k-p} \right) \mathbf{g}[p] \mathcal{U}^{(a)}(p, k) \\ &\quad + \sum_{pq} \frac{1}{2} \varepsilon_{q+p-k} \mathbf{g}[p] \mathcal{U}^{(a)}(p, k; q + p - k, q) \\ &= \sum_{pq} \left( \varepsilon_p + \frac{1}{2} \varepsilon_k + \frac{1}{2} \varepsilon_{q+p-k} + \frac{1}{2} J_{k-p} \right) \\ &\quad \times \mathbf{g}[p] \mathcal{U}^{(a)}(p, k; q + p - k, q). \end{aligned} \quad (73)$$

A convergence factor  $e^{i\omega_p 0^+}$  arises from the time ordering and is implied wherever necessary and the last line in both equations is valid provided the identity Eq. (72) is satisfied. Here  $\Lambda^{(a)} = \Lambda^{(2)} - \Lambda^{(3)}$  and  $\mathcal{U}^{(a)} = \mathcal{U}^{(2)} - \mathcal{U}^{(3)}$ .

With  $k = (\vec{k}, i\omega_k)$  and  $\omega_n = \pi(2n+1)k_B T$ , the Green's functions at a fixed  $\lambda$  read

$$\begin{aligned} \mathcal{G}(k) &= \mathbf{g}(k) \times \mu(k), \\ \mathbf{g}^{-1}[k] &= i\omega_n + \boldsymbol{\mu} - \varepsilon_k - \frac{1}{4} J_0 - \lambda Y_1(k) - \lambda \Phi(k), \\ \mu[k] &= 1 - \lambda \gamma + \lambda \Psi(k). \end{aligned} \quad (74)$$

The sum rule for the number of physical particles and the auxiliary fermions is given by

$$\sum_p \mu[p] \mathbf{g}[p] = \frac{n}{2}, \quad (75)$$

$$\sum_p \mathbf{g}[p] = \frac{n}{2}. \quad (76)$$

While the sum rule Eq. (75) clearly counts the number of physical electrons, the origin of the sum rule Eq. (76) for  $\mathbf{g}$  requires some discussion taken from Ref. 2. We recall that it is meant to enforce the Luttinger-Ward theorem of a conserved Fermi volume for the auxiliary fermions. By so doing and through the composition  $\mathcal{G} = \mathbf{g} \times \mu$ , it also preserves it for the physical fermions. While  $\boldsymbol{\mu}$  provides us with one obvious Lagrange multiplier to enforce one of the sum rules, the more subtle parameter  $u_0$ , introduced in Eq. (27), is required to enforce the second sum rule Eq. (91). Explicit expressions for  $\gamma, Y_1, \Phi, \Psi$  can be calculated at various orders in  $\lambda$  as demonstrated below.

TABLE II. Vertex functions. The theory requires three-point and four-point vertices. Their nomenclature (first column) and definition (second column) are given, along with the corresponding equation number in the paper.

Vertex	Defining Equation	Eq. No.
$\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]$	$-\frac{\delta}{\delta\mathcal{V}_j^{\sigma_c\sigma_d}}\mathbf{g}_{\sigma_a\sigma_b}^{-1}[i,m]$	Eq. (46)
$\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]$	$-\frac{\delta}{\delta\mathcal{V}_{j,k}^{\sigma_c\sigma_d}}\mathbf{g}_{\sigma_a\sigma_b}^{-1}[i,m]$	Eq. (48)
$\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]$	$\frac{\delta}{\delta\mathcal{V}_j^{\sigma_c\sigma_d}}\mu_{\sigma_a\sigma_b}[i,m]$	Eq. (46)
$\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]$	$\frac{\delta}{\delta\mathcal{V}_{j,k}^{\sigma_c\sigma_d}}\mu_{\sigma_a\sigma_b}[i,m]$	Eq. (48)

## VI. SUMMARIZING

Before proceeding to the iterative scheme, we collect all the relevant equations for convenience in Table I. The various vertex functions are found from relationships summarized in Table II.

It is worthwhile to provide one nontrivial example of the matrix notation. In component form note that  $\Phi[i,m]$  can be written out as

$$\begin{aligned}\Phi_{\sigma_i\sigma_m}[i,m] &= -t[i,\mathbf{j}]\sigma_i\sigma_1\mathbf{g}_{\sigma_1\sigma_2}[\mathbf{j},\mathbf{c}] \\ &\quad \times \left( \Lambda_{\tilde{\sigma}_i\tilde{\sigma}_1}^{\sigma_2\sigma_m}[\mathbf{c},m;i] + \frac{1}{2}\Lambda_{\tilde{\sigma}_i\tilde{\sigma}_1}^{\sigma_2\sigma_m}[\mathbf{c},m;\mathbf{j}] \right) \\ &\quad - \frac{1}{2}t[i,\mathbf{k}]\sigma_i\sigma_1\mathbf{g}_{\sigma_1\sigma_2}[i,\mathbf{c}]\Lambda_{\tilde{\sigma}_i\tilde{\sigma}_1}^{\sigma_2\sigma_m}[\mathbf{c},m;\mathbf{k},i] \\ &\quad + \frac{1}{2}J[i,\mathbf{k}]\sigma_i\sigma_1\mathbf{g}_{\sigma_1\sigma_2}[i,\mathbf{c}]\Lambda_{\tilde{\sigma}_i\tilde{\sigma}_1}^{\sigma_2\sigma_m}[\mathbf{c},m;\mathbf{k}].\end{aligned}$$

## VII. $\lambda$ EXPANSION AND THE ITERATIVE SCHEME

Taking functional derivatives with respect to  $\mathcal{V}$ , we generate a self-energy–vertex hierarchy of fermionic theory, paralleling the standard (i.e., canonical) theory, but with greater complexity due to the two kinds of vertex functions and self-energies. We describe the  $\lambda$  expansion and the iterative process next. The iterations are analogous to the skeleton diagram expansion in standard many-body theory, where  $\lambda$  plays the role of the interaction constant. Various objects are expanded in terms of  $\lambda$  and  $\mathbf{g}$ , while  $\mathbf{g}$  itself is left intact. Potentially confusing is the treatment of  $\mathbf{g}^{-1}$ , which is expanded in  $\lambda$  and  $\mathbf{g}$ , ignoring its obvious relationship as the inverse of  $\mathbf{g}$ . This becomes understandable when we recall that  $\mathbf{g}^{-1}$  is, apart from  $\mathbf{g}_0^{-1}$ , the Dyson self-energy of the auxiliary system, and is to be regarded as a functional of  $\mathbf{g}$ , as in the Luttinger-Ward functional Ref. 17. One example of this expansion may be

TABLE III. Iteration level  $p$  calculations. The auxiliary inverse Green's function  $\mathbf{g}^{-1}$  and the adaptive spectral weight  $\mu$  play the role of seed objects at the  $p$ th order. By computing them to  $p$ th order in the parameter  $\lambda$ , we obtain the vertex functions and the other variables listed in the second column to the same order as described in Eqs. (80)–(83).

Seed Object	Derived Objects
$[\mu[i,m]]_p$	$[\gamma[i],\gamma[i,m],Y_1[i,m],\mathcal{U}[a,b;c],\mathcal{U}[a,b;c,d]]_p$
$[\mathcal{U}[a,b;c],\mathcal{U}[a,b;c,d]]_p$	$[\Psi[i,m]]_p$
$[\mathbf{g}^{-1}[i,m]]_p$	$[\Lambda[a,b;c],\Lambda[a,b;c,d]]_p$
$[\Lambda[a,b;c],\Lambda[a,b;c,d]]_p$	$[\Phi[i,m]]_p$

TABLE IV. Iteration level step-up calculations. In proceeding upwards in the iterative process in Eq. (84) the computed  $(p+1)$ th-order objects are listed in the first column, and the  $p$ th-order objects needed are in the second column. Since  $\mathbf{g}^{-1}$  and  $\mu$  at a given level suffice to determine all other objects at that level through Table III, the iterative nature of the scheme becomes transparent.

Level $(p+1)$ Object	Required Level $p$ Objects
$[\mu[i,m]]_{p+1}$	$[\gamma[i],\Psi[i,m],\mathcal{U}[a,b;c],\mathcal{U}[a,b;c,d]]_p$
$[\mathbf{g}^{-1}[i,m]]_{p+1}$	$[Y_1[i,m],\Phi[i,m],\Lambda[a,b;c]]_p$

useful. Consider  $\gamma[i,m]$ ; we will expand it as

$$\begin{aligned}\gamma^{(k)}[i,m] &= \mathbf{g}[m,\mathbf{a}] \cdot \mu[\mathbf{a},i] = \mathbf{g}[m,\mathbf{a}] \cdot ([\mu[\mathbf{a},i]]_0 \\ &\quad + \lambda([\mu[\mathbf{a},i]]_1 + \lambda^2[\mu[\mathbf{a},i]]_2 + O(\lambda^3)),\end{aligned}\quad (77)$$

keeping  $\mathbf{g}$  intact, i.e., unexpanded in  $\lambda$ . A similar expansion is carried out also for  $\gamma[i]$ , leading to a correction of the high-frequency fall of coefficient  $c_0$  as noted above.

*Iterative process.* We now describe the various steps of the iteration process. First note that all variables (except  $\mathbf{g}$ ) are expanded as

$$\mathcal{A} = [\mathcal{A}]_0 + \lambda[\mathcal{A}]_1 + \lambda^2[\mathcal{A}]_2 + \cdots + \lambda^p[\mathcal{A}]_p + \cdots. \quad (78)$$

The iteration scheme can be summarized in the two following tables. Table III lists the seed objects needed at any order and gives the derived objects. Table IV lists the higher order objects and the needed lower level objects for stepping up.

(I) *Initialization at  $p=0$ .* The iterations require the following starting relations:

$$\begin{aligned}\mathbf{g}_0^{-1}[i,m] &= \left\{ [(\boldsymbol{\mu} - \partial_{t_i} - \frac{1}{4}J_0)\mathbb{1} - \mathcal{V}_i] \right. \\ &\quad \left. \times \delta[i,m] + t[i,m] - \mathcal{V}_{i,m} \right\}, \\ [\mu[i,f]]_0 &= \mathbb{1}\delta[i,f].\end{aligned}\quad (79)$$

(II) *Computation of derived objects at level  $p$  from Table I.* The set of equations requiring  $[\mu[i,m]]_p$ :

$$\begin{aligned}[\gamma[i]]_p &= [\mu^{(k)}[\mathbf{a},i]]_p \cdot \mathbf{g}^{(k)}[i,\mathbf{a}], \\ [\gamma[i,m]]_p &= [\mu^{(k)}[\mathbf{a},i]]_p \cdot \mathbf{g}^{(k)}[m,\mathbf{a}], \\ [\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]]_p &= \left( \frac{\delta}{\delta\mathcal{V}_j^{\sigma_c\sigma_d}} \right) [\mu_{\sigma_a\sigma_b}[i,m]]_p, \\ [\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]]_p &= \left( \frac{\delta}{\delta\mathcal{V}_{j,k}^{\sigma_c\sigma_d}} \right) [\mu_{\sigma_a\sigma_b}[i,m]]_p,\end{aligned}\quad (80)$$

$$[Y_1[i,m]]_p = t[i,m] \left[ \gamma[i] + \frac{1}{2}\gamma[m] \right] - \delta[i,m] \frac{1}{2} [J[i,\mathbf{k}]\gamma[\mathbf{k}] - t[i,\mathbf{k}]\gamma[\mathbf{k},i]]_p, \quad (81)$$

$$[\Psi[i,m]]_p = -t[i,\mathbf{j}]\xi^* \cdot \mathbf{g}[\mathbf{j},\mathbf{c}] \cdot \left( \mathcal{U}_*[\mathbf{c},m;i] + \frac{1}{2}\mathcal{U}_*[\mathbf{c},m;\mathbf{j}] \right)_p - \frac{1}{2}t[i,\mathbf{k}]\xi^* \cdot \mathbf{g}[i,\mathbf{c}] \cdot (\mathcal{U}_*[\mathbf{c},m;\mathbf{k},i])_p + \frac{1}{2}J[i,\mathbf{k}]\xi^* \cdot \mathbf{g}[i,\mathbf{c}] \cdot (\mathcal{U}_*[\mathbf{c},m;\mathbf{k}])_p. \quad (82)$$

The set of equations requiring  $[\mathbf{g}^{-1}[i,m]]_p$ :

$$[\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]]_p = - \left( \frac{\delta}{\delta\mathcal{V}_j^{\sigma_c\sigma_d}} \right) [\mathbf{g}_{\sigma_a\sigma_b}^{-1}[i,m]]_p, \\ [\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]]_p = - \left( \frac{\delta}{\delta\mathcal{V}_{j,k}^{\sigma_c\sigma_d}} \right) [\mathbf{g}_{\sigma_a\sigma_b}^{-1}[i,m]]_p, \\ [\Phi[i,m]]_p = -t[i,\mathbf{j}]\xi^* \cdot \mathbf{g}[\mathbf{j},\mathbf{c}] \cdot \left( \Lambda_*[\mathbf{c},m;i] + \frac{1}{2}\Lambda_*[\mathbf{c},m;\mathbf{j}] \right)_p - \frac{1}{2}t[i,\mathbf{k}]\xi^* \cdot \mathbf{g}[i,\mathbf{c}] \cdot (\Lambda_*[\mathbf{c},m;\mathbf{k},i])_p + \frac{1}{2}J[i,\mathbf{k}]\xi^* \cdot \mathbf{g}[i,\mathbf{c}] \cdot (\Lambda_*[\mathbf{c},m;\mathbf{k}])_p. \quad (83)$$

(III) *Level  $p$  to level  $(p+1)$ , step up equations:*

$$[\mu[i,m]]_{p+1} = -\delta[i,m][\gamma[i]]_p + [\Psi[i,m]]_p + [\xi^* \cdot \mathcal{V}_{i,\mathbf{a}} \cdot \mathbf{g}(\mathbf{a},\mathbf{b}) \cdot \mathcal{U}_*(\mathbf{b},m;i)]_p, \\ [\mathbf{g}^{-1}[i,m]]_{p+1} = [\gamma_i \cdot \mathcal{V}_{i,m} - \xi^* \cdot \mathcal{V}_{i,\mathbf{a}} \cdot \mathbf{g}(\mathbf{a},\mathbf{b}) \cdot \Lambda_*(\mathbf{b},m;i)]_p - [Y_1[i,m] + \Phi[i,m]]_p. \quad (84)$$

(IV) *If required level is reached, exit; otherwise return to step II.*

This iterative procedure can thus be applied to obtain equations for the Green's functions to any desired order. In practice the higher order terms grow very rapidly, as in the Feynman diagram series. However, as explained in the introduction, a low-order expansion is expected to capture already the significant features of extreme correlations, an important reason being that the range is finite and small; i.e.,  $\lambda \in [0,1]$ . In this work we will be content to work to  $O(\lambda^2)$  where all the relevant objects can be calculated explicitly.

*Second-order Green's function.* Having formulated the iterative process, we next apply this to obtain the second-order Green's functions. The calculations are detailed in Appendix B, and we directly present the first- and second-order results here. Displaying the so far hidden  $u_0$  coefficient, we write the complete set of equations to  $O(\lambda^2)$  from Eq. (B51)

and Eq. (B48):

$$\mathcal{G}[k] = \mathbf{g}(k) \times \mu[k], \\ \mu[k] = 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4} - \lambda^2 \sum_{p,q} (\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p} - u_0) \mathbf{g}[p] \mathbf{g}[q] \mathbf{g}[q+k-p] + O(\lambda^3), \quad (85)$$

$$\mathbf{g}^{-1}[k] = i\omega_n + \mu' - \left( 1 - \lambda n + \lambda^2 \frac{3n^2}{8} \right) \varepsilon_k + \lambda \sum_q \frac{1}{2} J_{k-q} \mathbf{g}[q] - \lambda^2 [\Phi(k)]_1 + O(\lambda^3), \quad (86)$$

$$[\Phi(k)]_1 = - \sum_{q,p} \mathbf{g}[q] \mathbf{g}[p] \mathbf{g}[k+q-p] \times (\varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + J_{k-p} - u_0) \times \left\{ \varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + \frac{1}{2}(J_{k-p} + J_{p-q}) - u_0 \right\}. \quad (87)$$

The shifted chemical potential  $\mu'$  is related to the physical (i.e., thermodynamical) chemical potential  $\mu$  and  $u_0$  through

$$\mu' = \mu - u_0 \frac{\lambda n}{2} \left( 1 - \frac{\lambda n}{4} \right) + \left[ J_0 \frac{\lambda n}{4} \left( 1 - \frac{\lambda n}{2} \right) + 2\lambda \left( 1 - \frac{\lambda n}{8} \right) \sum_q \varepsilon_q \mathbf{g}[q] \right]. \quad (88)$$

In using this expansion, one must first set  $\lambda \rightarrow 1$ . These expressions satisfy the shift theorem I.1 and shift theorem II, as one can verify by shifting  $\varepsilon_k$  and  $J_k$  by  $k$ -independent constants, and using  $\sum_q \mathbf{g}[q] = \frac{n}{2}$ . The self-energy from a Feynman diagram theory to second-order from  $H_{\text{eff}}$  in Eq. (27) matches the above expression for  $\mathbf{g}^{-1}$ . The required diagrams are shown in Fig. 4 up to second order where the zigzag line  $W_{\text{eff}}$  is defined in Fig. 2.

Apart from a single term (the expansion of  $Y_1$  in  $\lambda$ ), the expansion of the auxiliary Fermi liquid is largely "autonomous,"

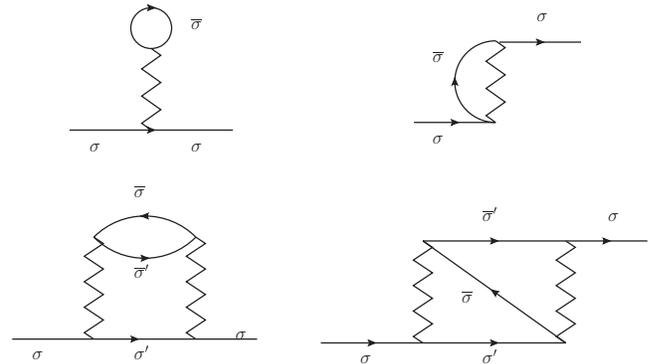


FIG. 4. The self-energy graphs to second order from  $W_{\text{eff}}$  and the effective Hamiltonian  $H_{\text{eff}}$ . These determine the  $\Phi$  self-energy.

i.e., proceeds without requiring the knowledge of  $\mu$ , and is represented in Feynman diagrammatic terms. The caparison term  $\mu$  has no obvious interpretation in terms of  $H_{\text{eff}}$ , but is easy to compute along lines similar to the ones shown here, and the full theory splices the two factors to yield  $\mathcal{G}$ , as described here.

A consistent first-order, i.e.,  $O(\lambda)$ , theory for  $\mathbf{g}^{-1}$  and  $\mu$  can be found after dropping all  $O(\lambda^2)$  terms. As it stands, we would get  $\mu = 1 - \lambda \frac{n}{2}$  to this order, and this would violate the Fermi surface volume theorem (Ref. 12). To recover from this, we may however set  $\mu[k]$  to unity instead. Formally this is achieved by adding  $\lambda(1 - \lambda) \frac{n}{2}$  to  $\mu[k]$  as discussed below Eq. (66), since this added term vanishes at both end points  $\lambda = 0$  and  $\lambda = 1$ . This procedure is within the permissible adjustments of the continuity argument, and at second order cancels out so that the quoted second-order result is unchanged. Further all vertices are unchanged since this is a static term. In this way the first-order theory can also be arranged to satisfy the Luttinger-Ward Fermi volume theorem. This theory has a band dispersion  $(1 - n)\varepsilon_k$  that shrinks in width by a factor  $(1 - n)$  as in the Gutzwiller-Brinkman-Rice theory,<sup>7,18</sup> with an enhanced effective mass  $m/m^* = (1 - n)$ . The second-order result presented here provides a more interesting and frequency-dependent correction to the Fermi gas.

In summary, the physical Green's function is obtained from

$$\mathcal{G}[k] = \mathbf{g}[k]\mu[k]. \quad (89)$$

The number of the physical electrons is fixed by the first sum rule,

$$\frac{n}{2} = \sum_k \mathcal{G}[k]e^{i\omega_n 0^+}, \quad (90)$$

while the auxiliary fermion satisfies an identical sum rule,

$$\frac{n}{2} = \sum_k \mathbf{g}[k]e^{i\omega_n 0^+}. \quad (91)$$

We can determine the two independent real parameters  $\boldsymbol{\mu}$  and  $u_0$  in order to satisfy both these equations simultaneously, and thus the role of  $u_0$  as a Lagrange multiplier, similar to that of  $\boldsymbol{\mu}$ , is now evident. It is also clear that the shifts of  $t$  or  $J$  can be absorbed in the two Lagrange multipliers  $\boldsymbol{\mu}$  and  $u_0$ . It is worth noting that the simplified ECFL model used in Ref. 2 and Ref. 9 can be obtained from Eq. (85) and Eq. (87) by throwing out the band energies and exchange energies in the coefficients of  $\mathbf{g}(q)\mathbf{g}(p)\mathbf{g}(k + q - p)$  while retaining  $u_0$ , so that the Lagrange multiplier of that approximation  $\Delta_0$  is related to  $u_0$ .

The role of the two sum rules in fixing the number of fermions and also the Luttinger-Ward Fermi surface is already discussed in Ref. 2 and above. We can add to that discussion with the help of the explicit functional forms found above. It should be noted from Eq. (85) and Eq. (87) that the functional derivatives

$$I[k, p] \equiv \frac{\delta [\Phi[k]]_1}{\delta \mathbf{g}[p]}, \quad J[k, p] \equiv \frac{\delta [\Psi[k]]_1}{\delta \mathbf{g}[p]} \quad (92)$$

are symmetric functions under  $k \leftrightarrow p$ . This symmetry therefore guarantees the existence of two Luttinger-Ward-type

functionals of the auxiliary Green's function  $\mathbf{g}$ ,

$$\begin{aligned} \Omega_\Phi[\mathbf{g}] &= -\frac{1}{4} \sum_{k,p,q,r} W(k,q;r,p)[W(k,q;r,p) \\ &\quad + W(k,q;p,r)]\mathbf{g}[k]\mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[r], \\ \Omega_\Psi[\mathbf{g}] &= \frac{1}{4} \sum_{k,p,q,r} W(k,q;r,p)\mathbf{g}[k]\mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[r], \end{aligned} \quad (93)$$

such that the two self-energies can be found from these functionals:

$$[\Phi[k]]_1 = \frac{\delta \Omega_\Phi}{\delta \mathbf{g}[k]}, \quad [\Psi[k]]_1 = \frac{\delta \Omega_\Psi}{\delta \mathbf{g}[k]}. \quad (94)$$

The form of these two functionals follows to this order from Eq. (87), and it is natural to conjecture that such functionals exist to all orders in  $\lambda$ . The existence of the  $\Omega_\Phi$  functional guarantees a (FS) volume-conserving Luttinger-Ward Fermi surface for the  $\mathbf{g}$  electrons, and the smooth behavior of  $\Psi(k)$  near this surface guarantees likewise for the physical electrons.

### VIII. WARD IDENTITIES

This theory admits Ward identities involving the vertices  $\Lambda$  and  $\mathcal{U}$  that guarantee current conservation in a similar fashion as Ref. 11. This is displayed with the help of sources, the charge potential  $u[m] = \sum_\sigma \mathcal{V}_m^{\sigma\sigma}$  and an added source  $v[m]$  coupling to the kinetic energy as

$$t[i, j] \rightarrow t[i, j](1 + v[j] - v[i]), \quad (95)$$

so that  $v[j] - v[i]$  acts as a discrete version of the Peierls phase factor of electromagnetic coupling in tight binding systems. We define

$$D_m \equiv \partial_{\tau_m} \frac{\delta}{\delta u[m]} - \frac{\delta}{\delta v[m]}, \quad (96)$$

so that the Ward identity expressing the conservation of current, from Ref. 11, reads

$$D_m \mathcal{G}[i, f] = (\delta[i, m] - \delta[f, m])\mathcal{G}[i, f]. \quad (97)$$

This is a discrete (Takahashi type) version of the usual Ward identity appropriate to the lattice Fermi system at hand, and electromagnetic coupling only requires the long-wavelength limit of this identity. We will define the ( $\mathcal{T}$ ) vertices (summing over  $\sigma$ )

$$\begin{aligned} \Lambda^{(\mathcal{T})}(i, j; m) &= -\frac{\delta}{\delta v[m]} \mathbf{g}_{\sigma\sigma}^{-1}[i, j]|_{u,v \rightarrow 0}, \\ \mathcal{U}^{(\mathcal{T})}(i, j; m) &= \frac{\delta}{\delta v[m]} \mu_{\sigma\sigma}[i, j]|_{u,v \rightarrow 0}. \end{aligned} \quad (98)$$

It is easy to see that the bare  $\tau$  vertices are given by differentiating  $\mathbf{g}_0^{-1}$  in Eq. (64) as

$$\begin{aligned} \lambda^{(\mathcal{T})}(i, j; m) &= t[i, j](\delta[i, m] - \delta[j, m]), \\ \lambda^{(\mathcal{T})}[p_1, p_2] &= \varepsilon_{p_1} - \varepsilon_{p_2}, \end{aligned} \quad (99)$$

while the singlet (i.e., density) vertices are already known from  $\Lambda^{(s)} = \sum_{\sigma\sigma'} \Lambda_{\sigma\sigma'}^{\sigma\sigma}$ . Note that the ( $\mathcal{T}$ ) type vertices are antisymmetric in  $i \leftrightarrow j$  or  $p_1 \leftrightarrow p_2$ .

Taking Fourier transforms in Eq. (97) and writing  $\mathcal{G} = \mathbf{g} \times \mu$ , we get the conservation law:

$$\begin{aligned} & (i\omega_{p_1} - i\omega_{p_2})(\mathbf{g}[p_1]\Lambda^{(s)}(p_1, p_2)\mathbf{g}[p_2]\mu[p_2] \\ & + \mathbf{g}[p_1]\mathcal{U}^{(s)}(p_1, p_2)) - (\mathbf{g}[p_1]\Lambda^{(T)}(p_1, p_2)\mathbf{g}[p_2]\mu[p_2] \\ & + \mathbf{g}[p_1]\mathcal{U}^{(T)}(p_1, p_2)) = \mathbf{g}[p_2]\mu[p_2] - \mathbf{g}[p_1]\mu[p_1]. \end{aligned} \quad (100)$$

Canceling out  $\mathbf{g}[p_1]\mathbf{g}[p_2]$  we get the Ward identity:

$$\begin{aligned} & (i\omega_{p_1} - i\omega_{p_2})(\Lambda^{(s)}(p_1, p_2)\mu[p_2] + \mathcal{U}^{(s)}(p_1, p_2)\mathbf{g}^{-1}[p_2]) \\ & - (\Lambda^{(T)}(p_1, p_2)\mu[p_2] + \mathcal{U}^{(T)}(p_1, p_2)\mathbf{g}^{-1}[p_2]) \\ & = \mathbf{g}^{-1}[p_1]\mu[p_2] - \mathbf{g}^{-1}[p_2]\mu[p_1]. \end{aligned} \quad (101)$$

With  $i\omega_n \rightarrow z_n$ , we rewrite this as

$$\mathcal{W}_g(p_1, p_2)\mu[p_2] + \mathbf{g}^{-1}[p_2]\mathcal{W}_\mu(p_1, p_2) = 0, \quad (102)$$

where we have defined the two Ward functions:

$$\begin{aligned} \mathcal{W}_g(p_1, p_2) &= (z_1 - z_2)\Lambda^{(s)}(p_1, p_2) - \Lambda^{(T)}(p_1, p_2) \\ & + \mathbf{g}^{-1}[p_2] - \mathbf{g}^{-1}[p_1], \\ \mathcal{W}_\mu(p_1, p_2) &= (z_1 - z_2)\mathcal{U}^{(s)}(p_1, p_2) - \mathcal{U}^{(T)}(p_1, p_2) \\ & + \mu[p_1] - \mu[p_2]. \end{aligned} \quad (103)$$

Since  $p_1$  and  $p_2$  are arbitrary, the two terms must vanish separately giving us the pair of Ward identities:

$$\mathcal{W}_g(p_1, p_2) = 0, \quad (104)$$

---


$$\Lambda^{(T)}[p_1, p_2] = (\varepsilon_{p_1} - \varepsilon_{p_2})(1 - \lambda n) - \lambda \sum_q \mathbf{g}[q]\Lambda^{(T)}(q, q + p_2 - p_1)\mathbf{g}[q + p_2 - p_1]\mathcal{F}(q, p_1, p_2) + O(\lambda^2),$$

$$\Lambda^{(s)}[p_1, p_2] = 1 - \lambda \sum_q \mathbf{g}[q]\Lambda^{(s)}(q, q + p_2 - p_1)\mathbf{g}[q + p_2 - p_1]\mathcal{F}(q, p_1, p_2) + O(\lambda^2), \quad (107)$$

$$\Lambda^{(t)}[p_1, p_2] = 1 + \lambda \sum_q \mathbf{g}[q]\Lambda^{(t)}(q, q + p_2 - p_1)\mathbf{g}[q + p_2 - p_1]\mathcal{F}(q, p_1, p_2) + O(\lambda^2),$$


---

where we use the shorthand  $\mathcal{F}(q, p_1, p_2) \equiv \{\varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_q + \varepsilon_{q+p_2-p_1} - u_0 + \frac{1}{2}(J_{p_1-p_2} + J_{q-p_1})\}$ . The term in underbrace receives an  $O(\lambda)$  contribution from differentiating the explicit  $v$  dependence of the transformed  $t[i, m] \rightarrow t[i, m](1 + v[m] - v[i])$  term in Eq. (B7). It is readily shown by examining the kernel of the integral equations that the solution for  $\Lambda^{(T)}(p_1, p_2)$  is antisymmetric under exchanging  $p_1 \leftrightarrow p_2$ , while  $\Lambda^{(s)}(p_1, p_2)$  and  $\Lambda^{(t)}(p_1, p_2)$  are symmetric.

These vertices are shown to be compatible with Ward identities to  $O(\lambda)$  if used with the first-order versions of the Green's functions Eq. (85) and Eq. (86),

$$\begin{aligned} \mathbf{g}[p] &= i\omega_n + \boldsymbol{\mu}' - (1 - \lambda n)\varepsilon_k \\ & + \frac{\lambda}{2} \sum_q J_{k-q}\mathbf{g}[q] + O(\lambda^2) \end{aligned} \quad (108)$$

$$\text{and } \mu[p] = 1,$$

by substituting in the expressions Eq. (104) and Eq. (105), and showing the self-consistency of this result. The details of this verification parallel the standard proof in QED and are omitted here. Note that  $\mu$  must be chosen to be unity rather than  $1 - \lambda \frac{n}{2}$

$$\mathcal{W}_\mu(p_1, p_2) = 0. \quad (105)$$

## IX. RANDOM PHASE APPROXIMATION

Since the Green's functions are known to  $O(\lambda^2)$ , we can take the derivatives of Eq. (B7) and Eq. (B8) to get vertices to this order. Here we calculate by taking the equations to  $O(\lambda)$  only, but assuming  $\frac{\delta}{\delta V}\mathbf{g} = \mathbf{g}\Lambda\mathbf{g}$  rather than  $\frac{\delta}{\delta V}\mathbf{g} = \mathbf{g}\mathbf{g}$ , thereby obtaining the analog of the RPA. Since the spin susceptibility is also of considerable interest, we will calculate the required vertices in the that channel as well. Summarizing the results we write linear integral equations for the  $\mathcal{U}$  vertices:

$$\begin{aligned} \mathcal{U}^{(T)}[p_1, p_2] &= -\lambda \sum_q \mathbf{g}[q]\Lambda^{(T)}(q, q + p_2 - p_1) \\ & \times \mathbf{g}[q + p_2 - p_1] + O(\lambda^2), \\ \mathcal{U}^{(s)}[p_1, p_2] &= -\lambda \sum_q \mathbf{g}[q]\Lambda^{(s)}(q, q + p_2 - p_1) \\ & \times \mathbf{g}[q + p_2 - p_1] + O(\lambda^2), \\ \mathcal{U}^{(t)}[p_1, p_2] &= \lambda \sum_q \mathbf{g}[q]\Lambda^{(t)}(q, q + p_2 - p_1) \\ & \times \mathbf{g}[q + p_2 - p_1] + O(\lambda^2), \end{aligned} \quad (106)$$

and similarly for the  $\Lambda$  vertices:

as discussed in the second paragraph below Eq. (86), although this choice is irrelevant to the verification of the Ward identity.

## X. TWO-PARTICLE RESPONSE

We are interested in the pair correlations of the density  $n_a = \sum_\sigma X_a^{\sigma\sigma}$  and the spin density  $S_a^z = \frac{1}{2} \sum_{\sigma_1, \sigma_2} \tau_{\sigma_1 \sigma_2}^z X_a^{\sigma_1 \sigma_2}$ , where  $\tau^z$  is the usual Pauli matrix. These can be obtained from taking the functional derivatives of the Green's function

$$\Upsilon_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}[i, j] = \frac{\delta}{\delta V_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}} \mathcal{G}_{\sigma_1 \sigma_2}[i^-, i], \quad (109)$$

and can be conveniently found from taking a limit of the three-site object  $\Upsilon_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(p, q; r) = \frac{\delta}{\delta V_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}} \mathcal{G}_{\sigma_1 \sigma_2}[p, q]$ . With the singlet and triplet objects denoted with a superscript  $\alpha = s, t$ , we note the following relationships with the standard charge and spin susceptibilities of interest:

$$\begin{aligned} \langle\langle n_a(\tau_a)n_b(\tau_b) \rangle\rangle &= n^2 - 2\Upsilon^{(s)}(a, b), \\ \langle\langle S_a^z(\tau_a)S_b^z(\tau_b) \rangle\rangle &= -\frac{1}{2}\Upsilon^{(t)}(a, b). \end{aligned} \quad (110)$$

Owing to the bosonic nature of the densities, we have the symmetry  $\Upsilon^{(\alpha)}(b, a) = \Upsilon^{(\alpha)}(a, b)$  from which the Fourier transform at  $Q \equiv (\bar{Q}, i\Omega_q)$  satisfies the relation:

$$\Upsilon^{(\alpha)}(Q) = \Upsilon^{(\alpha)}(-Q). \quad (111)$$

This symmetry can be used as another test of the consistency of any approximation.

The Green's function in Eq. (109) can be decomposed in to  $\mathbf{g}$  and  $\mu$  as before and we find

$$\begin{aligned} \Upsilon_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}(a, b; r) &= \frac{\delta}{\delta\Upsilon_r^{\sigma_3\sigma_4}} \{ \mathbf{g}_{\sigma_1\sigma_a}[a, \mathbf{a}] \mu_{\sigma_a\sigma_2}[\mathbf{a}, b] \}, \\ &= \{ \mathbf{g}[a, \mathbf{b}] \Lambda_{\sigma_3\sigma_4}(\mathbf{b}, \mathbf{s}; r) \mathbf{g}[\mathbf{s}, \mathbf{a}] \mu[\mathbf{a}, q] \}_{\sigma_1\sigma_2} \\ &\quad + \{ \mathbf{g}[a, \mathbf{b}] \mathcal{U}_{\sigma_3\sigma_4}(\mathbf{b}, b; r) \}_{\sigma_1\sigma_2}, \end{aligned} \quad (112)$$

where the vertex and  $\Upsilon$  carry upper spin indices that are part of the matrix product. Turning off the sources, we find the expressions for singlet and triplet response

$$\begin{aligned} \Upsilon^{(\alpha)}(a, b; r) &= \mathbf{g}[a, \mathbf{b}] \Lambda^{(\alpha)}(\mathbf{b}, \mathbf{s}) \mathbf{g}[\mathbf{s}, \mathbf{a}] \mu[\mathbf{a}, b] \\ &\quad + \mathbf{g}[a, \mathbf{b}] \mathcal{U}^{(\alpha)}(\mathbf{b}, b; r), \\ \Upsilon^{(\alpha)}(p_1, p_2) &= \mathbf{g}[p_1] \Lambda^{(\alpha)}(p_1, p_2) \mathbf{g}[p_2] \mu[p_2] \\ &\quad + \mathbf{g}[p_1] \mathcal{U}^{(\alpha)}(p_1, p_2), \end{aligned} \quad (113)$$

where  $\alpha = s, t$ . The definitional distinction between left and right derivatives leads to the asymmetry in the above equations making it necessary to test the consistency Eq. (111) term by term.

Using the zero-source limit notation from Ref. 11,

$$\begin{aligned} Q^{(1)} &= Q_{\sigma\sigma}^{\sigma\sigma}, \quad Q^{(2)} = Q_{\bar{\sigma}\bar{\sigma}}^{\sigma\sigma}, \quad Q^{(3)} = Q_{\sigma\bar{\sigma}}^{\sigma\bar{\sigma}}, \\ Q^{(a)} &= Q^{(2)} - Q^{(3)}, \quad Q^{(s)} = Q^{(1)} + Q^{(2)}, \\ Q^{(t)} &= Q^{(1)} - Q^{(2)} = Q^{(3)}. \end{aligned} \quad (114)$$

The charge  $\alpha = s$  and spin  $\alpha = t$  susceptibilities at finite  $Q \equiv (\bar{Q}, i\Omega_q)$  are given by setting  $p_2 \rightarrow p$  and  $p_1 \rightarrow p + Q$  and summing over  $p$ :

$$\begin{aligned} \Upsilon^{(\alpha)}(Q) &\equiv \sum_p \Upsilon^{(\alpha)}(p, p + Q) \\ &= \sum_p [ \mathbf{g}[p] \Lambda^{(\alpha)}(p, p + Q) \mathbf{g}[p + Q] \mu[p + Q] \\ &\quad + \mathbf{g}[p] \mathcal{U}^{(\alpha)}(p, p + Q) ]. \end{aligned} \quad (115)$$

These are exact expression for the susceptibilities, but as usual require a knowledge of the vertices and Green's functions to give practical results. We can now use the RPA vertices calculated in Sec. IX to give the corresponding expressions.

We denote the susceptibility of the auxiliary fermions as

$$\chi_{\Lambda}^{(\alpha)}(Q) \equiv - \sum_q \mathbf{g}[q] \Lambda^{(\alpha)}(q, q + Q) \mathbf{g}[q + Q], \quad (116)$$

and within RPA we note that  $\mu[p]$  is independent of  $p$ , and from Eq. (106) we denote that the  $\mathcal{U}$  vertices are functions of the momentum difference only:

$$\mathcal{U}^{(\alpha)}[p_1, p_2] = \lambda \xi_{\alpha} \chi_{\Lambda}^{(\alpha)}(p_2 - p_1), \quad (117)$$

where  $\xi_{\alpha}$  is 1 for  $\alpha = \text{singlet}$  and  $-1$  for  $\alpha = \text{triplet}$ . Therefore we can sum over the  $p$  dependence of the second term and

rewrite Eq. (115) as

$$(\Upsilon^{(\alpha)}(Q))_{RPA} = -C_{\alpha} \chi_{\Lambda}^{(\alpha)}(Q), \quad (118)$$

where  $C_{\alpha} = (\mu - \xi_{\alpha} \lambda \frac{n}{2})$ . It seems more appropriate to reset  $\mu = (1 - \lambda \frac{n}{2})$  from unity at this level, in order to recover the expected high-frequency behavior in the charge as well as spin channel, so that  $C_{\text{singlet}} = (1 - \lambda n) \rightarrow 1 - n$  and  $C_{\text{triplet}} = 1$ . The vertices  $\Lambda$  are to be computed from Eq. (107) and form a consistent set of equations for the two-particle response in the sense of the usual RPA.

The integral equations must be solved numerically. However in order to display some flavor of the results, we pursue this to the lowest order in  $\lambda$  by iteration, where explicit results can be obtained. Let us define a few frequently occurring generalized polarizability functions for convenience. We will now reinstate  $J_k \rightarrow J_k - u_0$ :

$$\begin{aligned} \chi_0(Q) &= - \sum_q \mathbf{g}[q] \mathbf{g}[q + Q], \\ \chi_1(Q) &= - \sum_q \mathbf{g}[q] \mathbf{g}[q + Q] \{ \varepsilon_q + \varepsilon_{q+Q} \}, \\ \chi_2(Q) &= \frac{1}{2} \sum_{r,p} \mathbf{g}[r] \mathbf{g}[r + Q] \mathbf{g}[p] \mathbf{g}[p + Q] J_{p-r}, \\ F(p + Q, p) &= \sum_r \mathbf{g}[r] \mathbf{g}[r + Q] \left\{ \varepsilon_p + \varepsilon_{p+Q} + \varepsilon_r + \varepsilon_{r+Q} \right. \\ &\quad \left. + \frac{1}{2} (J_Q + J_{p-r}) \right\}. \end{aligned} \quad (119)$$

Here  $\chi_0(Q)$  is the standard Lindhard function and is positive in the static limit as  $\bar{Q} \rightarrow 0$ , while the other functions are generalizations thereof.

The answers are

$$\begin{aligned} \Upsilon^{(s)}(Q) &= -(1 - \lambda n) \chi_0(Q) - \lambda [ 2\chi_0(Q) \chi_1(Q) \\ &\quad - (u_0 - \frac{1}{2} J_Q) \chi_0^2(Q) + \chi_2(Q) ], \\ \Upsilon^{(t)}(Q) &= -\chi_0(Q) + \lambda [ 2\chi_0(Q) \chi_1(Q) \\ &\quad + (\frac{1}{2} J_Q - u_0) \chi_0^2(Q) + \chi_2(Q) ]. \end{aligned} \quad (120)$$

It is clear that the role of  $u_0$  enhances the spin susceptibility while decreasing the charge susceptibility. To this order we see that the parity test Eq. (111) is satisfied by using the symmetries of the objects in Eq. (119).

Since the Green's function remains infinitely sharp within the RPA, its usefulness is limited, especially in view of the large frequency-dependent corrections with characteristic asymmetry seen in second-order results in Refs. 2, 9, and 10. A second-order version of RPA seems most desirable, although even without vertex corrections to second order, the single-particle spectral results are very interesting already. It also seems interesting to study phenomenologically the analog of the "bubble" diagram for purposes of extracting the optical conductivity; a scheme that reflects the width of the physical Green's function and satisfies the parity requirement Eq. (111) is given by

$$[\mathcal{U}(Q)]_{\text{phen}} = - \frac{1}{1 - n/2} \sum_q \mathcal{G}[q] \mathcal{G}[q + Q], \quad (121)$$

although this expression is not the result a systematic expansion of Eq. (115).

## XI. DISCUSSION AND CONCLUSIONS

We have described above a controlled technique of dealing with the  $t$ - $J$  model. This extremely correlated Fermi liquid theory is a strong-coupling approach, specifically designed to deal with a hard many-body problem. The considerations begin with the strong-coupling limit of the Hubbard model, leading to the  $t$ - $J$  model with a hard constraint of eliminated double occupancy. The Schwinger method gives us a crucial initial platform to deal with this problem. The ensuing exact functional differential equations are made tractable by the introduction of the exact *product ansatz*,  $\mathcal{G} = \mathbf{g} \times \mu$ , with  $\mathbf{g}$  a canonical Green's function of auxiliary electrons and  $\mu$  the caparison factor. The latter, in turn, is understood as an adaptive spectral weight balancing the requirements at the high- and low-frequency ends of the spectrum. Both objects are expanded in powers of a parameter  $\lambda$  that plays the role of fractional double occupancy. Thus  $\lambda = 1$  corresponds to complete elimination of double occupancy whereas  $\lambda < 1$  has some residual double occupancy. We thus replace the hard constraint of complete elimination of double occupancy by a softer one or partial removal. In order to provide a natural description of the canonical electrons, we introduce the effective Hamiltonian  $H_{\text{eff}}$ , depending parametrically on  $\lambda$ . In order to obey the shift theorems I and II, we find it obligatory to (re)introduce a Hubbard-type  $u_0$  parameter in this model. It also plays the role of a second chemical potential as explained above. The set of steps followed, in our starting as well as ending up with a Hubbard-type interaction, has a slightly circular feel to it. This recipe is perhaps best understood as a renormalization group type procedure, where the constraint of single occupancy is enforced incrementally and the density of doubly occupied sites is thinned out smoothly. The infinite starting value of  $U$  in the  $t$ - $J$  model is pushed downward to  $u_0$ , typically a fraction of the bandwidth from our numerical studies, albeit in a more general model  $H_{\text{eff}}$ , and is therefore amenable to a perturbative expansion. The form of the  $H_{\text{eff}}$  and the important role of the shift symmetries in validating the approximations is noteworthy. The hopping  $t_{ij}$  is elevated to an interaction constant of the model; this unfamiliar step is kept under check by requiring the two important shift invariances. The Schwinger equation Eq. (42) for  $\mathcal{G}$ , being an exact statement of the problem, provides us with a rigorous backdrop to the entire procedure. Further our procedure has the advantage of being systematically improvable through the iterative scheme developed here.

We can explore superconductivity at a qualitative level by studying the pairing instabilities of the auxiliary fermions given by  $H_{\text{eff}}$  via its BCS gap function  $\Delta(k)$ . In this first approximation, the physical electron order parameter  $\langle X_k^{\uparrow 0} X_{-k}^{\downarrow 0} \rangle$  is proportional to that of the auxiliary electrons  $\langle f_{\uparrow}^{\dagger}(k) f_{\downarrow}^{\dagger}(-k) \rangle$ , together with the single-occupancy constraint of vanishing upon summing over the wave vector  $k$ . Within a generalized Hartree Fock theory, retaining the self-energy correction to first order [as in Eq. (108)] as well as the pairing field average,

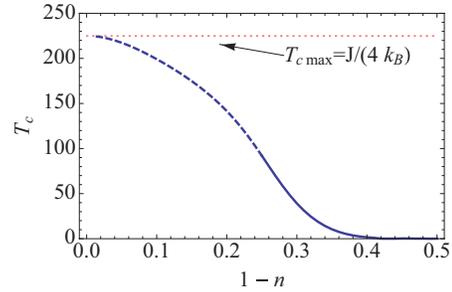


FIG. 5. (Color online) The transition temperature in kelvins, from solving Eq. (122) assuming  $t = 3000$  K and  $J = 900$  K. The solid line indicates the likely regime of validity of the  $O(\lambda)$  theory. Its dotted extension to lower hole density is speculative and is most likely to change with higher order corrections reflecting the nearby Mott insulating state. The dotted red line indicates the maximum  $T_c$  obtainable from this scheme, and is seen to depend solely upon the magnitude of  $J$ .

we obtain an equation for the gap function  $\Delta(k)$ :

$$\Delta(k) = \frac{1}{N_s} \sum_p \left\{ \varepsilon_k + \varepsilon_p - u_0 + \frac{1}{2} J_{k-p} \right\} \times \Delta(p) \frac{\tanh \beta E(p)/2}{E(p)}, \quad (122)$$

where  $E(p) = \sqrt{\Delta^2(p) + \xi_p^2}$ , and  $\xi_p = \varepsilon_p(1-n) - \frac{1}{2} \sum_q J_{q-p} n_q - \mu$ . In the computation below, we will neglect the numerically small  $J$  term in the single-particle energy. Other than  $u_0$  and the two single-particle energies in Eq. (122) required for satisfying the shift theorems I and II, this is the same equation as the one found within the resonating valence bond theory in Refs. 4, 19, and 20. The transition temperature for a  $d$ -wave state with a gap function  $\Delta(k) = \Delta_d[\cos(k_x) - \cos(k_y)]$  is obtained by solving Eq. (122) for the case of the nearest-neighbor square lattice  $t$ - $J$  model, with parameters indicated in the caption of Fig. 5.

It is straightforward to see that the  $T_c$  equation has a maximum scale of order  $J/(4k_B)$  as already noted in Ref. 19 and Ref. 20. This value is attained in this solution at a higher particle density, or equivalently, a lower hole density, than is warranted by the first approximation. The solid line represents a plausible regime of validity of this scheme.

The extended  $s$ -wave order is usually described by a gap function  $\Delta(k) = \Delta_{s,0} + \Delta_{s,1} [\cos(k_x) + \cos(k_y)]$ . The constant term  $\Delta_{s,0}$  leads to a finite probability of double occupancy, since it survives a wave vector sum. After it is dropped as per the above discussion, the assumed (purely extended)  $s$ -wave order is supported by the  $J$  term in the kernel of Eq. (122), but not by the  $u_0$ -dependent and single-particle energy terms. The latter thus do not play a role in determining  $T_c$  for either  $d$ -wave or  $s$ -wave orders despite their large magnitude relative to  $J$ .

A detailed calculation of the gap equation is planned for the pairing of physical particles, parallel to the  $O(\lambda^2)$  theory of the normal state. The finite-lifetime effects are then expected to become relevant. Such an improvement of the pairing scheme should yield a greater understanding of the balance between

the different orders and a greater range of validity in density than the schematic theory treated here.

### ACKNOWLEDGMENTS

I have benefited from discussions with T. Banks, G.-H. Gweon, D. Hansen, and E. Perepelitsky at Santa Cruz, and with P. W. Anderson, A. Georges, A. C. Hewson, G. Kotliar, and Y. Kuramoto elsewhere. This work was supported by DOE under Grant No. FG02-06ER46319.

### APPENDIX A: ATOMIC LIMIT $t = J = 0$

#### 1. Double-occupancy interpretation of $\lambda$ from the atomic limit

In order to understand the role of  $\lambda$  we study the atomic limit  $t, J \rightarrow 0$  where this parameter can be introduced into the physical Green's function in the form

$$\mathcal{G}[\lambda, i\omega_n] = \frac{1 - \lambda \frac{n}{2}}{i\omega_n + \mu}, \quad (\text{A1})$$

and study its dependence on  $\lambda$  in the interval  $[0, 1]$ . The chemical potential  $\mu$  can be calculated from the sum rule on the density  $n$  of the number of particles  $N$  with  $n = N/N_s$  and temperature  $T$  as

$$\mu = k_B T \ln \left( \frac{n}{2 - (1 + \lambda)n} \right). \quad (\text{A2})$$

Thermodynamics tells us that the entropy  $S$  can be expressed as

$$S(n) = -N_s \int_0^n dn' \frac{\partial \mu(n')}{\partial T}, \quad (\text{A3})$$

and since we know  $\mu$  from Eq. (A2) we obtain with  $y = (1 + \lambda)n$

$$\frac{S(n, \lambda)}{k_B N_s} = \frac{1}{1 + \lambda} \{ \ln 4 - y \ln n - (2 - y) \ln (2 - y) \}. \quad (\text{A4})$$

We see that its  $\lambda$  derivative  $\frac{1}{k_B N_s} \frac{\partial S}{\partial \lambda} = \frac{2}{(1 + \lambda)^2} [\frac{y}{2} + \ln(1 - \frac{y}{2})]$  is negative definite. Thus we see that the entropy at a fixed density interpolates monotonically between the free Fermi limit and the infinite  $U$  limits as  $\lambda$  ranges over its domain  $0 \leq \lambda \leq 1$ . The maximum allowed density is reduced from 2 to  $\frac{2}{1 + \lambda}$  and thus at  $\lambda = 1$  we have a maximum of one electron per site, as expected physically. Thus increasing  $\lambda$  from zero effectively removes the available states contributing to entropy; its role may be viewed as that of (continuous) removal of states. Thus for the equations of motion it is somewhat analogous to the role of Gutzwiller's parameter  $g$  in his projection operator  $\prod_i [1 - (1 - g)n_{i\uparrow}n_{i\downarrow}]$  at the wave function level.

In the atomic limit we can also calculate the entropy at a fixed density of doubly occupied sites  $d = \frac{1}{N_s} \sum_i n_{i\uparrow}n_{i\downarrow}$  as

$$\begin{aligned} \frac{S(n, d)}{k_B N_s} &= -d \ln d - (n - 2d) \ln \left( \frac{n}{2} - d \right) \\ &\quad - (1 + d - n) \ln(1 + d - n). \end{aligned} \quad (\text{A5})$$

An uncorrelated system corresponds to  $d = \frac{n^2}{4}$ , where the entropy Eq. (A5) is a maximum, while  $d = 0$  for the fully

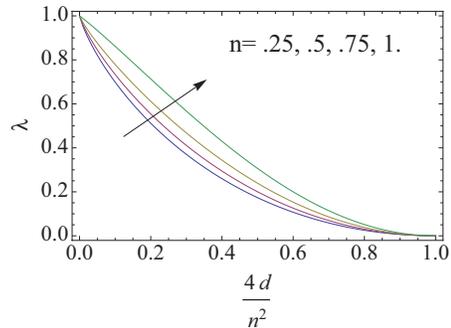


FIG. 6. (Color online) The parameter  $\lambda$  is determined in terms of the double occupancy  $d$  at various densities in the atomic limit as described in the text. The arrow indicates increasing density  $n$ . Note that the parameter  $d$  is scaled into the unit interval.

projected  $t$ - $J$  model. Comparing the two expressions for entropy Eq. (A4) and Eq. (A5), we can express  $\lambda$  in terms of  $d$  at any density. We have thus demonstrated that  $\lambda$  is a conjugate variable to the double-occupation density in this limit. Their explicit relationship is illustrated in Fig. 6.

#### 2. Expansion in $\lambda$ in the atomic limit

In the atomic limit we set  $t \rightarrow 0$  and  $J \rightarrow 0$  so that Eq. (67) and Eq. (68) become

$$\begin{aligned} \mathbf{g}[i, f] &= \mathbf{g}_0[i, f; \mu_a], \\ \mu[i, f] &= \delta[i, f] (1 - \lambda \gamma[i]) - \lambda \mu_b \mathbf{g}[i, \mathbf{f}] \cdot \mu[\mathbf{f}, f]. \end{aligned} \quad (\text{A6})$$

Here we split the chemical potential into two pieces  $\mu = \mu_a + \lambda \mu_b$ . Thus in this limit  $\mathbf{g}$  is the free Fermi Green's function independent of  $\lambda$ , and  $\mu_a$  is the free value  $\mu_a \rightarrow \mu_0$ , the latter determined from the noninteracting theory in terms of the number of particles. If we turn off the source  $\mathcal{V}$  the Fourier transforms can be taken as

$$\begin{aligned} \mathbf{g}[i\omega_n] &= \mathbf{g}_0[i\omega_n; \mu_a] = \frac{1}{i\omega_n + \mu_0}, \\ \mu[i\omega_n] &= \left( 1 - \lambda \frac{n}{2} \right) - \lambda \mu_b \mathbf{g}[i\omega_n] \mu[i\omega_n], \\ &= \frac{1 - \lambda \frac{n}{2}}{1 + \lambda \mu_b \mathbf{g}[i\omega_n]}. \end{aligned} \quad (\text{A7})$$

Thus the physical Green's function

$$\mathcal{G}[i\omega_n] = \frac{1 - \lambda \frac{n}{2}}{i\omega_n + \mu_0 + \lambda \mu_b}. \quad (\text{A8})$$

We fix the chemical potentials from the number sum rule as usual and thus

$$\begin{aligned} \frac{n}{2} &= \frac{1}{1 + e^{-\beta \mu_0}}, \\ \frac{n}{2} &= \left( 1 - \lambda \frac{n}{2} \right) \frac{1}{1 + e^{-\beta(\mu_0 + \lambda \mu_b)}}. \end{aligned} \quad (\text{A9})$$

We may then solve for  $\mu$ 's in terms of the density and obtain

$$\begin{aligned} \mu_0 &= k_B T \ln \left( \frac{n}{2 - n} \right), \\ \lambda \mu_b &= k_B T \ln \left( \frac{2 - n}{2 - (1 + \lambda)n} \right). \end{aligned} \quad (\text{A10})$$

Thus the chemical potential  $\mu_1$  has a power series representation

$$\mu_b = \sum_{m=0}^{\infty} \lambda^m \mu_b^{(m)} = k_B T \sum_{m=0}^{\infty} \frac{\lambda^m}{m+1} \left[ \frac{n}{2-n} \right]^{m+1}. \quad (\text{A11})$$

We see explicitly from Eq. (A11) that the  $\lambda$  expansion of the the atomic limit is an expansion in  $\lambda n/(2-n)$ , i.e., a density expansion as well.

## APPENDIX B: THE LOW-ORDER CALCULATIONS OF GREEN'S FUNCTIONS

### 1. Green's function to $O(\lambda)$

We evaluate the complete starting point of the hierarchy. We start with terms of  $O(\lambda^0)$  and end with  $[\mu]_1$  and  $[\mathbf{g}^{-1}]_1$ , which are the seeds for the  $O(\lambda)$  terms.

#### a. Seed terms and initialization

$$\begin{aligned} \mathbf{g}_0^{-1}[i,m] &= \{[(\mu - \partial_{\tau_i} - \frac{1}{4}J_0)\mathbb{1} - \mathcal{V}_i] \\ &\quad \times \delta[i,m] + t[i,m] - \mathcal{V}_{i,m}\}, \quad (\text{B1}) \\ [\mu[i,f]]_0 &= \mathbb{1}\delta[i,f]. \end{aligned}$$

Derived objects:

$$\begin{aligned} [\gamma[i]]_0 &= \mathbf{g}^{(k)}[i,i], \\ [\gamma[i,m]]_0 &= \mathbf{g}^{(k)}[m,i], \quad (\text{B2}) \\ [\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]]_0 &= 0, \\ [\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]]_0 &= 0. \end{aligned}$$

$$\begin{aligned} [Y_1[i,m]]_0 &= t[i,m](\mathbf{g}^{(k)}[i,i] + \frac{1}{2}\mathbf{g}^{(k)}[m,m]) \\ &\quad - \delta[i,m]\frac{1}{2}(J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] - t[i,\mathbf{j}]\mathbf{g}^{(k)}[i,\mathbf{j}]), \\ [\Psi[i,m]]_0 &= 0. \quad (\text{B3}) \end{aligned}$$

$$\begin{aligned} [\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]]_0 &= \delta_{\sigma_a\sigma_c}\delta_{\sigma_b\sigma_d}\delta[i,j]\delta[j,m], \\ [\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]]_0 &= \delta_{\sigma_a\sigma_c}\delta_{\sigma_b\sigma_d}\delta[i,j]\delta[m,k]\delta(\tau_j - \tau_k). \quad (\text{B4}) \end{aligned}$$

In the four-point vertex above, we have introduced the delta function  $\delta(\tau_j - \tau_k)$ , so that the labels  $i,m,j,k$  can be viewed as four independent space-time variables.

Thus

$$\begin{aligned} [\Phi[i,m]]_0 &= \delta[i,m]t[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},i] + \frac{1}{2}t[i,m]\mathbf{g}^{(k)}[m,m] \\ &\quad + \frac{1}{2}\delta[i,m]t[i,\mathbf{j}]\mathbf{g}^{(k)}[i,\mathbf{j}] - \frac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m]. \quad (\text{B5}) \end{aligned}$$

Combining the two we get

$$\begin{aligned} [Y_1[i,m]]_0 + [\Phi[i,m]]_0 &= \delta[i,m]t[i,\mathbf{j}](\mathbf{g}^{(k)}[i,\mathbf{j}] + \mathbf{g}^{(k)}[\mathbf{j},i]) \\ &\quad + t[i,m](\mathbf{g}^{(k)}[i,i] + \mathbf{g}^{(k)}[m,m]) - \frac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m] \\ &\quad - \delta[i,m]\frac{1}{2}J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}]. \quad (\text{B6}) \end{aligned}$$

### b. Stepping up and final Green's function to $O(\lambda)$

To first order in  $\lambda$  we collect the above results to obtain the Green's function

$$\begin{aligned} [\mathbf{g}^{-1}[i,m]]_1 &= \mathbf{g}^{(k)}[i,i] \cdot \mathcal{V}_{i,m} + \delta[i,m]\mathbf{g}^{(k)}[\mathbf{a},i] \cdot \mathcal{V}_{i,\mathbf{a}}^{(k)} \\ &\quad + \delta[i,m]\frac{1}{2}J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] + \frac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m] \\ &\quad - t[i,m](\mathbf{g}^{(k)}[i,i] + \mathbf{g}^{(k)}[m,m]) \\ &\quad - \delta[i,m]t[i,\mathbf{j}](\mathbf{g}^{(k)}[i,\mathbf{j}] + \mathbf{g}^{(k)}[\mathbf{j},i]) \quad (\text{B7}) \end{aligned}$$

and the caparison factor

$$[\mu[i,m]]_1 = -\mathbf{g}^{(k)}[i^-,i]\delta[i,m]. \quad (\text{B8})$$

The FTs of these on turning off the sources are found using  $\mathbf{g}[i^-,i] \rightarrow \frac{n}{2}$  as

$$\begin{aligned} [\mathbf{g}^{-1}[k]]_1 &= n\varepsilon_k - \frac{n}{2}u_0 + \frac{1}{2}\sum_q J_{k-q}\mathbf{g}[q] \\ &\quad + \left( \frac{1}{4}J_0n + 2\sum_q \varepsilon_q\mathbf{g}[q] \right), \quad (\text{B9}) \end{aligned}$$

$$[\mu[k]]_1 = -\frac{n}{2}. \quad (\text{B10})$$

The term  $-\frac{n}{2}u_0$  in Eq. (B9) arises when we reinstate  $J[i,j] \rightarrow J[i,j] - u_0\delta[i,j]$  in Eq. (B7). Let us note that under the shift Eq. (13), the first-order correction  $[\mathbf{g}^{-1}[k]]_1$  shifts by  $2nu_i + \frac{n}{2}u_j$ . Therefore this term is invariant under the shift theorem II and also the shift theorem I.1, provided  $u_0$  is simultaneously transformed as specified in Eq. (31).

### 2. Green's function to $O(\lambda^2)$

#### a. $\mu$ derived objects

We next start with seed terms of  $O(\lambda)$  calculated above and end with  $[\mu]_2$  and  $[\mathbf{g}^{-1}]_2$ :

$$[\mu[i,m]]_1 = -\mathbf{g}^{(k)}[i^-,i]\delta[i,m]. \quad (\text{B11})$$

Let us calculate the derived quantities from the above at the same level:

$$\begin{aligned} [\gamma[i]]_1 &= [\mu^{(k)}[\mathbf{a},i]]_1 \cdot \mathbf{g}^{(k)}[i,\mathbf{a}] = -\mathbf{g}[i,i] \cdot \mathbf{g}^{(k)}[i,i], \\ [\gamma[i,m]]_1 &= [\mu^{(k)}[\mathbf{a},i]]_1 \cdot \mathbf{g}^{(k)}[m,\mathbf{a}] = -\mathbf{g}[i,i] \cdot \mathbf{g}^{(k)}[m,i], \\ [Y[i,m]]_1 &= -t[i,m](\mathbf{g}[i,i] \cdot \mathbf{g}^{(k)}[i,i] + \frac{1}{2}\mathbf{g}[m,m] \cdot \mathbf{g}^{(k)}[m,m]) \\ &\quad + \frac{1}{2}\delta[i,m](J[i,\mathbf{j}]\mathbf{g}[\mathbf{j},\mathbf{j}] \cdot \mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] \\ &\quad - t[i,\mathbf{j}]\mathbf{g}[\mathbf{j},\mathbf{j}] \cdot \mathbf{g}^{(k)}[i,\mathbf{j}]). \quad (\text{B12}) \end{aligned}$$

Zero-source Fourier transforms:

$$\begin{aligned} ([\gamma[0]]_1)_{\nu \rightarrow 0} &= -\frac{n^2}{4}, \\ ([\gamma[k]]_1)_{\nu \rightarrow 0} &= -\frac{n}{2}\mathbf{g}[-k], \quad (\text{B13}) \\ ([Y_1[k]]_1)_{\nu \rightarrow 0} &= \frac{3n^2}{8}\varepsilon_k - \frac{n^2}{8}u_0 + \left( \frac{n^2}{8}J_0 + \frac{n}{4}\sum_q \varepsilon_q\mathbf{g}[q] \right). \end{aligned}$$

Here we reinstated  $J[i,j] \rightarrow J[i,j] - u_0\delta[i,j]$  in Eq. (B12) to obtain the  $-\frac{n^2}{8}u_0$  term in Eq. (B13).

*b.  $\mu$  derived vertices*

Next we calculate (using lowest-order functional derivatives)

$$\begin{aligned} [\mathcal{U}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[i,m;a]]_1 &= -\delta[i,m]\sigma_1\sigma_2\mathbf{g}_{\bar{\sigma}_2\sigma_3}[i,a]\mathbf{g}_{\sigma_4\bar{\sigma}_1}[a,i], \\ [\mathcal{U}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[i,m;a]]_1 \nu \rightarrow 0 &= -\delta[i,m]\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\mathbf{g}[i,a]\mathbf{g}[a,i]. \end{aligned} \quad (\text{B14})$$

At zero sources so with  $\mathcal{V} \rightarrow 0$

$$\begin{aligned} [\mathcal{U}(a)[i,m;a]]_1 &= -2\delta[i,m]\mathbf{g}[i,a]\mathbf{g}[a,i], \\ [\mathcal{U}^{(s)}[i,m;a]]_1 &= -\delta[i,m]\mathbf{g}[i,a]\mathbf{g}[a,i]. \end{aligned} \quad (\text{B15})$$

The zero-source Fourier transforms are as follows:

$$\begin{aligned} [\mathcal{U}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p_1,p_2]]_1 &= -\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\sum_q \mathbf{g}[q]\mathbf{g}[q+p_2-p_1], \\ [\mathcal{U}^{(a)}[p_1,p_2]]_1 &= -2\sum_q \mathbf{g}[q]\mathbf{g}[q+p_2-p_1]. \end{aligned} \quad (\text{B16})$$

Similarly we find for the four-index vertices

$$\begin{aligned} [\mathcal{U}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[i,m;a,b]]_1 &= -\delta[i,m]\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\mathbf{g}[i,a]\mathbf{g}[b,i]\delta(\tau_a-\tau_b), \\ [\mathcal{U}^{(a)}[i,m;a,b]]_1 \nu \rightarrow 0 &= -2\delta[i,m]\mathbf{g}[i,a]\mathbf{g}[b,i]\delta(\tau_a-\tau_b). \end{aligned} \quad (\text{B17})$$

The zero-source Fourier transforms are as follows:

$$\begin{aligned} [\mathcal{U}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p_1,p_2;p_3,p_4]]_1 &= -\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\delta_{p_1+p_4,p_2+p_3} \\ &\quad \times \mathbf{g}[p_3]\mathbf{g}[p_4], \\ [\mathcal{U}^{(a)}[p_1,p_2;p_3,p_4]]_1 &= -2\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4]. \end{aligned} \quad (\text{B18})$$

*c.  $\Psi$  to  $O(\lambda)$* 

We compute  $[\Psi]_1$  from these:

$$\begin{aligned} [\Psi(k)]_1 &= \sum_p \left( \varepsilon_p + \frac{1}{2}\varepsilon_k + \frac{1}{2}J_{k-p} \right) \mathbf{g}[p][\mathcal{U}^{(a)}(p,k)]_1 \\ &\quad + \sum_{pq} \frac{1}{2}\varepsilon_{q+p-k} \mathbf{g}[p][\mathcal{U}^{(a)}(p,k;q+p-k,q)]_1 \\ &= -\sum_{p,q} (\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p}) \\ &\quad \times \mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[q+k-p] \\ &= \sum_{p,q} W(k,q;q+k-p,p)\mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[q+k-p]. \end{aligned} \quad (\text{B19})$$

*d. Stepping up:  $\mu$  to  $O(\lambda^2)$* 

Stepping up, we calculate

$$[\mu[i,m]]_2 = -\delta[i,m]\mathbf{g}[i^-,i]\mathbf{g}^{(k)}[i^-,i] + [\Psi[i,m]]_1. \quad (\text{B20})$$

Hence at zero sources, the Fourier transform reads

$$\begin{aligned} [\mu[k]]_2 &= \frac{n^2}{4} - \sum_{p,q} (\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p}) \\ &\quad \times \mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[q+k-p]. \end{aligned} \quad (\text{B21})$$

Note that  $[\mu]_2, [\Psi]_1$  are invariant under all three shift theorems. It is clear that this is a more nontrivial application of the theorems than those in the lowest order.

*e.  $\mathbf{g}^{-1}$  derived objects*

Let us now start with  $[\mathbf{g}^{-1}[i,m]]_1$  given in Eq. (B7):

$$\begin{aligned} [\mathbf{g}^{-1}[i,m]]_1 &= [\mathbf{g}^{(k)}[i,i] \cdot \mathcal{V}_{i,m} + \delta[i,m]\mathbf{g}^{(k)}[\mathbf{a},i] \cdot \mathcal{V}_{i,\mathbf{a}}^{(k)}] \\ &\quad + [\delta[i,m]\frac{1}{2}J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] + \frac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m]] \\ &\quad \times [-t[i,m](\mathbf{g}^{(k)}[i,i] + \mathbf{g}^{(k)}[m,m]) \\ &\quad - \delta[i,m]t[i,\mathbf{j}](\mathbf{g}^{(k)}[i,\mathbf{j}] + \mathbf{g}^{(k)}[\mathbf{j},i])]. \end{aligned} \quad (\text{B22})$$

*f. Vertex functions to  $O(\lambda)$* 

Three-point vertex:

$$\begin{aligned} [\Lambda_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[i,m;a]]_1 &= -\left( \frac{\delta}{\delta\mathcal{V}_a^{\sigma_3\sigma_4}} \right) [\mathbf{g}_{\sigma_1\sigma_2}^{-1}[i,m]]_1 \\ &= [(I) + (II) + (III)]_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}, \end{aligned} \quad (\text{B23})$$

where the terms  $(I),(II),(III)$  refer to the three square bracketed terms in Eq. (B22). For the first term we calculate

$$\begin{aligned} (I)_{\sigma_3\sigma_4}^{\sigma_1\sigma_2} &= -\left( \frac{\delta}{\delta\mathcal{V}_a^{\sigma_3\sigma_4}} \right) [(\sigma_1\sigma_a)\mathbf{g}_{\bar{\sigma}_a\bar{\sigma}_1}[i,i]\mathcal{V}_{i,m}^{\sigma_a\sigma_2} \\ &\quad + \delta[i,m](\sigma_1\sigma_2)\mathbf{g}_{\bar{\sigma}_a\bar{\sigma}_1}[\mathbf{a},i]\mathcal{V}_{i,\mathbf{a}}^{\bar{\sigma}_2\bar{\sigma}_a}], \end{aligned} \quad (\text{B24})$$

$$\nu \rightarrow 0 = 0. \quad (\text{B25})$$

For all other terms we can use a simple calculation:

$$\begin{aligned} \mathcal{X}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p,q;r] &= \left( \frac{\delta}{\delta\mathcal{V}_r^{\sigma_3\sigma_4}} \right) \mathbf{g}_{\sigma_1\sigma_2}^{(k)}[p,q] \\ &= (\sigma_1\sigma_2)\mathbf{g}_{\bar{\sigma}_2\sigma_3}[p,r]\mathbf{g}_{\sigma_4\bar{\sigma}_1}[r,q], \\ \mathcal{X}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p,q;r] \nu \rightarrow 0 &= (\sigma_1\sigma_2)\delta_{\bar{\sigma}_2\sigma_3}\delta_{\bar{\sigma}_1\sigma_4}\mathbf{g}[p,r]\mathbf{g}[r,q], \\ \mathcal{X}^{(a)}[p,q;r] \nu \rightarrow 0 &= 2\mathbf{g}[p,r]\mathbf{g}[r,q]. \end{aligned} \quad (\text{B26})$$

Therefore

$$\begin{aligned} (II)^{(a)}[i,m;a] &= -\delta[i,m]J[i,\mathbf{j}]\mathbf{g}[\mathbf{j},a]\mathbf{g}[\mathbf{a},\mathbf{j}] \\ &\quad - J[i,m]\mathbf{g}[i,a]\mathbf{g}[\mathbf{a},m] \end{aligned} \quad (\text{B27})$$

and

$$\begin{aligned} (III)^{(a)}[i,m;a] &= 2[t[i,m](\mathbf{g}[i,a]\mathbf{g}[a,i] + \mathbf{g}[m,a]\mathbf{g}[a,m]) \\ &\quad + \delta[i,m]t[i,\mathbf{j}](\mathbf{g}[i,a]\mathbf{g}[\mathbf{a},\mathbf{j}] + \mathbf{g}[\mathbf{j},a]\mathbf{g}[a,i])]. \end{aligned} \quad (\text{B28})$$

Zero-source Fourier transforms read as

$$\begin{aligned} (II)^{(a)}[p_1, p_2] &= -J_{p_2-p_1} \sum_q \mathbf{g}[q] \mathbf{g}[q+p_2-p_1] - \sum_q J_{p_1-q} \mathbf{g}[q] \mathbf{g}[q+p_2-p_1], \\ (III)^{(a)}[p_1, p_2] &= -2 \sum_q \mathbf{g}[q] \mathbf{g}[q+p_2-p_1] \{\varepsilon_{p_2} + \varepsilon_{p_1} + \varepsilon_{q+p_2-p_1} + \varepsilon_q\}. \end{aligned} \quad (\text{B29})$$

Hence adding up we obtain

$$\begin{aligned} [\Lambda_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p_1, p_2]]_1 &= -(\sigma_1\sigma_2)\delta_{\bar{\sigma}_2\sigma_3}\delta_{\bar{\sigma}_1\sigma_4} \sum_{p_3, p_4} \delta_{p_1+p_4, p_2+p_3} \mathbf{g}[p_3] \mathbf{g}[p_4] \left\{ \varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + \frac{1}{2} (J_{p_1-p_2} + J_{p_1-p_3}) \right\} \\ &= \frac{1}{2} (\sigma_1\sigma_2)\delta_{\bar{\sigma}_2\sigma_3}\delta_{\bar{\sigma}_1\sigma_4} \sum_{p_3, p_4} \mathbf{g}[p_3] \mathbf{g}[p_4] [W(p_2, p_3, p_4, p_1) + W(p_2, p_3, p_1, p_4)], \end{aligned} \quad (\text{B30})$$

$$[\Lambda^{(a)}[p_1, p_2]]_1 = - \sum_{p_3, p_4} \mathbf{g}[p_3] \mathbf{g}[p_4] [W(p_2, p_3, p_4, p_1) + W(p_2, p_3, p_1, p_4)]. \quad (\text{B31})$$

Note that rotation invariance relations imply that since  $[\Lambda^{(1)}[p_1, p_2]]_1 = 0$ , we must have

$$[\Lambda^{(s)}[p_1, p_2]]_1 = -[\Lambda^{(t)}[p_1, p_2]]_1 = \frac{1}{2} [\Lambda^{(a)}[p_1, p_2]]_1. \quad (\text{B32})$$

*Four point vertex.* The calculation proceeds similarly:

$$[\Lambda_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[i, m; a, b]]_1 = - \left( \frac{\delta}{\delta \mathcal{V}_{a,b}^{\sigma_3\sigma_4}} \right) [\mathbf{g}_{\sigma_1\sigma_2}^{-1}[i, m]]_1 = [(IV) + (V) + (VI)]_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}. \quad (\text{B33})$$

Here the terms  $(IV)$ – $(VI)$  refer to the three square bracketed terms in Eq. (B22). For the first term we calculate with implicit  $\tau_a = \tau_b$ :

$$\begin{aligned} (IV)_{\sigma_3\sigma_4}^{\sigma_1\sigma_2} &= - \left( \frac{\delta}{\delta \mathcal{V}_{a,b}^{\sigma_3\sigma_4}} \right) [(\sigma_1\sigma_a)\mathbf{g}_{\bar{\sigma}_a\bar{\sigma}_1}[i, i] \mathcal{V}_{i,m}^{\sigma_a\sigma_2} + \delta[i, m] (\sigma_1\sigma_2)\mathbf{g}_{\bar{\sigma}_a\bar{\sigma}_1}[\mathbf{a}, i] \mathcal{V}_{i,\mathbf{a}}^{\bar{\sigma}_2\bar{\sigma}_a}] \\ &= -\delta[i, a] \delta[m, b] [(\sigma_1\sigma_3)\delta_{\sigma_2\sigma_4} \mathbf{g}_{\bar{\sigma}_3\bar{\sigma}_1}[i, i] - \delta[i, a] \delta[i, m] [(\sigma_1\sigma_2)\delta_{\bar{\sigma}_2, \sigma_3} \mathbf{g}_{\sigma_4\bar{\sigma}_1}[b, i]], \end{aligned} \quad (\text{B34})$$

$$\begin{aligned} (IV)_{\sigma_3\sigma_4}^{\sigma_1\sigma_2} \nu \rightarrow 0 &= -\delta[i, a] \delta[m, b] \delta_{\sigma_1\sigma_3} \delta_{\sigma_2\sigma_4} \mathbf{g}[i, i] - \delta[i, a] \delta[i, m] (\sigma_1\sigma_2) \delta_{\sigma_1\bar{\sigma}_4} \delta_{\sigma_2\bar{\sigma}_3} \mathbf{g}[b, i], \\ (IV)^{(a)} &= \underbrace{\delta[i, a] \delta[m, b] \mathbf{g}[i, i] - 2\delta[i, a] \delta[i, m] \mathbf{g}[b, i]}_{\text{dropped}}. \end{aligned} \quad (\text{B35})$$

This term is seen result in a violation of Eq. (49) and Eq. (72) for reasons discussed there and in the second remark below Eq. (68), and therefore is dropped below. We have carried it in the calculation, and demarcated it with the underbrace, in order to see its (minor) contribution explicitly before dropping it.

For all other terms we can use a simple calculation:

$$\begin{aligned} \mathcal{Y}_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p, q; r, s] &= \left( \frac{\delta}{\delta \mathcal{V}_{r,s}^{\sigma_3\sigma_4}} \right) \mathbf{g}_{\sigma_1\sigma_2}^{(k)}[p, q], \\ \mathcal{Y}^{(a)}[p, q; r, s]_{\nu \rightarrow 0} &= 2\mathbf{g}[p, r] \mathbf{g}[s, q] \delta(\tau_r - \tau_s). \end{aligned} \quad (\text{B36})$$

Therefore with implicit  $\tau_a = \tau_b$ :

$$(V)^{(a)}[i, m; a, b] = -\delta[i, m] J[i, \mathbf{j}] \mathbf{g}[\mathbf{j}, a] \mathbf{g}[b, \mathbf{j}] - J[i, m] \mathbf{g}[i, a] \mathbf{g}[b, m], \quad (\text{B37})$$

$$\begin{aligned} (VI)^{(a)}[i, m; a, b] &= 2[t[i, m](\mathbf{g}[i, a] \mathbf{g}[b, i] + \mathbf{g}[m, a] \mathbf{g}[b, m]) + \delta[i, m] t[i, \mathbf{j}] (\mathbf{g}[i, a] \mathbf{g}[b, \mathbf{j}] + \mathbf{g}[\mathbf{j}, a] \mathbf{g}[b, i])], \\ (IV)^{(a)}[i, m; a, b] &= \delta[i, a] \delta[m, b] \mathbf{g}[i, i] - 2\delta[i, a] \delta[i, m] \mathbf{g}[b, i], \\ (IV)^{(a)}[p_1, p_2, p_3, p_4] &= \underbrace{\delta_{p_1, p_3} \delta_{p_2, p_4} \mathbf{g}[0^-] - 2\delta_{p_1+p_4, p_2+p_3} \mathbf{g}[p_4]}_{\text{dropped}}, \end{aligned} \quad (\text{B38})$$

$$\begin{aligned} (V)^{(a)}[i, m; a, b] &= -\delta[i, m] J[i, \mathbf{j}] \mathbf{g}[\mathbf{j}, a] \mathbf{g}[b, \mathbf{j}] - J[i, m] \mathbf{g}[i, a] \mathbf{g}[b, m], \\ (V)^{(a)}[p_1, p_2, p_3, p_4] &= -(J_{p_2-p_1} + J_{p_1-p_3}) \mathbf{g}[p_3] \mathbf{g}[p_4] \delta_{p_1+p_4, p_2+p_3}, \end{aligned} \quad (\text{B39})$$

$$\begin{aligned} (VI)^{(a)}[i, m; a, b] &= 2[t[i, m](\mathbf{g}[i, a] \mathbf{g}[b, i] + \mathbf{g}[m, a] \mathbf{g}[b, m]) + \delta[i, m] t[i, \mathbf{j}] (\mathbf{g}[i, a] \mathbf{g}[b, \mathbf{j}] + \mathbf{g}[\mathbf{j}, a] \mathbf{g}[b, i])], \\ (VI)^{(a)}[p_1, p_2, p_3, p_4] &= -2\{\varepsilon_{p_2} + \varepsilon_{p_1} + \varepsilon_{p_3} + \varepsilon_{p_4}\} \mathbf{g}[p_3] \mathbf{g}[p_4] \delta_{p_1+p_4, p_2+p_3}. \end{aligned} \quad (\text{B40})$$

Hence

$$\begin{aligned} [\Lambda_{\sigma_3\sigma_4}^{\sigma_1\sigma_2}[p_1, p_2, p_3, p_4]]_1 = & -(\sigma_1\sigma_2)\delta_{\bar{\sigma}_2\sigma_3}\delta_{\bar{\sigma}_1\sigma_4}\delta_{p_1+p_4, p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4]\left\{\varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + \frac{1}{2}(J_{p_1-p_2} + J_{p_1-p_3})\right\} \\ & - \underbrace{\delta_{p_1, p_3}\delta_{p_2, p_4}\delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4}\mathbf{g}[0^-] - 2(\sigma_1\sigma_2)\delta_{\sigma_1\bar{\sigma}_4}\delta_{\sigma_2\bar{\sigma}_3}\delta_{p_1+p_4, p_2+p_3}\mathbf{g}[p_4]}_{\text{underbraces}}. \end{aligned} \quad (\text{B41})$$

Thus

$$\begin{aligned} [\Lambda^{(a)}[p_1, p_2, p_3, p_4]]_1 = & -2\delta_{p_1+p_4, p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4]\left\{\varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + \frac{1}{2}(J_{p_1-p_2} + J_{p_1-p_3})\right\} \\ & + \underbrace{\delta_{p_1, p_3}\delta_{p_2, p_4}\mathbf{g}[0^-] - 2\delta_{p_1+p_4, p_2+p_3}\mathbf{g}[p_4]}_{\text{underbraces}}. \end{aligned} \quad (\text{B42})$$

Comparing Eq. (B30) and Eq. (B42), we see that other than the term with underbraces, these vertices satisfy Eq. (49) or Eq. (72).

### g. $\Phi$ to $O(\lambda)$

We now assemble terms:

$$[\Phi(k)]_1 = \sum_p \left( \varepsilon_p + \frac{1}{2}\varepsilon_k + \frac{1}{2}J_{k-p} \right) \mathbf{g}[p][\Lambda^{(a)}(p, k)]_1 + \sum_{pq} \frac{1}{2}\varepsilon_{q+p-k}\mathbf{g}[p][\Lambda^{(a)}(p, k; q+p-k, q)]_1. \quad (\text{B43})$$

Let us rewrite this as ( $k \rightarrow p_2, p \rightarrow p_1, q \rightarrow p_4$ )

$$\begin{aligned} [\Phi(p_2)]_1 = & \sum_{p_1} \left( \varepsilon_{p_1} + \frac{1}{2}\varepsilon_{p_2} + \frac{1}{2}J_{p_1-p_2} \right) \mathbf{g}[p_1][\Lambda^{(a)}(p_1, p_2)]_1 + \sum_{p_1+p_4=p_2+p_3} \frac{1}{2}\varepsilon_{p_3}\mathbf{g}[p_1][\Lambda^{(a)}(p_1, p_2; p_3, p_4)]_1 \\ = & \underbrace{\frac{n}{4} \sum_{p_3} \varepsilon_{p_3}\mathbf{g}[p_3] - \sum_{p_1, p_4} \varepsilon_{p_1+p_4-p_2}\mathbf{g}[p_1]\mathbf{g}[p_4]}_{\text{underbraces}} \\ & - 2 \sum_{p_1+p_4=p_2+p_3} \mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4] \left( \varepsilon_{p_1} + \frac{1}{2}\varepsilon_{p_2} + \frac{1}{2}J_{p_1-p_2} \right) \left\{ \varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + \frac{1}{2}(J_{p_1-p_2} + J_{p_1-p_3}) \right\} \\ & - \sum_{p_1+p_4=p_2+p_3} \mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4]\varepsilon_{p_3} \left\{ \varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + \frac{1}{2}(J_{p_1-p_2} + J_{p_1-p_3}) \right\}. \end{aligned} \quad (\text{B44})$$

The first line with underbraces arises from the term  $\Lambda$  in Eq. (B39), or Eq. (B35) and Eq. (B42) which disobey the relation Eq. (49) or Eq. (72). It gives a static but momentum-dependent contribution, and we will drop it as discussed below Eq. (49) and in the second remark below Eq. (68). The rest are combined and rearranged to give

$$\begin{aligned} [\Phi(p_2)]_1 = & - \sum_{p_1+p_4=p_2+p_3} \mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4] \left( \varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + J_{p_1-p_2} \right) \left\{ \varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} + \varepsilon_{p_4} + \frac{1}{2}(J_{p_1-p_2} + J_{p_1-p_3}) \right\} \\ = & \frac{1}{2} \sum_{p_1, p_3, p_4} \mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4]W(p_2, p_3; p_4, p_1)[W(p_2, p_3; p_4, p_1) + W(p_2, p_3; p_1, p_4)], \end{aligned} \quad (\text{B45})$$

where in the first line we symmetrized further in  $p_1 \leftrightarrow p_4$ .

We can bring this into standard notation by sending  $p_2 \rightarrow k, p_1 \rightarrow p, p_3 \rightarrow q, p_4 \rightarrow k+q-p$ :

$$\begin{aligned} [\Phi(k)]_1 = & - \sum_{q, p} \mathbf{g}[q]\mathbf{g}[p]\mathbf{g}[k+q-p](\varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + J_{k-p}) \left\{ \varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + \frac{1}{2}(J_{k-p} + J_{p-q}) \right\}, \\ [\Phi(k)]_1 = & \frac{1}{2} \sum_{q, p} \mathbf{g}[q]\mathbf{g}[p]\mathbf{g}[k+q-p]W(k, q; q+k-p, p)[W(k, q; q+k-p, p) + W(k, q; p, q+k-p)]. \end{aligned} \quad (\text{B46})$$

### h. Stepping up and final Green's function to $O(\lambda^2)$

We are now in a position to put together the second-order result for  $\mathbf{g}^{-1}$  and also  $\mu$ . Recall that  $[\mathbf{g}^{-1}[k]]_2 = -[Y_1[k] + \Phi[k]]_1$ , where these variables are calculated in Eq. (B13) and Eq. (B46). Hence we can now compile the equations of the second-order theory with sources turned off:

$$\begin{aligned} \mathbf{g}^{-1}(k) = & \mathbf{g}_0^{-1}(k) + \lambda[\mathbf{g}^{-1}(k)]_1 + \lambda^2[\mathbf{g}^{-1}(k)]_2 + O(\lambda^3), \quad [\mathbf{g}^{-1}[k]]_0 = i\omega_n + \boldsymbol{\mu} - \varepsilon_k - \frac{1}{4}J_0, \\ [\mathbf{g}^{-1}[k]]_1 = & n\varepsilon_k - \frac{n}{2}u_0 + \frac{1}{2} \sum_q J_{k-q}\mathbf{g}[q] + \left( \frac{1}{4}J_0n + 2 \sum_q \varepsilon_q\mathbf{g}[q] \right), \\ [\mathbf{g}^{-1}[k]]_2 = & -\frac{3n^2}{8}\varepsilon_k + \frac{n^2}{8}u_0 - [\Phi(k)]_1 - \left( \frac{n^2}{8}J_0 + \frac{n}{4} \sum_q \varepsilon_q\mathbf{g}[q] \right). \end{aligned} \quad (\text{B47})$$

We thus see that all the computed  $[\mathbf{g}^{-1}]_j$  are invariant under the two shift theorems. Adding up terms to  $O(\lambda^2)$ ,

$$\begin{aligned} \mathbf{g}^{-1}[k] &= i\omega_n + \boldsymbol{\mu}' - \left(1 - \lambda n + \lambda^2 \frac{3n^2}{8}\right) \boldsymbol{\varepsilon}_k \\ &+ \lambda \sum_q \frac{1}{2} J_{k-q} \mathbf{g}[q] - \lambda^2 [\Phi(k)]_1 + O(\lambda^3), \end{aligned} \quad (\text{B48})$$

$$\begin{aligned} \boldsymbol{\mu}' &= \boldsymbol{\mu} - u_0 \frac{\lambda n}{2} \left(1 - \frac{\lambda n}{4}\right) + \left[ J_0 \frac{\lambda n}{4} \left(1 - \frac{\lambda n}{2}\right) \right. \\ &\left. + 2\lambda \left(1 - \frac{\lambda n}{8}\right) \sum_q \varepsilon_q \mathbf{g}[q] \right], \end{aligned} \quad (\text{B49})$$

with  $[\Phi(k)]_1$  defined in Eq. (B46) and a shifted chemical potential  $\boldsymbol{\mu}'$ . Note that both terms in square brackets in Eq. (B49) are independent of frequency and wave vector; the first ( $T$  independent) term may be safely ignored since it vanishes when we finally set  $J_0 \rightarrow 0$ , while the second term involving  $\sum_q \varepsilon_q \mathbf{g}[q]$  is expected to be weakly  $T$  dependent.

Similarly the caparison factor  $\mu$  is found to  $O(\lambda^2)$  as

$$\begin{aligned} \mu[k] &= 1 + \lambda [\mu[k]]_1 + \lambda^2 [\mu[k]]_2 + O(\lambda^3), \\ [\mu[k]]_1 &= -\frac{n}{2}, \quad [\mu[k]]_2 = \frac{n^2}{4} + [\Psi[k]]_1. \end{aligned} \quad (\text{B50})$$

Adding up terms to  $O(\lambda^2)$  we obtain

$$\mu[k] = 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4} + \lambda^2 [\Psi(k)]_1 + O(\lambda^3), \quad (\text{B51})$$

along with the definition in Eq. (B19).

<sup>1</sup>D. Hansen and B. S. Shastry, [arXiv:1211.0594](https://arxiv.org/abs/1211.0594) [cond-mat.str-el].

<sup>2</sup>B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011); **108**, 029702 (2012).

<sup>3</sup>B. Edegger, V. N. Muthukumar, and C. Gros, *Adv. Phys.* **56**, 927 (2007).

<sup>4</sup>P. W. Anderson, *Science* **235**, 1196 (1987).

<sup>5</sup>A. B. Harris and R. V. Lange, *Phys. Rev.* **157**, 295 (1967).

<sup>6</sup>F. C. Zhang and T. M. Rice, *Phys. Rev. B* **37**, 3759 (1988).

<sup>7</sup>M. Gutzwiller, *Phys. Rev. Lett.* **10**, 159 (1963).

<sup>8</sup>B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011).

<sup>9</sup>G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).

<sup>10</sup>B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).

<sup>11</sup>B. S. Shastry, *Phys. Rev. B* **81**, 045121 (2010).

<sup>12</sup>A simple example  $\mathcal{G} \sim c_0/(i\omega_n + \mu - \varepsilon_k)$  can be worked out in detail and illustrates this remark.

<sup>13</sup>A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

<sup>14</sup>G. D. Mahan, *Many-Particle Physics*, 3rd ed. (Kluwer-Plenum, 2000).

<sup>15</sup>The magnitude of the relative coefficient 4 in the condition for cancellation  $u_J = -4u_t$  is also consistent with a second procedure. In the latter, the shift of  $t_{ij}$  and  $J_{ij}$  is carried out *after taking the commutator of  $X_{i,\sigma}^{0\sigma}$  with  $H$* , e.g., in the expression Eq. (34), and the extra term in  $A_{i,\sigma_i}$  generated by this process is required to be zero. In general these two procedures can produce different coefficients, as in the minimal theory where one does not symmetrize the expressions. In such cases the coefficient is determined by the one appearing in the equation after taking the commutator, as in Eq. (34), since it propagates down the hierarchy of equations of motion. We report elsewhere the *minimal theory*, where the condition for cancellation is  $u_J = -2u_t$ , with a relative coefficient 2.

<sup>16</sup>Since finite- $T$  many-body theory formalism is of an isothermal rather than adiabatic character, where quantum numbers are unconstrained and consequently the Fermi surface changes its shape, the conventional usage of the term adiabatic continuity seems misplaced. It might be more appropriately replaced by the term “parametric continuity” or just “continuity.”

<sup>17</sup>J. M. Luttinger and J. C. Ward, *Phys. Rev.* **118**, 1417 (1960).

<sup>18</sup>W. Brinkman and T. M. Rice, *Phys. Rev. B* **2**, 4302 (1970).

<sup>19</sup>G. Baskaran, Z. Zhou, and P. W. Anderson, *Solid State Commun.* **63**, 973 (1987).

<sup>20</sup>G. Kotliar, *Phys. Rev. B* **37**, 3664 (1988).

# Extremely correlated Fermi liquids: Self-consistent solution of the second-order theory

Daniel Hansen and B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*

(Received 5 November 2012; revised manuscript received 10 May 2013; published 6 June 2013)

We present detailed results from a recent microscopic theory of extremely correlated Fermi liquids, applied to the  $t$ - $J$  model in two dimensions, developed recently by Shastry [Phys. Rev. Lett. **107**, 056403 (2011); Phys. Rev. B **87**, 125124 (2013)]. The second-order theory in the parameter  $\lambda$ , related to the density, is argued to be quantitatively valid in the overdoped regime for  $0 \leq n \lesssim 0.75$ , with  $n$  denoting the particle density. The calculation involves the self-consistent solution of equations for an auxiliary Fermi liquid Green's function and an adaptive spectral weight. We present numerical results at low as well as high  $T$ , at various low to intermediate densities in the normal phase, using a minimal set of band parameters relevant to the cuprate superconductors. We display the momentum space occupation function  $m_k$ , energy dispersion curves locating the peaks of spectral functions, the optical conductivity, relaxation rates for quasiparticles, and the electronic spectral functions on an *absolute scale*. The line shapes have an asymmetric shape and a broad background that is also seen in experiments, and our calculations validate approximate recent versions of the theory. The results also display the experimentally noted high-energy kink and provide an in-depth understanding of its origin and dependence on band parameters.

DOI: [10.1103/PhysRevB.87.245101](https://doi.org/10.1103/PhysRevB.87.245101)

PACS number(s): 71.10.-w, 71.10.Fd

## I. INTRODUCTION

The  $t$ - $J$  model describes the physics of very strongly interacting electrons, made especially difficult by the requirement of (at most) single occupancy of the lattice sites. It is the subject of many recent works in the context of the cuprate superconductors, and also other correlated systems such as sodium cobaltates. This problem is very hard since it precludes the application of standard perturbative methods. This conundrum has motivated a new strong-coupling approach, resulting in the theory of *extremely correlated Fermi liquids* (ECFLs).<sup>1,2</sup> Previous applications of the methodology of Ref. 1 to the cuprates has given encouraging results. These include spectral functions that compare very well with the experimental angle-resolved photoemission spectroscopy (ARPES) data,<sup>3-6</sup> providing natural explanations of the “high-energy kink,” and also the more subtle “low-energy kink” seen in experiments. The theory also has led to interesting predictions for the asymmetry of line shapes.<sup>6</sup>

The formalism initiated in Ref. 1 charted out an approach to the problem of the  $t$ - $J$  model using basic insights from Schwinger's powerful approach to field theory, using source fields to write down exact functional differential equations for the propagator. In the next crucial step, it was recognized that complexity arising from the noncanonical nature of the (projected) electrons can be circumvented by a product *ansatz*. This involves decomposing the propagator as the space time convolution of a *canonical electron propagator*, and an adaptive spectral weight factor termed the *caparison factor* satisfying coupled equations of motion. A recent work<sup>2</sup> develops this idea in a systematic fashion, emphasizing the role of expanding in a parameter  $\lambda$  ( $0 \leq \lambda \leq 1$ ), related to the particle density, or more closely to  $\lambda \sim (1 - \frac{4}{n^2}d)$ , where  $d$  is the double occupancy ( $0 \leq d \leq \frac{n^2}{4}$ ). It further explores the implications of a novel set of identities for the  $t$ - $J$  model, termed the *shift identities*. These simple but crucial identities provide an important constraint on the  $\lambda$  expansion. A method for generating a systematic set of equations for the propagator

to any orders in  $\lambda$  is given, along with explicit equations to second order in  $\lambda$  that manifestly obey the shift identity constraints. We will refer to this theory as (I) here and prefix equations of that paper with (I). A detailed numerical solution of this  $O(\lambda^2)$  ECFL propagator is the main focus of this work. We obtained and benchmark the results of these equations against known results, and thereby provide a solid platform for further developments of the method, as well as a validation of the phenomenological versions of ECFL. With the confidence gained by the benchmarking, we further study and report the hopping parameter sensitivity of the kink effect.

Broadly speaking, the  $O(\lambda^m)$  equations resemble the fully self-consistent  $m$ th-order skeleton diagram expansion of the standard Feynman-diagram-based theory, as described in standard texts,<sup>7-9</sup> but generalize to the case of extreme correlations. Summarizing the arguments in Refs. 1 and 2, a low-order theory in  $\lambda$  is already expected to capture features of extreme correlations. This perhaps initially surprising expectation arises in view of the non-Dysonian representation of the Green's function in terms of two self-energies  $\Phi$  and  $\Psi$ , within the ECFL formalism. The self-energy  $\Psi$  resides in the numerator of the Green's function, as in Eqs. (1) and (2). It plays the role of an adaptive spectral weight that balances the somewhat opposing requirements of the “high-energy” weight  $1 - \frac{n}{2}$  and the low-energy Luttinger theorem. The latter requires a greater magnitude of the numerator than  $1 - \frac{n}{2}$  to accommodate the particles into a Fermi surface (FS) with the same volume as in the Fermi gas. A further tactical advantage of this method is due to the finite range of variation of  $\lambda$ , namely,  $0 \leq \lambda \leq 1$ , that suffices to interpolate between the Fermi gas and the extreme correlation limit. This is in contrast to controlling the double occupancy  $d$  using a repulsive energy  $U$ , with its range of values  $0 \leq U \leq \infty$ . Experience shows that  $U$  must be tuned to a very large value  $U \gg |t|$  in order to achieve the same end, thereby invalidating low-order expansions in  $U$ . In summary, within the present formalism, a low-order theory in  $\lambda$  seems well worth examining in detail; this is our task here.

We note that apart from a few exact solutions in one dimension and some calculations for finite-sized systems (see below), we are aware of no systematic analytical calculations in higher dimensions, for the dynamics of the physically relevant spin- $\frac{1}{2}$  version of the  $t$ - $J$  model, working directly in the thermodynamic limit. An earlier body of work in Ref. 10 shares some of the objectives and features of our approach but is technically very different. It relies on an expansion in the inverse number of components  $\frac{1}{\mathcal{N}}$  and is thus somewhat removed from the physical case of interest, where  $\mathcal{N} = 2$ . Therefore while the importance of the  $t$ - $J$  model was understood many years ago, there has been little detailed comparison with the ARPES experiments until recently.<sup>3,4</sup> This gap is one of the main motivations for this (and our related) work. In this paper, we present a controlled calculation for the spectral functions of the  $t$ - $J$  model by solving the above  $O(\lambda^2)$  equations. We evaluate thermodynamical variables, the spectral functions, ARPES line shapes, and optical conductivity of the  $t$ - $J$  model. The ECFL formalism and the  $\lambda$  expansion method provides an in-built criterion to judge the validity of the expansion at any order. Using this criterion we argue that our present  $O(\lambda^2)$  calculations are valid in the high-hole-doping limit, known as the overdoped regime. Clearly this corresponds to low and intermediate *electron density*, since the hole doping is related to the particle density as  $x = 1 - n$ . Future work will be aimed at higher-order calculations in  $\lambda$  in order to enable us to address densities closer to optimal doping ( $n \sim 0.85$ ). The results are compared with other approximations as well as a few experiments. Needless to say, even in such an overdoped regime, experimental evidence points to the important role of strong correlations.<sup>11,12</sup>

While analytical methods beyond crude mean-field theories have been in short supply, there is a valuable body of numerical results for the  $t$ - $J$  model from exact diagonalization,<sup>13</sup> high-temperature series expansions,<sup>14</sup> variational wave functions,<sup>15-17</sup> and finite temperature Lanczos methods.<sup>18-21</sup> Noteworthy are the results of Ref. 19 from Prelovsek and co-workers, who handle the series expansion in inverse temperature in a stochastic fashion, thereby obtaining results down to fairly low temperatures. Owing to finite size effects and the inherent nature of the high- $T$  expansion, the results from this theory, although broadly comparable to ours, seem more grainy.

The Hubbard model for large on-site coupling  $U$  tends to the  $t$ - $J$  model [apart from  $O(t^2/U)$  correction terms], so the large  $U$  studies of this model are of interest. Quantum Monte Carlo methods, despite the difficulties associated with the sign problem, yield some valuable insights into the spectral features such as kinks.<sup>22</sup> We note that the dynamical mean-field theory (DMFT) for the Hubbard model<sup>23,24</sup> gives a numerically exact solution in high enough dimensions of the Hubbard model. Although the strong coupling (i.e.,  $U > W$ ) relevant to the  $t$ - $J$  model results is challenging, there is impressive progress overall. A recent DMFT study<sup>25</sup> at strong coupling obtains detailed spectral functions that are roughly comparable to what we find here for the  $t$ - $J$  model.

The ECFL formalism has several advantages, since it is essentially an analytical method with a computational aspect that is lightweight, in comparison with other methods listed above. The only present limitation is the density attainable with

the second-order theory. When possible, we present absolute scale results that are encouragingly close to experimental data with no other fitted parameters.

We finally note that the present  $O(\lambda^2)$  results for the location of the energy peaks has been recently tested in Ref. 26, against an independent theory with overlapping validity. Reference 26 studied the infinite-coupling Hubbard model in two dimensions by using a highly efficient computer program to generate a series expansion in hopping of the exact Green's function and its various moments to high order. The locations of the dispersion peaks can be estimated from these. These dispersion relations match *quantitatively* the ones found from the present theory, with  $J \rightarrow 0$  for the densities quoted in this paper. This suggests a high degree of reliability of the spectral functions discussed herein.

The plan of the paper is as follows: In Sec. II, we present a summary of the equations solved here from Sec. I. In Sec. III, we discuss the computational strategy and explain the scheme, using the fast Fourier transform method (FFT), so that the spectral functions can be computed efficiently. Section IV presents the detailed results of the calculation. Section V contains a summary and concluding comments. The Supplemental Material in Ref. 27 details the results for thermodynamics and the wave-function renormalization  $Z_k$ , and also gives further details of the computational method employed.

## II. SUMMARY OF THE $O(\lambda^2)$ THEORY

In the ECFL formalism developed in (I) (i.e. Ref. 2), the physical Green's function  $\mathcal{G}$  can be factored in the momentum space as

$$\mathcal{G}(k) = \mathbf{g}(k) \mu(k), \quad \text{where } (k) \equiv (\vec{k}, i\omega_k). \quad (1)$$

Here the caparison factor  $\mu(k)$  plays the role of an adaptive spectral weight, while  $\mathbf{g}(k)$  is the auxiliary canonical Fermion propagator. These objects are expanded in powers of a parameter  $\lambda$ , relating to density, and finally we set  $\lambda \rightarrow 1$ . As shown in Eqs. (I-83), (I-84), and (I-85), the second-order equations for the ECFL Green's function are as follows:

$$\mu(k) = 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4} + \lambda^2 \Psi(k), \quad (2)$$

$$\begin{aligned} \Psi(k) = & - \sum_{p,q} (\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p} - u_0) \\ & \times \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(q+k-p), \end{aligned} \quad (3)$$

$$\mathbf{g}^{-1}(k) = i\omega_n + \boldsymbol{\mu}' - \bar{\varepsilon}_k - \lambda^2 \bar{\Phi}(k), \quad (4)$$

$$\bar{\varepsilon}_k = \left(1 - \lambda n + \lambda^2 \frac{3n^2}{8}\right) \varepsilon_k + \lambda \sum_q \frac{1}{2} J_{k-q} \mathbf{g}(q), \quad (5)$$

$$\begin{aligned} \bar{\Phi}(k) = & - \sum_{q,p} \mathbf{g}(q) \mathbf{g}(p) \mathbf{g}(k+q-p) \\ & \times (\varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + J_{k-p} - u_0) \\ & \times \left\{ \varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + \frac{1}{2}(J_{k-p} + J_{p-q}) - u_0 \right\}, \end{aligned} \quad (6)$$

where  $\sum_k \equiv \frac{1}{\beta N_s} \sum_{\vec{k}, \omega_n}$ , with  $N_s$  being the number of lattice sites and  $\beta$  is inverse temperature. These expressions for the Green's function satisfy the "shift invariances" described in

Ref. 2, i.e., any uniform shift in  $\varepsilon_k$  or  $J_k$  can be absorbed in  $\mu'$  and  $u_0$  such that the spectral function is invariant. These second-order equations are the lowest-order ones where nontrivial frequency dependence arises and are the focus of this work. Below we discuss in detail the criterion for the quantitative validity of the present second-order expansion.

As written here,  $\mu(k)$  and  $\mathbf{g}(k)$  have acquired a variety of static terms as well as frequency-dependent terms called  $\Psi$  and  $\bar{\Phi}$ , respectively. This is written with a slight change of notation  $[\Phi(k)]_1 \rightarrow \bar{\Phi}(k)$  from (I-85), and we have introduced the effective band energy  $\bar{\varepsilon}_k$  in Eq. (5) that gets a static contribution from shrinking of the bare energies  $\varepsilon_k$ , as well as from the exchange energy  $J$ . The role of the parameter  $u_0$  as a second chemical potential is described below. All terms are understood to be correct up to  $O(\lambda^2)$ , and hence possess corrections of  $O(\lambda^3)$  that are ignored here.

The number of the physical electrons is fixed by the number sum rule:

$$\frac{n}{2} = \sum_k \mathcal{G}(k) e^{i\omega_n 0^+}. \quad (7)$$

In order for  $\mathcal{G}$  to satisfy the Luttinger volume theorem, the auxiliary Fermions described by  $\mathbf{g}$  must be equal in number and therefore satisfy a *second* sum rule:

$$\frac{n}{2} = \sum_k \mathbf{g}(k) e^{i\omega_n 0^+}. \quad (8)$$

In contrast to canonical theories, here we have two independent sum-rule constraints requiring two Lagrange multipliers. The first Lagrange multiplier  $\mu'$  is a standard chemical potential in that it sits next to the band energies  $\bar{\varepsilon}_k$  in the denominator of  $\mathbf{g}$ . A second Lagrange multiplier  $u_0$  arises naturally in the ECFL formalism, thanks to the role of the shift identities, as shown in (I). The  $u_0$  term has a role similar to that of the Hubbard  $U$  in the effective Hamiltonian in (I). It controls the broadening of the spectral function through the magnitude of  $\Phi$  and  $\Psi$ . Neither of these Lagrange multipliers is the physical thermodynamic chemical potential of the grand canonical ensemble. The physical chemical potential  $\mu_{\text{phys}}$ , denoted by  $\mu$ , can be obtained as a function of  $\mu'$  and  $u_0$ , as shown in Eq. (179) of (I):

$$\begin{aligned} \mu &= \mu' + u_0 \frac{\lambda n}{2} \left(1 - \frac{\lambda n}{4}\right) \\ &\quad - \left[ J_0 \frac{\lambda n}{4} \left(1 - \frac{\lambda n}{2}\right) + 2\lambda \left(1 - \frac{\lambda n}{8}\right) \sum_q \varepsilon_q \mathbf{g}(q) \right] \\ &\quad + O(\lambda^3). \end{aligned} \quad (9)$$

We now discuss the criterion for validity of equations to a second order in  $\lambda$ . As stated above, dropping terms of  $O(\lambda^3)$  in Eqs. (2)–(6) limits the regime of validity of these calculation to densities not too close to unity. To see this, note from Eq. (2) that this theory would give a high-frequency

behavior of  $\mathcal{G} \sim \frac{c_0}{i\omega}$  with  $c_0 = 1 - \frac{n}{2} + \frac{n^2}{4}$ , rather than the exact value  $c_0 = 1 - \frac{n}{2}$ , thus introducing an error. This slight error in the high-frequency physics is a result of keeping a few terms in the expansion in  $\lambda$ . Note, however, that the low-frequency physics encoded by the Luttinger-Ward sum rule is untouched by this and is exactly obeyed to each order in  $\lambda$ . Thus at  $n \sim 0.78$  we have an error of  $\frac{n^2}{4-2n} \sim 25\%$  in the high-frequency spectral weight in this theory, a value somewhat beyond where we can push this approximation. The  $O(\lambda^3)$  terms are expected to extend the range of this approximation to higher particle densities.

### III. COMPUTATION OF SPECTRAL FUNCTIONS

#### A. Definitions

Computationally, it is expedient to employ a spectral function notation as described for example in Ref. 9. The Matsubara frequency object  $\mathcal{G}(k, i\omega_n)$  is analytically continued to the real axis and we define as follows:

$$\rho_{\mathcal{G}}(k, \omega) = -\frac{1}{\pi} \text{Im} [\mathcal{G}(k, i\omega_n \rightarrow \omega + i0^+)]. \quad (10)$$

This object is the spectral function, denoted in most experimental literature by  $A(k, \omega)$ . The real part of the analytically continued function can be obtained by a Hilbert transform

$$\text{Re } \mathcal{G}(k, \omega) = \text{P.V.} \int_{-\infty}^{\infty} \frac{\rho_{\mathcal{G}}(k, \nu)}{\omega - \nu} d\nu. \quad (11)$$

An analogous definition is given for spectral representation  $\rho_{\mathbf{g}}(k, \nu)$ ,  $\rho_{\bar{\Phi}}(k, \nu)$ ,  $\rho_{\Psi}(k, \nu)$  used for  $\mathbf{g}$ ,  $\bar{\Phi}$ ,  $\Psi$ , etc., and hence, the full set of equations above can be rewritten in terms of these spectral functions. Since  $\mathcal{G}$  is a product as in Eq. (1), we note that within the  $O(\lambda^2)$  theory

$$\begin{aligned} \rho_{\mathcal{G}}(k, \omega) &= \rho_{\mathbf{g}}(k, \omega) \left(1 - \frac{n}{2} + \frac{n^2}{4} + \text{Re } \Psi(k, \omega)\right) \\ &\quad + \rho_{\Psi}(k, \omega) \text{Re } \mathbf{g}(k, \omega), \end{aligned} \quad (12)$$

so the two sum rules Eq. (7) and Eq. (8) can be written as

$$\begin{aligned} \frac{n}{2} &= \sum_k \int d\omega \rho_{\mathbf{g}}(k, \omega) f(\omega), \\ \frac{n^2}{4} \left(1 - \frac{n}{2}\right) &= - \sum_k \int d\omega f(\omega) (\rho_{\mathbf{g}}(k, \omega) \text{Re } \Psi(k, \omega) \\ &\quad + \text{Re } \mathbf{g}(k, \omega) \rho_{\Psi}(k, \omega)), \end{aligned} \quad (13)$$

where  $f(\omega) = [1 + \exp(\beta\omega)]^{-1}$  and  $\bar{f}(\omega) = 1 - f(\omega)$ . The auxiliary spectral function is in the usual Dysonian form,

$$\rho_{\mathbf{g}}(k, \omega) = \frac{\rho_{\bar{\Phi}}(k, \omega)}{\{\omega + \mu' - \bar{\varepsilon}_k - \text{Re } \bar{\Phi}(k, \omega)\}^2 + (\pi \rho_{\bar{\Phi}})^2}. \quad (14)$$

Using Eqs. (1)–(6), we express the spectral functions for  $\Psi$  and  $\bar{\Phi}$  as

$$\begin{aligned} \rho_{\bar{\Phi}}(k, \omega) &= \frac{1}{N_s^2} \sum_{pq} \int d\nu_1 d\nu_2 \rho_{\mathbf{g}}(p, \nu_1) \rho_{\mathbf{g}}(q, \nu_2) \rho_{\mathbf{g}}(p+q-k, \nu_1 + \nu_2 - \omega) \\ &\quad \times \{f(\nu_1) f(\nu_2) \bar{f}(\nu_1 + \nu_2 - \omega) + \bar{f}(\nu_1) \bar{f}(\nu_2) f(\nu_1 + \nu_2 - \omega)\} \\ &\quad \times (\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p} - u_0) \left\{ \varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + \frac{1}{2}(J_{k-p} + J_{k-q}) - u_0 \right\}, \end{aligned} \quad (15)$$

$$\begin{aligned}
\rho_{\Psi}(k, \omega) &= \frac{1}{N_s^2} \sum_{pq} \int dv_1 dv_2 \rho_{\mathbf{g}}(p, v_1) \rho_{\mathbf{g}}(q, v_2) \rho_{\mathbf{g}}(p+q-k, v_1+v_2-\omega) \\
&\quad \times \{f(v_1) f(v_2) \bar{f}(v_1+v_2-\omega) + \bar{f}(v_1) \bar{f}(v_2) f(v_1+v_2-\omega)\} \\
&\quad \times (\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p} - u_0).
\end{aligned} \tag{16}$$

These frequency integrals are solved by discretizing frequency over a finite window that is wide enough to capture the finite support of the spectral functions. In Ref. 27 we outline how this is accomplished efficiently with FFTs and implemented in an iterative process.

## IV. RESULTS

### A. Physical variables

The computational program has several parameters that can be varied. These include the tight-binding band structure (through hopping parameters  $t$ ,  $t'$  etc.), the spin coupling  $J$ , density, and temperature. For the parameters of the model, we focus on a minimal model with the nearest-neighbor hopping  $t \sim 3000$  K and  $J \sim 900$  K, and all longer-range hopping parameters are zero. These values are chosen to match the bandwidth of the cuprates. However, at the bare level, this produces an electronlike Fermi surface near half filling, remaining closed around the  $\Gamma = (0,0)$  point in the Brillouin zone (BZ). This is in contrast to the ARPES reconstructed FS of, say, *BISSCO* displaying a holelike surface. Nonetheless, this minimal parameter set exhibits a variety of features in common with the cuprates, most notably, a broad incoherent spectrum at high negative frequency. Interestingly, we find that the distribution of incoherent weight at high frequencies is very sensitive to the bare hopping parameters. For this reason, when we look the high-energy features, we will explore their dependence in the second neighbor hopping parameter  $t'$ , also including a fine-tuned tight-binding fit of *BISSCO* from Ref. 28.

### B. Other parameters in the programs

The program can be implemented on lattices of various size and spatial dimension. For a given choice of these parameters an appropriate choice must be made for computational grid. This includes the lattice size as well as the discretized frequency grid. We look at converged spectral functions for a wide variety of these parameters.

The majority of the following results were performed on a square lattice with dimension  $L \times L$  with  $L = 36$ , and periodic boundary conditions are imposed. We therefore work in a momentum representation with an  $L \times L$ -sized  $k$  grid of points  $k_{i,j} = \frac{\pi}{aL}(i,j)$ , where  $1 \leq i, j \leq L$  and the lattice parameter is  $a = 3.82$  Å. The spectral functions have compact support, extending to  $|\omega| \lesssim 8 \times t$ . We choose a frequency range  $-\frac{1}{2}\omega_c \leq \omega \leq \frac{1}{2}\omega_c$ , with  $\omega_c = 30 \times t$ , a range that is sufficient to capture the full range of the spectral functions. We discretize this frequency range in  $N_{\omega} = 3000$  bins each of width  $\Delta\omega = \frac{\omega_c}{N_{\omega}} = 0.01t = 30$  K.  $\Delta\omega$  is the lowest resolvable frequency scale in the calculation, so it is prudent to disallow

any spectral features from becoming any sharper than this scale. Therefore we introduce the convergence factor  $\eta_{\min} = \Delta\omega$ . It serves as a lower limit on the width of spectral features. Thus in the Dysonian form of  $\rho_{\mathbf{g}}$  [Eq. (14)] we set  $\rho_{\Phi} \rightarrow \rho_{\Phi} + \frac{\eta}{\pi}$ .

### C. Frequency independent variables

We now proceed to study the FS in this theory, starting with the momentum occupation function  $m_k$  of the Gutzwiller projected fermions:

$$m_k \equiv \langle \hat{C}_{k\sigma}^{\dagger} \hat{C}_{k\sigma} \rangle = \int_{-\infty}^{\infty} \rho_{\mathcal{G}}(\vec{k}, \omega) f(\omega) d\omega. \tag{17}$$

A sharp drop in this function helps to locate the FS at low  $T$ . This can be compared with the Luttinger-Ward surface, defined by a sign change in  $\text{Re } \mathcal{G}(k, 0)$ , also given in terms of the spectral function by

$$\text{Re } \mathcal{G}(\vec{k}, 0) = P.V. \int_{-\infty}^{\infty} \frac{\rho_{\mathcal{G}}(\vec{k}, \omega) d\omega}{\omega}. \tag{18}$$

At  $T = 0$  the FS in  $\vec{k}$  space is traced out by  $\text{Re } \mathcal{G}^{-1}(\vec{k}, 0) = 0$ , as dictated by the Luttinger-Ward sum rule. The momentum distribution  $m_k$  is plotted in Fig. 1 at  $T = 130$  K and  $T = 605$  K for various densities along three principle directions of the BZ. The Luttinger-Ward zero crossings  $\text{Re } \mathcal{G}^{-1}(\vec{k}, 0) = 0$  are depicted by dashed vertical lines. There is a close correspondence between these crossings and the point where  $m_k = 0.5$ , similar to that noted previously by Stephan and Horsch<sup>13</sup> in an exact diagonalization study. Since this correspondence is not on any rigorously firm basis, it is difficult to do more than to list the conditions for its approximate validity. Using high-temperature expansions for the  $t$ - $J$  model, Singh and Glenister<sup>14</sup> found the FS to be that of the Fermi gas by various criteria, and noted that the condition  $m_{k_F} \sim 0.5$  is only satisfied approximately at high  $T$ . At higher temperature where the quasi-particle (QP) near the FS have been significantly broadened, we find that the condition  $m_{k_F} \sim 0.5$  is still reliable, in agreement with Ref. 13.

In Fig. 1, a point of considerable interest is the spillover of the occupation to the regions in  $k$  space that are unoccupied in the Fermi gas, as noted in various variational wave-function studies of the  $t$ - $J$  model already.<sup>15-17</sup> From Eq. (17) we note that the magnitude of  $m_k$  for momenta  $k > k_F$  provides an estimate of the spectral weight  $\rho_{\mathcal{G}}(k, \omega)$  at occupied energies at low  $T$ . In early analyses of ARPES data, the significance of this piece of information was not always realized, and often substantial spectral weight was discarded as belonging to some unspecified background. Only recent studies such as Ref. 3 have taken note of the significance of the background.

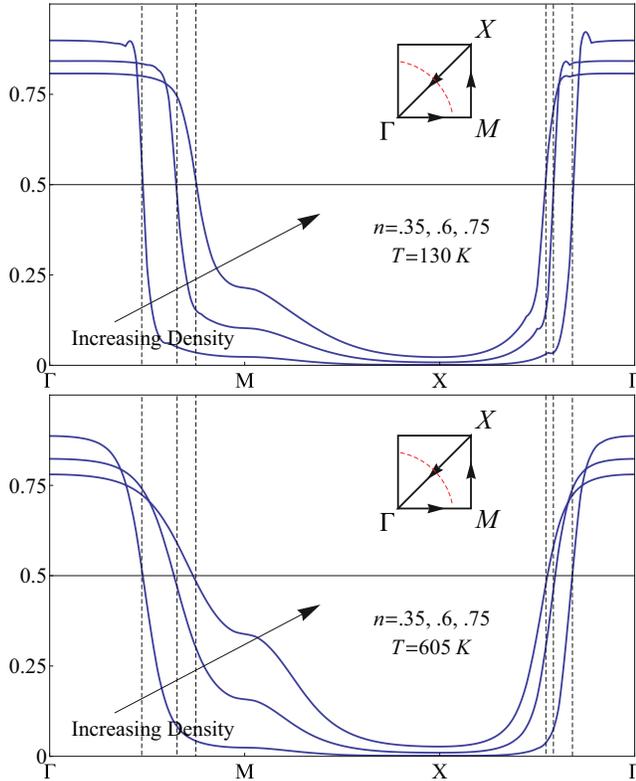


FIG. 1. (Color online) The momentum distribution function  $m_k$  is plotted along three principle lines of the BZ. The left and right figures are at 130 and 605 K, respectively. In each case the FS is the same as in the noninteracting problem. The Luttinger-Ward crossing  $\text{Re } \mathcal{G}^{-1}(\vec{k}, 0) = 0$  is indicated for each density by the vertical dashed lines. For each density and each temperature the Luttinger-Ward crossings correspond well with the condition  $m_k = \frac{1}{2}$ .

#### D. Various excitation energies

The spectra obtained here contain sharp peaks as well as substantial incoherent background due to extreme correlations. The QP weight  $Z_k$  is discussed in the Supplemental Material.<sup>27</sup> To understand the effect of the many-body renormalizations, it is fruitful to study three dispersion relations defined in Ref. 4:

$$\begin{aligned} \bar{\epsilon}_k &= \left(1 - n + \frac{3n^2}{8}\right) \epsilon_k + \frac{1}{2} \sum_q J_{k-q} m_q, \\ E_k &= \bar{\epsilon}_k - \mu' + \text{Re } \bar{\Phi}(k, E_k), \\ E_k^* &= \max[\rho_{\mathcal{G}}(k, \omega) : \omega]. \end{aligned} \quad (19)$$

Here  $\bar{\epsilon}_k$  defines the bare energy times its static renormalization, while  $E_k$  locates the vanishing point for the real part of the auxiliary Green's function  $\mathbf{g}$ , thereby defining the Luttinger-Ward surface through a change of sign.  $E_k^*$  locates the highest peak of the physical Green's function  $\mathcal{G}$ , and hence defines QP excitations, provided they are sufficiently sharp. ARPES experiments performed with constant  $k$ , termed the energy distribution curves (EDCs), locate  $E_k^*$  as the peak locations; thus  $E_{EDC}(k) \leftrightarrow E_k^*$ . On the other hand, the momentum distribution curves (MDCs) are obtained by fixing  $\omega$  and by scanning  $k$ . The so-obtained peak locations yield the fourth dispersion spectrum  $E_{MDC}$ . To obtain  $E_{MDC}$  in practice, one

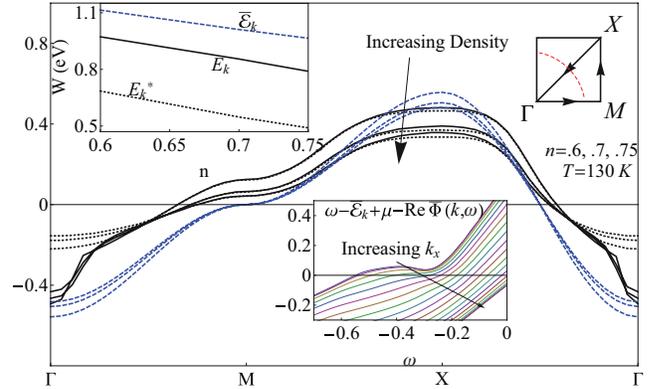


FIG. 2. (Color online)  $T = 130$  K. The three dispersions defined in Eq. (19) are plotted along principle directions for three different densities. The upper insets show the bandwidths of the dispersions as a function of the density. The bare bandwidth is 2 eV, but each of these dispersions shrinks compared to that scale. The bandwidth renormalization due to  $\text{Re } \bar{\Phi}$  in Eq. (19) is  $k$  dependent, and so  $E_k$  has a different shape than  $\epsilon_k$ . Note that  $E_k \sim \epsilon_k^*$  near the FS. However,  $E_k^*$  differs from  $E_k$  near the  $\Gamma$  point for each of the densities. The lower inset shows the evolution of the real part of the denominator of  $\mathbf{g}(k, \omega)$  with  $\omega$  to illustrate the origin of the difference between  $E_k$  and  $E_k^*$ . In the inset  $E_k$  is determined by the zero crossings of the curves. At low  $k$  notice that a relatively flat feature develops with a shallow minimum near  $\omega = -0.3$  eV. The minimum corresponds to the peak  $E_k^*$ . For increasing  $k$ , the flat feature quickly disappears and the zero crossing moves quickly upward in frequency, producing the observed kink in  $E_k$ .

may invert the MDC peak locations through

$$k^*(\omega) = \max[\rho_{\mathcal{G}}(k, \omega) : k], \quad E_{MDC}(k) = \text{Inverse of } k^*(E). \quad (20)$$

It is worth mentioning that the high-energy kink (or the waterfall) is experimentally defined as the peeling off of the  $E_{MDC}(k)$  from the  $E_{EDC}(k) = E_k^*$  spectra.<sup>29</sup>

In Fig. 2 we illustrate the density dependence of the three dispersions in Eq. (19). The inset shows the bandwidths,  $W(n)$ , of the three dispersions as a function of the density. Note that the bare bandwidth of  $\epsilon_k$  is 2 eV for both cases. Near the FS we see that  $E_k \approx E_k^*$ , but they differ near the  $\Gamma$  point where  $E_k^*$  and  $E_{MDC}$  are also split off from each other, satisfying the above operational definition of the high-energy kink. We now discuss the origin of these splittings.

Although  $E_k$  is not directly experimentally relevant, it plays a significant role in the theory, so we first comment on the splitting between  $E_k$  and  $E_k^*$  near the  $\Gamma$  point. Since  $E_k$  is defined as the root of  $\text{Re } \mathbf{g}^{-1}(k, E_k) = 0$ , we plot  $\omega + \mu' - \bar{\epsilon}_k - \text{Re } \bar{\Phi}(k, \omega)$  at various  $k$  as a function of  $\omega$  in the inset of Fig. 2. A strong  $\omega$  dependence of  $\text{Re } \bar{\Phi}(k, \omega)$  causes a flattening of the curves near the zero crossing between  $-0.6$  and  $-0.3$  eV, and this causes the  $E_k$  to fall rapidly with  $k$  in the main figure, Fig. 2. Just as  $E_k$  breaks away from  $E_k^*$ , so also does  $E_{MDC}$ , resulting in the kink. This is shown most clearly in the left panel of Fig. 3, where the spectral function is depicted as a color density plot with the dispersions ( $E_k, E_k^*, E_{MDC}$ ) overlaid. Near the  $\Gamma$  point where  $k = (0, 0)$  the QP becomes incoherent and the bulk of its spectral weight is

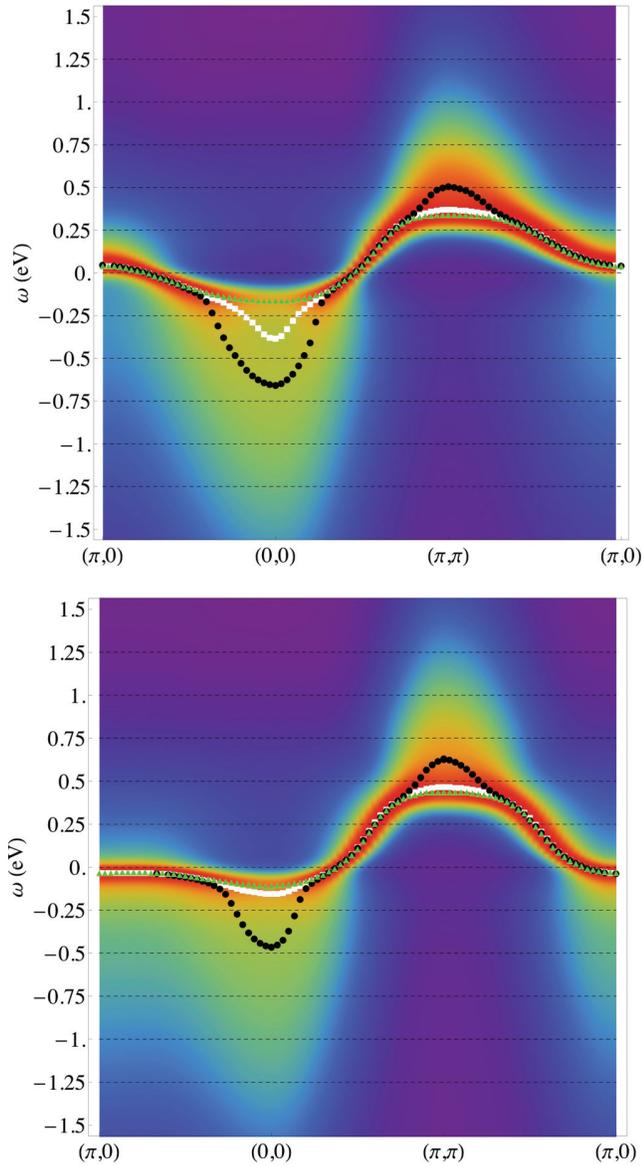


FIG. 3. (Color online)  $L = 60$ ,  $(n, T) = (0.75, 300)$  K. Density plot of  $A(k, \omega)$  of the minimal model (top) and the refined model Ref. 28 (bottom). (Here and below, red denotes high intensity and blue denotes low intensity).  $E_k$ ,  $E_k^*$ , and  $E_{MDC}(k)$  spectra are white, green, and black, respectively. Near  $k_F$  we see that the three spectra coincide. In the region near  $k = (0, 0)$ ,  $E_{MDC}(k)$  is at a significantly higher energy scale than  $E_k$  or  $E_k^*$ , signifying the high-energy kink (waterfall) effect. Also, the EDC peak loses weight in this regime. A new feature arises at near  $k = (\pi, \pi)$  resembling an inverted waterfall.

spread out to high negative frequencies. In this region  $E_{MDC}$  differs considerably from  $E_k^*$  and recovers the scale of the bare dispersion  $\epsilon_k$ . The right panel of Fig. 3 shows the spectral function as calculated using the tight binding parameters of *BiSSCO* given in Ref. 28. These parameters result in a holelike FS around the  $\Gamma$  point, unlike the minimal model with an electronlike FS. However, we observe in Fig. 3 that the high-energy kink occurs for both sets of parameters.

The occurrence of the high-energy kink is understandable as a straightforward consequence of additional broad peaks in the

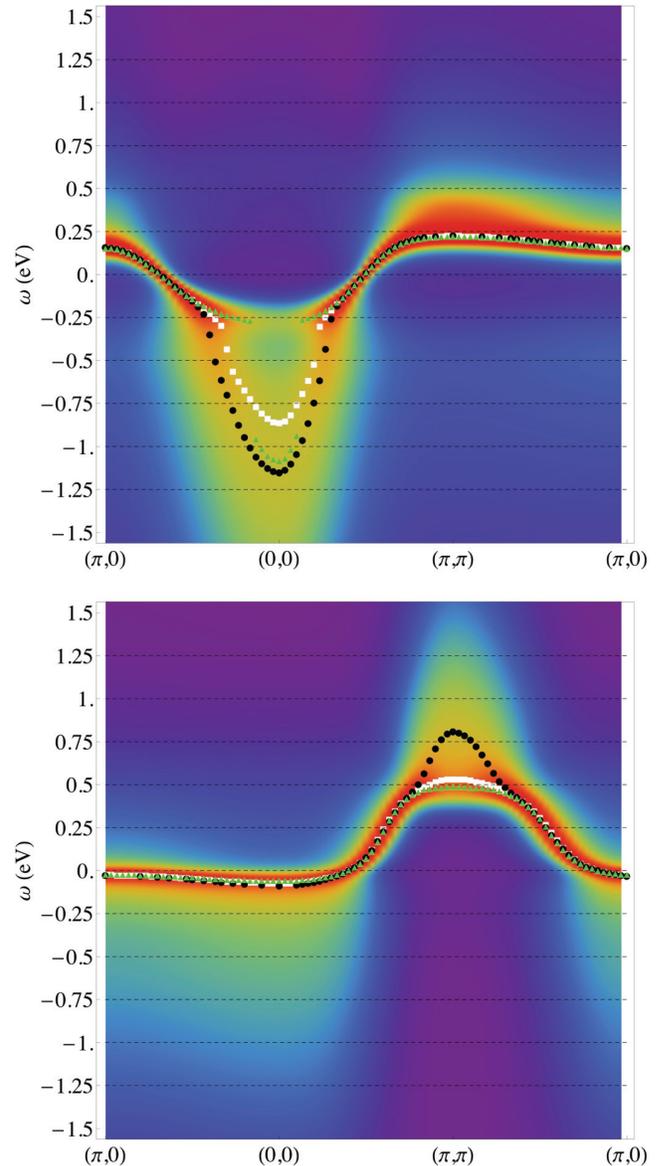


FIG. 4. (Color online)  $L = 60$ ,  $(n, T) = (0.75, 300)$  K. (Top)  $t'/t = 0.4$  is used to model electron-doped high- $T_c$  superconductors. The kink feature is prominent here. (Bottom) Uses  $t'/t = -0.4$  to crudely model a holelike FS. In this case the kink near  $(0, 0)$  is lost, unlike in Fig. 3, correlating with a flat (bare) band dispersion.

spectral function, separated from the quasiparticle-type peaks. In an energy range where they exist, these are particularly effective in dominating  $E_{MDC}$  and less prominent in  $E_{EDC}$ , therefore resulting in the separation between these dispersions.

While the qualitative picture of the kinks is reasonably clear, it is not immediately clear what accounts for the slightly different magnitude of the scale of the high-energy kink in Fig. 3. In Fig. 4 we show density plots of the spectral function with  $t'/t = \pm 0.4$ . The case  $t' = 0.4 \times t$  on the left has greater curvature at the band bottom and is identified with the phenomenology of the electron-doped cuprates (Refs. 22,30). The QP peaks lose most of their weight, unlike in the minimal case. The resulting scale of the drop in the waterfall is bigger

than in the minimal case, and correlates well with experimental observations in Ref. 31.

We note in Fig. 4 (right) that the case  $t'/t = -0.4$  has no measurable waterfall near the  $\Gamma$  point. The background at negative frequency is essentially featureless, and the QP peaks maintain their spectral weight. However, at positive frequencies, an inverted waterfall-like feature develops near  $k = (\pi, \pi)$ . This particular parametrization is often invoked to rectify the electronlike curvature of the minimal model ( $t' = 0$ ), but ends up giving a very flat band bottom at  $\Gamma$ . This is unlike the more sophisticated band parameters in Ref. 28, where the curvature is also holelike, and now the band regains significant curvature at its bottom, resulting in the observed kink.

### E. Detailed spectral line shapes (EDCs)

In this section, we present detailed line shapes for the spectral function. In an earlier work,<sup>3</sup> we have compared the results of the simplified ECFL formalism. These included some phenomenological inputs, with the experimental data at somewhat higher particle densities  $n \sim 0.85$ , and found remarkably good agreement with the line shapes. We are content in this work to present the results at lower particle densities, but *from a microscopic calculation* of ECFL. This is made possible by solving the  $O(\lambda^2)$  equations in Eq. (6) numerically. The line shapes obtained here have a similar general nature as the ones in Ref. 3, giving support to that work. However, as one expects from a lower-density situation, we find somewhat less dynamical asymmetry about zero energy. More detailed comparison with data near optimal doping with the microscopic ECFL theory must await the solution of the third- or higher-order equations, where the criterion for validity discussed above [see paragraph following Eq. (8)] is satisfied more closely than here.

Let us first examine the local density of states (LDOS) at  $n = 0.75$  for both cases at low  $T$  in Fig. 5. A prominent feature is that the main peak is much narrower than in the bare LDOS. Furthermore, there is a long tail extending to (negative)

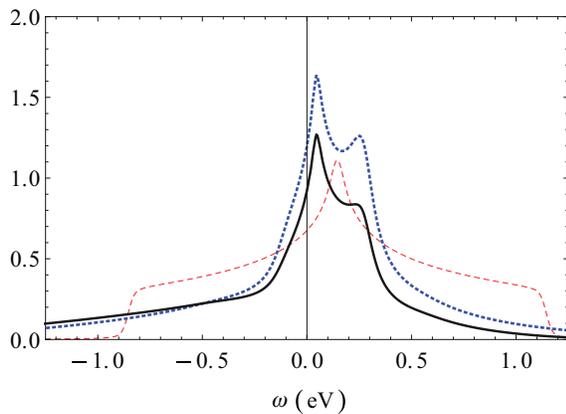


FIG. 5. (Color online)  $n = 0.75$ . The LDOS of the physical  $\mathcal{G}$  (auxiliary  $\mathbf{g}$ ) is in black (dotted blue), and the bare DOS is the dashed red curve. The renormalized band displays narrowing, and a long tail at  $\omega < 0$ . The LDOS develops a second spectral peak for  $\omega > 0$  from a strongly  $k$ -dependent feature in the self-energy.

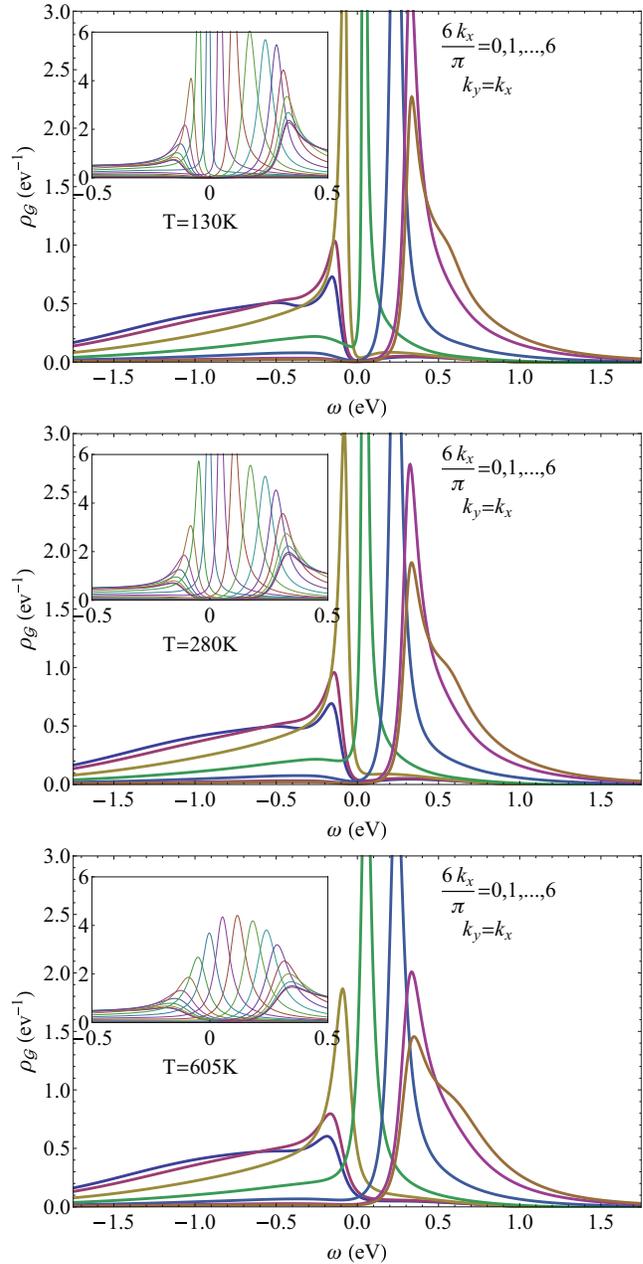


FIG. 6. (Color online)  $n = 0.75$ . The spectral function  $\rho_{\mathcal{G}} [= A(k, \omega)]$  at several  $k$  points along the  $\langle 11 \rangle$  direction and  $T$ . We used  $L_x = 36$ ; the insets show all positive  $k_x$ 's and the main figures display a third of the allowed  $k_x$ 's. The inset in each case zooms out to reveal the heights. The linewidth near  $k_F$  is seen to be strongly effected by rising  $T$ ; the incoherent parts have very little  $T$  dependence. The tails exhibit a secondary broad peak near  $\omega = -0.4$  eV, giving rise to the high-energy kink (waterfall).

frequencies, much greater than those seen in the bare LDOS. Finally, we note that the LDOS acquires a second peak at positive frequency. This peak arises due to some  $k$ -dependent features in the self-energy (discussed below), resulting in sharper QP at positive frequency.

We next discuss Fig. 6, displaying the nodal spectral function at three different temperatures. The lines are quite

sharp near  $k_F$  but broaden out rapidly away from  $k_F$ . The insets give an idea of the change of spectral density with temperature. Notably, there is a secondary local maximum for  $k$  near the  $\Gamma$  point near  $\omega = -0.4$  eV. This second peak is responsible for the waterfall discussed above and is also contained in the models used in Refs. 3,5. As discussed above in connection with kinks, its microscopic origin is sensitive to tight binding parameters. It is also noteworthy that lines with  $k > k_F$ , though broader than at  $k_F$ , are sharper than those with  $k < k_F$ .

Finally, we note that while the self-energy is strongly  $k$  dependent it is not anisotropic. Consequently, the EDC line shapes look similar at different parts of the FS, at least to  $O(\lambda^2)$ . In the regime of validity of this theory, namely, the (hole) overdoped region, the cuprates do not display a strong anisotropy either.

### F. Optical conductivity

The optical conductivity  $\sigma(\Omega)$  is computed within the lowest approximation of (I) here by discarding the vertex corrections and working with the auxiliary  $\mathbf{g}$ :

$$\text{Re}\sigma(\Omega) = \frac{1}{\Omega} \sum_k v_k^2 \int \rho_{\mathbf{g}}(k, \omega) \rho_{\mathbf{g}}(k, \Omega + \omega) d\omega \times [f(\omega) - f(\Omega + \omega)]. \quad (21)$$

The imaginary part of the conductivity can be obtained by a Hilbert transform of the real part. In this purely  $t$ - $J$  calculation we must be careful how we interpret the imaginary part of  $\sigma$ . A more realistic calculation should include contributions from the upper Hubbard band and from charge-transfer processes that are significant at high frequencies; these are discarded in the  $t$ - $J$  model. For our current purposes we will discuss two kinds of relaxation rates. First we compute a momentum-averaged rate  $1/\tau_\sigma$  extracted from the low-frequency behavior  $\sigma(\omega)$  using

$$\frac{1}{\tau_\sigma} = \frac{4}{\pi} \int_0^{1/\tau_\sigma} \text{Re} \sigma(\omega) / \sigma(0) d\omega, \quad (22)$$

where the prefactor is chosen to yield the usual rate for a Lorentzian shape. This convenient definition is designed to be insensitive to the shape of  $\sigma(\omega)$ . Secondly, we look at the momentum-resolved scattering lifetimes, defined as the inverse width of the ARPES line shape at the Fermi momentum. These scattering rates are displayed in Fig. 7. We find that the  $1/\tau$  curves from ARPES and the conductivity have essentially the same temperature dependence, apart from a factor of  $O(1)$ . The  $1/\tau$  rises quadratically at low temperature, in accordance with the standard Fermi liquid (FL) picture, crossing over to a linear dependence at a fairly low-temperature scale.

In Fig. 8, we display the computed optical conductivity  $\text{Re} \sigma(\omega)$  at various  $T$  for  $n = 0.75$ , and also the phase angle  $\theta = \tan^{-1}(\frac{\sigma''(\omega)}{\sigma'(\omega)})$  on an absolute scale. The rapid fall of the optical conductivity at low  $T$  is rapidly filled in at low  $\omega$ , and the phase angle falls off with  $\omega$  at about  $4000 \text{ cm}^{-1}$ . At optimum doping, the phase angle is known experimentally to be flat in  $\omega$  over a much larger range,<sup>32</sup> and differs from the present calculation, whose validity is confined to overdoping. Experimental measurements in the overdoped case of the phase

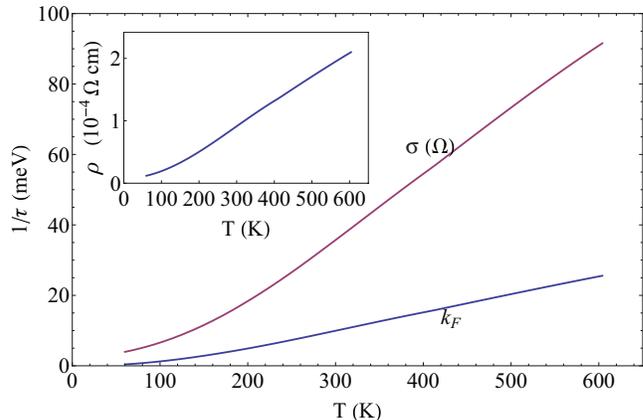


FIG. 7. (Color online)  $n = 0.75$ . The QP relaxation rate at the FS along  $\langle 11 \rangle$  obtained from  $\rho_{\Sigma}(k, E_k^*)$ , and the rate obtained from the optical conductivity as in Eq. (22). The  $T^2$  behavior of an FL is visible at low temperature, crossing over at a modest temperature ( $\sim 150$ ) K, partly due to the shrinking bandwidth, as seen directly in Fig. 2. The inset shows the dc resistivity obtained from the inverse of Eq. (21). It similarly displays a  $T^2$  behavior crossing over to a linear behavior, as well as a lack of saturation that persists to higher  $T$  than shown.

angle would be useful in benchmarking theories in regimes such as the present one. For the real part, such a comparison is possible. In Fig. 9 we display the  $\text{Re} \sigma(\omega)$  curves along with optical conductivity measurements published by Puchkov *et al.*<sup>12</sup> for an overdoped thallium compound. We note that in the overdoped regime, the computed conductivity matches quite well with experiments (to within a factor 2 on the vertical scale).

A further interesting aspect of the resistivity obtained from this ECFL formalism lies in the high-temperature limit. A lack of resistivity saturation has been observed in numerical treatments of strongly coupled models, as in a recent DMFT work.<sup>25</sup> These results are in qualitative agreement with resistivity measurements in the cuprates and other strongly

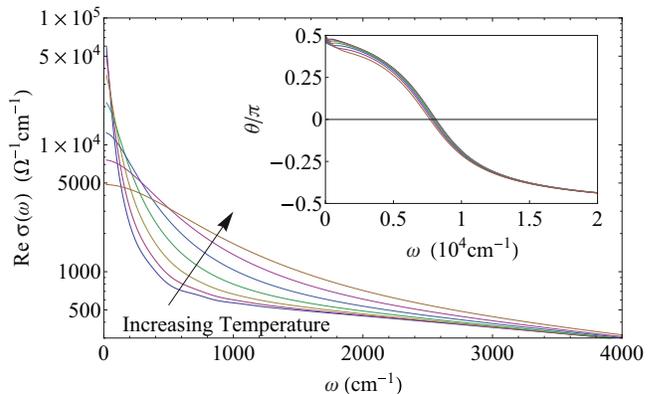


FIG. 8. (Color online)  $n = 0.75$ ,  $T = 60, 90, 130, 190, 280, 410, 605$  K. The optical conductivity is calculated on an absolute scale and illustrates how increasing  $T$  rapidly fills up the regime  $200 \leq \omega \leq 1000 \text{ cm}^{-1}$ . The rise of conductivity at very low  $\omega$  is also inferred from the dc resistivity displayed in Fig. 7. The phase of the complex  $\sigma$  falls off rapidly beyond  $4000 \text{ cm}^{-1}$ .

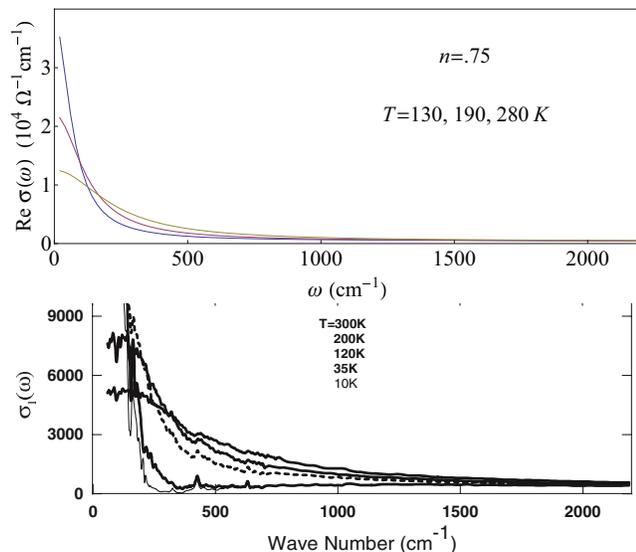


FIG. 9. (Color online) An explicit comparison of optical conductivity with measurements of Puchkov *et al.* from Ref. 12 with the author's kind permission. The data pertains to an overdoped thallium-based cuprate with  $T_c = 23$  K, with a density  $n \approx 0.75$ . We note the similarity of magnitude and variation with  $\omega$  and  $T$ . It is worth noting (to be reported elsewhere) that the vertical scale can be brought into better agreement with an adjusted hopping, as with the Fermi velocity.

correlated compounds. The ECFL theory leads to a similar result and provides a simple picture for its origin in terms of the second Lagrange multiplier  $u_0$ . As discussed in the Supplemental Material,<sup>27</sup> both  $\mu'$  and  $u_0$  rise linearly with  $T$  at high temperature. Due to the explicit appearance of  $u_0$  in the expressions for  $\Phi$  and  $\Psi$ , the magnitude of the self-energies also grows continuously with temperature via  $u_0$ , resulting in a monotonic broadening of the spectral function. This broadening is insensitive to the Mott-Ioffe-Regel (MIR) saturation expected in weakly correlated metals, and leads to a nonsaturating resistivity at high  $T$ , as we observe in the inset of Fig. 7.

### G. Self-energies

We now display the self-energies that are involved in calculating the spectral functions. In Fig. 10 we display  $\rho_{\bar{\Phi}}$  and  $\rho_{\Psi}$ . Both functions exhibit the  $\omega^2$  behavior close to zero, as one finds for a weakly interacting FL self-energy. Unlike conventional FLs, the magnitude of the quadratic term is strongly  $k$  dependent. From these functions and the associated real parts we can construct a Dyson-Mori (D-M) self-energy, defined through the equation

$$\mathcal{G} = \frac{a_{\mathcal{G}}}{x - \Sigma}, \quad (23)$$

where  $a_{\mathcal{G}}$  is the total spectral weight of the physical  $\mathcal{G}$  and  $x = \omega + \mu' - \bar{\epsilon}_k$  such that

$$\Sigma = x + \frac{a_{\mathcal{G}}}{a_{\mathcal{G}} + \Psi} (\bar{\Phi} - x). \quad (24)$$

In Fig. 11 we plot the computed imaginary part of the D-M self-energy,  $\rho_{\Sigma}$ . It exhibits a similar magnitude and  $k$  dependence at

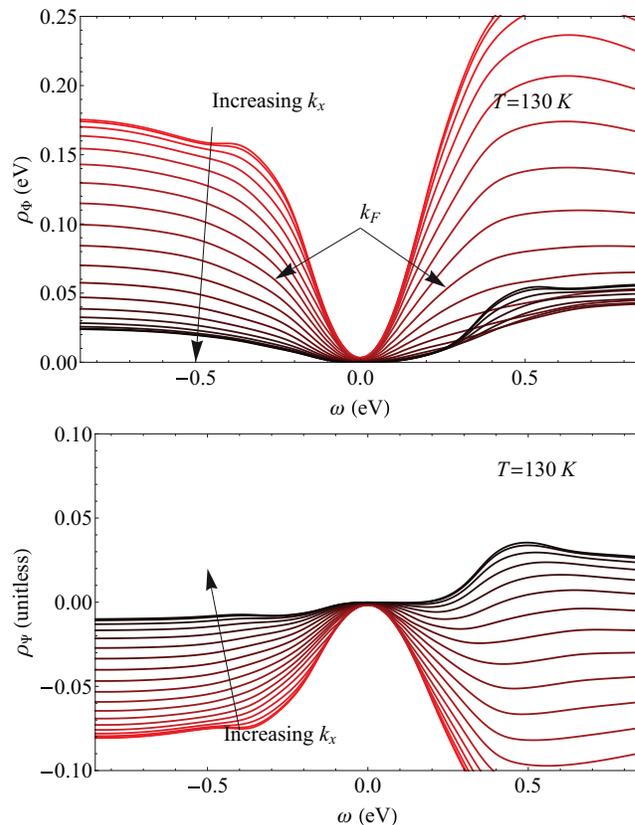


FIG. 10. (Color online)  $(n, T) = (0.75, 130)$  K. The spectral functions for the two self-energies  $\Phi$  and  $\Psi$ , i.e.,  $\rho_{\bar{\Phi}}$  (top) and  $\rho_{\Psi}$  (bottom), at several  $k$  points along the  $\langle 11 \rangle$  direction. Both are roughly quadratic and symmetric at low frequency but have a strongly  $k$ -dependent curvature. In the plot of  $\rho_{\bar{\Phi}}$ , the minimum width  $\eta$  chops off the bottom of the low-frequency minimum.

low frequency to that in  $\rho_{\bar{\Phi}}$ . However, large asymmetries begin to appear at intermediate frequencies. It is interesting that at positive frequency the function is considerably smaller than at negative frequencies, a feature that has already been noted for simplified versions of the ECFL<sup>4,5</sup> and also in a recent DMFT study of the Hubbard model.<sup>25</sup> In this calculation, however, we see an interplay between the momentum and frequency dependencies. In particular, we see that at positive frequency  $0 < \omega \lesssim 200$  meV,  $\rho_{\Sigma}$  is strongly  $k$  dependent, so that particlelike excitations near  $k = (\pi, \pi)$  are long-lived while those inside the FS suffer a large damping. This is very different from weakly coupled or local theories such as DMFT, where the scattering rate is determined by frequency alone. We note that this self-energy does not differentiate between nodal and antinodal directions, but rather, the  $k$  dependence arises only through  $\epsilon_{\bar{k}}$ , so that the scattering rate is constant along the FS.

The low-frequency asymmetry is usefully described as an FL-like quadratic dependence modified by a cubic term. The right panel of Fig. 11 shows low-frequency ( $|\omega| \leq 75$  meV) fit parameters of  $\rho_{\Sigma}$  as a function of  $k$ , exhibiting a marked softening of the quadratic coefficient  $b$ . The final effect on the relaxation rate  $\Gamma(k) = \rho_{\Sigma}(k, E_k^*)$ , displayed in Fig. 12, is

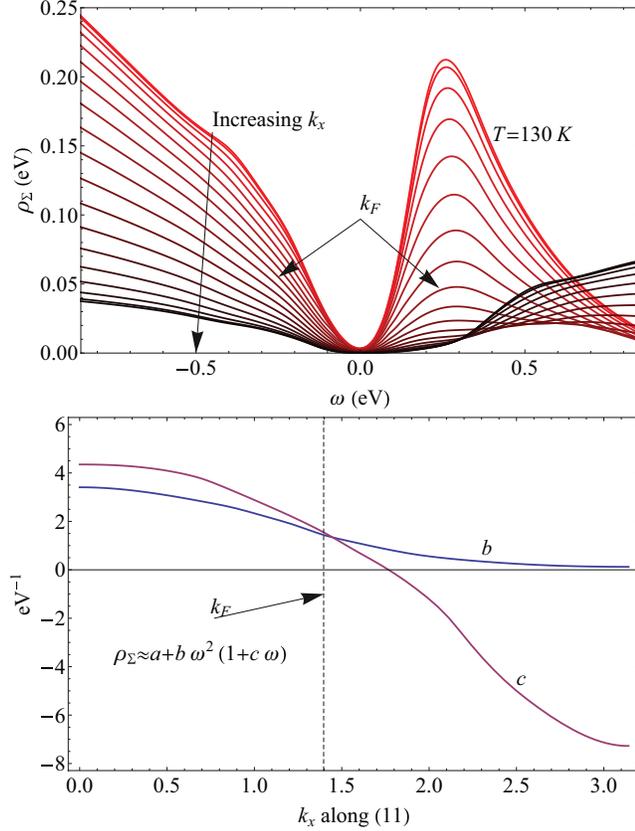


FIG. 11. (Color online)  $(n, T) = (0.75, 130)$  K. (Top) The spectral function  $\rho_\Sigma$  of the Dyson-Mori self-energy  $\Sigma$  from Eq. (24), at several  $k$  points along the  $\langle 11 \rangle$  direction. As with  $\rho_\Phi$ ,  $\rho_\Sigma$  has inherited a strong  $k$  dependence. (Bottom)  $k$  dependence of the fit parameters from  $\rho_\Sigma = a + b\omega^2(1 + c\omega)$  at low frequencies  $|\omega| \leq 75$  meV. Observe the softening of the quadratic coefficient with increased  $k$ . The cubic term  $\rho_\Sigma \propto \omega^3$  produces particle-hole asymmetry, as argued in Ref. 6, and grows in magnitude with increasing  $k$  beyond  $k_F$ .

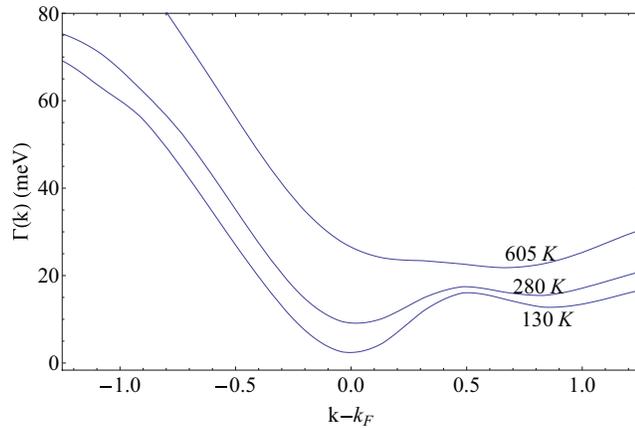


FIG. 12. (Color online)  $n = 0.75$ . The decay rate of QP near  $k_F$  along the nodal line from  $\Gamma_k = \rho_\Sigma(k, E_k^*)$ . The strong and  $T$ -dependent asymmetry makes quasiparticles longer lived at  $k > k_F$ . With increased  $T$  the minimum of  $\Gamma$  moves to  $k > k_F$ .

summarized by the expression

$$\Gamma(k) \sim b_f \left( 1 - \left| \frac{b'_f}{b_f} \right| (k - k_F) \right) V_F^2 (k - k_F)^2, \quad (25)$$

where  $b_f$  ( $b'_f$ ) is the coefficient (derivative of the coefficient) at the Fermi momentum, and  $V_F$  is the Fermi velocity. The cubic term in  $k - k_F$  is a significant correction to the leading term from Fermi liquid theory, resulting in longer-lived quasiparticles outside the Fermi surface, as compared to quasiholes inside the Fermi surface. Furthermore, the  $T$  dependence of  $\Gamma$  is stronger at  $k < k_F$ . At the highest temperature shown, the longest lived quasiparticles drift somewhat away from  $k_F$ . In Fig. 7, we also display the  $T$  dependence of the single-particle relaxation rate  $\Gamma(k)$ . This rate shows a crossover at a reduced scale to linear in  $T$  behavior, about  $\sim 150$  K, as compared to  $T_\mu \sim 400$  K, detailed in the Supplemental Material.<sup>27</sup>

## V. CONCLUDING REMARKS

In summary, we have presented the results of a systematic low-density expansion for the  $t$ - $J$  model using the recently developed formalism of *extremely correlated Fermi liquids*, discussed in Refs. 1 and 2. This calculation complements the phenomenological theory in Ref. 3, where the line shapes at optimal doping are successfully modeled, using a very small number of parameters. Here we calculate from first principles, assuming only the value of  $J$  and the hopping  $t$ , and where possible, quote results on an absolute scale. The second order in  $\lambda$  equations studied here, valid for  $n \lesssim 0.75$ , are somewhat removed from the most interesting regime of optimal doping. Nevertheless, the computed forms of the twin self-energies found here indeed have the character assumed in the phenomenological ECFL studies; also, the resulting spectral functions have line shapes that are skewed towards negative  $\omega$ . This feature is ultimately a consequence of Gutzwiller projection, as argued in Ref. 1, and captures a striking characteristic of the experimental data.

The salient points from our study may be summarized as follows:

(i) The momentum occupation function  $m_k = \langle \hat{C}_k^\dagger \hat{C}_k \rangle$  is calculated along the nodal direction at various  $T$  and densities, where it indicates a large spillover for  $k > k_F$ . This spillover quantifies the smooth part of spectral weight at  $\omega < 0$  for wave vectors  $k > k_F$  and is of potential use in calibrating ARPES studies.

(ii) The spectral functions  $A(k, \omega)$  at various  $k$  values and different temperatures displays a non-Lorentzian form, with a pronounced skew towards occupied energies  $\omega < 0$ . This results in spectra resembling those seen in most experiments in cuprates and emerges as a natural consequence of the Gutzwiller projection, i.e., very strong correlations.

(iii) The dispersion relations  $E_{MDC}(k)$  and  $E_{EDC}(k)$  are deduced from the peaks of  $A(k, \omega)$  and display considerable band narrowing due to correlations. They further split apart near  $\vec{k} \sim (0, 0)$ , i.e., the  $\Gamma$  point, resulting in a high-energy kink, quite similar to that seen in experiments. The splitting between these peaks is due to a prominent broad second maximum in the spectral function, away from the quasiparticle peak. A high sensitivity of the high-energy kink to the

bare band parameters is found, with flatband dispersions eliminating the kinks.

(iv) The ECFL results for the optical conductivity and the phase angle are reported on an absolute scale, and the real part is in quite reasonable proximity of experimental data. Better agreement should be possible with tuning the available band parameters, although we have not explored this here.

(v) The resistivity is calculated as a function of  $T$  at various densities and found to be nonsaturating in its  $T$  dependence, analogous to the resistivity seen in experiments. The absence of saturation is easy to understand within the ECFL formalism. The magnitude of the self-energy grows indefinitely due to its

dependence on the second chemical potential  $u_0$  and leads to a growing resistivity from the Kubo formula.

(vi) The single-particle decay rate  $\Gamma(k, T)$  is reported at various  $k$  and  $T$ . It is smaller for  $k > k_F$  than for  $k < k_F$  due to a strong correction to Fermi liquid behavior, leading to spectral lines that are narrower than for  $k < k_F$ .

#### ACKNOWLEDGMENTS

This work was supported by the DOE under Grant No. FG02-06ER46319. We thank Gey-Hong Gweon for stimulating discussions and Ehsan Khatami for helpful comments on the manuscript.

- 
- <sup>1</sup>B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).  
<sup>2</sup>B. S. Shastry, *Phys. Rev. B* **87**, 125124 (2013).  
<sup>3</sup>G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).  
<sup>4</sup>K. Matsuyama and G.-H. Gweon, arXiv:1212.0299.  
<sup>5</sup>B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011).  
<sup>6</sup>B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).  
<sup>7</sup>P. Nozières, *Theory of Interacting Fermi Systems* (W. A. Benjamin, Amsterdam, 1964).  
<sup>8</sup>A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).  
<sup>9</sup>G. D. Mahan, *Many-Particle Physics* (Plenum Press, New York, 1981).  
<sup>10</sup>M. Bejas, A. Greco, and A. Foussats, *Phys. Rev. B* **73**, 245104 (2006); E. Cappelluti and R. Zeyher, *ibid.* **59**, 6475 (1999).  
<sup>11</sup>A. Pasupathy, A. Pushp, K. Gomes, C. Parker, J. Wen, Z. Xu, G. Gu, Y. Ando, and A. Yazdani, *Science* **320**, 196 (2008), see especially Figs. 1 A, 1B, and 4A.  
<sup>12</sup>A. V. Puchkov, D. Basov, and T. Timusk, *J. Phys.: Condens. Matter* **8**, 10049 (1996).  
<sup>13</sup>W. Stephan and P. Horsch, *Phys. Rev. Lett.* **66**, 2258 (1991).  
<sup>14</sup>R. R. P. Singh and R. L. Glenister, *Phys. Rev. B* **46**, 14313 (1992).  
<sup>15</sup>C. Gros, R. Joynt, and T. M. Rice, *Phys. Rev. B* **36**, 381 (1987).  
<sup>16</sup>B. Edegger, V. N. Muthukumar, and C. Gros, *Adv. Phys.* **56**, 927 (2007).  
<sup>17</sup>A. Paramekanti, M. Randeria, and N. Trivedi, *Phys. Rev. Lett.* **87**, 217002 (2001).  
<sup>18</sup>M. M. Zempljic and P. Prelovsek, *Phys. Rev. B* **72**, 075108 (2005).  
<sup>19</sup>M. M. Zempljic, P. Prelovsek, and T. Tohyama, *Phys. Rev. B* **76**, 012502 (2007); *Phys. Rev. Lett.* **100**, 036402 (2008).  
<sup>20</sup>J. Kokalj and P. Prelovsek, *Phys. Rev. B* **75**, 045111 (2007).  
<sup>21</sup>J. Jaklic and P. Prelovsek, *Adv. Phys.* **49**, 1 (2000).  
<sup>22</sup>B. Moritz, F. Schmitt, W. Meevasana, S. Johnston, E. M. Motoyama, M. Greven, D. H. Lu, C. Kim, R. T. Scalettar, Z.-X. Shen, and T. P. Devereaux, *New J. Phys.* **11**, 093020 (2009).  
<sup>23</sup>W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989).  
<sup>24</sup>A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).  
<sup>25</sup>X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **100**, 086401 (2013).  
<sup>26</sup>E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, and B. S. Shastry, *Phys. Rev. B* **87**, 161120(R) (2013).  
<sup>27</sup>See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.87.245101> for details of the results for thermodynamics and the wave-function renormalization  $Z_k$ , and further details of the computational method employed.  
<sup>28</sup>R. S. Markiewicz, S. Sahrakorpi, M. Lindroos, H. Lin, and A. Bansil, *Phys. Rev. B* **72**, 054519 (2005).  
<sup>29</sup>J. Graf, G.-H. Gweon, K. McElroy, S. Y. Zhou, C. Jozwiak, E. Rotenberg, A. Bill, T. Sasagawa, H. Eisaki, S. Uchida, H. Takagi, D.-H. Lee, and A. Lanzara, *Phys. Rev. Lett.* **98**, 067004 (2007); W. Meevasana *et al.*, *Phys. Rev. B* **75**, 174506 (2007).  
<sup>30</sup>L. Hozoi, M. S. Laad, and P. Fulde, *Phys. Rev. B* **78**, 165107 (2008).  
<sup>31</sup>S. R. Park, D. J. Song, C. S. Leem, C. Kim, C. Kim, B. J. Kim, and H. Eisaki, *Phys. Rev. Lett.* **101**, 117006 (2008).  
<sup>32</sup>D. van der Marel, H. J. A. Molegraaf, J. Zaanen, Z. Nussinov, F. Carbone, A. Damascelli, H. Eisaki, M. Greven, P. H. Kes, and M. Li, *Nature (London)* **425**, 271 (2003).

# Extremely Correlated Fermi Liquids: Self consistent solution of the second order theory.

## Supplementary Material

Daniel Hansen and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, CA 95064, USA*

(Dated: May 10, 2013)

### THERMODYNAMICS

In Fig. (13) we display the density dependence of the two Lagrange multipliers,  $\mu'$  and  $u_0$ , as well as the physical chemical potential,  $\mu$ , as defined in Eq. (9) of the main paper. The plot illustrates a few points regarding the nature of  $\mathbf{g}$  and  $\mathcal{G}$ , and their differences. First, we see that  $\mu \approx \mu'$  in the low density limit. With increased density,  $u_0$  grows monotonically causing the development of large asymmetric tails in the spectral background. Furthermore, with the growth of both density and  $u_0$ ,  $\mu$  grows faster than  $\mu'$ . While  $\mu'$  approaches 0, a limit expected for a half-filled FL with particle-hole symmetry,  $\mu$  approaches a limit which is comparable to the renormalized bandwidth, i.e. the top of the band. This difference therefore also signals the reduction of spectral weight in  $\mathcal{G}$  relative to  $\mathbf{g}$ , a symptom of the main constraint of this theory, namely the removal of states with double occupancy.

The temperature dependence of  $(\mu', \mu, u_0)$  is shown below in Fig. (14) revealing the temperature scale for the thermodynamics of the ECFL at a single density ( $n = .75$ ). At the lowest temperatures we expect that the chemical potential goes as  $\mu'(T) = \mu'(0) - bT^2$  (to focus on the quadratic term we subtract off the  $T = 0$  intercepts in Fig. (14)). We can define a temperature scale  $T_{\mu'}$  which sets the strength of the  $O(T^2)$  term according to

$$T_{\mu'} = \sqrt{\left| \frac{\mu'(0)}{b} \right|}. \quad (\text{S.1})$$

where we have used the auxiliary ‘‘chemical potential’’  $\mu'$  because it is most closely related to the FL aspects of the spectral function.  $T_{\mu'}$  is plotted in the inset as a function of the density  $n$ . Note the reduction of the scale of  $T_{\mu'}$  as we approach half filling. At high  $T$  ( $> 600K$ ),  $\mu'$  rises linearly with  $T$  in typical FL fashion.

We also observe in Fig. (14) that the temperature dependence of  $\mu$  comes from  $\mu'$  as well as that of  $u_0$  and hence, unlike in a simple FL, is substantially driven by many body effects even at low  $T$ . Furthermore, the scale of the temperature variation is larger than would be expected for weakly coupled systems. We note the large variation of the chemical potential with temperature,  $\Delta\mu \sim 20meV$  on heating from 100K to 300K, and about half this number for the change for  $\Delta\mu'$ . This large variation should be readily measurable in ARPES, and appears to have been overlooked in most studies so far.

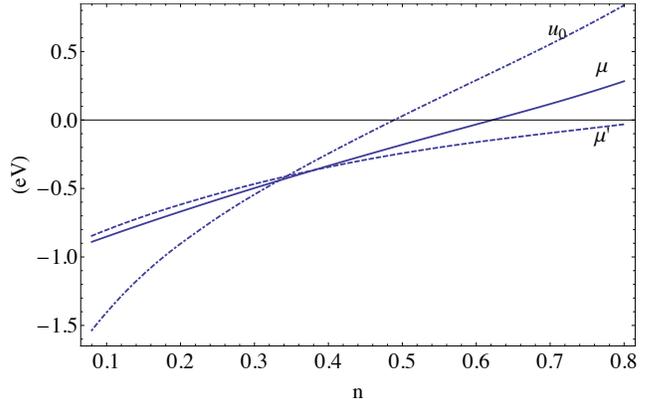


FIG. 13.  $T = 280K$ . The density dependence of the various chemical potentials, as defined in Eq. (9) of the main paper. The potentials  $\mu'$  and  $u_0$ , and the net physical chemical potential  $\mu$ , dashed, dot-dashed, and solid, respectively. The potential  $\mu'$  approaches 0 as  $n \rightarrow 1$ , i.e. half filling, while the potential  $u_0$  rises monotonically with density. The physical chemical potential  $\mu$  becomes positive at high densities. This is natural for a state where spectral weight has been removed by the single occupancy constraint, so that  $n=1$  corresponds to a filled rather than half-filled band. The scale of  $\mu$  and  $u_0$  corresponds to the scale of the renormalized bandwidth in the high density limit.

Unlike  $\mu'$ ,  $u_0$  does not follow the standard behavior of a FL chemical potential. Whereas the low  $T$  temperature dependence of  $\mu'$  arises from small changes in occupation at low frequency only,  $u_0$  feels *all frequencies* due to its explicit appearance in  $\bar{\Phi}$  and  $\Psi$ . The scaling with  $T$  is therefore difficult to predict by a Sommerfeld type argument. Numerically, we observe linear- $T$  behavior at low  $T$  and at high  $T$  separated by a minima at an intermediate  $T$ .  $u_0$  asymptotes to a  $T$  linear behavior for  $T \gtrsim 650K$ .

### QUASIPARTICLE WEIGHT

The spectra obtained here contain sharp peaks, and also substantial background due to extreme correlations which can be quantified after some care is taken in defining a suitable  $Z_k$ . In a conventional FL the QP weight is defined by  $Z_k = \frac{1}{1 - \frac{\partial \Sigma}{\partial \omega}(k_F, 0)}$  where  $\Sigma$  is the Dyson self energy. This definition does not immediately work for us as we do not have a conventional Dysonian form

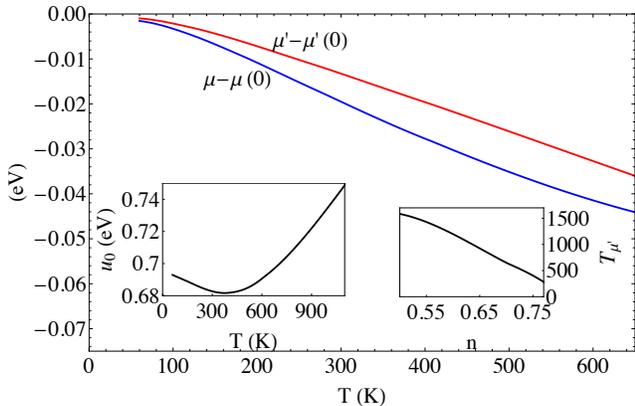


FIG. 14.  $\mu'$  and  $\mu$  versus  $T$  at  $n = .75$ . An effective Fermi temperature  $T_{\mu'}$ , found from Eq. (S.1), is shown as a function of the density. Note that the scale of variation of the  $\mu'$ ,  $\Delta\mu' \sim 10$  meV in heating from  $T = 100\text{K}$  to  $T = 300\text{K}$ , is quite large and is potentially observable in ARPES. A shrunk overall energy scale, seen most clearly in the reduction of the QP bandwidth, is ultimately responsible for this sensitivity. The temperature dependence of  $u_0$  is shown in each case to be non-monotonic with a minimum at finite temperature. At high temperature  $u_0$  rises linearly with  $T$ , similarly to  $\mu'$ .

for our  $\mathcal{G}$ . To obtain an appropriate definition for  $Z_k$  we note that the ECFL Greens function can be written in a Dyson-Mori form

$$\mathcal{G} = \frac{a_G + \Psi}{x - \Phi} = \frac{a_G}{x - \Sigma_{DM}} \quad (\text{S.2})$$

where  $a_G = 1 - \frac{n}{2}$  exactly and  $x = i\omega + \mu - \bar{\epsilon}_k$ . However in the present approximation  $a_G = 1 - \frac{n}{2} + \frac{n^2}{4}$  owing to the second order in  $\lambda$  approximation. In analogy to the standard FL we now define

$$Z_k = \frac{a_G}{1 - \frac{\partial}{\partial \omega} \Sigma_{DM}}. \quad (\text{S.3})$$

While it may be tempting to drop the factor  $a_G$  from Eq. (S.3), it represents an important piece of physics in the larger context and therefore must be retained. To elaborate this, note that the full spectral function of canonical electrons, e.g. in a large  $U$  Hubbard model, would have features at the scale of  $U$  that correspond to the upper Hubbard band, and are thrown out in the  $t$ - $J$  model thereby isolating the lower Hubbard band. Thus in a comprehensive canonical theory, the (low) value of  $Z_k$  representing a faint QP feature found with the present definition, would be compensated by a large background piece with net weight  $1 - Z_k$  contained partly in the lower Hubbard band, and partly in the upper Hubbard band that lies outside the domain of the  $t$ - $J$  model. It is therefore Eq. (S.3) that can be compared to the values found in experiments, and also in studies of the Hubbard model with large  $U$ .

Alternatively, if we define the Dyson self energy according to  $\mathcal{G} = \frac{1}{x - \Sigma'}$ , the quasiparticle weight would once again be given by the more familiar expression  $Z'_k = \frac{1}{1 - \frac{\partial \Sigma'}{\partial \omega}}$ , and moreover the computed  $Z'_k$  would be the same as in Eq. (S.3). Including  $a_G$  in the definition of  $\Sigma$  is a convenience employed for two reasons. First, it separates the effects of high and low energy physics formally. The factor  $a_G$  in Eq. (S.2) represents an overall depletion of the lower Hubbard band (due to excitations at scale  $U$  in a Hubbard model) and the term  $\frac{\partial \Sigma_{DM}}{\partial \omega}$  in the denominator accounts for the depletion of the QP by low frequency FL effects. Secondly, introducing  $a_G$  results in a self energy  $\Sigma_{DM}$  that reaches a constant at high frequency, unlike the Dyson form given above with  $Z'$  which grows indefinitely.

Numerics are performed at finite temperature and the  $T=0$  value of  $\frac{\partial \Sigma}{\partial \omega}$  is obtained by extrapolation of the finite  $T$  data. It is found that  $Z_k$  falls quickly with density as plotted in Fig. (15).  $Z_k$  is always somewhat less than  $1 - n$  decreasing roughly linearly with density. At the highest densities, near the limits of this theory, it begins to flatten. Thus, a Mott transition is absent in the  $O(\lambda^2)$  theory.

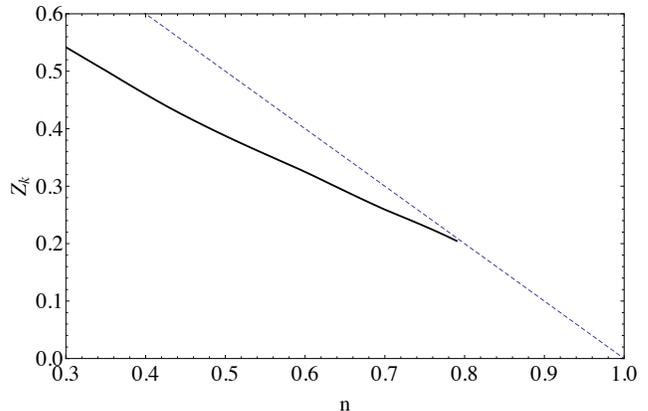


FIG. 15. The QP weight  $Z_{k_F}$  is plotted as a function of density.  $Z_{k_F}$  decreases with increased density, similar to  $(1-n)$  (which is depicted as a dotted line) but somewhat less. The incoherent spectral contribution is therefore already  $\sim 4$  times greater than the QP part at  $n \sim .7$ , and this ratio appears to increase further near half filling.

## FLOWCHART OF THE ITERATIVE PROCESS

Here we summarize the process by which numerical self consistency is achieved. The self consistency loop proceeds as follows.

1. Initialize all quantities to those of the Fermi gas:  $\mu' = \mu_0$ ,  $\rho_\Psi = \rho_\Phi = 0$ ,  $u_0$  is set to a plausible value of  $2t$ .
2. Build  $\rho_g$  from latest instance of  $\mu'$ ,  $\bar{\Phi}$ .

3. Calculate  $\rho_{\bar{\Phi}}$  from latest instance of  $\rho_{\mathbf{g}}$  and  $u_0$ . Obtain the real part via Hilbert transform.
4. Calculate new  $\boldsymbol{\mu}'$  using a bisection root finder.
5. Repeat steps 2-4 until  $\boldsymbol{\mu}'$  and  $\rho_{\bar{\Phi}}$  have converged to specified tolerance.
6. Calculate  $\rho_{\Psi}$  from latest  $\rho_{\mathbf{g}}$  and  $u_0$ . Obtain real part through Hilbert transform.
7. Calculate  $\sum_k \Psi(k)\mathbf{g}(k)$  and recalculate  $u_0$  with a root finder.
8. Return to Step 2 and repeat loop. Continue to the next step only when  $u_0$  has converged to specified tolerance.
9. Calculate  $\rho_{\mathcal{G}}$ .

The most computationally expensive step in this loop is the double integration for  $\rho_{\bar{\Phi}}$ . If computed by a direct summation the computational time required would scale as  $N_s^2 N_\omega^2$ . Furthermore, this slow step is on the inner most loop so it is repeated many times to find self consistent values of  $\boldsymbol{\mu}$  and  $u_0$ . This leads to unacceptably slow convergence for any reasonable system size. Noting that the summation has the form of a convolution we can make use of FFT routines to calculate  $\rho_{\bar{\Phi}}$  with linear scaling in  $N_s N_\omega$ . This allows us to reach significantly larger systems and lower temperatures than would be possible by a direct approach. The next bottleneck in this flowchart is the calculation of the Hilbert transforms. These can also be made fast through a judicious use of FFT routines. Thus, by using this approach we obtain a scheme which can calculate the full frequency and momentum dependence of  $\rho_{\mathcal{G}}$  for lattices of substantial size,  $N_s \sim 2000$  at temperatures as low as 30K.

Is it useful to discuss the tolerances set on the Lagrange multipliers.  $\boldsymbol{\mu}$  is obtained to a relative precision of  $10^{-5}$ . This is significantly more accurate than is required to satisfy the particle sum rule to within a tenth of a percent. However, we find empirically that this strict convergence criterion for  $\boldsymbol{\mu}$  can not be satisfied until the spectral function  $\rho_{\bar{\Phi}}$  is also well converged. Thus, if  $\boldsymbol{\mu}$  has successfully converged to this tolerance we can be sure the  $\rho_{\bar{\Phi}}$  is also well converged. The convergence criterion on  $u_0$  requires the sum rule  $\sum \Psi(k)g(k) = \frac{n^2}{4}(1 - \frac{n}{2})$  to be satisfied to less than  $10^{-4}$ . Again, this is overkill as it concerns the particle density alone. However,  $u_0$  appears explicitly in  $\bar{\Phi}$  and  $\Psi$ . The chosen convergence criterion is such that the final  $u_0$  lands within .01t of the exact value. This range is comparable to the smallest scales in our calculation, namely the frequency resolution  $\Delta\omega$  and an implicit level broadening scale  $\eta \rightarrow \Delta\omega$ .

To exactly calculate the spectral functions, it is important to capture the entire range of the relevant frequencies. For non-interacting Fermions in 2 spatial dimensions this requires a frequency window no larger than 8t. However, the spectral functions  $\rho_{\bar{\Phi}}$  and  $\rho_{\Psi}$  have long tails

which extend to much higher frequency even though the renormalized QP bandwidth may be much narrower than the bare band. Thus it is important to determine empirically what range of frequency is sufficient to capture the full support of the spectral function. As mentioned before, we employ a frequency grid which extends over the range  $|\omega| < 15t$ , nearly four times the bare bandwidth, and find that this suffices to capture the support of all functions that arise.

### FAST FOURIER TRANSFORMS FOR EVALUATING CONVOLUTIONS

The use of FFTs vastly reduces the time taken to compute the frequency and momentum sums. Each term of  $\rho_{\bar{\Phi}}$  and  $\rho_{\Psi}$  is a convolution of 3  $\mathbf{g}$ 's and has a form which is very similar to the particle-hole bubble diagrams familiar from a second-order perturbation treatment of the Hubbard model in  $U/t$ .

$$\Sigma(k)_{2nd} \sim \sum_{pq} G(p)G(q)G(p+q-k); \quad \Sigma(i, j) \propto G(i, j)^2 G(j, i) \quad (\text{S.4})$$

where  $i = \vec{R}_i, \tau_i$  is a space time point. The convolution in Fourier space is a simple product in the space time domain and hence the real space version is advantageous. This is the well-known core idea of the FFT technique, where the time savings arise since the Fourier transforms are performed in  $N \log N$  steps rather than  $N^2$  (here  $N = N_s N_\omega$ ). The ECFL "self energies",  $\Psi$  and  $\Phi$  have the same frequency convolution structure of  $\Sigma(k)_{2nd}$  which appears only through the frequency arguments of  $\mathbf{g}$ . However, the  $\rho_{\bar{\Phi}}$  and  $\rho_{\Psi}$  equations suffer from the presence of momentum-dependent decorations which render them not technically convolutions. Nonetheless we can use FFT routines to solve these summations. The strategy is to break up the integral into elementary pieces that do have the form of a convolution. We then avoid the need to do one large integral with quadratic complexity by doing many ( $\approx 70$ ) small FFT's of linear complexity.

To accomplish this we define several  $\mathbf{g}\mathbf{g}$  correlation functions which are similar to particle-hole bubbles.

$$\begin{aligned} \chi_0(Q) &= \sum_q \mathbf{g}(q)\mathbf{g}(q+Q); \quad \chi_1(Q) = \sum_q \varepsilon_q \mathbf{g}(q)\mathbf{g}(q+Q) \\ \chi_2(Q) &= \sum_q \varepsilon_{q+Q} \mathbf{g}(q)\mathbf{g}(q+Q); \quad \chi_3(Q) = \sum_q \varepsilon_{q+Q}^2 \mathbf{g}(q)\mathbf{g}(q+Q) \end{aligned}$$

each of which is a convolution in both frequency and momentum with a spectral function which can be calculated by FFT in linear time.

With these correlation functions every  $\mathbf{g}\mathbf{g}\mathbf{g}$  term (except one to be discussed later) found in  $\Psi$  and  $\bar{\Phi}$  can be written in the form

$$B_{\mathbf{g}\mathbf{g}\mathbf{g}}(k) = F_1(k) \sum_p F_2(p)\mathbf{g}(p)\chi_n(p-k)F_3(p-k). \quad (\text{S.5})$$

where  $F_1, F_2$ , and  $F_3$  are each functions of momentum only and their arguments are carefully matched with the arguments of  $B_{\mathbf{g}\mathbf{g}\mathbf{g}}$ ,  $\mathbf{g}$ , and  $\chi_n$  as they appear in the integral such that all factors fit the form of a convolution in both momentum and frequency. In this way we can massage every term of  $\rho_{\overline{\mathbf{g}}}$  into a convolution of one  $\mathbf{g}$  and a  $\chi_n$  rather than three  $\mathbf{g}$ 's as originally written. There is one term in this problem which cannot be treated in this way because the argument matching of Eq. (S.5) cannot be achieved in such a simple way. This problem term looks like

$$\overline{\Phi}_{JJ}(k) = \sum_{pq} J_{q-k} J_{p-k} \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k).$$

Nonetheless, this term can be treated by the FFT approach if the factor  $J_{q-k}$  is broken up using angle addition identities. This is accomplished without difficulty because the locality of  $J_{ij}$  ensures that  $J_k$  is composed of a small number of Fourier components.

In defining the Fourier transforms, we need to extend the frequency functions to infinity, since it is only then that the frequency convolutions become products in the time domain. Recall that our frequency integrals have been discretized onto  $N_\omega$  frequency bins which cover the support of our spectral functions. In extending the dis-

cretized frequency summations to infinity, we follow the standard procedure of padding the  $N_\omega$  frequency bins with an equal number of frequency bins with value zero. By a simple exercise one can verify that padding finite data in this way allows an application of the periodic FFT in such a way that the result of the infinite transform is reproduced. No such considerations are required for the momentum sums which are by definition periodic and discrete, making them naturally suited to treatment by FFT.

The Hilbert transform is formally a convolution and can therefore be solved with the advantages of the FFT routines. Once again, however, we face the problem that this convolution is a non-periodic frequency integral. Furthermore, the Hilbert kernel  $\frac{1}{\omega}$ , unlike other spectral functions with a compact support, falls off very slowly at large frequencies so the padding trick from the  $\mathbf{g}\chi$  convolutions will not work well in this case. It is found that the use of FFT's to calculate a Hilbert transform will always introduce some error. Fortunately, this error can be controlled by increasing the length of padding used. In our code we use a frequency padding of  $8N_\omega$  for the Hilbert transforms. This relegates the error of the real parts of the various functions to very high frequency, far beyond the compact support of the spectral functions. The error introduced is therefore negligible.

## Electronic spectral properties of the two-dimensional infinite- $U$ Hubbard model

Ehsan Khatami,<sup>1,2</sup> Daniel Hansen,<sup>1</sup> Edward Perepelitsky,<sup>1</sup> Marcos Rigol,<sup>3</sup> and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, California 95064, USA*

<sup>2</sup>*Department of Physics, Georgetown University, Washington D.C., 20057 USA*

<sup>3</sup>*Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

(Received 11 March 2013; published 29 April 2013)

A strong-coupling series expansion for the Green's function and the extremely correlated Fermi liquid (ECFL) theory are used to calculate the moments of the electronic spectral functions of the infinite- $U$  Hubbard model. Results from these two complementary methods agree very well at both low densities, where the ECFL solution is the most accurate, and at high to intermediate temperatures, where the series converge. We find that a modified first moment, which underestimates the contributions from the occupied states and is accessible in the series through the time-dependent Green's function, best describes the peak location of the spectral function in the strongly correlated regime. This is examined by the ECFL results at low temperatures, where it is shown that the spectral function is largely skewed towards the occupied states.

DOI: [10.1103/PhysRevB.87.161120](https://doi.org/10.1103/PhysRevB.87.161120)

PACS number(s): 71.10.Fd

### I. INTRODUCTION

A long-standing theme in the dynamics of strongly interacting systems is the reconstruction of dynamics from the knowledge of the first few moments.<sup>1</sup> Its appeal lies in the relative ease with which these moments can be computed, in contrast to computing the complete dynamical correlation functions. The method of moments works well in cases where the qualitative features of the correlation functions are somewhat understood by other arguments, including conservation laws in the case of spin dynamics. In the important problem of the strong-coupling Hubbard model, the moments are dominated by the energy scale  $U$ ,<sup>2</sup> the on-site repulsive Coulomb interaction, and hence rendered useless. In contrast, for the  $t$ - $J$  model embodying extreme correlations, i.e.,  $U \rightarrow \infty$  at the very outset, a better prospect exists. The moments are blind to the scale of  $U$ , since it does not occur in the Hamiltonian, and therefore one expects them to be meaningful in determining the broad features of the dynamics. With this in mind, we study a simple version of the  $t$ - $J$  model by focusing on  $J = 0$ , which is identical to the  $U = \infty$  Hubbard model, thereby making more tools available for the analysis. As we show in what follows, we have developed the capability to compute the moments of the electron spectral function of this model by utilizing series expansions.<sup>3,4</sup> Experiments using angle-resolved photoemission spectroscopy (ARPES)<sup>5-8</sup> directly measure this spectral function, providing an added impetus.

An independent source of information about the electronic spectral function is the recent analytical theory of extremely correlated Fermi liquids (ECFL). This theory has been developed in recent publications,<sup>9,10</sup> and several results of the model pertaining to the detailed line shapes find close agreement with experiment.<sup>5</sup> On the calculational front, the theory provides a systematic methodology for computation, and the initial low order implementation yields the single-electron spectral function for particle densities in the range  $0 \leq n \lesssim 0.7$ . The line shapes of this calculation for  $n \geq 0.5$  display a characteristic skewed shape found in the experimental curves in ARPES, as detailed in Ref. 10. The computed spectra are available at any temperature (high or low), and the only limitation at present

is the inability to access the regime close to half filling with density greater than  $n \sim 0.75$ . Given the inherent complexity of the newly developed ECFL formalism, the possibility of an objective cross-check using series expansions is a very attractive one, and here we provide a comparison.

We compute and compare the moments of the  $t$ - $J$  model with  $J = 0$  in two dimensions by utilizing a series expansion<sup>11</sup> and the ECFL theory. The two techniques are largely complementary. While they individually run into difficulties in different regimes, namely, at low temperatures for the series expansion and high densities for the ECFL, there is sufficient overlap in densities and temperatures where *both methods* give reliable results. This provides us with a unique opportunity to test the validity of the answers. For ECFL, this provides a stringent test of the resulting moments by comparing with the series expansion. For the series expansion, the availability of an analytical theory and hence, of the entire spectrum, is of great advantage in interpreting the distinctions between three types of moments that can be computed [see Eq. (7) below]. We find that especially at high densities, the line shape of the spectral function is skewed towards occupied energies,  $\omega \leq 0$ , therefore the spectral peak (SP) location (the maximum location in the energy distributed curves) is best estimated by the first moment of a modified function with dominant contribution from unoccupied states.

In the rest of this Rapid Communication, we first explain how the series expansion and ECFL results are obtained (Sec. II). In Sec. III, we compare the results from the two methods, and discuss our findings. A summary follows in Sec. IV.

### II. PRELIMINARIES

#### A. Definitions of computed coefficients

We denote the imaginary-time Green's function for the  $U = \infty$  Hubbard model, or equivalently, the  $t$ - $J$  model with  $J = 0$ , as  $\mathcal{G}(i, \tau_i; j, \tau_j) = -\langle T_\tau \hat{C}_{i\sigma}(\tau_i) \hat{C}_{j\sigma}^\dagger(\tau_j) \rangle$ , where  $T_\tau$  is the time-ordering operator and  $\langle \cdot \rangle$  denotes the thermal expectation value. We thus study the limit of extreme correlations. The operators are Gutzwiller-projected Fermi objects and related to the Hubbard  $X$  operators as  $\hat{C}_{i\sigma} \equiv X_i^{0\sigma}$ , etc. As usual,<sup>12</sup>

this object is a function of the time difference  $\tau \equiv \tau_i - \tau_j$ , and we will study its spatial Fourier transform  $\mathcal{G}(k, \tau)$ . Our study begins with the following expansions:

$$\mathcal{G}(k, \tau > 0) = (-1) \sum_{m=0}^{\infty} (-1)^m \frac{\tau^m}{m!} a_m(k), \quad (1)$$

$$\mathcal{G}(k, \tau < 0) = \sum_{m=0}^{\infty} (-1)^m \frac{\tau^m}{m!} b_m(k), \quad (2)$$

where the coefficients  $a_m$  are computed analytically as a series in the hopping amplitude  $t$ . The series expansion can be carried out to the fourth order by hand,<sup>13</sup> and pushed to the eighth order by a highly efficient computer program<sup>11</sup> based on Metzner's linked-cluster formalism.<sup>14</sup> This order is the limit achievable by currently available supercomputers. Using antiperiodic boundary conditions,  $\mathcal{G}(\tau - \beta) = -\mathcal{G}(\tau)$ , we obtain Eq. (2) from Eq. (1). Here  $\beta = 1/(k_B T)$  is the inverse temperature (we set  $t = 1$  as the unit of energy, and  $k_B = 1$ ). Therefore, the main calculation focuses on Eq. (1). Its Fourier series in Matsubara frequencies,  $\omega_n = (2n + 1)\pi/\beta$ , is obtained from  $\mathcal{G}(k, i\omega_n) = \int_0^\beta e^{i\omega_n \tau} \mathcal{G}(k, \tau) d\tau$ . The spectral function at momentum  $k$  and for the real frequency  $\nu$  is denoted by  $\rho_G(k, \nu)$  and determines the Green's function through the relation  $\mathcal{G}(k, i\omega_n) = \int_{-\infty}^{+\infty} \frac{\rho_G(k, \nu)}{i\omega_n - \nu} d\nu$ . At high frequencies  $\omega_n$ , we have an expansion

$$\mathcal{G}(k, i\omega_n) = \sum_{m=0}^{\infty} \frac{c_m(k)}{(i\omega_n)^{m+1}},$$

involving the ‘‘symmetric’’ coefficient,  $c_m(k)$  (see below). The time domain Green's function is also given in terms of the spectral function by the important representation

$$\mathcal{G}(k, \tau) = \int_{-\infty}^{+\infty} d\nu \rho_G(k, \nu) e^{-\nu \tau} [\Theta(-\tau) f(\nu) - \Theta(\tau) \bar{f}(\nu)], \quad (3)$$

where

$$f(\nu) = \frac{1}{1 + e^{\beta \nu}} \quad \text{and} \quad \bar{f}(\nu) = \frac{1}{1 + e^{-\beta \nu}}. \quad (4)$$

The three sets of coefficients  $\alpha_m$  (i.e.,  $a_m$ ,  $b_m$ , and  $c_m$ ) are easily seen to originate from the spectral function convoluted by a different filter function  $\chi(\nu)$  [respectively,  $\bar{f}(\nu)$ ,  $f(\nu)$ , 1] as

$$\alpha_m(k) = \int_{-\infty}^{\infty} \nu^m \chi(\nu) \rho_G(k, \nu) d\nu. \quad (5)$$

Using this and the identity  $f + \bar{f} = 1$ , we see that the symmetric coefficients satisfy the important relation

$$c_m(k) = a_m(k) + b_m(k). \quad (6)$$

### B. Definition of moments

Equation (5) gives the power integrals of the effective spectral function  $\chi(\nu) \rho_G(\nu)$ , and naturally leads to three sets of moments at each  $k$ ,  $\varepsilon_m^X(k) = \alpha_m(k)/\alpha_0(k)$ . Thus, the moments can be obtained from the coefficients  $a_m, b_m, c_m$ , and contain complementary information as we discuss below. We assign them suggestive names

$$\varepsilon_m^>(k) = \frac{a_m(k)}{a_0(k)}, \quad \varepsilon_m^<(k) = \frac{b_m(k)}{b_0(k)}, \quad \varepsilon_m^0(k) = \frac{c_m(k)}{c_0(k)}, \quad (7)$$

the *greater*, *lesser*, and *symmetric* moments, respectively.<sup>15</sup> The superscripts in the notations  $\varepsilon^>$  and  $\varepsilon^<$  signify that the contribution to these energy moments comes predominantly from the weight of the spectral function that lies *above* or *below* the chemical potential, and hence the unoccupied or occupied states. The coefficients at  $m = 0$  have special meanings: By computing the anticommutator of  $\hat{C}$  and  $\hat{C}^\dagger$ , and taking its average we find  $c_0(k) \equiv c_0 = 1 - \frac{n}{2}$  in this model. The coefficient  $b_0(k)$  is also the momentum distribution function,

$$m_\sigma(k) = \langle \hat{C}_{k\sigma}^\dagger \hat{C}_{k\sigma} \rangle = b_0(k). \quad (8)$$

Using Eq. (6), we find  $a_0(k) = 1 - \frac{n}{2} - m(k)$ .

In this work, we study only the first moments, i.e.,  $m = 1$ . We argue below that these give an estimate of the quasiparticle spectrum for a given  $k$ . It is particularly useful to study all three moments separately since they exhibit different behavior, and the comparison with the spectra of ECFL gives a clearer understanding of their differences, as we discuss below.

### C. Summary of relevant ECFL results

In Ref. 10, the formalism of ECFL for general  $J$  is implemented to second order in the variable  $\lambda$ , which is closely related to the density. A self-consistent argument indicates that the calculation in Ref. 10 is valid for densities  $n \lesssim 0.7$ . It has no limitation on the temperature or system size, since it is essentially an analytical theory—resembling the skeleton graph expansion theories of standard models in structure. We note that the ECFL assumes a specific type of Fermi liquid with strong asymmetric corrections,<sup>9</sup> and the reasonable similarity to the series data, as we will see in Sec. III, suggests that this conclusion is fairly safe, at least for high enough temperatures. At low temperatures, there could be other instabilities that are hard to capture with the series analysis, and the present versions of the ECFL.

The full spectral function  $\rho_G(k, \nu)$  is computed and its moments (for the case of  $J = 0$ ) are readily available for comparison with those from the series expansion. Also available in this work is the location of the SPs  $\varepsilon^{\text{SP}}(k)$ , when they exist, the momentum distribution function, etc. It is therefore possible to compute various dispersion curves, relating the different characteristic energies (i.e., moments) to wave vectors, and to compare them with the true SP dispersion. The benchmarking of these moments provides us with valuable insight for interpreting the series data, where the SPs are not available, but the moments are.

## III. RESULTS

In Fig. 1, we plot the symmetric first moment  $\varepsilon_1^0(k)$  as a function of momentum at  $T = 0.77$  for five different densities  $n = 0.2, 0.5, 0.7, 0.8$ , and  $0.9$ . We find excellent agreement between the results from the series and the ECFL for  $n = 0.2$  for all the momenta around the irreducible wedge of the Brillouin zone. At higher densities up to  $n = 0.7$  (beyond which the ECFL results are not quoted), the agreement is still very good, except around the zone corner, where the disagreement grows as the density increases.<sup>18</sup> The results for the series are obtained from Padé approximations as the bare results show divergent behavior at  $T < 1$ . The number of terms

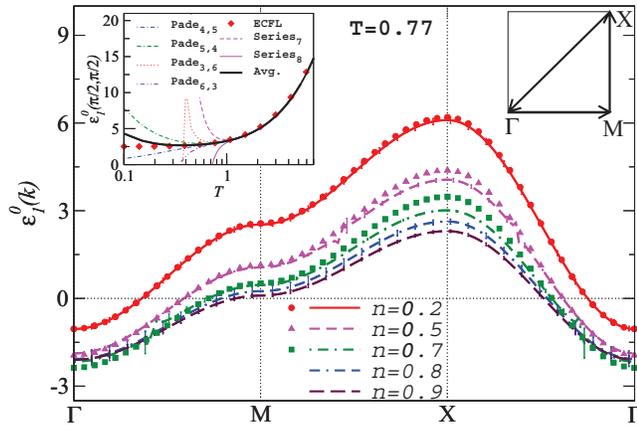


FIG. 1. (Color online) The first symmetric moment  $\varepsilon_1^0(k)$  at  $T = 0.77$  vs momentum around the irreducible wedge of the Brillouin zone (the path is shown in the right inset). Lines and symbols for  $n \leq 0.7$  are from ECFL calculations. Left inset:  $\varepsilon_1^0(k)$  for  $n = 0.2$  at  $k = (\pi/2, \pi/2)$  from the ECFL (diamonds), up to orders seven and eight of the series (labeled Series<sub>7</sub> and Series<sub>8</sub>), and up to the eighth order after various Padé approximations, vs temperature on a logarithmic scale. The numbers in the subscripts of “Padé” labels represent the order of the polynomial in the numerator and in the denominator of the Padé ratio, respectively. “Avg.” denotes the average between Padé<sub>(4,5)</sub> and Padé<sub>(5,4)</sub>. In the main panel, the results for the series are either the average between Padé<sub>(4,5)</sub> and Padé<sub>(5,4)</sub> or Padé<sub>(5,5)</sub> and Padé<sub>(5,4)</sub> with the “error bars” defined as the differences between the two.<sup>17</sup>

in the series is large enough to justify the utilization of Padé approximations in order to extend the convergence to lower temperatures. A comparison of several of these approximations with the ECFL results for a (low) density of  $n = 0.2$  is shown in the inset of Fig. 1. In that case, we see that the agreement between the two methods extends to temperatures as low as  $T = 0.3$  using Padé approximations.

The greater moment  $\varepsilon_1^>(k)$  is plotted in Fig. 2(a) at the same temperature and densities as in Fig. 1. For  $\varepsilon_1^>(k)$ , the overall agreement between the series expansions and the ECFL results for all  $n \leq 0.7$  is better than for  $\varepsilon_1^0(k)$ , especially around the  $X$  point. We also note that  $\varepsilon_1^>(k)$  exhibits a more intriguing behavior than  $\varepsilon_1^0(k)$ . One of the prominent features of the former, seen in Fig. 2(a), is the significant narrowing of the band by increasing the density. In Fig. 2(b), we plot the bandwidth [i.e.,  $\max(\varepsilon_1^>) - \min(\varepsilon_1^>)$ ] from the series as a function of density at  $T = 1.52, 1.00$ , and  $0.77$ . It appears that the bandwidth deviates from a linear dependence on  $n$  by decreasing the temperature, and saturates for  $n \rightarrow 1$  at a nonzero value that decreases towards zero with decreasing  $T$ . Close to  $n = 1$  at  $T = 0.77$ , we find a weaker agreement between different Padé approximations, leading to larger error bars. The version of ECFL in Ref. 10 cannot be used to study this effect as the high-density region  $n \sim 1$  is beyond its regime of validity.

Another interesting feature of  $\varepsilon_1^>(k)$  [Fig. 2(a)] is the change in sign of its slope near the  $\Gamma$  point as the density increases towards unity. To better study this feature, in Fig. 2(c), we report only the results along the nodal direction. We find that for  $n \gtrsim 0.7$ , the greater moment initially decreases as

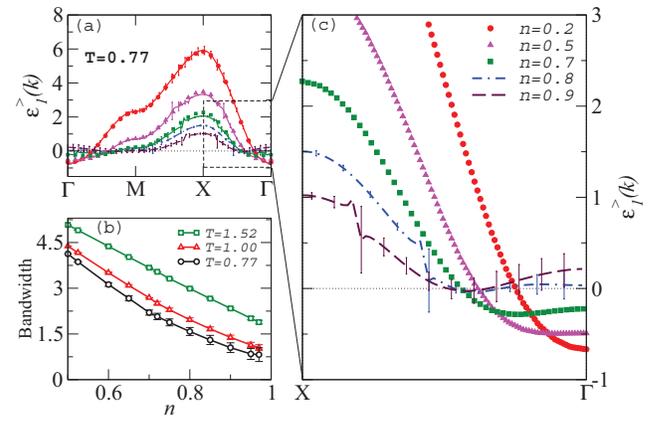


FIG. 2. (Color online) (a) The first greater moment  $\varepsilon_1^>(k)$  at  $T = 0.77$  vs momentum for the same path around the irreducible wedge of the Brillouin zone as in Fig. 1. Lines and symbols are also the same as in Fig. 1. (b) The bandwidth of  $\varepsilon_1^>(k)$ , defined as the difference between its maximum and minimum values at momenta shown in panel (a), vs density for  $T = 1.52, 1.00$ , and  $0.77$ . Panel (c) zooms in the results in panel (a) for  $k$  along the nodal direction. The two methods more or less agree with each other, within the error bars, in this window for  $n \leq 0.7$ , and therefore, we show only the ECFL results for the latter cases.

the momentum increases from zero, leading to a negative curvature, or effective mass, at the  $\Gamma$  point. This feature becomes more pronounced as we increase the density, or decrease the temperature (see Fig. 3). These results hint at a possible reconstruction of the Fermi surface, i.e., the negative mass persisting and extending in  $k$  space so as to reach the Fermi momentum. The appearance of such a hole pocket in the (hole) underdoped regime, could be of interest in ARPES

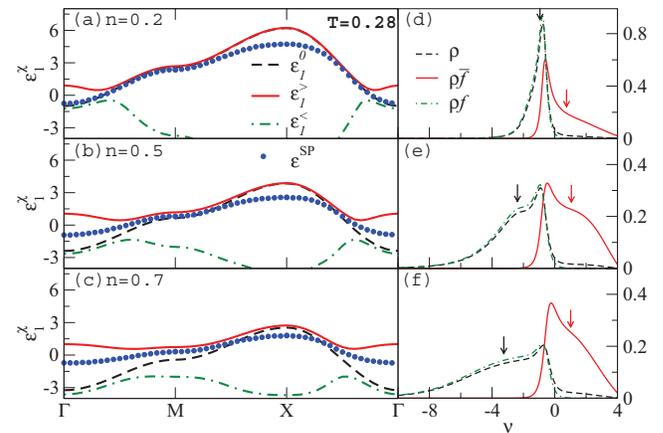


FIG. 3. (Color online) Comparison of the SP location  $\varepsilon_1^{\text{SP}}(k)$  (symbols) and the three moments from ECFL at  $T = 0.28$  and for (a)  $n = 0.2$ , (b)  $n = 0.5$ , and (c)  $n = 0.7$ . Right panels show the corresponding spectral functions and their products to  $\bar{f}(\omega)$  and  $f(\omega)$  at  $\Gamma$  for the same densities shown in the left panels. Dark (light) arrows show the values of  $\varepsilon_1^0$  ( $\varepsilon_1^>$ ). At low densities, the SP location is estimated well by the first symmetric moment. At higher density, the spectral function is skewed and the greater moment, which is calculated for the spectral function after most of its weight in the negative frequency region is cut off, provides a better estimate.

and quantum oscillation studies. However, establishing this firmly requires higher order terms in the series, and is therefore difficult.

So far, we have seen that for intermediate temperatures and at relatively small densities, the ECFL agrees extremely well with the results of the series expansion. But, unlike the series expansion, ECFL is not limited to high temperatures at those densities and can be used to study the moments, and more importantly, the real-frequency spectral functions, at much lower temperatures. Therefore, we focus on the ECFL results at  $n = 0.2, 0.5$ , and  $0.7$ , and at a reduced temperature of  $T = 0.28$ , a temperature at which the series do not converge. In Figs. 3(a)–3(c), we plot  $\varepsilon_1^0(k)$ ,  $\varepsilon_1^>(k)$ , and  $\varepsilon_1^<(k)$  from the ECFL, along with  $\varepsilon^{\text{SP}}(k)$ , obtained from the spectral functions, at different momenta. We find that in the physically interesting region of low temperatures and high densities, where correlation effects are strongest, the location of the SP is generally better estimated by the greater moment than by the symmetric, or the lesser one [see Fig. 3(c)].

The spectral functions shown in Figs. 3(d)–3(f) help us understand why this is the case. There, we plot the spectral functions  $\rho_G(k, \omega)$ ,  $\rho_G(k, \omega)\bar{f}(\omega)$ , and  $\rho_G(k, \omega)f(\omega)$ , corresponding to the three moments at  $k = (0, 0)$ , where the differences between the moments are the most pronounced, vs frequency. At  $n = 0.2$ , there exists a relatively sharp quasiparticle peak in  $\rho_G$  whose location matches the first symmetric moment (marked by a dark arrow) very well.  $\varepsilon_1^>(k)$ , on the other hand, falls slightly to the right of the quasiparticle peak (marked by a light-colored arrow) as most of the spectral weight in negative frequencies is cut off after multiplying  $\rho_G$  by  $\bar{f}(\omega)$  [see Eq. (5)]. Also, since there is very little spectral weight in the positive frequency side,  $\varepsilon_1^<(k)$  is very close in value to  $\varepsilon_1^0(k)$ . As the density is increased to  $n = 0.5$ , the spectral function is skewed as a result of correlations. In this case, at small  $k$ , there is much more spectral weight on the left of the SP than on the right, causing the symmetric moment to be smaller than  $\varepsilon^{\text{SP}}(k)$ . This feature becomes more

significant at a higher density of  $n = 0.7$ , where almost all of the spectral weight is in the negative frequency side. As a result, multiplying  $\rho_G$  by  $\bar{f}(\omega)$  helps in neglecting the excess weight on the left side of the SP. Hence,  $\varepsilon_1^>(k)$ , which is readily available from the series at even higher densities, may be used as an indicator of  $\varepsilon^{\text{SP}}(k)$  using this insight from the ECFL spectra.

#### IV. SUMMARY

We employ two complementary methods, namely, a strong-coupling series expansion and the ECFL, to calculate the moments of the spectral functions for the infinite- $U$  Hubbard model. Unveiling the basic physics of the model is benefited by the complementarity of those approaches. Furthermore, the series expansion results provide the first independent check of the ECFL theory, which has been self-consistently established. At intermediate temperatures and low densities, where the results from both methods are available, we find very good agreement between the two. Unlike ECFL, the series is not limited to small densities and, by increasing the density in the series to near half filling, we find interesting features in the dispersion of the moment with dominant contributions from unoccupied states (the greater moment). These include a significant narrowing of its band as well as hints of Fermi-surface reconstruction. Unlike the series, the ECFL is not limited to high temperatures and, by exploring the ECFL results at lower temperatures, we find that the greater moment better describes the location of the SP as the density increases. This is understood based on the skewing of the spectral functions in the negative frequency region in the strongly correlated regime.

#### ACKNOWLEDGMENTS

This work was supported by DOE under Grant No. FG02-06ER46319 (B.S.S., D.H., and E.P.), and by NSF under Grant No. OCI-0904597 (E.K. and M.R.).

<sup>1</sup>H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965); **34**, 399 (1965); M. Dupuis, *ibid.* **37**, 502 (1967); K. Tomita and H. Mashiyama, *ibid.* **51**, 1312 (1974).

<sup>2</sup>W. Nolting, *Z. Phys.* **255**, 25 (1972); J. J. Deisz, D. W. Hess, and J. W. Serene, *Phys. Rev. B* **66**, 014539 (2002).

<sup>3</sup>*Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London, 1974), Vol. 3.

<sup>4</sup>W. Brauneck, *Z. Phys. B* **28**, 291 (1977); K. Kubo, *Prog. Theor. Phys.* **64**, 758 (1980), <http://ptp.oxfordjournals.org/content/64/3/758.abstract>; R. R. P. Singh and R. L. Glenister, *Phys. Rev. B* **46**, 14313 (1992); **46**, 11871 (1992); W. O. Putikka, M. U. Luchini, and T. M. Rice, *Phys. Rev. Lett.* **68**, 538 (1992); W. O. Putikka, M. U. Luchini, and R. R. P. Singh, *ibid.* **81**, 2966 (1998).

<sup>5</sup>G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011); K. Matsuyama and G. H. Gweon, [arXiv:1212.0299](https://arxiv.org/abs/1212.0299).

<sup>6</sup>T. Valla, A. V. Fedorov, P. D. Johnson, B. O. Wells, S. L. Hulbert, Q. Li, G. D. Gu, and N. Koshizuka, *Science* **285**, 2110 (1999).

<sup>7</sup>C. G. Olson, R. Liu, D. W. Lynch, R. S. List, A. J. Arko, B. W. Veal, Y. C. Chang, P. Z. Jiang, and A. P. Paulikas, *Phys. Rev. B* **42**, 381 (1990); T. Yoshida *et al.*, *J. Phys. Condens. Matter* **19**, 125209 (2007).

<sup>8</sup>A. Kaminski *et al.*, *Phys. Rev. B* **69**, 212509 (2004)

<sup>9</sup>B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011); **108**, 029702 (2012); *Phys. Rev. B* **84**, 165112 (2011); **87**, 125124 (2013); *Phys. Rev. Lett.* **109**, 067004 (2012); *Phys. Rev. B* **81**, 045121 (2010).

<sup>10</sup>D. Hansen and B. S. Shastry, [arXiv:1211.0594](https://arxiv.org/abs/1211.0594).

<sup>11</sup>E. Khatami, E. Perepelitsky, B. S. Shastry, and M. Rigol (unpublished).

<sup>12</sup>A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

<sup>13</sup>E. Perepelitsky and B. S. Shastry (unpublished).

<sup>14</sup>W. Metzner, *Phys. Rev. B* **43**, 8549 (1991).

<sup>15</sup>B. Farid, *Philos. Mag.* **84**, 909 (2004).

<sup>16</sup>Note that the zeroth-order terms in the expansion for  $\varepsilon_1^0(k)$  is proportional to  $\beta^{-1}$ . Hence,  $\text{Padé}_{(l,m)}$  with  $l + m = 9$  will match the

series order by order exactly up to the eighth order. Nevertheless, we find that  $\text{Padé}_{\{5,5\}}$ , for which one assumes that the coefficient of the ninth-order term in the series is zero, often results in less spurious features than with  $\text{Padé}_{\{4,5\}}$ , and therefore is used instead of the latter for  $n = 0.2$  and  $0.9$ . On the other hand, since  $\varepsilon_1^>(k)$  is itself a ratio of two polynomials, either of the above two Padé approximants is equally valid. In this case (Fig. 2), we use the average of  $\text{Padé}_{\{5,4\}}$  and  $\text{Padé}_{\{5,5\}}$  for  $n = 0.8$  and  $\text{Padé}_{\{5,4\}}$  and  $\text{Padé}_{\{4,5\}}$  for the rest.

<sup>17</sup>There is no error per se in the calculation of the coefficients of terms in the series. The so-called error bars are merely a measure of the

convergence limit for the Padé approximations at low temperatures, where the bare results show divergent behavior. They do not represent statistical or particular systematic errors.

<sup>18</sup>We may take the curves of  $\varepsilon_1^0(k)$ , or more accurately,  $\varepsilon_1^>(k)$  as estimates of the SP dispersion  $\varepsilon^{\text{SP}}(k)$ , after shifting them by a constant chosen to pass them through zero energy at the Fermi momentum (as in Figs. 1 and 2). The magnitudes of the shift constants are on the scale seen in Figs. 3(d)–3(f) as the separation between the peak locating the  $\varepsilon^{\text{SP}}(k)$  and the arrows locating the moments.

# Extremely correlated Fermi liquid study of the $U = \infty$ Anderson impurity model

B. Sriram Shastry and Edward Perepelitsky

*Physics Department, University of California, Santa Cruz, California 95064, USA*

Alex C. Hewson

*Department of Mathematics, Imperial College, 180 Queen's Gate, London, SW7 2BZ, United Kingdom*

(Received 12 July 2013; published 12 November 2013)

We apply the recently developed *extremely correlated Fermi liquid* (ECFL) theory to the Anderson impurity model, in the extreme correlation limit  $U \rightarrow \infty$ . We develop an expansion in a parameter  $\lambda$ , related to  $n_d$ , the average occupation of the localized orbital, and find analytic expressions for the Green's functions to  $O(\lambda^2)$ . These yield the impurity spectral function and also the self-energy  $\Sigma(\omega)$  in terms of the two self-energies of the ECFL formalism. The imaginary parts of the latter have roughly symmetric low-energy behavior ( $\propto \omega^2$ ), as predicted by Fermi liquid theory. However, the inferred impurity self-energy  $\Sigma''(\omega)$  develops asymmetric corrections near  $n_d \rightarrow 1$ , leading in turn to a strongly asymmetric impurity spectral function with a skew towards the occupied states. Within this approximation, the Friedel sum rule is satisfied but we overestimate the quasiparticle weight  $z$  relative to the known exact results, resulting in an overbroadening of the Kondo peak. Upon scaling the frequency by the quasiparticle weight  $z$ , the spectrum is found to be in reasonable agreement with numerical renormalization group results over a wide range of densities.

DOI: [10.1103/PhysRevB.88.205108](https://doi.org/10.1103/PhysRevB.88.205108)

PACS number(s): 71.10.Fd, 75.30.Mb, 71.28.+d

## I. INTRODUCTION AND MOTIVATION

The *extremely correlated Fermi liquid* (ECFL) theory has been recently developed to understand the physics of correlations in the limit of infinite  $U$  and applied to the  $t$ - $J$  model in Refs. 1 and 2. Here, we apply the ECFL theory to the problem of the spin- $\frac{1}{2}$  Anderson impurity model (AIM) at  $U = \infty$ . The ECFL theory is based on a systematic expansion of the formally exact Schwinger equations of motion of the model for the (Gutzwiller) projected electrons in powers of a parameter  $\lambda$ . This parameter is argued to be related to  $n$  the density of particles in the  $t$ - $J$  model, and in the same spirit, to  $n_d$  the average impurity level occupancy in the Anderson model considered here. Thus, at low enough densities of particles, the complete description of the system, including its dynamics, is expected in *quantitative* terms, with just a few terms in the  $\lambda$  expansion. Presently, the theory to  $O(\lambda^2)$  has been evaluated for the  $t$ - $J$  model,<sup>2</sup> and higher-order calculations in  $\lambda$  valid up to higher densities could be carried out in principle. We thus envisage systematically cranking up the density from the dilute limit, until we hit singularities arising from phase transitions near  $n \sim 1$ .<sup>3</sup> This represents a possible road map for solving one of the hard problems of condensed matter physics and is exciting for that reason.

We apply the ECFL theory equations to  $O(\lambda^2)$  to the AIM model in this work. This problem was introduced by Anderson<sup>4</sup> in 1961, and has been a fertile ground where several fruitful ideas and powerful techniques have been developed, and tested against experiments in Kondo, mixed valency, and heavy-fermion systems. These include the renormalization group ideas, from the intuitive poor man's scaling of Anderson<sup>5,6</sup> to the powerful numerical renormalization group (NRG) of Wilson,<sup>7</sup> Krishna-murthy *et al.*,<sup>8</sup> and more recent work in Refs. 9 and 10. A comprehensive review of the AIM and many popular techniques used to study it, such as the large- $N$  expansion,<sup>11,12</sup> slave particles,<sup>13</sup> and the Bethe *ansatz*,<sup>14</sup> can be found in Ref. 15. In the AIM, the Wilson renormalization

group method provides an essentially exact solution of the crossover from weak to strong coupling, without any intervening singularity in the coupling constant. As emphasized in Refs. 16–18, the ground state is asymptotically a Fermi liquid at all densities. This implies that as a function of the density  $n_d$  (at any  $U$ ), the Fermi liquid ground state evolves smoothly without encountering any singularity, from the low-density limit (the empty orbital limit) to the intermediate-density limit (the mixed valent regime), and finally through to the very high-density limit (Kondo regime). In view of the nonsingular evolution in density, the AIM provides us with an ideal problem to benchmark the basic ECFL ideas discussed above.

The current understanding of the AIM model from Refs. 8, 16, and 17 is that Fermi liquid ground state and its attendant excitation spectrum are reached in the asymptotic sense, i.e., at low enough energies and  $T$ . Our present study of this model is somewhat broader. We wish to understand the excitations of the model in an enlarged region, in order to additionally obtain an estimate of the magnitude of corrections to the asymptotic behavior. To motivate this remark, note that the ECFL formalism yields an asymmetry in the excitations and the spectral functions of the  $t$ - $J$  model for sufficiently high densities, with a pronounced skew towards  $\omega < 0$ , arising fundamentally from Gutzwiller projection. This skew can be interpreted as an asymmetric correction to the leading particle-hole-symmetric excitation spectrum of that model<sup>19</sup> [e.g., corrections to  $\Sigma''(\omega) \sim \{\omega^2 + (\pi k_B T)^2\}$  behavior of the Fermi liquid of the form  $\Sigma''(\omega) \sim \omega^3$ ]. Such corrections have been argued to be of central importance in explaining the anomalous line shapes in the angle-resolved photoemission spectra of high- $T_c$  superconductors in the normal state.<sup>19,20</sup> Therefore, it is useful and important to understand the line shape and self-energy asymmetry in controlled calculations of the Anderson model with infinite  $U$ , which shares the local Gutzwiller constraint with the  $t$ - $J$  model on a lattice. A necessary condition for substantial asymmetry of the type

seen in ECFL at  $U = \infty$  appears to be a large  $U$ , and hence is difficult to find from a perturbative expansion in  $U$  of the type pioneered in Ref. 16. The study of the infinite- $U$  limit of the AIM is therefore particularly interesting in the present context. AIM studies of the spectral functions<sup>21–24</sup> using NRG have become available in recent years. We will compare our results with some of these calculations later in this paper.

In this paper, we use the ECFL machinery<sup>2</sup> to obtain the exact Schwinger equation of motion for the  $d$ -electron Green's function and represent it in terms of two self-energies. These are further expanded in a series in the parameter  $\lambda$  mentioned above, and the equations to second order are arrived at. These involve a second chemical potential  $u_0$  that contributes to a shift in the location of the localized energy level, bringing it closer to the chemical potential of the conduction electrons. The rationale for introducing this second chemical potential is similar to that in the  $t$ - $J$  model; the AIM possesses a shift invariance identified in Eq. (11). Maintaining this invariance to different orders in  $\lambda$  is possible only if we introduce  $u_0$ . The second-order equations are studied numerically, and the solution for the spectral function is compared with the NRG results.

Since we expect some readers to be interested in the AIM more than in the  $t$ - $J$  model, we provide a fairly self-contained description of the ECFL method used here for the AIM. In this spirit, let us note that a direct interpretation of the parameter  $\lambda$  as a partial projector can be found from a simple calculation of the atomic limit, with the parameter  $\lambda$  thrown in. Reference 2 (Appendix A1 and especially Fig. 6) explicitly shows that the double occupancy goes from its maximum to zero as  $\lambda$  varies from 0 to 1. Further, it is useful to view the  $\lambda$  parameter in an operator sense, by writing a partially projected ( $d$ -orbital) fermion operator  $\hat{f}_\sigma^\dagger(\lambda) = (1 - \lambda f_{\bar{\sigma}}^\dagger f_{\bar{\sigma}}) f_\sigma^\dagger$  and its adjoint (here  $\bar{\sigma} = -\sigma$ ). The operator  $\hat{f}_\sigma^\dagger(\lambda)$  interpolates between the unprojected Fermi operator  $f_\sigma^\dagger$  at  $\lambda = 0$ , and the Gutzwiller projected Hubbard operator  $X_i^{\sigma 0}$  at  $\lambda = 1$ . The Hamiltonian is written in terms of  $\hat{f}_\sigma^\dagger(\lambda)$ ,  $\hat{f}_\sigma(\lambda)$ , and expanding in  $\lambda$  gives an effective Hamiltonian that generates the auxiliary Green's function  $\mathbf{g}$  below. As explained in Ref. 2, the second (comparison) part also has an expansion in  $\lambda$  that follows from the Schwinger equation and the product form Eq. (12). At the end, we set the parameter  $\lambda \rightarrow 1$  in the formal equations, and only then begin the actual computations. Therefore, the primary use of the parameter  $\lambda$  is to count the relative orders of the terms that are higher than quadratic in the Fermi operators.

Let us first present an overview of the formal equations; the AIM model impurity Green's function  $\mathcal{G}$  is written in the presence of time- and spin-dependent potentials  $\mathcal{V}$ , i.e., the sources of Schwinger, and their exact Schwinger equation of motion obtained in Eqs. (3) and (5). In terms of the auxiliary Green's function  $\mathbf{g}$  and the comparison function  $\mu$ , we introduce a convolution ansatz  $\mathcal{G} = \mathbf{g} \cdot \mu$  in Eq. (12). Two types of vertices are introduced by taking the functional derivatives  $\frac{\delta}{\delta \mathcal{V}}$  of  $\mathbf{g}^{-1}$  and  $\mu$ , and in terms of these, we find exact Schwinger equations of motion for  $\mathbf{g}$  and  $\mu$  in Eq. (14), which are expressed in terms of the two self-energies [Eq. (15)]. The auxiliary Green's function and the physical impurity Green's function satisfy the number sum rules in Eqs. (19) or (23), and are then expanded in the parameter  $\lambda$  discussed above. An important shift invariance of the AIM is noted in the equations

of motion, and a second chemical potential  $u_0$  introduced as a second Lagrange multiplier to satisfy the two sum rules. To second order  $O(\lambda^2)$ , explicit equations are written out in Eqs. (30)–(33). The Friedel sum rule in this scheme is written out in Eq. (43), and shown to be satisfied exactly at  $T = 0$ . The numerical solution of the equations is performed through a spectral representation of the variables  $\mathbf{g}$ , and in Eq. (50) the two self-energies are written out in terms of these. These are compared with the NRG results at the same densities and results for the spectral functions of the Green's function and the self-energy are compared in detail, both before and after scaling with  $z$ , the quasiparticle weight.

## II. ECFL THEORY OF ANDERSON IMPURITY MODEL

### A. Model and equations for the Green's function

We consider the Anderson impurity model in the limit  $U \rightarrow \infty$  given by the following Hamiltonian:

$$H = \sum_{\sigma} \epsilon_d X^{\sigma\sigma} + \sum_{k\sigma} \epsilon_k n_{k\sigma} + \frac{1}{\sqrt{\Omega}} \sum_{k\sigma} (V_k X^{\sigma 0} c_{k\sigma} + V_k^* c_{k\sigma}^\dagger X^{0\sigma}), \quad (1)$$

where  $\Omega$  is the box volume, and we have set the Fermi energy of the conduction electrons to zero. Here,  $X^{ab} = |a\rangle\langle b|$  is the Hubbard projected electron operator with  $|a\rangle$  describing the empty orbital, and the two singly occupied states  $a = 0, \pm\sigma$ . We study the impurity Green's function

$$\mathcal{G}_{\sigma_i\sigma_f}(\tau_i, \tau_f) = -\langle\langle X^{\sigma_i 0}(\tau_i) X^{\sigma_f 0}(\tau_f) \rangle\rangle, \quad (2)$$

with  $T_\tau$  the imaginary-time ordering symbol, the definition for an arbitrary time-dependent operator  $\mathcal{Q}$ :  $\langle\langle \mathcal{Q} \rangle\rangle = \langle \text{Tr} T_\tau e^{-\mathcal{A}} \mathcal{Q} \rangle / \langle \text{Tr} T_\tau e^{-\mathcal{A}} \rangle$ , and with the Schwinger source term  $\mathcal{A} = \int_0^\beta d\tau \mathcal{V}^{\sigma_1\sigma_2}(\tau) X^{\sigma_1\sigma_2}(\tau)$ , involving a bosonic time-dependent potential  $\mathcal{V}$ . Often, we abbreviate  $\mathcal{V}(\tau_i) \rightarrow \mathcal{V}_i$ . As usual, this potential is set to zero at the end of the calculation. In this paper, expressions such as  $\mathcal{G}(\tau_i, \tau_f)$  and  $\mathcal{V}$  are understood as  $2 \times 2$  matrices in spin space. We assume a constant hybridization  $V_k = V_0$ , and a (flat) band of half-width  $D$  with constant density of states  $\rho(\epsilon) = \rho_0 \theta(D - |\epsilon|)$  with  $\rho_0 = \frac{1}{2D}$ .

Taking the time derivative of Eq. (2), we obtain the Schwinger equation of motion (EOM)

$$\begin{aligned} & \{(\partial_{\tau_i} + \epsilon_d)\mathbb{1} + \mathcal{V}_i\} \mathcal{G}(\tau_i, \tau_f) \\ &= -\delta(\tau_i - \tau_f) \times [\mathbb{1} - \gamma(\tau_i)] \\ & \quad - \frac{1}{\sqrt{\Omega}} [\mathbb{1} - \gamma(\tau_i) + \mathcal{D}_i] \cdot \sum_k V_k G(k, \tau_i; \tau_f), \end{aligned} \quad (3)$$

where  $\gamma(\tau_i) = \mathcal{G}^{(k)}(\tau_i^-, \tau_i)$  following Ref. 1, Eq. (35), or more explicitly in terms of spin indices as  $\gamma_{\sigma_i\sigma_f}(\tau_i) = \sigma_i\sigma_f \mathcal{G}_{\bar{\sigma}_f\bar{\sigma}_i}(\tau_i, \tau_i^+)$ , and with  $\bar{\sigma} = -\sigma$ . In the following, we abbreviate  $\gamma(\tau_i) \rightarrow \gamma_i$ . Here, we introduced the mixed Green's function  $G_{\sigma_i\sigma_f}(k, \tau_i; \tau_f) = -\langle\langle c_{k\sigma_i}(\tau_i) X^{\sigma_f 0}(\tau_f) \rangle\rangle$ , and a functional derivative operator  $(\mathcal{D}_i)_{\sigma_i\sigma_j} = (\sigma_i\sigma_j) \delta/\delta \mathcal{V}^{\sigma_i\bar{\sigma}_j}(\tau_i)$ . In the ECFL formalism,<sup>1</sup> Eq. (3) and similar equations are to be understood as matrix equations in spin space. Following the Schwinger technique, the higher-order Green's functions have been expressed in terms of the source

functional derivatives of the basic ones; an example illustrates this:  $\sigma_i \sigma_j \langle \langle X^{\sigma_i \sigma_j} Q \rangle \rangle = (\gamma_i - \mathcal{D}_i) \langle \langle Q \rangle \rangle$ . Proceeding further, we take a time derivative of the mixed Green's function to find

$$(\partial_{\tau_i} + \epsilon_k) G(k, \tau_i; \tau_f) = -\frac{1}{\sqrt{\Omega}} V_k^* \mathcal{G}(\tau_i, \tau_f), \quad (4)$$

so combining with Eq. (3) we find the exact EOM for the localized electron Green's function:

$$\begin{aligned} & \{(\partial_{\tau_i} + \epsilon_d) \mathbb{1} + \mathcal{V}_i\} \mathcal{G}(\tau_i, \tau_f) \\ &= -\delta(\tau_i - \tau_f) (\mathbb{1} - \gamma_i) \\ & \quad - (\mathbb{1} - \gamma_i + \mathcal{D}_i) \cdot \Delta(\tau_i - \tau_j) \cdot \mathcal{G}(\tau_j, \tau_f), \end{aligned} \quad (5)$$

with the convention that the time label in bold letters  $\tau_j$  is to be integrated over  $\in [0, \beta]$ . The conduction band enters through the usual ( $\mathcal{V}$ -independent) function

$$\Delta(\tau_i - \tau_j) = -\frac{\mathbb{1}}{\Omega} \sum_k |V_k|^2 (\partial_{\tau_i} + \epsilon_k)^{-1} \delta(\tau_i - \tau_j), \quad (6)$$

with a Fourier transform

$$\Delta(i\omega_n) = \frac{1}{\Omega} \sum_k \frac{|V_k|^2}{i\omega_n - \epsilon_k} = V_0^2 \int \frac{\rho(\epsilon) d\epsilon}{i\omega_n - \epsilon}. \quad (7)$$

We will require below its analytic continuation  $i\omega_n \rightarrow \omega + i\eta$ :

$$\Delta(\omega + i\eta) = \Delta_R(\omega) - i\Gamma(\omega); \quad (8)$$

$$\Gamma(\omega) = \pi V_0^2 \rho(\omega); \quad \Delta_R(\omega) = \frac{\Gamma_0}{\pi} \ln \frac{|\omega + D|}{|\omega - D|}. \quad (9)$$

Here,  $\Gamma_0 = \pi V_0^2 \rho_0$ . We now use the noninteracting Green's function

$$\mathbf{g}_0^{-1}(\tau_i, \tau_f) = -[\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i)] \delta(\tau_i - \tau_f) - \Delta(\tau_i, \tau_f), \quad (10)$$

and rewrite the fundamental equation of motion Eq. (5) as

$$\begin{aligned} & \{\mathbf{g}_0^{-1}(\tau_i, \tau_j) + (\gamma_i - \mathcal{D}_i) \cdot \Delta(\tau_i - \tau_j)\} \cdot \mathcal{G}(\tau_j, \tau_f) \\ &= (\mathbb{1} - \gamma_i) \delta(\tau_i - \tau_f). \end{aligned} \quad (11)$$

Let us note an important *shift invariance* of Eqs. (11) and (10). If we consider a transformation  $\Delta(\tau) \rightarrow \Delta(\tau) + u_\tau \times \delta(\tau)$  with an arbitrary  $u_\tau$ , it is possible to show that Eq. (11) is unchanged, except for a shift of  $\epsilon_d$  by  $-u_\tau$ . The added term  $u_\tau (\gamma_i - \mathcal{D}_i) \mathcal{G}(\tau_i, \tau_f)$  vanishes upon using the Pauli principle and the Gutzwiller projection applicable to operators *at the same time instant*. We use this shift invariance below to introduce a second chemical potential. In the ECFL theory, we use a product *ansatz*

$$\mathcal{G}(\tau_i, \tau_f) = \mathbf{g}(\tau_i, \tau_j) \cdot \mu(\tau_j, \tau_f), \quad (12)$$

where  $\mu(\tau_i, \tau_j)$  is the caparison factor, and we use this in Eq. (11). It is useful to introduce two vertex functions  $\Lambda_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(\tau_n, \tau_m; \tau_i) = -\frac{\delta}{\delta \mathcal{V}_i^{\sigma_3 \sigma_4}} \mathbf{g}_{\sigma_1 \sigma_2}^{-1}(\tau_n, \tau_m)$  and  $\mathcal{U}_{\sigma_3 \sigma_4}^{\sigma_1 \sigma_2}(\tau_n, \tau_m; \tau_i) = \frac{\delta}{\delta \mathcal{V}_i^{\sigma_3 \sigma_4}} \mu_{\sigma_1 \sigma_2}(\tau_n, \tau_m)$  as usual, and suppressing the time indices, we note  $\frac{\delta}{\delta \mathcal{V}} \cdot \mathbf{g} = \mathbf{g} \cdot \Lambda \cdot \mathbf{g}$ . We now use the chain rule and Eq. (12) to write  $\mathcal{D} \Delta \cdot \mathcal{G} = \mathcal{D} \Delta \cdot \mathbf{g} \cdot \mu = \xi^* \Delta \cdot \mathbf{g} \cdot \Lambda_* \cdot \mathbf{g} \cdot \mu + \xi^* \Delta \cdot \mathbf{g} \cdot \mathcal{U}_*$ , with the matrix  $\xi_{\sigma \sigma'} = \sigma \sigma'$ . The \* symbol from Ref. 1 is illustrated in component form by an example:  $\dots \xi_{\sigma_a \sigma_b}^* \dots \delta / \delta \mathcal{V}^* = \dots \sigma_a \sigma_b \dots \delta / \delta \mathcal{V}^{\sigma_a \sigma_b}$ , or

in terms of the vertex functions  $\dots \xi_{\sigma_a \sigma_b}^* \dots \Lambda_*^{\sigma' \sigma''} \dots = \dots \sigma_a \sigma_b \dots \Lambda_{\sigma_a \sigma_b}^{\sigma' \sigma''} \dots$ , with the upper indices of  $\Lambda$  governed by the rules of the matrix multiplication. Following Ref. 1 we define the linear operator  $\mathbf{L}(i, j) = \xi^* \Delta(i, \mathbf{j}) \cdot \mathbf{g}(\mathbf{j}, j) \cdot \frac{\delta}{\delta \mathcal{V}_i^*}$ . We can now collect these definitions to rewrite  $\mathcal{D} \Delta \cdot \mathcal{G} = \xi^* \Delta \cdot \mathbf{g} \cdot \Lambda_* \cdot \mathbf{g} \cdot \mu + \xi^* \Delta \cdot \mathbf{g} \cdot \mathcal{U}_* = \Phi \cdot \mathbf{g} \cdot \mu + \Psi$ , and define the two self-energies:

$$\begin{aligned} \Phi(i, j) &= -\mathbf{L}(i, \mathbf{r}) \cdot \mathbf{g}^{-1}(\mathbf{r}, j) = \xi^* \Delta(i, \mathbf{j}) \cdot \mathbf{g}(\mathbf{j}, \mathbf{k}) \cdot \Lambda_*(\mathbf{k}, j; i); \\ \Psi(i, j) &= \mathbf{L}(i, \mathbf{r}) \cdot \mu(\mathbf{r}, j) = \xi^* \Delta(i, \mathbf{j}) \cdot \mathbf{g}(\mathbf{j}, \mathbf{k}) \cdot \mathcal{U}_*(\mathbf{k}, j; i). \end{aligned} \quad (13)$$

Summarizing, we may rewrite the exact EOM Eq. (11) symbolically:

$$\{\mathbf{g}_0^{-1} + \gamma \Delta - \Phi\} \cdot \mathbf{g} \cdot \mu = (\mathbb{1} - \gamma) \delta + \Psi. \quad (14)$$

This equation is split into two parts by requiring  $\mathbf{g}$  to be canonical:

$$\mathbf{g}^{-1} = \{\mathbf{g}_0^{-1} + \gamma \Delta - \Phi\} \quad \text{and} \quad \mu = (\mathbb{1} - \gamma) \delta + \Psi, \quad (15)$$

bringing it into the standard form in the ECFL theory.<sup>1</sup> Using Eq. (13), we note that the formal solutions of Eq. (15) are  $\mathbf{g}^{-1} = (\mathbb{1} - \mathbf{L})^{-1} \cdot (\mathbf{g}_0^{-1} + \gamma \Delta)$  and  $\mu = (\mathbb{1} - \mathbf{L})^{-1} \cdot (\mathbb{1} - \gamma) \delta$ . We introduce the resolvent kernel  $\mathcal{L}$  using the identity  $(\mathbb{1} - \mathbf{L})^{-1} = \mathbb{1} + \mathcal{L}$  where  $\mathcal{L} = (\mathbb{1} - \mathbf{L})^{-1} \cdot \mathbf{L}$ . In terms of the resolvent, we see that

$$\Phi = \mathcal{L} \cdot (-\mathbf{g}_0^{-1} - \gamma \Delta) \quad \text{and} \quad \Psi = -\mathcal{L} \cdot \gamma \delta. \quad (16)$$

Therefore, distributing the action of  $\mathcal{L}$  over the two terms, we can rewrite

$$\Phi = \chi + \Psi \Delta, \quad (17)$$

with

$$\chi = \mathcal{L} \cdot (-\mathbf{g}_0^{-1}). \quad (18)$$

Therefore, the self-energy  $\Phi$  breaks up into two parts, as in Eq. (17). Note that in Eq. (16), the expressions  $\gamma \Delta$  and  $\gamma \delta$  involve multiplication at equal times, whereas in Eq. (17),  $\Psi \Delta$  implies a convolution in time. The two Green's functions satisfy the pair of sum rules

$$\mathbf{g}(\tau, \tau^+) = \frac{n_d}{2}; \quad \mathcal{G}(\tau, \tau^+) = \frac{n_d}{2}, \quad (19)$$

where  $n_d$  is the number of electrons on the  $d$  orbital  $n_d = \sum_{\sigma} \langle X^{\sigma \sigma} \rangle$ .

In the context of the  $t$ - $J$  model in Ref. 2, the sum rule for  $\mathbf{g}$  is necessary to satisfy the Luttinger-Ward theorem. If we use the representation  $\hat{f}_{\sigma}^{\dagger}(\lambda) = (1 - \lambda f_{\sigma}^{\dagger} f_{\sigma}) f_{\sigma}^{\dagger}$  for the correlated electrons, this constraint is understandable as the constraint on the number of ‘‘uncorrelated’’ fermions  $\langle f_{\sigma}^{\dagger} f_{\sigma} \rangle$ , which must agree with the number of physical (correlated) electrons  $\langle \hat{f}_{\sigma}^{\dagger} \hat{f}_{\sigma} \rangle$ . Similarly, in the present case, this constraint is needed to fulfill the Friedel sum rule. We also remark that the self-energy  $\Psi$ , unlike its counterpart  $\Phi$ , is dimensionless, and thus interpreted as an adaptive spectral weight.<sup>2</sup>

## B. Zero source limit

Upon turning off the sources, all objects become functions of only  $\tau_i - \tau_f$  and may therefore be Fourier transformed to Matsubara frequency space. By Fourier transforming

Eqs. (12), (15), and (17) and using  $\gamma \rightarrow \frac{n_d}{2}$  we obtain the following expressions in frequency space:

$$\begin{aligned}\mathcal{G}(i\omega_n) &= \mathbf{g}(i\omega_n) \cdot \mu(i\omega_n), \\ \mu(i\omega_n) &= 1 - \frac{n_d}{2} + \Psi(i\omega_n), \\ \mathbf{g}^{-1}(i\omega_n) &= i\omega_n - \epsilon_d - \Delta(i\omega_n)\mu(i\omega_n) - \chi(i\omega_n).\end{aligned}\quad (20)$$

Alternately, this result can be rewritten in terms of the Dyson-Mori self-energy representation as

$$\mathcal{G}(i\omega_n) = \frac{1 - \frac{n_d}{2}}{i\omega_n - \epsilon_d - \left(1 - \frac{n_d}{2}\right)\Delta(i\omega_n) - \Sigma_{\text{DM}}(i\omega_n)} \quad (21)$$

and

$$\begin{aligned}\Sigma_{\text{DM}}(i\omega_n) + \epsilon_d - i\omega_n \\ = \frac{1 - \frac{n_d}{2}}{1 - \frac{n_d}{2} + \Psi(i\omega_n)} [\chi(i\omega_n) + \epsilon_d - i\omega_n].\end{aligned}\quad (22)$$

The sum rules (19) are

$$\sum_{i\omega_n} \mathcal{G}(i\omega_n) e^{i\omega_n \eta} = \frac{n_d}{2}; \quad \sum_{i\omega_n} \mathbf{g}(i\omega_n) e^{i\omega_n \eta} = \frac{n_d}{2}. \quad (23)$$

These are satisfied at a fixed  $n_d$  using two Lagrange multipliers: the localized state energy  $\epsilon_d$  and the second chemical potential  $u_0$  introduced in Eq. (25). We observe that the usual Dysonian self-energy  $\Sigma_{\text{AM}}(i\omega_n)$  defined through the usual Dyson equation (valid for finite  $U$ )  $G^{-1} = i\omega_n - \epsilon_d - \Delta(i\omega_n) - \Sigma_{\text{AM}}(i\omega_n)$  in the infinite- $U$  limit can be obtained from

$$\Sigma_{\text{AM}}(i\omega_n) = \frac{2}{2 - n_d} \Sigma_{\text{DM}}(i\omega_n) + \frac{n_d}{2 - n_d} (\epsilon_d - i\omega_n). \quad (24)$$

The unlimited growth with  $\omega_n$  makes this self-energy somewhat inconvenient to deal with, and therefore motivated the introduction of the Dyson-Mori object, which is better behaved in this regard. After analytic continuation  $i\omega_n \rightarrow \omega + i0^+$ , the imaginary part of  $\Sigma_{\text{AM}}$  is well behaved and finite as  $\omega \rightarrow \infty$ . It is obtained from the NRG method and compared with the relevant ECFL functions after scaling by  $1 - \frac{n_d}{2}$  as in Eq. (24). We notice that the density  $n_d$  appears explicitly in the expressions for the Green's functions, and must therefore be calculated self-consistently from Eq. (23). This feature is quite natural in the present approach since Eq. (3) for the Green's function contains  $\gamma$  and therefore  $n_d$  explicitly.

### C. Introducing $\lambda$ and $u_0$ into the equations

Summarizing the work so far: Eqs. (15)–(17) follow from Eq. (11) upon using the product ansatz (12), and are exact equations. In order to get concrete results, we proceed by introducing two parameters into the equations. (i) The parameter  $\lambda \in [0, 1]$  multiplies certain terms shown in Eq. (25), allowing a density-type expansion, and continuously connects the uncorrelated Fermi system  $\lambda = 0$  to the extremely correlated case  $\lambda = 1$ . (ii) The second parameter  $u_0$  is introduced as shown in Eq. (25). It is the second chemical potential used to enforce the shift identities of the exact equation (11).

Equation (11) now becomes

$$\left\{ \mathbf{g}_0^{-1} + \lambda(\gamma - \mathcal{D}) \left( \Delta - \frac{u_0}{2} \delta \right) \right\} \cdot \mathcal{G} = (\mathbb{1} - \lambda\gamma)\delta. \quad (25)$$

As a consequence, in Eq. (14) to Eq. (18) we set  $\gamma \rightarrow \lambda\gamma$ ,  $\Psi \rightarrow \lambda\Psi$ , and  $\Phi \rightarrow \lambda\Phi$ , or  $\chi \rightarrow \lambda\chi$ . Second, in Eq. (14) to Eq. (18) we set  $\Delta(\tau_i, \tau_f) \rightarrow \Delta(\tau_i, \tau_f) - \frac{u_0}{2} \delta(\tau_i - \tau_f)$ . Note that there is no shift of Eq. (10) implied in Eq. (25).

We write Eq. (15) with  $\lambda$  inserted explicitly and the understanding that  $\Delta(\tau_i, \tau_f)$  has been shifted as<sup>25</sup>

$$\begin{aligned}\mathbf{g}^{-1}(\tau_i, \tau_f) &= \mathbf{g}_0^{-1}(\tau_i, \tau_f) + \lambda\gamma(\tau_i)\Delta(\tau_i, \tau_f) - \lambda\Phi(\tau_i, \tau_f), \\ \mu(\tau_i, \tau_f) &= \delta(\tau_i - \tau_f)[\mathbb{1} - \lambda\gamma(\tau_i)] + \lambda\Psi(\tau_i, \tau_f),\end{aligned}\quad (26)$$

where the two self-energies are given in terms of the vertex functions as

$$\begin{aligned}\Phi(\tau_i, \tau_f) &= \xi^* \Delta(\tau_i, \tau_j) \cdot \mathbf{g}(\tau_j, \tau_k) \cdot \Lambda_*(\tau_k, \tau_f; \tau_i), \\ \Psi(\tau_i, \tau_f) &= \xi^* \Delta(\tau_i, \tau_j) \cdot \mathbf{g}(\tau_j, \tau_k) \cdot U_*(\tau_k, \tau_f; \tau_i).\end{aligned}\quad (27)$$

On switching off the sources, these expressions can be spin resolved and expressed as  $\Phi = \Delta \mathbf{g} \Lambda^{(a)}$  and  $\Psi = \Delta \mathbf{g} U^{(a)}$ , with the same time labels as above, and with the usual spin decomposition  $\Lambda^{(a)} = \Lambda_{\sigma\bar{\sigma}}^{\sigma\sigma} - \Lambda_{\sigma\bar{\sigma}}^{\sigma\bar{\sigma}}$ .

### D. $\lambda$ expansion

We note that we can obtain the equations of motion for the Anderson model from the equations of motion for the  $t$ - $J$  model by making the following substitutions and replacing all space-time variables with just time<sup>26</sup>:

$$\begin{aligned}t[i, f] &\rightarrow -\Delta(\tau_i, \tau_f); \quad \epsilon_k \rightarrow \Delta(i\omega_k), \\ J &\rightarrow 0, \quad \mu \rightarrow -\epsilon_d.\end{aligned}\quad (28)$$

The  $\lambda$  expansion for the Anderson model is therefore analogous to the one for the  $t$ - $J$  model in Ref. 2 and the large- $d$   $t$ - $J$  model in Ref. 26, and can be obtained from them by making the substitutions in Eq. (28) and changing all frequency momentum four-vectors to just frequency. For completeness, Appendix A provides a brief derivation (in time domain) of the following equations. Denoting

$$a_G = 1 - \lambda \frac{n_d}{2} + \lambda^2 \frac{n_d^2}{4}, \quad (29)$$

and the frequently occurring object

$$\mathcal{R} = \mathbf{g}(i\omega_p) \mathbf{g}(i\omega_q) \mathbf{g}(i\omega_p + i\omega_q - i\omega_n),$$

we obtain to  $O(\lambda^2)$  the expressions

$$\mathcal{G}(i\omega_n) = \mathbf{g}(i\omega_n) \mu(i\omega_n), \quad \mu(i\omega_n) = a_G + \lambda\Psi(i\omega_n), \quad (30)$$

$$\mathbf{g}^{-1}(i\omega_n) = i\omega_n - \epsilon'_d - \left( \Delta(i\omega_n) - \frac{u_0}{2} \right) \mu(i\omega_n) - \lambda\chi(i\omega_n), \quad (31)$$

$$\begin{aligned}\chi(i\omega_n) &= -\lambda \sum_{p,q} [2\Delta(i\omega_p) - u_0] \\ &\quad \times \left[ \Delta(i\omega_p + i\omega_q - i\omega_n) - \frac{u_0}{2} \right] \mathcal{R},\end{aligned}\quad (32)$$

$$\Psi(i\omega_n) = -\lambda \sum_{p,q} [2\Delta(i\omega_p) - u_0] \mathcal{R}. \quad (33)$$

The energy  $\epsilon'_d$  is given by collecting the static terms in  $\Phi$  as

$$\epsilon'_d = \epsilon_d + u_0 \left( \lambda \frac{n_d}{2} - \lambda^2 \frac{n_d^2}{8} \right) + \frac{u_0}{2} a_G - \lambda \sum_{i\omega_p} \Delta(i\omega_p) \mathbf{g}(i\omega_p). \quad (34)$$

The shift theorem is satisfied by all the terms separately since we have taken care to form expressions of the type  $\Delta - \frac{u_0}{2}$ . As discussed in Ref. 2, the shift theorems mandate the introduction of  $u_0$ , and its availability, in addition to  $\epsilon_d$ , enables us to fix the pair of sum rules (19). As explained, we must set  $\lambda \rightarrow 1$  before using these expressions.

Within the  $O(\lambda^2)$  theory, the total spectral weight of the Green's function is  $a_G$  rather than the exact value  $1 - \frac{n_d}{2}$ . This is understood as the incomplete projection to single occupancy leading to an excess in the total number of states available to the system. In order to ensure that  $\Sigma_{\text{DM}}(\omega)$  retain the feature of being finite as  $\omega \rightarrow \infty$ , it must be slightly redefined (to  $\hat{\Sigma}_{\text{DM}}$ ) in the  $O(\lambda^2)$  theory:

$$G(\omega) = \frac{a_G}{\omega - \epsilon'_d - a_G \Delta(\omega) - \hat{\Sigma}_{\text{DM}}(\omega)}, \quad (35)$$

where

$$\epsilon''_d \equiv \epsilon'_d - \frac{u_0}{2} a_G. \quad (36)$$

Using Eqs. (30) and (31), we can relate  $\hat{\Sigma}_{\text{DM}}(\omega)$  to  $\chi(\omega)$  and  $\Psi(\omega)$ :

$$\hat{\Sigma}_{\text{DM}}(\omega) + \epsilon'_d - \omega = \frac{a_G}{a_G + \Psi(\omega)} [\chi(\omega) + \epsilon'_d - \omega]. \quad (37)$$

Since  $\Psi(\omega), \chi(\omega) \rightarrow 0$  as  $\omega \rightarrow \infty$ , we see explicitly that  $\hat{\Sigma}_{\text{DM}}(\omega)$  remains finite in this limit. Just as in the case of  $\text{Im } \Sigma_{\text{DM}}(\omega)$ ,  $\text{Im } \hat{\Sigma}_{\text{DM}}(\omega)$  is related to  $\text{Im } \Sigma_{\text{AM}}(\omega)$  by a multiplicative constant ( $1 - \frac{n_d}{2}$  and  $a_G$ , respectively), and therefore their spectra are identical apart from this multiplicative constant. Comparing Eqs. (21) and (35), we see that the latter is obtained from the former with the substitutions

$$\Sigma_{\text{DM}}(\omega) \rightarrow \hat{\Sigma}_{\text{DM}}(\omega); \quad \epsilon_d \rightarrow \epsilon''_d; \quad 1 - \frac{n_d}{2} \rightarrow a_G. \quad (38)$$

Keeping these substitutions in mind, we will now only use  $\Sigma_{\text{DM}}(\omega)$  from the exact theory, with the understanding that the same expressions hold for  $\hat{\Sigma}_{\text{DM}}(\omega)$  in the  $O(\lambda^2)$  theory as long as the substitutions in Eq. (38) are made.

### E. Friedel sum rule at $T = 0$

At  $T = 0$ , the Friedel sum rule<sup>27-29</sup> plays an important role in the AIM, parallel to that of the Luttinger-Ward volume theorem in Fermi liquids. In Ref. 29, the original form of the Friedel sum rule is written in terms of  $\eta_\sigma(\omega)$ , the phase shift of the conduction electron with spin  $\sigma$  at energy  $\omega$ :

$$\eta_\sigma(\omega) = \frac{1}{2i} \ln [\mathcal{G}_\sigma(\omega + i0^+) \mathcal{G}_\sigma^{-1}(\omega - i0^+)], \quad (39)$$

where the logarithm is chosen with a branch cut along the positive real axis, so that  $0 \leq \eta \leq \pi$ . The Friedel sum rule is then written as

$$\eta_\sigma(\omega = 0) = \frac{\pi n_d}{2}. \quad (40)$$

This theorem is proven for the AIM at finite  $U$  (Ref. 29) by adapting the argument of Luttinger and Ward<sup>30</sup> with an implicit assumption of a nonsingular evolution in  $U$  from 0. We assume that the Friedel sum rule also holds in the extreme correlation limit  $U \rightarrow \infty$ . Using the Dyson-Mori representation (21) to compute the phase shift in Eq. (39), we may rewrite this as

$$n_d = 1 - \frac{2}{\pi} \tan^{-1} \left[ \frac{\epsilon_d + \text{Re } \Sigma_{\text{DM}}(0)}{\Gamma_0 (1 - \frac{n_d}{2})} \right], \quad (41)$$

with  $\epsilon_d + \text{Re } \Sigma_{\text{DM}}(0) > 0$ , in the physical case of  $0 \leq n_d \leq 1$ . It is easily seen<sup>32</sup> that this form is equivalent to the standard statement of the Friedel sum rule<sup>15</sup>

$$\rho_G(0) = \frac{1}{\pi \Gamma_0} \sin^2 \left( \frac{\pi n_d}{2} \right). \quad (42)$$

Within the approximation of the  $\lambda$  expansion, the Friedel sum rule implies a relationship between the values of the two self-energies at zero frequency:

$$n_d = 1 - \frac{2}{\pi} \tan^{-1} \left[ \frac{\epsilon'_d - \frac{u_0}{2} \mu(0) + \chi(0)}{\Gamma_0 \mu(0)} \right]. \quad (43)$$

This can be obtained by using the substitutions from Eq. (38) in Eq. (41), and using Eqs. (37), (36), and (30).

We can also record a result for the auxiliary density of states  $\rho_g(\omega = 0)$ , analogous to Eq. (42) here. It follows from Eq. (47), with the Fermi liquid type assumption of vanishing of  $\rho_\psi(0)$  at  $T = 0$ , and reads as

$$\rho_g(0) = \frac{1}{\pi \Gamma_0 \mu(0)} \sin^2 \left( \frac{\pi n_d}{2} \right). \quad (44)$$

We check the validity of the Friedel sum rule within the  $\lambda$  expansion in both the forms (42) and (43). In doing so, we are thus testing if the strategy of the two ECFL sum rules [Eq. (23)] enforces the Friedel sum rule, in a situation that is essentially different from that in finite- $U$  theories so that the central result of Luttinger and Ward<sup>30</sup> is not applicable in any obvious way.

### F. Computation of spectral function

In computing the spectral function, we follow the approach taken in Ref. 2, in which the spectral function is calculated for the  $O(\lambda^2)$  ECFL theory of the  $t$ - $J$  model. Our calculation is made simpler due to the absence of any spatial degrees of freedom, but more complicated by the presence of the frequency-dependent factor  $\Delta(i\omega_n)$ . We define the various spectral functions and the relationships between them. These expressions are analogous to those in Sec. III A of Ref. 2:

$$Q(i\omega_n) = \int_{-\infty}^{\infty} dv \frac{\rho_Q(v)}{i\omega_n - v}, \quad (45)$$

where  $Q$  can stand for any object such as  $\mathcal{G}$ ,  $\mathbf{g}$ ,  $\chi$ ,  $\Sigma_{\text{DM}}$ , or  $\Psi$ . Therefore, after analytic continuation  $i\omega_n \rightarrow \omega + i0^+$ ,

$$\rho_Q(\omega) \equiv -\frac{\text{Im}}{\pi} Q(\omega + i0^+) \quad \text{and} \quad \text{Re } Q(\omega) = \mathcal{H}[\rho_Q](\omega), \quad (46)$$

where for any real density  $\rho_Q(\omega)$  the Hilbert transform is denoted as  $\mathcal{H}[\rho_Q](\omega) = P \int_{-\infty}^{\infty} dv \frac{\rho_Q(v)}{\omega - v}$ . From Eq. (33), we

TABLE I. The bare impurity level  $\epsilon_d$  as well as the quasiparticle weight  $z$  are displayed for the ECFL and the NRG calculations for all values of the density. Additionally, the theoretical value for the Friedel sum rule as well as the ECFL deviation from it are displayed.

$n_d$	$\rho_{G,\text{ECFL}}(0)$	$\epsilon_{d,\text{ECFL}}$	$\epsilon_{d,\text{NRG}}$	$z_{\text{ECFL}}$	$z_{\text{NRG}}$
0.35	8.69 + 1.8%	-0.003	-0.003	0.753	0.697
0.441	13.0 + 1.1%	-0.010	-0.009	0.661	0.567
0.536	17.7 + 0.73%	-0.015	-0.015	0.559	0.416
0.6	20.8 + 0.41%	-0.019	-0.018	0.489	0.312
0.7	25.3 + 0.62%	-0.024	-0.024	0.388	0.169
0.777	28.1 + 0.26%	-0.031	-0.029	0.314	0.081
0.834	29.7 + 0.20 %	-0.037	-0.035	0.265	0.035

find that

$$\rho_G(\omega) = \rho_g(\omega)[a_G + \text{Re } \Psi(\omega)] + \rho_\psi(\omega)\text{Re } g(\omega). \quad (47)$$

With  $f(\omega) = \frac{1}{1+e^{\beta\omega}}$  and  $\bar{f}(\omega) = 1 - f(\omega)$ , the two sum rules (23) read as

$$\int_{-\infty}^{\infty} f(\omega)\rho_g(\omega)d\omega = \frac{n_d}{2}, \quad \int_{-\infty}^{\infty} \bar{f}(\omega)\rho_g(\omega)d\omega = \frac{n_d}{2}. \quad (48)$$

We also note  $\rho_\Delta(\omega) = \frac{\Gamma(\omega)}{\pi}$ . It is useful to define a mixed (composite) density

$$\rho_M(x) = \rho_g(x)\left(\Delta_R(x) - \frac{u_0}{2}\right) + \rho_\Delta(x)\text{Re } g(x), \quad (49)$$

so that we can integrate (or sum) the internal frequencies in Eq. (33) efficiently (see Appendix B), and write the two relevant complex self-energies (with  $\omega \equiv \omega + i0^+$ ) as

$$\Psi(\omega) = -2\lambda \int_{u,v,w} \frac{\rho_M(u)\rho_g(v)\rho_g(w)}{\omega - u - v + w} \times [f(u)f(v)\bar{f}(w) + \bar{f}(u)\bar{f}(v)f(w)],$$

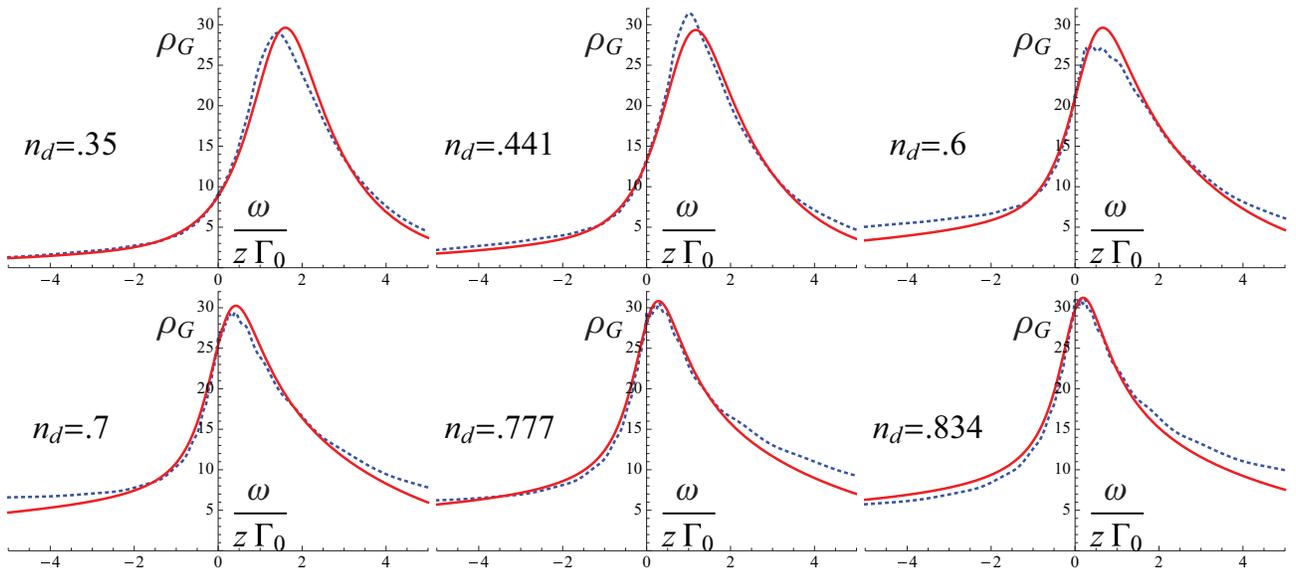


FIG. 1. (Color online) The spectral density for the physical Green's function versus  $\frac{\omega}{\Gamma_0 z}$  for densities of  $n_d = 0.35, 0.441, 0.6, 0.7, 0.777, 0.834$ . The red curve is the ECFL calculation, while the blue curve is the NRG calculation.

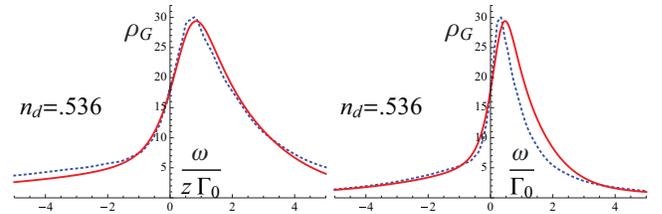


FIG. 2. (Color online) The spectral density for the physical Green's function for the density of  $n_d = 0.536$ . For the plot on the left, both the ECFL and NRG curves are plotted versus  $\frac{\omega}{z\Gamma_0}$ . Since ECFL has a larger  $z$  value, the absolute scale of the  $\omega$  axis differs for the two curves. For the plot on the right, both ECFL and NRG are plotted versus  $\frac{\omega}{\Gamma_0}$  and hence the ECFL peak is too wide.

$$\chi(\omega) = -2\lambda \int_{u,v,w} \frac{\rho_M(u)\rho_g(v)\rho_M(w)}{\omega - u - v + w} \times [f(u)f(v)\bar{f}(w) + \bar{f}(u)\bar{f}(v)f(w)]. \quad (50)$$

In these expressions,  $u, v, w$  are understood to be real variables, and using Eq. (46) we can extract the real and imaginary parts of  $\Psi$  and  $\chi$  in terms of the spectral functions.

### III. RESULTS

The following explicit results were obtained after setting  $\lambda = 1$  in the equations noted above. We calculated the spectral functions  $\rho_G$ ,  $\rho_\Sigma$ ,  $\rho_\chi$ , and  $\rho_\psi$  using the values  $D = 1$ ,  $\Gamma_0 = 0.01$ , and  $T = 0$ . The zero-temperature limit is easily achieved in the ECFL theory by setting all of the Fermi functions to step functions. We expect that the spectral function calculated within the ECFL  $O(\lambda^2)$  theory will be accurate through a density of approximately  $n_d = 0.6$ , or perhaps at best  $n_d \sim 0.7$ . As discussed in Refs. 1 and 2, this is the main limitation of the low-order  $\lambda$  results: the theory begins to have substantial corrections as we increase  $n_d$  towards unity. The source of this error estimate is the high-frequency behavior

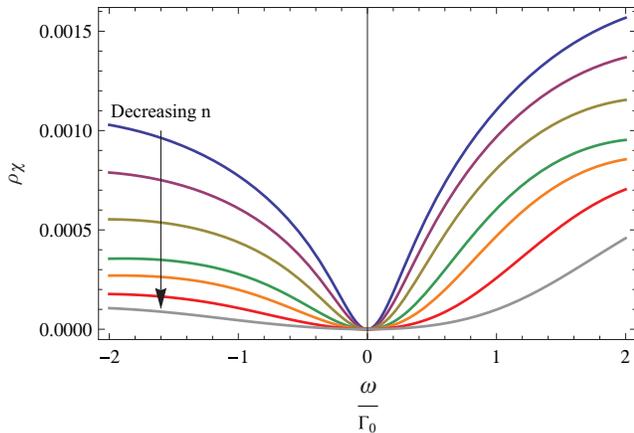


FIG. 3. (Color online) The spectral function for  $\chi$  for densities of  $n_d = 0.834, 0.777, 0.7, 0.6, 0.536, 0.441, 0.35$ .

within the  $\lambda$  expansion of the Green's function (33)  $\mathcal{G} \sim \frac{a_{\mathcal{G}}}{i\omega}$ ; this deviates from the known exact behavior  $\mathcal{G} \sim \frac{1-n_d/2}{i\omega}$ . The error grows with increasing density, but we expect to have reasonable results even at  $n_d = 0.7$ .

In Table I, we show the results for the spectral function at zero energy in terms of the percentage deviation from the Friedel sum rule (42), demonstrating that the ECFL satisfies the Friedel sum rule to a high degree of accuracy. We specify the occupation number  $n_d$  and show the values of the energy level  $\epsilon_d$  and quasiparticle weight  $z$  calculated within the ECFL and NRG calculations. The values of  $\epsilon_d$  are in good agreement between the two calculations, while there is a discrepancy in  $z$  which becomes more pronounced at higher densities. While the error in the scale of  $z$  as  $n_d \rightarrow 1$  is expected from the low order in  $\lambda$  aspect of the theory, we should keep in mind that the *shape* of the spectral function, and also the imaginary part of the self-energy, is another matter altogether. We display below these objects after scaling the frequency with  $z$ : this captures the *shape* of the spectra and isolates the discrepancy to a single number, namely, the magnitude of  $z$ . The admittedly nontrivial problem of the magnitude of  $z$  must await a more satisfactory resolution involving the treatment of higher-order terms in  $\lambda$ .

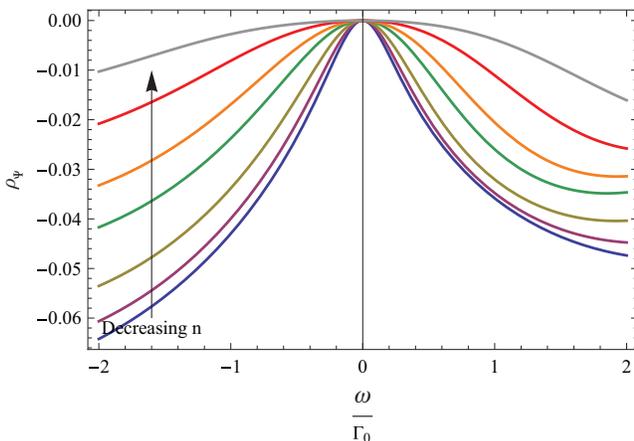


FIG. 4. (Color online) The spectral function for  $\Psi$  for densities of  $n_d = 0.834, 0.777, 0.7, 0.6, 0.536, 0.441, 0.35$ .

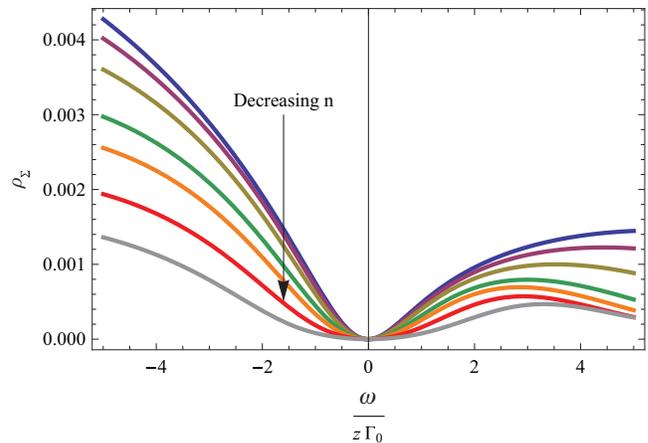


FIG. 5. (Color online) The spectral function for the Dyson-Mori self-energy for densities of  $n_d = 0.834, 0.777, 0.7, 0.6, 0.536, 0.441, 0.35$ . The curvature of the quadratic minimum becomes larger with increasing density.

In Fig. 1 we display the spectral functions at the indicated densities, indicating a smooth evolution with density. The Kondo or Abrikosov-Suhl resonance at positive frequencies becomes sharper as we increase density and moves closer to  $\omega = 0$ . If the raw ECFL and NRG spectral functions are compared (as in right panel of Fig. 2 for  $n_d = 0.536$ ), one finds that the peak in the ECFL spectral function is too broad. This overbroadening becomes worse at larger densities and better at lower densities. However, it can be understood well in terms of the elevated value of  $z$  for ECFL at higher densities. Hence, before doing the comparison, as in Fig. 1, we first rescale the  $\omega$  axis for both the ECFL and NRG spectral functions by the appropriate  $z$  (as in the left panel of Fig. 2 for  $n_d = 0.536$  and in Fig. 1 for the other densities). They are then found to be in good agreement up to surprisingly high values of  $n_d$ , suggesting that the ECFL theory captures the *shape* (but not the scale) of the spectral functions and their asymmetry in a very natural fashion. We also found good agreement with the NRG spectral functions in Ref. 24. The ECFL spectral function  $\rho_G$  is constructed out of the two spectral functions  $\rho_\chi$  and  $\rho_\Psi$  that are shown at various densities in Figs. 3 and 4, exhibiting Fermi liquid type quadratic frequency dependence at low  $\omega$ .

In Fig. 5 we present the density evolution of the spectral function for the Dyson-Mori self-energy [see Eq. (22)]. This exhibits a remarkable similarity to the analogous spectral density for the  $t$ - $J$  model in the limit of high dimensions<sup>33</sup> and the Hubbard model at large  $U$ .<sup>34</sup>

#### IV. CONCLUSION

In this work, we have applied the ECFL formalism at the simplest level, using the  $O(\lambda^2)$  equations, to the Anderson impurity model with  $U \rightarrow \infty$ . In this formalism, the two self-energies of the ECFL theory  $\Psi$  and  $\chi$  are calculated using a skeleton expansion in the auxiliary Green's function  $\mathbf{g}$ . This is analogous to the skeleton expansion for the Dyson self-energy  $\Sigma$ , in standard Feynman-Dyson perturbation theory applicable to the case of finite  $U$ . These two self-energies determine  $\mathbf{g}$  as well as the physical  $\mathcal{G}$ , leading to a self-consistent solution. We

obtained the equations to second order and solved them numerically at  $T = 0$ . We found that at low enough  $\omega$ , the ECFL self-energies have symmetric spectra of the type predicted by Fermi liquid theory (see Figs. 3 and 4). Combining them through the ECFL functional form (22) generates a nontrivial self-energy with an asymmetric spectrum displayed in Fig. 5. It therefore appears that functional form (22) has the potential to generate realistic and nontrivial spectral densities, starting with rather simple components. The availability of convenient and natural analytical expressions is seen to provide a distinct advantage of the ECFL formalism. Formally exact techniques such as the NRG involve steps that are not automatically endowed with these, but rather rely on analytic continuation or other equivalent techniques.

The physical spectral function for the impurity site is obtained from the above pair of ECFL self-energies, and displays a Kondo or Abrikosov-Suhl resonance. This feature becomes more narrow and the spectrum becomes more skewed towards the occupied side of the peak with increasing density. However, the computed quasiparticle  $z$  in the present calculation is considerably larger than the exact value  $z \propto e^{-1/2(1-n_d)}$ , as  $n_d \rightarrow 1$ ,<sup>31</sup> i.e. into the Kondo regime. This large  $z$  makes it impossible for the  $O(\lambda^2)$  version of ECFL presented here to address the Kondo regime  $n_d \rightarrow 1$ . It results in the masking of a small (and broad) peak at  $\omega \sim \epsilon_d$ , found in our NRG spectral functions, as we approach the Kondo limit. Both real and imaginary parts of the computed  $\Sigma_{\text{DM}}(\omega)$  are larger than their NRG counterparts in that regime, thereby precluding a peak.

To place this result in context, we observe that the same level of approximation of ECFL, applied to the lattice problem of the  $d \rightarrow \infty$ ,  $U \rightarrow \infty$  Hubbard model in Ref. 33 (see Fig. 12), *does* show a lower Hubbard band peak in the spectral function. This difference presumably arises from the robust value of  $z \sim (1-n)$  in the lattice model, arising from Gutzwiller physics; it is much larger than the exponentially small value  $z \propto e^{-1/2(1-n_d)}$  in the AIM. Therefore, the fractional error made by the  $O(\lambda^2)$  ECFL calculation is smaller in the lattice model compared to the AIM.

The location of the peak is set by  $\epsilon_d + \Sigma_{\text{DM}}(0)$  [Eq. (21)]. Using Eq. (41), we can see that this quantity must decrease with increasing density. This is consistent with the expectation that the location of the peak will approach  $\omega = 0$  as  $n_d \rightarrow 1$ . This can also be understood from the need to have more spectral weight when  $\omega \leq 0$ , to yield a higher value of  $n_d$ . We found that the ECFL spectrum satisfies the Friedel sum rule [Eq. (42)] to a high degree of accuracy, and that ECFL yields values of  $\epsilon_d$  in good agreement with the NRG values at all densities (see Table I).

As mentioned above, the ECFL calculation to  $O(\lambda^2)$  overestimates the value of the quasiparticle weight  $z$  as compared with the NRG and the exact asymptotic result  $z \propto e^{-1/2(1-n_d)}$  as  $n_d \rightarrow 1$ ,<sup>31</sup> the difference becoming more significant with increasing density. This also leads to an overbroadening of the peak in the ECFL spectrum at higher densities. This is consistent with the fact that the  $\lambda$  expansion of the ECFL is a low-density expansion and the current calculation has only been carried out to  $O(\lambda^2)$ . Nevertheless, after rescaling the  $\omega$  axis for both the ECFL and NRG spectra by their respective values of  $z$ , we find good quantitative agreement between the

two as in Fig. 1. In Fig. 2, we illustrate the comparison between scaled and unscaled spectral functions at a typical density. We find similarly good agreement with the NRG calculation from Ref. 24. This implies that the ECFL theory has the correct *shape* of the spectra built into it quite naturally.

Finally, we note that the computed spectral functions exhibit a remarkable similarity to the analogous spectral density for the  $t$ - $J$  model in the limit of high dimensions<sup>33</sup> and the Hubbard model at large  $U$ .<sup>34</sup>

## ACKNOWLEDGMENTS

This work was supported by DOE under Grant No. FG02-06ER46319. We thank H. R. Krishna-murthy for helpful comments.

## APPENDIX A: CALCULATING THE SELF-ENERGIES IN THE $O(\lambda^2)$ THEORY

The calculation follows the procedure given in Ref. 2. A few comments are provided to make the connections explicit: the zeroth-order vertices are common to Ref. 2 [Eqs. (B3) and (B14)] and the first-order  $\mathcal{U}$  is common to Eq. (B15). The first-order vertex  $[\Delta]_1$  can be found parallel to Eqs. (B23)–(B28) in Ref. 2 from differentiating

$$[\mathbf{g}^{-1}(i, f)]_1 = \Delta(i, f)\mathbf{g}^{(k)}(i, i) + \delta(i, f)\Delta(i, \mathbf{a})\mathbf{g}^{(k)}(\mathbf{a}, f), \quad (\text{A1})$$

as

$$[\Lambda^{(a)}(i, m; j)]_1 = -2\Delta(i, m)\mathbf{g}(i, j) \cdot \mathbf{g}(j, i) - 2\delta(i, m)\Delta(i, \mathbf{k})\mathbf{g}(\mathbf{k}, j) \cdot \mathbf{g}(j, i). \quad (\text{A2})$$

Here, the bold labels are integrated over. From this we construct, the time-domain self-energies

$$\Psi(i, f) = -2\lambda\Delta(i, \mathbf{k})\mathbf{g}(\mathbf{k}, f) \cdot \mathbf{g}(i, f) \cdot \mathbf{g}(f, i) \quad (\text{A3})$$

and

$$\begin{aligned} \Phi(i, f) = & -\delta(i, f)\Delta(i\mathbf{k})\mathbf{g}(\mathbf{k}i) \\ & - 2\lambda\Delta(i\mathbf{j})\mathbf{g}(\mathbf{j}\mathbf{k}) \cdot \Delta(\mathbf{k}f)\mathbf{g}(\mathbf{k}i) \cdot \mathbf{g}(i\mathbf{k}) \\ & - 2\lambda\Delta(i\mathbf{j})\mathbf{g}(\mathbf{j}f) \cdot \Delta(f\mathbf{k})\mathbf{g}(\mathbf{k}i) \cdot \mathbf{g}(if). \end{aligned} \quad (\text{A4})$$

After shifting  $\Delta(i, f) \rightarrow \Delta(i, f) - \frac{u_0}{2}\delta(i, f)$  and Fourier transforming these, we obtain Eqs. (33) and (34). These expressions for the self-energies are correct to  $O(\lambda)$  and lead to expression for  $\mathbf{g}^{-1}$  and  $\mu$  which are correct to  $O(\lambda^2)$ .  $\chi$  can be extracted from  $\Phi$  as indicated in the text.

## APPENDIX B: FREQUENCY SUMMATIONS

An efficient method to perform the frequency sums is to work with the time-domain formulas (A3) and (A4) until the final step where Fourier transforms are taken. We note the representation for the Green's function

$$\mathbf{g}(\tau) = \int_x \rho_{\mathbf{g}}(x)e^{-\tau x}[\theta(-\tau)f(x) - \theta(\tau)\bar{f}(x)], \quad (\text{B1})$$

so that we can easily compound any pair that arises by dropping the cross products  $\theta(\tau)\theta(-\tau)$  and using  $\theta(\tau)^2 = \theta(\tau)$ . An

example illustrates this procedure:

$$\mathbf{g}(\tau)\mathbf{g}(-\tau) = - \int_{x,y} \rho_{\mathbf{g}}(x)\rho_{\mathbf{g}}(y)e^{-\tau(x-y)} \times [\theta(-\tau)f(x)\bar{f}(y) + \theta(\tau)\bar{f}(x)f(y)]. \quad (\text{B2})$$

We also need to deal with the convolution of pairs of functions

$$\begin{aligned} X(\tau) &= \int_{-\beta}^{\beta} d\bar{\tau} \mathbf{g}(\bar{\tau}) \left[ \Delta(\tau - \bar{\tau}) - \frac{u_0}{2} \delta(\tau - \bar{\tau}) \right] \\ &= \int_x \rho_M(x) e^{-\tau x} [\theta(-\tau)f(x) - \theta(\tau)\bar{f}(x)], \quad (\text{B3}) \end{aligned}$$

where the density  $\rho_M(x)$  is defined in Eq. (49). This equation in turn is easiest to prove by transforming into a product in the Matsubara frequency space, simplifying using partial fractions, and then transforming back to time domain. We next note that Eqs. (A3) and (A4) imply

$$\begin{aligned} \Psi(\tau) &= -2\lambda X(\tau)\mathbf{g}(\tau)\mathbf{g}(-\tau), \\ \chi(\tau) &= -2\lambda X(\tau)X(-\tau)\mathbf{g}(\tau), \end{aligned} \quad (\text{B4})$$

so that taking Fourier transforms is simplest if we first multiply out as in Eq. (B2), leading to Eq. (50).

- 
- <sup>1</sup>B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).  
<sup>2</sup>B. S. Shastry, *Phys. Rev. B* **87**, 125124 (2013); D. Hansen and B. S. Shastry, *ibid.* **87**, 245101 (2013).  
<sup>3</sup>Aficionados will recognize this as the case for the virial expansion in statistical mechanics, where a low-density expansion can often be found for very strongly interacting systems. The critical density (where a phase transition occurs) is a natural boundary for such an expansion. If one knows *a priori* that there is no critical density in a given problem such as the Anderson model, then this expansion can be pushed to arbitrarily high densities. In the very different context of the theory of integrable systems in one dimension, this is precisely the reason why the *asymptotic Bethe ansatz* works very well, even yielding exact results in regimes where its validity seems *a priori* only approximate, as explained in Bill Sutherland, *Beautiful Models* (World Scientific, Singapore, 2004).  
<sup>4</sup>P. W. Anderson, *Phys. Rev.* **124**, 41 (1961).  
<sup>5</sup>P. W. Anderson, *Phys. Rev.* **164**, 352 (1967); P. W. Anderson and G. Yuval, *Phys. Rev. Lett.* **23**, 89 (1969).  
<sup>6</sup>F. D. M. Haldane, *Phys. Rev. Lett.* **40**, 416 (1978).  
<sup>7</sup>K. G. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975).  
<sup>8</sup>H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, *Phys. Rev. B* **21**, 1044 (1980).  
<sup>9</sup>A. C. Hewson, A. Oguri, and D. Meyer, *Eur. Phys. J. B* **40**, 177 (2004); K. Edwards, A. C. Hewson, and V. Pandis, *Phys. Rev. B* **87**, 165128 (2013).  
<sup>10</sup>D. E. Logan, M. P. Eastwood, and M. A. Tusch, *J. Phys. Condens. Matter* **10**, 2673 (1998); M. T. Glossop and D. E. Logan, *ibid.* **14**, 6737 (2002).  
<sup>11</sup>N. Bickers, *Rev. Mod. Phys.* **59**, 845 (1987).  
<sup>12</sup>N. Read and D. M. Newns, *J. Phys. C: Solid State Phys.* **16**, L1055 (1983); N. Read, *ibid.* **18**, 2651 (1985).  
<sup>13</sup>P. Coleman, *Phys. Rev. B* **29**, 3035 (1984); S. E. Barnes, *J. Phys. F: Met. Phys.* **6**, 1375 (1976).  
<sup>14</sup>N. Andrei, *Phys. Rev. Lett.* **45**, 379 (1980); P. B. Wiegmann, *Sov. Phys.–JETP* **31**, 392 (1980).  
<sup>15</sup>A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, UK, 1993).  
<sup>16</sup>K. Yosida and K. Yamada, *Prog. Theor. Phys. Suppl.* **46**, 244 (1970).  
<sup>17</sup>P. Nozières, *J. Low. Temp. Phys.* **17**, 31 (1974).  
<sup>18</sup>A. C. Hewson, *Phys. Rev. Lett.* **70**, 4007 (1993).  
<sup>19</sup>B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).  
<sup>20</sup>G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).  
<sup>21</sup>H. O. Frota and L. N. Oliveira, *Phys. Rev. B* **33**, 7871 (1986).  
<sup>22</sup>O. Sakai, Y. Shimuzu, and T. Kasuya, *J. Phys. Soc. Jpn.* **58**, 162 (1989).  
<sup>23</sup>T. A. Costi and A. C. Hewson, *J. Phys.: Condens. Matter* **5**, L361 (1993).  
<sup>24</sup>T. A. Costi, J. Kroha, and P. Wölfle, *Phys. Rev. B* **53**, 1850 (1996). We thank the authors for providing us with the digital versions of their results.  
<sup>25</sup>In the notation of Ref. 1, Eq. (58), this corresponds to writing  $Y_1(\tau_i, \tau_f) = -\gamma(\tau_i)\Delta(\tau_i, \tau_f)$ .  
<sup>26</sup>E. Perepelitsky and B. S. Shastry, *Ann. Phys. (NY)* **338**, 283 (2013).  
<sup>27</sup>J. Friedel, *Can. J. Phys.* **54**, 1190 (1956).  
<sup>28</sup>J. S. Langer and V. Ambegaokar, *Phys. Rev.* **164**, 498 (1967).  
<sup>29</sup>D. C. Langreth, *Phys. Rev.* **150**, 516 (1966).  
<sup>30</sup>J. M. Luttinger and J. C. Ward, *Phys. Rev.* **118**, 1417 (1960).  
<sup>31</sup>J. W. Rasul and A. C. Hewson, *J. Phys. C: Solid State Phys.* **17**, 3337 (1984).  
<sup>32</sup>To recover Eq. (42), we may use Eq. (21) and the Fermi liquid assumption of  $\text{Im} \Sigma_{\text{DM}}(0) = 0$  so that  $\rho_G(0) = \frac{1}{\pi} \frac{\Gamma_0(1-\frac{u_d}{2})^2}{\Gamma_0^2(1-\frac{u_d}{2})^2 + [\epsilon_d + \text{Re} \Sigma_{\text{DM}}(0)]^2}$ , and combine with Eq. (41).  
<sup>33</sup>R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, arXiv:1309.5284.  
<sup>34</sup>X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013)

# Extremely correlated Fermi liquid theory meets dynamical mean-field theory: Analytical insights into the doping-driven Mott transition

R. Žitko,<sup>1,2</sup> D. Hansen,<sup>3</sup> E. Perepelitsky,<sup>3</sup> J. Mravlje,<sup>1</sup> A. Georges,<sup>4,5,6</sup> and B. S. Shastry<sup>3</sup>

<sup>1</sup>*Jozef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

<sup>2</sup>*Faculty for Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia*

<sup>3</sup>*Physics Department, University of California Santa Cruz, California 95064, USA*

<sup>4</sup>*Centre de Physique Théorique, École Polytechnique, CNRS, 91128 Palaiseau Cedex, France*

<sup>5</sup>*Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*

<sup>6</sup>*DPMC-MaNEP, Université de Genève, CH-1211 Genève, Switzerland*

(Received 20 September 2013; revised manuscript received 9 December 2013; published 30 December 2013)

We consider a doped Mott insulator in the large dimensionality limit within both the recently developed extremely correlated Fermi liquid (ECFL) theory and the dynamical mean-field theory (DMFT). We show that the general structure of the ECFL sheds light on the rich frequency dependence of the DMFT self-energy. Using the leading Fermi liquid form of the two key auxiliary functions introduced in the ECFL theory, we obtain an analytical ansatz, which provides a good quantitative description of the DMFT self-energy down to hole doping level  $\delta \simeq 0.2$ . In particular, the deviation from Fermi liquid behavior and the corresponding particle-hole asymmetry developing at a low-energy scale are well reproduced by this ansatz. The DMFT being exact at large dimensionality, our study also provides a benchmark of the ECFL in this limit. We find that the main features of the self-energy and spectral line shape are well reproduced by the ECFL calculations in the  $O(\lambda^2)$  minimal scheme, for not too low doping level  $\delta \gtrsim 0.3$ . The DMFT calculations reported here are performed using a state-of-the-art numerical renormalization-group impurity solver, which yields accurate results down to an unprecedentedly small doping level  $\delta \lesssim 0.001$ .

DOI: [10.1103/PhysRevB.88.235132](https://doi.org/10.1103/PhysRevB.88.235132)

PACS number(s): 71.10.Ay, 71.10.Fd, 71.30.+h

## I. INTRODUCTION

Strong electronic correlations constitute one of the major challenges in condensed-matter physics and continue to inspire new theoretical approaches. In search for novel functionalities, new materials are being synthesized on a regular basis, giving the field a steady impetus. Significant progress in the understanding of electronic correlations has been achieved from the dynamical mean-field theory (DMFT), in which the self-energy is assumed to be momentum independent (see Ref. 1 for a review). This theory becomes exact in the limit of infinite dimensionality.

The situation in low dimensions has further challenges relating to the  $k$  dependence of the self-energy, and thus new methods for strongly correlated electrons continue to be developed. One promising approach is Shastry's extremely correlated Fermi liquid theory (ECFL), developed in a recent series of papers.<sup>2-5</sup> This theory starts from the infinite- $U$  limit and is based on the Schwinger equation of motion for Gutzwiller projected electrons, these noncanonical objects requiring special attention. The theory leads to a set of analytical expressions that are in principle exact. So far, solutions of the second-order expansion of these expressions in a partial projection parameter  $\lambda$  are available. They can be obtained for any lattice by an iterative process analogous to the skeleton diagram method. The ECFL theory expressions have been successfully applied to account for the angle resolved photoemission spectroscopy (ARPES) line shapes of cuprate superconductors<sup>6,7</sup> in the normal state.

In this work, we perform a comparative study of these two methods. We use as a test bed the single-band doped Hubbard model at strong coupling  $U$ , in the limit of large

dimensionality. This limit leads to simplifications in the ECFL theory, which we introduce here (the details of the formalism are provided elsewhere<sup>8</sup>). The comparison focuses on the frequency dependence of the self-energy and single-particle spectral line shapes, and their evolution as the Mott insulator is approached by reducing the doping level  $\delta$  [defined in Eq. (5)].

The first outcome of the present work is that, by looking at the DMFT results within an ECFL perspective, we are able to obtain new analytical insights into the DMFT description of the doping-driven Mott transition. Within the DMFT, the single-particle self-energy  $\Sigma(\omega)$  displays a rich and complex frequency dependence. This has been known for some time (see, e.g., Ref. 9 for a recent study), but is further investigated in the present work down to unprecedentedly low doping levels  $\delta \lesssim 0.001$  using a state-of-the-art numerical renormalization group (NRG) solution of the DMFT equations. Local Fermi liquid behavior  $\text{Im}\Sigma \propto \omega^2 + (\pi T)^2$  is obeyed only below a very low energy scale. Above this energy scale, a marked particle-hole asymmetry develops, a feature that is beyond the Fermi liquid theory. Furthermore, the strong suppression of spectral weight in the intermediate range of energies separating the quasiparticle peak from the lower Hubbard band corresponds to a marked quasipole in the self-energy.

We show that all of these features can be well reproduced by constructing an analytical ansatz for the one-particle self-energy, which is directly motivated by the ECFL construction. The ECFL introduces two key quantities,  $\Psi$  and  $\chi$ , which play the role of auxiliary self-energies in the Schwinger construction. The proposed analytical ansatz is obtained by retaining only the dominant Fermi liquid terms in the low-frequency expansion of these auxiliary quantities. This is found to provide a satisfactory fit of the DMFT results for

doping levels  $\delta \gtrsim 0.2$ . Hence, quite remarkably, the marked deviations from Fermi liquid behavior, and the particle-hole asymmetry in the physical single-particle self-energy, can be accounted for by an underlying Fermi liquid form of the ECFL auxiliary quantities. For very large  $U$ , and especially for very small doping levels, additional structures appear in the DMFT results, which are not present in this simplest ECFL ansatz, and presumably require additional terms beyond the Fermi liquid ones in the auxiliary functions  $\Psi$  and  $\chi$ .

Another synergistic outcome of our study is that, because the DMFT provides an exact solution in the limit of large dimensionality, it can be used to benchmark the ECFL in this limit. We present here quantitative results obtained within the minimal scheme implementation of the ECFL in high dimensions,<sup>8</sup> giving rise to an expansion to order  $\lambda^2$  in the projection parameter  $\lambda$ . We find that the main features of the self-energy and the spectral line shape are well reproduced by the  $O(\lambda^2)$  ECFL calculations, on a semiquantitative level, for not too low doping  $\delta \gtrsim 0.3$ . Improvement will require further developments of the ECFL approach. Since the DMFT is able to handle any finite  $U$ , while the ECFL construction is motivated by the very large  $U$  limit, this comparison also sheds light on the adiabatic connection between the regime of moderate and extreme correlations.

We emphasize that ECFL can be used on two different levels. On one level, it provides a functional form for the physical Green's function and the corresponding self-energy in terms of the auxiliary ECFL self-energies  $\Psi(\omega)$  and  $\chi(\omega)$ . By assuming the simplest Fermi liquid form for these two self-energies over a certain frequency range centered around  $\omega = 0$ , we successfully fit the physical self-energy obtained through DMFT in this frequency range for  $\delta \gtrsim 0.2$ . This is remarkable since the frequency range used is substantially larger than the characteristic frequency at which the physical self-energy begins to deviate from Fermi liquid behavior, and even encompasses the quasipole in the physical self-energy at negative frequencies. This phenomenological approach to ECFL is the one used in the first five sections of the paper, and the results of this fit are displayed in Figs. 8, 9, and 10. On the second level, ECFL provides a microscopic theory by which one can obtain concrete results for  $\Psi$  and  $\chi$  via an expansion in the projection parameter  $\lambda$ . In the remainder of the paper, the results obtained from the  $O(\lambda^2)$  theory are benchmarked against the results obtained from DMFT, which are exact in the limit of infinite dimensions. In the long run, further combined use of the ECFL and the DMFT approaches could lead to a better understanding of the momentum dependence of the self-energy that becomes important in lower dimensions.

The paper is organized as follows. After defining the model in Sec. II, the general structure of the ECFL formalism is reviewed in Sec. III. In Sec. IV, we present detailed DMFT results for the hole-doped Hubbard model using high-precision Wilson's NRG as a solver. In Sec. V, the DMFT self-energies are interpreted in light of the ECFL-motivated analytical expressions. The second part of the paper is devoted to the  $O(\lambda^2)$  ECFL minimal implementation. The basic equations and their simplification in infinite dimensions are established in Sec. VI, and in Sec. VII a quantitative comparison is made to the DMFT results.

## II. MODEL

We study the Hubbard model defined by the Hamiltonian

$$H = \sum_{k\sigma} \varepsilon_k c_{k\sigma}^\dagger c_{k\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \quad (1)$$

where  $\varepsilon_k$  is the bare band dispersion relation obtained by Fourier transforming the hopping matrix. In this study we consider a doped Hubbard model with nearest-neighbor hopping on a Bethe lattice, with semicircular density of states:

$$\rho_0(\varepsilon) = \frac{2}{\pi D^2} \sqrt{D^2 - \varepsilon^2}, \quad (2)$$

where  $D$  is the half bandwidth, and thus any sum over the band energy can be converted to an integral as:

$$\frac{1}{N} \sum_k \mathcal{A}(\varepsilon_k) \rightarrow \int_{-D}^D d\varepsilon \rho_0(\varepsilon) \mathcal{A}(\varepsilon). \quad (3)$$

We note that the Fermi energy  $\varepsilon_F$  satisfies

$$\sin^{-1} \left( \frac{\varepsilon_F}{D} \right) + \left( \frac{\varepsilon_F}{D} \right) \sqrt{1 - \left( \frac{\varepsilon_F}{D} \right)^2} = -\frac{\pi}{2} (1 - n), \quad (4)$$

and vanishes near  $n \sim 1$  as  $\varepsilon_F = -\frac{\pi}{4} (1 - n) D$ . The hole doping level  $\delta$  is related to the particle density  $n$  ( $n = N/N_{\text{sites}}$ ) as:

$$\delta = 1 - n. \quad (5)$$

We will use  $\rho_Q(\omega)$  as a shorthand notation for the spectral function associated with any relevant quantity  $Q(i\omega_n)$  (Green's function, self-energy):

$$\rho_Q(\omega) = -\frac{1}{\pi} \text{Im}[Q(\omega + i0^+)]. \quad (6)$$

## III. ECFL: GENERAL FRAMEWORK

### A. ECFL formalism

The ECFL methodology has been discussed extensively in recent literature;<sup>2,3</sup> here we highlight only the aspects that are of relevance to this work. ECFL deals with Gutzwiller-projected states obtained in the limit of  $U \rightarrow \infty$ , with the no-double-occupancy constraint built into the electron operators, leading to the well-known  $t$ - $J$  model. This results in a noncanonical theory, where familiar Feynman diagram methods fail due to the absence of Wick's theorem. The ECFL formalism is an exact alternative to the Feynman diagram technique. Instead it works with the Schwinger equations of motion for the projected electrons. It provides results for the electronic Green's functions that describe the physics of the low-energy sector in the problem, namely the dynamics of the quasiparticle (QP) states near the Fermi energy and of the lower Hubbard band (LHB).

For our purposes, we need to express the ECFL theory in the large-dimensionality limit. The related technical problems outlined in Ref. 10 (paragraph 3) have been recently solved<sup>8</sup> by analyzing the infinite-dimensional limit of the Schwinger equations of motion in the ECFL.<sup>11</sup>

The exact mapping of the momentum-independent self-energy of the infinite-dimensional Hubbard model onto that

of a self-consistent Anderson impurity model<sup>12</sup> provides a roadmap for a suitable formulation of the ECFL equations in this limit.

In the simplest version of the ECFL theory,<sup>13</sup> the physical (i.e., projected) electronic Green's function is expressed as a product of an auxiliary Green's function  $\mathbf{g}(k)$  and a caparison term denoted in the present work as  $\tilde{\mu}(k)$ . Thus,

$$\mathcal{G}(k) = \tilde{\mu}(k) \times \mathbf{g}(k), \quad (7)$$

where  $k \equiv (\vec{k}, i\omega_k)$  and  $\omega_k$  is a fermionic Matsubara frequency. Here  $\mathbf{g}(k)$  is a Fermi liquidlike Green's function

$$\mathbf{g}^{-1}(k) = i\omega_k + \mu - \left(1 - \frac{n}{2}\right) \varepsilon_k - \Phi(k), \quad (8)$$

and  $\mu$  is the chemical potential. The factor  $\tilde{\mu}(k)$  (here distinguished from the chemical potential  $\mu$  by the tilde), plays the role of an adaptive spectral weight, and is given by

$$\tilde{\mu}(k) = 1 - \frac{n}{2} + \Psi(k), \quad (9)$$

with  $\Psi(k)$  vanishing at infinite frequency. The functions  $\Phi(k)$  and  $\Psi(k)$  are the twin self-energies in the theory, and are exactly defined as the appropriate functional derivatives of  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  respectively.<sup>2,3</sup> The term  $\tilde{\mu}(k)$  is termed the caparison (i.e., dressing) factor, since it provides a second layer of renormalization to the propagator  $\mathbf{g}$ , which is already dressed by  $\Phi$ . Both Green's functions satisfy an identical number sum rule  $\sum_{k, \omega_k} \mathcal{G}(k) = \frac{n}{2} = \sum_{k, \omega_k} \mathbf{g}(k)$ ; this enables us to satisfy the Luttinger-Ward volume theorem.

In the large- $d$  limit, a further simplification can be established:<sup>8</sup>  $\Psi$  is independent of  $\vec{k}$  and  $\Phi$  is decomposable into two  $\vec{k}$ -independent functions,

$$\Psi(k) = \Psi(i\omega_k), \quad (10)$$

$$\Phi(k) = \chi(i\omega_k) + \varepsilon_k \Psi(i\omega_k), \quad (11)$$

i.e., the two frequency-dependent (but  $\vec{k}$ -independent) functions  $\chi$  and  $\Psi$  determine the Green's function.

The single-electron physical (Dyson) self-energy  $\Sigma$  is defined from the single-electron Green's function  $\mathcal{G}$  in the usual manner, as (using the analytic continuation  $i\omega_k \rightarrow \omega + i\eta$ ,  $\eta = 0^+$ ):

$$\mathcal{G}(k, \omega + i\eta) = \frac{1}{\omega + i\eta + \mu - \varepsilon_k - \Sigma(\omega + i\eta)}. \quad (12)$$

Within the large-dimensional ECFL, the Dyson self-energy  $\Sigma$  can be related to  $\Psi$  and  $\chi$  as follows:

$$\Sigma(\omega + i\eta) - \mu - \omega = \frac{\chi(\omega + i\eta) - \mu - \omega}{1 - \frac{n}{2} + \Psi(\omega + i\eta)}. \quad (13)$$

We see that the Dyson self-energy is manifestly momentum independent in this limit. Note also that, as seen from (13), its real part grows linearly with  $\omega$  as  $\omega \rightarrow \infty$ . This is a consequence<sup>14</sup> of the Gutzwiller projection in the  $U \rightarrow \infty$  limit. At finite  $U$ , this behavior is regularized at high-enough frequency and  $\Sigma$  goes to a constant.

For a concrete implementation, the ECFL formalism allows for a perturbative expansion in a projection parameter  $\lambda \in [0, 1]$ , ultimately identified with the double-occupancy

density.<sup>3</sup> The theory to  $O(\lambda^2)$  is expected to be quantitatively accurate for densities up to  $n \lesssim 0.7$ .<sup>15</sup> We postpone the description of these equations to Sec. VI, but note an important general insight gained from examining and evaluating such an expansion<sup>2,3,15</sup>; the two self-energies  $\chi$  and  $\Psi$  have simple Fermi liquid functional forms, with a dissipative part that is quadratic in  $\omega$ , at sufficiently low energies (see Fig. 11 in Sec. VI). This insight is used in the following to obtain a low-energy expansion for the Green's function.

## B. Low-frequency expansion of self-energies and Green's functions

In this section we derive the low-frequency behavior of the Green's function and self-energy within the ECFL. We obtain an analytical expression, which will be used to interpret and fit the DMFT results in Sec. V. We show in particular how a characteristic particle-hole asymmetry in the Dyson self-energy is generated even when the expansion of  $\text{Im}\chi$  and  $\text{Im}\Psi$  is limited to the particle-hole symmetric lowest-order Fermi liquid terms.

Indeed, as mentioned above, the first few terms of a systematic  $\lambda$  expansion of the ECFL equations indicate that the self-energies  $\Psi$  and  $\chi$  are very similar functions and resemble the self-energy of a Fermi liquid at low enough  $T, \omega$ , with suitable scale constants.

For low  $\omega$  and low  $T$ , up to a low-frequency cutoff scale  $\Omega_c$ , so that  $|\omega| \leq \Omega_c \ll D$ , we define (with  $k_B = 1$ )

$$\mathcal{R}(\omega, T) \equiv \pi[\omega^2 + (\pi T)^2], \quad (14)$$

and write a Fermi liquidlike expansion for the complex ECFL self-energies:

$$\Psi(\omega) \sim \Psi_0 + c_\Psi \omega + \frac{i}{\gamma_\Psi} \mathcal{R}(\omega, T) + \Psi_{\text{rem}}(\omega), \quad (15)$$

$$\chi(\omega) \sim \chi_0 - c_\chi \omega - \frac{i}{\Omega_\chi} \mathcal{R}(\omega, T) + \chi_{\text{rem}}(\omega), \quad (16)$$

where

$$c_\Psi = 2 \frac{\Omega_c}{\gamma_\Psi} \quad \text{and} \quad c_\chi = 2 \frac{\Omega_c}{\Omega_\chi}. \quad (17)$$

$\Omega_\chi$  and  $\gamma_\Psi$  are parameters that determine the curvatures of the two imaginary parts, with  $\Omega_\chi$  having the dimensions of energy, while  $\gamma_\Psi$  has the dimensions of the square of energy. Consequently,  $c_\Psi$  has the dimensions of an inverse energy, while  $c_\chi$  is dimensionless. The terms  $\Psi_{\text{rem}}(\omega)$  and  $\chi_{\text{rem}}(\omega)$  in Eq. (15) represent the remainders containing the leading corrections to the Fermi liquid behavior, of the type  $O(\omega^2)$  for the real part and  $O(\omega^3)$  for the imaginary parts of these functions. In the initial analysis below, we simply ignore these terms. They can be readily incorporated to find a systematic improvement of the fits, and lead to further corrections to the low-frequency behavior of the imaginary part of the Dyson self-energy beyond the terms considered here. Note that in Eq. (17) the constants  $c_\Psi$  and  $c_\chi$  also receive contributions from higher terms beyond the quadratic. Hence these approximate relations become exact if we retain the imaginary terms only to quadratic order, i.e., assuming

$\rho_\Psi \equiv -\frac{\mathcal{R}(\omega, T)}{\pi \gamma_\Psi}$  and  $\rho_\chi \equiv \frac{\mathcal{R}(\omega, T)}{\pi \Omega_\chi}$ . In general, however,  $c_\Psi$  and  $c_\chi$  can be considered as additional free parameters.

Some further remarks about this expansion are called for.

(i) Expressions (15) and (16) are of the standard Fermi liquid type (symmetric in  $\omega$  for the imaginary parts). Nonetheless, when processed through the ECFL formalism, they lead to important contributions to the imaginary part of the Dysonian self-energy which are *antisymmetric in  $\omega$* . Revealing the origin of this important asymmetry is one of the main strengths of the ECFL analysis.

(ii) We shall find that as we approach half filling,  $\Omega_\chi$  and  $\gamma_\Psi$  turn out to be similar functions of the electron density, in view of their parallel role in the two self-energies within the  $\lambda$  expansion. In the analysis below, we will find that as  $n \rightarrow 1$ , it is consistent to choose  $\Omega_\chi, \gamma_\Psi \sim \delta$ , where  $\delta = 1 - n$ , so that the Mott insulating limit is reached smoothly.

(iii) The energy scale  $\Omega_c$ , which determines the range of frequencies where the quadratic behavior of  $\text{Im } \Psi$  and  $\text{Im } \chi$  applies, is itself a function of the density. It shrinks linearly with  $\delta$  as  $n \rightarrow 1$ , and therefore  $c_\Psi$  and  $c_\chi$  are finite as  $n \rightarrow 1$ . We should note that these are leading-order assumptions in  $\delta$ .

Thus we find at low  $T, \omega$ :

$$\begin{aligned} & \mathcal{G}(k, \omega + i\eta) \\ & \sim \frac{\alpha_0 + c_\Psi \omega + \frac{i}{\gamma_\Psi} \mathcal{R}}{\omega(1 + c_\chi) + \mu + \frac{i}{\Omega_\chi} \mathcal{R} - \varepsilon_k \{ \alpha_0 + c_\Psi \omega + \frac{i}{\gamma_\Psi} \mathcal{R} \}}, \end{aligned} \quad (18)$$

where we have introduced

$$\alpha_0 \equiv 1 - \frac{n}{2} + \Psi_0, \quad (19)$$

$$A(k, \omega) \sim \left( \frac{\alpha_0^2}{\pi \Omega_\Sigma} \right) \frac{\mathcal{R} \left( 1 - \frac{\omega}{\Delta} \right)}{\{ (1 + c_\chi - c_\Psi \varepsilon_k) \omega - \alpha_0 (\varepsilon_k - \varepsilon_F) \}^2 + \{ \alpha_0^2 \mathcal{R}^2 / \Omega_\Sigma^2 \}}, \quad (24)$$

where

$$\Omega_\Sigma \equiv \alpha_0 \frac{\Omega_\chi \gamma_\Psi}{\gamma_\Psi - \varepsilon_F \Omega_\chi}, \quad (25)$$

$$\Delta \equiv \frac{\alpha_0^2 \gamma_\Psi \Omega_\chi}{\Omega_\Sigma \{ (1 + c_\chi) \Omega_\chi - c_\Psi \gamma_\Psi \}}. \quad (26)$$

In terms of the wave-function renormalization factor  $Z$ ,

$$A(k, \omega) \sim \left( \frac{Z^2}{\pi \Omega_\Sigma} \right) \frac{\mathcal{R} \left( 1 - \frac{\omega}{\Delta} \right)}{\{ \omega - Z(\varepsilon_k - \varepsilon_F) \}^2 + \{ Z^2 \mathcal{R}^2 / \Omega_\Sigma^2 \}}. \quad (27)$$

We thus obtain the following final form for the low-energy expression of the Dysonian self-energy:

$$\begin{aligned} \text{Im} \Sigma(\omega) & \sim -\frac{\mathcal{R}}{\Omega_\Sigma} \frac{1 - \frac{\omega}{\Delta}}{\{ 1 + \omega c_\Psi / \alpha_0 \}^2 + \mathcal{R}^2 / (\alpha_0 \gamma_\Psi)^2}, \\ \text{Re} \Sigma(\omega) & \sim \alpha_0 \varepsilon_F + \omega \\ & \quad - \frac{\{ \varepsilon_F + \omega(1 + c_\chi) / \alpha_0 \} (1 + \omega c_\Psi / \alpha_0) + \mathcal{R}^2 / (\alpha_0^2 \Omega_\chi \gamma_\Psi)}{\{ 1 + \omega c_\Psi / \alpha_0 \}^2 + \mathcal{R}^2 / (\alpha_0 \gamma_\Psi)^2}, \end{aligned} \quad (28)$$

where we recall that  $\mathcal{R}(\omega, T)$  is defined in Eq. (14). These expressions, and in particular that of  $\text{Im} \Sigma$ , are among the key

and  $\chi_0$  has been absorbed into the chemical potential  $\mu$ . The entire momentum dependence is contained in  $\varepsilon_k$ . At  $T = 0$  and  $\omega = 0$  we must require  $\mathcal{G}^{-1}(k_F, 0) = 0$ , so we need to set

$$\mu = \alpha_0 \varepsilon_F. \quad (20)$$

At low  $\omega + i\eta$  and a fixed  $\vec{k}$ , we can write a useful expression

$$\mathcal{G}^{-1} \sim -\varepsilon_k + \frac{\omega(1 + c_\chi) + \alpha_0 \varepsilon_F + \frac{i}{\Omega_\chi} \mathcal{R}}{\alpha_0 + c_\Psi \omega + \frac{i}{\gamma_\Psi} \mathcal{R}} \quad (21)$$

and therefore

$$\Sigma(\omega + i\eta) \sim \alpha_0 \varepsilon_F + \omega - \frac{\omega(1 + c_\chi) + \alpha_0 \varepsilon_F + \frac{i}{\Omega_\chi} \mathcal{R}}{\alpha_0 + c_\Psi \omega + \frac{i}{\gamma_\Psi} \mathcal{R}}. \quad (22)$$

Note that we adjusted the self-energy so that  $\text{Re } \Sigma(0) = \mu - \varepsilon_F$ , thereby placing the zero-energy pole at the Fermi momentum. We now extract the wave-function renormalization factor  $Z$  from  $Z^{-1} = \frac{\partial}{\partial \omega} \mathcal{G}^{-1}(k, \omega)|_{\omega=0}$  as

$$Z = \frac{\alpha_0}{1 + c_\chi - \varepsilon_F c_\Psi}. \quad (23)$$

Using the above expansion we find the spectral function  $\rho_{\mathcal{G}}(k, \omega)$  [or equivalently  $A(k, \omega)$ , as denoted in the experimental literature], at low  $\omega$  and  $k \sim k_F$

results of the present paper, and will be used below in order to fit and interpret the DMFT data.

If we take the  $\Psi_{\text{rem}}$  and  $\chi_{\text{rem}}$  terms in Eq. (15) and Eq. (16) into account, then Eq. (28) receives higher-order polynomial corrections in  $\omega$  in both the numerator and denominator. Let us also note that Eq. (27) is of the form of a phenomenological version of the ECFL theory that has been recently tested against experimental data with considerable success, in some cases after adding a constant times  $(\varepsilon_k - \varepsilon_F)$  in the numerator.<sup>2,6,7</sup>

### C. Low-doping limit $n \rightarrow 1$

At  $T = 0$  and near half filling we get  $\varepsilon_F = -\frac{\pi}{4}\delta D$  from Eq. (4). Further, from the single assumption that  $\Psi_0 = -\frac{n}{2}$  near half filling,<sup>16</sup> we find that the self-energy and wave function renormalization factor scale correctly with  $\delta$  as  $\delta \rightarrow 0$ . This assumption gives  $\alpha_0 = \delta$ , and a scaling of various energy scales with  $\delta$ . In particular, we find from the equations that  $\Omega_\chi \sim \gamma_\Psi \sim \Omega_c \sim \delta$ . This, together with Eq. (17), leads to  $c_\Psi \sim O(1)$  and  $c_\chi \sim O(1)$ . This is consistent with the scaling behavior described in remarks (ii) and (iii) in the previous section. Keeping the dominant terms in Eqs. (23), (25), and (26), we find that

$$Z = \frac{\alpha_0}{1 + c_\chi}, \quad (29)$$

$$\Omega_\Sigma = \alpha_0 \Omega_\chi, \quad (30)$$

$$\Delta = \frac{\alpha_0 \gamma_\Psi}{\{(1 + c_\chi)\Omega_\chi - c_\Psi \gamma_\Psi\}}. \quad (31)$$

Near half filling ( $\delta \rightarrow 0$ ), we define

$$Z = \delta \bar{Z}, \quad (32)$$

$$\Omega_\Sigma = \delta^2 \bar{\Omega}_\Sigma, \quad (33)$$

$$\Delta \equiv \delta \bar{\Delta}. \quad (34)$$

All objects with an overline, such as  $\bar{Q}$ , are determined to be finite as  $\delta \rightarrow 0$ . Eq. (32) is expected on general grounds near the insulating limit: to leading order  $\mathcal{G}^{-1}(k, \omega) = \varepsilon_F - \varepsilon_k + \frac{\omega}{Z} + O(\omega^2)$ , and therefore the propagating solutions correspond to quasiparticles with an energy dispersion  $Z(\varepsilon_k - \varepsilon_F)$  that shrinks to zero at the insulating point  $n = 1$ . We find here that this occurs as a linear function of  $\delta$ . Eq. (33), together with Eq. (28), implies that at small  $\omega \sim O(\delta)$ , the imaginary part of the self-energy has a finite value. Further combined with a cutoff  $\Omega_c \sim O(\delta)$ , it gives  $\text{Re } \Sigma \sim \omega - 2\omega \frac{\Omega_c}{\Omega_\Sigma}$ , which is then consistent with the linear vanishing of  $Z$  in Eq. (32). Equation (34) shows that the particle-hole asymmetry in the spectral function increases as we approach half filling. Finally, we see that the spectral density  $\text{Im}\Sigma$  becomes a scaling function of  $\omega/\delta$  at low doping levels.

## IV. DOPED MOTT INSULATOR: SINGLE-SITE DMFT

The dynamical mean-field theory<sup>1</sup> is based on the fact<sup>17</sup> that in the limit of a large number of dimensions  $d$  the self-energy becomes a momentum-independent local quantity,  $\Sigma(\mathbf{k}, \omega) \rightarrow \Sigma(\omega)$ . This implies that the bulk problem for

$d \rightarrow \infty$  coincides with the problem of an interacting impurity embedded in an appropriate noninteracting bath.<sup>12</sup> DMFT formulates a practical prescription for finding this effective impurity problem and the self-consistency equation. For the Hubbard model, the corresponding impurity problem is the single-impurity Anderson model, which can be efficiently solved with the numerical-renormalization group (NRG) method.<sup>18–22</sup>

### A. NRG method

The NRG calculations have been performed with the discretization parameter  $\Lambda = 2$  using the discretization scheme with reduced systematic artifacts described in Ref. 22. Furthermore, the twist averaging over  $N_z = 8$  different discretization meshes has been used to reduce the oscillatory NRG discretization artifacts.<sup>23</sup> The truncation cutoff in the NRG was set in the energy space at  $10\omega_N$  (here  $\omega_N$  is the characteristic energy at the  $N$ th NRG step); such results are well converged with respect to the truncation. The U(1) charge conservation and SU(2) spin rotational invariance symmetries have been used explicitly. The raw spectral data (weighted  $\delta$  peaks) were collected in bins on a logarithmic mesh with 1000 bins per frequency decade, then the broadening scheme from Ref. 24 with  $\alpha = 0.2$  has been used to obtain the continuous representation of the spectral functions. To calculate the self-energy  $\Sigma$ , we have used the procedure<sup>25</sup> based on the following exact relation from equations of motion:

$$\Sigma_\sigma(z) = \frac{\langle\langle [d_\sigma, H_{\text{int}}]; d_\sigma^\dagger \rangle\rangle_z}{\langle\langle d_\sigma; d_\sigma^\dagger \rangle\rangle_z} = \frac{U \langle\langle n_{\bar{\sigma}} d_\sigma; d_\sigma^\dagger \rangle\rangle_z}{\langle\langle d_\sigma; d_\sigma^\dagger \rangle\rangle_z}. \quad (35)$$

Here  $d_\sigma$  is the impurity annihilation operator, while  $H_{\text{int}}$  is the interaction part of the Hamiltonian. The two correlators in this expression were computed using the full-density-matrix NRG algorithm.<sup>24,26</sup> To accelerate the convergence of the DMFT self-consistency loop, the Broyden mixing algorithm has been used.<sup>27</sup> This technique is particularly important to ensure the convergence at small doping as the Mott transition is approached. The Broyden solver has been used both to control the chemical potential to obtain the desired band filling and to apply the DMFT self-consistency equations.<sup>27</sup>

When performing the calculations in the large- $U$  limit, it is important to note that the upper Hubbard band (UHB) is outside the NRG discretization energy window. The correlator  $F_\sigma(z) = \langle\langle n_{\bar{\sigma}} d_\sigma; d_\sigma^\dagger \rangle\rangle_z$  receives a contribution

$$F_{\text{UHB}}(z) = \frac{w_{\text{UHB}}}{z - (\epsilon_d + U)} \xrightarrow{U \rightarrow \infty} \frac{w_{\text{UHB}}}{-U} \quad (36)$$

from the UHB, where  $w_{\text{UHB}}$  is the total weight of the upper Hubbard band, which in the  $U \rightarrow \infty$  limit is equal to  $n/2$ . The correlator  $F(z)$  in Eq. (35) is multiplied by a factor  $U$ , thus the UHB contribution to the numerator in the  $U \rightarrow \infty$  limit is equal to  $-w_{\text{UHB}}$ . It is crucial to correct the raw numerical results by making this subtraction when the UHB is outside the discretization window, otherwise the causality is very strongly violated. [No such subtraction is necessary for the correlator  $G_\sigma(z) = \langle\langle d_\sigma; d_\sigma^\dagger \rangle\rangle_z$ , because the UHB only makes an  $O(1/U)$  contribution to the denominator.]

An analysis of the convergence of the NRG results with respect to the variation of various parameters in the method is presented in Appendix B.

## B. DMFT results

### 1. Scaling of quasiparticle weight $Z$ vs doping level $\delta$

In this work, we only consider paramagnetic solutions. At low temperatures, the DMFT equations, depending on the strength of the interaction  $U$  and the doping  $\delta$ , give either an insulating or a metallic Fermi liquid state. The key quantity characterizing the metallic state is the quasiparticle residue

$$Z = \left( 1 - \left. \frac{\partial \text{Re}\Sigma(\omega)}{\partial \omega} \right|_{\omega=0} \right)^{-1}. \quad (37)$$

At  $\delta = 0$  (half-filled system), a metallic solution is found for  $U < U_c$  (this critical value of  $U$  is often denoted  $U_{c2}$  in the DMFT literature—the spinodal of the metallic solution<sup>1</sup>—and will be denoted  $U_c$  here for simplicity). For  $U > U_c$ , the DMFT equations only have a unique insulating solution. At  $U_c/D = 2.918$ , a Mott metal-insulator transition takes place, with characteristics similar to the Brinkman-Rice picture<sup>28</sup> in that the quasiparticle weight  $Z$  vanishes continuously and the quasiparticle effective mass diverges.

Away from half filling, i.e., for any  $\delta \neq 0$ , the solution is always metallic ( $Z > 0$ ); the Mott insulator ( $Z = 0$ ) only exists exactly at half filling (for  $U > U_{c1}$ , the spinodal of the insulator).<sup>1</sup> As  $\delta \rightarrow 0$  for  $U > U_c$ ,  $Z$  diminishes and vanishes at  $\delta = 0$ . This doping-driven Mott transition is illustrated in Fig. 1. In Figs. 1(a)–1(d) we plot the results of  $Z$  vs  $\delta$  for a set of values of  $U$ . It is seen that, when considered over a broad doping range  $\delta \lesssim 0.5$  [Figs. 1(a) and 1(b)], the overall doping dependence of  $Z$  is fairly linear at intermediate values of  $U/D$ , while at strong coupling (large  $U/D$ ) a marked curvature is seen (approximately fit by a power-law with exponent close to 1.4).

A plot of  $Z/\delta$  vs  $\delta$  focusing on the low-doping region [Figs. 1(c) and 1(d)] reveals, however, that the asymptotic low- $\delta$  behavior is actually linear,  $Z \propto \delta$  (except close to the multicritical point  $U = U_c, \delta = 0$  where sizable corrections are found). This is indeed the behavior expected within the Gutzwiller approach<sup>28–30</sup>: Figures 1(c)–1(d) thus confirm that DMFT obeys this mean-field behavior. The prefactor of this linear dependence is also decreasing with  $U$  in reasonable agreement with the Gutzwiller estimate<sup>30</sup>  $\sim (1 - U_c/U)^{-1/2}$ . Note that the results displayed here extend previous studies to much lower doping levels ( $\delta \lesssim 0.001$ ) than previously reported in the literature, due to the improvements in the NRG methodology.

### 2. Self-energy and spectral function: overview of the main structures

We now address the properties of the self-energy  $\Sigma(\omega)$  and one-particle spectral function in more detail.

An overview plot in Fig. 2(a) shows the main features in the local spectral function  $A(\omega)$  and in the imaginary part of  $\Sigma(\omega)$  in a broad frequency range.  $\text{Im}\Sigma$  has two very pronounced and sharp resonances (quasipoles). These resonances are responsible for the suppression of the spectral

weight in  $A(\omega)$  between the QP peak and the LHB and UHB, respectively. They are correspondingly positioned close to the minima of  $A(\omega)$ . In contrast to the half-filled case, where these resonances are symmetrically positioned on each side of  $\omega = 0$  at a scale<sup>31</sup>  $\propto \pm\sqrt{Z}$ , their locations in the doped case are no longer symmetric and will be discussed below. In addition, there are two broad humps in  $\text{Im}\Sigma$  in the frequency ranges associated with the two Hubbard bands. As  $U$  increases towards very large values at fixed doping, the UHB moves to higher frequencies, while the LHB and QP band gradually converge to their high- $U$  asymptotic form. This convergence is, however, rather slow and the spectra start to very closely agree with the asymptotic ones only for  $U$  on the order of  $100D$ .

In Fig. 2(b) we plot a closeup on the low-energy structures, i.e., the QP band and its vicinity. We notice that the Fermi liquid quadratic behavior of  $\text{Im}\Sigma$  is limited to a very narrow frequency interval, much smaller than the width of the QP peak itself. We also see (Fig. 2 and Fig. 5) that at low doping level,  $\text{Im}\Sigma$  develops a marked particle-hole asymmetry. These deviations to Fermi liquid behavior are discussed below in a more quantitative manner.

One of the goals of this work is to provide an analytical account of the complex frequency dependence of the self-energy that we just summarized. It should be kept in mind that the ECFL theory that we are going to use for this purpose works with the  $U = \infty$  model, which begins by throwing out the UHB altogether and deals only with the LHB. Thus the comparison carried out later in this paper refers only to the LHB and QP sector (no double occupancies), containing the interesting low-energy physics of the problem.

### 3. Dynamical particle-hole asymmetry

The local spectral function and self-energy are displayed on Fig. 3 for  $U = \infty$ , at two different doping levels (a small and large one, for comparison). An immediately apparent feature of these plots is the large asymmetry between holelike ( $\omega < 0$ ) and particlelike ( $\omega > 0$ ) excitations.

Indeed, for  $\omega < 0$ ,  $|\text{Im}\Sigma|$  increases rapidly from  $\omega = 0$  in order to connect with the negative-energy quasipole. The detailed form of this increase is somewhat different depending on the doping level. At large doping it is approximately parabolic, in continuity with the low- $\omega$  FL  $\sim \omega^2$  dependence. In contrast, at small doping, the low- $\omega$  parabolic dependence evolves into a more linearlike increase at higher frequency. The local spectral function also displays an almost complete suppression of the spectral weight between the QP peak and the LHB at low doping level, while this suppression is only modest at higher doping.

In contrast, for  $\omega > 0$ ,  $|\text{Im}\Sigma|$  rapidly flattens out after its initial FL increase. It has a plateaulike behavior with a broad maximum at large and intermediate doping level (the maximum is sharper at smaller doping). Overall,  $|\text{Im}\Sigma|$  remains much smaller at  $\omega > 0$  than at  $\omega < 0$ .

This asymmetry also reflects into the QP peak in the local spectral function  $A(\omega)$ , i.e., the local  $\rho_G(\omega)$ , which has a very asymmetric line shape. The decrease from its  $\omega = 0$  value  $A(0)$  is much faster on the  $\omega < 0$  side, in accordance with the large  $|\text{Im}\Sigma|$ . The detailed form of the line shape on the more

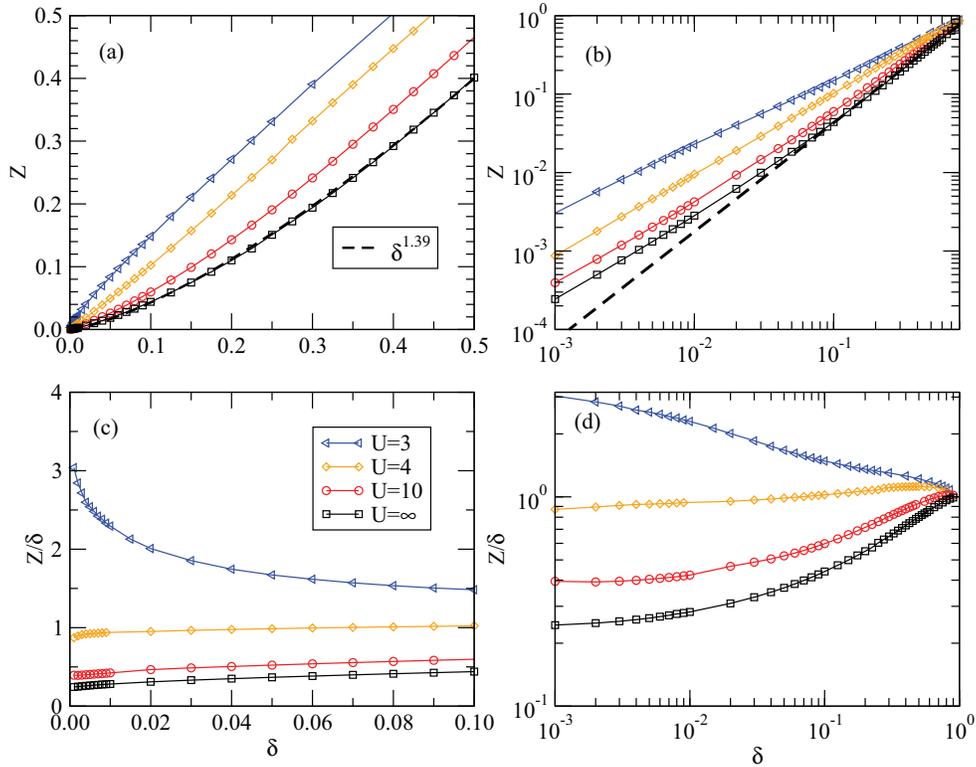


FIG. 1. (Color online) Doping-driven Mott transition within the DMFT: approach to the Mott insulating state for  $U > U_c = 2.918D$  with decreasing doping  $\delta$ . Top panels: doping dependence of  $Z$  for a range of  $\delta$  on a lin-lin (a) and on a log-log (b) plot. The dashed line is a fit to a power-law function  $Z = \delta^\alpha$  with  $\alpha = 1.39$ . Bottom panels:  $Z/\delta$  vs  $\delta$ (c), and  $Z/\delta$  vs  $\delta$  on a log-log plot (d).

extended  $\omega > 0$  side is different at lower and higher doping levels, with a convex and concave shape, respectively.

Finally, the particle-hole asymmetry has a very distinctive signature in the momentum-resolved spectra  $A(\varepsilon, \omega)$ , which are displayed in Fig. 4. It is seen there that the dispersion of the QP peak deviates from its low-energy form  $\omega_{QP} = Z(\varepsilon - \varepsilon_F)$  much more rapidly on the  $\omega > 0$  side, where a stronger dispersion closer to that of the bare band is rapidly found. This is mostly due to the distinct behavior of the real part  $\text{Re}\Sigma$  for positive and negative frequencies (shown later in Fig. 6). This finding, which is also supported by the ECFL results as discussed below, is one of the main predictions of our work. It calls for the development of momentum-resolved spectroscopies for unoccupied states (the dark side that is not directly accessible to ARPES). The physical significance of this dark side has also been recently pointed out in cluster-DMFT studies of the two-dimensional Hubbard model.<sup>32</sup>

#### 4. $\omega/Z$ scaling

Close to the Mott transition, all low-frequency properties are expected to scale with  $Z$ , i.e., be described by scaling functions<sup>1,33,34</sup> of  $\omega/Z$ . This is indeed the case, as demonstrated in Fig. 5 in which good data collapse is obtained in the lowest frequency range when plotted vs.  $\omega/Z$ . However, we also clearly observe that the scaling is limited to the asymptotic region of very small frequencies.

On Figs. 5(b) and 5(c), one can compare the evolution of the shape of the local spectral function, discussed above, as

a function of the doping level. One sees that the QP peak becomes increasingly asymmetric at very low doping. We also observe that the LHB has some internal structure, quite similar to that observed at half filling as the correlation-driven Mott transition is approached from the metallic side.<sup>22,35–37</sup>

In Fig. 6, the real and imaginary part of the self-energy are plotted against  $\omega/\delta D$ . The different panels cover different frequency ranges. While a better collapse of the different curves at very low frequency was obtained above when using  $\omega/ZD$  as a scaling variable, it is seen from the plots in a broader frequency range that the overall structures of the self-energy obey rather good scaling properties with respect to  $\omega/\delta D$ . For example, the sharp peak (quasipole) structure at  $\omega < 0$  in  $\text{Im}\Sigma$  is seen to be located at a frequency proportional to doping level ( $\omega_{\text{peak}} \simeq -0.7\delta D$ ). This peak in  $\text{Im}\Sigma$  is associated with the suppression of the spectral weight between the QP peak and the LHB in the spectral function. Correspondingly, it is associated with a resonancelike structure in  $\text{Re}\Sigma$ .

#### 5. Deviation from the low-frequency Fermi liquid behavior

From Figs. 6(b), (c) (for  $\text{Re}\Sigma$ ) and 6(e), (f) (for  $\text{Im}\Sigma$ ), one can visualize the low-frequency deviations from Fermi liquid behavior. The latter is indicated by the dashed straight and parabolic lines on this figure:  $\text{Re}\Sigma - \text{Re}\Sigma(0) = \omega(1 - 1/Z)$  and  $\text{Im}\Sigma \propto -(\omega/Z)^2$ .

When visualized on an intermediate frequency scale [Figs. 6(b) and 6(e)] it is seen that deviation from the FL

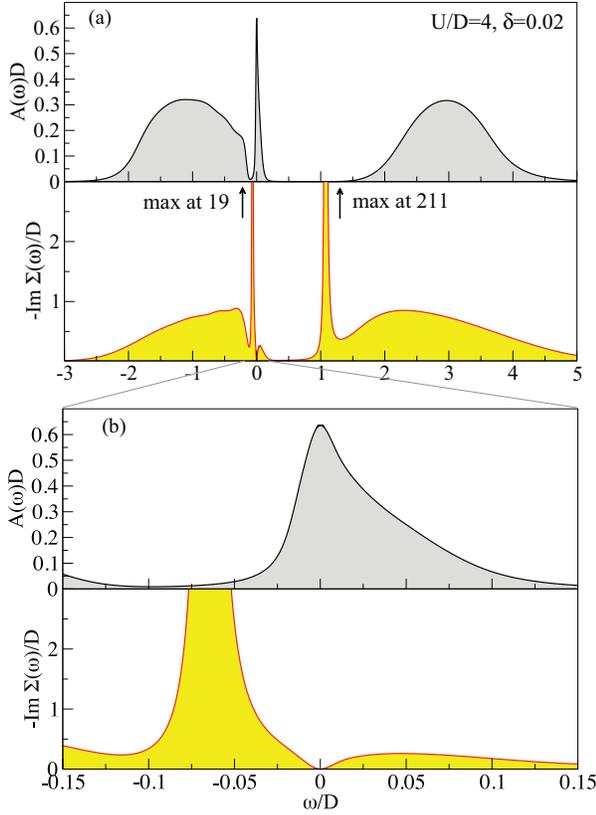


FIG. 2. (Color online) (a) Overview plot showing the full structure of spectra on high-frequency scales (lower Hubbard band, quasiparticle band, upper Hubbard band) at finite  $U = 4D$ . We show the DMFT local spectral function (top panel) and the imaginary part of the self-energy (bottom panel). (b) Closeup on the quasiparticle band at low frequencies.

behavior is more apparent on the  $\omega > 0$  side, in accordance with the particle-hole asymmetry discussed above and as pointed out in previous studies.<sup>9,38</sup>  $\text{Re}\Sigma$  deviates from linearity and flattens upwards for  $\omega > 0$ , resulting in the bending of the

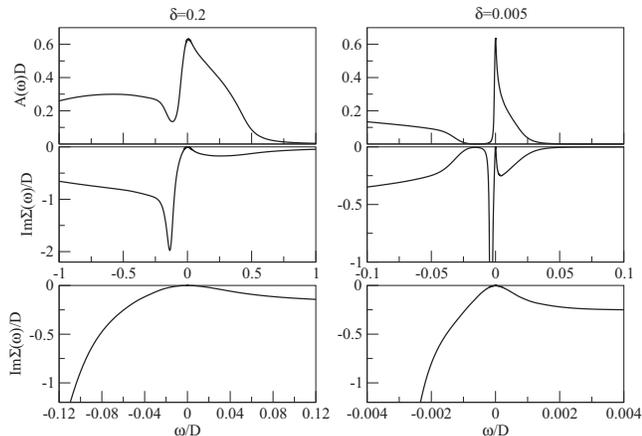


FIG. 3. Local spectral function and the imaginary part of the self-energy for large doping  $\delta = 0.2$  (left) and small doping  $\delta = 0.005$  (right), for  $U = \infty$ .

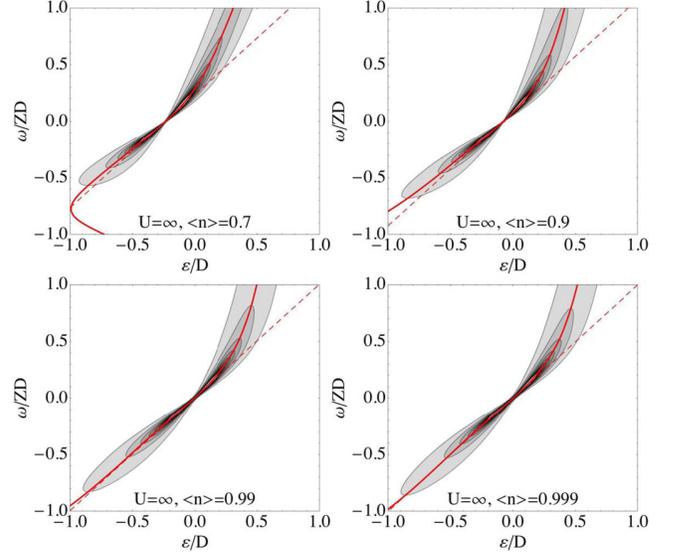


FIG. 4. (Color online) Intensity plots of the momentum ( $\epsilon_k$ -) resolved spectral function  $A(\epsilon, \omega)$  for  $U = \infty$  at four different doping levels, plotted as a function of  $\epsilon/D$  and  $\omega/ZD$ . The plain line locates the solution of the QP pole equation  $\omega + \mu - \text{Re}\Sigma(\omega) - \epsilon = 0$  (neglecting  $\text{Im}\Sigma$ ). By definition of the QP excitations, this line has slope unity (cf. dashed line) at low- $\omega$  when plotted in this manner since  $\omega_{\text{QP}} = Z(\epsilon - \epsilon_F)$  (i.e.,  $v_F^* = Zv_F$ ) within the DMFT.

dispersion of  $\omega > 0$  quasiparticles towards the noninteracting bare dispersion, displayed above on Fig. 4. Accordingly, the deviations from parabolic behavior in  $\text{Im}\Sigma(\omega)$  are much more pronounced on the positive frequency side.

Zooming further on the low-frequency range [Figs. 6(c) and 6(f)] allows one to locate more quantitatively the deviation from the FL behavior. At  $U = \infty$ , it is seen to occur at  $\omega_{\text{FL}}^* \simeq 0.1 ZD$ , which is of order  $0.025\delta D$  to  $0.05\delta D$  depending on  $\delta$ . In agreement with previous studies<sup>9</sup> at finite  $U$ , the scale below which FL is found to apply is seen to be a very low one. It is one order of magnitude smaller than the Brinkman-Rice scale  $\approx \delta D$ , which corresponds to scaling the bare bandwidth by the (inverse) of the effective mass. When converted to a temperature scale, the Brinkman-Rice scale roughly corresponds to the temperature at which QP excitations disappear altogether (and the resistivity approaches the Mott-Ioffe-Regel limit),<sup>9</sup> but it should not be identified with the much lower scale associated with deviations from FL behavior.

The low-frequency zooms in Figs. 6(c) and 6(f) actually reveal that the deviations from FL behavior are seen both on the  $\omega < 0$  and  $\omega > 0$  side, at similar scales  $\pm\omega_{\text{FL}}^*$ . This scale corresponds to a low-energy kink in  $\text{Re}\Sigma$ . The corresponding low-energy kink in the QP dispersion<sup>39–41</sup> is actually visible upon close examination of Fig. 4.

As seen on Figs. 5(b) and 5(c), the full collapse of the data is limited to very low frequencies. Two kinds of deviations from the universal behavior can be recognized. On the negative frequency side, at moderate doping the deviations occur at the onset of the Hubbard band as the quasiparticle peak is not clearly separated from the LHB. On the positive frequency side, the different curves deviate from each other also in the small doping limit. Comparing the two lowest dopings, for

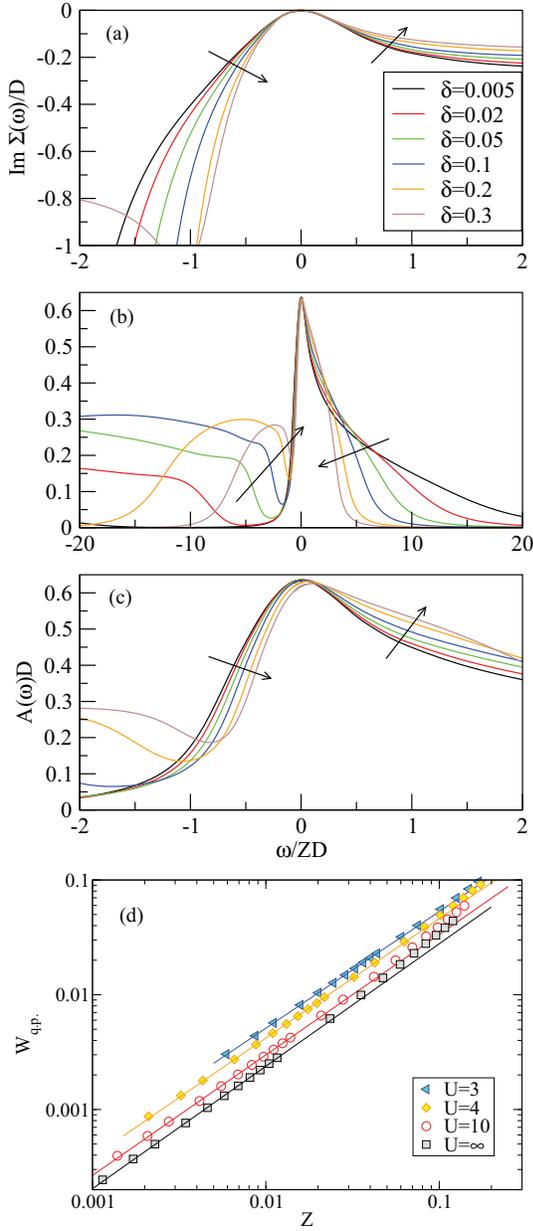


FIG. 5. (Color online) (a) Imaginary part of the self-energy  $\Sigma$  versus the rescaled frequency  $\omega/Z$  for  $U = \infty$ . (b), (c) Corresponding local spectral functions. Note that when plotted vs  $\omega/Z$ , the peak related to the onset of the LHB moves to the left with diminishing  $\delta$ , as seen in (b). The results for finite  $U$  are qualitatively very similar. The arrows indicate the direction of the increasing value of  $\delta$ . (d) Quasiparticle weight as a function of  $Z$ .

instance, reveals the excess of the spectral weight for the lower doping curve. This suggests that the quasiparticle peak is not fully characterized by the renormalization factor  $Z$  alone: the quasiparticle peak weight  $W_{\text{q.p.}}$  and  $Z$  are not necessarily simply proportional. This question is best addressed at very low doping, when the quasiparticle peak is well separated from the LHB. We can then extract  $W_{\text{q.p.}}$  by integrating the spectral function between the two local minima in  $A(\omega)$ . The results are plotted in Fig. 5(d). We find that at low  $\delta$ ,  $W_{\text{q.p.}}$

and  $Z$  are related by a power-law  $W_{\text{q.p.}} = Z^\gamma$ , with  $\gamma$  close to one, but not exactly 1. More specifically,  $\gamma$  is found to be  $U$  dependent:  $\gamma = 1.017$  for  $U = 3$ ,  $\gamma = 1.039$  for  $U = 4$ ,  $\gamma = 1.049$  for  $U = 10$ , and  $\gamma = 1.067$  for  $U = \infty$ .

#### 6. Charge compressibility: absence of phase separation

For some types of the (noninteracting) conduction-band density of states, there can be phase separation near half filling.<sup>42</sup> We verify that this is not the case for the Bethe lattice by plotting the band filling  $n$  as a function of the chemical potential  $\mu$  in Fig. 7(a) for  $U = \infty$ . The dependence is monotonous, thus all solutions are physically stable with positive charge compressibility  $\kappa = \partial n / \partial \mu$ . We also plot the quasiparticle residue  $Z$  as a function of the chemical potential  $\mu$  [Fig. 7(b)]. The charge compressibility  $\kappa$  as a function of the band filling [Fig. 7(c)] has a maximum near quarter filling. For smaller  $n$ , the decrease is due to the particular form of the non-interacting DOS (semicircular function). For larger  $n$ ,  $\kappa$  drops to zero as the Mott transition is approached. The asymptotic behavior is a power law  $\delta^\beta$  with  $\beta \approx 1/5$ .

### V. DOPED MOTT INSULATOR: AN ECFL PERSPECTIVE ON THE DMFT

In this section we make use of the general structure of the self-energy resulting from the ECFL in order to interpret, fit, and better understand the complex frequency dependence of the DMFT self-energies. The emphasis will be on the intermediate frequency range, which encompasses both the vicinity of  $\omega = 0$  and of the quasipole (sharp peak) in  $\text{Im}\Sigma$  on the negative frequency side at  $\omega_{\text{peak}} \simeq -0.7\delta D$ . We focus on intermediate doping levels, which turns out to be the range where the ECFL applies best, rather than on very low doping. For these reasons, we can use as a scaling variable:

$$x \equiv \omega/\delta \quad (38)$$

#### A. ECFL line shapes: main features

The low-frequency ECFL analysis from Sec. II gives a simple expression, Eq. (28), for  $\text{Im}\Sigma$  at  $T = 0$ . Using the overline convention of Eqs. (32), (33), and (34) to denote variables that remain finite as  $\delta \rightarrow 0$ , e.g.,  $\bar{P} = P/\delta$ , we rewrite Eq. (28) as

$$\text{Im}\Sigma = -\frac{x^2}{\bar{\Omega}_1} \frac{1 - x/\bar{\Delta}_1}{(1 + x/\bar{\Delta}_2)^2 + x^4/\bar{\Omega}_2^2}. \quad (39)$$

This ansatz function is determined by two variables with the meaning of curvature  $\bar{\Omega}_{1,2}$  that are simply related to  $\bar{\Omega}_\Sigma$  and  $\gamma_\psi$ , respectively, and by two parameters, which adjust the asymmetry  $\bar{\Delta}_{1,2}$  that are related to  $\bar{\Delta}$  and  $c_\psi$ .

The numerator of the expression describes a parabolic dependence, with a cubic correction term. This ansatz function has a peak (quasipole) at frequency  $x = -\bar{\Delta}_2$  for a finite  $\bar{\Omega}_2$ , turning into a true pole when  $\bar{\Omega}_2 \rightarrow \infty$ . The low-frequency asymmetry of the self-energy, important for the low-temperature thermoelectric properties, arises through a combination of the terms present in the numerator and denominator. Expanded to cubic order in frequency, the ansatz gives

$$\text{Im}\Sigma = -x^2/\bar{\Omega}_1 [1 - (1/\bar{\Delta}_1 + 2/\bar{\Delta}_2)x]. \quad (40)$$

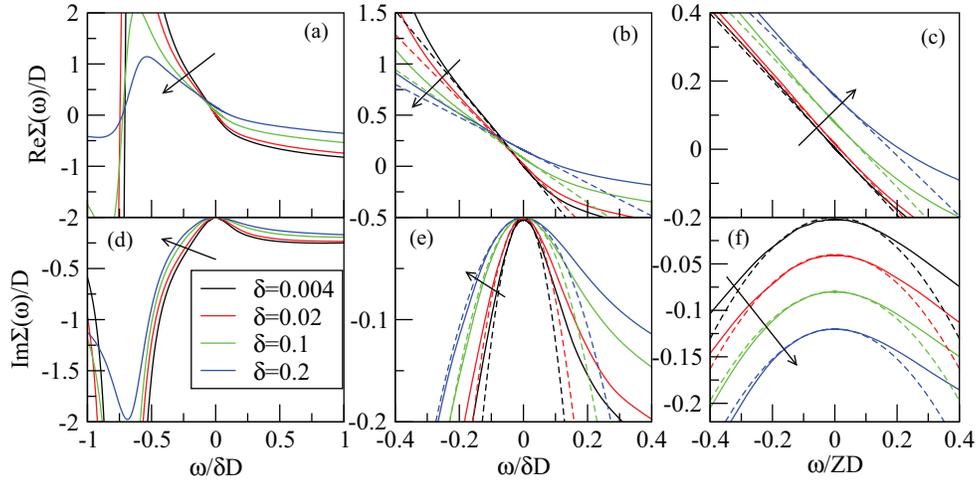


FIG. 6. (Color online) Real and imaginary parts of the DMFT self-energy  $\Sigma(\omega)$  for a range of doping for  $U = \infty$ , as a function of  $\omega/\delta D$  (left and center) or as a function of  $\omega/ZD$  (right). The arrows indicate the direction of the increasing value of  $\delta$ . The linear (for real parts) and parabolic (for imaginary parts) are performed in the frequency range  $[-0.05 : 0.05]\delta D$ . In (f), the data is vertically offset for clarity.

The ansatz function and its evolution as the parameters are varied is illustrated in Fig. 8.

Summarizing, the ansatz function Eq. (39) contains a parabolic dependence, multiplied by a function with a sharp peak at negative frequencies; therefore, it can be expected to describe the coarse structure found in the DMFT very well already at this order including *only* the Fermi liquid structure of the underlying functions  $\Psi$  and  $\chi$ .

### B. ECFL fits of the DMFT self-energy

The DMFT results for  $U = \infty$  self-energy for a range of doping levels are presented in Fig. 9 together with fits to Eq. (39). At large doping, the ansatz function describes the DMFT data remarkably well: the low-frequency dependence and the main shape of the self-energy are fully reproduced.

At smaller doping, the quasipole at negative frequencies becomes very sharp and  $\text{Im}\Sigma$  in DMFT develops a nipplelike structure at low-frequency, with semilinear frequency dependence at negative frequencies. These two features (quasipole and nipple) cannot be simultaneously well described by the simplest ECFL ansatz function in a broad frequency range. Given that the ansatz has a structure that already contains the pole, the fits in a broad frequency window are more meaningful. The DMFT data can still be described successfully, but terms beyond the lowest-order Fermi liquid form in  $\Psi$  and  $\chi$  need to be retained. Work along these lines

to reproduce the precise shape and to analyze its physical contents should be possible.

The evolution with doping of the fitting parameters is shown in Fig. 10. The first observation is that the fitting parameters (except for  $\bar{\Delta}_1$  to be discussed below) do not depend substantially on the doping, hence validating our assumptions stated above.

The second observation is that  $\bar{\Omega}_1$  is usually found to be very close to  $\bar{\Omega}_2$ . This supports the conclusions of the  $\lambda^2$  analysis (to be discussed in the next section), which also finds that  $\gamma_\Psi$  is close to  $\Omega_\chi$ .

The third observation is that the bulk of the asymmetry does actually not come from the explicitly cubic term (parametrized by  $1/\Delta_1$ ), but rather from  $1/\Delta_2$ . Hence, it is the presence of the quasipole at negative frequency that is responsible for the strong particle-hole asymmetry. This is seen most explicitly in the doping range  $\delta = 0.2-0.3$ , where  $1/\Delta_1$  almost vanishes and where, furthermore, the data is excellently described by the ansatz function from Eq. (39). The physical content of this observation might be that the low-frequency particle-hole asymmetry is (at least at not too low dopings) directly related to the presence of the LHB at much higher frequency scales, and thus ultimately to the strong-correlation physics. This observation is consistent with the picture emerging from the ECFL, where the asymmetry is a consequence of the Gutzwiller projection.

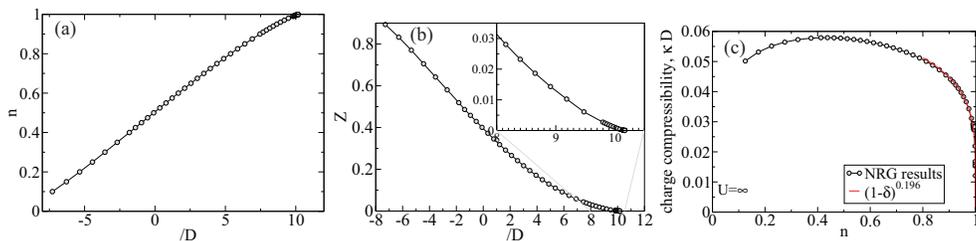


FIG. 7. (Color online) (a) Band filling  $n$  vs chemical potential  $\mu$  for  $U = \infty$ . (b) Quasiparticle residue  $Z$  vs. chemical potential  $\mu$  for  $U = \infty$ . (c) Charge compressibility for  $U = \infty$ .

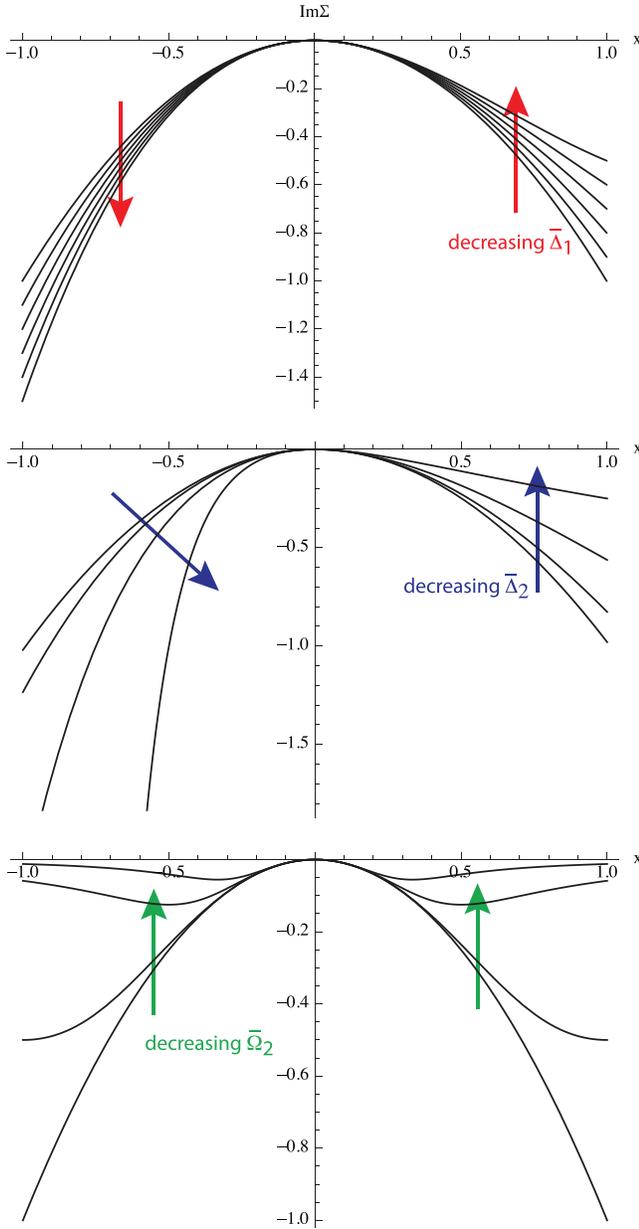


FIG. 8. (Color online) Role of the parameters in the ECFL ansatz function, Eq. (39), used for fitting in Figs. 9 and 10.

### C. Summary

To summarize, the ECFL-derived ansatz, Eq. (39), which retains only the lowest-order Fermi liquid terms in  $\Psi$  and  $\chi$ , describes the rather complex frequency dependence of the DMFT data remarkably well at low to intermediate frequencies and for not too small doping levels. Importantly, retaining only the Fermi liquid (hence particle-hole symmetric) terms in these ECFL self-energies already yields a marked particle-hole asymmetry in the physical electron Dysonian self-energy. The ansatz also describes the pole at negative frequencies, associated with the onset of the LHB. Whereas the Fermi liquid behavior only applies at extremely low frequencies in the Dysonian self-energy, the Fermi liquid concepts can still be used over a much broader frequency range when

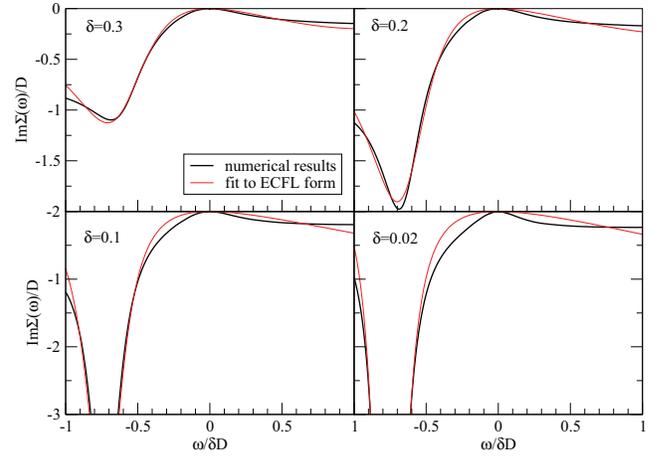


FIG. 9. (Color online)  $\text{Im}\Sigma(\omega)$  vs rescaled frequency  $\omega/\delta D$  and fits to Eq. (39) for  $U = \infty$ . The ansatz describes the DMFT self-energy at moderate dopings remarkably well. The nipple structure that becomes pronounced at small doping signals that  $\chi$  and  $\Psi$  develop non-Fermi liquid corrections.

proper auxiliary quantities are considered, within the broader framework provided, e.g., by the ECFL theory. At lower doping levels, however, the DMFT results display structures (nipple), which signal the increasing importance of corrections beyond the dominant Fermi liquid terms in  $\Psi, \chi$ .

## VI. ECFL: EXPANSION TO $O(\lambda^2)$

### A. Summary of equations

We now summarize the results of the  $O(\lambda^2)$  expansion of the ECFL equations in Ref. 8, which are then computed and compared with the DMFT results. We note that the ECFL reformulation of the Dyson self-energy  $\Sigma(\omega)$  into the ECFL auxiliary self-energies  $\Psi(\omega)$  and  $\chi(\omega)$  is exact. Therefore, if

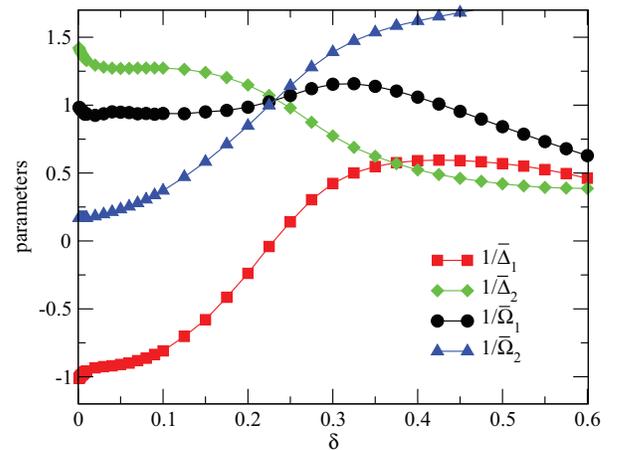


FIG. 10. (Color online) Parameters in the fit function, Eq. (39), for a range of doping  $\delta$  at  $U = \infty$ . For a broad fitting energy range, the fitting parameters are smooth as a function of  $\delta$ , thus the fitting procedure is well defined. At dopings where the fits work best (that is around  $\delta = 0.25$ ),  $\bar{\Omega}_1$  is found to be close to  $\bar{\Omega}_2$ .

one could perform the  $\lambda$  expansion to infinite order in  $\lambda$ , one would obtain the exact answer for these auxiliary self-energies, and consequently the Dyson self-energy. The resulting Dyson self-energy would then agree exactly with the one obtained through DMFT for the case of infinite  $U$ . Our aim here is to benchmark the lowest nontrivial order of the  $\lambda$  expansion against the exact DMFT results. Note that in the  $d \rightarrow \infty$  limit, in the paramagnetic phase, the single-particle properties of the  $t$ - $J$  model are identical to those of the  $U = \infty$  Hubbard model. In other words, as long as antiferromagnetic correlations are short ranged,  $J$  does not enter single-particle properties in the  $d = \infty$  limit. Accordingly, the only coupling constant entering the simplified ECFL equations is the hopping (band dispersion) itself, and not the superexchange.

In the  $O(\lambda^2)$  scheme, the explicit density factors  $1 - \frac{n}{2}$  that occur in Eq. (8) and (9) are replaced by the rule

$$1 - \frac{n}{2} \rightarrow a_G \equiv 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4} + O(\lambda^3). \quad (41)$$

The second rule is that the explicit self-energy expressions in these equations are multiplied by  $\lambda$ . As an illustration of these rules, we write Eq. (8) and (9) as

$$\mathbf{g}^{-1}(k) = i\omega + \mu - a_G \varepsilon_k - \lambda \Phi(k), \quad (42)$$

$$\tilde{\mu}(k) = a_G + \lambda \Psi(k). \quad (43)$$

The factor  $\lambda$  is set to 1 before actually computing with these formulas. The two self-energy functions  $\Phi(k)$  and  $\Psi(k)$  satisfy the equations to second order in  $\lambda$ :

$$\begin{aligned} \Psi(i\omega_k) &= -\lambda \sum_{pq} (\varepsilon_p + \varepsilon_q - u_0) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k), \\ \Phi(k) &= \left( \varepsilon_k - \frac{u_0}{2} \right) \Psi(i\omega_k) + \chi(i\omega_k) - u_0 \left( \lambda \frac{n^2}{8} - \frac{n}{2} \right) \\ &\quad - \sum_p \varepsilon_p \mathbf{g}(p) \\ \chi(i\omega_k) &= -\lambda \sum_{pq} \left( \varepsilon_{p+q-k} - \frac{u_0}{2} \right) (\varepsilon_p + \varepsilon_q - u_0) \\ &\quad \times \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k). \end{aligned} \quad (44)$$

All equations in Eq. (44) are implicitly understood to have  $O(\lambda^3)$  corrections, so that the  $\mathbf{g}$  and  $\tilde{\mu}$  pieces of  $\mathcal{G}$  in Eq. (7) are correct to the stated order. As expected, the functions  $\Psi, \chi$  depend on the frequency but not the momentum  $\vec{k}$ . Both Green's functions satisfy an identical number sum rule  $\sum_{k, \omega_n} \mathcal{G}(k) = \frac{n}{2} = \sum_{k, \omega_n} \mathbf{g}(k)$ , and the theory has two chemical potentials necessary to impose these, namely  $\mu$  and  $u_0$ . As discussed in Ref. 3, the second chemical potential  $u_0$  arises from the requirement of satisfying a shift invariance in the theory. The shift transformation in the present model acts as  $\varepsilon_p \rightarrow \varepsilon_p + c$ . This transformation shifts the center of gravity of the band; it is absorbable in  $u_0$ , and thus rendered inconsequential. We can easily verify that  $\chi$  and  $\Psi$  are independently shift invariant. Combining the expressions,

we write

$$\mu' \equiv \mu + u_0 \left( \lambda^2 \frac{n^2}{8} - \lambda \frac{n}{2} \right) - \frac{u_0}{2} a_G + \lambda \sum_p \varepsilon_p \mathbf{g}(p), \quad (45)$$

$$\mathbf{g}^{-1}(k) = i\omega + \mu' - \left( \varepsilon_k - \frac{u_0}{2} \right) \{ a_G + \lambda \Psi(i\omega_k) \} - \lambda \chi(i\omega_k).$$

The Green's function is then found by combining Eqs. (45), (44), and (43) in the expression Eq. (7).

## B. Setting up the computation

To set up the computation, we write a local Green's function with weight  $m = 0, 1, \dots$  using Eq. (45) as

$$\begin{aligned} \mathbf{g}_{\text{loc}, m}(i\omega_k) &\equiv \sum_{\vec{k}} \mathbf{g}(k) (\varepsilon_{\vec{k}})^m \\ &= \int_{-D}^D d\varepsilon \rho_0(\varepsilon) \frac{\varepsilon^m}{i\omega + \mu' - (a_G + \Psi)(\varepsilon - \frac{u_0}{2}) - \chi}, \end{aligned} \quad (46)$$

where  $\chi$  and  $\Psi$  are functions of frequency  $i\omega_k$  but not the energy  $\varepsilon$ . We find that both  $\mathbf{g}_{\text{loc}, 0}$  and  $\mathbf{g}_{\text{loc}, 1}$  are needed to compute the frequency-dependent self-energy. Similarly, a local  $\mathcal{G}$  can be defined, and the number sum rules can be written as  $\frac{n}{2} = \sum_{i\omega} \mathbf{g}_{\text{loc}, 0}(i\omega)$  and  $\frac{n}{2} = \sum_{i\omega} \mathcal{G}_{\text{loc}, 0}(i\omega)$ . Where necessary, the usual convergence factor  $e^{i\omega 0^+}$  is inserted. The two  $\vec{k}$ -independent functions  $\Psi$  and  $\chi$  in Eq. (44) can be written in a compact way if we first define a function with three indices ( $m_1 m_2 m_3$ ) from the weight factors:

$$\begin{aligned} I_{m_1 m_2 m_3}(i\omega) &= -\frac{1}{\beta^2} \sum_{v_1, v_2} \mathbf{g}_{\text{loc}, m_1}(i v_1) \\ &\quad \times \mathbf{g}_{\text{loc}, m_2}(i v_2) \mathbf{g}_{\text{loc}, m_3}(i v_1 + i v_2 - i\omega). \end{aligned} \quad (47)$$

After continuation  $i\omega \rightarrow \omega + i\eta$ , and for all values of the indices, the low frequency and temperature  $I(i\omega)$  is a Fermi liquid-like self-energy with an imaginary part  $\propto [\omega^2 + (\pi k_B T)^2]$ . We can now rewrite Eq. (44) as:

$$\begin{aligned} \Psi(i\omega) &= -u_0 I_{000}(i\omega) + 2I_{010}(i\omega) \\ \chi(i\omega) &= -\frac{u_0}{2} \Psi(i\omega) - u_0 I_{001} + 2I_{011}(i\omega). \end{aligned} \quad (48)$$

Clearly, Eqs. (48) and (46) along with the definition (47) and the number sum rules form a self-consistent set of equations that can be solved iteratively on a computer. The Dyson self-energy and the spectral function can be computed in terms of these quantities using Eq. (13)

## C. Auxiliary and Dyson self-energies to $O(\lambda^2)$

In Fig. 11 we present  $\rho_\chi$ ,  $\rho_\Psi$ , and  $\rho_\Sigma$  (from top to bottom).  $\rho_\Psi$  and  $\rho_\chi$  have similar frequency dependence and a Fermi liquid form is obeyed much more accurately than what is found for the Dyson self-energy  $\rho_\Sigma$ . This supports the ansatz that we employed above. In particular, the auxiliary self-energies are more particle-hole symmetric; most of the particle-hole asymmetry follows from the structure of ECFL equations. This signals that the Fermi liquid concept has validity outside of the canonical Fermi liquid behavior.

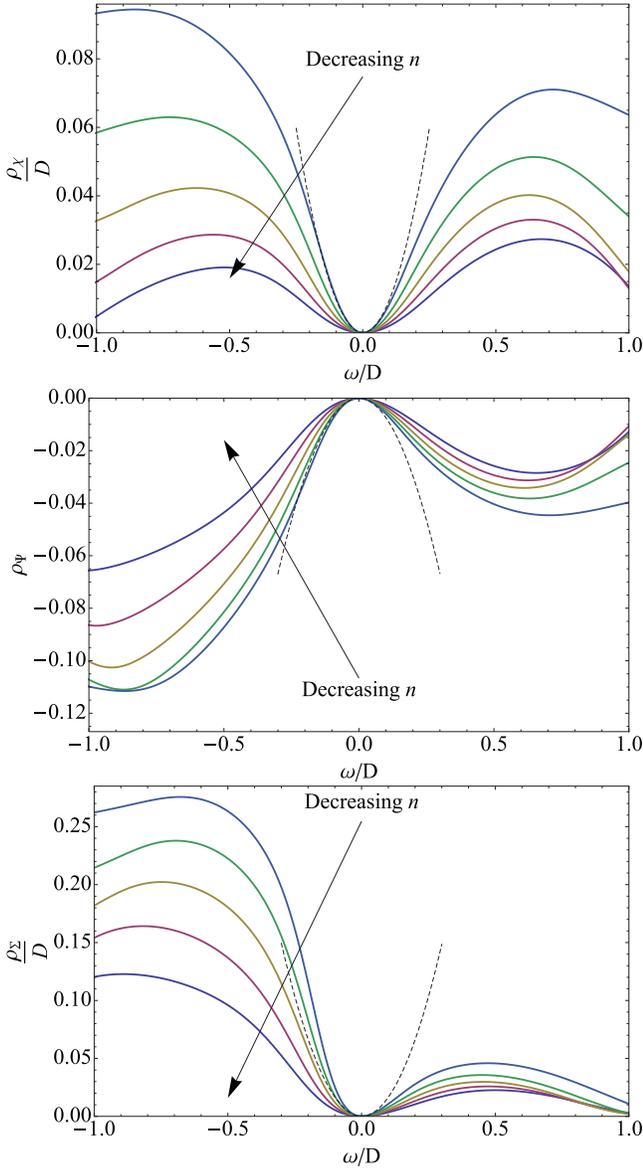


FIG. 11. (Color online) Imaginary parts (spectral densities) of the auxiliary self-energies  $\chi$  and  $\Psi$ , and the Dyson self-energy  $\Sigma$  within the  $O(\lambda^2)$  ECFL. The dotted lines are parabolic fits at the highest density. Recall that typical Fermi liquid-type spectral functions exhibit a parabolic and therefore particle-hole symmetric behavior over a large energy range. From these fits one observes that the auxiliary functions  $\Psi$  and  $\chi$  have a Fermi liquid form over a wider energy range than the Dyson self-energy  $\Sigma$ .

## VII. DETAILED COMPARISON OF $O(\lambda^2)$ ECFL RESULTS TO DMFT

### A. Effective density of the ECFL spectral functions and its phenomenological adjustment

The  $O(\lambda^2)$  equations of ECFL discussed here give a high- $\omega$  limiting behavior  $\mathcal{G} \sim \frac{a_g}{\omega}$ , differing from the exact form  $\mathcal{G} \sim \frac{1-n}{\omega}$  due to the replacement  $1 - \frac{n}{2} \rightarrow a_g \equiv 1 - \lambda n/2 + \lambda^2 n^2/4$  as per the rules of the calculation. This effect is due

to the incomplete projection of the  $O(\lambda^2)$  treatment of the ECFL equation of motion. At  $n \sim 0.75$  the error in the high-frequency weight is 22.5%.

A phenomenological scheme for adjusting for this feature defines an effective density  $n_{\text{eff}}$ , using the ratio of particle addition and removal states as the relevant metric, so that  $\frac{n}{1-n+n^2/4} = \frac{n_{\text{eff}}}{1-n_{\text{eff}}}$ , thus yielding

$$n_{\text{eff}} = \frac{n}{1 + \frac{n^2}{4}}. \quad (49)$$

Clearly higher-order calculations would have a corresponding mapping between the two densities. For several of the comparisons below, agreement is greatly improved by plotting the results of ECFL as a function of  $n_{\text{eff}}$ .

### B. Comparison between $O(\lambda^2)$ -ECFL and DMFT

We find that the computed values of the quasiparticle weight  $Z$  from ECFL are close to the  $U/D = 4$  DMFT curve, we detail this in Appendix A, where the momentum distribution is also shown. This is suggestive of an analogy between the two incompletely projected theories. In particular, making  $U$  finite, and truncating the  $\lambda$  expansion at second order, both introduce some double occupancy into the system. It is therefore not surprising that the  $U/D = 4$  DMFT results agree better with the  $O(\lambda^2)$ -ECFL than with the  $U/D = \infty$  DMFT results. However the limitations of the  $O(\lambda^2)$  calculation within ECFL preclude obtaining reliable results for doping levels smaller than  $\delta \approx 0.25$ .

#### 1. Spectral line shapes

In Fig. 12 we compare the ECFL and the DMFT results at  $U = 4D$  and  $U = \infty$  for the  $\epsilon$ -resolved spectral functions at two values of the band energy,  $\epsilon_k = -D$  and  $\epsilon_k = \epsilon_F$ . In general, the agreement is encouraging. At  $\epsilon_k = -D$  (left panel), DMFT has a deeper minimum between the QP and the secondary feature at high binding energy than is seen in ECFL, but the position of the ECFL peaks agrees well with that of the DMFT peaks. At  $\epsilon_F$  (right panel) the QP are of similar width but have different values of  $Z$ , as discussed above. The background of width  $\sim D$  lies over essentially the same frequency range for all three calculations, and has a peak at  $\omega = -0.5D$ , approximately the same position for each data set. However, the height of the peak is less pronounced for the ECFL than the DMFT. At positive frequencies the spectral functions are in excellent agreement. Plotting the spectral function as a function of the scaled frequency  $\omega/ZD$  improves the agreement in the position and width of the quasiparticle, as illustrated in the more sensitive self-energy curves in Fig. 13. We note that the scaled ECFL curves agree well with the DMFT curves even for density  $n \sim 0.8-0.9$  for scaled frequency  $|\omega| \leq 0.5DZ$ . We find this agreement surprising in view of our criterion discussed above, placing  $n \sim 0.75$  as the limiting density.

The physical spectral function  $A(\omega)$ , when displayed as a color intensity plot using the scaled frequency  $\omega/ZD$  as in Fig. 14, further emphasizes this similarity. At this level of description, the  $U = 4D$  DMFT curve and the

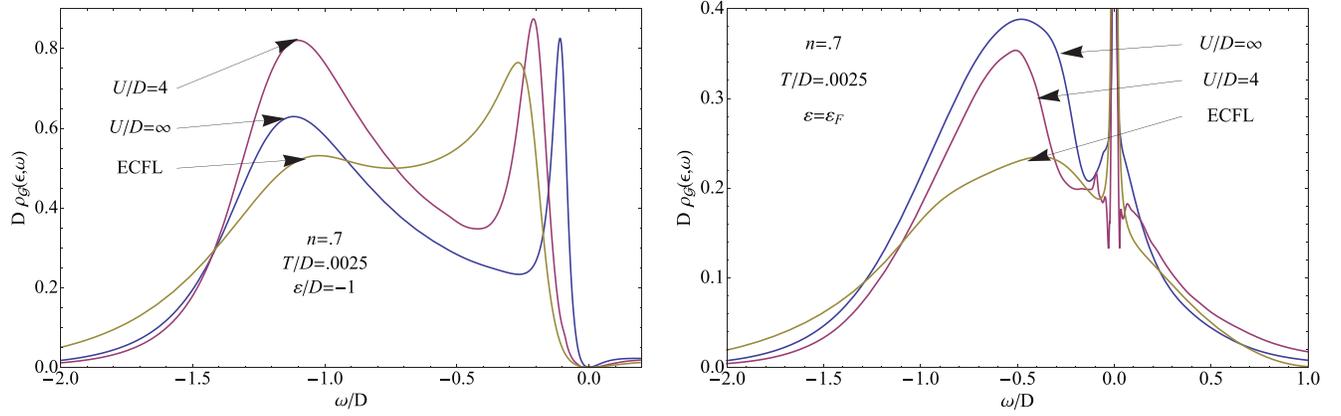


FIG. 12. (Color online) Spectral functions within the DMFT ( $U = \infty$  and  $U/D = 4$ ) and ECFL at two typical energies  $\varepsilon = -D$  and  $\varepsilon = \varepsilon_F$ , with  $n = 0.7$  and  $T = 0.0025D$ . The location of the quasiparticle peak near  $\omega \sim 0$  and the broad secondary peak for  $\omega < 0$  are common to both calculations. While there are subtle differences, especially in the magnitudes of the secondary peaks, the main features of the three calculations match at high and low frequency.

$O(\lambda^2)$  calculation look almost identical. In particular, as clear from this figure, both theories indicate that the quasiparticle peak becomes rapidly *more dispersive* as one moves to positive energies, corresponding to unoccupied states (i.e.,

the effective Fermi velocity increases as compared to its low-energy value and becomes closer to the band value). As discussed above (Sec. IV, Fig. 4), this is one of the primary common conclusions of both theories, which could be tested

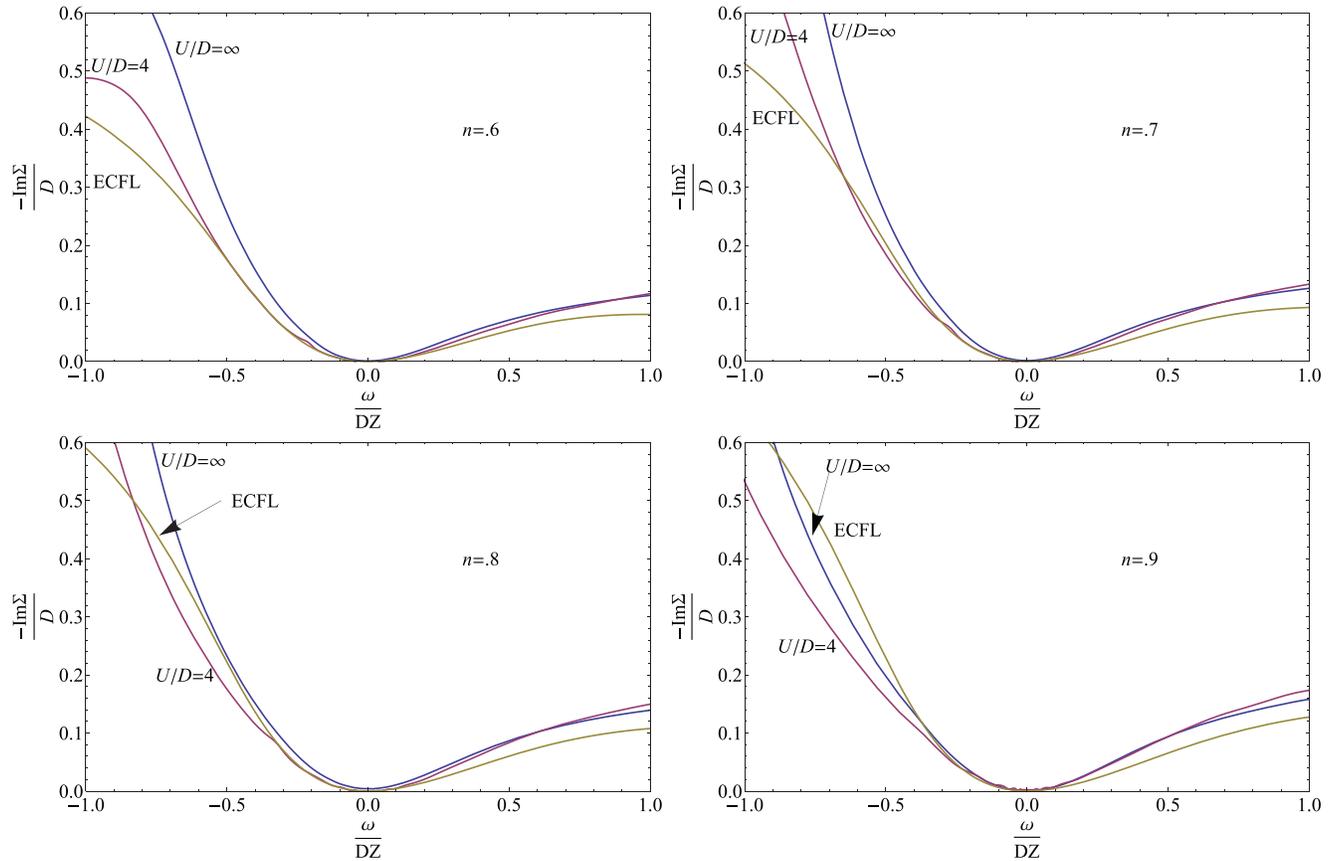


FIG. 13. (Color online) Spectral function (imaginary part) of the Dyson self-energy  $\Sigma$  versus the scaled variable  $\omega/(DZ)$  in the ECFL theory at order  $\lambda^2$ , and the DMFT at two values of  $U$ . The ECFL predicts a value of  $Z$  which is too large at low doping, and significant  $U$  dependence creates differences between the  $U/D = 4$  and  $U = \infty$  results of the DMFT. Nonetheless, all three cases overlap well at low frequencies when plotted against the scaled frequency. Surprisingly, this agreement survives to densities far beyond the expected range of the current version of the ECFL.

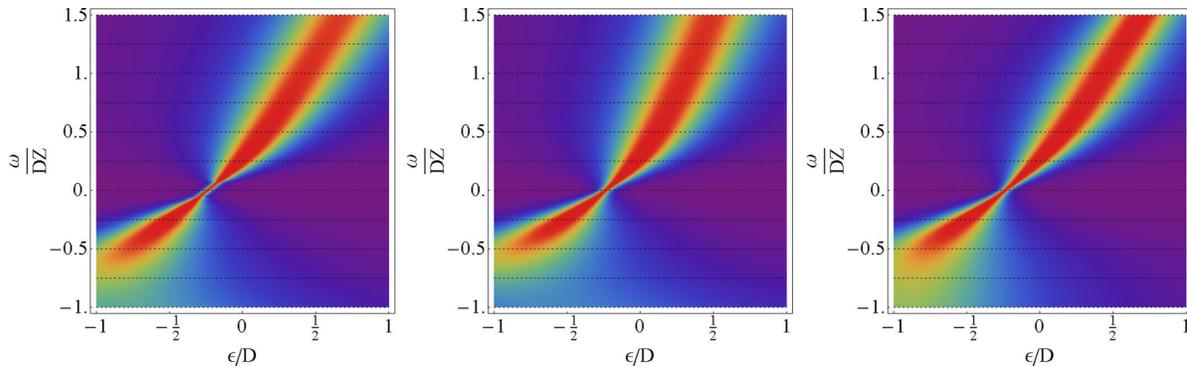


FIG. 14. (Color online) Physical spectral function  $A(\epsilon, \omega)$ . From left to right:  $U = 4$  DMFT,  $U = \infty$  DMFT, and ECFL with  $n = 0.7$  and  $T/D = 0.0025$ . Hot colors represent high intensity, while darker blue represents low intensity. Noting from the left panel of Fig. 12 that the QP band has a slightly different width in each calculation, we plot the spectral function here as a function of  $\frac{\omega}{DZ}$ . This brings the low-energy (QP) features of the spectral function into impressive agreement, indicating that  $Z$ , rather than  $\delta$ , is the fundamental energy scale of the extremely correlated state.

in future experiments able to probe the unoccupied states in a momentum-resolved manner.

In view of the remarkable similarity between the different theories, as seen in Fig. 13 and Fig. 14, it appears that the  $O(\lambda^2)$  version of ECFL has the correct *shape* of the spectra built into it, but requires a correction for a too large value of the QP factor  $Z$ . This is the main conclusion of this work regarding the benchmarking of the ECFL.

### VIII. CONCLUSION AND PROSPECTS

In this work we have presented a detailed comparison between the DMFT and the ECFL theories, applied to the doped Hubbard model at large as well as infinite  $U$ , in the limit of infinite dimensions (Bethe lattice with infinite coordination).

Our approach here is twofold. On the one hand, we have used the general structure of the Green's function and self-energy in the ECFL theory to obtain a useful analytical ansatz, which reproduces quite well the rich and complex frequency dependence of the DMFT self-energy at not too low doping level. This ansatz relies on the lowest-order Fermi liquid expansion of the two auxiliary ECFL self-energies  $\Psi$  and  $\chi$ . Quite remarkably, the marked deviations from the Fermi liquid form and the particle-hole asymmetry found in the physical single-particle self-energy can be accounted for by this underlying Fermi liquid form of auxiliary quantities. In turn, the deviations observed between the DMFT results and this lowest-order ansatz at lower doping levels emphasize the need for corrections to FL behavior in  $\Psi, \chi$  within the ECFL. This part of our study thus provides useful analytical insights into the DMFT description of the doping-driven Mott transition.

On the other hand, we have used the DMFT results (obtained here with a high-accuracy NRG solver) as a benchmark of the ECFL theory. Specifically, we have solved numerically the  $O(\lambda^2)$  ECFL equations, appropriately simplified in the limit of large dimensions. For not too low doping levels, where this  $O(\lambda^2)$  scheme is applicable, we found that the spectral properties agree well provided the comparison is made as a function of the scaled frequency  $\omega/ZD$ , with  $Z$  the

quasiparticle weight. A similar situation arises in comparing the ECFL method for the Anderson impurity model, where  $Z$  is rapidly suppressed as the Kondo limit is approached.<sup>43</sup> This adjustment of the frequency scale compensates the known weakness of the  $O(\lambda^2)$  theory in obtaining  $Z$  quantitatively, and enables, to some extent, a preview of the results of the planned higher-order calculations in the ECFL projection parameter  $\lambda$ .

From a physics point of view, we now summarize the most significant insights provided by our study.

Doped Mott insulators are found to be characterized by a marked particle-hole *dynamical asymmetry*, as emphasized in recent ECFL<sup>44</sup> and DMFT<sup>9</sup> studies. In the case of hole doping, particlelike ( $\omega > 0$ ) excitations are longer lived than holelike ( $\omega < 0$ ) ones, leading to more resilient electronlike quasiparticles.<sup>9</sup> This dynamical asymmetry has physical implications for the spectral line shapes<sup>9,44</sup> as well as thermopower.<sup>9,45</sup> The asymmetric terms in the low-frequency expansion of the self-energy signal deviations from the Fermi liquid theory, which are usually ignored in weak-coupling studies. They become large at low hole doping and strong coupling, as demonstrated here in considerable detail, thus confirming the proposal made originally in Ref. 44.

Due to the importance of this asymmetry, we found that the energy vs momentum dispersion of the quasiparticle state quickly deviates on the  $\omega > 0$  side from its low-energy value (associated with the renormalized effective Fermi velocity). The deviation is towards a weaker dispersion, closer to the bare band value. This is a prediction of both ECFL and DMFT, which could be tested experimentally once momentum-resolved spectroscopies are developed in order to address unoccupied states (the dark side for photoemission).

Regarding ARPES line shapes, we also emphasize that the recent successful comparison<sup>6,7</sup> between the ECFL and the experimental ARPES line shapes in the optimally doped and overdoped cuprates along the nodal direction can just as well be interpreted as the similar success of the DMFT interpretation of these line shapes. The adjustment of the momentum dependence of the comparison factor for different systems in Refs. 6 and 7 hints at the importance of the momentum dependence of the self-energy. This momentum

dependence is already present in the ECFL in two dimensions, and also emerges from cluster DMFT calculations.

Further comparison between the nature of the momentum dependence in both theories is to be addressed in future work. More generally, we believe that this work lays the foundation of a useful program where the momentum-dependent self-energies can be reliably computed and expressed in simple analytic forms. While cluster DMFT methods can already provide some answers to this important problem, the ECFL theory readily treats low dimensions and the momentum dependence. In order to get further solid results, the current limitation of the ECFL to the somewhat overdoped regime needs to be overcome. This limitation arises from the low order of the expansion in  $\lambda$ , and brute-force higher-order calculations in  $\lambda$  are planned. In this task, the insights gained from the present comparison with DMFT, are invaluable.

### ACKNOWLEDGMENTS

A.G., J.M., and B.S.S. acknowledge useful discussions with Xiaoyu Deng at the initial stage of this project. The work at UCSC was supported by DOE-BES under Grant No. DE-FG02-06ER46319. A.G. and J.M. acknowledge the support of Ecole Polytechnique and Collège de France. R.Ž. and J.M. acknowledge the support of the Slovenian Research Agency (ARRS) under Program P1-0044.

### APPENDIX A: QUASIPARTICLE OCCUPATION AND $Z_k$ IN THE $O(\lambda^2)$ ECFL

The momentum distribution function, Fig. 15, shows good agreement with the DMFT at a density  $n = .7$ , and the large spillover for  $k > k_F$  is of the same scale in both sets of calculation. Its importance in estimating the background spectrum in ARPES is well known, so this is already a reasonably reliable common result. It is also interesting that at the Fermi momentum, the magnitude of the distribution function is close to  $\frac{1}{2}$  in both calculations, as argued in the literature.<sup>15,46</sup>

In Fig. 16 we compare the quasiparticle weight  $Z$  in ECFL and DMFT. The  $O(\lambda^2)$  ECFL result has some similarity to the Gutzwiller approximation<sup>47</sup> result  $2\delta/(1 + \delta)$  in the limited density range of validity. However it does not seem to vanish in any obvious way, if we extrapolate by eye to higher density  $n$ , highlighting its main weakness in the current state of development, but plotted against the effective density it becomes comparable to the  $U/D = 4$  DMFT curve over a limited range.

### APPENDIX B: NRG IMPURITY SOLVER CONVERGENCE AT SMALL DOPING

In order to obtain well-converged spectral functions using the NRG impurity solver at low doping  $\delta$ , several parameters in the method need to be appropriately tuned. Their choice affects both low-frequency and high-frequency parts of the spectral functions. In addition, it significantly affects the numerical requirements—both the duration of each NRG calculation and the number of the DMFT cycles until self-consistency. Very close to the Mott transition, obtaining fully converged results

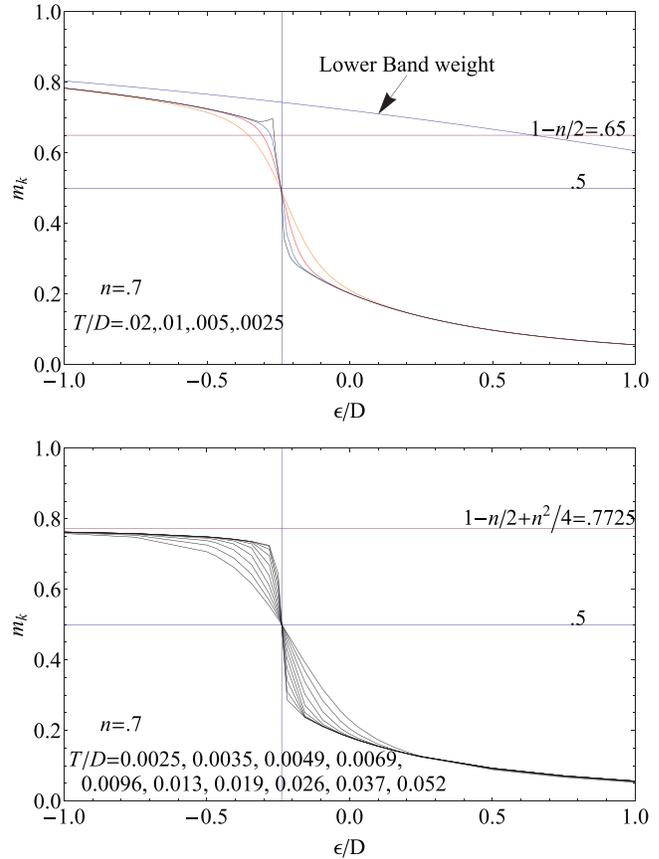


FIG. 15. (Color online) Momentum occupation versus  $\varepsilon$  within the DMFT at  $U/D = 4$  (top) and the ECFL (bottom).

becomes computationally very expensive (several hundreds of DMFT cycles) even with Broyden acceleration.<sup>22</sup> In this section, we explore the effects of different choices on the quasiparticle residue  $Z$  (low-frequency property), Fig. 17, and on the shape of the LHB (high-frequency property), Fig. 18.

We first explore the choice of the discretization scheme (i.e., how the coefficients of the Wilson chain are computed based from the input hybridization function). We compare the discretization scheme (denoted as Z) proposed by R. Žitko and Th. Pruschke in Ref. 22, which corrects the systematic discretization errors near band edges present both in the conventional discretization scheme (Y), Ref. 48, and in the improved scheme by V. Campo and L. Oliveira (C), Ref. 49. At low frequencies, one observes excellent overlap of the results, as indicated in Fig. 17(a). This is in line with the common wisdom that the NRG is a reliable method for low-frequency properties, having good spectral resolution in the vicinity of the Fermi level where the discretization grid is condensed. For this reason, the choice of the discretization scheme has little effect on  $Z$ . At high frequencies, however, one can clearly observe the systematic artifacts present in schemes Y and C: the LHB presents spurious (nonphysical) structure at the outer edge, which is not present in the results of scheme Z, see Fig. 18(a). Recent comparisons of the NRG (using scheme Z) and continuous-time quantum Monte Carlo at finite temperatures have established that the NRG

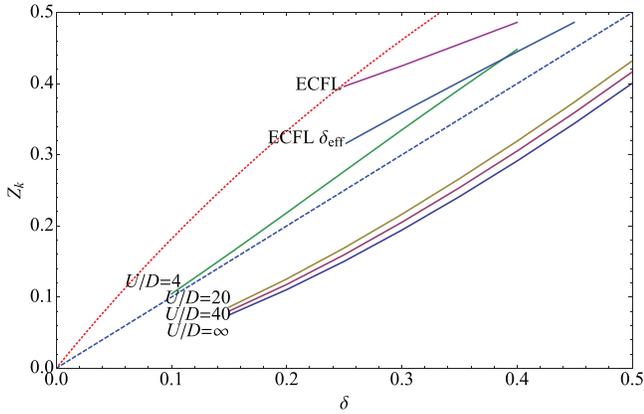


FIG. 16. (Color online) Quasiparticle weight  $Z$  as a function of hole doping  $\delta$  (or  $\delta_{\text{eff}} = 1 - n_{\text{eff}}$ , the effective hole doping) from the  $O(\lambda^2)$  version of ECFL, and from DMFT for various values of  $U$ . More detailed DMFT results are in Fig. 1. The blue dashed line represents  $Z = \delta$ , the simplest  $U = \infty$  slave-boson estimate, as a guide to the eye. The dotted red line represents the Gutzwiller approximation result  $Z = \frac{2\delta}{1+\delta}$ .

(using scheme Z) is, in fact, a rather reliable method also for high-frequency/finite-temperature properties. On the other hand, the NRG using schemes Y or C is expected to exhibit more pronounced systematic errors at high frequencies and at finite temperatures.

The second important choice concerns the value of the discretization parameter  $\Lambda$ , which controls the coarseness of the logarithmic grid. The standard choice is  $\Lambda = 2$ , which is suitable to obtain well converged results at both low and high frequencies, see Figs. 17(b) and 18(b). In fact, the results

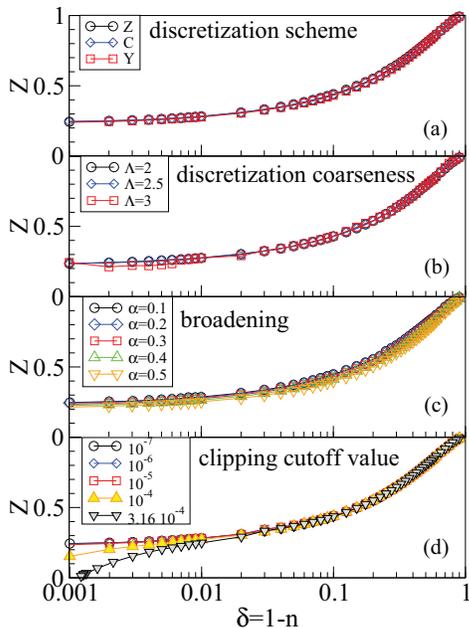


FIG. 17. (Color online) Quasiparticle residue  $Z$  for the  $U = \infty$  Hubbard model as a function of doping  $\delta$ . We compare the DMFT results for different choices of the NRG impurity solver parameters.

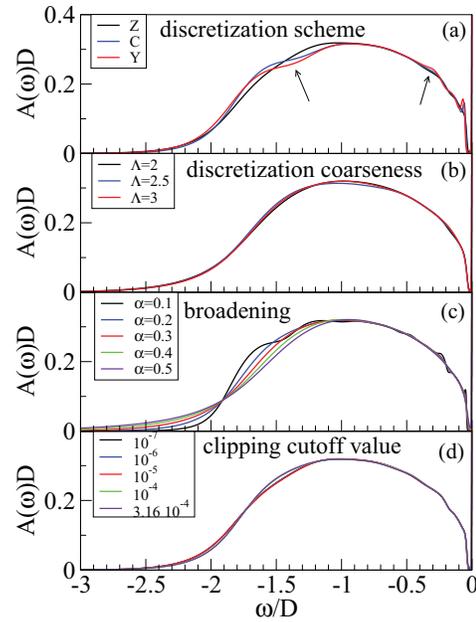


FIG. 18. (Color online) Lower Hubbard band part of the local spectral function  $A(\omega)$  for the  $U = \infty$  Hubbard model at low doping,  $\delta = 0.01$ . In (a), the schemes C and Y produce spurious features at the outer band edge and at  $\omega = 0.3D$  (indicated by arrows). In addition, the (expected) feature at the inner band edge at  $\omega = 0.1D$  is overemphasized in C and Y schemes. In (d), we note that the curves for clipping values  $10^{-7}$ ,  $10^{-6}$ , and  $10^{-5}$  nearly overlap, while those for  $\geq 10^{-4}$  exhibit some deviations.

do not change much even when going to somewhat higher  $\Lambda = 2.5$ , while for  $\Lambda = 3$  we start to observe some systematic deviations at very low doping  $\delta$ . We have also performed some test calculations for smaller values  $\Lambda = 1.9, 1.8, 1.7$ ; the results differ little from those for  $\Lambda = 2$  while being significantly more computationally expensive to produce.

We now consider the broadening parameter  $\alpha$ , which controls how the raw spectral function in the form of a set of weighted  $\delta$  peaks is processed to obtain a smooth continuous representation. Too small values lead to spurious oscillations, too high values to overbroadening. These effects are nicely illustrated by the results for the LHB part of the spectral function in Fig. 18(c). The long high-frequency tail of the LHB for increasing  $\alpha$  is a clear overbroadening effect, while the oscillatory features for  $\alpha = 0.1$  are a discretization artifact. At low frequencies, the QP residue  $Z$  converges as  $\alpha$  is decreased, see Fig. 17(c). We find that for  $\alpha \leq 0.1$ , the results practically overlap, while for  $\alpha = 0.2$  (the value used for most calculations in this work), the deviation from the asymptotic value is of order one percent. For large values of broadening (as commonly done in NRG calculations),  $Z$  is underestimated. This is because the spectral weight is more spread around as  $\alpha$  increases, thus less weight remains in the QP peak. Based on these results, we find that  $\alpha = 0.2$  is a good compromise.

Finally, we discuss a subtle issue, which becomes important at very small dopings. The NRG discretization has difficulties if in the hybridization function there are extended regions of very low values. In particular, this leads to very slow approach

to the self-consistency. For this reason, it is convenient to use a small, but finite cutoff value for the hybridization function to clip the input hybridization function to some minimum value at all frequencies. It is important, however, to choose this value so that the results are not perturbed. We find that using too high cutoff leads to incorrect  $Z$  vs  $\delta$  behavior at low doping (a downturn), see Fig. 17(d). The effect is thus similar

to overbroadening, since the spectral weight shifts from the quasiparticle peak to the region between the LHB and the QP peak, where the clipping is applied (for small  $\delta$ , where the LHB and the QP peak no longer overlap, but rather the QP becomes an isolated spectral peak in the gap). There is also some effect of clipping on the LHB itself, Fig. 18(d). Again, this effect is analogous to overbroadening.

- <sup>1</sup>A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- <sup>2</sup>B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).
- <sup>3</sup>B. S. Shastry, *Phys. Rev. B* **87**, 125124 (2013).
- <sup>4</sup>B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011).
- <sup>5</sup>E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, and B. S. Shastry, *Phys. Rev. B* **87**, 161120 (2013).
- <sup>6</sup>G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).
- <sup>7</sup>Kazue Matsuyama and G.-H. Gweon, *Phys. Rev. Lett.* **111**, 246401 (2013).
- <sup>8</sup>E. Perepelitsky and B. S. Shastry, *Ann. Phys. (N.Y.)* **338**, 283 (2013).
- <sup>9</sup>X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- <sup>10</sup>B. S. Shastry, *Phys. Rev. Lett.* **108**, 029702 (2012).
- <sup>11</sup>An early comment on Ref. 2 [K. Matho, *Phys. Rev. Lett.* **108**, 029702 (2012)] and the response to this comment [B. S. Shastry, *Phys. Rev. Lett.* **108**, 029702 (2012)] are noteworthy. In these works the momentum dependence of the Dyson self-energy in the limit of large  $d$  was discussed, before the exact simplifications arising from that limit were known precisely. As shown in Ref. 8 and in the present work, ECFL does provide a momentum-independent Dyson self-energy in this limit.
- <sup>12</sup>A. Georges and G. Kotliar, *Phys. Rev. B* **45**, 6479 (1992).
- <sup>13</sup>The minimal version of ECFL is obtained by setting to zero all correlation functions satisfying the Gutzwiller constraint, whereas the symmetrized version adds in some terms that are known to be zero, in order to make the effective Hamiltonian manifestly Hermitian,<sup>3</sup> (p. 11).
- <sup>14</sup>To avoid the linear growth with  $\omega$  of the self-energy, one may define the Dyson-Mori self-energy, as in Ref. 4 with a factor  $1 - \frac{n}{2}$  in the numerator of the Green's function. However in the present work, since DMFT directly yields the Dyson self-energy, we simplify by quoting the results for this object directly.
- <sup>15</sup>D. Hansen and B. S. Shastry, *Phys. Rev. B* **87**, 245101 (2013).
- <sup>16</sup>A more complete discussion of the approach to the insulating limit within ECFL is not yet available. It is possible to also achieve a vanishing quasiparticle weight near half-filling in other ways as well, such as letting  $c_x$  diverge as  $\frac{1}{\delta}$ . We prefer the current scheme for its simplicity and economy.
- <sup>17</sup>W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989).
- <sup>18</sup>K. G. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975).
- <sup>19</sup>H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, *Phys. Rev. B* **21**, 1003 (1980).
- <sup>20</sup>R. Bulla, *Phys. Rev. Lett.* **83**, 136 (1999).
- <sup>21</sup>R. Bulla, T. Costi, and T. Pruschke, *Rev. Mod. Phys.* **80**, 395 (2008).
- <sup>22</sup>R. Žitko and T. Pruschke, *Phys. Rev. B* **79**, 085106 (2009).
- <sup>23</sup>R. Žitko, *Phys. Rev. B* **84**, 085142 (2011).
- <sup>24</sup>A. Weichselbaum and J. von Delft, *Phys. Rev. Lett.* **99**, 076402 (2007).
- <sup>25</sup>R. Bulla, A. C. Hewson, and T. Pruschke, *J. Phys.: Condens. Matter* **10**, 8365 (1998).
- <sup>26</sup>R. Peters, T. Pruschke, and F. B. Anders, *Phys. Rev. B* **74**, 245114 (2006).
- <sup>27</sup>R. Žitko, *Phys. Rev. B* **80**, 125125 (2009).
- <sup>28</sup>W. F. Brinkman and T. M. Rice, *Phys. Rev. B* **2**, 4302 (1970).
- <sup>29</sup>D. Vollhardt, *Rev. Mod. Phys.* **56**, 99 (1984).
- <sup>30</sup>P. Nozières, *La méthode de Gutzwiller*, Lecture Notes at Collège de France, 1986 (unpublished).
- <sup>31</sup>M. J. Rozenberg, G. Kotliar, and X. Y. Zhang, *Phys. Rev. B* **49**, 10181 (1994).
- <sup>32</sup>S. Sakai, S. Blanc, M. Civelli, Y. Gallais, M. Cazayous, M.-A. Measson, J. S. Wen, Z. J. Xu, G. D. Gu, G. Sangiovanni *et al.*, *Phys. Rev. Lett.* **111**, 107001 (2013).
- <sup>33</sup>G. Moeller, Q. Si, G. Kotliar, M. Rozenberg, and D. S. Fisher, *Phys. Rev. Lett.* **74**, 2082 (1995).
- <sup>34</sup>R. Bulla, T. Pruschke, and A. C. Hewson, *Physica B Condens. Matt.* **259**, 721 (1999).
- <sup>35</sup>M. Karski, C. Raas, and G. S. Uhrig, *Phys. Rev. B* **72**, 113110 (2005).
- <sup>36</sup>M. Karski, C. Raas, and G. S. Uhrig, *Phys. Rev. B* **77**, 075116 (2008).
- <sup>37</sup>Žiga Osolin and R. Žitko, *Phys. Rev. B* **87**, 245135 (2013).
- <sup>38</sup>M. M. Zemljič, P. Prelovšek, and T. Tohyama, *Phys. Rev. Lett.* **100**, 036402 (2008).
- <sup>39</sup>K. Byczuk, M. Kollar, K. Held, Y.-F. Yang, I. A. Nekrasov, T. Pruschke, and D. Vollhardt, *Nature Phys.* **3**, 168 (2007).
- <sup>40</sup>P. Grete, S. Schmitt, C. Raas, F. B. Anders, and G. S. Uhrig, *Phys. Rev. B* **84**, 205104 (2011).
- <sup>41</sup>K. Held, R. Peters, and A. Toschi, *Phys. Rev. Lett.* **110**, 246402 (2013).
- <sup>42</sup>M. Eckstein, M. Kollar, M. Potthoff, and D. Vollhardt, *Phys. Rev. B* **75**, 125103 (2007).
- <sup>43</sup>B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *Phys. Rev. B* **88**, 205108 (2013).
- <sup>44</sup>B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).
- <sup>45</sup>K. Haule and G. Kotliar, in *Properties and Applications of Thermoelectric Materials*, NATO Science for Peace and Security Series B: Physics and Biophysics, edited by V. Zlatić and A. C. Hewson (Springer, The Netherlands, 2009), pp. 119–131.
- <sup>46</sup>R. R. P. Singh and R. L. Glenister, *Phys. Rev. B* **46**, 14313 (1992).
- <sup>47</sup>T. M. Rice and K. Ueda, *Phys. Rev. Lett.* **55**, 995 (1985).
- <sup>48</sup>M. Yoshida, M. A. Whitaker, and L. N. Oliveira, *Phys. Rev. B* **41**, 9403 (1990).
- <sup>49</sup>V. L. Campo and L. N. Oliveira, *Phys. Rev. B* **72**, 104432 (2005).



Contents lists available at ScienceDirect

## Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)

# Extremely correlated Fermi liquids in the limit of infinite dimensions



Edward Perepelitsky\*, B. Sriram Shastry

Physics Department, University of California, Santa Cruz, CA 95064, USA

## HIGHLIGHTS

- Infinite-dimensional  $t$ - $J$  model ( $J = 0$ ) studied within new ECFL theory.
- Mapping to the infinite  $U$  Anderson model with self consistent hybridization.
- Single particle Green's function determined by two local self energies.
- Partial projection through control variable  $\lambda$ .
- Expansion carried out to  $O(\lambda^2)$  explicitly.

## ARTICLE INFO

### Article history:

Received 6 September 2013

Accepted 17 September 2013

Available online 25 September 2013

### Keywords:

High dimensions

 $t$ - $J$  model

Hubbard model

Extremely Correlated Fermi Liquid model

Strongly correlated electrons

## ABSTRACT

We study the infinite spatial dimensionality limit ( $d \rightarrow \infty$ ) of the recently developed Extremely Correlated Fermi Liquid (ECFL) theory (Shastry 2011, 2013) [17,18] for the  $t$ - $J$  model at  $J = 0$ . We directly analyze the Schwinger equations of motion for the Gutzwiller projected (i.e.  $U = \infty$ ) electron Green's function  $\mathcal{G}$ . From simplifications arising in this limit  $d \rightarrow \infty$ , we are able to make several exact statements about the theory. The ECFL Green's function is shown to have a momentum independent Dyson (Mori) self energy. For practical calculations we introduce a partial projection parameter  $\lambda$ , and obtain the complete set of ECFL integral equations to  $O(\lambda^2)$ . In a related publication (Zitko et al. 2013) [23], these equations are compared in detail with the dynamical mean field theory for the large  $U$  Hubbard model. Paralleling the well known mapping for the Hubbard model, we find that the infinite dimensional  $t$ - $J$  model (with  $J = 0$ ) can be mapped to the infinite- $U$  Anderson impurity model with a self-consistently determined set of parameters. This mapping extends individually to the auxiliary Green's function  $g$  and the caparison

\* Corresponding author. Tel.: +1 408 891 4906.

E-mail address: [eperepel@ucsc.edu](mailto:eperepel@ucsc.edu) (E. Perepelitsky).

factor  $\mu$ . Additionally, the optical conductivity is shown to be obtainable from  $\mathcal{G}$  with negligibly small vertex corrections. These results are shown to hold to each order in  $\lambda$ .

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

### 1.1. Motivation

The Hubbard model (HM) with the Hamiltonian:

$$H = - \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu \sum_i n_i, \quad (1)$$

has attracted great theoretical interest in condensed matter physics, and is also a fairly realistic model of strongly correlated materials such as the cuprates. While the small  $\frac{U}{t}$  limit is well described by standard Fermi-Liquid theory [1,2], the large and intermediate  $\frac{U}{t}$  (strongly correlated) cases are much less well understood. Considerable progress has been made by considering the HM in the limit of infinite dimensions [3–11]. One important result is that the Dyson self energy, defined by inverting the expression for the electron Green's function  $\mathcal{G}$ :

$$\mathcal{G}(k) = \frac{1}{i\omega_k + \mu - \epsilon_k - \Sigma_D(k)}, \quad (2)$$

becomes momentum independent in this limit [3–6]. Two other important results are the self-consistent mapping of the infinite dimensional HM onto the Anderson Impurity model (AIM), detailed in [8] (Dynamical Mean Field Theory), and the vanishing of the vertex corrections in the optical conductivity [10,11], so that the two particle response is obtainable from the single particle Green's function. The Dynamical Mean Field Theory (DMFT) provides a means for doing reliable numerical calculations for the Hubbard model, at any value of  $U$  and has continued to provide new, and interesting results [12,13].

A different approach to understanding strong correlations is to consider the extreme correlation limit, where one sets  $U \rightarrow \infty$  at the outset. In this case, the Hilbert space is Gutzwiller projected so that only single occupancy is allowed on each lattice site. One such extremely correlated model, the  $t$ - $J$  model, consists of taking the  $U \rightarrow \infty$  limit of the Hubbard model (the  $t$  part of the model) and adding on a nearest neighbor anti-ferromagnetic coupling term (the  $J$  part of the model). The  $t$  model studied here, is obtained by dropping the  $J$  term and thus is identical to the  $U = \infty$  limit of the HM. It has been argued by Anderson [14] that the  $t$ - $J$  model describes the physics of the cuprates, thereby providing an impetus for its detailed study. The Hamiltonian for this model can be written in terms of the Hubbard  $X$  operators as [15]

$$H = - \sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i\sigma} X_i^{\sigma\sigma} + \frac{1}{2} \sum_{ij\sigma} J_{ij} X_i^{\sigma\sigma} + \frac{1}{4} \sum_{ij\sigma_1\sigma_2} J_{ij} \{X_i^{\sigma_1\sigma_2} X_j^{\sigma_2\sigma_1} - X_i^{\sigma_1\sigma_1} X_j^{\sigma_2\sigma_2}\}. \quad (3)$$

The operator  $X_i^{ab} = |a\rangle\langle b|$  takes the electron at site  $i$  from the state  $|b\rangle$  to the state  $|a\rangle$ , where  $|a\rangle$  and  $|b\rangle$  are one of the three allowed states  $|\uparrow\rangle$ ,  $|\downarrow\rangle$ , or  $|-\rangle$ . Our present goal is to obtain a formally exact solution of the above  $t$  model in the limit of large dimensions by studying its equations of motion. This is designed to be methodologically independent of the available DMFT solution of the HM with  $U = \infty$ , and can be compared with it.

Our object of study is the Green's function written as

$$\mathcal{G}_{\sigma_1\sigma_2}(i, f) = -\langle T_\tau X_i^{0\sigma_1}(\tau_i) X_f^{\sigma_2 0}(\tau_f) \rangle, \quad (4)$$

where the angular brackets indicate the usual thermal average. Due to the non-canonical commutation relations of the  $X$  operators, the high frequency limit of Green's function is  $\frac{1-n}{i\omega_n}$  rather than  $\frac{1}{i\omega_n}$  as in the canonical case. To avoid linear growth of the self-energy in the high frequency limit [15], the Dyson self-energy must be redefined to the Dyson–Mori self energy [16] as in:

$$\mathcal{G}(k) = \frac{1 - \frac{n}{2}}{i\omega_k + \mu - \epsilon_k \left(1 - \frac{n}{2}\right) - \Sigma_{DM}(k)}. \quad (5)$$

Just as is the case for  $\Sigma_D$  in the finite- $U$  Hubbard model,  $\Sigma_{DM}$  is finite as  $i\omega \rightarrow \infty$  in the  $t$ - $J$  model.

Shastry has recently introduced a novel and promising approach for calculating correlation functions within the  $t$ - $J$  model based on Schwinger's formulation of field theory [15,17,18]. This has culminated in the theory of the Extremely Correlated Fermi Liquid (ECFL) [17,18]. This theory has been successfully benchmarked against: line shapes from (ARPES) experiments [19,20], high-temperature series [21] and the numerical renormalization group (NRG) calculations for the Anderson impurity model [22]. A recent theoretical benchmarking related to this work is the comparison with DMFT calculations for the large  $U$  Hubbard model in a concurrent publication [23], with the formulas found here. Indeed the main motivation of the present paper is to obtain results in the limit of large  $d$  for the same model, the  $t$ - $J$  model (at  $J = 0$ ) or equivalently the  $U = \infty$  Hubbard model by two different methods, the ECFL and the DMFT, allowing such a comparison.

In the ECFL theory, the physical Green's function  $\mathcal{G}(k)$  is factored into a canonical auxiliary Green's function  $\mathbf{g}(k)$  and an adaptive spectral  $\mu(k)$ , where  $k = (\vec{k}, i\omega_k)$ :

$$\mathcal{G}(k) = \mathbf{g}(k) \times \mu(k). \quad (6)$$

These two factors are in turn written in terms of two self-energies,  $\Phi(k)$  and  $\Psi(k)$ .

$$\mathbf{g}^{-1}(k) = i\omega_k + \mu - (1 - n/2)\epsilon_k - \Phi(k), \quad (7)$$

$$\mu(k) = 1 - \frac{n}{2} + \Psi(k). \quad (8)$$

Here  $\Phi(k)$  plays the role of a Dyson self-energy for the canonical Green's function  $\mathbf{g}(k)$ , and  $\Psi(k)$  is a frequency-dependent correction to  $\mu(k)$  from its high frequency value of  $1 - \frac{n}{2}$ .  $\Phi$  and  $\Psi$  are then given in terms of the vertices (i.e. functional derivatives w.r.t. the source of the  $\mathbf{g}^{-1}$  and  $\mu$ ) as will be described below, leading to a closed set of Schwinger differential equations (the ECFL equations of motion). These equations are in general intractable since there is no obvious small parameter, and therefore to enable practical calculations, an expansion is carried out in a partial projection parameter  $\lambda$ . Here  $\lambda$  interpolates between the free Fermi gas and the  $t$ - $J$  model. The meaning of  $\lambda$  as a partial projection parameter is detailed in [18], and may be summarized in the mapping  $X_i^{\sigma 0} \rightarrow f_{i\sigma}^\dagger (1 - \lambda n_{i\bar{\sigma}})$ , where  $f_{i\sigma}$  is a canonical electron operator. Thus at  $\lambda = 0$  we have canonical electrons, whereas at  $\lambda = 1$  we have the fully projected electrons.

In this work, our aim is to combine the two approaches, namely to consider the ECFL in the limit of infinite spatial dimensions. In this limit,  $J \rightarrow 0$ , and the infinite-dimensional  $t$ - $J$  model becomes the infinite dimensional infinite  $U$  Hubbard model (see Section 6A of Ref. [23] for a brief discussion of this). It is not clear a priori, whether or not the aforementioned results, valid for the infinite dimensional finite- $U$  Hubbard model, carry over to the infinite dimensional  $t$ - $J$  model. The possible conflict arises from the fact that in the case of the former, the ratio  $\frac{U}{d} \rightarrow 0$ , while in the case of the latter,  $\frac{U}{d} \rightarrow \infty$ . This question was raised in Ref. [24], pointing to the ECFL solution of the infinite dimensional  $t$ - $J$  model as a source of resolution. Working directly with the infinite- $U$  Hamiltonian (Eq. (3) with  $J = 0$ ), and using the corresponding ECFL equations of motion, we are able to address this challenging task and to show that the two limits  $U \rightarrow \infty$  and  $d \rightarrow \infty$  do in fact commute.

Moreover, we are able to determine the structure of the ECFL objects  $\Phi(k)$  and  $\Psi(k)$  in the limit of infinite dimensions. Such structural information has already been used to fit numerical results obtained through DMFT calculations to a convenient and flexible functional form [23]. Finally, we are able to elucidate the nature of the  $\lambda$  expansion in the large  $d$  limit. For readers who might be more interested in the results than the methodology, we provide a detailed summary of our results at the outset.

### 1.2. Results in the limit of infinite dimensions

We show that in the large  $d$  limit, the two self energies  $\Phi(k)$  and  $\Psi(k)$  simplify in the following way.

$$\Psi(k) = \Psi(i\omega_k), \quad (9)$$

$$\Phi(k) = \chi(i\omega_k) + \epsilon_k \Psi(i\omega_k). \quad (10)$$

These in turn show that the Dyson–Mori self energy behaves as

$$\Sigma_{DM}(k) = \Sigma_{DM}(i\omega_k) = \frac{(i\omega_k + \mu)\Psi(i\omega_k) + \left(1 - \frac{n}{2}\right)\chi(i\omega_k)}{1 - \frac{n}{2} + \Psi(i\omega_k)}, \quad (11)$$

and is therefore local in the limit of infinite dimensions. We show that to each order in the  $\lambda$  expansion,  $\Psi(i\omega_k)$  and  $\chi(i\omega_k)$  are each a product of an arbitrary number of factors, each of which take on the form  $\sum_{\vec{p}} g(\vec{p}, i\omega_p) \epsilon_{\vec{p}}^m$ , with  $m$  equal to zero or one, and with arbitrarily complex frequency dependence of the individual factors.

We show that just as in the finite  $U$  case [10,11], the optical conductivity is given by the expression

$$\sigma^{\alpha\beta}(\omega) = \frac{2}{i\omega} \sum_{\vec{p}, i\omega_p} g(\vec{p}, i\omega_p) v_{\vec{p}}^{\alpha} v_{\vec{p}}^{\beta} [g(\vec{p}, \omega + i\eta + i\omega_p) - g(\vec{p}, i\eta + i\omega_p)], \quad (12)$$

where  $v_{\vec{p}}^{\alpha}$  is the component of the velocity in the  $\alpha$  direction (Eq. (39)). We show that this formula can be applied at each order of the  $\lambda$  expansion.

We show that there is a self consistent mapping between the ECFL theory of the infinite-dimensional  $t$ - $J$  model and the ECFL theory of the infinite- $U$  Anderson impurity model (AIM) [22]. This mapping is similar in spirit to the mapping first discussed by Georges and Kotliar for the Hubbard model [8], but is made directly in the infinite  $U$  limit here. In this mapping,  $\mathbf{g}_{i,i}[\tau_i, \tau_f]$  and  $\mu_{i,i}[\tau_i, \tau_f]$  of the  $t$ - $J$  model are mapped to the objects  $\mathbf{g}[\tau_i, \tau_f]$  and  $\mu[\tau_i, \tau_f]$  of the Anderson model, written with the same symbols, but without the spatial or momentum labels. This mapping is valid under the self-consistency condition

$$\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(\vec{k}) = \sum_{\vec{k}} \frac{|V_{\vec{k}}|^2}{i\omega_n - \tilde{\epsilon}_{\vec{k}}} \mathbf{g}(i\omega_k), \quad (13)$$

where  $\epsilon_{\vec{k}}$  is the dispersion of the lattice in the  $t$ - $J$  model, and  $V_{\vec{k}}$  and  $\tilde{\epsilon}_{\vec{k}}$  are the hybridization and dispersion of the bath respectively in the Anderson impurity model. This self-consistency condition is shown to be equivalent to the standard self-consistency condition from DMFT [8,9]. We also show that the mapping holds to each order in  $\lambda$  under the same self-consistency condition. We note that this implies that ECFL computations for the infinite-dimensional  $t$ - $J$  model can be done with a DMFT-like self-consistency loop involving ECFL computations for the AIM. However, since the  $\lambda$  expansion provides integral equations which are relatively straightforward to solve numerically, this is not necessary as the  $t$ - $J$  model equations can be solved directly.

### 1.3. Outline of the paper

The paper is structured as follows. In Section 2, some basic facts about lattice sums in the limit of large dimensions and the ECFL equations of motion as well as the  $\lambda$  expansion are reviewed. Additionally, the spatial dependence of various standard and ECFL specific objects in the limit of large dimensions is stated. Finally, we introduce a class of local functions denoted as class- $L$  functions; these turn out to play a central role in the ECFL in the limit of large dimensions. In Sections 3.1 and 3.2, Eqs. (9) and (10) are proven in general and to each order in  $\lambda$ , and the locality of the Dyson–Mori self energy is shown as a consequence. In Section 3.3, Eq. (12) is shown to hold in general and to each order in  $\lambda$ . In Section 3.4, the ECFL self-consistent integral equations are derived to  $O(\lambda^2)$  in the large- $d$  limit. Finally, in Section 4, the ECFL of the infinite dimensional  $t$ - $J$  model is mapped onto the ECFL of the infinite- $U$  AIM under the self-consistency condition (Eq. (13)). This is done in general and to each order in  $\lambda$ .

## 2. Preliminaries

### 2.1. Spatial dependence of lattice sums in large $d$ dimensions

We take the hopping to nearest neighbor sites on the  $d$ -dimensional hypercube. In this case, it is well known [4] that  $t_{ij} \rightarrow \frac{1}{\sqrt{2d}} t_0$  with  $t_0$  of  $O(1)$ . We would like to exploit the smallness of individual  $t_{ij}$ 's, these can only contribute (after multiplying with another like object), if one of the indices is summed over the  $d$ -neighbors as in the simplest example  $\sum_j t_{ij}^2 = t_0^2$ . Extending this argument further, for a pair of sites  $(i, m)$  located at a (Manhattan metric) distance  $r_{im}$  on the hypercube, suppose there are two objects  $W_{i,m}$  and  $V_{i,m}$  who both have the dependence on  $r_{im}$ :  $V_{i,m}; W_{i,m} \sim O\left(\frac{1}{(\sqrt{d})^{r_{im}}}\right)$ . Then it follows that

$$W_{i,\mathbf{n}} V_{\mathbf{n},m} \sim O\left(\frac{1}{(\sqrt{d})^{r_{im}}}\right). \quad (14)$$

Here, and in the rest of the paper, bold and repeated indices are summed and/or integrated over. This relation can be understood by first considering the case that the site  $\mathbf{n}$  is on one of the shortest paths between  $i$  and  $m$ . In this case,  $r_{i\mathbf{n}} + r_{\mathbf{n}m} = r_{im}$  proving the relation. If,  $\mathbf{n}$  is a certain distance  $r_o$  off of a shortest path, then  $r_{i\mathbf{n}} + r_{\mathbf{n}m} = r_{im} + 2r_o$ . This introduces an extra factor of  $\frac{1}{d^{r_o}}$  into the lattice sum in Eq. (14). However, this factor is exactly cancelled by the  $d^{r_o}$  choices for the site  $\mathbf{n}$ . In this argument, the number of shortest paths between  $i$  and  $m$  is taken to be  $O(1)$ .

### 2.2. ECFL equations of motion and the $\lambda$ expansion

The ECFL equations of motion for the finite dimensional  $t$ - $J$  model can be found in Ref. [18]. There is some freedom in how these equations are written because one may add terms to them which vanish identically in the exact solution, but play a non-trivial role when implementing approximations (such as the  $\lambda$  expansion). We denote the version of these equations with no added terms the minimal theory, and the version containing the added terms the symmetrized theory (since the added terms make the resulting expressions symmetric in a certain sense). In Ref. [18], the ECFL equations of motion for the symmetrized theory are derived, and the added terms required to go from the minimal theory to the symmetrized theory are singled out. The ECFL equations for the minimal theory, which are the ones used in this paper and in Ref. [23], can therefore be obtained from those in Ref. [18] by dropping these extra terms.

Setting  $J \rightarrow 0$  (as discussed in Section 1.1), we write the minimal theory ECFL equations of motion in expanded form:

$$\begin{aligned} \mathbf{g}^{-1}[i, m] &= (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] + t[i, m] (1 - \lambda \gamma[i]) \\ &\quad + \lambda t[i, \mathbf{j}] \boldsymbol{\xi}^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \Lambda_*[\mathbf{n}, m; i], \\ \mu[i, m] &= (1 - \lambda \gamma[i]) \delta[i, m] - \lambda t[i, \mathbf{j}] \boldsymbol{\xi}^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i], \end{aligned} \quad (15)$$

where  $\mathcal{V}_i \equiv \mathcal{V}_i(\tau_i)$  is the Bosonic Schwinger source function, and we have used the notation  $\delta[i, m] = \delta_{i,m} \delta(\tau_i - \tau_m)$  and  $t[i, m] = t_{i,m} \delta(\tau_i - \tau_m)$ . These exact relations give the required objects  $\mathbf{g}$  and  $\mu$  in terms of the vertex functions. Here we also note that the local (in space and time) Green's function  $\gamma[i]$ , and the vertices  $\Lambda[n, m; i]$  and  $\mathcal{U}[n, m; i]$ , are defined as

$$\begin{aligned} \gamma[i] &= \mu^{(k)}[\mathbf{n}, i^+] \cdot \mathbf{g}^{(k)}[i, \mathbf{n}]; & \Lambda[n, m; i] &= -\frac{\delta}{\delta \mathcal{V}_i} \mathbf{g}^{-1}[n, m]; \\ \mathcal{U}[n, m; i] &= \frac{\delta}{\delta \mathcal{V}_i} \mu[n, m], \end{aligned} \quad (16)$$

where we have used the notation  $M_{\sigma_1, \sigma_2}^{(k)} = \sigma_1 \sigma_2 M_{\bar{\sigma}_2, \bar{\sigma}_1}$  to denote the time reversed matrix  $M^{(k)}$  of an arbitrary matrix  $M$ . These exact relations give the vertex functions in terms of the objects  $\mathbf{g}$  and  $\mu$ . The vertices defined above ( $\Lambda$  and  $\mathcal{U}$ ) have four spin indices, those of the object being differentiated and those of the source. For example,  $\mathcal{U}_{\sigma_a \sigma_b}^{\sigma_1 \sigma_2}[n, m; i] = \frac{\delta}{\delta v_i^{\sigma_a \sigma_b}} \mu_{\sigma_1 \sigma_2}[n, m]$ . In Eq. (15),  $\xi_{\sigma_a \sigma_b} = \sigma_a \sigma_b$ , and  $*$  indicates that these spin indices should also be carried over (after being flipped) to the bottom indices of the vertex, which is also marked with a  $*$ . The top indices of the vertex are given by the usual matrix multiplication. An illustrative example is useful here:  $(\xi^* \cdot \mathbf{g}[j, \mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i])_{\sigma_1 \sigma_2} = \sigma_1 \sigma_a \mathbf{g}_{\sigma_a, \sigma_b}[j, \mathbf{n}] \frac{\delta}{\delta v_i^{\sigma_1 \sigma_a}} \mu_{\sigma_b, \sigma_2}[\mathbf{n}, m]$ . Finally, in order to ensure that the shift identities (Ref. [18]) are satisfied, the substitution  $t_{ij} \rightarrow t_{ij} + \frac{u_0}{2} \delta_{ij}$  is made, where  $u_0$  is the second chemical potential. For the sake of clarity, this substitution will be ignored in the proofs given below, although they are easily generalized to account for it. This generalization is discussed at the end of Section 3.1.

The  $\lambda$  expansion is obtained by expanding Eqs. (15) and (16) iteratively in the continuity parameter  $\lambda$ . The  $\lambda = 0$  limit of these equations is the free Fermi gas. Therefore, a direct expansion in  $\lambda$  will lead to a series in  $\lambda$  in which each term is made up of the hopping  $t_{ij}$  and the free Fermi gas Green's function  $\mathbf{g}_0[i, f]$ . As is the case in the Feynman series, this can be reorganized into a skeleton expansion in which only the skeleton graphs are kept and  $\mathbf{g}_0[i, f] \rightarrow \mathbf{g}[i, f]$ . However, one can also obtain the skeleton expansion directly by expanding Eqs. (15) and (16) in  $\lambda$ , but treating  $\mathbf{g}[i, f]$  as a zeroth order (i.e. unexpanded) object in the expansion. This expansion is carried out to second order for the finite-dimensional case in Ref. [18]. In doing this expansion, one must evaluate the functional derivative  $\frac{\delta \mathbf{g}}{\delta v}$ . This is done with the help of the following useful formula which stems from the product rule for functional derivatives:

$$\frac{\delta \mathbf{g}[i, m]}{\delta v_r} = \mathbf{g}[i, \mathbf{x}] \cdot \Lambda[\mathbf{x}, \mathbf{y}, r] \cdot \mathbf{g}[\mathbf{y}, m]. \quad (17)$$

This is an exact formula and will be used extensively in the arguments given below. Within the  $\lambda$  expansion, the LHS is evaluated to a certain order in  $\lambda$  by taking the vertex  $\Lambda$  on the RHS to be of that order in  $\lambda$ .

### 2.3. Leading order spatial dependence of various objects

All objects may be expanded in the inverse square root of the number of spatial dimensions  $d$ . The lowest order term in the physical Green's function  $\mathcal{G}[i, f]$  must be at least  $O\left(\frac{1}{(\sqrt{d})^{r_{if}}}\right)$ . This must be so because it takes at least  $r_{if}$  hops to get from the site  $i$  to the site  $f$ . Any terms that contribute to  $\mathcal{G}[i, f]$  at higher order than  $O\left(\frac{1}{(\sqrt{d})^{r_{if}}}\right)$  are neglected in the large  $d$  limit. In a similar vein, the lowest order term in  $\mathbf{g}[i, f]$ ,  $\mathbf{g}^{-1}[i, f]$ ,  $\mu[i, f]$ ,  $\Lambda[i, f; r]$ , and  $\mathcal{U}[i, f; r]$  must be at least  $O\left(\frac{1}{(\sqrt{d})^{r_{if}}}\right)$ . Furthermore, using the real space version of Eqs. (6) and (14), we see that any terms of higher order than this in  $\mathbf{g}[i, f]$  and  $\mu[i, f]$  will result in a higher order term in  $\mathcal{G}[i, f]$  and may therefore be neglected as well. Finally, using matrix inversion in the space–time indices, we see that higher order terms may also be dropped from  $\mathbf{g}^{-1}[i, f]$  as these will lead to higher order terms in  $\mathbf{g}[i, f]$ , and using Eq. (15), higher order terms may be dropped from  $\Lambda[i, f; r]$ , and  $\mathcal{U}[i, f; r]$  as these will lead to higher order terms in  $\mathbf{g}^{-1}[i, f]$  and  $\mu[i, f]$  respectively. In summary, in all objects:  $\mathcal{G}[i, f]$ ,  $\mathbf{g}[i, f]$ ,  $\mathbf{g}^{-1}[i, f]$ ,  $\mu[i, f]$ ,  $\Lambda[i, f; r]$ , and  $\mathcal{U}[i, f; r]$ , terms of higher order than  $O\left(\frac{1}{(\sqrt{d})^{r_{if}}}\right)$  may be neglected in the large  $d$  limit.

We also note that the correlation function  $\Pi_{\alpha\beta}[i, f]$  appearing in Eq. (40) must be at least  $O\left(\frac{1}{d^{r_{if}}}\right)$ . This is due to the fact that unlike the creation and destruction operators which appear in the Green's function, the current operators appearing in this correlation function conserve particle number. Hence, one must hop from site  $i$  to site  $f$  and back, which takes  $2 \times r_{if}$  hops. Any terms that contribute to  $\Pi_{\alpha\beta}[i, f]$  at higher order than  $O\left(\frac{1}{d^{r_{if}}}\right)$  are neglected in the large  $d$  limit.

## 2.4. Class $L$ functions

For the arguments given below, we need to define a class of *localized functions*, denoted as class  $L$  functions. A class  $L$  function  $L_i$  has three properties.

- (a)  $L_i \sim O\left(\frac{1}{d^0}\right)$ .
- (b)  $L_i$  is a function of only one site  $i$ , and an arbitrary number of time variables. Upon turning off the sources, it becomes translationally invariant, but an arbitrary function of frequencies.
- (c) The  $\mathcal{V}$  source derivative of  $L_i$  is also localized:

$$\frac{\delta}{\delta \mathcal{V}_i} L_j = \delta_{ij} L'_i, \quad (18)$$

with  $L'_i$  again a class- $L$  function.

Our proofs deal with functions that turn out to be of this class. Iterating property (c), the following equation must hold for any positive integer  $s$ .

$$\frac{\delta}{\delta \mathcal{V}_{r_1}} \cdots \frac{\delta}{\delta \mathcal{V}_{r_s}} L_i = \delta_{ir_1} \cdots \delta_{ir_s} \frac{\delta}{\delta \mathcal{V}_i(\tau_{r_1})} \cdots \frac{\delta}{\delta \mathcal{V}_i(\tau_{r_s})} L_i. \quad (19)$$

In the presence of the current source  $\kappa$  (Eq. (42)), class  $L$  functions acquire one additional property (d): consider a typical contribution to  $\Pi_{\alpha\beta}[i, f]$  (Eq. (44)) denoted by  $O_{if}$

$$O_{if} = W_{f,\mathbf{x}} \frac{\delta}{\delta \kappa_i^\alpha} (L_{\mathbf{x}}) V_{\mathbf{x},f}, \quad (20)$$

where the functions  $V_{\mathbf{x},f}$ ,  $W_{f,\mathbf{x}} \sim O\left(\frac{1}{(\sqrt{d})^{|\mathbf{x}|}}\right)$ . Then, neglecting terms of higher order than  $O\left(\frac{1}{d^{|\mathbf{x}|}}\right)$  in  $O_{if}$ ,  $\sum_{i=f} O_{if} \rightarrow 0$  as  $\mathcal{A} \rightarrow 0$ . Again iterating property (c) and using property (d), the following must hold for any nonnegative integer  $s$ :

$$\sum_{i=f} \left( W_{f,\mathbf{x}} \frac{\delta}{\delta \kappa_i^\alpha} \frac{\delta}{\delta \mathcal{V}_{\mathbf{x}}(\tau_{r_1})} \cdots \frac{\delta}{\delta \mathcal{V}_{\mathbf{x}}(\tau_{r_s})} (L_{\mathbf{x}}) V_{\mathbf{x},f} \right)_{\mathcal{A} \rightarrow 0} = 0. \quad (21)$$

## 3. Limit of large dimensionality through the ECFL equations of motion

### 3.1. Simplification of the ECFL self energies

We use notation in which we indicate spatial dependence by subscripts, so that  $\mathbf{g}[i, j] \rightarrow \mathbf{g}_{i,j}[\tau_i, \tau_j]$ , and recall that  $t[i, j] = t_{i,j} \delta(\tau_i - \tau_j)$ ,  $\delta[i, j] = \delta_{i,j} \delta(\tau_i - \tau_j)$ ,  $\delta[\tau_i, \tau_j] = \delta(\tau_i - \tau_j)$  etc. After some inspection of Eqs. (15) and (16) in the limit of high dimension, we make an Ansatz – to be proven below – namely

$$\begin{aligned} \mathbf{g}^{-1}[i, m] &= (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] + t[i, m] (1 - \lambda \gamma[i]) - \lambda \delta_{i,m} \chi_i[\tau_i, \tau_m] \\ &\quad + \lambda t_{i,m} \Psi_i[\tau_i, \tau_m], \\ \mu[i, m] &= \delta[i, m] (1 - \lambda \gamma[i]) + \lambda \delta_{i,m} \Psi_i[\tau_i, \tau_m], \end{aligned} \quad (22)$$

where  $\Psi_i[\tau_i, \tau_m]$ ,  $\chi_i[\tau_i, \tau_m]$ , and  $\gamma[i]$  are class  $L$  functions. We will prove Eq. (22) by assuming that it is true, and then showing that this assumption is consistent with the equations of motion (Eqs. (15) and (16)). This argument will consist of a loop which begins with Eq. (22). Then, substituting this equation into Eq. (16), we will derive a certain form for  $\Lambda$ ,  $\mathcal{U}$ , and  $\gamma$ . Finally, substituting these objects into Eq. (15), and using simplifications which occur in the large  $d$  limit, we will complete the loop and arrive back at Eq. (22).

Substituting Eq. (22) into Eq. (16), we find that the vertices and  $\gamma[i]$  have the following form.

$$\begin{aligned} \Lambda[n, m; i] &= \delta_{i,n} \delta_{i,m} A_i[\tau_n, \tau_m; \tau_i] + \delta_{i,n} t_{n,m} B_i[\tau_n, \tau_m; \tau_i], \\ \mathcal{U}[n, m; i] &= -\delta_{i,n} \delta_{i,m} B_i[\tau_n, \tau_m; \tau_i], \\ \gamma[i] &= (1 - \lambda \gamma^{(k)}[i]) \mathbf{g}^{(k)}[i, i] + \lambda \Psi_i^{(k)}[\tau_j, \tau_i] \mathbf{g}_{ii}^{(k)}[\tau_i, \tau_j], \end{aligned} \quad (23)$$

where we defined two new functions:

$$\begin{aligned} A_i[\tau_n, \tau_m; \tau_i] &= \delta[\tau_i, \tau_n] \delta[\tau_i, \tau_m] \mathbb{1} + \lambda \frac{\delta}{\delta \mathcal{V}_i} \chi_i[\tau_n, \tau_m], \\ B_i[\tau_n, \tau_m; \tau_i] &= \lambda \delta[\tau_n, \tau_m] \frac{\delta}{\delta \mathcal{V}_i} \gamma_i[\tau_n] - \lambda \frac{\delta}{\delta \mathcal{V}_i} \Psi_i[\tau_n, \tau_m]. \end{aligned} \quad (24)$$

Here  $A_i$  and  $B_i$  are class  $L$  functions since they inherit this property from  $\Psi_i$ ,  $\chi_i$ , and  $\gamma[i]$  by functional differentiation. Substituting Eq. (23) into Eq. (15) and comparing with Eq. (22),

$$\begin{aligned} \chi_i[\tau_i, \tau_m] &= -t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_n] \cdot A_{i,*}[\tau_n, \tau_m; \tau_i], \\ \Psi_i[\tau_i, \tau_m] &= t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_n] \cdot B_{i,*}[\tau_n, \tau_m; \tau_i]. \end{aligned} \quad (25)$$

If we can now show that  $\chi_i$ ,  $\Psi_i$ , and  $\gamma[i]$  as defined in Eqs. (23) and (25) are Class  $L$  functions, we will have justified our Ansatz and therefore we will have proven all of the above equations. To do this, we must show that  $\mathbf{g}_{ii}[\tau_i, \tau_m]$  and  $t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_m]$  are Class  $L$  functions. Taking their functional derivatives we obtain:

$$\begin{aligned} \frac{\delta}{\delta \mathcal{V}_r} t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_m] &= t_{i,j} \mathbf{g}_{j,r}[\tau_i, \tau_k] A_r[\tau_k, \tau_i; \tau_r] \mathbf{g}_{r,i}[\tau_i, \tau_m] \\ &\quad + t_{i,j} \mathbf{g}_{j,r}[\tau_i, \tau_k] B_r[\tau_k, \tau_i; \tau_r] t_{r,i} \mathbf{g}_{i,i}[\tau_i, \tau_m], \end{aligned} \quad (26)$$

and

$$\begin{aligned} \frac{\delta}{\delta \mathcal{V}_r} \mathbf{g}_{i,i}[\tau_i, \tau_m] &= \mathbf{g}_{i,r}[\tau_i, \tau_k] A_r[\tau_k, \tau_i; \tau_r] \mathbf{g}_{r,i}[\tau_i, \tau_m] \\ &\quad + \mathbf{g}_{i,r}[\tau_i, \tau_k] B_r[\tau_k, \tau_i; \tau_r] t_{r,i} \mathbf{g}_{i,i}[\tau_i, \tau_m]. \end{aligned} \quad (27)$$

Using Eq. (14), the terms on the RHS of Eqs. (26) and (27) survive the large  $d$  limit if and only if  $r = i$ . Moreover, upon making the substitution  $r \rightarrow i$ , we see that the RHS is made up of the same objects that appear on the LHS of the equations (as well as the class  $L$  functions  $A$  and  $B$ ). Therefore, this argument can be iterated to any number of derivatives acting on  $t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_m]$  or  $\mathbf{g}_{i,i}[\tau_i, \tau_m]$  (as required by Eq. (19)), which are therefore class  $L$  functions. Thus, we have shown the self-consistency of our ansatz Eq. (22).

The above results hold for any value of  $\lambda$ , since the proof was done with  $\lambda$  present in all of the equations. In the bare expansion, this would imply that they also hold to each order in  $\lambda$ . However, this line of reasoning is not as straightforward in the skeleton expansion because each order in the skeleton expansion contains contributions from all orders in the bare expansion. Nonetheless, the above results do hold to each order in  $\lambda$  in the skeleton expansion. In proving this, we shall shed more light on the nature of the objects  $\Psi_i$ ,  $\chi_i$ ,  $\gamma[i]$ ,  $A_i$ , and  $B_i$ . In particular, we will show that they satisfy a certain explicit form stated below in Eq. (28). We will do this using an inductive argument, in which we will assume that they have this form through a certain order in  $\lambda$ , and then substituting this form into the equations of motion, will show that it must hold for the next order.

We now use the symbol  $R_i$  as a proxy for either of the two functions  $\mathbf{g}_{i,i}[\tau_n, \tau_m]$  or  $t_{i,j} \mathbf{g}_{j,i}[\tau_n, \tau_m]$  where the time indices are arbitrary. *Inductive hypothesis:* through  $n$ th order in  $\lambda$ , Eqs. (22) and (23) hold. Through  $n - 1$ st order in  $\lambda$ , the objects  $\Psi_i$ ,  $\chi_i$ , and  $\gamma[i]$ , and through  $n$ th order, the objects  $A_i$  and  $B_i$ , (all denoted below by the generic object  $L_i$ ) can be written as the following product (multiplied by some delta functions in time variables):

$$(L_i)^{(n)} = \lambda^n (R_i)^m, \quad (28)$$

where  $m$  is arbitrary. We first examine the base case of zeroth order. In this case,

$$A_i^{(0)}[\tau_n, \tau_m; \tau_i] = \delta[\tau_i, \tau_n]\delta[\tau_i, \tau_m]; \quad B_i^{(0)}[\tau_n, \tau_m; \tau_i] = 0. \quad (29)$$

Clearly the hypothesis is satisfied. Now, we prove the inductive step. Explicitly displaying the order in  $\lambda$  of all objects, the equations for  $\chi$ ,  $\Psi$ , and  $\gamma$  (Eqs. (25) and (23)) become

$$\begin{aligned} \chi_i^{(n)}[\tau_i, \tau_m] &= -t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_n] \cdot A_{i,*}^{(n)}[\tau_n, \tau_m; \tau_i], \\ \Psi_i^{(n)}[\tau_i, \tau_m] &= t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_n] \cdot B_{i,*}^{(n)}[\tau_n, \tau_m; \tau_i], \\ \gamma^{(n)}[i] &= -\lambda \gamma^{(k)(n-1)}[i] \mathbf{g}^{(k)}[i, i] + \lambda \Psi_i^{(k)(n-1)}[\tau_j, \tau_i] \mathbf{g}_{ii}^{(k)}[\tau_i, \tau_j]. \end{aligned} \quad (30)$$

By the inductive hypothesis,  $\chi_i^{(n)}$ ,  $\Psi_i^{(n)}$ , and  $\gamma^{(n)}[i]$  have the required form. The equations for  $A$  and  $B$  (Eq. (24)) become

$$\begin{aligned} A_i^{(n+1)}[\tau_n, \tau_m; \tau_i] &= \lambda \left( \sum_{r \leq n} \frac{\delta}{\delta \mathcal{V}_i} \chi_i^{(r)}[\tau_n, \tau_m] \right)^{(n)}, \\ B_i^{(n+1)}[\tau_n, \tau_m; \tau_i] &= \lambda \delta[\tau_n, \tau_m] \left( \sum_{r \leq n} \frac{\delta}{\delta \mathcal{V}_i} \gamma_i^{(r)}[\tau_n] \right)^{(n)} - \lambda \left( \sum_{r \leq n} \frac{\delta}{\delta \mathcal{V}_i} \Psi_i^{(r)}[\tau_n, \tau_m] \right)^{(n)}. \end{aligned} \quad (31)$$

To see that  $A^{(n+1)}$  and  $B^{(n+1)}$  have the required form we note that for all  $l \leq n$ ,

$$\begin{aligned} \left( \frac{\delta}{\delta \mathcal{V}_r} t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_m] \right)^{(l)} &= t_{i,j} \mathbf{g}_{j,r}[\tau_i, \tau_k] A_r^{(l)}[\tau_k, \tau_i; \tau_r] \mathbf{g}_{r,i}[\tau_i, \tau_m] \\ &\quad + t_{i,j} \mathbf{g}_{j,r}[\tau_i, \tau_k] B_r^{(l)}[\tau_k, \tau_i; \tau_r] t_{r,i} \mathbf{g}_{i,i}[\tau_i, \tau_m], \end{aligned} \quad (32)$$

and

$$\begin{aligned} \left( \frac{\delta}{\delta \mathcal{V}_r} \mathbf{g}_{i,i}[\tau_i, \tau_m] \right)^{(l)} &= \mathbf{g}_{i,r}[\tau_i, \tau_k] A_r^{(l)}[\tau_k, \tau_i; \tau_r] \mathbf{g}_{r,i}[\tau_i, \tau_m] \\ &\quad + \mathbf{g}_{i,r}[\tau_i, \tau_k] B_r^{(l)}[\tau_k, \tau_i; \tau_r] t_{r,i} \mathbf{g}_{i,i}[\tau_i, \tau_m]. \end{aligned} \quad (33)$$

In the limit of large dimensions,  $r \rightarrow i$ . We can therefore (using the inductive hypothesis) write the RHS of Eqs. (32) and (33) as  $\lambda^l (R_i)^m$ . Applying Eq. (28) (which has been shown to hold for  $\chi_i^{(n)}$ ,  $\Psi_i^{(n)}$ , and  $\gamma^{(n)}[i]$ ) to Eq. (31), we may write

$$\begin{aligned} A_i^{(n+1)} &= \sum_{r=0}^n \lambda^{r+1} \left( \frac{\delta}{\delta \mathcal{V}_i} (R_i)^m \right)^{(n-r)}, \\ B_i^{(n+1)} &= \sum_{r=0}^n \lambda^{r+1} \left( \frac{\delta}{\delta \mathcal{V}_i} (R_i)^m \right)^{(n-r)}. \end{aligned} \quad (34)$$

Eq. (34), in conjunction with Eqs. (32) and (33), shows that  $A_i^{(n+1)}$  and  $B_i^{(n+1)}$  have the required form. This completes the proof.

Since  $t_{i,j}$  is independent of the source, the substitution  $t_{i,j} \rightarrow t_{i,j} + \frac{u_0}{2} \delta_{i,j}$  can be made directly into all of the above equations. The only problem that could potentially arise involves Eqs. (26) and (27), where the large  $d$  simplifications are actually used. However, one can check that this substitution does not affect the simplifications. Therefore, this substitution merely adds the term  $\lambda \frac{u_0}{2} \delta_{i,m} \Psi_i[\tau_i, \tau_m] - \lambda \frac{u_0}{2} \delta[i, m] \gamma[i]$  to  $\mathbf{g}^{-1}[i, m]$ , and everywhere replaces the local function  $t_{i,j} \mathbf{g}_{j,i}[\tau_n, \tau_m]$  with the local function  $t_{i,j} \mathbf{g}_{j,i}[\tau_n, \tau_m] + \frac{u_0}{2} \mathbf{g}_{i,i}[\tau_n, \tau_m]$ . This can be seen explicitly in the  $O(\lambda^2)$  equations in Section 3.4, and does not change the general structure of the solution.

### 3.2. The zero source limit

Setting the sources to zero, the system becomes translationally invariant so that all objects can be written in momentum space. Additionally,  $\gamma[i] \rightarrow \frac{n}{2}$ . Then, the above results can be summed up in the following formulae (in which we set  $\lambda = 1$ ):

$$\begin{aligned} \mathbf{g}^{-1}(k) &= i\omega_k + \boldsymbol{\mu} - \varepsilon_k \left(1 - \frac{n}{2}\right) - \chi(i\omega_k) - \varepsilon_k \Psi(i\omega_k), \\ \mu(k) &= 1 - \frac{n}{2} + \Psi(i\omega_k), \end{aligned} \quad (35)$$

where  $\Psi(i\omega_k)$  and  $\chi(i\omega_k)$  are the two momentum independent self-energies of the ECFL in infinite dimensions. In terms of these self-energies, the physical Green's function is written as

$$\mathcal{G}(k) = \frac{1 - \frac{n}{2} + \Psi(i\omega_k)}{i\omega_k + \boldsymbol{\mu} - \varepsilon_k \left(1 - \frac{n}{2}\right) - \chi(i\omega_k) - \varepsilon_k \Psi(i\omega_k)}. \quad (36)$$

Comparing with the standard form of the Green's function in terms of the Dyson–Mori self energy

$$\mathcal{G}(k) = \frac{1 - \frac{n}{2}}{i\omega_k + \boldsymbol{\mu} - \varepsilon_k \left(1 - \frac{n}{2}\right) - \Sigma_{DM}(k)}, \quad (37)$$

we see the momentum independence of the Dyson–Mori self energy  $\Sigma_{DM}(k) = \Sigma_{DM}(i\omega_k)$ , and

$$\Sigma_{DM}(i\omega_k) = \frac{(i\omega_k + \boldsymbol{\mu})\Psi(i\omega_k) + \left(1 - \frac{n}{2}\right)\chi(i\omega_k)}{1 - \frac{n}{2} + \Psi(i\omega_k)}. \quad (38)$$

### 3.3. Conductivity in the limit of large dimensions

It is well known that for the finite- $U$  Hubbard model in the limit of large dimensions, for zero wave vector, vertex corrections can be neglected in the current–current correlation function [10,9]. This simple observation allows one to express the optical conductivity in terms of the single particle Green's function as in Eq. (50). We show that this is also the case for the infinite dimensional  $t$ - $J$  model. Moreover, a question of practical importance for the purpose of calculating the optical conductivity within the framework of ECFL, is whether or not Eq. (50) can be applied at each order in the  $\lambda$  expansion (as is done in Ref. [23]). We show that it can be applied and is the correct procedure. First, we define the relevant objects.

The Schrödinger picture current operator for site  $j$  in the direction  $\alpha$  is defined as follows:

$$J_j^\alpha = i \sum_{k\sigma} v_{k,j}^\alpha \chi_k^{\sigma 0} \chi_j^{0\sigma}; \quad v_{k,j}^\alpha = t_{k,j} (\vec{R}_k - \vec{R}_j)_\alpha, \quad (39)$$

so that  $v$  is a velocity. Using the notation  $J_i^\alpha[i] = J_i^\alpha(\tau_i)$ ;  $\tilde{J}^\alpha[i] = J^\alpha[i] - \langle J^\alpha[i] \rangle$ , we define the correlation function  $\Pi_{\alpha\beta}[i, f]$  and its Fourier transform as

$$\begin{aligned} \Pi_{\alpha\beta}[i, f] &= \langle T_\tau \tilde{J}^\alpha[i] \tilde{J}^\beta[f] \rangle; \\ \Pi_{\alpha\beta}(\vec{q}, i\Omega_n) &= \int_0^\beta d(\tau_i - \tau_f) e^{i\Omega_n(\tau_i - \tau_f)} \sum_{i-f} e^{-i\vec{q} \cdot (\vec{R}_i - \vec{R}_f)} \Pi_{\alpha\beta}[i, f]. \end{aligned} \quad (40)$$

The optical conductivity can be given in terms of this object as

$$\sigma^{\alpha\beta}(\omega) = \frac{1}{i\omega - \eta} \left[ \Pi_{\alpha\beta}(\vec{0}, \omega + i\eta) - \Pi_{\alpha\beta}(\vec{0}, i\eta) \right], \quad (41)$$

where  $\eta = 0^+$ . We would like to express the object  $\Pi_{\alpha\beta}[i, f]$  as a functional derivative of the Green's function. To this end, we add a source which couples to the current operator

$$\mathcal{A} \rightarrow \mathcal{A} + \sum_{j\alpha} \int_0^\beta d\tau \kappa_j^\alpha(\tau) J_j^\alpha(\tau). \quad (42)$$

In terms of the  $\kappa$  source derivative of the Green's function, and using the definitions  $v^\alpha[i, j] = v_{i,j}^\alpha \delta(\tau_i - \tau_j)$ ;  $\kappa_i^\alpha = \kappa_i^\alpha(\tau_i)$ ,  $\Pi_{\alpha\beta}[i, f]$  is given as

$$\Pi_{\alpha\beta}[i, f] = -i \text{Tr} \left( \frac{\delta}{\delta \kappa_i^\alpha} \mathcal{G}[f, \mathbf{j}] v^\beta[\mathbf{j}, f^+] \right)_{\mathcal{A} \rightarrow 0}, \quad (43)$$

where the trace is over the spin degrees of freedom only. We expand the RHS of this equation using Eq. (17) (which holds equally well for the  $\kappa$  source derivative), finally obtaining an expression for  $\Pi_{\alpha\beta}[i, f]$  in terms of the  $\kappa$  source derivatives of  $\mathbf{g}^{-1}$  and  $\mu$ .

$$\begin{aligned} \Pi_{\alpha\beta}[i, f] = & i \text{Tr} \left( \mathbf{g}[f, \mathbf{x}] \frac{\delta}{\delta \kappa_i^\alpha} \mathbf{g}^{-1}[\mathbf{x}, \mathbf{y}] \mathbf{g}[\mathbf{y}, \mathbf{k}] \mu[\mathbf{k}, \mathbf{j}] v^\beta[\mathbf{j}, f^+] \right)_{\mathcal{A} \rightarrow 0} \\ & - i \text{Tr} \left( \mathbf{g}[f, \mathbf{k}] \frac{\delta}{\delta \kappa_i^\alpha} \mu[\mathbf{k}, \mathbf{j}] v^\beta[\mathbf{j}, f^+] \right)_{\mathcal{A} \rightarrow 0}. \end{aligned} \quad (44)$$

We now consider how the additional source Eq. (42) affects the ECFL equations of motion (Eqs. (15) and (16)). The source enters into the equations of motion in the same way as the Hamiltonian does, via its commutator with the destruction operator,  $X_i^{0\sigma}$ . Moreover, the source has the same form as the Hamiltonian, with the hopping in the kinetic energy replaced by the velocity in the current operator. Therefore, the additional source affects the equations of motion only through the substitution

$$t[i, f] \rightarrow t[i, f] - i \sum_\alpha \kappa_f^\alpha v^\alpha[i, f]. \quad (45)$$

Thus, the new equations of motion can be read off from Eq. (15) as

$$\begin{aligned} \mathbf{g}^{-1}[i, m] = & (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] + \left( t[i, m] - i \sum_\alpha \kappa_m^\alpha v^\alpha[i, m] \right) \\ & \times (1 - \lambda \gamma[i]) + \lambda \left( t[i, \mathbf{j}] - i \sum_\alpha \kappa_j^\alpha v^\alpha[i, \mathbf{j}] \right) \boldsymbol{\xi}^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \mathcal{A}_*[\mathbf{n}, m; i], \quad (46) \\ \mu[i, m] = & (1 - \lambda \gamma[i]) \delta[i, m] - \lambda \left( t[i, \mathbf{j}] - i \sum_\alpha \kappa_j^\alpha v^\alpha[i, \mathbf{j}] \right) \boldsymbol{\xi}^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i]. \end{aligned}$$

Since there is no source derivative with respect to  $\kappa$  in the equations of motion and  $v^\alpha[i, f]$  is of the same order in  $\frac{1}{\sqrt{d}}$  as  $t[i, f]$ , all of the results derived in Section 3.1 continue to hold after making the substitution in Eq. (45). In particular, we showed that  $\mathbf{g}^{-1}[i, m]$  and  $\mu[i, m]$  have the following form (Eq. (22)):

$$\begin{aligned} \mathbf{g}^{-1}[i, m] = & (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] - \lambda \delta_{i,m} \chi_i[\tau_i, \tau_m] + \left( t[i, m] - i \sum_\alpha \kappa_m^\alpha v^\alpha[i, m] \right) \\ & \times (1 - \lambda \gamma[i]) + \lambda \left( t_{i,m} - i \sum_\alpha \kappa_m^\alpha v_{i,m}^\alpha \right) \Psi_i[\tau_i, \tau_m], \\ \mu[i, m] = & \delta[i, m] (1 - \lambda \gamma[i]) + \lambda \delta_{i,m} \Psi_i[\tau_i, \tau_m], \end{aligned} \quad (47)$$

where  $\chi_i$ ,  $\Psi_i$ , and  $\gamma[i]$  have properties (a)–(c) of class  $L$  functions (Section 2.4), and are defined by Eqs. (22) through (25). We shall now further assume that they also satisfy property (d) (Eq. (21))

and show that this assumption is consistent with their definitions. This, in turn, will allow us to demonstrate the validity of Eq. (50).

Our task is then to show that  $\chi_i$ ,  $\Psi_i$ , and  $\gamma[i]$ , as defined in the last line of Eqs. (23) and (25), satisfy Eq. (21). By Eq. (24),  $A_i$  and  $B_i$  satisfy Eq. (21) since they inherit this property from  $\chi_i$ ,  $\Psi_i$ , and  $\gamma[i]$ . It remains to show that  $\mathbf{g}_{\mathbf{x},\mathbf{x}}[\tau_n, \tau_m]$  and  $(t_{\mathbf{x},\mathbf{j}} - i \sum_{\alpha} \kappa_{\mathbf{j}}^{\alpha}(\tau_n) v_{\mathbf{x},\mathbf{j}}^{\alpha}) \mathbf{g}_{\mathbf{j},\mathbf{x}}[\tau_n, \tau_m]$  (the time indices are arbitrary) satisfy this equation.

Defining the notation  $w_{i,f}(\tau_i) \equiv t_{i,f} - i \sum_{\alpha} \kappa_f^{\alpha}(\tau_i) v_{i,f}^{\alpha}$ , and using (the  $\kappa$  source derivative version of) Eq. (17) as well as Eq. (47), we find that

$$\begin{aligned} \left( \frac{\delta}{\delta \kappa_i^{\alpha}} w_{\mathbf{x},\mathbf{j}}(\tau_n) \mathbf{g}_{\mathbf{j},\mathbf{x}}[\tau_n, \tau_m] \right)_{\mathcal{A} \rightarrow 0} &= -i \delta[\tau_i, \tau_n] v_{\mathbf{x},i}^{\alpha} \mathbf{g}_{i,\mathbf{x}}[\tau_i, \tau_m] + i t_{\mathbf{x},\mathbf{j}} \mathbf{g}_{\mathbf{j},\mathbf{a}}[\tau_n, \tau_{\mathbf{a}}] \\ &\quad \times (1 - \lambda \gamma[\mathbf{a}] \delta[\tau_{\mathbf{a}}, \tau_i] + \lambda \Psi_{\mathbf{a}}[\tau_{\mathbf{a}}, \tau_i]) v_{\mathbf{a},i}^{\alpha} \mathbf{g}_{i,\mathbf{x}}[\tau_i, \tau_m] \\ &\quad + \lambda t_{\mathbf{x},\mathbf{j}} \mathbf{g}_{\mathbf{j},\mathbf{a}}[\tau_n, \tau_{\mathbf{a}}] \frac{\delta}{\delta \kappa_i^{\alpha}} (\gamma[\mathbf{a}] \delta[\tau_{\mathbf{a}}, \tau_{\mathbf{b}}] - \Psi_{\mathbf{a}}[\tau_{\mathbf{a}}, \tau_{\mathbf{b}}]) \\ &\quad \times t_{\mathbf{a},\mathbf{b}} \mathbf{g}_{\mathbf{b},\mathbf{x}}[\tau_{\mathbf{b}}, \tau_m] + \lambda t_{\mathbf{x},\mathbf{j}} \mathbf{g}_{\mathbf{j},\mathbf{a}}[\tau_n, \tau_{\mathbf{a}}] \\ &\quad \times \frac{\delta}{\delta \kappa_i^{\alpha}} (\chi_{\mathbf{a}}[\tau_{\mathbf{a}}, \tau_{\mathbf{b}}]) \mathbf{g}_{\mathbf{a},\mathbf{x}}[\tau_{\mathbf{b}}, \tau_m], \end{aligned} \quad (48)$$

where the RHS is also evaluated in the  $\mathcal{A} \rightarrow 0$  limit. We now substitute this into Eq. (21) (with  $s = 0$ ). The last two terms must vanish by assumption (where  $\mathbf{a}$  has taken the place of  $\mathbf{x}$ ). The first term contains two paths from  $i$  to  $f$ , both via  $\mathbf{x}$ . Hence, this term must vanish in the large  $d$  limit unless  $\mathbf{x} = i$  or  $\mathbf{x} = f$ . The former also vanishes since  $v_{i,i}^{\alpha} = 0$  while the latter must vanish due to the sum over  $i - f$  and the odd parity of  $v_{i,f}^{\alpha}$ . The same reasoning applies to the second term except that in this term the  $\mathbf{x} = i$  case vanishes by the odd parity of  $v_{i,f}^{\alpha}$ . Hence, we have shown that  $(t_{\mathbf{x},\mathbf{j}} - i \sum_{\alpha} \kappa_{\mathbf{j}}^{\alpha}(\tau_n) v_{\mathbf{x},\mathbf{j}}^{\alpha}) \mathbf{g}_{\mathbf{j},\mathbf{x}}[\tau_n, \tau_m]$  satisfies Eq. (21) with  $s = 0$ . A completely analogous argument shows that this is also the case for  $\mathbf{g}_{\mathbf{x},\mathbf{x}}[\tau_n, \tau_m]$ . Using Eqs. (26) and (27) (in particular the fact that the RHS is made up of the same objects as the LHS), the above argument can be used to show that the result holds for any value of  $s$ . Thus, we have demonstrated the self-consistency of our ansatz (Eq. (21)).

Substituting Eq. (47) into Eq. (44), and using Eq. (21), we find that

$$\sum_{i-f} \Pi_{\alpha\beta}[i, f] = \sum_{i-f} \text{Tr} (\mathcal{G}[f, \mathbf{k}] v^{\alpha}[\mathbf{k}, i] \mathcal{G}[i, \mathbf{j}] v^{\beta}[\mathbf{j}, f^+])_{\mathcal{A} \rightarrow 0}. \quad (49)$$

Substituting this equation into Eq. (41), the optical conductivity may be expressed as

$$\sigma^{\alpha\beta}(\omega) = \frac{2}{i\omega} \sum_{\vec{p}, i\omega_p} \mathcal{G}(\vec{p}, i\omega_p) v_p^{\alpha} v_p^{\beta} [\mathcal{G}(\vec{p}, \omega + i\eta + i\omega_p) - \mathcal{G}(\vec{p}, i\eta + i\omega_p)]. \quad (50)$$

We now want to prove that this result holds to each order in  $\lambda$ . We do this via an inductive argument, in which we assume that through  $n$ th order in  $\lambda$ ,  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} L_{\mathbf{x}} \right)_{\mathcal{A} \rightarrow 0}^{(n)}$  (where  $L_i$  can be  $\Psi_i$ ,  $\chi_i$ , or  $\gamma[i]$ ) satisfies a certain explicit form (Eq. (51)), and then show that this form holds for  $n+1$ st order. We then plug Eq. (47) into  $\sum_{i-f} \Pi_{\alpha\beta}[i, f]$  (Eq. (44)), and use the explicit form of  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} L_{\mathbf{x}} \right)_{\mathcal{A} \rightarrow 0}^{(n)}$  to simplify the resulting expressions, thereby proving Eqs. (49) and (50) to each order in  $\lambda$ .

For the reason given below (Eq. (46)), we are free to use any of the results from Section 3.1, after making the substitution in Eq. (45). We define  $X_i$  to be a product of local functions of the type in Eq. (28) (i.e.  $X_i = (R_i)^m$ ) and  $Y_{i,f}$  to be a proxy for either  $\mathbf{g}_{i,f}[\tau_n, \tau_m]$  or  $t_{i,\mathbf{j}} \mathbf{g}_{\mathbf{j},f}[\tau_n, \tau_m]$  where the time indices are again arbitrary. *Inductive hypothesis:* through  $n$ th order in  $\lambda$ , the  $\kappa$  source derivative of the objects  $\Psi_i$ ,  $\chi_i$ , and  $\gamma[i]$  (denoted below by the generic symbol  $L_i$ ) can be written as

$$\begin{aligned} \left( \frac{\delta}{\delta \kappa_i^{\alpha}} L_{\mathbf{x}} \right)_{\mathcal{A} \rightarrow 0}^{(n)} &= \lambda^n X_{\mathbf{x}} Y_{\mathbf{x},\mathbf{x}_1} X_{\mathbf{x}_1} Y_{\mathbf{x}_1,\mathbf{x}_2} X_{\mathbf{x}_2} \cdots X_{\mathbf{x}_{m-1}} Y_{\mathbf{x}_{m-1},\mathbf{x}_m} X_{\mathbf{x}_m} \\ &\quad \times v_{\mathbf{x}_m,i}^{\alpha} Y_{i,\mathbf{x}_{m-1}} X_{\mathbf{x}_{m-1}} \cdots X_{\mathbf{x}_1} Y_{\mathbf{x}_1,\mathbf{x}} X_{\mathbf{x}}, \end{aligned} \quad (51)$$

where the number  $m$  is arbitrary. In the base case of zeroth order, the objects  $\Psi_i$ ,  $\chi_i$ , and  $\gamma[i]$  are

$$\begin{aligned} \Psi_i^{(0)}[\tau_i, \tau_m] &= 0; & \gamma^{(0)}[i] &= \mathbf{g}^{(k)}[i, i]; \\ \chi_i^{(0)}[\tau_i, \tau_m] &= - \left( t_{i,j} - i \sum_{\alpha} \kappa_j^{\alpha}(\tau_i) v_{ij}^{\alpha} \right) \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_i] \delta[\tau_i, \tau_m]. \end{aligned} \quad (52)$$

We note that  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} w_{x,j}(\tau_n) \mathbf{g}_{j,x}[\tau_n, \tau_m] \right)_{\mathcal{A} \rightarrow 0}^{(l)}$  is given by Eq. (48) with the appropriate objects on the RHS evaluated to the appropriate order in  $\lambda$ . An analogous formula holds for  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} \mathbf{g}_{x,x}[\tau_n, \tau_m] \right)_{\mathcal{A} \rightarrow 0}^{(l)}$ . Using these formulas with  $l = 0$  shows that the hypothesis is satisfied for the base case.

We now prove the inductive step. Eq. (28) continues to hold with  $t_{i,j} \rightarrow w_{i,j}(\tau_n)$  (the time index is again arbitrary). Therefore, using the notation  $\tilde{R}_i = [R_i]_{t_{i,j} \rightarrow w_{i,j}(\tau_n)}$ , we may write

$$\left( \frac{\delta}{\delta \kappa_i^{\alpha}} L_x \right)_{\mathcal{A} \rightarrow 0}^{(n+1)} = \sum_{r=0}^{n+1} \lambda^r \left( \frac{\delta}{\delta \kappa_i^{\alpha}} (\tilde{R}_x)^m \right)_{\mathcal{A} \rightarrow 0}^{(n+1-r)}. \quad (53)$$

Substituting the formulas for  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} w_{x,j}(\tau_n) \mathbf{g}_{j,x}[\tau_n, \tau_m] \right)_{\mathcal{A} \rightarrow 0}^{(l)}$  and  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} \mathbf{g}_{x,x}[\tau_n, \tau_m] \right)_{\mathcal{A} \rightarrow 0}^{(l)}$  (Eq. (48)) for  $l \leq n + 1$  into Eq. (53), and using the inductive hypothesis, shows that  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} \Psi_x \right)_{\mathcal{A} \rightarrow 0}^{(n+1)}$ ,  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} \chi_x \right)_{\mathcal{A} \rightarrow 0}^{(n+1)}$ , and  $\left( \frac{\delta}{\delta \kappa_i^{\alpha}} \gamma[x] \right)_{\mathcal{A} \rightarrow 0}^{(n+1)}$  all have the desired form (Eq. (51)). Thus, Eq. (51) holds to all orders in  $\lambda$ .

Substituting Eq. (47) into  $\sum_{i-f} \Pi_{\alpha\beta}[i, f]$  (Eq. (44)), and using Eq. (51), the only non vanishing terms are those which involve a derivative of the explicit factor  $(t_{x,y} - i \sum_{\alpha} \kappa_y^{\alpha} v_{x,y}^{\alpha})$  from Eq. (47). The other terms vanish due to the following reasoning. Upon substituting Eq. (51), in each of these terms there are two paths from  $i$  to  $f$ , both of which pass through the point  $\mathbf{x}$  as well as the points  $\mathbf{x}_1 \dots \mathbf{x}_{m-1}$  in Eq. (51). Hence, in the large  $d$  limit, all of these points must be chosen to be either  $i$  or  $f$  for these terms to be non vanishing. Then, if we choose  $\mathbf{x}_{m-1} = i$ , the term vanishes due to parity, while if we choose  $\mathbf{x}_{m-1} = f$ , the term vanishes due to parity combined with the sum  $\sum_{i-f}$ . Therefore, after making these simplifications, we find that Eq. (49) and consequently Eq. (50) hold to each order in  $\lambda$ .

### 3.4. $O(\lambda^2)$ theory in the limit of large dimensions

To obtain self-consistent integral equations to any order in  $\lambda$  for the objects  $\mathbf{g}^{-1}[i, f]$  and  $\mu[i, f]$ , we expand Eqs. (22) through (25) iteratively in  $\lambda$ , and set the sources to zero. Once the sources are set to zero, the system becomes translationally invariant in both space and time and we may express the equations in momentum/frequency space. Using the definitions

$$\mathbf{g}_{\text{loc},m}(i\omega_k) \equiv \sum_{\bar{k}} \mathbf{g}(k) \epsilon_{\bar{k}}^m, \quad (54)$$

$$I_{m_1 m_2 m_3}(i\omega_k) \equiv - \sum_{\omega_p, \omega_q} \mathbf{g}_{\text{loc},m_1}(i\omega_q) \mathbf{g}_{\text{loc},m_2}(i\omega_p) \mathbf{g}_{\text{loc},m_3}(i\omega_q + i\omega_p - i\omega_k), \quad (55)$$

the resulting equations to  $O(\lambda^2)$  are

$$a_{\mathcal{G}} \equiv 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4}, \quad (56)$$

$$\mathbf{g}^{-1}(k) = i\omega_k + \boldsymbol{\mu}' - a_{\mathcal{G}} \left( \epsilon_k - \frac{u_0}{2} \right) - \lambda \left( \epsilon_{\bar{k}} - \frac{u_0}{2} \right) \Psi(i\omega_k) - \lambda \chi(i\omega_k), \quad (57)$$

$$\mu(i\omega_k) = a_{\mathcal{G}} + \lambda \Psi(i\omega_k), \quad (58)$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} - u_0 \left( \lambda \frac{n}{2} - \lambda^2 \frac{n^2}{8} \right) + \lambda \sum_p \varepsilon_p \mathbf{g}(p) - a_g \frac{u_0}{2}, \quad (59)$$

$$\Psi(i\omega_k) = -\lambda u_0 I_{000}(i\omega_k) + 2\lambda I_{010}(i\omega_k), \quad (60)$$

$$\chi(i\omega_k) = -\frac{u_0}{2} \Psi(i\omega_k) - u_0 \lambda I_{001}(i\omega_k) + 2\lambda I_{011}(i\omega_k). \quad (61)$$

Before solving the equations, one must set  $\lambda = 1$ . The two Lagrange multipliers  $\boldsymbol{\mu}$  and  $u_0$  are determined by the two sum rules:

$$\sum_k \mathbf{g}(k) = \frac{n}{2}; \quad \sum_k g(k) = \frac{n}{2}. \quad (62)$$

The objects  $\mathbf{g}_{\text{loc},m}(i\omega_k)$  are given by an appropriate integral over the non-interacting density of states of a function composed of the two self energies  $\chi(i\omega_k)$  and  $\Psi(i\omega_k)$  and the energy  $\epsilon$  (Eq. (57)). Therefore, these constitute a self-consistent set of equations for the two self energies. These equations have been solved numerically and compared to DMFT calculations in Ref. [23].

#### 4. Anderson model

A word is needed at this point on the notation used, since similar looking symbols represent quite different objects in the  $t$ - $J$  model and the AIM. We use the functions  $g(\{\tau_j\})$ ,  $\mathbf{g}(\{\tau_j\})$ ,  $\mu(\{\tau_j\})$  or  $g(\{i\omega_j\})$ ,  $\mathbf{g}(\{i\omega_j\})$ ,  $\mu(\{i\omega_j\})$  and the related vertex functions for the impurity site of the AIM as well, but distinguish them from the  $t$ - $J$  model variables by dropping the spatial or momentum labels. Therefore in an equation such as Eq. (88), the object on the left (right) hand side corresponds to the  $t$ - $J$  model (AIM).

##### 4.1. Equations of motion for the Anderson model

In DMFT [8,9], the local Green's function of the infinite-dimensional finite- $U$  Hubbard model is mapped onto the impurity Green's function of the finite- $U$  AIM, with a self-consistently determined set of parameters. Using the ECFL equations of motion for both models, we show that the same mapping can be made between the infinite-dimensional  $t$ - $J$  model and the infinite- $U$  AIM. Further, we show that this mapping also extends to the auxiliary Green's function  $\mathbf{g}$ , and the caparison factor  $\mu$  individually. In this section, we briefly review the ECFL theory of the AIM [22], and we establish the mapping in the following section.

Consider the AIM in the limit  $U \rightarrow \infty$  which has the following Hamiltonian:

$$H = \sum_{\sigma} \epsilon_d X^{\sigma\sigma} + \sum_{k\sigma} \tilde{\epsilon}_k n_{k\sigma} + \sum_{k\sigma} (V_k X^{\sigma 0} c_{k\sigma} + V_k^* c_{k\sigma}^{\dagger} X^{0\sigma}), \quad (63)$$

where we have set the Fermi energy of the conduction electrons to be zero. The impurity Green's function is given by the following expression:

$$g_{\sigma_i\sigma_f}[\tau_i, \tau_f] = -\langle\langle X^{0\sigma_i}(\tau_i) X^{\sigma_f 0}(\tau_f) \rangle\rangle. \quad (64)$$

The ECFL solution of the Anderson model is presented in Ref. [22]. The impurity Green's function is factored into the auxiliary Green's function and the caparison factor:

$$g[\tau_i, \tau_f] = \mathbf{g}[\tau_i, \tau_j] \cdot \mu[\tau_j, \tau_f]. \quad (65)$$

The equations of motion for  $\mathbf{g}$  and  $\mu$  can be written as

$$(\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i))\mathbf{g}[\tau_i, \tau_f] = -\delta(\tau_i - \tau_f) - (1 - \lambda\gamma[\tau_i]) \cdot \Delta[\tau_i, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_f] \\ - \lambda \xi^* \Delta[\tau_i, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_{\mathbf{x}}] \cdot \Lambda_*[\tau_{\mathbf{x}}, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}[\tau_{\mathbf{y}}, \tau_f], \quad (66)$$

$$\mu[\tau_i, \tau_f] = \delta(\tau_i - \tau_f)(\mathbb{1} - \lambda\gamma[\tau_i]) + \lambda \xi^* \cdot \Delta[\tau_i, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_{\mathbf{x}}] \cdot \mathcal{U}_*[\tau_{\mathbf{x}}, \tau_f; \tau_i], \quad (67)$$

where the conduction band enters through the ( $\mathcal{V}$  independent) function

$$\Delta[\tau_i, \tau_f] = -\mathbb{1} \sum_k |V_k|^2 (\partial_{\tau_i} + \tilde{\epsilon}_k)^{-1} \delta(\tau_i - \tau_f). \quad (68)$$

We have also made use of the following definitions:

$$\begin{aligned} \Lambda[\tau_n, \tau_m; \tau_i] &= -\frac{\delta}{\delta \mathcal{V}(\tau_i)} \mathbf{g}^{-1}[\tau_n, \tau_m]; & \mathcal{U}[\tau_n, \tau_m; \tau_i] &= \frac{\delta}{\delta \mathcal{V}(\tau_i)} \mu[\tau_n, \tau_m]; \\ \gamma[\tau_i] &= \mu^{(k)}[\tau_n, \tau_i^+] \cdot \mathbf{g}^{(k)}[\tau_i, \tau_n]. \end{aligned} \quad (69)$$

#### 4.2. Mapping of the $t$ - $J$ model onto the Anderson model in infinite dimensions

Now let us consider the  $t$ - $J$  model in the limit of infinite dimensions. Inverting Eq. (15), the equations of motion for  $\mathbf{g}_{i,i}[\tau_i, \tau_f]$  and  $\mu_{i,i}[\tau_i, \tau_f]$  are

$$\begin{aligned} (\partial_{\tau_i} - \mu + \mathcal{V}_i(\tau_i)) \mathbf{g}_{i,i}[\tau_i, \tau_f] &= -\delta(\tau_i - \tau_f) + (1 - \lambda \gamma[i]) \cdot t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_f] \\ &\quad + \lambda t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_x] \cdot A_{i,*}[\tau_x, \tau_y; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_y, \tau_f] \\ &\quad + \lambda t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_x] \cdot B_{i,*}[\tau_x, \tau_y; \tau_i] \cdot t_{i,y} \mathbf{g}_{y,i}[\tau_y, \tau_f], \end{aligned} \quad (70)$$

$$\mu_{i,i}[\tau_i, \tau_f] = (1 - \lambda \gamma[i]) \delta(\tau_i - \tau_f) + \lambda t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_x] \cdot B_{i,*}[\tau_x, \tau_f; \tau_i]. \quad (71)$$

By mapping  $\mathbf{g}_{i,i}[\tau_i, \tau_f]$  and  $\mu_{i,i}[\tau_i, \tau_f]$  onto  $\mathbf{g}[\tau_i, \tau_f]$  and  $\mu[\tau_i, \tau_f]$  of the AIM, we would like to show that the equations of motion of the AIM (Eqs. (66) and (67)) and those of the infinite dimensional  $t$ - $J$  model (Eqs. (70) and (71)) map onto each other. To do this, we need the analog of the object  $\mathbf{g}^{-1}[\tau_i, \tau_f]$  of the AIM in the  $t$ - $J$  model. We denote this new object by  $\mathbf{g}_{\text{loc},i}^{-1}[\tau_i, \tau_f]$  and define it to be the temporal inverse of the local auxiliary Green's function:

$$\mathbf{g}_{i,i}[\tau_i, \tau_j] \cdot \mathbf{g}_{\text{loc},i}^{-1}[\tau_j, \tau_f] = \delta(\tau_i - \tau_f). \quad (72)$$

Note that  $\mathbf{g}_{\text{loc},i}^{-1}[\tau_i, \tau_f] \neq \mathbf{g}_{i,i}^{-1}[\tau_i, \tau_f]$ . We also define the corresponding vertex:

$$\Lambda_{\text{loc},i}[\tau_n, \tau_m; \tau_i] = -\frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{\text{loc},i}^{-1}[\tau_n, \tau_m]. \quad (73)$$

We now make use of the following identity:

$$\Lambda_{\text{loc},i}[\tau_x, \tau_y; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_y, \tau_f] = A_i[\tau_x, \tau_y; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_y, \tau_f] + B_i[\tau_x, \tau_y; \tau_i] \cdot t_{i,y} \mathbf{g}_{y,i}[\tau_y, \tau_f]. \quad (74)$$

This identity is easily proven by considering  $\frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{i,i}[\tau_x, \tau_f]$ :

$$\frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{i,i}[\tau_x, \tau_f] = \mathbf{g}_{i,i}[\tau_x, \tau_j] \Lambda_{\text{loc},i}[\tau_j, \tau_y; \tau_i] \mathbf{g}_{i,i}[\tau_y, \tau_f]. \quad (75)$$

The LHS can also be expressed as

$$\begin{aligned} \frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{i,i}[\tau_x, \tau_f] &= \mathbf{g}_{i,i}[\tau_x, \tau_j] \cdot (A_i[\tau_j, \tau_y; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_y, \tau_f] \\ &\quad + B_i[\tau_j, \tau_y; \tau_i] \cdot t_{i,y} \mathbf{g}_{y,i}[\tau_y, \tau_f]). \end{aligned} \quad (76)$$

Left multiplying the above 2 equations by  $\mathbf{g}_{\text{loc},i}^{-1}$ , we recover the identity Eq. (74). Substituting this identity into Eq. (70), we obtain

$$\begin{aligned} (\partial_{\tau_i} - \mu + \mathcal{V}_i(\tau_i)) \mathbf{g}_{i,i}[\tau_i, \tau_f] &= -\delta(\tau_i - \tau_f) + (1 - \lambda \gamma[i]) \cdot t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_f] \\ &\quad + \lambda t_{i,j} \xi^* \cdot \mathbf{g}_{j,i}[\tau_i, \tau_x] \cdot \Lambda_{\text{loc},i*}[\tau_x, \tau_y; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_y, \tau_f]. \end{aligned} \quad (77)$$

We are now ready to map the  $t$ - $J$  model onto the Anderson model. To do this, we map the local objects  $\mathbf{g}_{i,i}[\tau_i, \tau_f]$  and  $\mu_{i,i}[\tau_i, \tau_f]$  of the  $t$ - $J$  model to the objects  $\mathbf{g}[\tau_i, \tau_f]$  and  $\mu[\tau_i, \tau_f]$  of the Anderson model. We also map  $\mu$  to  $-\epsilon_d$ . The following mappings also follow as a consequence of these.

$$\begin{aligned} \gamma[i] &\rightarrow \gamma[\tau_i]; & \Lambda_{\text{loc},i}[\tau_n, \tau_m; \tau_i] &\rightarrow \Lambda[\tau_n, \tau_m; \tau_i]; \\ B_i[\tau_n, \tau_m; \tau_i] &\rightarrow -\mathcal{U}[\tau_n, \tau_m; \tau_i]. \end{aligned} \quad (78)$$

Comparing Eq. (77) with Eq. (66) and Eq. (71) with Eq. (67), we see that the equations of motion map onto each other if the following constraint is satisfied:

$$t_{i,j} \mathbf{g}_{j,i}[\tau_i, \tau_f] = -\Delta[\tau_i, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_f]. \quad (79)$$

### 4.3. Mapping to each order in $\lambda$

The  $O(\lambda^2)$  equations for the infinite-dimensional  $t$ - $J$  model and infinite- $U$  AIM are solved numerically in Ref. [23] and Ref. [22] respectively. This can in principle be done to higher orders in  $\lambda$  as well, and it is therefore interesting to know if the mapping from the previous section holds to each order in  $\lambda$ . We show that it does, and give a simple prescription for obtaining the ECFL integral equations for one model from those of the other one (Eq. (83)).

We review the  $\lambda$  expansion for the Anderson model from Ref. [22]. There, Eqs. (66) and (67) are written as

$$\begin{aligned} \mathbf{g}^{-1}[\tau_i, \tau_f] &= -(\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i))\delta(\tau_i - \tau_f) - (1 - \lambda\gamma[\tau_i]) \cdot \Delta[\tau_i, \tau_f] \\ &\quad - \lambda\xi^* \Delta[\tau_i, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_x] \cdot \Lambda_*[\tau_x, \tau_f; \tau_i], \end{aligned} \quad (80)$$

$$\mu[\tau_i, \tau_f] = \delta(\tau_i - \tau_f)(\mathbb{1} - \lambda\gamma[\tau_i]) + \lambda\xi^* \cdot \Delta[\tau_i, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_x] \cdot \mathcal{U}_*[\tau_x, \tau_f; \tau_i]. \quad (81)$$

The  $\lambda$  expansion is obtained in the same way as for the  $t$ - $J$  model, by iterating the equations in  $\mathbf{g}^{-1}$  and  $\mu$  and keeping track of explicit powers of  $\lambda$ . The details to  $O(\lambda^2)$  can be found in Ref. [22]. To relate this to the  $\lambda$  expansion for the infinite-dimensional  $t$ - $J$  model, recall from Eq. (28) that to each order in  $\lambda$ ,  $\Psi_i$ ,  $\chi_i$ ,  $\gamma[i]$ ,  $A_i$ , and  $B_i$  can be written as a product of the functions  $\mathbf{g}_{i,i}[\tau_n, \tau_m]$  and  $t_{i,j}\mathbf{g}_{j,i}[\tau_n, \tau_m]$ . We can now state our *inductive hypothesis*: through  $n$ th order in  $\lambda$ , the  $\lambda$  expansion for the Anderson model has the form

$$\begin{aligned} \mathbf{g}^{-1}[\tau_i, \tau_m] &= -(\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i))\delta[\tau_i, \tau_m] - \lambda\chi[\tau_i, \tau_m] \\ &\quad - (1 - \lambda\gamma[\tau_i])\Delta[\tau_i, \tau_m] - \lambda\Psi[\tau_i, \tau_j]\Delta[\tau_j, \tau_m], \\ \mu[\tau_i, \tau_m] &= \delta[\tau_i, \tau_m](1 - \lambda\gamma[\tau_i]) + \lambda\Psi[\tau_i, \tau_m], \\ \Lambda[\tau_n, \tau_m; \tau_i] &= A[\tau_n, \tau_m; \tau_i] - B[\tau_n, \tau_j; \tau_i]\Delta[\tau_j, \tau_m], \\ \mathcal{U}[\tau_n, \tau_m; \tau_i] &= -B[\tau_n, \tau_m; \tau_i], \end{aligned} \quad (82)$$

where through  $n$ th order in  $\lambda$ , the objects  $A[\tau_n, \tau_m; \tau_i]$  and  $B[\tau_n, \tau_m; \tau_i]$ , and through  $n-1$ st order in  $\lambda$ , the objects  $\gamma[\tau_i]$ ,  $\chi[\tau_i, \tau_m]$ , and  $\Psi[\tau_i, \tau_m]$ , can be obtained from their infinite dimensional  $t$ - $J$  model counterparts via the substitution

$$\mathbf{g}_{i,i}[\tau_n, \tau_m] \rightarrow \mathbf{g}[\tau_n, \tau_m]; \quad \mu \rightarrow -\epsilon_d; \quad t_{i,j}\mathbf{g}_{j,i}[\tau_n, \tau_m] \rightarrow -\Delta[\tau_n, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_m]. \quad (83)$$

We first examine the base case of zeroth order:

$$A^{(0)}[\tau_n, \tau_m; \tau_i] = \delta[\tau_i, \tau_n]\delta[\tau_i, \tau_m]; \quad B^{(0)}[\tau_n, \tau_m; \tau_i] = 0. \quad (84)$$

Comparing with Eq. (29), the hypothesis clearly holds. We now prove the inductive step. Eq. (69) together with Eqs. (80) through (82) implies the following:

$$\begin{aligned} \chi^{(n)}[\tau_n, \tau_m] &= \xi^* \Delta[\tau_n, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_x] \cdot A_*^{(n)}[\tau_x, \tau_m; \tau_n], \\ \Psi^{(n)}[\tau_n, \tau_m] &= -\xi^* \Delta[\tau_n, \tau_j] \cdot \mathbf{g}[\tau_j, \tau_x] \cdot B_*^{(n)}[\tau_x, \tau_m; \tau_n], \\ A^{(n+1)}[\tau_n, \tau_m; \tau_i] &= \lambda \left( \frac{\delta}{\delta\mathcal{V}(\tau_i)} \chi[\tau_n, \tau_m] \right)^{(n)}, \end{aligned}$$

$$B^{(n+1)}[\tau_n, \tau_m; \tau_i] = \lambda \delta[\tau_n, \tau_m] \left( \frac{\delta}{\delta \mathcal{V}(\tau_i)} \gamma[\tau_n] \right)^{(n)} - \lambda \left( \frac{\delta}{\delta \mathcal{V}(\tau_i)} \Psi[\tau_n, \tau_m] \right)^{(n)},$$

$$\gamma^{(n)}[\tau_i] = -\lambda \gamma^{(k)(n-1)}[\tau_i] \mathbf{g}^{(k)}[\tau_i, \tau_i] + \lambda \Psi^{(k)(n-1)}[\tau_i, \tau_i] \mathbf{g}^{(k)}[\tau_i, \tau_i]. \quad (85)$$

Comparing with Eq. (30), we see that  $\chi^{(n)}[\tau_n, \tau_m]$ ,  $\Psi^{(n)}[\tau_n, \tau_m]$ , and  $\gamma^{(n)}[\tau_i]$  have the desired form. We also note that

$$\left( \frac{\delta}{\delta \mathcal{V}(\tau_r)} \mathbf{g}[\tau_i, \tau_m] \right)^{(l)} = \mathbf{g}[\tau_i, \tau_x] (A^{(l)}[\tau_x, \tau_y; \tau_r] - B^{(l)}[\tau_x, \tau_j; \tau_r] \Delta[\tau_j, \tau_y]) \mathbf{g}[\tau_y, \tau_m]. \quad (86)$$

Comparing this with Eqs. (26) and (27), we see that by the inductive hypothesis, the mapping Eq. (83) continues to hold through order  $l \leq n$  even after both sides have been acted on with a functional derivative. Furthermore, in evaluating  $A^{(n+1)}[\tau_n, \tau_m; \tau_i]$  and  $B^{(n+1)}[\tau_n, \tau_m; \tau_i]$  using Eq. (85), we will at most need to set  $l = n$  in Eq. (86). Finally, comparing Eq. (85) with Eq. (31), we see that  $A^{(n+1)}[\tau_n, \tau_m; \tau_i]$  and  $B^{(n+1)}[\tau_n, \tau_m; \tau_i]$  have the desired form. Thus, we have proven our inductive hypothesis.

Setting the sources to zero, and Fourier transforming Eq. (82), we may write ( $\lambda \rightarrow 1$ ,  $\gamma[\tau_i] \rightarrow \frac{n_d}{2} \equiv \frac{n}{2}$ )

$$\mathbf{g}^{-1}(i\omega_k) = i\omega_k - \epsilon_d - \left(1 - \frac{n}{2}\right) \Delta(i\omega_k) - \chi(i\omega_k) - \Delta(i\omega_k) \Psi(i\omega_k),$$

$$\mu(i\omega_k) = 1 - \frac{n}{2} + \Psi(i\omega_k). \quad (87)$$

Comparing with Eq. (35), it immediately follows that under the mapping Eq. (83),  $\mu_{i,i}(i\omega_k) \rightarrow \mu(i\omega_k)$ . Furthermore, multiplying both sides of the equation for  $\mathbf{g}^{-1}(k)$  by  $\mathbf{g}(k)$ , summing over  $\vec{k}$ , and using the mapping Eq. (83), it follows that  $\mathbf{g}_{i,i}(i\omega_k) \rightarrow \mathbf{g}(i\omega_k)$ . Therefore, the ECFL solution of the infinite dimensional  $t$ - $J$  model maps onto the ECFL solution of the AIM to each order in  $\lambda$  as long as the following self-consistency condition is satisfied:

$$\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) = \sum_{\vec{k}} \frac{|V_{\vec{k}}|^2}{i\omega_n - \tilde{\epsilon}_{\vec{k}}} \mathbf{g}(i\omega_k). \quad (88)$$

This mapping and self-consistency condition can be understood by referring back to DMFT. In DMFT [9], the physical Green's function  $\mathcal{G}_{i,f}(i\omega_k)$  is determined for any separation of  $i$  and  $f$  by the local Green's function  $G_{i,i}(i\omega_k)$  or equivalently the local self energy  $\Sigma(i\omega_k)$ . The impurity Green's function of the Anderson model  $\mathcal{G}(i\omega_k)$  can be set equal to  $G_{i,i}(i\omega_k)$  as long as  $\tilde{\epsilon}_k$  and  $V_k$  satisfy a self-consistency condition relating them to  $\mathcal{G}(i\omega_k)$  (see Eqs. (13) and (15) of Ref. [9]). In the ECFL mapping, the auxiliary Green's function  $\mathbf{g}_{i,f}(i\omega_k)$  is determined for any separation of  $i$  and  $f$  by the local auxiliary Green's function  $\mathbf{g}_{i,i}(i\omega_k)$  and by the local comparison factor  $\mu_{i,i}(i\omega_k)$ , or equivalently by the two local self energies  $\Psi(i\omega_k)$  and  $\chi(i\omega_k)$ .  $\mu_{i,f}(i\omega_k)$  is itself local and related simply to  $\Psi(i\omega_k)$ . The impurity auxiliary Green's function of the Anderson model  $\mathbf{g}(i\omega_k)$  can be set equal to  $\mathbf{g}_{i,i}(i\omega_k)$  and the comparison factor of the Anderson model  $\mu(i\omega_k)$  set equal to  $\mu_{i,i}(i\omega_k)$  as long as  $\tilde{\epsilon}_k$  and  $V_k$  satisfy the self-consistency condition (Eq. (88)). We now show that Eq. (88) can be put into the form of Eqs. (13) and (15) of Ref. [9]. Using Eq. (35) the LHS can be written as

$$\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) = \frac{-1}{1 - \frac{n}{2} + \Psi(i\omega_k)} [1 - (i\omega_k + \mu - \chi(i\omega_k)) \mathbf{g}(i\omega_k)]. \quad (89)$$

Using Eqs. (2), (5), (35), (38) and the relation  $\mathcal{G}(i\omega_k) = \mathbf{g}(i\omega_k) \cdot \mu(i\omega_k)$ , the above equation becomes

$$\Sigma_D(i\omega_k) + \frac{1}{\mathcal{G}(i\omega_k)} - (i\omega_k + \mu) = - \sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) \frac{1}{\mathbf{g}(i\omega_k)}. \quad (90)$$

Substituting Eq. (88) into the RHS of the above equation, we recover Eqs. (13) and (15) from Ref. [9].

## 5. Conclusion

In this work we provide a detailed analysis of the simplifications arising from the large dimensionality limit of the  $t$ - $J$  model, and have given the first few terms in the  $\lambda$  series that leads to practically usable results. It is clear that the formal result of a local Dysonian self energy is already implied by the large  $d$  results for the Hubbard model reviewed in Ref. [9], if we take the limit of infinite  $U$ ; that is indeed another description of the model studied here. However it must be kept in mind that the present calculation starts with the infinite  $U$  limit already taken, and thus provides a non trivial check on the uniqueness of the limit of  $U \rightarrow \infty$  and  $d \rightarrow \infty$ , i.e. its independence on the order of these two limits. Also the present work uses the novel ECFL methodology that rests on a different set of tools from the ones usually used to study the Hubbard model and its large dimensional limit. We use the Schwinger equations of motion, as opposed to the usual Feynman–Wick theory, and we have obtained analytical results that do not rely on Wick’s theorem.

Summarizing, we have considered the ECFL theory for the  $t$ - $J$  model ( $J = 0$ ) by establishing the simplifications that arise in the equations of motion in the limit of large dimensions. The auxiliary Green’s function  $\mathbf{g}(k)$  and the caparison factor  $\mu(k)$  can be written in terms of two local self energies  $\Psi(i\omega_k)$  and  $\chi(i\omega_k)$  as in Eq. (35). This insight into the structural form of the physical Green’s function  $\mathcal{G}(k)$  has been used in a concurrent publication (Ref. [23]), to benchmark and compare the ECFL and DMFT calculations. The ECFL integral equations in the large  $d$  limit, derived here to  $O(\lambda^2)$ , have been solved numerically in Ref. [23], and their solution compares favorably with DMFT results. It can be seen explicitly from these equations that Eq. (35) holds to second order in  $\lambda$ , with  $\Psi(i\omega_k)$  and  $\chi(i\omega_k)$  written as a product of the functions  $\mathbf{g}_{\text{loc},m}(i\omega_k)$  (Eq. (54)) with  $m = 0$  or  $m = 1$ . This continues to hold to each order in  $\lambda$ . We have analyzed the optical conductivity and have shown that it is given by Eq. (50) in general and to each order in  $\lambda$ . We have separately also studied the ECFL theory of the infinite- $U$  AIM[22], and have shown that there is a mapping between the ECFL of the infinite dimensional  $t$ - $J$  model and the ECFL of the AIM with a self-consistently determined set of parameters (Eq. (88)). This mapping holds to each order in  $\lambda$  and there is a simple prescription for obtaining the ECFL integral equations for one model from those of the other (Eq. (83)).

In conclusion this work provides a solid foundation for the study of the  $t$ - $J$  model, and in particular for the ECFL formalism, in the limit of infinite dimensions, by providing exact statements about the  $k$  dependence of self energies, the absence of vertex corrections in computing the conductivity and finally in yielding a systematic expansion in the parameter  $\lambda$  that enables a quantitative comparison with other methods as in Ref. [23].

## Acknowledgments

We acknowledge stimulating discussions with A. Georges, H. Krishnamurthy and D. Vollhardt. We also thank A. Garg and J. Otsuki for helpful comments. This work was supported by DOE under grant no. FG02-06ER46319.

## References

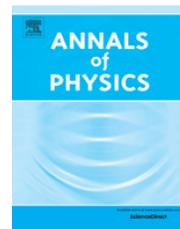
- [1] L.D. Landau, *Sov. Phys.—JETP* 3 (1957) 920.
- [2] L.D. Landau, *Sov. Phys.—JETP* 5 (1957) 101.
- [3] Y. Kuramoto, T. Watanabe, *Physica B* 148 (1987) 80.
- [4] W. Metzner, D. Vollhardt, *Phys. Rev. Lett.* 62 (1989) 324.
- [5] E. Müller-Hartmann, *Z. Phys. B* 74 (1989) 507.
- [6] W. Metzner, *Phys. Rev. B* 43 (1991) 8549.
- [7] F.J. Ohkawa, *J. Phys. Soc. Japan* 60 (1991) 3218;  
F.J. Ohkawa, *Progr. Theoret. Phys. Suppl.* 106 (1991) 95.
- [8] A. Georges, G. Kotliar, *Phys. Rev. B* 45 (1992) 6479.
- [9] A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, *Rev. Modern Phys.* 68 (1) (1996) 13–125.
- [10] A. Khurana, *Phys. Rev. Lett.* 64 (1990).
- [11] V. Zlatić, B. Horvatić, *Solid State Commun.* 75 (1990) 263.
- [12] X. Deng, J. Mravlje, R. Zitko, M. Ferrero, G. Kotliar, A. Georges, *Phys. Rev. Lett.* 110 (2013) 086401.
- [13] J. Otsuki, D. Vollhardt, *Phys. Rev. Lett.* 110 (2013) 196407.

- [14] P.W. Anderson, *Science* 235 (1987) 1196;  
The Theory of Superconductivity, Princeton University Press, Princeton, NJ, 1997;  
unpublished. arXiv:0709.0656.
- [15] B.S. Shastry, *Phys. Rev. B* 81 (2010) 045121.
- [16] B.S. Shastry, *Phys. Rev. B* 84 (2011) 165112.
- [17] B.S. Shastry, *Phys. Rev. Lett.* 107 (2011) 056403.
- [18] B.S. Shastry, *Phys. Rev. B* 87 (2013) 125124.
- [19] G.-H. Gweon, B.S. Shastry, G.D. Gu, *Phys. Rev. Lett.* 107 (2011) 056404.
- [20] K. Matsuyama, G.-H. Gweon, 2013. arXiv:1212.0299.
- [21] E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, B.S. Shastry, *Phys. Rev. B* 87 (2013) 161120.
- [22] B.S. Shastry, E. Perepelitsky, A.C. Hewson, 2013. arXiv:1307.3492;  
D. Hansen, B.S. Shastry, *Phys. Rev. B* 87 (2013) 245101.
- [23] R. Zitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, B.S. Shastry, 2013, arXiv:1309.5284, in press.
- [24] B.S. Shastry, *Phys. Rev. Lett.* 108 (2012) 029702.



Contents lists available at ScienceDirect

## Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)

# Theory of extreme correlations using canonical Fermions and path integrals

**B. Sriram Shastry***Physics Department, University of California, Santa Cruz, CA 95064, United States***H I G H L I G H T S**

- Spectral function of the Extremely Correlated Fermi Liquid theory at low energy.
- Electronic origin of low energy kinks in energy dispersion.
- Non Hermitian representation of Gutzwiller projected electrons.
- Analogy with Dyson–Maleev representation of spins.
- Path integral formulation of extremely correlated electrons.

**A R T I C L E I N F O***Article history:*

Received 19 December 2013

Accepted 5 February 2014

Available online 14 February 2014

*Keywords:**t*–*J* model

Gutzwiller projection

Energy dispersion

Angle resolved photoemission

Path integrals

Schwinger equations of motion

**A B S T R A C T**

The *t*–*J* model is studied using a novel and rigorous mapping of the Gutzwiller projected electrons, in terms of canonical electrons. The mapping has considerable similarity to the Dyson–Maleev transformation relating spin operators to canonical Bosons. This representation gives rise to a non Hermitian quantum theory, characterized by minimal redundancies. A path integral representation of the canonical theory is given. Using it, the salient results of the extremely correlated Fermi liquid (ECFL) theory, including the previously found Schwinger equations of motion, are easily re-derived. Further, a transparent physical interpretation of the previously introduced auxiliary Greens function and the ‘caparison factor’, is obtained.

The low energy electron spectral function in this theory, with a strong intrinsic asymmetry, is summarized in terms of a few expansion coefficients. These include an important emergent energy scale  $\Delta_0$  that shrinks to zero on approaching the insulating state, thereby making it difficult to access the underlying very low energy Fermi liquid behavior. The scaled low frequency ECFL spectral function, related simply to the Fano line shape, has a peculiar energy dependence unlike that of a Lorentzian. The resulting energy

*E-mail addresses:* [sriram@physics.ucsc.edu](mailto:sriram@physics.ucsc.edu), [b.sriram.shastry@gmail.com](mailto:b.sriram.shastry@gmail.com).<http://dx.doi.org/10.1016/j.aop.2014.02.005>

0003-4916/© 2014 Elsevier Inc. All rights reserved.

dispersion obtained by maximization is a hybrid of a massive and a massless Dirac spectrum  $E_Q^* \sim \gamma Q - \sqrt{\Gamma_0^2 + Q^2}$ , where the vanishing of  $Q$ , a momentum type variable, locates the kink minimum. Therefore the quasiparticle velocity interpolates between  $(\gamma \mp 1)$  over a width  $\Gamma_0$  on the two sides of  $Q = 0$ , implying a kink there that strongly resembles a prominent low energy feature seen in angle resolved photoemission spectra (ARPES) of cuprate materials. We also propose novel ways of analyzing the ARPES data to isolate the predicted asymmetry between particle and hole excitations.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

The intensely studied  $t$ - $J$  model is often regarded as the effective low energy Hamiltonian for describing several observed phenomena in cuprate superconductors [1]. Here the  $U \rightarrow \infty$  limit is presupposed, and hence the Hilbert space is restricted to a maximum of single occupancy at each site, i.e. Gutzwiller projected [2]. A few words on the choice of the  $t$ - $J$  model are relevant here. The implied infinite  $U$  limit eliminates high energy ( $U$  scale) electronic states, known as the upper Hubbard band states. The residual low energy ( $\lesssim 100$  meV scale) excitations are probed by sensitive spectroscopies and transport phenomena, making the  $t$ - $J$  model suitable for our task. At reasonably high  $U$ , say comparable to the band width in a Hubbard model, this elimination of the upper Hubbard band must already occur in part. Therefore the limit  $U \rightarrow \infty$  must be regarded as a useful mathematical idealization of the very strong, or extreme correlation phenomenon. The resulting Gutzwiller projected electron operators, denoted by Hubbard's convenient notation of  $X$  operators [3], are rendered non canonical. The non-canonical nature of the electrons precludes the Wick's theorem underlying the Feynman diagram approach, whereby leading to the fundamental difficulty of the  $t$ - $J$  model, namely the impossibility of a straightforward Feynman type perturbative expansion. This situation leads to enormous calculational difficulties, so that systematic and controlled analytical calculations with this model have been very difficult.

In a series of recent papers [4,6–8,5,9,10], we have shown that it is possible to overcome some of these difficulties by using alternate methods based on Schwinger's treatment of field theory with time dependent potentials. This idea yields exact equations of motion for the electron Greens function. These equations have the nature of functional differential equations, and provide a powerful launching pad for various approximations. The specific approximation pursued is a systematic expansions in a parameter  $\lambda$  related to double occupancy. Using this we have presented an analytical theory of the normal state of the  $t$ - $J$  model termed the extremely correlated Fermi liquid (ECFL) theory. An interesting feature of the theory is that we find a non-Dysonian representation of the projected electron Greens function. This is a significant structural departure from the usual field theories, and arises in a most natural fashion. The Greens function is determined by a *pair* of self energies, denoted by  $\Phi(\vec{k}, i\omega_n)$  and  $\Psi(\vec{k}, i\omega_n)$ , instead of the standard Dyson self energy  $\Sigma(\vec{k}, i\omega_n)$  (see Eq. (21) below). The latter can be reconstructed from the pair by a simple inversion. Starting with rather simple pairs of self energies, it is found that non trivial complexity is introduced into the Dyson self energy through this inversion process. Explicit self consistent calculations in parameter  $\lambda$  have been carried out to  $O(\lambda^2)$  so far, and yield reliable results for electron densities  $0 \leq n \lesssim .7$ . The detailed dynamical results of the ECFL theory have been benchmarked against independent theories in overlapping domains; e.g. against high temperature series results in Ref. [11]. The ECFL theory has been shown to have a momentum independent Dyson self energy in the limit of infinite dimensions [10]. This enables benchmarking against the dynamical mean field theory (DMFT) in Ref. [9]. Importantly, the results from the ECFL theory for the spectral function compare well with a large  $U$  Hubbard model solved by the DMFT method, and not just infinite  $U$ . The ECFL theory has also been benchmarked in Ref.

[12] against the exact solution of the asymmetric  $U = \infty$  Anderson impurity model, obtained from the numerical renormalization group study of Krishnamurthy, Wilson and Wilkins [13]. In addition, a detailed comparison between the data on cuprate superconductors at optimal filling and the theoretical photoemission spectral lines of the ECFL theory has been carried out in Refs. [14,15], where excellent agreement is found. In all cases studied, the comparisons with ECFL are good, and seem to indicate the utility of this approach.

The ECFL formalism could initially seem somewhat unfamiliar, in view of its reliance on the analysis of the Schwinger equations of motion. This analysis was originally used to derive the main constituents of the theory, namely the auxiliary Greens function and the two self energies (detailed below). This type of analysis is somewhat removed from the toolkit of “standard” many body physics courses, and hence might obstruct a ready visualization of these objects. One goal of the present work is to show that these results are (A) minimal, i.e. having least redundancy, and (B) available more transparently. The latter follows from an important and novel *hat removal theorem*, leading to a compact mapping of the Hubbard operators to canonical Fermions. The mapping is given in Eq. (1) and described further in Section 3.2, leading to a path integral formulation (Section 7). It is possible that such a simplified presentation could lead to improved strategies for devising approximate methods, especially close to the insulating state.

This method rests on an exact replacement rule for the Hubbard  $X$  operators in terms of the canonical Fermi operators

$$X_i^{0\sigma} \rightarrow C_{i\sigma}, \quad X_i^{\sigma 0} \rightarrow C_{i\bar{\sigma}}^\dagger (1 - N_{i\bar{\sigma}}), \quad X_i^{\sigma\sigma'} \rightarrow C_{i\sigma}^\dagger C_{i\sigma'}. \quad (1)$$

This replacement rule is shown to be exact when “right-operating” on states which satisfy the Gutzwiller constraint. This replacement is similar in spirit to the Dyson–Maleev representation [16, 17], where spin operators are expressed in terms of canonical Bosonic operators. With the advantage of this representation, most steps in the ECFL theory, such as the factorization of the Greens function into an auxiliary Greens function, the two self energies and the caparison function (see Eqs. (18), (19) and (21)) becomes very intuitive.

The analogy can be pushed further to establish a parallel between the  $\lambda$  parameter of the ECFL theory, and the small parameter of the Dyson–Maleev [16,17] theory, namely the inverse spin  $\frac{1}{2s}$ . Finally we are able to make contact with the illuminating work of Harris, Kumar, Halperin and Hohenberg [18]. In a detailed work these authors computed the Greens function of the spins for two sublattice antiferromagnet using the Dyson–Maleev scheme and extracted the lifetime of the magnons of the theory. We find that their calculation contains the precise Bosonic counterparts of the auxiliary Greens function and the second self energy  $\Psi$  defining the “caparison function” of the ECFL theory (see Eqs. (18), (19) and (21)). Unlike the spin problem with variable number of excitations, the  $t$ – $J$  model has a fixed number of particles. Hence there are significant new elements in the ECFL theory involving the imposition of the Luttinger–Ward volume theorem, as discussed later.

A few comments on the canonical description of the equations of motion are appropriate. The general problem is to represent a time evolution of a state of the  $t$ – $J$  model

$$[\psi]_{final}' = Q_M' \dots Q_2' \cdot Q_1' \cdot [\psi]_{initial}', \quad (2)$$

where the primed states and operators are in the  $t$ – $J$  model Hilbert space defined with the three allowed states at each site as usual (see Section 2.1 for details). The operators  $Q_j'$  may be thought of as the exponential of the  $t$ – $J$  Hamiltonian:  $Q_j' \sim e^{-it_j H_{tj}}$  written in terms of the projected operators. Since the algebra of the projected electrons is very inconvenient, one seeks a reframing of the problem into a canonical space. This involves mapping the states, the Hamiltonian and all other operators of the original theory, into the unconstrained Hilbert space of two Fermions at each site. This canonical space is of course described by the usual Fermi operators  $C_{j\uparrow}, C_{j\downarrow}$  and their adjoints. This gives us an enlarged space with four states per site, with one redundant state corresponding to double occupancy, eliminated using Gutzwiller projection. There are various possibilities for doing this elimination leading to the different theories in literature. This includes the popular slave Boson or slave Fermion technique [19–21], where additional degrees of freedom, over and above the already enlarged 4 dimensional local state space, are introduced and finally eliminated as best as possible. This handling of the redundancy leads to gauge theories for the  $t$ – $J$  model that are reviewed in Ref. [21].

In the enlarged state space let us block diagonalize the state space into physical and unphysical states and write the projection operator as

$$[\psi] = \begin{bmatrix} \psi^{ph} \\ \psi^{un} \end{bmatrix}; \quad \hat{P}_G = \begin{bmatrix} \mathbb{1}^{ph} & 0 \\ 0 & 0 \end{bmatrix}, \quad (3)$$

where  $\mathbb{1}^{ph}$  is the identity operator in the physical space. The relevant operators in the theory  $Q_j$  e.g. the Hamiltonian, the creation operators or the destruction operators, are now written in terms of the canonical Fermions:

$$Q_j = \begin{bmatrix} Q_j^{pp} & Q_j^{pu} \\ Q_j^{up} & Q_j^{uu} \end{bmatrix}. \quad (4)$$

The next goal of the construction is to ensure that a state resulting from the application of a sequence of operators on a projected state remains in the projected space, i.e.

$$[\psi]_{final} = Q_M \dots Q_2 \cdot Q_1 \cdot \hat{P}_G \cdot [\psi]_{initial}, \quad (5)$$

and  $[\psi]_{final} = \hat{P}_G \cdot [\psi]_{final}$ . If this condition is not ensured, the projector has to be introduced at all intermediate time slices, thus making the calculations intractable. A sufficiency condition for this is the commutation  $[Q_j, \hat{P}_G] = 0$  for all  $j$ . The slave Boson–Fermion technique uses the conservation of the Gutzwiller constraint by writing a suitable version of the Hamiltonian. This enables the use of a time independent Lagrange multiplier, as demonstrated in the work of Read and Newns [20]. In Section 3.1, we display a compact Hermitian representation that also achieves this, without however adding further states (beyond the four states) into the problem.

While the commutation condition  $[Q_j, \hat{P}_G] = 0$  is sufficient, it is not necessary, and a much less restrictive condition can be found. We note that if the operators  $Q_j$  have a vanishing  $Q_j^{up}$  then the product in Eq. (5) remains in the physical sector with

$$[\psi]_{final} = \begin{bmatrix} Q_M^{pp} \dots Q_2^{pp} \cdot Q_1^{pp} \cdot \psi_{initial}^{ph} \\ 0 \end{bmatrix}. \quad (6)$$

The property of a commuting projection operator  $[Q_j, \hat{P}_G] = 0$ , requires that  $Q_j^{pu} = 0$  as well as  $Q_j^{up} = 0$ , whereas the vanishing property of the unphysical components noted in Eq. (6) requires only  $Q_j^{up} = 0$ . Then  $Q_j^{pu}$  as well as  $Q_j^{uu}$  are quite arbitrary. With this property, all the  $Q_j$  operators in Eq. (4) are block triangular

$$Q_j = \begin{bmatrix} Q_j^{pp} & Q_j^{pu} \\ 0 & Q_j^{uu} \end{bmatrix}. \quad (7)$$

In more formal terms the sufficiency condition with least constraints that leads to Eq. (6) (via the block triangularity Eq. (7)) is

$$(\mathbb{1} - \hat{P}_G) \cdot Q_j \cdot \hat{P}_G = 0. \quad (8)$$

This condition is also expressible as  $[Q_j, \hat{P}_G] \cdot \hat{P}_G = 0$ ; a conditional vanishing of the commutator, when right operating on projected states. This observation provides some intuition for why Eq. (8) is sufficient in the present context. In view of the block triangular operators in Eq. (7), the adjoint property, namely of representing conjugate operators by their matrix Hermitian conjugates, is lost in this representation. This is seen clearly in Eq. (1), where the first two operators are mutual adjoints in the defining representation, but not so in the canonical basis. In general this situation is expected to lead to non Hermitian Hamiltonians. The non Hermitian representation in Eq. (64) and Section 3.2 implements this idea and therefore leads to the most efficient canonical theory. We show that it exactly matches the minimal theory, found from the minimal description of the  $t$ - $J$  model in terms of the Hubbard  $X$  operators and the Schwinger equations of motion. It is notable that the Gutzwiller

projection operator does not appear *explicitly* in the equations of motion, although it does play a crucial role in the canonical theory, and is at the root of its difficulty.

The plan of the paper is as follows. In Sections 2.1–2.3 we review the Schwinger equations of motion for the  $t$ - $J$  model, and the ingredients of the recent method developed for a systematic expansion in a parameter  $\lambda$ . In Section 2.4 we summarize the general form of the Greens function at low frequencies near the Fermi surface, and obtain the prototypical spectral function of the theory. We summarize in Section 2.5 a kink in the electronic dispersion that arises from the theory, and seems to be closely related to that seen in many photoemission experiments. We also present simple but important ideas for analyzing photoemission data, with a view to isolating important feature of asymmetry predicted by the ECFL theory.

In Section 3 we formulate the “best possible” representation of the Hubbard operators in terms of canonical Fermions, as discussed above. Section 3.1 summarizes the well known representation and Section 3.2 implements the block triangular idea to obtain a non Hermitian method with least redundancy. Sections 3.3 and 3.4 give further details of the Hamiltonian in this representation and the proof of the antiperiodic temporal boundary conditions necessary for defining the new framework.

In Section 4, the above non Hermitian representation is used to analyze the nature of the Greens function of projected electrons. Quite remarkably this process also yields the Greens function as a convolution of an auxiliary Greens function and a caparison function, in complete parallel to that obtained from the Schwinger method employed in Sections 2.2 and 2.3. In Section 5 we generalize the above representation to define  $\lambda$  Fermions where the Gutzwiller projection is only partial, and becomes full at  $\lambda = 1$ . The equations of motion from these Fermions are shown to be those obtained in the  $\lambda$  expansion of Section 2.3.

In Section 6 we display a close analogy between the non Hermitian representation of the Gutzwiller projected electrons and the well known Dyson–Maleev representation of spin operators in terms of canonical Bosons. This connection also provides further meaning of the small parameter  $\lambda$  in the Fermion theory, as a parallel of the expansion parameter  $\frac{1}{25}$  of the Dyson–Maleev theory. A connection with the work of Harris, Kumar, Halperin and Hohenberg (HKHH) [18] is noted, who invented a method for computing the lifetime of spin waves in antiferromagnets, with considerable overlap with our representation of the Greens function with two self energies.

In Section 7, we cast the canonical theory in terms of Fermionic path integrals, and show how the exact Schwinger equations of motion can be obtained directly from this representation, thereby validating all the links in the argument. The subtle role of the Gutzwiller projection operator is explored, it does not appear explicitly in the equations of motion and yet plays an important role in the theory. In Section 8 we summarize the main points of the paper.

In Appendix A we summarize the derivation of the minimal equations of motion from the Schwinger viewpoint. In Appendices B–D we provide the details of the coherent state path integrals and the implementation of the Gutzwiller projection. In Appendix E we provide a more detailed interpretation of the caparison function in terms of a change of variable of the source fields.

## 2. Summary of the ECFL theory for the $t$ - $J$ model

### 2.1. The $t$ - $J$ model preliminaries

The well studied  $t$ - $J$  model is a two component Fermi system on a lattice, defined on the restricted subspace of three local states, obtained by excluding all doubly occupied configurations. The allowed states are  $|a\rangle$  with  $a = 0, \uparrow, \downarrow$ , and the double occupancy state  $|\uparrow\downarrow\rangle$  is removed by the (Gutzwiller) projection operator. These Gutzwiller projected electron operators are denoted, in the convenient notation due to Hubbard, as  $X_i^{a,b} = |a\rangle\langle b|$ . Its Hamiltonian  $H_{tJ}$  is expressed in terms of the  $X$  operators so that the single occupancy constraint is explicit. Summing over repeated spin indices we write

$$H_{tJ} = H_t + H_J,$$

$$H_t = - \sum_{ij} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_i X_i^{\sigma\sigma},$$

$$H_J = \frac{1}{2} \sum_{ij} J_{ij} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} X_i^{\sigma\sigma} X_j^{\sigma'\sigma'} \right). \quad (9)$$

In computing the Green's functions we add two kinds of Schwinger sources to the Hamiltonian; the anticommuting Grassmann pair  $J, J^*$  coupling to electron creation and destruction operators, and the commuting potential  $\mathcal{V}$ , coupling to the charge as well as spin density. These sources serve to generate compact Schwinger equations of motion (EOM), and are set to zero at the end. Explicitly we write

$$\hat{\mathcal{A}}_S = \sum_i \int_0^\beta \hat{\mathcal{A}}_S(i, \tau) d\tau, \quad (10)$$

$$\hat{\mathcal{A}}_S(i, \tau) = [X_i^{\sigma 0}(\tau) J_{i\sigma}(\tau) + J_{i\sigma}^*(\tau) X_i^{0\sigma}(\tau)] + \mathcal{V}_i^{\sigma'\sigma}(\tau) X_i^{\sigma'\sigma}(\tau),$$

and all time dependences are as in  $Q(\tau) = e^{\tau H_{tJ}} Q e^{-\tau H_{tJ}}$ . The generating functional of Green's functions of the  $t$ - $J$  model is

$$Z[J, J^*, \mathcal{V}] \equiv \text{Tr}_{tJ} e^{-\beta H_{tJ}} T_\tau \left( e^{-\hat{\mathcal{A}}_S} \right). \quad (11)$$

It reduces to the standard partition function on turning off the indicated source terms. The Green's functions for positive times  $0 \leq \tau_j \leq \beta$ , are defined through the Martin–Schwinger prescription [22,23]:

$$\mathcal{G}_{\sigma\sigma'}(i\tau_i, f\tau_f) = - \frac{\left\langle T_\tau \left( e^{-\hat{\mathcal{A}}_S} X_i^{0\sigma}(\tau_i) X_f^{\sigma'0}(\tau_f) \right) \right\rangle}{\langle T_\tau e^{-\hat{\mathcal{A}}_S} \rangle}. \quad (12)$$

The functional  $Z$  conveniently yields the Green's functions upon taking functional derivatives with respect to the sources, e.g.

$$\mathcal{G}_{\sigma\sigma'}(i\tau_i, f\tau_f) = \left( \frac{1}{Z} \frac{\delta^2 Z}{\delta J_{i\sigma}^*(\tau_i) \delta J_{f\sigma'}(\tau_f)} \right), \quad (13)$$

where the sources are turned off at then end. We note that  $n_\sigma$ , the number of particles per site, is determined from the number sum rule:

$$n_\sigma = \mathcal{G}_{\sigma\sigma}(i\tau^-, i\tau), \quad (14)$$

and  $\mu$  the chemical potential is fixed by this constraint.

## 2.2. The Schwinger equations of motion

The detailed theory of the  $t$ - $J$  model developed so far [4,6] uses the Schwinger equations of motion. Since these equations play a fundamental role in the theory, we summarize next the equations of motion and their extension, obtained by introducing a parameter  $\lambda$ . We relegate to [Appendix A](#) the derivation of the “minimal theory” equations. In the minimal theory, the most compact set of Schwinger equations are established, and some redundant terms from Ref. [4] are omitted. This minimal version of the theory is important for the purposes of the present paper, since our goal in this paper is to recover these from a canonical formalism.

As the Schwinger school has [22,24,25] emphasized, a field theory is rigorously determined by its equations of motion plus the boundary conditions. We can also establish alternate descriptions such as path integrals formulations, from the requirement that they reproduce these equations of motion—we present an example of this approach in Section 7.2. In terms of the original description of the  $t$ - $J$  model involving the Hubbard  $X$  operators, the Schwinger equation of motion is a partial differential equation in time and also a functional differential equation involving the derivatives with respect to a Bosonic source:

$$\left( \mathbf{g}_{0,\sigma_i,\sigma_j}^{-1}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) - Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j) \right) \times \mathcal{G}_{\sigma_j\sigma_f}(j\tau_j, f\tau_f) = \delta_{if} \delta(\tau_i - \tau_f) (\delta_{\sigma_i\sigma_f} - \gamma_{\sigma_i\sigma_f}(i\tau_i)), \quad (15)$$

where  $\mathbf{g}_0$  is the noninteracting Green's function (Eq. (136)),  $\hat{X}$  is a functional derivative operator (Eq. (130)),  $\gamma$  is the local Green's function obtained from  $\mathcal{G}$  as  $\gamma_{\sigma_a\sigma_b}(i\tau_i) = \sigma_a\sigma_b\mathcal{G}_{\bar{\sigma}_b\bar{\sigma}_a}(i\tau_i^-, i\tau_i)$  (see Eq. (135)) and  $Y_1$  is the band hopping times  $\gamma$  (Eq. (131)); further details can be found in Appendix A. This equation has been written down in Refs. [4,6]: Antiperiodic boundary conditions with respect to both times (as in Eqs. (75) and (76)), and the number sum-rule (14) together with the equation of motion (15), specify the theory completely.

### 2.3. The $\lambda$ expansion, the shift identities and second chemical potential $u_0$

The idea of introducing a parameter into the EOM (15) becomes quite natural when we observe the Schwinger EOM for the Hubbard model closely. These can be written schematically as  $(\mathbf{g}_0^{-1} - U\delta/\delta\mathcal{V} - UG) \cdot G = \delta \mathbb{1}$ . By scaling the interaction  $U \rightarrow \lambda U$ , with a parameter  $\lambda$  ( $0 \leq \lambda \leq 1$ ), the interacting theory is connected continuously to the Fermi gas by tuning  $\lambda$  from 1 to 0. The standard perturbative expansion can be organized by counting the various powers of  $\lambda$ , setting  $\lambda = 1$  at the end before evaluating the expressions [26]. Below in Section 5 we provide a more microscopic argument for introducing the  $\lambda$  parameter in the Hubbard  $X$  operators directly, this method leads back to the equations found here.

In the corresponding equation for the  $t$ - $J$  model (15), we observe that the Green's function differs from that for the free Fermi gas through two terms on the left hand side, exactly as in the Hubbard model, but also through one term on the right hand side. Scaling these three terms by  $\lambda$ , we rewrite (15) schematically as:

$$\left(\mathbf{g}_0^{-1} - \lambda\hat{X} - \lambda Y_1\right) \cdot \mathcal{G} = \delta (\mathbb{1} - \lambda\gamma). \quad (16)$$

The strategy of the perturbative  $\lambda$  expansion method is to build up the solution of this equation at  $\lambda = 1$  through a suitable expansion in  $\lambda$ , starting from the free Fermi limit  $\lambda = 0$ . Thus  $\lambda < 1$  corresponds to the admixture of a finite fraction of double occupancy that vanishes at  $\lambda = 1$ . Insights from sum rules, the skeleton graph expansion and the physics of the Hubbard sub bands has played a major role in formulating a systematic  $\lambda$  expansion described in detail in Refs. [4,6].

Within this approach it is also necessary to add a term  $\lambda u_0 \sum_i N_{i\uparrow} N_{i\downarrow}$  to the Hamiltonian, and a corresponding term to the EOM, so that the  $\hat{X}$  and  $Y_1$  in Eq. (16) are suitably redefined. Here  $u_0$  is an extra Hubbard interaction type parameter that is *determined by a sum rule* as explained below. At  $\lambda = 1$  such a term makes no difference since the double occupancy is excluded. This parameter  $u_0$  also enables us to enforce a simple but crucial symmetry of the  $t$ - $J$  model—the *shift invariance*, noted in Ref. [6]. This invariance arises from the twofold function of the hopping in the  $t$ - $J$  model when expressed in terms of the canonical operators, of providing hopping as well as the four Fermion (interaction) terms. Therefore in an exact treatment, adding a constant times the identity matrix to the hopping matrix:  $t_{ij} \rightarrow t_{ij} + \text{const} \times \delta_{ij}$ , shifts the center of gravity band innocuously. In approximate implementations it has the unphysical effect of also adding to the interaction (i.e. four Fermion type) terms. Such a change must therefore be compensated by an adjustable parameter that can soak up this additive constant. Indeed  $u_0$  provides precisely this type of a parameter. It also plays the role of a second chemical potential  $u_0$  [6] to fix the number of Fermions in the auxiliary Green's function  $\mathbf{g}$  through  $n_\sigma = \mathbf{g}_{\sigma\sigma}(i\tau^-, i\tau)$ , while the thermodynamical chemical potential  $\mu$  (residing in the non interacting  $\mathbf{g}_0^{-1}$ ), is fixed by the number sum rule  $n_\sigma = \mathcal{G}_{\sigma\sigma}(i\tau^-, i\tau)$  (Eq. (14)). Enforcing this shift invariance to each order in the  $\lambda$  expansion plays an important “watchdog” role on the  $\lambda$  expansion, in addition to other standard constraints such as the Ward identities.

To summarize some key points of the  $\lambda$  expansion, we first decompose the Greens function into the space time convolution of an auxiliary Greens function and a caparison function as:

$$\mathcal{G} = \mathbf{g} \cdot \mu. \quad (17)$$

With this the operator in Eq. (16) acts on the two factors of Eq. (17), and breaks into two equations upon using the *ansatz* that  $\mathbf{g}$  has a canonical structure  $(\mathbf{g}_0^{-1} - \lambda\hat{X} - \lambda Y_1) \cdot \mathbf{g} = \delta \mathbb{1}$ . The  $\lambda$  expansion [6] is then an iteration scheme that proceeds by an expansion of the caparison function  $\mu(k)$  and  $Y_1$

**Table 1**

A flowchart of the ECFL theory as developed in Refs. [4,6]. See Sections 2.2 and 2.3 for a detailed description.

Step(I)	Step(II)	Step(III)	Step(IV)	Step(V)	Step(VI)	Step(VII)
Green's function $\mathcal{G}$ in terms of Hubbard operators.	Exact Schwinger equations of motion for $\mathcal{G}$ .	Product expression into canonical part $\mathbf{g}$ and adaptive spectral weight (caparison) part $\mu(k)$ .	Exact equations for $\mathbf{g}(k)$ and $\mu(k)$ .	Introduction of interpolating parameter $\lambda$ connecting the Fermi gas to the extreme correlation limit.	Shift invariance requires second chemical potential $u_0$ . Same sum rule for both Greens functions so that Fermi surface volume is conserved.	Successive orders in $\lambda$ expansion satisfying shift invariance for practical calculations.
$\mathcal{G}$	$\partial_\tau \mathcal{G}$	$\mathcal{G}(k) = \mathbf{g}(k)\mu(k)$		$0 \leq \lambda \leq 1$	$\sum \mathcal{G} = \sum \mathbf{g} = \frac{n}{2}$	

( $Y_1 = t\gamma$ ) in powers of  $\lambda$ . Dyson's skeleton graph idea is implemented by keeping the auxiliary  $\mathbf{g}$  intact (i.e. unexpanded in  $\lambda$ ), while all other variables are expanded in powers of  $\lambda$  and  $\mathbf{g}$ , thereby obtaining self consistent equations for  $\mathbf{g}$  and the vertex functions. Successive levels of approximation are obtained by retaining increasing powers of  $\lambda$ . At each approximation level, we set  $\lambda = 1$  before actually evaluating the expressions, and implement the antiperiodic boundary conditions (75), (76), and the number sum-rule  $n_\sigma = \mathcal{G}_{\sigma\sigma}(i\tau^-, i\tau)$  (Eq. (14)).

Elaborating on the representation (17) of  $\mathcal{G}$ , we note that the  $\gamma$  term on the right hand side of (16) is due to the non canonical anticommutator of the projected Fermi operators. As noted in Ref. [4], this term contains the essential difficulty of the  $t$ - $J$  problem, having no parallel in the (canonical) Hubbard type models. After turning off the sources, in the momentum–frequency space we can further introduce two self energies  $\Psi(k, i\omega)$ , and  $\Phi(k, i\omega)$  with

$$\mu(\vec{k}, i\omega_n) = 1 - \frac{n}{2} + \Psi(\vec{k}, i\omega_n) \quad (18)$$

$$\mathbf{g}^{-1}(\vec{k}, i\omega_n) = \mathbf{g}_0^{(-1)}(\vec{k}, i\omega) - \Phi(\vec{k}, i\omega_n), \quad (19)$$

where the constant  $\frac{n}{2}$  in Eq. (18) is fixed by the condition that  $\Psi$  vanishes at infinite frequency. The auxiliary Greens function satisfies a second sumrule analogous to Eq. (14), written in the Fourier domain:

$$(k_B T) \sum_{k,n} e^{i\omega_n 0^+} \mathbf{g}_{\sigma\sigma}(k, i\omega_n) = n_\sigma. \quad (20)$$

Clearly the same sumrule holds for  $\mathcal{G}_{\sigma\sigma}(k, i\omega_n)$ . Eq. (17) can now be written explicitly in the non-Dysonian form proposed in Refs. [4,5]

$$\mathcal{G}(\vec{k}, i\omega) = \frac{1 - \frac{n}{2} + \Psi(\vec{k}, i\omega)}{\mathbf{g}_0^{(-1)}(\vec{k}, i\omega) - \Phi(\vec{k}, i\omega)}. \quad (21)$$

As argued in [4,6,8,9], simple Fermi liquid type self energies  $\Psi$  and  $\Phi$  can, in the combination above, lead to highly asymmetric (in frequency) Dyson self energies from the structure of Eq. (21), thus providing a considerable tactical advantage in describing extreme correlations. We further discuss the physical meaning of this decomposition and the twin self energies in Section 4. Table 1 provides an overview of the various steps in the construction of the theory.

#### 2.4. $\mathcal{G}(\vec{k}, i\omega_n)$ and the low energy spectral function in ECFL theory

We summarize here the low temperature low energy theory near the Fermi surface that follows from the general structure of Eq. (21) in terms of a small number of parameters, upon assuming that the two self energies have a Fermi liquid behavior at low energies. In the limit of large dimensions, a similar exercise gives a very interesting spectral function that matches the exact solution of the

$U = \infty$  Hubbard model found from the dynamical mean field theory (DMFT) [9]. The presentation below generalizes that to include a momentum dependence that is absent in high dimensions, and is supplemented by a discussion of the behavior of the various coefficients as the density of electrons  $n$  approaches unity, or equivalently the hole density  $\delta \rightarrow 0$ .

The Dyson self energy can be inferred from a simple inversion, and has a strong set of corrections to the Fermi liquid theory that we delineate here. We assume here a Fermi liquid type state that survives the limit of small hole density  $\delta \rightarrow 0$ . In reality at very small  $\delta$  several other broken symmetry states would compete and presumably win over the liquid state, so that this Fermi liquid state would be metastable. Its characteristics are of interest and hence we proceed to describe these.

We study Eq. (21) by analytically continuing  $i\omega \rightarrow \omega + i0^+$  and write

$$\mathbf{g}_0^{(-1)}(\vec{k}, i\omega) = \omega + \mu - \left(1 - \frac{n}{2}\right) \varepsilon_k. \quad (22)$$

Let us define  $\hat{k}$  as the *normal deviation* from the Fermi surface i.e.  $\hat{k} = (\vec{k} - \vec{k}_F) \cdot \vec{k}_F / |\vec{k}_F|$ , and the frequently occurring Fermi liquid function

$$\mathcal{R} = \pi \{\omega^2 + (\pi k_B T)^2\}. \quad (23)$$

We carry out a low frequency expansion as follows:

$$1 - \frac{n}{2} + \Psi(\vec{k}, \omega) = \alpha_0 + c_\psi (\omega + v_\psi \hat{k} v_f) + i\mathcal{R}/\gamma_\psi + O(\omega^3), \quad (24)$$

where  $\alpha_0 = 1 - \frac{n}{2} + \Psi_0$  is the constant term at the Fermi surface, and a similar expansion for  $\Phi(\vec{k}, \omega)$  so that

$$\omega + \mu - \left(1 - \frac{n}{2}\right) \varepsilon_k - \Phi(k, \omega) = (1 + c_\phi) \left(\omega - v_\phi \hat{k} v_f + i\mathcal{R}/\Omega_\phi + O(\omega^3)\right), \quad (25)$$

where  $v_f = (\partial_k \varepsilon_k)_{k_F}$  is the *bare* Fermi velocity. The expansion coefficients above are in principle functions of the location of  $\vec{k}_F$  on the Fermi surface, and have suitable dimensions to ensure that  $\Psi$  is dimensionless and  $\Phi$  is an energy. The dimensionless velocity renormalization constants  $v_\phi$  and  $v_\psi$  capture the momentum dependence normal to the Fermi surface, arising from the two respective self energies. The Greens function near the Fermi surface can now be written as

$$\mathcal{g}(\vec{k}, \omega) \sim \frac{z_0}{\alpha_0} \left( \frac{\alpha_0 + c_\psi (\omega + v_\psi \hat{k} v_f) + i\mathcal{R}/\gamma_\psi}{\omega - v_\phi \hat{k} v_f + i\mathcal{R}/\Omega_\phi} \right) \quad (26)$$

where  $z_0 = \alpha_0/(1 + c_\phi)$  is the net quasiparticle renormalization constant. The spectral function can be computed from  $A(\vec{k}, \omega) = -\frac{1}{\pi} \Im m \mathcal{g}(\vec{k}, \omega + i0^+)$  in the ECFL form of a Fermi liquid function times a caparison function  $\mu(k, \omega)$  as follows:

$$A(\vec{k}, \omega) = \frac{z_0}{\pi} \frac{\Gamma_0}{(\omega - v_\phi \hat{k} v_f)^2 + \Gamma_0^2} \times \mu(k, \omega), \quad (27)$$

where the (Fermi liquid) width function

$$\Gamma_0(\hat{k}, \omega) = \eta + \frac{\pi(\omega^2 + (\pi k_B T)^2)}{\Omega_\phi}, \quad (28)$$

with an extra phenomenological parameter  $\eta$  required to describe elastic scattering [14] in impure systems. The caparison function is

$$\mu(\hat{k}, \omega) = 1 - \frac{\omega}{\Delta_0} + \frac{v_0 \hat{k} v_f}{\Delta_0}, \quad (29)$$

where we introduced an important (emergent) low energy scale combining the other parameters:

$$\Delta_0 = \alpha_0 \frac{\gamma_\psi}{\Omega_\phi - c_\psi \gamma_\psi} \quad (30)$$

and the dimensionless momentum dependence coefficient

$$\nu_0 = (\nu_\psi \gamma_\psi c_\psi + \nu_\phi \Omega_\phi) / (\Omega_\phi - c_\psi \gamma_\psi). \quad (31)$$

A cutoff  $\theta(\mu(\hat{k}, \omega))$  is implicit in Eq. (29), so that the function  $\mu(\hat{k}, \omega)$  is assumed to be zero at large positive frequencies as discussed in Ref. [4]. The five final parameters defining the spectral function (27) are  $z_0, \nu_0, \nu_\phi, \Omega_\phi, \Delta_0$ . For fitting experimental data, it may be best to think of them as adjustable parameters that determine the line shapes, their asymmetries and also features in the spectral dispersions. In addition the  $\eta$  parameter is needed to describe impurities that are not contained in the microscopic theory. In the early fit [14] the total number of free parameters is even smaller—just two instead of five. The corrections to the Landau Fermi liquid theory are encapsulated in the caparison factor, which contains a correction term that is odd in frequency and seems to be ultimately responsible for the asymmetric appearance of the line shapes [14,8].

For reference we note that in the limit of high dimensions [9], the coefficient of the momentum dependent term  $\nu_0$  vanishes in Eq. (27), while the earlier fits to experiments in [14], it is non zero, and in modified fits [15] its magnitude is varied to get a good description of the constant energy cuts of the data.

It is useful to consider the approach to the Mott insulating limit, where the parameters behave in a specific fashion to satisfy the expected behavior. We consider the limit of density  $\delta \rightarrow 0$ , and a frequency scale  $0 \leq |\omega| < \omega_c \sim \delta t$ , where the above expression (27) may be expected to work. For reference, it is useful to note that in this limiting case, the widely used Gutzwiller–Brinkman–Rice theory [2,27] gives the quasiparticle propagator as:

$$G_{GBR}(\vec{k}, \omega) \sim \frac{z}{\omega - z \hat{k} v_f}, \quad (32)$$

where  $z$  vanishes linearly with  $\delta$  as  $z = 2\delta/(1 + \delta)$ . This leads to a delta function spectral weight  $A_{GBR} = z \delta(\omega - z \hat{k} v_f)$ . In contrast Eq. (27) provides the spectral function at non zero  $T$  and  $\omega$ .

As  $n \rightarrow 1$  in Eq. (24) we expect that the constant  $\Psi_0 \rightarrow -\frac{n}{2}$ , in order to reach the Mott insulating limit continuously. This implies that  $\alpha_0 \propto \delta$  in this regime, and this drives the various other coefficients as well. We summarize the expected behavior of the above five coefficients

$$\begin{aligned} z_0 &\rightarrow \bar{z}_0 \times \delta; & \Delta_0 &\rightarrow \bar{\Delta}_0 \times \delta; & \Omega_\phi &\rightarrow \bar{\Omega}_\phi \times \delta; \\ \nu_0 &\rightarrow \bar{\nu}_0 \times \delta; & \nu_\phi &\rightarrow \bar{\nu}_\phi \times \delta; \end{aligned} \quad (33)$$

by using an overline for denoting a non vanishing limit of the stated variable [9,28]. The scaling of the velocity constants  $\nu$  is guided by the results in high dimensions, and ensure that the dispersing quasiparticles have a vanishing bandwidth as we approach the insulator—as emphasized by Brinkman and Rice [27]. From this we find that the ECFL spectral function (27) satisfies a simple homogeneity (i.e. scaling) relation valid in the low energy regime for a scale parameter  $s$ :

$$A(\hat{k}, s \omega | s T, s \delta) = A(\hat{k}, \omega | T, \delta), \quad (34)$$

where the dependence on the temperature and hole density are made explicit. The momentum variable does not scale with  $s$  due to the assumed behavior of the  $\nu$ 's. The scaling holds for  $\eta = 0$ , and generalizes to a non zero values if we scale  $\eta \rightarrow s \eta$ . This scaling relation describes a Fermi liquid including significant corrections to Fermi liquid theory through the caparison function. It rests upon the specific behavior for the coefficients as the density varies near the insulating state, unlike other generalized scaling relations that have been proposed in literature Ref. [29] for non Fermi liquid states. If set  $s \times \delta = \delta_0$  with say  $\delta_0 \lesssim .5$ , then the ratio  $\frac{\delta_0}{\delta} \gg 1$  and we infer

$$A(\hat{k}, \omega | T, \delta) \sim A\left(\hat{k}, \omega \frac{\delta_0}{\delta} \left| T \frac{\delta_0}{\delta}, \delta_0\right.\right), \quad (35)$$

relating the low hole density system to an overdoped (i.e. high hole density) system at a high effective temperature. This relation provides basic intuition for why the  $t$ - $J$  model, near the insulating limit behaves almost like a classical liquid, unless one fine tunes parameters very close to the  $T = 0, \omega = 0$  limit.

## 2.5. Electronic origin of the low energy kink and further tests of dynamical asymmetry

In this section we summarize the origin of the important low energy *kink* feature of the dispersion relation obtained in the ECFL theory. Since a similar feature is seen in the experiments on angle resolved photoemission studies (ARPES) of various groups [30–32,14], it is worth clarifying the purely electronic origin of this feature within the ECFL theory. A higher (binding) energy kink is also seen and is well understood in terms of the behavior of the self energy over a greater range [5,9], and is not pursued here. Rather we focus on the low energy kink seen around  $-.05$  eV in several compounds [30–32,14], and finds a natural interpretation within ECFL.

We also present a few experimentally testable features relating to *dynamical asymmetry*, i.e. the asymmetric in  $\omega$  correction to the Fermi liquid theory contained in ECFL, arising from the caparison function in Eq. (27).

Let us assume that  $|\omega| \ll \Gamma_0$  at low enough frequency relative to  $T$  so that we may treat  $\Gamma_0$  as a constant. We may then bring Eq. (27) to an interesting form studied in Ref. [5] by defining variables

$$\begin{aligned}\epsilon &= \frac{\omega - v_\phi \hat{k} v_f}{\Gamma_0} \\ \sinh u_k &= \frac{\Delta_0 + (v_0 - v_\phi) \hat{k} v_f}{\Gamma_0},\end{aligned}\quad (36)$$

so that the spectral function reduces to the standard form occurring in the ECFL theory:

$$A(u_k, \epsilon) = A_0 \frac{\sinh u_k - \epsilon}{1 + \epsilon^2} \times \theta(\sinh u_k - \epsilon) \quad (37)$$

with  $A_0 = \frac{z_0}{\Delta_0}$ . This expression is valid for small enough  $\epsilon$  [5,4], and can be viewed as the (weighted) sum of the real and imaginary parts of a simple damped oscillator with a scaled susceptibility  $\chi(\epsilon) = 1/(\epsilon + i)$ . It is interesting to note that the scaled spectral function (37) can be related to the (scaled) Fano line shape

$$A_{Fano}(q_f, \epsilon) \propto \frac{(q_f + \epsilon)^2}{(1 + \epsilon^2)}. \quad (38)$$

This spectrum is often considered with the Fano parameter  $q_f > 0$ , it is highlighted by a vanishing at negative energies  $\epsilon = -q_f$ , representing the destructive interference of a scattering amplitude with a background term arising from a continuum of states. However we can flip the sign of  $q_f$  and by choosing  $q_f = -e^{u_k}$ , we can relate these through

$$A(u_k, \epsilon) \propto (A_{Fano}(-e^{u_k}, \epsilon) - A_{Fano}(-e^{u_k}, \infty)). \quad (39)$$

For the purpose of representing ARPES spectral functions, the scaled spectral function (37) gains an advantage over the Fano line shape (38) by the absence of a background at large  $|\epsilon|$ . In relating them via Eq. (39), the background term in the Fano process is killed, while its interference with the peak is retained.

Unlike the simple Lorentzian obtained at  $u_k \rightarrow \infty$ , the energy variable enters the numerator as well as the denominator in both Eq. (37) and the Fano shape. This feature gives rise to the characteristic skew to the ECFL spectrum. The spectral function can be maximized with respect to the frequency at a fixed  $\hat{k}$ , yielding the energy distribution curve (EDC) dispersion  $E_k^*$ , or with respect to  $\hat{k}$  at a fixed frequency  $\omega$ , giving the momentum distribution curve (MDC) dispersion  $E_k$ . Let us introduce the convenient variables

$$r = \frac{v_0}{v_\phi}, \quad (40)$$

giving the ratio of the two velocity factors. The ratio  $r = 0$  in the limit of high dimensions [9]. In the simplified ECFL analysis in [4,14], we find  $r > 1$  due to the suppression of  $v_\phi$  relative to  $v_0$  by

a quasiparticle renormalization factor  $z_{FL}$ . We see below that the magnitude and sign of  $(r - 1)$  play a significant role in determining the location of the kink, and its observability in ARPES respectively. We also introduce a (linear in  $\hat{k} v_f$ ) energy variable:

$$Q(\hat{k}) = \Delta_0 + (v_0 - v_\phi) \hat{k} v_f. \quad (41)$$

In terms of these, the two dispersions are obtained as

$$E(k) = \frac{1}{2-r} \left( v_\phi \hat{k} v_f + \Delta_0 - \sqrt{r(2-r) \Gamma_0^2 + Q^2} \right), \quad (42)$$

$$E^*(k) = \left( v_0 \hat{k} v_f + \Delta_0 - \sqrt{\Gamma_0^2 + Q^2} \right). \quad (43)$$

Simplifying the notation, both energy dispersions are of the form  $E \sim \gamma Q - \sqrt{Q^2 + M^2}$ , i.e. the hybrid of a massless and a massive Dirac spectrum. As  $Q$  varies from  $-\infty$  to  $\infty$ , the energy crosses over from  $(\gamma + 1)Q$  to  $(\gamma - 1)Q$ , thus exhibiting a knee or a kink near  $Q \sim 0$ , with its sharpness determined by the “mass term”. The mass term in the MDC spectrum depends on the ratio  $r$ , and this generally leads to a smaller magnitude. Upon turning off the decay rate  $\Gamma_0$ , both the EDC and MDC spectra reduce to the expected spectrum  $\varepsilon_k = v_\phi \hat{k} v_f$ , arising from the pole of the auxiliary Greens function in Eq. (21). These expressions illustrate an unusual feature of this theory: the two dispersions are influenced by the emergent energy scale  $\Delta_0$ , as well as the width  $\Gamma_0$  (Eq. (28)).

The above dispersions exhibit an interesting *kink feature* at  $Q = 0$  in both spectra. The condition  $Q = 0$  locates the kink momentum as

$$(\hat{k} v_f)_{kink} = \frac{\Delta_0}{v_\phi(1-r)}, \quad (44)$$

it corresponds to occupied momenta provided  $r > 1$ , we will confine to this case below. For the other case  $r < 1$ , a kink would arise in the unoccupied side, for this reason we do not pursue it here. For  $|Q| \gg \Gamma_0$ , the two dispersions asymptotically become  $E^*(k) \sim (v_0 + (v_0 - v_\phi) \text{sign}(\hat{k})) \hat{k} v_f$  and  $E(k) \sim \frac{1}{2-r} (v_\phi + (v_0 - v_\phi) \text{sign}(\hat{k})) \hat{k} v_f$ . Hence these spectra exhibit a change in velocity (i.e. slope) around  $Q \sim 0$  of magnitude  $2(v_0 - v_\phi)v_f$  for the EDC and the usually larger  $\frac{2}{2-r}(v_0 - v_\phi)v_f$  for the MDC spectrum. The change in slope of the spectrum occurs over a range  $\Delta Q \propto \Gamma_0$ , thus becoming sharper as  $\Gamma_0$  decreases.

The value of the EDC energy at the kink is found by substituting  $Q = 0$  and gives

$$E^*(k_{kink}) = -\frac{r}{r-1} \Delta_0 - \Gamma_0. \quad (45)$$

The MDC spectrum shows a kink for  $2 \geq r \geq 1$  at the same momentum (44), with energy

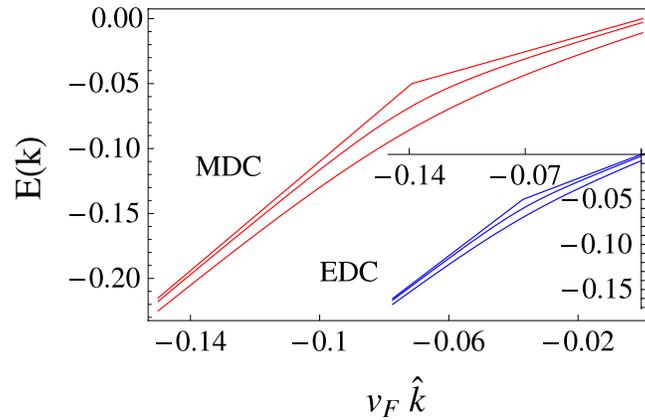
$$E(k_{kink}) = -\frac{1}{r-1} \Delta_0 - \Gamma_0 \sqrt{\frac{r}{2-r}}, \quad (46)$$

this feature is sharper than in the EDC spectrum since the effective damping is smaller.

When  $r > 2$ , the MDC energy is real only for  $|\hat{k} v_f| < (|\hat{k} v_f|)_{cutoff}$ , where the (negative) momentum

$$(\hat{k} v_f)_{cutoff} = (\hat{k} v_f)_{kink} + \frac{\Gamma_0}{v_\phi(r-1)} \sqrt{r(r-2)}.$$

For  $\hat{k} v_f$  beyond the cut off, the root becomes complex implying the loss of a clear peak in the MDC spectrum. Thus the spectrum “fades” before reaching the kink momentum (44). Therefore in this case, the kink is less than ideal, unlike the EDC kink or the MDC kink for  $1 \leq r \leq 2$ , which should be visible on both sides of the kink momentum. From Eq. (33) we may extract the hole density dependence of all the kink parameters, while  $\Gamma_0$ , determining the kink width, is given in Eq. (28).



**Fig. 1.** A kink feature in the MDC dispersion relation  $E(k)$  from Eq. (42) and in the inset from the EDC dispersion  $E^*(k)$  (Eq. (43)) with parameters  $\Delta_0 = .025$  eV,  $v_0 = 1.05$ ,  $v_\phi = 0.7$  and three values of  $\Gamma_0 = 0., .01, .02$  in eV from top to bottom. The kink is more pronounced in the MDC curve as discussed in text.

We observe in Fig. 1 that the kink becomes sharp when  $\Gamma_0$  decreases. The MDC curves display a sharper kink than the EDC curves, this is easy to understand since the effective damping is smaller in this case, and also the net change in velocity across the kink is greater, as discussed above. From Eq. (28) we see various parameters that control  $\Gamma_0$ , in case of laser ARPES, it is argued [14] that  $\eta$  is small so we expect to see sharper kinks in this setup. Further, as  $T$  drops below  $T_c$ , the d-wave superconductor has gapless excitations along the nodal direction  $\langle 11 \rangle$ , and the quasiparticles seen in this case are sharper. Theoretical considerations [33] show that in the superconducting state, a reduction in the available gapless states responsible for the linewidth implies a reduction of  $\Gamma_0$  and hence to a sharper kink.

We next discuss the feature of *dynamical asymmetry* in the spectra. It is also important to note that the ECFL spectral function (27) has an unusual correction to the standard Fermi liquid part, embodied in the caparison function  $\mu(k, \omega)$ . This function is odd in frequency, thus disturbing the particle hole symmetry of the Fermi liquid part, and it grows in importance as we approach the insulating state since  $\Delta_0 \rightarrow \delta\Delta_0$  as indicated in Eq. (33). It is also interesting that the spectral line shape in the calculation of Anderson and Casey (AC) [34] as well as Doniach and Sunjic (DS) [35] also have such odd in  $\omega$  corrections to the Fermi liquid part. In fact the AC result may be viewed as the vanishing of the scale  $\Delta_0 \propto k_B T$  so that the ground state is non Fermi liquid like. At finite  $T$  and  $\omega$  the AC and DS theories are parallel with the ECFL line shapes regarding the asymmetry as remarked in Ref. [8], and we wish to make a few comments about the experimental tests for such an asymmetry, going beyond standard measures such as the skewness factor.

DS [35] make the interesting point that the asymmetry is best isolated by looking at the inverse of the spectral function in a plot of

$$\frac{1}{A(k, \omega)} \text{ vs } (\omega - E_k^*)^2, \quad (47)$$

where  $E_k^*$  is the peak location in the EDC. With this plot, a Fermi liquid yields two coincident straight lines above and below  $E_k^*$ , whereas an asymmetric contribution, as in Eq. (27) or the DS line shape [35], would split into two distinct non linear curves, from below and above  $E_k^*$ . The inversion of the spectral function is an interesting device, since it refocuses attention on the asymmetric parts. For very similar reasons Ref. [4] (Fig. 1 inset) also advocates plotting the inverse of the spectral function. On the other hand an untrained examination of the EDC curves invariably focuses on the close proximity of the peaks of  $A(k, \omega)$ , these are arguably the least interesting part of the asymmetry story!

In fact armed with the explicit knowledge of the spectral function of the ECFL theory in Eq. (27), we can aim to do better in establishing the asymmetry and in determining the various parameters. We first redefine the frequency by subtracting off the EDC peak value

$$\tilde{\omega}_k = \omega - E_k^*, \quad (48)$$

so that the spectral peak occurs at  $\tilde{\omega}_k = 0$ . The inverse spectral function can be computed as a function of  $\tilde{\omega}_k$  and reads:

$$\frac{A(k, E_k^*)}{A(k, E_k^* + \tilde{\omega}_k)} = 1 + \frac{e^{u_k}}{2\Gamma_0} \times \frac{\tilde{\omega}_k^2}{\Gamma_0 \cosh(u_k) - \tilde{\omega}_k}, \quad (49)$$

where the peak value of the spectral function at  $\tilde{\omega}_k = 0$  is:

$$A(k, E_k^*) = \frac{A_0}{2} e^{u_k}. \quad (50)$$

We next construct the object  $\mathcal{Q}(\tilde{\omega}_k)$  from Eq. (49) by subtracting unity and cross multiplying:

$$\mathcal{Q}(\tilde{\omega}_k) = \frac{\tilde{\omega}_k^2}{A(k, E_k^*)/A(k, E_k^* + \tilde{\omega}_k) - 1}. \quad (51)$$

This variable is designed to be a  $\tilde{\omega}_k$  independent constant in a simple Fermi liquid with a Lorentzian line shape (i.e. Eq. (27) without the caparison function  $\mu$ ). Here  $\mathcal{Q}$  has dimensions of the square of energy, and when plotted against  $\tilde{\omega}_k$  in the small range surrounding zero i.e.  $|\tilde{\omega}_k| \leq \Gamma_0$  it exhibits a linearly decreasing behavior with  $\tilde{\omega}_k$  within the ECFL spectral function (27)

$$\mathcal{Q}(\tilde{\omega}_k) = \Gamma_0^2 (1 + e^{-2u_k}) - (2\Gamma_0 e^{-u_k}) \tilde{\omega}_k. \quad (52)$$

Note that this function is flat for the usual Fermi liquid state without asymmetric corrections, since in this case  $u_k \rightarrow +\infty$ . If found in data, this linear in  $\tilde{\omega}$  behavior is the distinctive aspect of the asymmetric line shapes. We can then read off various physical quantities once the curve of  $\mathcal{Q}(\tilde{\omega}_k)$  versus  $\tilde{\omega}_k$  is obtained. For this purpose we need the intercept  $\mathcal{Q}(0)$  and the slope near the origin  $(d\mathcal{Q}(\tilde{\omega}_k)/d\tilde{\omega}_k)_0$ . Clearly the  $\mathcal{Q}(\tilde{\omega}_k)$  function will deviate from a straight line sufficiently far from  $\tilde{\omega}_k = 0$ , and it will also be contaminated with background terms as well as noise. However, with high quality data this procedure could be useful in inverting the data to fit simple functional forms, and to make decisive tests of the predictions of the theories containing asymmetry, namely the DS and AC theories as well as ECFL.

### 3. Exact formulation in terms of a canonical Fermions

We will next rewrite this in canonical Fermi representation in *an enlarged Hilbert space* where double occupancy is permitted, and the singly occupied states form a subspace. We regard the physical subspace of states  $|\Psi\rangle$  as those that satisfy the condition of single occupancy, i.e.  $\hat{D}|\Psi\rangle = 0$  with the double occupancy operator  $\hat{D}$  is given by:

$$\hat{D} = \sum_i \hat{D}_i, \quad \hat{D}_i \equiv C_{i\uparrow}^\dagger C_{i\uparrow} C_{i\downarrow}^\dagger C_{i\downarrow}, \quad (53)$$

and  $C_{i\sigma}$  and  $C_{i\sigma}^\dagger$  denote the canonical Fermionic destruction and creation operators. The unphysical states contain one or more doubly occupied states. In terms of these, the Gutzwiller projector over all sites is written as:

$$\hat{P}_G = \prod_i (1 - \hat{D}_i). \quad (54)$$

This projection operator can be introduced into a partition function to deal with unphysical states, as we show below.

The next goal (see Table 1) is to write the most efficient representation in the enlarged space of the  $t$ - $J$  model Green's functions, in terms of the canonical operators and the projection operator. As pointed out in the introduction, we note that pairs of operators that are mutual adjoints in the  $t$ - $J$  model (e.g.  $X_i^{0\sigma} = (X_i^{\sigma 0})^\dagger$ ), are allowed to be represented by operators that violate this adjoint property. The main result of this section is that this possibility leads to the most compact canonical theory; we term it the non-Hermitian theory. However we first warmup with a short summary of the more obvious Hermitian theory, which sets the stage for the main result.

### 3.1. A Hermitian canonical representation with redundancy

Projected Fermi operators distinguished by the hats can be written in a familiar construction [36]

$$\begin{aligned}\tilde{C}_{i\sigma} &= C_{i\sigma} (1 - N_{i\bar{\sigma}}) \\ \tilde{C}_{i\sigma}^\dagger &= C_{i\sigma}^\dagger (1 - N_{i\bar{\sigma}}),\end{aligned}\quad (55)$$

where  $N_{i\sigma} = C_{i\sigma}^\dagger C_{i\sigma}$ , and  $N_i = \sum_{\sigma} N_{i\sigma}$ , with the property that these conserve the number of doubly occupied sites locally:

$$[\tilde{C}_{i\sigma}, \hat{D}_i] = 0, \quad [\tilde{C}_{i\sigma}^\dagger, \hat{D}_i] = 0 \quad (56)$$

and therefore also globally i.e. with  $\hat{D}$  in place of  $\hat{D}_i$ . It implies that any Hamiltonian written in terms of these operators with hats commutes with the individual  $\hat{D}_i$  as well as the global  $\hat{D}$ , and thus conserves the local symmetry of the model. Therefore acting within the physical subspace of states, (55) provide a faithful realization of the  $X_i^{ab}$  operators as  $X_i^{0\sigma} \leftrightarrow \tilde{C}_{i\sigma}$  and  $X_i^{\sigma 0} \leftrightarrow \tilde{C}_{i\sigma}^\dagger$ , and clearly satisfies the mutual adjoint property. We are also interested in the product of two  $X$ 's in order to represent the kinetic energy term of the effective Hamiltonian below. The optimal choice is seen to be

$$X_i^{\sigma 0} X_j^{0\sigma} \leftrightarrow C_{i\sigma}^\dagger C_{j\sigma} (1 - N_{i\bar{\sigma}} - N_{j\bar{\sigma}}). \quad (57)$$

While the choice

$$X_i^{\sigma 0} X_j^{0\sigma} \leftrightarrow \tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma} \quad (58)$$

is also a faithful representation, it contains an extra term  $C_{i\sigma}^\dagger C_{j\sigma} N_{i\bar{\sigma}} N_{j\bar{\sigma}}$ , over and above (57), which is redundant since (57) already commutes with (54).

Using (57) we write a canonical expression for the Hamiltonian

$$H_{ij} \rightarrow \hat{H}_{eff} = \hat{H}_t + \hat{H}_J, \quad (59)$$

with

$$\begin{aligned}\hat{H}_t &= \hat{T}_{eff} - \mu \sum_i N_{i\sigma}, \\ \hat{T}_{eff} &= - \sum_{ij\sigma} t_{ij} C_{i\sigma}^\dagger C_{j\sigma} (1 - N_{i\bar{\sigma}} - N_{j\bar{\sigma}}),\end{aligned}\quad (60)$$

we call this as the *symmetrized kinetic energy* in view of its obvious symmetry under the exchange  $i \leftrightarrow j$ , and write  $\hat{H}_J \rightarrow \frac{1}{2} \sum_{ij} J_{ij} (\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} N_i N_j)$ , with the spin and number operators written in terms of  $C$ 's and  $C^\dagger$ 's without hats (since the occupancy of a site is unaffected by the exchange term). We easily verify that

$$[\hat{H}_{eff}, \hat{D}] = 0 = [\hat{H}_{eff}, \hat{P}_G], \quad (61)$$

therefore if we start with a state satisfying  $\hat{D}|\Psi\rangle = 0$ , i.e. in the singly occupied subspace, the resultant state  $H_{eff}|\Psi\rangle$  remains in this subspace; we do not create doubly occupied states. We note that (61) implies that the operator (54) is invariant under time evolution through  $H_{eff}$ :

$$\hat{P}_G(\tau) = \hat{P}_G(0). \quad (62)$$

The partition functional as in (11), now defined with arbitrary  $\tau_0$ :

$$Z = \text{Tr} e^{-\beta \hat{H}_{eff}} T_\tau (e^{-\hat{A}_S} \hat{P}_G(\tau_0)), \quad (63)$$

where the trace (unlike that in Eq. (11)), is over the entire canonical basis, i.e. includes doubly occupied states. For the observables as well as the source terms  $\hat{A}_S$ , we use the replacement rules:

$$X_i^{0\sigma} \rightarrow \tilde{C}_{i\sigma}, \quad X_i^{\sigma 0} \rightarrow \tilde{C}_{i\sigma}^\dagger, \quad X_i^{\sigma\sigma'} \rightarrow C_{i\sigma}^\dagger C_{i\sigma'}, \quad (64)$$

to convert arbitrary expressions involving  $X_i^{ab}$  into those with the  $\tilde{C}, \tilde{C}^\dagger$ . Note that the density or spin density type variables are replaced by the canonical operators without a hat, since these commutes with the local  $\hat{D}_i$ .

We can compute the Green's functions in the enlarged (canonical) basis from

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\frac{\left\langle T_\tau \left( e^{-\hat{A}_S} \tilde{C}_{i\sigma_i}(\tau_i) \tilde{C}_{f\sigma_f}^\dagger(\tau_f) \hat{P}_G(\tau_0) \right) \right\rangle}{\left\langle T_\tau \left( e^{-\hat{A}_S} \hat{P}_G(\tau_0) \right) \right\rangle}, \quad (65)$$

evaluated [23] at  $\hat{A}_S \rightarrow 0$ . This relation can be replaced by differentiating the partition functional (63) with the Fermi sources  $J, J^*$ . Using the commutation of  $\hat{P}_G$  or  $\hat{D}$  with all operators and (62), we are free at this stage to locate place  $\hat{P}_G$  at any specific time, without affecting the results. This formulation of the theory has parallels with the path integral representation of the electromagnetic field (QED) in the temporal gauge, where the scalar potential is chosen to be zero (i.e.  $\phi(rt) = 0$ ) [37,38]. In this case the Gauss's law condition  $\nabla \cdot \vec{E}(r, t) = 0$  needs to be imposed at each time slice. However upon using  $[H, \vec{E}] = \vec{\nabla} \times \vec{B}$ , this object commutes with the Hamiltonian  $[H, \vec{\nabla} \cdot \vec{E}] = 0$ , and therefore it suffices to impose this condition at the initial time. The situation has a clear analogy with Eq. (63), where it suffices to insert the projection operator at the initial time.

### 3.2. The hat removal rule and optimal non-Hermitian theory

The non-Hermitian theory arises when we inspect closely expressions of the type in (63), with the time  $\tau_0$  chosen as the earliest time  $0^-$ . The general argument has been given in the introduction, we consider its specific application to the present problem next. Discretizing the time variables and expanding, we obtain a series containing expression of the type

$$\text{const} \times \sum \langle i | Q_1(\tau_1) \dots Q_m(\tau_m) \hat{P}_G | i \rangle,$$

so that the first operator from the right  $Q_m(\tau_m)$  acts upon a state which is Gutzwiller projected. Now the creation operators contained in the  $Q(\tau)$ 's are defined with the hats (see (55)) ensuring that they never create doubly occupied states. Next observe that destroying a particle cannot *create* a doubly occupied site. Therefore it cannot take a projected state out of this subspace! Therefore the operator  $\tilde{C}_{i\sigma}$  can as well be replaced by the destruction operator  $C_{i\sigma}$  *without a hat*. We can iterate this argument for the next operator, which also acts on a Gutzwiller projected state, and so forth, leading to the hat removal rules. In this argument, we may replace the operator's  $Q(\tau_m)$  by any expressions involving the destruction operators as well as creation operators with hats (as in (55)), and the same argument holds. More formally we may summarize by saying that the destruction operator *conditionally commutes* with the projection operator, when right-operating on projected states:

$$[C_{i\sigma}, \hat{P}_G] \hat{P}_G = 0, \quad (66)$$

although  $[C_{i\sigma}, \hat{P}_G] \neq 0$ , as one readily checks. Thus the commutator lives in an orthogonal subspace to that spanned by the Gutzwiller projected states. This property also extends to arbitrary functions  $\hat{f}(\hat{f} \equiv \hat{f}\{C_{i\sigma}\}, \{\tilde{C}_{j\sigma'}^\dagger\})$  of the operators:

$$[\hat{f}, \hat{P}_G] \hat{P}_G = 0. \quad (67)$$

This property is just a rewriting of the important block triangularity condition of the operators noted in Eq. (8) leading to Eq. (7). We will make frequent use of this expression below.

We now turn to implementing this observation. Let us write the partition functional

$$Z = \text{Tr} e^{-\beta \hat{H}_{\text{eff}}} T_\tau \left( e^{-\hat{A}_S} \hat{P}_G(0^-) \right), \quad (68)$$

and introduce the important abbreviation for averages:

$$\langle\langle A(\tau_1) B(\tau_2) \dots \rangle\rangle \equiv \frac{1}{Z} \text{Tr} e^{-\beta \hat{H}_{\text{eff}}} T_\tau \left( e^{-\hat{A}_S} A(\tau_1) B(\tau_2) \dots \hat{P}_G(0^-) \right), \quad (69)$$

where notice that we located the projector at the *initial time*, by bringing it under the time ordering symbol.

We now state the crucial *hat-removal rule*: in all expressions of the type (68) and (69), the hats on *all destruction operators* can be removed

$$\tilde{C}_{i\sigma}(\tau) \rightarrow C_{i\sigma}(\tau), \quad (70)$$

leaving the result unchanged. Notice that this rule can also be applied to  $H_{\text{eff}}$ , and the source terms  $\hat{A}_S$  containing the destruction operators  $C_{i\sigma}$ . Note that the *creation operators* cannot be ‘un-hatted’ in this fashion since these do create a doubly occupied site. Summarizing, we can use instead of (64), the more compact non-Hermitian rule

$$X_i^{0\sigma} \rightarrow C_{i\sigma}, \quad X_i^{\sigma 0} \rightarrow \tilde{C}_{i\sigma}^\dagger = C_{i\sigma}^\dagger (1 - N_{i\bar{\sigma}}), \quad X_i^{\sigma\sigma'} \rightarrow C_{i\sigma}^\dagger C_{i\sigma'}. \quad (71)$$

We thus rewrite the sources (10) as:

$$\hat{A}_S(i, \tau) = \left( \tilde{C}_{i\sigma}^\dagger(\tau) J_{i\sigma}(\tau) + J_{i\sigma}^*(\tau) C_{i\sigma}(\tau) \right) + \mathcal{V}_i^{\sigma'\sigma}(\tau) C_{i\sigma'}^\dagger(\tau) C_{i\sigma}(\tau), \quad (72)$$

and the Green's function with imaginary time  $0 \leq \tau_i, \tau_j \leq \beta$  is therefore written as:

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle\langle C_{i\sigma_i}(\tau_i) \tilde{C}_{f\sigma_f}^\dagger(\tau_f) \rangle\rangle, \quad (73)$$

analogous to (65) but with an unprojected destruction operator. We will show below that this is the most useful and compact expression for the Green's function. To complete the description of this theory, we turn to the task of specifying the Hamiltonian, and obtain the boundary conditions on the time variables. The last task is somewhat nontrivial since the projection operator does not commute with the other operators.

### 3.3. Hamiltonian in the symmetrized and minimal theories

In order to represent the Hamiltonian, the spin operators of the exchange part  $H_J$  are unambiguously expressed in terms of the  $C_{i\sigma}$  and  $C_{i\sigma}^\dagger$  operators without hats as in (71), since they preserve the occupation of a site. For the kinetic energy we could choose to work with (60), and thereby gain some advantage of dealing with a Hermitian Hamiltonian. This leads to the equations of motion termed the *the symmetrized theory* in Ref. [6]. Alternately we can implement the hat removal rule for the kinetic energy as well:

$$\hat{T}_{\text{eff}} = - \sum_{ij\sigma} t_{ij} \tilde{C}_{i\sigma}^\dagger C_{j\sigma}. \quad (74)$$

This minimal version of the kinetic energy is clearly non-Hermitian. However, it has exactly the same action as the symmetrized version (59), when right-operating on the physical Gutzwiller projected states, as proved above. This leads to equations of motion of the *minimal theory* noted in Ref. [6] and elaborated upon in Refs. [10,9]. For completeness, we provide in Section 7.2 a brief derivation of these equations for the minimal case, using the above canonical representation, in place of the Schwinger equations.

### 3.4. Kubo–Martin–Schwinger antiperiodic boundary conditions

In working with the expressions (68), (71) and (73), we have assumed that all the times  $\tau_j$  are positive and satisfy  $0 \leq \tau_j \leq \beta$ . The Green's function (12) satisfies the Kubo–Martin–Schwinger (KMS) anti-periodic boundary conditions [39]

$$\mathcal{G}(a\tau_i = 0, b\tau_f) = -\mathcal{G}(a\tau_i = \beta, b\tau_f), \quad (75)$$

$$\mathcal{G}(a\tau_i, b\tau_f = 0) = -\mathcal{G}(a\tau_i, b\tau_f = \beta), \quad (76)$$

**Table 2**

A summary of the representations of the Green's functions. The non-Hermitian minimal theory provides the most compact set of equations of motion, which are identical to those from the Hubbard–Gutzwiller theory in the second column. The absence of the adjoint property for the non-Hermitian theory arises from the asymmetric hat removal between the destruction and creation operators in the first two rows of the last column.

	Hubbard–Gutzwiller theory	(Canonical) Hermitian theory	(Canonical) Non-Hermitian theory
Operators:	$X^{\sigma 0}$ $X^{0\sigma}$ $X_i^{\sigma\sigma'}$	$\tilde{C}_\sigma^\dagger = C_\sigma^\dagger (1 - N_{\bar{\sigma}})$ $\tilde{C}_\sigma = C_\sigma (1 - N_{\bar{\sigma}})$ $C_{i\sigma}^\dagger C_{i\sigma'}$	$\tilde{C}_\sigma^\dagger = C_\sigma^\dagger (1 - N_{\bar{\sigma}})$ $C_\sigma$ $C_{i\sigma}^\dagger C_{i\sigma'}$
Partition Functional: $Z$	$\text{Tr}_{ij} e^{-\beta H_{ij}} T_\tau \left( e^{-\hat{\lambda}s} \right)$	$\text{Tr} e^{-\beta \hat{H}_{\text{eff}}} T_\tau \left( e^{-\hat{\lambda}s} \hat{P}_G(\tau_0) \right);$ Arbitrary time $\tau_0$ ( $0 \leq \tau_0 \leq \beta$ ).	$\text{Tr} e^{-\beta \hat{H}_{\text{eff}}} T_\tau \left( e^{-\hat{\lambda}s} \hat{P}_G(0^-) \right)$
Green's function: $-\hat{g}(1, 1')$	$\langle T_\tau (e^{-\hat{\lambda}s} X_1^{0\sigma} X_1^{\sigma'0}) \rangle$	$\langle T_\tau (e^{-\hat{\lambda}s} \tilde{C}_\sigma(1) \tilde{C}_{\sigma'}^\dagger(2) \hat{P}_G(\tau_0)) \rangle$ Arbitrary time $\tau_0$ ( $0 \leq \tau_0 \leq \beta$ ).	$\langle T_\tau (e^{-\hat{\lambda}s} C_\sigma(1) \tilde{C}_{\sigma'}^\dagger(2) \hat{P}_G(0^-)) \rangle$
Remarks:	$H = H^\dagger$ in the defining representation.	Symmetrized theory $H = H^\dagger$	(i) Symmetrized theory: $\hat{H}_{\text{eff}} = H_{\text{eff}}^\dagger$ (ii) Minimal theory: $\hat{H}_{\text{eff}} \neq \hat{H}_{\text{eff}}^\dagger$

where the fixed time  $\tau_f$  ( $\tau_i$ ) in the first (second) equations is assumed to satisfy  $0 \leq \tau \leq \beta$ . These conditions are usually proven by using the cyclic invariance of the trace [24], and translates easily to the canonical representation (65), with  $\tilde{C}$  and  $\tilde{C}^\dagger$  replacing the  $X$  operators (64).

In using the non-Hermitian representation (71) as in (73), we cannot use cyclicity of trace since the operator  $\tilde{C}$  does not commute with  $P_G$ . Remarkably enough, the conditional commutativity (66) and (67) suffices to guarantee the required antiperiodicity. In physical terms these proofs follow from the observation made above, the creation operators with hats, and destruction operators (without hats) preserve a Gutzwiller projected state within that subspace.

For simplicity we present the case with sources turned off i.e.  $\mathcal{A} \rightarrow 0$ , the more general case follows by a similar argument. From the definitions of the Green's functions, Eq. (76) is true since  $\text{Tr} \left( e^{-\beta H_{\text{eff}}} C_{a\sigma}(\tau_i) [\tilde{C}_{b\sigma'}^\dagger(0), \hat{P}_G] \right)$  vanishes identically from Eq. (56).

In order to prove that Eq. (75) remains true, we need to show that the expression

$$\text{Tr} \left( e^{-\beta H_{\text{eff}}} \tilde{C}_{b\sigma'}^\dagger(\tau_f) [C_{a\sigma}(0), \hat{P}_G] \right) \quad (77)$$

vanishes, despite the non vanishing of the commutator in the expression. For this purpose, we utilize the conditional commutator (66) to write  $[C_{a\sigma}(0), \hat{P}_G] = [C_{a\sigma}(0), \hat{P}_G](\mathbb{1} - \hat{P}_G)$ . We next use cyclicity of trace and the simple identity (for any  $\hat{Q}$ ):  $\text{Tr} \left( (\mathbb{1} - \hat{P}_G) \hat{Q} \hat{P}_G \right) = 0$ , to write the required expression (77) in the form

$$\text{Tr} \left( (\hat{P}_G - \mathbb{1}) e^{-\beta H_{\text{eff}}} \tilde{C}_{b\sigma'}^\dagger(\tau_f) \hat{P}_G C_{a\sigma}(0) \right). \quad (78)$$

Using  $(\hat{P}_G)^2 = \hat{P}_G$ , we rewrite this as:

$$(\hat{P}_G - \mathbb{1}) e^{-\beta H_{\text{eff}}} \tilde{C}_{b\sigma'}^\dagger(\tau_f) \hat{P}_G = [\hat{P}_G, e^{-\beta H_{\text{eff}}} \tilde{C}_{b\sigma'}^\dagger(\tau_f)] \hat{P}_G.$$

This expression vanishes on using the conditional commutator (67), thereby proving the required result (75).

The two canonical theories providing an exact mapping of the original theory are summarized in Table 2.

#### 4. The auxiliary Green's function and the caparison function using canonical Fermions

We next discuss the rationale for decomposing the Green's function into an auxiliary Greens function and a caparison function as in Ref. [4], using a simple argument from the exact formula

(73). This important part of the theory is also encountered in Section 6. In its simplest version this decomposition can be illustrated using the minimal theory, where the averages are defined as in Eq. (68), with the projection operator pinned at the initial time. We recall the Green's function from Eq. (73)  $\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle\langle C_{i\sigma_i}(\tau_i)\tilde{C}_{f\sigma_f}^\dagger(\tau_f)\rangle\rangle$ , with the averages from Eq. (69). Expanding the  $\tilde{C}^\dagger$  operator this becomes

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle\langle C_{i\sigma_i}(\tau_i)C_{f\sigma_f}^\dagger(\tau_f)\rangle\rangle + \langle\langle C_{i\sigma_i}(\tau_i)C_{f\sigma_f}^\dagger(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle. \quad (79)$$

We next define the auxiliary Green's function as:

$$\mathbf{g}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) = -\langle\langle C_{i\sigma_i}(\tau_i)C_{j\sigma}^\dagger(\tau_j)\rangle\rangle, \quad (80)$$

and regarding the spin, space and time indices as matrix indices with a matrix inverse  $\mathbf{g}^{-1}$ . By separating the disconnected and connected parts ( $_c$ ) of the second term in (79) we write

$$\langle\langle C_{i\sigma_i}(\tau_i)C_{f\sigma_f}^\dagger(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle = -\mathbf{g}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f)\langle N_{f\bar{\sigma}_f}(\tau_f)\rangle + \langle\langle C_{i\sigma_i}(\tau_i)C_{f\sigma_f}^\dagger(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle_c. \quad (81)$$

The connected part is written in terms of a second self energy  $\Psi$  defined as

$$\Psi_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = \mathbf{g}_{\sigma_i\sigma_k}^{-1}(i\tau_i, \mathbf{k}\tau_k) \times \langle\langle C_{\mathbf{k}\sigma_k}(\tau_k)C_{f\sigma_f}^\dagger(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle_c, \quad (82)$$

and assembling these we rewrite (79) as the product relation [4]

$$\begin{aligned} \mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) &= \mathbf{g}_{\sigma_i\sigma_k}(i\tau_i, \mathbf{k}\tau_k)\mu_{\sigma_k\sigma_f}(\mathbf{k}\tau_k, f\tau_f), \\ \mu_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) &= \delta(\text{if}) (1 - \langle N_{\bar{\sigma}_i}(\tau_i)\rangle) + \Psi_{\sigma_i\sigma_f}(i\tau_i, f\tau_f). \end{aligned} \quad (83)$$

There is a slight ambiguity in defining the two objects  $\mathbf{g}$  and  $\mu$ , since we have the freedom of adding a common function to the two parts of Eq. (79) that cancels out in the physical Greens function. Apart from this, we expect that the two objects in Eq. (83) are exactly equivalent to the auxiliary Greens function and the caparison factor in Eqs. (17)–(19) as found from the Schwinger method.

We observe from the expression (82) that if the averages are (temporarily) computed in a standard Feynman Dyson theory, then  $\Psi$  is essentially the self energy of a Hubbard type model, made dimensionless by dropping an explicit interaction constant  $U$ . Indeed this is the key observation made in Ref. [4], on the basis of the  $\lambda$  expansion, where the two self energies are argued to be generically Fermi liquid-like and similar to each other. An energy scale ( $\Delta$ ) emerges from a ratio of their imaginary parts, and controls the significant asymmetry seen in the spectral functions.

## 5. The $\lambda$ -Fermions

A natural question is whether Eq. (16), explicitly containing the parameter  $\lambda$ , can arise in a microscopic theory where  $\lambda$  enters in a fundamental way, as opposed to the “engineering approach” in Section 2.3. A set of  $\lambda$ -Fermi operators are defined below, as generalized version of the non-Hermitian representation (71) with a parameter  $\lambda \in [0, 1]$  providing a continuous interpolation between the free Fermi and extremely correlated limits:

$$\begin{aligned} X_i^{\sigma 0}(\lambda) &\rightarrow C_{i\sigma}^\dagger(1 - \lambda C_{i\bar{\sigma}}^\dagger C_{i\bar{\sigma}}) \\ X_i^{0\sigma}(\lambda) &\rightarrow C_{i\sigma} \\ X_i^{\sigma\sigma'}(\lambda) &\rightarrow C_{i\sigma}^\dagger C_{i\sigma'}. \end{aligned} \quad (84)$$

Clearly  $\lambda = 0$  gives us back the canonical Fermion operators, whereas  $\lambda = 1$  gives the Gutzwiller projected Hubbard  $X$  operators [3] as in (71), provided the states are Gutzwiller projected. A feature of this representation is the loss of the adjoint property, i.e.  $(X_i^{\sigma 0}(\lambda))^\dagger \neq X_i^{0\sigma}(\lambda)$ , unless  $\lambda = 0$ .

These operators satisfy a  $\lambda$  dependent (graded) Lie algebra with fundamental brackets that are partly Fermionic and partly Bosonic. Using the canonical anticommutation relations of the  $C, C^\dagger$  operators, we work out the fundamental Fermionic bracket:

$$\{X_i^{0\sigma_i}(\lambda), X_j^{\sigma_j 0}(\lambda)\} = \delta_{ij}\{\delta_{\sigma_i\sigma_j} - \lambda \sigma_i\sigma_j X_i^{\bar{\sigma}_i\bar{\sigma}_j}(\lambda)\}. \quad (85)$$

We next evaluate the fundamental Bosonic bracket

$$[X_i^{0\sigma_i}(\lambda), X_j^{\sigma_j\sigma_k}(\lambda)] = \delta_{ij}\delta_{\sigma_i\sigma_j}X_i^{0\sigma_k}(\lambda) \quad (86)$$

$$[X_i^{\sigma_i 0}(\lambda), X_j^{\sigma_j\sigma_k}(\lambda)] = -\delta_{ij}\delta_{\sigma_i\sigma_k}X_i^{\sigma_j 0}(\lambda). \quad (87)$$

Here (87) requires a brief calculation [40] invoking the Pauli principle vanishing of  $C_\sigma^\dagger C_\sigma^\dagger \rightarrow 0$ . On the other hand (86) is elementary, due to the absence of  $\lambda$  in both sides of the equation. At  $\lambda = 1$  these reduce to the relevant subset of the Hubbard algebra [3] found from the fundamental definition  $X_i^{ab} = |a\rangle\langle b|$ .

The representation (84) does not at general  $\lambda$  reproduce the “half bracket”, or product relations expected for projection operators. We find that

$$X_i^{\sigma 0}(\lambda)X_i^{0\sigma'}(\lambda) \neq X_i^{\sigma\sigma'}(\lambda), \quad (88)$$

$$X_i^{0\sigma'}(\lambda)X_i^{\sigma 0}(\lambda) \neq X_i^{00}\delta_{\sigma\sigma'}. \quad (89)$$

The exceptions are at  $\lambda = 0$ , where it is trivially true, and non trivially at  $\lambda = 1$ , where Gutzwiller projection of the allowed states does restore this property when right-operating on the projected states. In the Green’s functions below, we will equate the averages of both sides of Eq. (88). This equality of the averages acts as the number constraint and fixes the chemical potential  $\mu$ . In doing so, the average of Eq. (89) is not constrained and takes on a suitable value determined by the anticommutation relation (85).

This representation can be used to define a many-body problem where the  $\lambda$  dependent EOMs for the Green’s functions constructed from (84) can be written down. Observe that the EOMs for the Green’s functions only require the use of (85) and the Heisenberg equations of motion, and in turn these arise from the basic Lie commutators (anticommutators) of the type given in (86) and (87). The calculation does not ever require the use of product relations of the type (88). It then follows that we can replace the  $t$ - $J$  Hamiltonian and the operators in the original theory by their  $\lambda$ -versions, i.e. replacing  $X_i^{ab} \rightarrow X_i^{ab}(\lambda)$ , and thereby obtain equations that yield (16). This procedure then provides a (continuous) interpolation between the free Fermi and extremely correlated regimes by varying  $\lambda$  from 0 to 1. Let us first demonstrate this by a brief calculation.

### 5.1. The $\lambda$ -Fermion theory equations of motion

Using the  $\lambda$  Fermions, we define the Green’s function as

$$\mathcal{G}_{\sigma_i\sigma_f}^{(\lambda)}(i\tau_i, f\tau_f) = -\langle T_\tau X_i^{0\sigma_i}(\tau_i, \lambda) X_f^{\sigma_f 0}(\tau_f, \lambda) \rangle_{(\lambda)} \quad (90)$$

where with arbitrary  $\hat{A}$

$$\begin{aligned} \langle \hat{A} \rangle_\lambda &\equiv -\frac{\text{Tr} e^{-\beta H_{\text{eff}}(\lambda)} T_\tau \left( e^{-\hat{\mathcal{A}}_S(\lambda)} \hat{A} \right)}{Z(\lambda)}, \\ Z(\lambda) &= \text{Tr} e^{-\beta H_{\text{eff}}(\lambda)} T_\tau \left( e^{-\hat{\mathcal{A}}_S(\lambda)} \right). \end{aligned} \quad (91)$$

In this expression  $H_{\text{eff}}(\lambda)$  is given by Eq. (92) and  $\hat{\mathcal{A}}_S(\lambda)$  is obtained from (10), with the replacement  $X_i^{ab} \rightarrow X_i^{ab}(\lambda)$ :

$$\begin{aligned} H_{\text{eff}}(\lambda) &= -\sum_{ij} t_{ij} X_i^{\sigma 0}(\lambda) X_j^{0\sigma}(\lambda) - \mu \sum_i N_{i\sigma} + \lambda \frac{1}{2} \sum_{ij} J_{ij} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} N_{i\sigma} N_{j\sigma'} \right) \\ &\quad + u_0 \lambda \sum_i N_{i\uparrow} N_{i\downarrow}, \end{aligned} \quad (92)$$

where  $u_0$  is now the “second chemical potential”. The scaling of the  $J$  term with  $\lambda$  is optional, and done here so that we obtain the Fermi gas at  $\lambda = 0$ . Using Eq. (84), we see that this Hamiltonian is linear in  $\lambda$  and interpolates between the free Fermi gas and the fully interacting model, when acting on suitably projected states. The equation of motion of  $\mathcal{G}^{(\lambda)}$  can be obtained using the commutation relations (85)–(87), the calculation is parallel to that in Appendix A. In brief, Eqs. (132) and (133) are unchanged by working with  $X(\lambda)$ 's, and in place of Eq. (137) we obtain

$$\begin{aligned} \mathbf{g}_{0,\sigma_i,\sigma_j}^{-1}(i\tau_i, j\tau_j) \mathcal{G}_{\sigma_j\sigma_f}^{(\lambda)}(j\tau_j, f\tau_f) &= \delta(\tau_i - \tau_f) \delta_{ij} (1 - \lambda \gamma_{\sigma_i\sigma_f}(i\tau_i)) \\ &\quad - \lambda \sum_{j\sigma_j} t_{ij}(\sigma_i\sigma_j) \left\langle T_\tau \left( X_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i) X_j^{0\sigma_j}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \right) \right\rangle_{(\lambda)} \\ &\quad + \frac{1}{2} \sum_{j\sigma_j} J_{ij}(\sigma_i\sigma_j) \left\langle T_\tau \left( X_j^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i) X_i^{0\sigma_j}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \right) \right\rangle_{(\lambda)} \\ &\quad - \frac{1}{2} \lambda u_0 \sum_{\sigma_j} (\sigma_i\sigma_j) \left\langle T_\tau \left( X_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i) X_i^{0\sigma_j}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \right) \right\rangle_{(\lambda)}, \end{aligned} \quad (93)$$

where the  $\lambda$  dependence of the  $X$  operators is implicit. The higher order Green's functions may be expressed as functional derivatives with respect to the Bosonic source  $\mathcal{V}$ , in the same fashion as in Appendix A. The exchange term  $J_{ij}$  does not pick up a factor of  $\lambda$  through the EOM since it conserves double occupancy. We can choose to additionally scale it with  $\lambda$  as  $J_{ij} \rightarrow \lambda J_{ij}$ , so that at  $\lambda = 0$  we obtain the Fermi gas. This choice seems reasonable in the liquid phase of the electrons, in other phases it is easy enough to recover from this scaling if needed. To save writing the  $u_0$  term is absorbed as  $J_{ij} \rightarrow J_{ij} - u_0 \delta_{ij}$ , with this the resulting equation is

$$\begin{aligned} &\left( \mathbf{g}_{0,\sigma_i,\sigma_j}^{-1}(i\tau_i, j\tau_j) - \lambda \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) - \lambda Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j) \right) \times \mathcal{G}_{\sigma_j\sigma_f}^{(\lambda)}(j\tau_j, f\tau_f) \\ &= \delta_{if} \delta(\tau_i - \tau_f) (\delta_{\sigma_i\sigma_f} - \lambda \gamma_{\sigma_i\sigma_f}(i\tau_i)). \end{aligned} \quad (94)$$

The constitutive relation determining the chemical potential is taken as

$$\begin{aligned} n_{i\sigma} &= \langle X_i^{\sigma 0}(\tau, \lambda) X_i^{0\sigma}(\tau^-, \lambda) \rangle_{(\lambda)}, \\ &= \mathcal{G}_{\sigma\sigma}^{(\lambda)}(i, \tau^-, \tau), \end{aligned} \quad (95)$$

rather than  $n_{i\sigma} = \langle X_i^{\sigma 0}(\tau, \lambda) X_i^{0\sigma}(\tau, \lambda) \rangle_{(\lambda)}$  [41]. This limiting process corresponds to enforcing the half bracket relation Eq. (88) as an average. (95) is exact for the fully projected operators where  $\lambda = 1$ , while for other values of  $\lambda$  it is guided by the requirement of continuity in  $\lambda$ . In the same spirit, we express the function  $\gamma$  in Eq. (16) as

$$\gamma_{\sigma\sigma'}(i\tau) = \sigma\sigma' \mathcal{G}_{\bar{\sigma}\bar{\sigma}'}(i\tau^-, i\tau), \quad (96)$$

while the direct computation using Eq. (85) would yield identical times, rather than the split times in Eq. (96). An iteration scheme for solving these equations using ideas of the skeleton expansion is detailed in Refs. [6,7], and hence we skip the details.

A very simple example can be given to illustrate the role of  $\lambda$  and  $u_0$ , where the skeleton expansion is avoided. Let us consider the atomic limit of the  $\lambda$ -Fermions theory. We consider the Hamiltonian  $H_0 = -\mu \sum_{\sigma} N_{\sigma} + \lambda u_0 N_{\uparrow} N_{\downarrow}$  with  $u_0 \geq 0$ . The Green's function in Eq. (90) can be calculated easily using the EOM technique as:

$$\mathcal{G}(i\omega_n) = \frac{1 - n_{\bar{\sigma}}}{i\omega_n + \mu} + \frac{(1 - \lambda)n_{\bar{\sigma}}}{i\omega_n + \mu - \lambda u_0}. \quad (97)$$

At  $\lambda = 0$  or 1, this yields the exact atomic limit result, and provides a smooth interpolation between these limits. The positive energy pole at  $\lambda u_0 - \mu$  does not contribute to the occupancy for a sufficiently large  $u_0$  and low  $T$ . In the more realistic case with non zero hopping discussed in Refs. [6,7], the energy  $u_0$  is non-trivially fixed by a second sum rule (20), and the iteration procedure is more

complex, involving the skeleton expansion. While the atomic limit example is quite explicit, it does not generalize in any simple way to the case of finite hopping, and therefore is somewhat trivial.

We next remark on some consequences of the  $\lambda$  expansion in the intermediate region  $\lambda < 1$ , that follow from general principles. Let us first summarize the high frequency limit of the Green's functions. When  $i\omega_n \rightarrow \infty$ , the local Green's function falls off as  $\mathcal{G}(i\omega_n) \rightarrow a_G/i\omega_n$ . Here the constant  $a_G = \langle \{\hat{C}, \hat{C}^\dagger\} \rangle$ , with  $\hat{C}, \hat{C}^\dagger$  the two appropriate operators involved in  $\mathcal{G}$ , it is a measure of the total fraction of states. In the Hubbard model  $a_G = 1$ , since we have canonical operators, and implicitly  $|\omega_n| \gg U$  as well. However for the  $t$ - $J$  model we obtain  $a_G = (1 - n/2)$ , with a net deficit of  $n/2$  states from the Hubbard model. This deficit is accounted for by the upper Hubbard band that is ignored in the  $t$ - $J$  model. The lower Hubbard band thus contains a fraction  $1 - n/2$  of all the states, of which we account for  $n/2$  as the occupied states (with two spin projections available), and  $1 - n$  as the unoccupied part of the lower Hubbard band. These  $1 - n$  states are available for charge excitations in the  $t$ - $J$  model, and freeze out towards the insulating limit. Summarizing, in this picture we have  $n/2$  occupied and  $1 - n$  unoccupied states in the lower Hubbard band, and  $n/2$  states at high energy of  $O(U)$ .

In the  $\lambda$  expansion, from Eq. (16) we have  $a_G = 1 - \lambda\gamma$ , where  $\gamma$  is further expanded in  $\lambda$ . On enforcing the number sum rule (95) we find that the effective number of states described by this theory can be decomposed into  $n/2$  occupied states and  $(1 - n) + (n/2 - \lambda\gamma)$  unoccupied states. These are to be taken as the low energy sector of a fiduciary Hamiltonian. The fraction  $(n/2 - \lambda\gamma)$  vanishes only when  $\lambda = 1$  and is otherwise an unspecified surplus of states in the low energy sector. An unbalanced state count of this type is to be expected when we have non-unitary evolution. Indeed in the second order  $\lambda$  expansion carried out numerically, a similar excess of states is found [7, Section (2), last paragraph]. Another related consequence is that the spectral function positivity, requiring unitary evolution, can no longer be guaranteed in finite orders of the  $\lambda$  expansion. This feature is well recognized in Ref. [4], where it is noted that the occupied states with  $\omega < 0$  are essentially unaffected by this problem.

## 6. Analogy with the Dyson–Maleev representation of spin operators

The non-Hermitian representation in Eq. (71) of the Gutzwiller projected electron operators, when used with the averaging in Eq. (69), was shown in Section 3.2 to provide an exact mapping of the  $t$ - $J$  model. Reflecting on this result, the author realized recently that the mapping Eq. (71) is the Fermionic analog of the Dyson–Maleev representation for spin operators [16,17], used to understand spin wave interactions in magnets (see Table 3).

With the advantage of hindsight, this connection seems natural. The Gutzwiller projected electronic  $X^{ab}$  operators defined by Hubbard [3], generate a non canonical algebra of Fermions that is (partly) given in Eqs. (85)–(87) with  $\lambda = 1$ . On the other hand the spin operators provide the best studied non canonical Bosonic algebras. The spins are not quite Bosons, they are equivalent to “hard core” Bosons—with infinite on site repulsion, in parallel to the infinite  $U$  in the extremely correlated electron problem. In order to avoid dealing with the infinite energy of the hard core, several other representations of spins were invented, such as the Holstein Primakoff method [44]. Dyson's use of a non-Hermitian representation provides the most compact canonical description of the spin operators. In fact it is analogous to the non-Hermitian mapping of the Fermionic Gutzwiller problem in Eq. (71).

Dyson's representation, later streamlined by Maleev [17], may be written with  $n_i = b_i^\dagger b_i$  as

$$\begin{aligned} S_i^+ &= (2s) b_i^\dagger \left(1 - \frac{n_i}{2s}\right) \\ S_i^- &= b_i \\ S_i^z + s &= n_i, \end{aligned} \tag{98}$$

where  $\vec{S}_i \cdot \vec{S}_i = s(s + 1)$  and  $b_i, b_i^\dagger$  are canonical Bose operators. The Boson vacuum state  $b_i|vac\rangle = 0$  is mapped as  $|vac\rangle \leftrightarrow |\downarrow, \downarrow, \downarrow \dots \downarrow\rangle$ , so that the action of  $b_i^\dagger$  creates spin reversals. Their number is cut off such that  $n_i \leq (2s)$ , thereby defining the physical states. Under these conditions Eq. (98) is

**Table 3**

A comparison of the Dyson–Maleev representation for spins and the non-Hermitian representation (71) for two component Fermions  $\sigma = \pm 1$  with  $\bar{\sigma} = -\sigma$ . At  $\lambda = 1$  the Fermion mappings provide a faithful representation of Gutzwiller projected Fermi operators  $X_i^{ab}$ , acting to the right on states with single occupancy, since their action produces states that remain in this space. The representation is non self adjoint, i.e. its left operation on Dirac bra states is not faithful. The situation has an exact parallel in the Dyson–Maleev representation. The Dyson projection operator  $\hat{P}_D$  for integer  $2s$  and the Gutzwiller projection operator  $\hat{P}_G$  at  $\lambda = 1$ , play a similar role in filtering out unphysical states. The role of the parameter  $\lambda$  away from 0, 1 is similar to that of  $\frac{1}{2s}$ , extending the Dyson–Maleev representation to spin values that are neither integer or half integer. The last three rows show the auxiliary Green’s function, the caparison function and the second self energy in terms of the Bosons from Eqs. (100) and (101). These follow from the work of Harris, Kumar, Halperin and Hohenberg [18] adapted to the ferromagnet. The corresponding Fermionic objects are discussed in Section 4 and detailed in Eqs. (82) and (83).

	Spins: The Dyson–Maleev mapping		Fermions: The non-Hermitian mapping	
Destruction operator	$S_i^-$	$b_i$	$X_i^{0\sigma}$	$C_{i\sigma}$
Creation operator	$S_i^+$	$(2s) b_i^\dagger (1 - \frac{n_i}{2s})$	$X_i^{\sigma 0}$	$C_{i\sigma}^\dagger (1 - \lambda N_{i\bar{\sigma}})$
Density operator(s)	$S_i^z + s$	$n_i = b_i^\dagger b_i$	$X_i^{\sigma\sigma'}$	$C_{i\sigma}^\dagger C_{i\sigma'}$
Projection operator	$\hat{P}_D$	$\prod_i \{ \sum_{m=0}^{2s} \delta_{n_i, m} \}$	$\hat{P}_G$	$\prod_i (1 - N_{i\uparrow} N_{i\downarrow})$ , for $\lambda = 1$
Vacuum	$ \downarrow\downarrow\dots\downarrow\rangle$	$ 00\dots 0\rangle$	$ \text{Vac}\rangle$	$ 00\dots 0\rangle$
Small parameter & its range	$\frac{1}{2s}$	$\frac{1}{2s} \in [0, 1]$	$\lambda$	$\lambda \in [0, 1]$
Auxiliary Green’s function		$\mathbf{g}(i, j) = -\langle\langle b_i b_j^\dagger \rangle\rangle$		$\mathbf{g}(i, j) = -\langle\langle C_{i\sigma} C_{j\sigma}^\dagger \rangle\rangle$
Caparison function		$\mu(i, j) = \delta_{ij} (1 - \frac{1}{2s} \langle n_j \rangle) + \frac{1}{2s} \Psi(i, j)$		$\mu(i, j) = \delta_{ij} (1 - \lambda \gamma) + \lambda \Psi(i, j)$
Second Self energy $\Psi$		$\Psi(i, j) = \mathbf{g}^{-1}(i, \mathbf{a}) \langle\langle b_{\mathbf{a}} b_j^\dagger n_j \rangle\rangle_c$		$\Psi(i, j) = \mathbf{g}^{-1}(i, \mathbf{a}) \langle\langle C_{\mathbf{a}\sigma} C_{j\sigma}^\dagger N_{j\bar{\sigma}} \rangle\rangle_c$

shown to provide a faithful representation of the angular momentum operators, when right-operating on physical states. Under the action of the operators in (98), the physical states form an invariant subspace of the extended Bose Hilbert space, and are selected by projection. The Dyson projection operator  $\hat{P}_D$  acts on the Bose state space and leaves the physical states unchanged while annihilating states with  $n_i > (2s)$ .

It is now evident that the Dyson–Maleev representation has a strong formal similarity to the minimal representation (71). The Dyson projector  $\hat{P}_D$  plays a role parallel to that of the Gutzwiller projector  $\hat{P}_G$  in (71) in our theory. The parallel further deepens in the path integral representation of the Fermions that we discuss below. The interesting work of Douglass [42], following Langer’s [43] path integral program for Bosons, employs the projection operator  $\hat{P}_D$  in the same spirit to our usage below.

The work of Harris, Kumar, Halperin and Hohenberg (HKHH) [18] extended Dyson’s method to two sublattice antiferromagnets, and provided a non trivial generalization to study the lifetime of the excitations. Details of the ECFL formalism turn out to have points of overlap with those in HKHH that are worth noting. In particular HKHH decompose the physical Green’s function into a space time convolution of two parts. These parts are precisely the Bosonic analogs of the ECFL breakup of the physical Green’s function, into an auxiliary Green’s function  $\mathbf{g}(k)$  and a caparison function  $\mu(k)$ , as detailed in Ref. [4] and in Section 4.

The computation of the Green’s function by HKHH [18] was carried out for the two sublattice antiferromagnet. In order to avoid dealing with the added complexity of the two sublattice problem, we translate their method to the Dyson problem of the dynamical Green’s function of the ferromagnet. We use a notation that brings out the close parallel with the product *ansatz* used in ECFL [4].

The calculation, paraphrasing that of HKHH, proceeds as follows. In order to compute the imaginary time Green’s function  $\mathcal{G}(i, j) = -\langle\langle S_i^- S_j^+ \rangle\rangle$  with the shorthand spacetime notation  $i \equiv (r_i, \tau_i)$ , the repeated index summation (integration) convention and denoting the averages as  $\langle\langle Q \rangle\rangle = \text{Tr}(e^{-\beta H})$

$T_\tau QP_D)/\text{Tre}^{-\beta H}P_D$ , we write from (98)

$$\frac{1}{2s} \mathcal{G}(i, j) = - \left\langle\left\langle b_i b_j^\dagger \left(1 - \frac{1}{2s} n_j\right) \right\rangle\right\rangle. \quad (99)$$

Separating out the disconnected part we write  $\langle\langle b_i b_j^\dagger n_j \rangle\rangle = \langle\langle b_i b_j^\dagger \rangle\rangle \langle n_j \rangle + \langle\langle b_i b_j^\dagger n_j \rangle\rangle_c$ , and defining the auxiliary Green's function  $\mathbf{g}(i, j) = -\langle\langle b_i b_j^\dagger \rangle\rangle$  as well as its inverse through  $\mathbf{g}(i, \mathbf{k}) \mathbf{g}^{-1}(\mathbf{k}, j) = \delta(i, j)$ , we arrive at

$$\frac{1}{2s} \mathcal{G}(i, j) = \mathbf{g}(i, j) \left(1 - \frac{1}{2s} \langle n_j \rangle\right) + \frac{1}{2s} \mathbf{g}(i, \mathbf{k}) \Psi(\mathbf{k}, j), \quad (100)$$

$$\Psi(i, j) = \mathbf{g}^{-1}(i, \mathbf{a}) \langle\langle b_{\mathbf{a}} b_j^\dagger n_j \rangle\rangle_c. \quad (101)$$

We use a notation with sums over repeated bold indices everywhere. We can rewrite (100) as a convolution of the auxiliary Green's function  $\mathbf{g}$  and a caparison function  $\mu$ , in the form  $\frac{1}{2s} \mathcal{G}(i, j) = \mathbf{g}(i, \mathbf{k}) \mu(\mathbf{k}, j)$ , where  $\mu(i, j) = \delta_{ij} \left(1 - \frac{1}{2s} \langle n_j \rangle\right) + \frac{1}{2s} \Psi(i, j)$ . The auxiliary Green's function is defined in terms of its own self energy  $\Phi$  through the usual Dyson equation  $\mathbf{g}^{-1}(i, j) = \mathbf{g}_0^{-1}(i, j) - \Phi(i, j)$ . Thus the physical Green's function  $\mathcal{G}$  is determined in terms of the two self energies  $\Phi(k, \omega)$  and  $\Psi(k, \omega)$ . Written in  $(k, i\omega)$  space, this is identical to the functional form in ECFL theory Eq. (21)!

The corresponding Fermionic objects are discussed in Section 4 and detailed in Eqs. (82) and (83). On comparing the two we recognize that the structure of Eqs. (100) and (101) is the exact parallel of the ECFL theory for the Green's function written in the notation of Ref. [4]. In the HKHH paper, the objects evaluated amount to these two ECFL self energies, by the correspondence  $\Psi(k, \omega) \leftrightarrow \Lambda(k, \omega)$  (see [18, Eq. (C10)]), and  $\Phi(k, \omega) \leftrightarrow \Sigma(k, \omega)$  (see [18, Eq. (2.22)]). It is worth noting further that the role of the parameter  $\lambda$  in the ECFL theory is in close parallel to that of  $\frac{1}{2s}$  in the magnon problem. Expansions in these two “small parameters” serve to organize the calculations.

The product ansatz in ECFL [4,6] was originally arrived at in Ref. [4] by analyzing the Schwinger equations and insisting on a canonical Green's function to be factored out from the physical  $\mathcal{G}$ . The calculation of HKHH, on the other hand, was through a different route using insights from the Feynman diagrams applied to the four Boson operators in (99). It is satisfying that the two independent calculations, one for Gutzwiller projected Fermions and the other for hard-core Bosons, lead to such a close parallel, expressed most naturally in the twin self energy representation (100) and (101).

A few additional comments on the role of the projection operator in the two problems are relevant here. Dyson demonstrated in his non-Hermitian representation that magnon interactions at low temperatures lead to  $T^4$  type corrections to the magnetization of the ideal spin wave theory. He argued that the projection operator  $\hat{P}_D$  is largely irrelevant in the ferromagnet, and provided an estimate of corrections to the low  $T$  behavior arising from this neglect. For the antiferromagnet, HKHH similarly argued that the projector is unimportant at low  $T$ , and gave an estimate of the expected corrections. The corrections are larger than in the ferromagnet, and yet smaller than most quantities of interest at low  $T$ . The density of excitations is small at low  $T$  in the magnetic problem, and thus provides a basis for ignoring the projection operator. However in the Fermion problem studied here, the particle density is never too small in the interesting regime, and hence the projection operator must be respected. Interestingly enough, the projector does not explicitly appear in the Schwinger EOM (16), but it does determine the choice of the correct constitutive relation (14). Thus the projection operator plays a significant role in enforcing the Luttinger–Ward theorem [45] for the volume of the Fermi surface.

Another major difference between the Fermionic and the spin problems is the role of the second Lagrange multiplier  $u_0$ , when the parameter  $\lambda < 1$ . In the Fermi problem, it is essential to change the Hamiltonian by adding the term  $\lambda u_0 \sum_i N_{i\uparrow} N_{i\downarrow}$ , in addition to replacing the projected  $X_i^{ab}$  by  $X_i^{ab}(\lambda)$ . This is required in order to satisfy the shift identities, and as explained in Ref. [6], the parameter  $u_0$  is fixed by a number sum rule on the auxiliary Green's function. The problem of magnetic excitations does not have a counterpart to this term. However, we can imagine extending the Dyson–Maleev and HKHH formalism to an extremely correlated Bose liquid with a fixed number of Bosons, e.g.  ${}^4\text{He}$  on a suitable substrate giving rise to a lattice model with hard core repulsion. In such a case, a corresponding theory parallel to ECFL can be developed, requiring both the shift identities and a second Lagrange multiplier  $u_0$  disfavoring multiple occupancy to satisfy these.

## 7. Path integrals

### 7.1. Canonical path integral representation

We now introduce path integrals to represent the partition functional (68), wherein the operators are replaced by anticommuting c-numbers, i.e. the Grassmann variables. We will keep the discussions to a minimum since excellent references can be consulted for details [46–49]. We map the operators as  $C_{i\sigma} \rightarrow c_{i\sigma}$ ,  $\tilde{C}_{i\sigma} \rightarrow \tilde{c}_{i\sigma} \equiv c_{i\sigma}(1 - c_{i\bar{\sigma}}^* c_{i\bar{\sigma}})$ ,  $C_{i\sigma}^\dagger \rightarrow c_{i\sigma}^*$ ,  $\tilde{C}_{i\sigma}^\dagger \rightarrow \tilde{c}_{i\sigma}^* \equiv c_{i\sigma}^*(1 - c_{i\bar{\sigma}}^* c_{i\bar{\sigma}})$ . The time dependence is dealt with using a standard Trotter decomposition of the non commuting pieces [49]. Handling the Gutzwiller projector is discussed below and in Appendix C. It is understood that when the Trotter index  $M$  is finite, we have a discretized time representation, so that when  $M \rightarrow \infty$ , we obtain the continuous time path integrals. We work initially with the discrete time version since somewhat subtle identities such as the Pauli principle and the Gutzwiller projection identities can be verified explicitly. We now write the partition functional  $Z$  (68), in terms of Grassmann variables at discrete times  $c_{i\sigma}(\tau_j)$  and  $c_{i\sigma}^*(\tau_j)$ , and a global integration over all Grassmann variables with the conventional definition [49]:

$$Z^{(M)}[J^*, J, \mathcal{V}] = \int_{\mathcal{C}} P_G(\tau_1, \tau_0) e^{-\mathcal{A}_{Tot}^{(M)}},$$

$$\mathcal{A}_{Tot}^{(M)} = \mathcal{A}_0^{(M)} + \mathcal{A}_S^{(M)} + \mathcal{A}_t^{(M)} + \mathcal{A}_J^{(M)}. \quad (102)$$

We detail the various contributions next; the free Fermi term is given by

$$\frac{1}{\Delta\tau} \mathcal{A}_0^{(M)} = \sum_{j=0}^{M-1} [c_{i\sigma}^*(\tau_{j+1}) \delta_{\tau_j} c_{i\sigma}(\tau_j) - \mu n_{j\sigma}(\tau_j)], \quad (103)$$

with the finite difference operator  $\delta_{\tau_j}$  defined through

$$\delta_{\tau_j} F(\tau_j) \equiv \frac{1}{\Delta\tau} \{F(\tau_{j+1}) - F(\tau_j)\}. \quad (104)$$

As  $M \rightarrow \infty$ , we note that  $\delta_{\tau_j}$  reduces to the derivative operator  $\partial_{\tau}$ , and we obtain the integral  $\mathcal{A}_0 = \int_0^\beta d\tau c_{i\sigma}^*(\tau)(\partial_{\tau} - \mu)c_{i\sigma}(\tau)$ , and in that limit  $Z^{(M)} \rightarrow Z[J^*, J, \mathcal{V}]$ . The source term  $\mathcal{A}_S^{(M)}(\tau_{j+1}, \tau_j)$  obtained from (72) is given by

$$\mathcal{A}_S^{(M)} = \sum_i [\tilde{c}_{i\sigma}^*(\tau_{j+1}) J_{i\sigma}(\tau_{j+1}) + J_{i\sigma}^*(\tau_{j+1}) c_{i\sigma}(\tau_j)] + [\mathcal{V}_i^{\sigma'\sigma}(\tau_{j+1}) c_{i\sigma'}^*(\tau_{j+1}) c_{i\sigma}(\tau_j)]. \quad (105)$$

As in (72), the projected variable with a hat appears in the creation operator and nowhere else in this expression. The Hamiltonian (59) gives rise to two parts of the action. The hopping term is given by

$$\mathcal{A}_t^{(M)} = \Delta\tau \sum_j T_{eff}(\tau_j) \rightarrow \int_0^\beta d\tau T_{eff}(\tau), \quad (106)$$

with  $T_{\text{eff}}$  from Eq. (60) or Eq. (74):

$$\begin{aligned} T_{\text{eff}}^{\text{Sym}}(\tau_j) &= - \sum_{lm\sigma} t_{lm} c_{l\sigma}^*(\tau_{j+1}) c_{m\sigma}(\tau_j) \times (1 - n_{l\bar{\sigma}}(\tau_j) - n_{m\bar{\sigma}}(\tau_j)), \\ T_{\text{eff}}^{\text{Min}}(\tau_j) &= - \sum_{lm\sigma} t_{lm} c_{l\sigma}^*(\tau_{j+1}) c_{m\sigma}(\tau_j) \times (1 - n_{l\bar{\sigma}}(\tau_j)), \end{aligned} \quad (107)$$

where (107) corresponds to the symmetrized theory of (64) and (107) to the minimal version of (71). The exchange part of the action is given by

$$\begin{aligned} \mathcal{A}_j^{(M)} &= \Delta\tau \sum_j H_j(\tau_j) \rightarrow \int_0^\beta d\tau H_j(\tau), \\ H_j(\tau_j) &\equiv -\frac{1}{4} \sum_{lm} J_{lm} \sigma_1 \sigma_2 \times c_{l\sigma_1}^*(\tau_{j+1}) c_{m\bar{\sigma}_1}^*(\tau_{j+1}) c_{m\bar{\sigma}_2}(\tau_j) c_{l\sigma_2}(\tau_j). \end{aligned} \quad (108)$$

Where possible we simplify the notation by dropping the superscript  $M$ ; most expressions provide sufficient context for this and there should be no confusion. Thus we will write  $\mathcal{G}_{\sigma\sigma'}^{(M)}(a\tau_i, b\tau_f) \rightarrow \mathcal{G}_{\sigma\sigma'}(a\tau_i, b\tau_f)$  and  $Z^{(M)} \rightarrow Z$  etc. below. When no confusion is likely we will refer to  $Z[J^*, J, \mathcal{V}]$  as simply  $Z$ , and also abbreviate terms such as  $\mathcal{H}_{\text{eff}}(\tau_{j+1}, \tau_j)$  to  $\mathcal{H}_{\text{eff}}(\tau_j)$  or even more simply to  $\mathcal{H}_{\text{eff}}$ . Eq. (102) is almost in the form of a canonical partition function for unprojected electrons, but with an important difference. The extra term in the integration measure is the Gutzwiller projector written in Grassmann variables. These variables arise *at the initial and next time instant only* and the rest of the time variables have only the standard measure of unity. Explicitly we find

$$P_G(\tau_1, \tau_0) \equiv \prod_{i=1}^{N_s} (1 - c_{i\uparrow}^*(\tau_1) c_{i\uparrow}(\tau_0) c_{i\downarrow}^*(\tau_1) c_{i\downarrow}(\tau_0)), \quad (109)$$

it has all creation (destruction) operators at  $j = 1$  ( $j = 0$ ), and  $N_s$  is the number of sites. In Appendix D, we summarize the Pauli principle and Gutzwiller identities obeyed by the present coherent state representation, these represent an important aspect of the strong correlation problem. We will also recycle the notation of (63) for the average in this distribution of any function  $Q$  of the Grassmann variables:

$$\langle\langle Q \rangle\rangle = \frac{\|Q\|}{Z}, \quad \text{with } \|Q\| = \int_c P_G(\tau_1, \tau_0) e^{-\mathcal{A}} Q, \quad (110)$$

a useful abbreviation (110), and drop the superscript ( $M$ ). This representation of the path integral with a constraining projection factor at only the initial time has a resemblance to the that in the canonical quantization of the electromagnetic field in the temporal gauge [37,38], as already noted in the introduction. The Green's functions follow from Eq. (73) using  $\delta/\delta J(\tau_j) \rightarrow \frac{1}{(\Delta\tau)} d/dJ(\tau_j)$  [50]:

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = \frac{1}{Z} \|\tilde{c}_{f\sigma_f}^*(\tau_f) c_{i\sigma}(\tau_i)\|. \quad (111)$$

## 7.2. Equations of motion from path integral representation

In this section we obtain the Schwinger equations of motion of ECFL (see Ref. [6] and especially Appendix A Eq. (138)), directly from the path integral representation given above thus providing a non trivial check on the representation. To obtain Eq. (138), we initially set the Fermionic sources to zero, the Bosonic sources are turned off at the very end. The equations of motion are most easily found using a Grassmann integration identity:

$$\int_c P_G(\tau_1, \tau_0) \frac{\delta}{\delta c_{i\sigma}^*(\tau_{i+1})} \left[ \tilde{c}_{f\sigma_f}^*(\tau_{f+1}) e^{-\mathcal{A}_{\text{Tot}}} \right] = 0. \quad (112)$$

This identity is a straight forward generalization of the theorem on vanishing of a total derivative [46], including a non trivial measure  $P_G$  (109) where the time arguments are greater than all time

arguments in (109), i.e.  $i, f \geq 1$ . It is proved by the usual logic for Grassmann variables; the derivative  $\frac{\delta}{\delta c_{i\sigma_i}^*(\tau_{i+1})}$  is in addition to an integration over  $c_{i\sigma_i}^*(\tau_{i+1})$  contained in the overall integration. We next recall that the highest possible degree of a polynomial in any Grassmann variable is unity. The above expression vanishes upon further noting that Grassmann integration and Grassmann differentiation are identical. The same identity is valid if we replace  $\tilde{c}_{f\sigma_f}^*(\tau_{f+1})$  by any other allowed Grassmann variable  $U$ , subject to the double occupancy restriction, and similarly with  $V$  (see Ref. [51]). In summary, an abstract equation of motion, following from  $\int P_G \frac{\delta}{\delta V} (Ue^{-A_{Tot}}) = 0$  and Fermionic  $U, V$  reads

$$\left\| \frac{\delta U}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_S}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_0}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_t}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_J}{\delta V} \right\| = 0. \quad (113)$$

### 7.3. Equation for $\mathcal{G}_{\sigma_i\sigma_f}(i, f)$

Our first task is to find an equation for the Green's function [50]—we use (113) with  $U = \tilde{c}_{f\sigma_f}^*(\tau_f)$  and  $V = c_{i\sigma_i}^*(\tau_i)$ . We compute the various pieces of (113) next.

Denoting

$$\widehat{\gamma}_{\sigma_i\sigma_f}(i) \equiv \sigma_i\sigma_f c_{i\sigma_i}^*(\tau_{i+1}) c_{i\sigma_f}(\tau_i), \quad (114)$$

and using the convention that repeated spin indices are summed over, we obtain the first result:

$$\frac{\delta}{\delta c_{i\sigma_i}^*(\tau_{i+1})} \tilde{c}_{f\sigma_f}^*(\tau_{f+1}) = \delta_{\tau_i\tau_f} \delta_{if} \{ \delta_{\sigma_i\sigma_f} - \widehat{\gamma}_{\sigma_i\sigma_f}(i) \}. \quad (115)$$

We obtain

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_0}{\delta c_{i\sigma_i}^*(\tau_{i+1})} = \delta_{\tau_i} c_{i\sigma_i}(\tau_i) - \boldsymbol{\mu} c_{i\sigma_i}(\tau_i), \quad (116)$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_S}{\delta c_{i\sigma_i}^*(\tau_{i+1})} = \mathcal{V}_i^{\sigma_i\sigma_j}(\tau_{i+1}) c_{i\sigma_j}(\tau_i) + \{ \delta_{\sigma_i\sigma_j} - \widehat{\gamma}_{\sigma_i\sigma_j}(i) \} J_{i\sigma_j}(\tau_{i+1}), \quad (117)$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_t^{Sym}}{\delta c_{i\sigma_i}^*(\tau_{i+1})} = -t_{ij} c_{j\sigma_i}(\tau_i) + t_{ij} [ \widehat{\gamma}_{\sigma_i\sigma_j}(i\tau_i) c_{j\sigma_j}(\tau_i) + c_{j\sigma_i}^* c_{j\sigma_i} c_{j\sigma_i} + c_{j\sigma_i}^* c_{i\sigma_i} c_{i\sigma_i} ],$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_t^{Min}}{\delta c_{i\sigma_i}^*(\tau_{i+1})} = -t_{ij} c_{j\sigma_i}(\tau_i) + t_{ij} \widehat{\gamma}_{\sigma_i\sigma_j}(i\tau_i) c_{j\sigma_j}(\tau_i)$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_J}{\delta c_{i\sigma_i}^*(\tau_{i+1})} = -\frac{1}{2} J_{ij} \sigma_i \sigma_j c_{j\sigma_i}^*(\tau_{i+1}) c_{j\sigma_j}(\tau_i) c_{i\sigma_j}(\tau_i), \quad (118)$$

We combine the two terms as:

$$\frac{1}{\Delta\tau} \frac{\delta (\mathcal{A}_t + \mathcal{A}_J)}{\delta c_{i\sigma_i}^*(\tau_{i+1})} = - \sum_j t_{ij} c_{j\sigma_i}(\tau_i) + A_{i\sigma_i}(\tau_{i+1}, \tau_i), \quad (119)$$

the first (linear) term in Fermions is separated out in this expression, and  $A_{i\sigma_i}$ , detailed below in Eq. (120), is obtained by combining all the *three Fermion* contributions in Eqs. (117) and (118). In the minimal case we get from Eqs. (107), (108), (114) and (115)

$$A_{i\sigma_i}^{Min} = t_{ij} \widehat{\gamma}_{\sigma_i\sigma_j}(i\tau_i) c_{j\sigma_j}(\tau_i) - \frac{1}{2} J_{ij} \widehat{\gamma}_{\sigma_i\sigma_j}(j\tau_i) c_{i\sigma_j}(\tau_i), \quad (120)$$

in agreement with Eq. (22) of Ref. [6], and the symmetrized case is obtained in a similar way. Combining these (with  $J \rightarrow 0$ ) we get the EOM in discrete time space:

$$\begin{aligned} & \left\{ \boldsymbol{\mu} - \delta_{\tau_i} - \mathcal{V}_i^{\sigma_i \sigma_j}(\tau_{i+1}) \right\} \delta_{i,j} + t_{ij} \left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_{f+1}) c_{j\sigma_j}(\tau_i) \right\| - \left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_{f+1}) A_{i\sigma_i}(\tau_{i+1}, \tau_i) \right\| \\ & = \delta_{if} \left\| (\delta_{\sigma_i \sigma_f} - \hat{\gamma}_{\sigma_i \sigma_f}(i)) \right\| \frac{\delta_{\tau_i, \tau_f}}{\Delta \tau}. \end{aligned} \quad (121)$$

We next take the continuum limit in time  $\tau_i$ ; with  $\Delta \tau \rightarrow 0$ , and using  $\frac{\delta_{\tau_i, \tau_f}}{\Delta \tau} \rightarrow \delta(\tau_i - \tau_f)$ , and using the non interacting Fermi Green's function from (136), and implementing the basic Schwinger identity for representing higher order correlation functions as source derivatives:

$$\left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_f) A_{i\sigma_i}(\tau_i) \right\| = \hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) \left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_f) c_{j\sigma_j}(\tau_j) \right\| \quad (122)$$

where  $\hat{X}$  is a functional derivative operator defined more completely below in (130). With this preparation we can rewrite Eq. (121) as

$$\left( \mathbf{g}_{0, \sigma_i, \sigma_j}^{-1}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) \right) \left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_f) c_{j\sigma_j}(\tau_j) \right\| = \delta_{if} \left\| (\delta_{\sigma_i \sigma_f} - \hat{\gamma}_{\sigma_i \sigma_f}(i)) \right\| \delta(\tau_i - \tau_f). \quad (123)$$

Here and elsewhere since  $\tau_j$  repeats in product, it is assumed to be integrated between  $0 \leq \tau_j \leq \beta$ , this rule is analogous to the spin index summation rule. We next divide by  $Z$ , use (110) to define the Green's function, and also define

$$Y_{1\sigma_i \sigma_j}(i\tau_i, j\tau_j) = \frac{1}{Z} (\hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) Z), \quad (124)$$

to rewrite (123) in the same form as Eq. (138)

$$\begin{aligned} & \left( \mathbf{g}_{0, \sigma_i, \sigma_j}^{-1}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) - Y_{1\sigma_i \sigma_j}(i\tau_i, j\tau_j) \right) \\ & \times \mathcal{G}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) = \delta_{if} \delta(\tau_i - \tau_f) \left[ \delta_{\sigma_i \sigma_f} - \gamma_{\sigma_i \sigma_f}(i) \right], \end{aligned} \quad (125)$$

where

$$\gamma_{\sigma_i \sigma_f}(i) \equiv \langle\langle \hat{\gamma}_{\sigma_i \sigma_f}(i) \rangle\rangle. \quad (126)$$

This is readily seen to be identical to the direct definition given before in (135). We next use  $\tilde{i} \equiv (i, \tau_i, \sigma_i)$  as an abbreviation for the (space, time, spin) indices, and use the repeated index summation convention. Here summation stands for spin and spatial sums, and temporal integrals in the standard intervals. With this we can write  $\hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) \leftrightarrow \hat{X}_{\tilde{i}\tilde{j}}$ , and similarly for  $\mathbf{g}_0^{-1}$ ,  $\mathcal{G}$  and  $Y_1$ . The variable (126) is local and needs the extra definition  $\gamma_{\sigma_i \sigma_f}(i\tau_i) \delta_{if} \delta(\tau_i - \tau_f) \leftrightarrow \gamma_{\tilde{i}\tilde{f}}$  and also denote  $\delta_{if} \delta_{\sigma_i \sigma_f} \delta(\tau_i - \tau_f) \leftrightarrow \delta_{\tilde{i}\tilde{f}}$ . With these, the matrix product form of Eq. (123) reads:

$$\left( \mathbf{g}_{0, \tilde{i}\tilde{j}}^{-1} - \hat{X}_{\tilde{i}\tilde{j}} - Y_{1\tilde{i}\tilde{j}} \right) \mathcal{G}_{\tilde{j}\tilde{f}} = (\delta_{\tilde{i}\tilde{f}} - \gamma_{\tilde{i}\tilde{f}}). \quad (127)$$

This is exactly the form of the Schwinger equation for the Green's function obtained from the continuous time Heisenberg equations of motion (138) in [4,6], using the above abbreviation convention.

In order to obtain an expression for  $\hat{X}$ , we note a useful relationship involving the action on the partition functional (102) of the operator  $D_{\sigma_i \sigma_j}(i) \equiv \sigma_i \sigma_j \delta / \delta \mathcal{V}_i^{\sigma_i \sigma_j}$  (from Eq. (39) of Ref. [6])

$$D_{\sigma_i \sigma_j}(i) Z[\mathcal{V}] = -\left\| \hat{\gamma}_{\sigma_i \sigma_j}(i) \right\|, \quad (128)$$

so that:

$$\left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_{f+1}) A_{i\sigma_i}^{Min} \right\| = -t_{ij} D_{\sigma_i \sigma_j}(i) \left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_{f+1}) c_{j\sigma_j}(\tau_i) \right\| + \frac{1}{2} J_{ij} D_{\sigma_i \sigma_j}(j) \left\| \tilde{\mathcal{C}}_{f\sigma_f}^*(\tau_{f+1}) c_{i\sigma_i}(\tau_i) \right\|, \quad (129)$$

and comparing with Eq. (122) we conclude

$$\hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) = \delta(\tau_i - \tau_j)(-t_{ij}D_{\sigma_i\sigma_j}(i) + \delta_{ij} \sum_k \frac{1}{2}J_{ik}D_{\sigma_i\sigma_j}(k\tau_i)), \quad (130)$$

where the derivative  $D_{\sigma_i\sigma_j}(k\tau_i)$  is at spatial site  $k$  and time  $\tau_i$ . The corresponding  $Y_1$  (with a similar convention as above) in Eq. (124) is

$$Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j) = -\delta(\tau_i - \tau_j) \left( -t_{ij}\mathcal{V}_{\sigma_i\sigma_j}(i) + \delta_{ij} \sum_k \frac{1}{2}J_{ik}\mathcal{V}_{\sigma_i\sigma_j}(k\tau_i) \right). \quad (131)$$

Analogous expressions for the symmetrized case, for  $A_i$ ,  $\hat{X}$  and  $Y_1$  parallel to Eqs. (130), (131) and (120), can be obtained by using the symmetrized version (top line) of Eq. (117). This expression agrees with Eq. (43) of Ref. [6], and their minimal version obtained after dropping the second and fourth term. We have thus verified that the exact equations of motion are obtained from the path integral representation outlined here, constituting a non trivial check on the formalism.

## 8. Conclusions

In this work we have presented a simpler method to obtain the ECFL theory that complements the Schwinger method used earlier. This new method brings an important analogy to the Dyson–Maleev theory to attention, and this connection helps us to get a different perspective on the main results of ECFL, in particular the novel non-Dysonian representation of the Greens function is placed on a firm foundation. The path integral method is used to set up an alternate quantum field theory with a non Hermitian Hamiltonian, and it is proven to be valid by reproducing the Schwinger equations of motion.

We draw particular attention to the scaling result for the spectral function (34) and (35) in Section 2.4. Here the low energy spectral function is shown to satisfy a simple relation involving the hole density that throws light on the ever shrinking regime of validity of the Landau Fermi liquid, as we approach the insulating state. Finally the discussion of the alternate ways to analyze the ARPES line shapes discussed in Section 2.5 should be of interest to the ARPES community, as also the discussion of the electronic origin of a kink in the EDC energy dispersion.

## Acknowledgments

This work was supported by DOE under Grant No. FG02-06ER46319. I thank T. Banks and O. Narayan for helpful discussions on the path integral representation. I am grateful to P.W. Anderson, P. Coleman, A. Georges, G.H. Gweon, A. Hewson, P.D. Johnson, G. Kotliar, H.R. Krishnamurthy, E. Perepelitsky and T.M. Rice for helpful comments and valuable discussions.

## Appendix A. Summary of the minimal theory and its Schwinger equations of motion

In order to make the discussions reasonably self contained, we provide a brief discussion of the minimal equations of motion for the Green's function. These are obtained through the usual Schwinger method used in Refs. [4,6]. These equations are a subset of the ones given in Ref. [6], and can be obtained by omitting certain extra terms therein, which were added to satisfy a symmetry property. We term these equations are *the minimal theory*, since no terms are added or dropped, and the expressions are not reducible by any other argument. We also indicate the generalization to include the parameter  $\lambda$  in these equations, to facilitate comparing with the equations in this work.

Using the Hamiltonian (9) we note the important commutator (given in Ref. [6]):

$$[H_{ij}, X_i^{0\sigma_i}] = \sum_j t_{ij}X_j^{0\sigma_i} + \mu X_i^{0\sigma_i} - \sum_{j\sigma_j} t_{ij}(\sigma_i\sigma_j)X_i^{\bar{\sigma}_i\bar{\sigma}_j} X_j^{0\sigma_j} + \frac{1}{2} \sum_{j \neq i} J_{ij}(\sigma_i\sigma_j)X_j^{\bar{\sigma}_i\bar{\sigma}_j} X_i^{0\sigma_j}. \quad (132)$$

Temporarily ignoring the Fermionic sources:

$$[\hat{\mathcal{A}}_S(i\tau_i), X_i^{0\sigma_i}] = -\mathcal{V}_i^{\sigma_i\sigma_j} X_i^{0\sigma_j}, \quad (133)$$

and combining with the Heisenberg equation of motion, we see that the Green's function satisfies the EOM

$$\begin{aligned} \partial_{\tau_i} \mathcal{G}_{\sigma_i \sigma_f}(i, f) &= -\delta(\tau_i - \tau_f) \delta_{ij} (1 - \gamma_{\sigma_i \sigma_f}(i\tau_i)) \\ &\quad - \left\langle T_{\tau} \left( e^{-\hat{\mathcal{A}}_S} [H_{ij} + \hat{\mathcal{A}}_S(i, \tau_i), X_i^{0\sigma_i}(\tau_i)] X_f^{\sigma_f 0}(\tau_f) \right) \right\rangle \end{aligned} \quad (134)$$

where the local Green's function

$$\gamma_{\sigma_a \sigma_b}(i\tau_i) = \sigma_a \sigma_b \mathcal{G}_{\bar{\sigma}_b \bar{\sigma}_a}(i\tau_i^-, i\tau_i). \quad (135)$$

Substituting and using the Fermi gas (i.e. free) Green's function:

$$\mathbf{g}_{0, \sigma_i, \sigma_j}^{-1}(i\tau_i, j\tau_j) = \left\{ \delta_{\sigma_i \sigma_j} [\delta_{ij}(\boldsymbol{\mu} - \partial_{\tau_i}) + t_{ij}] - \delta_{ij} \mathcal{V}_i^{\sigma_i \sigma_j}(\tau_i) \right\} \delta(\tau_i - \tau_j), \quad (136)$$

we obtain (using the repeated index summation and integration convention of Ref. [6])

$$\begin{aligned} \mathbf{g}_{0, \sigma_i, \sigma_j}^{-1}(i\tau_i, j\tau_j) \mathcal{G}_{\sigma_j \sigma_f}(j\tau_j, f\tau_f) &= \delta(\tau_i - \tau_f) \delta_{ij} (1 - \gamma_{\sigma_i \sigma_f}(i\tau_i)) \\ &\quad - \sum_{j\sigma_j} t_{ij}(\sigma_i \sigma_j) \left\langle T_{\tau} \left( X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) X_j^{0\sigma_j}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \right) \right\rangle \\ &\quad + \frac{1}{2} \sum_{j\sigma_j} J_{ij}(\sigma_i \sigma_j) \left\langle T_{\tau} \left( X_j^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) X_i^{0\sigma_j}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \right) \right\rangle. \end{aligned} \quad (137)$$

We next express the higher order Green's function in terms of the derivatives of the lower one to obtain the Schwinger EOM:

$$\begin{aligned} &\left( \mathbf{g}_{0, \sigma_i, \sigma_j}^{-1}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) - Y_{1\sigma_i \sigma_j}(i\tau_i, j\tau_j) \right) \mathcal{G}_{\sigma_j \sigma_f}(j\tau_j, f\tau_f) \\ &= \delta_{ij} \delta(\tau_i - \tau_f) (\delta_{\sigma_i \sigma_f} - \gamma_{\sigma_i \sigma_f}(i\tau_i)), \end{aligned} \quad (138)$$

where we used the functional derivative operator

$$D_{\sigma_i \sigma_j}(i\tau_i) = \sigma_i \sigma_j \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i)} \quad (139)$$

and the composite derivative operator

$$\hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) = \delta(\tau_i - \tau_j) \left( -t_{ij} D_{\sigma_i \sigma_j}(i\tau_i) + \delta_{ij} \sum_k \frac{1}{2} J_{ik} D_{\sigma_i \sigma_j}(k\tau_i) \right), \quad (140)$$

where the derivative  $D_{\sigma_i \sigma_j}(k\tau_i)$  is at spatial site  $k$  and time  $\tau_i$ . The corresponding  $Y_1$  (with a similar convention as above) in Eq. (124) is

$$Y_{1\sigma_i \sigma_j}(i\tau_i, j\tau_j) = -\delta(\tau_i - \tau_j) \left( -t_{ij} \gamma_{\sigma_i \sigma_j}(i\tau_i) + \delta_{ij} \sum_k \frac{1}{2} J_{ik} \gamma_{\sigma_i \sigma_j}(k\tau_i) \right). \quad (141)$$

Eqs. (138), (140) and (141) define the minimal theory. For reference we note that Ref. [6] gives these equations, and goes on to add terms that account for the symmetrized theory with a Hermitian  $H_{\text{eff}}$ . We also note that Eq. (138) can be generalized to include the  $\lambda$  parameter by scaling  $\hat{X}_{\sigma_i \sigma_j}, Y_{1\sigma_i \sigma_j}, \gamma_{\sigma_i \sigma_j} \rightarrow \lambda \hat{X}_{\sigma_i \sigma_j}, \lambda Y_{1\sigma_i \sigma_j}, \lambda \gamma_{\sigma_i \sigma_j}$ .

## Appendix B. Coherent state definitions

We use standard anticommuting Grassmann variables [46] to represent the canonical Fermions  $C$  and  $C^\dagger$  for each spin and site. In brief we note the anticommuting property  $\{c_i, c_j^*\} = 0 = \{c_i, c_j\} = \{c_i^*, c_j^*\} = \{c_i, C_j^\dagger\}$ . Suppressing  $j$  and spin index the Fermi coherent states are given as usual by:

$$\begin{aligned} |c\rangle &= e^{-c C^\dagger} |vac\rangle = (1 - c C^\dagger) |vac\rangle \\ \langle c| &= \langle vac| e^{-C c^*} = \langle vac| (1 - C c^*) \\ \langle c|c'\rangle &= 1 + c^* c' = e^{c^* c'}, \end{aligned} \quad (142)$$

where  $|vac\rangle$  is the vacuum state. We use the abbreviation to denote coherent state integrals:

$$\int_c = \int dc^* dc. \quad (143)$$

The basic integrals are

$$\begin{aligned} \int_c (1, c^*, c, cc^*) &= (0, 0, 0, 1) \\ \int_c e^{-c^* c} &= 1. \end{aligned} \quad (144)$$

The completeness relation reads:

$$\int_c e^{-c^* c} |c\rangle \langle c| = |vac\rangle \langle vac| + C^\dagger |vac\rangle \langle vac| C \equiv \mathbb{1}, \quad (145)$$

and the trace over Fermionic variables is given by:

$$\text{Tr}A = \int_c e^{-c^* c} \langle -c|A|c\rangle. \quad (146)$$

## Appendix C. Path integral representation

We now introduce path integrals to represent Eq. (68) leading to (102). Towards this end let us write

$$\begin{aligned} \beta \hat{H}_{\text{eff}} + \hat{A}_S &= \int_0^\beta \hat{\mathcal{H}}(\tau) d\tau \\ \hat{\mathcal{H}}(\tau) &\equiv \hat{H}_{\text{eff}} + \sum_i \hat{A}_S(i, \tau). \end{aligned} \quad (147)$$

The integral is represented by a finite sum over  $M$  intervals, and the limit  $M \rightarrow \infty$  taken at the end, thus

$$\int_0^\beta \hat{\mathcal{H}}(\tau) d\tau \rightarrow \lim_{M \rightarrow \infty} \Delta\tau \sum_{j=1, M} \hat{\mathcal{H}}(\tau_j), \quad (148)$$

where we defined

$$\begin{aligned} \tau_j &= \Delta\tau \times j = \frac{j\beta}{M}, \\ \Delta\tau &= \frac{\beta}{M}, \quad j = 1, M. \end{aligned} \quad (149)$$

Thus with  $\hat{\mathcal{H}}(j) \equiv \hat{\mathcal{H}}(\tau_j)$  arranged to be in normal ordered form (creation operators to the left of the destruction operators) we write Trotters formula for the exponential

$$\begin{aligned} Z^{(M)} &= \int_{c(0)} e^{-c_{i\sigma}^*(0)c_{i\sigma}(0)} \langle -c(0) | e^{-\Delta\tau \hat{\mathcal{H}}(\tau_M)} e^{-\Delta\tau \hat{\mathcal{H}}(\tau_{M-1})} \dots e^{-\Delta\tau \hat{\mathcal{H}}(\tau_2)} e^{-\Delta\tau \hat{\mathcal{H}}(\tau_1)} \hat{P}_G | c(0) \rangle \\ &= \int_c e^{-\sum_{j=1}^M c_{i\sigma}^*(j)c_{i\sigma}(j)} \langle c(M) | e^{-\Delta\tau \hat{\mathcal{H}}(\tau_M)} | c(M-1) \rangle \dots | c(2) \rangle \langle c(2) | \\ &\quad \times e^{-\Delta\tau \hat{\mathcal{H}}(\tau_2)} | c(1) \rangle \langle c(1) | e^{-\Delta\tau \hat{\mathcal{H}}(\tau_1)} \hat{P}_G | c(0) \rangle. \end{aligned} \quad (150)$$

Anti periodic boundary conditions are used:  $c(\tau_M) = -c(\tau_0)$  and we set at each time slice  $\tau_j$  the coherent state  $|c(j)\rangle = \prod_{i\sigma} |c_{i\sigma}(\tau_j)\rangle$  as a global product over all sites and both spins, and the symbol  $\int_c$  represents integration over all the sites spins and time slices. The site index  $i$  and spin  $\sigma$  are implicitly summed over. Recall that  $\hat{P}_G$  is brought to the extreme right of the product. We calculate as usual:

$$\begin{aligned} \langle c(\tau_{j+1}) | c(\tau_j) \rangle &= e^{c^*(\tau_{j+1})c(\tau_j)} \\ \langle c(\tau_{j+1}) | e^{-\Delta\tau \hat{\mathcal{H}}(\tau_{j+1})} | c(\tau_j) \rangle &\equiv e^{c^*(\tau_{j+1})c(\tau_j) - \Delta\tau \mathcal{H}(\tau_{j+1}, \tau_j)} + O(\Delta\tau^2) \\ \mathcal{H}(\tau_{j+1}, \tau_j) &\equiv \frac{\langle c(\tau_{j+1}) | \hat{\mathcal{H}}(\tau_{j+1}) | c(\tau_j) \rangle}{\langle c(\tau_{j+1}) | c(\tau_j) \rangle}. \end{aligned} \quad (151)$$

The last term needs careful attention, we note

$$\begin{aligned} \langle c(\tau_1) | e^{-\Delta\tau \hat{\mathcal{H}}(\tau_1)} \hat{P}_G | c(\tau_0) \rangle &= \langle c(\tau_1) | (1 - \Delta\tau \hat{\mathcal{H}}(\tau_1)) \hat{P}_G | c(\tau_0) \rangle + O(\Delta\tau^2) \\ &= \langle c(\tau_1) | c(\tau_0) \rangle (1 - \Delta\tau \mathcal{H}(\tau_1, \tau_0)) P_G(\tau_1, \tau_0) + O(\Delta\tau^2) \\ &= \langle c(\tau_1) | c(\tau_0) \rangle e^{-\Delta\tau \mathcal{H}(\tau_1, \tau_0)} P_G(\tau_1, \tau_0) + O(\Delta\tau^2), \end{aligned} \quad (152)$$

where Eq. (109) details the expression for  $P_G(\tau_1, \tau_0)$ , it contains variables at the initial and next time instant only. Combining all terms, we get the expression (102). We have thrown out terms of  $O(\Delta\tau)^2$  in obtaining Eq. (102), and hence it is important to keep track of the Pauli principle identities, discussed above in Eqs. (156) and (158). Note that for arbitrary  $\tau$

$$\langle \tau | C_{i\sigma}^\dagger | \tau_j \rangle = c_{i\sigma}^*(\tau) e^{c_{i\sigma}^*(\tau)c_{i\sigma}(\tau_j)} = -\frac{\delta}{\delta c_{i\sigma}(\tau_j)} \langle \tau | \tau_j \rangle. \quad (153)$$

In view of this relation we note the following mappings:

$$\begin{aligned} \langle \psi | C_{i\sigma}^\dagger | \tau_j \rangle &\rightarrow -\frac{\delta}{\delta c_{i\sigma}(\tau_j)} \langle \psi | \tau_j \rangle, \\ \langle \psi | C_{i\sigma} | \tau_j \rangle &\rightarrow c_{i\sigma}(\tau_j) \langle \psi | \tau_j \rangle, \\ \langle \tau_j | C_{i\sigma}^\dagger | \psi \rangle &\rightarrow c_{i\sigma}^*(\tau_j) \langle \tau_j | \psi \rangle, \\ \langle \tau_j | C_{i\sigma} | \psi \rangle &\rightarrow \frac{\delta}{\delta c_{i\sigma}^*(\tau_j)} \langle \tau_j | \psi \rangle, \end{aligned} \quad (154)$$

Let us show how the commutation works here:

$$(CC^\dagger + C^\dagger C) | c \rangle = \left( -C \frac{\delta}{\delta c} + C^\dagger c \right) | c \rangle = \left( \frac{\delta}{\delta c} C - c C^\dagger \right) | c \rangle = \left( \frac{\delta}{\delta c} c + c \frac{\delta}{\delta c} \right) | c \rangle = | c \rangle. \quad (155)$$

#### Appendix D. Pauli and Gutzwiller exclusion identities

It is worth highlighting a few conventions about (102) and related expressions. These are designed to retain some of the most important features of strongly interacting electrons *on a lattice*. In contrast, in a theory of electrons in the continuum, these constraints are of no special consequence since coincident spatial points have a measure of zero. We first discuss the Pauli principle related rules

referring to the same spin species, and then the Gutzwiller projection related rules relating to opposite spin species, these are operative when two electronic operators have coincident space and time coordinates.

- (I) When two coincident times in a product of operators *on the same lattice site* and same spin arise, we follow the convention of immediate evaluation of the product. By evaluation, we understand that the product of two similar Grassmann variables is set to zero, and for dissimilar Grassmann variables (e.g.  $c$  and  $c^*$ ) at a common time, both of them are integrated out immediately. This leads to the basic set of *Pauli exclusion identities* at equal times as one easily verifies:

$$\begin{aligned} c_{i\sigma}(\tau_j)c_{i\sigma}^*(\tau_j) &\rightarrow 1 \\ c_{i\sigma}^*(\tau_j)c_{i\sigma}(\tau_j) &\rightarrow -1 \\ c_{i\sigma}(\tau_j)c_{i\sigma}(\tau_j) &\rightarrow 0 \\ c_{i\sigma}^*(\tau_j)c_{i\sigma}^*(\tau_j) &\rightarrow 0. \end{aligned} \quad (156)$$

- (II) We denote the number operator as

$$n_{i\sigma}(\tau_j) \equiv c_{i\sigma}^*(\tau_{j+1})c_{i\sigma}(\tau_j), \quad (157)$$

where we observe that the  $c^*$  has the immediately later time argument than that of  $c$ , this comes about from representing  $\langle j+1|C^\dagger C|j\rangle = c^*(\tau_{j+1})c(\tau_j) \times \langle j+1|j\rangle$ . Using this we will verify the *second set* of Pauli exclusion identities

$$\begin{aligned} n_{i\sigma}(\tau_{j+1})n_{i\sigma}(\tau_j) &= n_{i\sigma}(\tau_j) \\ c_{i\sigma}^*(\tau_{j+1})n_{i\sigma}(\tau_j) &= 0 \\ n_{i\sigma}(\tau_j)c_{i\sigma}^*(\tau_j) &= c_{i\sigma}^*(\tau_{j+1}) \\ n_{i\sigma}(\tau_j)c_{i\sigma}(\tau_j) &= 0 \\ c_{i\sigma}(\tau_j)n_{i\sigma}(\tau_{j-1}) &= c_{i\sigma}(\tau_{j-1}). \end{aligned} \quad (158)$$

(III) We next obtain the important *Gutzwiller exclusion identity*. Calling the  $i$ th term in the product (109) as  $P_G^{(i)}(\tau_1, \tau_0)$ ; we see that

$$\begin{aligned} n_{i\uparrow}(\tau_1)n_{i\downarrow}(\tau_1)P_G^{(i)}(\tau_1, \tau_0) &= n_{i\uparrow}(\tau_1)n_{i\downarrow}(\tau_1) \\ -c_{i\uparrow}^*(\tau_2)c_{i\uparrow}(\tau_0)c_{i\downarrow}^*(\tau_2)c_{i\downarrow}(\tau_0) &\sim 0. \end{aligned} \quad (159)$$

The last line follows upon expanding  $\tau_2(\equiv \tau_1 + \Delta\tau)$  about  $\tau_1$ . The assumption that terms of  $O(\Delta\tau)$  are negligible is implicit in the entire path integral formulation. This shows that the double occupancy type terms  $n_{i\uparrow}(\tau_1)n_{i\downarrow}(\tau_1)$  that occur at any site lead to vanishing contribution, thus enforcing Gutzwiller projection. We can extend this argument to other times  $\tau_j \geq \tau_1$ :

$$c_{i\uparrow}(\tau_j)c_{i\downarrow}(\tau_j) \dots P_G^{(i)}(\tau_1, \tau_0) = 0, \quad (160)$$

where the dots indicate contributions from intermediate times. These contributions, after Grassmann integration over the terms at intermediate times, must necessarily end up with  $\dots c_{i\uparrow}(\tau_1)c_{i\downarrow}(\tau_1)P_G^{(i)}(\tau_1, \tau_0)$ . Expanding this factor leads to  $c_{i\uparrow}(\tau_1)c_{i\downarrow}(\tau_1) - c_{i\uparrow}(\tau_0)c_{i\downarrow}(\tau_0)$ , and therefore vanishes to  $O(\Delta\tau)$ , as in the argument in (159).

## Appendix E. Interpreting the caparison factor in the Schwinger method

Within the Schwinger method, or the related path integral formulation given above, the decomposition of  $\mathcal{G}$  is best done by rescaling the source terms by a factor determined through a self consistent argument given next. A convenient method is to work in the presence of the Fermionic sources, which allows us to start with the non vanishing average of a Fermi operator:

$$\xi_{i\sigma_i}(\tau_i) \equiv \frac{1}{Z} \|X_i^{0\sigma_i}(\tau_i)\| = \frac{1}{Z} \|c_{i\sigma_i}(\tau_i)\|, \quad (161)$$

and further abbreviate  $\xi_{i\sigma_i}(\tau_i) \leftrightarrow \xi_{\bar{i}}$ . A creation variable average  $\xi_{\bar{i}}^* \equiv \frac{1}{Z} \|X_i^{\sigma_i 0}(\tau_i)\|$  is also useful. The variable  $\xi_{\bar{i}}$  satisfies the functional differential equation that we study next:

$$(\mathbf{g}_{0,\bar{i},\bar{j}}^{-1} - \hat{X}_{\bar{i}\bar{j}})(Z\xi_{\bar{j}}) = Z(\delta_{\bar{i},\bar{k}} - \gamma_{\bar{i},\bar{k}})J_{\bar{k}}, \quad (162)$$

or using Eq. (124)

$$(\mathbf{g}_{0,\bar{i},\bar{j}}^{-1} - \hat{X}_{\bar{i}\bar{j}} - Y_{1\bar{i}\bar{j}}) \xi_{\bar{j}} = (\delta_{\bar{i},\bar{k}} - \gamma_{\bar{i},\bar{k}})J_{\bar{k}}. \quad (163)$$

This equation can be arrived at within the path integral representation (73), by using a variant of (112) after omitting the Fermionic creation type variable in the square bracket; and of course with a non vanishing Fermi source term. Alternately we can take the Heisenberg equations of motion in terms of the original expressions in terms of the  $X$  operators (11) and (13). The agreement between the two methods can be checked easily, and provides a strong check on the path integral formulation.

The Green's function is found from a variant of Eq. (13):

$$\mathcal{G}_{\bar{i}\bar{j}} - \xi_{\bar{j}}^* \xi_{\bar{i}} = \frac{\delta \xi_{\bar{i}}}{\delta J_{\bar{j}}}, \quad (164)$$

and on taking the limit  $J \rightarrow 0, J^* \rightarrow 0$ , all the single Fermi expectations  $\xi, \xi^*$  vanish. Taking the  $J$  derivative of (164), we see that (127) follows, so we will work with this equation from here.

The main objective from this point onwards, is to cast Eq. (163) or Eq. (127) into a form where the expressions on right are in the canonical form, i.e. where the time dependent  $\gamma$  term is gotten rid of in favor of a suitable constraint [4,52,53]. The occurrence of the factor  $\mathbb{1} - \gamma$  multiplying the source  $J$  in Eq. (163) suggests that one should scale the source  $J$  by a suitable time dependent factor to obtain new sources  $\mathcal{J}$ . The factor can be adjusted self consistently, so as to extract a canonical Green's function. Thus we scale

$$J_{\bar{i}} = (\mu^{-1})_{\bar{i}\bar{j}} \mathcal{J}_{\bar{j}}, \quad (165)$$

so that

$$\frac{\delta}{\delta J_{\bar{j}}} = \frac{\delta}{\delta \mathcal{J}_{\bar{k}}} \mu_{\bar{k}\bar{j}}, \quad (166)$$

where  $(\mu^{-1})$  is the *matrix inverse* of  $\mu$ . These equations have inverses that are easily obtained. The matrix  $\mu$  is dependent on the Fermi sources only indirectly, and this dependence may be neglected since the Fermi sources are turned off in the sequel. However it is allowed to be a functional of the Bosonic sources  $\mathcal{V}$ , thereby giving us considerable flexibility in defining it, we must also then be careful in locating it relative to the operator  $\hat{X}$ , since it contains derivatives with respect to  $\mathcal{V}$ .

In view of Eq. (166) we obtain the product relation

$$\begin{aligned} \mathcal{G}_{\bar{i}\bar{j}} &= \mathbf{g}_{\bar{i}\bar{k}} \mu_{\bar{k}\bar{j}} + \xi_{\bar{j}}^* \xi_{\bar{i}}, \\ \mathbf{g}_{\bar{i}\bar{k}} &= \frac{\delta \xi_{\bar{i}}}{\delta \mathcal{J}_{\bar{k}}}. \end{aligned} \quad (167)$$

The goal is to choose  $\mu$  such that the so defined  $\mathbf{g}$  satisfies a canonical equation, i.e. the analog of Eq. (127), but without the  $\gamma$  term on the right. For this purpose we can differentiate Eq. (163) with the scaled source field  $\mathcal{J}_{\bar{k}}$ , taking care to observe the non commutation of  $\mathcal{J}_{\bar{k}}$  with the derivative term  $\hat{X}$ . This process yields the equations:

$$(\mathbf{g}_{0,\bar{i},\bar{j}}^{-1} - \hat{X}_{\bar{i}\bar{j}} - Y_{1\bar{i}\bar{j}}) \mathbf{g}_{\bar{j}\bar{k}} = (\delta_{\bar{i},\bar{j}} - \gamma_{\bar{i},\bar{j}})(\mu^{-1})_{\bar{j}\bar{k}} - \left[ \hat{X}_{\bar{i}\bar{j}}(\mu^{-1})_{\bar{i}\bar{k}} \right] \mathbf{g}_{\bar{j}\bar{m}} \mu_{\bar{m}\bar{i}}, \quad (168)$$

where the square brackets demarcate the terms acted upon, by the derivative operators in  $\hat{X}$ . Eq. (168) exhibits a separation of variables, all the dependence on  $\mu$  is confined to the right hand side, and hence we set both sides to the identity matrix:

$$(\mathbf{g}_{0,\bar{i},\bar{j}}^{-1} - \hat{X}_{\bar{i}\bar{j}} - Y_{1\bar{i}\bar{j}}) \mathbf{g}_{\bar{j}\bar{k}} = \delta_{\bar{i}\bar{k}}, \quad (169)$$

and

$$(\delta_{\bar{i}\bar{j}} - \mathcal{V}_{\bar{i}\bar{j}})(\mu^{-1})_{\bar{j}\bar{k}} - \left[ \hat{X}_{\bar{i}\bar{j}}(\mu^{-1})_{\bar{i}\bar{k}} \right] \mathbf{g}_{\bar{j}\bar{m}} \mu_{\bar{m}\bar{l}} = \delta_{\bar{i}\bar{k}}. \quad (170)$$

Multiplying through with  $\mu$  and using

$$\left[ \hat{X}_{\bar{i}\bar{j}}(\mu^{-1})_{\bar{i}\bar{k}} \right] \mu_{\bar{m}\bar{l}} = - \left[ \hat{X}_{\bar{i}\bar{j}} \mu_{\bar{m}\bar{l}} \right] (\mu^{-1})_{\bar{i}\bar{k}}, \quad (171)$$

we rewrite Eq. (170) as

$$\mu_{\bar{i}\bar{f}} = (\delta_{\bar{i}\bar{f}} - \mathcal{V}_{\bar{i}\bar{f}}) + \mathbf{g}_{\bar{j}\bar{k}} \left[ \hat{X}_{\bar{i}\bar{j}} \mu_{\bar{k}\bar{f}} \right]. \quad (172)$$

We next show that Eqs. (169) and (170) can be rewritten in terms of the two self energies  $\Phi$  and  $\Psi$  used in Refs. [4,6]. We need the relation analogous to Eq. (171) to simplify Eq. (169):

$$\left[ \hat{X}_{\bar{i}\bar{j}} \mathbf{g}_{\bar{j}\bar{k}} \right] = - \mathbf{g}_{\bar{j}\bar{k}} \left[ \hat{X}_{\bar{i}\bar{j}} \mathbf{g}_{\bar{k}\bar{l}}^{-1} \right] \mathbf{g}_{\bar{i}\bar{k}}. \quad (173)$$

Therefore we write the two equations as

$$\begin{aligned} (\mathbf{g}_{0,\bar{i}\bar{j}}^{-1} - \Phi_{\bar{i}\bar{j}} - Y_{1\bar{i}\bar{j}}) \mathbf{g}_{\bar{j}\bar{k}} &= \delta_{\bar{i}\bar{k}}, \\ \mu_{\bar{i}\bar{f}} &= (\delta_{\bar{i}\bar{f}} - \mathcal{V}_{\bar{i}\bar{f}}) + \Psi_{\bar{i}\bar{f}}, \end{aligned} \quad (174)$$

where the two self energies  $\Phi$  and  $\Psi$  are functions obtained by iteration, and have a finite limit on turning off the Bosonic source  $\mathcal{V}$ . These are obtained from the above as

$$\begin{aligned} \Phi_{\bar{i}\bar{j}} &= - \mathbf{g}_{\bar{m}\bar{k}} \left[ \hat{X}_{\bar{i}\bar{m}} \mathbf{g}_{\bar{k}\bar{j}}^{-1} \right] \\ \Psi_{\bar{i}\bar{f}} &= \mathbf{g}_{\bar{j}\bar{k}} \left[ \hat{X}_{\bar{i}\bar{j}} \mu_{\bar{k}\bar{f}} \right]. \end{aligned} \quad (175)$$

Using the definition of  $\hat{X}$  and of various vertex functions, we can verify that these are precisely the pair of equations that we obtained in Refs. [4,6] for the two self energies of  $\mathcal{G}$ .

## References

- [1] P.W. Anderson, *Science* 235 (1987) 1196.
- [2] M.C. Gutzwiller, *Phys. Rev. Lett.* 10 (1963) 159; *Phys. Rev.* 137 (1965) 1726.
- [3] J. Hubbard, *Proc. R. Soc. Lond. A* 285 (1965) 542.
- [4] B.S. Shastry, arXiv:1102.2858, 2011; *Phys. Rev. Lett.* 107 (2011) 056403; *Phys. Rev. Lett.* 108 (2012) 029702.
- [5] B.S. Shastry, arXiv:1104.2633 ; *Phys. Rev. B* 84 (2011) 165112; *Phys. Rev. B* 86 (2012) 079911(E).
- [6] B.S. Shastry, arXiv:1207.6826, 2012; *Phys. Rev. B* 87 (2013) 125124.
- [7] D. Hansen, B.S. Shastry, arXiv:1211.0594, 2012; *Phys. Rev. B* 87 (2013) 245101.
- [8] B.S. Shastry, arXiv:1110.1032, 2011; *Phys. Rev. Lett.* 109 (2012) 067004.
- [9] R. Zitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, B.S. Shastry, arXiv:1309.5284, 2013.
- [10] E. Perepelitsky, B.S. Shastry, arXiv:1309.5373, 2013; *Ann. Phys.* 338 (2013) 283–301.
- [11] E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, B.S. Shastry, arXiv:1303.2657 [cond-mat.str-el], 2013; *Phys. Rev. B* 87 (2013) 161120(R).
- [12] B.S. Shastry, E. Perepelitsky, A.C. Hewson, arXiv:1307.3492 [cond-mat.str-el]; *Phys. Rev. B* 88 (2013) 205108.
- [13] H.R. Krishnamurthy, K. Wilson, J.W. Wilkins, *Phys. Rev. B* 21 (1980) 1044.
- [14] G.-H. Gweon, B.S. Shastry, G.D. Gu, arXiv:1104.2631, 2011; *Phys. Rev. Lett.* 107 (2011) 056404.
- [15] K. Matsuyama, G.H. Gweon, *Phys. Rev. Lett.* 111 (2013) 246401.
- [16] F.J. Dyson, *Phys. Rev.* 102 (1956) 1217–1230.
- [17] S.V. Maleev, *Sov. Phys.—JETP* 6 (1958) 776.
- [18] A.B. Harris, D. Kumar, B.I. Halperin, P.C. Hohenberg, *Phys. Rev. B* 3 (1971) 961.
- [19] S. Barnes, *J. Phys. F* 6 (1976) 1375;  
P. Coleman, *Phys. Rev. B* 29 (1984) 3035;  
G. Kotliar, A. Ruckenstein, *Phys. Rev. Lett.* 57 (1986) 1362.
- [20] N. Read, D.M. Newns, *J. Phys. C* 16 (1983) 3273;  
N. Read, D.M. Newns, *J. Phys.* 16 (1983) L1055;  
N. Read, *J. Phys. C* 18 (1985) 2651.
- [21] P.A. Lee, N. Nagaosa, X.G. Wen, *Rev. Modern Phys.* 78 (2006) 17.
- [22] P. Martin, J. Schwinger, *Phys. Rev.* 115 (1959) 1342.

- [23] The thermodynamic averages are denoted by  $\langle A \rangle = \frac{1}{Z} \text{Tr}_{ij} e^{-\beta H} A$ . Here  $H = H_{ij}$  or  $H = H_{\text{eff}}$  in the original description or the canonical theory description respectively.
- [24] L. Kadanoff, G. Baym, *Quantum Statistical Mechanics*, Perseus Books, Cambridge, 1962.
- [25] L. Kadanoff, P. Martin, *Phys. Rev.* 124 (1961) 670.
- [26] In most textbook examples of perturbation theory, one scales the interaction term with  $\lambda$ , since typically the interaction strength is an experimentally determined number (e.g. the Coulomb energy in hydrogenic problems). Introducing  $\lambda$  is a device to count the various orders of perturbation theory, and therefore a formal small parameter. A true small parameter, replacing it in many physical problems, is found by taking the ratio of characteristic energies of two parts of the Hamiltonian.
- [27] W.F. Brinkman, T.M. Rice, *Phys. Rev. B* 2 (1970) 4302.
- [28] In Ref. [9] a similar exercise is carried for the large dimensional case, extending to a somewhat greater (occupied) range of  $\omega$ . That analysis utilizes the greater information available in that case.
- [29] T. Senthil, *Phys. Rev. B* 78 (2008) 035103.
- [30] P.D. Johnson, T. Valla, A.V. Fedorov, Z. Yusof, B.O. Wells, Q. Li, A.R. Moodenbaugh, G.D. Gu, N. Koshizuka, C. Kendziora, Sha Jian, D.G. Hinks, *Phys. Rev. Lett.* 87 (2001) 177007.
- [31] A. Lanzara, P.V. Bogdanov, X.J. Zhou, S.A. Kellar, D.L. Feng, E.D. Lu, T. Yoshida, H. Eisaki, A. Fujimori, K. Kishio, J.-I. Shimoyama, T. Noda, S. Uchida, Z. Hussain, Z.-X. Shen, *Nature* 412 (2001) 510.
- [32] A. Kaminski, M. Randeria, J.C. Campuzano, M.R. Norman, H. Fretwell, J. Mesot, T. Sato, T. Takahashi, K. Kadowaki, *Phys. Rev. Lett.* 86 (2001) 1070.
- [33] B.S. Shastry, Superconducting state of the  $t$ - $J$  model in the ECFL theory (2014) to be published.
- [34] P.A. Casey, P.W. Anderson, *Phys. Rev. Lett.* 106 (2011) 097002;  
P.A. Casey, J.D. Koralek, N.C. Plumb, D.S. Dessau, P.W. Anderson, *Nat. Phys.* 4 (2008) 210;  
P.W. Anderson, *Nat. Phys.* 2 (2006) 626.
- [35] S. Doniach, M. Sunjic, *J. Phys. C* 3 (1970) 285.
- [36] This representation is very well known, to the point that it is hard to identify its first usage. Early usage is found in A.B. Harris and R.V. Lange, *Phys. Rev.* 157, 295 (1967) in Eq. (4.18 and 4.19a) and quite possibly to earlier papers.
- [37] P.A.M. Dirac, *Lectures on Quantum Mechanics*, Yeshiva Univ.-Belfer Graduate School of Science, New York, 1964, p. 32.
- [38] J.F. Willemssen, *Phys. Rev. D* 17 (1978) 574;  
M. Creutz, I.J. Muzinich, T.N. Tudron, *Phys. Rev. D* 19 (1979) 531;  
M. Creutz, *Ann. Phys.* 117 (1979) 471–483.
- [39] R. Kubo, *J. Phys. Soc. Japan* 12 (1957) 570;  
P.C. Martin, J. Schwinger, *Phys. Rev.* 115 (1959) 1342.
- [40] With all operators at the same (undisplayed) site, we can show that the commutator  $[C_{\sigma_1}^\dagger (1 - \lambda N_{\bar{\sigma}_1}), C_{\sigma_2}^\dagger C_{\sigma_3}]$  reduces to  $-\delta_{\sigma_1 \sigma_3} C_{\sigma_2}^\dagger (1 - \lambda N_{\bar{\sigma}_2})$  as follows. By straightforward expansion of the commutator we get  $-\delta_{\sigma_1 \sigma_3} C_{\sigma_2}^\dagger + \lambda C_{\sigma_2}^\dagger (\delta_{\bar{\sigma}_1 \sigma_3} N_{\bar{\sigma}_1} + \delta_{\sigma_2 \bar{\sigma}_3} (\delta_{\bar{\sigma}_1 \sigma_2} - \delta_{\bar{\sigma}_1 \sigma_3}) C_{\sigma_1}^\dagger C_{\sigma_3})$ . Here introduced the extra term  $\delta_{\sigma_2 \bar{\sigma}_3}$  using  $1 = \delta_{\sigma_2 \bar{\sigma}_3} + \delta_{\sigma_2 \sigma_3}$ , since the number of components is only two, and the vanishing of the other delta term. On using  $C_{\sigma_2}^\dagger C_{\sigma_2}^\dagger = 0$ , the last term then simplifies to  $\lambda \delta_{\sigma_1 \sigma_3} C_{\sigma_2}^\dagger (N_{\bar{\sigma}_2} \delta_{\bar{\sigma}_1 \sigma_2} + N_{\bar{\sigma}_1})$ , and its second piece can be simplified as  $(N_{\bar{\sigma}_2} \delta_{\bar{\sigma}_1 \sigma_2} + N_{\bar{\sigma}_1}) = (N_{\sigma_2} \delta_{\bar{\sigma}_1 \sigma_2} + N_{\bar{\sigma}_2})$ . Using  $C_{\sigma_2}^\dagger C_{\sigma_2}^\dagger = 0$ , the required result follows.
- [41] This alternative choice, upon using Eq. (85) leads to

$$n_{i\sigma} = \frac{1}{\lambda} \{1 - \mathcal{G}_{\bar{\sigma}\bar{\sigma}}^{(\lambda)}(i\tau^- i\tau) + \mathcal{G}_{\bar{\sigma}\bar{\sigma}}^{(\lambda)}(i\tau^+ i\tau)\}.$$

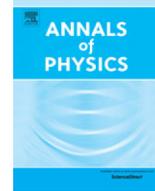
However, this prescription does not connect continuously to the correct constitutive relation at  $\lambda = 1$ , and is therefore rejected.

- [42] K.H. Douglass, *Ann. Phys.* 62 (1971) 383;  
K.H. Douglass, *Ann. Phys.* 64 (1971) 396.
- [43] J.S. Langer, *Phys. Rev.* 167 (1968) 183;  
J.S. Langer, *Phys. Rev.* 184 (1969) 219.
- [44] T. Holstein, H. Primakoff, *Phys. Rev.* 58 (1940) 1098;  
D.C. Mattis, *Theory of Magnetism*, Springer-Verlag, Berlin, 1981.
- [45] J.M. Luttinger, J.C. Ward, *Phys. Rev.* 118 (1960) 1417.
- [46] J. Zinn Justin, *Quantum Field Theory and Critical Phenomena*, Clarendon Press, Oxford, 1993, see Eq. (5.25).
- [47] R. Shankar, *Rev. Modern Phys.* 66 (1994) 129.
- [48] V.N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics*, D. Riedel and Co, Dordrecht, 1983, esp. Chapters 2 and 6.
- [49] J. Negele, H. Orland, *Quantum Many Particle Systems*, Perseus Books, Reading, Mass, 1998, esp. Chapter 2.
- [50] Since we are interested in  $M \rightarrow \infty$  at the end, we will simplify the notation from here by writing  $\tilde{c}_{f\sigma_f}^*(\tau_{f+1}) \rightarrow \tilde{c}_{f\sigma_f}^*(\tau_f)$ .
- [51] Regarding  $V$ , we need to exercise some care, since Eq. (154) shows that a derivative with respect to  $c_{i\sigma}$  is related to multiplying with  $c^*$ , rather than by.
- [52] B.S. Shastry, *Phys. Rev. B* 81 (2010) 045121.
- [53] It is worth mentioning one possibility that was tried earlier before discussing a more practical solution found recently in Ref. [4]. In Ref. [52] a combination of two transformations is made: (1) Eq. (163) is left multiplied by  $(\mathbb{1} - \gamma)^{-1}$  and (2) redefine  $\xi \rightarrow (\mathbb{1} - \gamma)\xi'$ . Together, these transformations give an exact formulation with a canonical Green's function  $G_{i,f} = \frac{\delta \xi'_i}{\delta \bar{J}_f}$  satisfying a canonical equation, when acted upon by an operator  $(\mathbb{1} - \gamma)^{-1}(\mathbf{g}_0 - \hat{X} - Y_1)(\mathbb{1} - \gamma)$ . A more thorough investigation of this scheme than in [52], especially with the addition of the  $\lambda$  parameter, seems to be a worthwhile goal for the future.



Contents lists available at ScienceDirect

## Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)

## Erratum

## Erratum to: “Theory of extreme correlations using canonical Fermions and path integrals” [Ann. Phys. 343 (2014) 164–199]



B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, CA 95064, United States*

## ARTICLE INFO

*Article history:*

Received 9 August 2016

Accepted 15 August 2016

Available online 18 August 2016

*Keywords:*

t-J model

Gutzwiller projection

Energy dispersion

Angle resolved photoemission

Path integrals

Schwinger equations of motion

## ABSTRACT

Erratum for paper “Theory of extreme correlations using canonical Fermions and path integrals”, B.S. Shastry, Ann. Phys. 343, 164–199 (2014) <http://dx.doi.org/10.1016/j.aop.2014.02.005> AOP69618.

The following errors should be corrected in the paper.

- (1) Page 172 below Eq. (22).

Replace

$$\hat{k} = (\vec{k} - \vec{k}_F) \cdot \vec{k}_F / |\vec{k}_F|$$

by

$$\hat{k} = (\vec{k} - \vec{k}_F) \cdot \nabla \varepsilon_k / |\nabla \varepsilon_k|.$$

- (2) Page 175 Second paragraph.

Replace

“Upon turning off ...”

by

“For  $Q > 0$ , upon turning off ...”.

DOI of original article: <http://dx.doi.org/10.1016/j.aop.2014.02.005>.

E-mail address: [sriram@physics.ucsc.edu](mailto:sriram@physics.ucsc.edu).

<http://dx.doi.org/10.1016/j.aop.2016.08.015>

(3) Page 175 Third line after Eq. (44).

Replace

$\text{sign}(\hat{k})$

by

$\text{sign}(-Q)$ .

(4) Page 175 Fourth line after Eq. (44).

Replace

$\text{sign}(\hat{k})$

by

$\text{sign}(-Q)$ .

(5) Eq. (45) should read

$$E^*(k_{\text{kink}}) = -\frac{1}{r-1} \Delta_0 - \Gamma_0.$$

### Acknowledgment

The work at UCSC was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award # DE-FG02-06ER46319.

## Linked-cluster expansion for the Green's function of the infinite- $U$ Hubbard model

Ehsan Khatami,<sup>1,2</sup> Edward Perepelitsky,<sup>1</sup> Marcos Rigol,<sup>3</sup> and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>*Department of Physics, University of California, Santa Cruz, California 95064, USA*

<sup>2</sup>*Department of Physics, University of California, Davis, California 95616, USA*

<sup>3</sup>*Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

(Received 3 December 2013; revised manuscript received 16 April 2014; published 2 June 2014)

We implement a highly efficient strong-coupling expansion for the Green's function of the Hubbard model. In the limit of extreme correlations, where the onsite interaction is infinite, the evaluation of diagrams simplifies dramatically enabling us to carry out the expansion to the eighth order in powers of the hopping amplitude. We compute the finite-temperature Green's function analytically in the momentum and Matsubara frequency space as a function of the electron density. Employing Padé approximations, we study the equation of state, Kelvin thermopower, momentum distribution function, quasiparticle fraction, and quasiparticle lifetime of the system at temperatures lower than, or of the order of, the hopping amplitude. We also discuss several different approaches for obtaining the spectral functions through analytic continuation of the imaginary frequency Green's function, and show results for the system near half filling. We benchmark our results for the equation of state against those obtained from a numerical linked-cluster expansion carried out to the eleventh order.

DOI: [10.1103/PhysRevE.89.063301](https://doi.org/10.1103/PhysRevE.89.063301)

PACS number(s): 02.60.-x, 71.10.Fd, 71.27.+a

### I. INTRODUCTION

In 1991, Metzner put forth an algorithm to compute the finite-temperature Green's function of the Fermi-Hubbard model [Eq. (1)] through a linked-cluster strong-coupling expansion [1]. His approach offers a relatively straightforward implementation on a computer, which is particularly useful today given the enormous improvements in computer power in the past two decades. The Metzner formalism further simplifies in the limit of extreme correlations, as the onsite repulsion  $U$  tends to infinity. In this paper, we implement his approach to obtain analytical expressions for the single-particle Green's function in that limit through eighth order in the expansion parameter  $\beta t$ , where  $\beta$  is the inverse temperature and  $t$  is the hopping amplitude of the electrons on the lattice.

In another recent development, the extremely correlated Fermi liquid theory (ECFL) [2] addresses this important limit through the use of Schwinger's source formulation of field theory. One of the significant physical ideas to come out of this theory is the presence of particle-hole asymmetry in the spectral densities of the single-particle Green's function and the Dyson-Mori self-energy [2–8]. This asymmetry, which has also been observed in dynamical mean-field theory studies of the Hubbard model [7,9,10], becomes more pronounced as the density approaches half-filling, i.e., as  $n \rightarrow 1$ . The asymmetry has implications for understanding the magnitude and sign of the Seebeck coefficient near the Mott insulating limit [10–12] and for explaining the anomalous line shapes of angle-resolved photoemission spectroscopy experiments [13] in strongly correlated materials.

In a recent work [14], the present authors (with Hansen) used the series expansion method to successfully benchmark the ECFL results for the spectral function [5], in their common regime of applicability. The currently available [ $O(\lambda^2)$ ] self-consistent solution of the ECFL is valid for  $n \lesssim 0.75$ . Additionally, the insight afforded by the aforementioned particle-hole asymmetry enabled us to construct a suitably modified first moment of the spectral function, providing a good estimate for the location of the quasiparticle peak. This

moment reduces the contribution from the occupied side of the spectrum relative to the unoccupied side, leading to a sharper location of the peaks. Therefore, using the series expansion to calculate this moment, we were able to study the dispersion of the quasiparticle energy and, as a result, the evolution of the Fermi surface in the limit  $n \rightarrow 1$ , i.e., beyond the density regime currently accessible to the  $O(\lambda^2)$  version of the ECFL.

Here, we expand upon our previous findings and perform analytic continuation to obtain the full spectral functions. Direct analytic continuation of finite series, however, leads to unphysical results, e.g., negative spectral functions can arise due to the truncation of the series. This is a well-studied problem with known resolutions [15,16]. Therefore, and in particular, to ensure the positivity of spectral densities, we either take advantage of a transformation that guarantees this positivity, or assume an approximate form for the spectral functions, which comes out of the ECFL. We find a good agreement between results from the two approaches, which capture the expected features of the spectra discussed above.

Using strong-coupling expansions, there have been several earlier studies of the thermodynamics and time-independent correlations of the Hubbard and related models [17]. However, strong-coupling expansions for the *time-dependent correlations* are very rare [16,18,19]. In Ref. [16], the authors carried out a strong-coupling expansion for the Green's function to fifth order in  $\beta t$  for the finite- $U$  Hubbard model. Here, the simplifications arising from the  $U \rightarrow \infty$  limit allow us to go to eighth order in  $\beta t$ . This provides us with the opportunity to employ Padé approximations and study several static and dynamic quantities, such as the equation of state, momentum distribution function, the quasiparticle fraction, and lifetime at temperatures lower than the hopping amplitude, where the direct sums in the series do not converge. We also take advantage of the state-of-the-art numerical linked-cluster expansions (NLCEs) [20], developed recently for the  $t$ - $J$  model, and set the exchange interaction  $J$  to 0, to gauge our low-temperature equation of state obtained from the Padé approximations.

The organization of the paper is as follows: In Secs. II and III, we review the Metzner formalism and detail its numerical implementation. In Sec. IV, we provide analytical expressions for the Green's function and the Dyson-Mori self-energy in momentum and Matsubara frequency space as a function of the density. In Sec. V, we discuss the convergence of the series both before and after the use of Padé approximations. Additionally, using the series, we report results for the time-dependent local Green's function, the equation of state, Kelvin thermopower, the quasiparticle weight at the Fermi surface, momentum occupation number, and quasiparticle lifetime and spectral functions at different points along the irreducible wedge of the Brillouin zone. We summarize our results in Sec. VII.

## II. THE MODEL

In the strong-coupling limit, the Hubbard Hamiltonian is written as the sum of the unperturbed local Hamiltonian  $H_0$ , and a perturbation  $H_1$  that accounts for hopping of electrons between the sites of the lattice,

$$H = H_0 + H_1, \quad (1)$$

where

$$\begin{aligned} H_0 &= U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu \sum_{i\sigma} n_{i\sigma}, \\ H_1 &= - \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma}. \end{aligned} \quad (2)$$

Here,  $c_{i\sigma}$  ( $c_{i\sigma}^\dagger$ ) annihilates (creates) a fermion with spin  $\sigma$  on site  $i$ ,  $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$  is the number operator,  $U$  is the onsite repulsive Coulomb interaction,  $\mu$  is the chemical potential, and  $t_{ij}$  is the hopping matrix element between sites  $i$  and  $j$ . As discussed in the following, we allow for nearest-neighbor hopping only, namely,  $t_{ij} = t$  if  $i$  and  $j$  are nearest neighbors, and  $t_{ij} = 0$  otherwise.

## III. METZNER'S APPROACH FOR COMPUTING THE GREEN'S FUNCTION

We start by describing the Metzner formalism before turning our focus to topics related to its computational implementation in the limit of extreme correlations. Following the conventions in Ref. [1], we define the finite-temperature single-particle Green's function as

$$G_{\sigma jj'}(\tau - \tau') = -\langle T_\tau c_{j\sigma}(\tau) c_{j'\sigma}^\dagger(\tau') \rangle, \quad (3)$$

where  $\langle \dots \rangle$  denotes the thermal average with respect to  $H$ ,  $T_\tau$  denotes the imaginary time-ordering operator, and the creation and annihilation operators in the Heisenberg representation are expressed as

$$\begin{aligned} c_{j\sigma}^\dagger(\tau) &= e^{H\tau} c_{j\sigma}^\dagger e^{-H\tau}, \\ c_{j\sigma}(\tau) &= e^{H\tau} c_{j\sigma} e^{-H\tau}, \end{aligned} \quad (4)$$

where  $0 \leq \tau \leq \beta$  is an imaginary time variable.

To derive a perturbative expansion for  $G_{\sigma jj'}(\tau - \tau')$ , we switch to the interaction representation, where the time evolution of the operators is governed by the unperturbed

Hamiltonian  $H_0$ . The Green's function can then be expressed as

$$G_{\sigma jj'}(\tau - \tau') = -\langle T_\tau c_{j\sigma}(\tau) c_{j'\sigma}^\dagger(\tau') \mathcal{S} \rangle_0 / \langle \mathcal{S} \rangle_0, \quad (5)$$

where the expectation values  $\langle \dots \rangle_0$  are taken with respect to the unperturbed Hamiltonian, and  $\mathcal{S}$  is given by

$$\mathcal{S} = T_\tau \exp \left[ \int_0^\beta d\tau \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger(\tau) c_{j\sigma}(\tau) \right]. \quad (6)$$

Next, by expanding the exponential in Eq. (6), both the numerator and the denominator of Eq. (5) can be written as perturbative series expansions in  $t$ . As detailed in Ref. [1], every term of the expansions can be written in terms of cumulants (connected many-particle Green's functions) of the unperturbed system, denoted by  $C_m^0$  ( $m$  indicates the number of creation or destruction operators in the cumulant). Due to the local nature of the unperturbed Hamiltonian, the cumulants are site diagonal, i.e., the only nonzero ones are those whose site variables are the same, and they can therefore be indexed by site. Due to the translational invariance of the Hamiltonian, an order  $m$  cumulant at site  $i$  is independent of  $i$  and is a function of only the time and spin indices of the  $m$  creation, and  $m$  destruction operators acting on  $i$ , i.e.,  $C_{mi}^0 \equiv C_m^0(\tau_1 \sigma_1, \dots, \tau_m \sigma_m | \tau'_1 \sigma'_1, \dots, \tau'_m \sigma'_m)$ . As we will see in the following, this invariance is a major advantage of the present method. Using it, each term in the expansion can be written as a product of a spatial part and a temporal part, which may then be evaluated independently.

The terms in the expansion for  $\langle \mathcal{S} \rangle_0$  can be evaluated using a diagrammatic approach, where each diagram consists of vertices, and directed lines connecting the vertices. Each vertex represents a site on the lattice, and each line represents a hopping process between two sites. Furthermore, the spatial sums reduce to calculating free multiplicities of graphs when embedded on the lattice. This computationally inexpensive part of the algorithm can be carried out independently of the most expensive part (taking the time integrals), for *any* lattice geometry.

The expectation value in the numerator of Eq. (5) can be calculated the same way as  $\langle \mathcal{S} \rangle_0$ , except that any graph in the former contains two additional *external* lines, one entering the site  $j'$  at time  $\tau'$  and one exiting the site  $j$  at time  $\tau$ . Consequently, in the lattice sums, one has to "fix" the vertices to which the external lines attach to be the sites  $j$  and  $j'$  on the lattice with the desired separation between them.

Another important feature of this method is the fact that the spatial sums are unrestricted (different vertices are allowed to be on the same lattice site), and therefore it can be verified that the contributions of disconnected diagrams are products of the contributions of their connected components (the linked-cluster theorem holds). Hence, the disconnected diagrams in the numerator of Eq. (5) are canceled by the denominator, and  $G_{\sigma jj'}(\tau, \tau')$  is given as the sum of the contributions of only the connected graphs in its numerator.

Further details of the method are given in Ref. [1] and will not be repeated here. The rules mentioned in Ref. [1] for generating the graphs and evaluating their contributions are reproduced below.

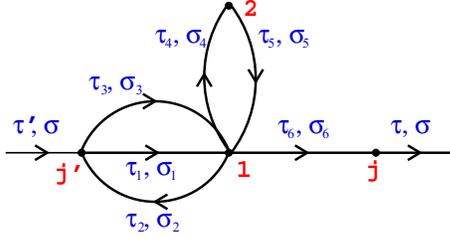


FIG. 1. (Color online) Diagram of a sample graph in the sixth order (with six internal and two external lines). The time and spin indices of lines are ordered according to rule (iii). To calculate the contribution of this graph, we need to insert  $C_2^0$ ,  $C_3^0$ ,  $C_1^0$ , and  $C_1^0$  for vertices  $j'$ , 1, 2, and  $j$ , respectively, for the time integral and the spin sum, and  $(t_{j'1})^3(t_{12})^2t_{1j}$  for the spatial sum. The symmetry factor is 2 since exchanging lines that correspond to  $\tau_1\sigma_1$  and  $\tau_3\sigma_3$  does not change the topology of the graph.

### A. Rules for calculating the one-particle Green's function diagrammatically

(i) Draw all topologically distinct diagrams: vertices connected by directed lines such that the number of entering and exiting lines at each vertex is the same. The graphs consist of the internal lines that connect two vertices as well as two *external* lines that enter a vertex and exit a vertex. The order to which each graph contributes is equal to the number of internal lines it has.

(ii) Label each line with a time and spin index, and each vertex by a lattice index. The vertex that has the entering external line is labeled by  $j'$  and the vertex that has the exiting external line is labeled by  $j$ .

(iii) Order the lines by defining a path that starts from the entering external line at vertex  $j'$ , goes through all of the vertices, and ends with the exiting external line at  $j$ . Figure 1 shows an example of such a graph in the sixth order.

(iv) Insert a factor of  $(-t_{ij})$  for each line that connects vertex  $i$  to vertex  $l$ .

(v) Insert  $C_m^0(\tau_1\sigma_1, \dots, \tau_m\sigma_m | \tau'_1\sigma'_1, \dots, \tau'_m\sigma'_m)$  for each vertex that has  $m$  entering lines labeled  $\tau'_1\sigma'_1, \dots, \tau'_m\sigma'_m$  and  $m$  exiting lines labeled  $\tau_1\sigma_1, \dots, \tau_m\sigma_m$ , such that  $\tau_i\sigma_i$  corresponds to the next line after  $\tau'_i\sigma'_i$  according to the ordering defined in (iii). This will ensure that there are no fermion loops in the diagram.

(vi) Determine the symmetry factor of the graph, which is the number of permutations of labeled lines and vertices that do not change its topology.

(vii) To calculate the contribution of the graph, integrate each *internal* time index between 0 and  $\beta$ , sum over the *internal* spatial and spin indices, and divide the result by the symmetry factor. As an example, the contribution of the graph  $c$  in Fig. 1 is

$$\begin{aligned}
 W(c) = & \frac{1}{2} \sum_{1,2} (t_{j'1})^3 (t_{12})^2 t_{1j} \int_0^\beta d\tau_1 \dots \int_0^\beta d\tau_6 \\
 & \times \sum_{\sigma_1 \dots \sigma_6} C_2^0(\tau_1\sigma_1, \tau_3\sigma_3 | \tau'\sigma, \tau_2\sigma_2) \\
 & \times C_3^0(\tau_2\sigma_2, \tau_4\sigma_4, \tau_6\sigma_6 | \tau_1\sigma_1, \tau_3\sigma_3, \tau_5\sigma_5) \\
 & \times C_1^0(\tau_5\sigma_5 | \tau_4\sigma_4) C_1^0(\tau\sigma | \tau_6\sigma_6). \quad (7)
 \end{aligned}$$

(viii) To obtain the  $l$ th-order contribution to the Green's function, add the contributions  $W(c)$ , of all the graphs with  $l$  internal lines:

$$G^{(l)} = \sum_{c \in \text{order } l} W(c). \quad (8)$$

In this scheme, the only zeroth-order graph consists of a vertex and the two external lines. In the first order, the only possible topology has two vertices, each having an external line, and a single internal line connecting them. In higher orders, the number of vertices can vary from two to  $l+1$ , where  $l$  denotes the order, depending on the topology. The topologically distinct graphs up to the fourth order are shown in Fig. 4 of Ref. [1].

### B. Computational implementation

We have implemented a computer program to perform all of the steps described in Sec. III A for the infinite- $U$  case. In this limit, since no double occupancy is allowed, the calculation of the cumulants simplifies drastically. This enables us to carry out the expansion to eighth order. In this subsection, we explain some of the details of this implementation at each step.

#### 1. Generation of topologically distinct graphs

To generate all topologically distinct diagrams in step (i) above, we need to have a way of uniquely identifying them in a computer program. For this, we use the concept of adjacency matrices. The elements of a  $m \times m$  adjacency matrix, where  $m$  is the number of vertices, represent the connections between every two vertices. For instance, for a graph with undirected lines between vertices, the  $(i, j)$  element can be an integer that simply counts the number of lines between vertices  $i$  and  $j$ . Here, since the lines are directed, we use a generalization of this matrix where every element is replaced by an array of size two. The first element of this array (we call it the left element) represents the number of incoming lines from vertex  $i$  to vertex  $j$  while the second element (or the right element) represents the number of outgoing lines from vertex  $j$  to vertex  $i$ .

One has to note that a topologically distinct graph cannot be uniquely represented by such an adjacency matrix since different labelings of the vertices, while not altering the topology, lead to different adjacency matrices. Therefore, one has to devise an algorithm to pick only one, out of  $m!$  possibilities, of the labelings of a graph to be able to establish a one-to-one correspondence between the graphs and its adjacency matrix. This can be done, for example, through *sorting* the adjacency matrix; by assigning the first row (column) to the vertex that possesses the largest number of lines, and so on. Alternatively, in our case, we can more simply employ the order of vertices that results from rule (iii) above.

After defining the mapping between the adjacency matrices and graphs in the computer algorithm, we generate graphs with  $m$  vertices by considering all possible numbers for the elements of the  $m \times 2m$  adjacency matrix, subject to the following two constraints: First, the number of incoming and outgoing lines at each vertex have to be the same, so, if we subtract the sum of left elements and the sum of right elements at each row (column) the result has to be zero. Second, the total

number of lines in the graph (or the sum of all elements of the matrix, divided by 2) should be equal to the desired order in the expansion. Note that, in this strong-coupling expansion, there is no line that leaves a vertex and then enters the same vertex, i.e., the diagonal elements of all adjacency matrices are zero.

## 2. Cumulants

We obtain cumulants to any order by taking functional derivatives of the generating functional with respect to Grassmann variables as described in Refs. [1,21]. As a result, a cumulant of order  $l$  is written in terms of the local unperturbed Green's function (UGF) of the same order  $G_l^0$ , as well as lower

order UGFs. In Appendix A, we show this expansion for the first few cumulants. The calculation of the cumulants then reduces to the evaluation of the UGFs, which, for order  $l$ , is the expectation value of  $2l$  time-ordered creation and annihilation operators with respect to the unperturbed Hamiltonian. For our case of the infinite- $U$  limit, since no double occupancy is allowed, a creation operator can only be followed by an annihilation operator and vice versa. Hence, the Green's function can assume only two distinct values depending on whether a creation or an annihilation operator is on the right side of the time-ordered product of operators. The two values are, respectively,  $(1 - \rho)$  and  $\frac{\rho}{2}$ , where  $\rho = \frac{2e^{\beta\mu}}{1+2e^{\beta\mu}}$  is the density in the atomic limit. For example, we end up with the following terms for the first two orders:

$$G_1^0(\tau_1\sigma_1|\tau_1'\sigma_1') = \langle T_\tau c_{j\sigma_1'}^\dagger(\tau_1')c_{j\sigma_1}(\tau_1) \rangle = e^{\mu(\tau_1-\tau_1')}\delta_{\sigma_1\sigma_1'} \left[ \frac{\rho}{2}\Theta(\tau_1'-\tau_1) - (1-\rho)\Theta(\tau_1-\tau_1') \right], \quad (9)$$

$$\begin{aligned} G_2^0(\tau_1\sigma_1, \tau_2\sigma_2|\tau_1'\sigma_1', \tau_2'\sigma_2') &= \langle T_\tau c_{j\sigma_1'}^\dagger(\tau_1')c_{j\sigma_1}(\tau_1)c_{j\sigma_2'}^\dagger(\tau_2')c_{j\sigma_2}(\tau_2) \rangle \\ &= e^{\mu(\tau_1+\tau_2-\tau_1'-\tau_2')} \sum_{qp} (-1)^q (-1)^p \left[ \frac{\rho}{2}\delta_{q\sigma_2 p\sigma_1'}\delta_{q\sigma_1 p\sigma_2'}\Theta(p\tau_1'-q\tau_1)\Theta(q\tau_1-p\tau_2')\Theta(p\tau_2'-q\tau_2) \right. \\ &\quad \left. + (1-\rho)\delta_{q\sigma_2 p\sigma_2'}\delta_{q\sigma_1 p\sigma_1'}\Theta(q\tau_1-p\tau_1')\Theta(p\tau_1'-q\tau_2)\Theta(q\tau_2-p\tau_2') \right], \quad (10) \end{aligned}$$

where the sum runs over permutations  $p$  and  $q$  of the time and spin indices of the primed and unprimed variables, respectively, and  $\Theta$  is the usual step function.

## 3. Free multiplicities

The spatial sums are performed for a specific lattice geometry. We have calculated them on the square lattice. In the computer program, we define a large enough lattice where we can fit any cluster with a number of sites at least twice as large as the maximum number of vertices in our largest order graphs. We then assign vertices  $j'$  (where an external line enters) and  $j$  (where an external line exits) to two lattice sites with a given displacement between them. The next part of the algorithm involves finding the number of possibilities for assigning the rest of the vertices to lattice sites. This can be done by following the path we have defined for each graph in rule (iii) to go from vertex  $j'$  to  $j$ . We start from vertex  $j'$  and in each step, we move to the next vertex in the list and assign a site to it. We ensure that if we come back to a vertex in the graph, we also come back to the corresponding site on the lattice. However, since we are calculating free multiplicities, we can assign the same lattice site to multiple vertices wherever the topology of the graph allows for it. In Table I, we show the number of topologically distinct graphs in each order, along with the number of graphs that have nonzero contributions on bipartite geometries, and the sum of free multiplicities for all graphs in each order for the (0,0) and (1,0) separations, up to the 10th order.

This computationally inexpensive process can be repeated for all possible separations (the maximum separation is set by the largest order considered). They can then be used to calculate the Fourier transform of the Green's function into the momentum space.

## 4. Time integrals

As seen in Sec. III B 2, the cumulants for the infinite- $U$  Hubbard model consist of products of only step functions and exponentials in the internal and external imaginary times. After multiplying several cumulants to obtain the contribution of a graph, we typically end up with a huge number of terms, each consisting of the product of a set of step functions of the time variables, the exponentials associated with the external times (the exponentials associated with the internal times cancel), Kronecker delta functions of the spin indices, and a function

TABLE I. Total number of topologically distinct graphs (second column), number of graphs that have nonzero multiplicity on bipartite geometries (third column), and the sum of multiplicities of all graphs for the smallest separations for which they have nonzero multiplicity (fourth column) at each order. The smallest separation for graphs with even number of lines (in even orders) is  $r_{j'} - r_j = (0,0)$ , and for graphs in odd orders is considered to be  $r_{j'} - r_j = (1,0)$ .

Order	Topologically distinct	Used for bipartite	$\sum$ Multiplicities
0	1	1	1
1	1	1	1
2	2	2	8
3	5	4	18
4	14	10	164
5	41	22	458
6	130	59	4240
7	431	146	13 544
8	1512	425	130 516
9	5542	1136	448 211
10	21 236	3497	4 408 216

of  $\rho$ . As mentioned before, one of the main advantages of our approach is that the time integrals over internal time variables can be taken independently of the spatial sums (free multiplicity calculations). We choose  $\tau' = 0$  without loss of generality since the Green's function is a function of  $\tau - \tau'$  and  $G(\tau - \tau' < 0)$  can be obtained from  $G(\tau - \tau' > 0)$  using the antiperiodicity of the Green's function in imaginary time [22]. To see how the time integrals are evaluated, we proceed with the following example. Suppose that one of the terms that belongs to a graph in the third order can be written as

$$\mathcal{I}(\tau) = \int_0^\beta \int_0^\beta \int_0^\beta d\tau_1 d\tau_2 d\tau_3 \Theta(\tau_1 - \tau_3) \Theta(\tau - \tau_3). \quad (11)$$

Note that in the above example, we have a smaller number of step functions in the integrand than typically expected for a term in the third order. However, the above combination is a perfectly valid one as the products of step functions are often simplified given that  $\Theta^n(x) = \Theta(x)$  for any nonzero  $n$ . The integral over  $\tau_2$  yields a factor  $\beta$  as there is no restriction on  $\tau_2$ . The remaining integrals are nonzero if  $\tau_1 > \tau_3$  and  $\tau > \tau_3$ . But, the latter condition does not uniquely determine the position of  $\tau_1$  relative to  $\tau$  in the  $[0, \beta]$  interval. Therefore, we consider the two possibilities,  $\tau > \tau_1$  and  $\tau < \tau_1$ , and rewrite the integral of Eq. (11) as

$$\begin{aligned} \mathcal{I}(\tau) &= \beta \int_0^\beta \int_0^\beta d\tau_1 d\tau_3 \\ &\times [\Theta(\tau_1 - \tau) \Theta(\tau - \tau_3) + \Theta(\tau - \tau_1) \Theta(\tau_1 - \tau_3)]. \end{aligned} \quad (12)$$

Note that for any value of  $\tau_1$  and  $\tau_3$ , only one of the terms in the integral in Eq. (12) is nonzero, justifying the equality. At this point, we can use the known results for the types of integrals in Eq. (12) (see Appendix B), leading to  $\beta[\tau(\beta - \tau) + \frac{\tau^2}{2}]$ .

Computationally, the two distinct possibilities for the ordering of times in the above example can be found by

generating all of the permutations of the time indices, and for each permutation, examining whether every step function in the product is nonzero. If that is the case, a multiplication of step functions corresponding to that permutation is inserted as the integrand.

### 5. Symmetry factor

Calculating the symmetry factor of each graph is straightforward in the framework of adjacency matrices. First, we note that the symmetry factor is proportional to the factorials of elements of the adjacency matrix in its upper triangle as they correspond to the number of permutations of directed lines that do not change the topology of the graph. Second, in order to find the symmetry factor related to those permutations of labeled vertices that leave the graph topology intact, we simply generate all the  $m!$  matrices that correspond to different orderings of vertex labels and find how many of them are the same as the original matrix. We then multiply this number by the factorials calculated in the first step to obtain the symmetry factor of the graph.

## IV. ANALYTICAL RESULTS

After evaluating the contribution of each diagram in a particular order by multiplying its free multiplicity for a given separation, time integral, and the spin sum, and dividing it by the symmetry factor, we add all of those contributions for that order to form the Green's function in terms of the atomic density  $\rho$ , the imaginary time  $\tau$ ,  $t$ ,  $\mu$ , and  $\beta$ . By calculating the spatial sums for all possible separations for each graph and performing a Fourier transformation on the space and imaginary time, one can express the Green's function in terms of the momentum  $k$ , and the Matsubara frequency  $\omega_n$ . Below, we show the resulting Green's function in the first four orders [23]:

$$\begin{aligned} G_\sigma^{(0)}(z, k) &= \frac{1 - \frac{\rho}{2}}{z}, \\ G_\sigma^{(1)}(z, k) &= \frac{(1 - \frac{\rho}{2})^2 \epsilon_k}{z^2}, \\ G_\sigma^{(2)}(z, k) &= \frac{(1 - \frac{\rho}{2})^3 \epsilon_k^2}{z^3} + \frac{(4 - \rho)\rho(1 - \frac{\rho}{2})t^2}{z^3} - \frac{2\beta(\rho - 1)\rho t^2}{z^2} + \frac{\beta^2 \rho [(3 - 2\rho)\rho - 1]t^2}{z}, \\ G_\sigma^{(3)}(z, k) &= \frac{(1 - \frac{\rho}{2})^4 \epsilon_k^3}{z^4} - \frac{7(\rho - 4)\rho(\rho - 2)^2 t^2 \epsilon_k}{16z^4} + \frac{3\beta(\rho - 1)\rho(\rho - 2)t^2 \epsilon_k}{2z^3} + \frac{\beta^2(\rho - 1)\rho[\rho(7\rho - 19) + 8]t^2 \epsilon_k}{4z^2}, \\ G_\sigma^{(4)}(z, k) &= \frac{(1 - \frac{\rho}{2})^5 \epsilon_k^4}{z^5} + \frac{5(\rho - 4)\rho(\rho - 2)^3 t^2 \epsilon_k^2}{16z^5} - \frac{\rho\{\rho[(\rho - 8)\rho - 152] + 240\}(\rho - 2)t^4}{8z^5} \\ &\quad - \frac{\beta(\rho - 1)\rho(\rho - 2)^2 t^2 \epsilon_k^2}{z^4} + \frac{\beta(\rho - 1)\rho[\rho(4\rho + 11) - 16]t^4}{z^4} - \frac{\beta^2(\rho - 1)\rho[\rho(5\rho - 14) + 6](\rho - 2)t^2 \epsilon_k^2}{4z^3} \\ &\quad + \frac{\beta^2(\rho - 1)\rho\{2\rho(5\rho - 24) + 43\} - 16\}t^4}{2z^3} - \frac{\beta^3(\rho - 1)\rho[\rho(97\rho - 100) + 18]t^4}{6z^2} \\ &\quad - \frac{\beta^4(\rho - 1)\rho\{\rho[388\rho - 591] + 236\} - 18\}t^4}{24z} \\ &\quad \vdots, \end{aligned} \quad (13)$$

where  $z = i\omega_n + \mu$ , and  $\epsilon_k = -2t[\cos(k_x) + \cos(k_y)]$ . Note that in this format, the Green's function is written in terms of the atomic density  $\rho$  or equivalently the chemical potential  $\mu$ , and not the true density for the many-body system,  $n = 1 + G_{jj\sigma}(\tau - \tau' = 0^+, \mu)$  [25]. By definition,  $n$ , too, can be written as an expansion in the hopping (using the expansion for the local Green's function). However, we would like to treat  $n$  as a parameter and rewrite the Green's function in terms of it. In that case, the chemical potential can no longer remain constant and we have to solve for it order by order in terms of  $n$  and  $t$ :  $\mu = \mu^{(0)} + \mu^{(2)} + \mu^{(4)} \dots$  where

$$n - 1 = G_{jj\sigma}^{(0)}(0^+, \mu^{(0)}) + G_{jj\sigma}^{(2)}(0^+, \mu^{(0)}) + \frac{dG_{jj\sigma}^{(0)}(0^+, \mu)}{d\mu} \Big|_{\mu=\mu^{(0)}} \mu^{(2)} + G_{jj\sigma}^{(4)}(0^+, \mu^{(0)}) + \frac{dG_{jj\sigma}^{(2)}(0^+, \mu)}{d\mu} \Big|_{\mu=\mu^{(0)}} \mu^{(2)} \\ + \frac{1}{2} \frac{d^2 G_{jj\sigma}^{(0)}(0^+, \mu)}{d\mu^2} \Big|_{\mu=\mu^{(0)}} (\mu^{(2)})^2 + \frac{dG_{jj\sigma}^{(0)}(0^+, \mu)}{d\mu} \Big|_{\mu=\mu^{(0)}} \mu^{(4)} + \dots \quad (14)$$

Inverting this equation for  $\mu$  in terms of  $n$ , we obtain

$$\mu^{(0)} = \frac{1}{\beta} \log \frac{n}{2(1-n)}, \quad \mu^{(2)} = 2(2n-1)t^2\beta, \quad \mu^{(4)} = \frac{1}{12}(6+n(n-4)(1+4n))t^4\beta^3 \\ \vdots \quad (15)$$

Finally, by inserting these back into the expansion for the momentum- and frequency-dependent Green's function order by order, we end up with the following terms for up to the fourth order [26]:

$$G_{\sigma}^{(0)}(z, k) = \frac{1 - \frac{n}{2}}{z}, \\ G_{\sigma}^{(1)}(z, k) = \frac{(1 - \frac{n}{2})^2 \epsilon_k}{z^2}, \\ G_{\sigma}^{(2)}(z, k) = \frac{(1 - \frac{n}{2})^3 \epsilon_k^2}{z^3} + \frac{[2(n-2) - n][2(n-1) - n]nt^2}{2z^3} - \frac{[2(n-1) + n]t^2\beta}{z^2}, \\ G_{\sigma}^{(3)}(z, k) = \frac{(1 - \frac{n}{2})^4 \epsilon_k^3}{z^4} - \frac{7[2(n-2) - n]n(2-n)^2 t^2 \epsilon_k}{16z^4} - \frac{[2(n-1) - n][2(2-3n) + (n-1)n]t^2 \beta \epsilon_k}{2z^3} - \frac{(n-1)^2 n^2 t^2 \beta^2 \epsilon_k}{4z^2}, \\ G_{\sigma}^{(4)}(z, k) = + \frac{(1 - \frac{n}{2})^5 \epsilon_k^4}{z^5} + \frac{5[2(n-2) - n][2(n-1) - n]^3 n t^2 \epsilon_k^2}{16z^5} + \frac{[2(n-1) - n]n(-n^3 + 8n^2 + 152n - 240)t^4}{8z^5} \\ + \frac{\{4(n-1)n - 6[2(n-1) + n]\}(2-n)^2 t^2 \beta \epsilon_k^2}{8z^4} \\ + \frac{n\{2[-3n^2 + 6(2n-3)n + 4(n-1)(9n-10)] + (1-n)n(4n+2)\}t^4 \beta}{2z^4} \\ + \frac{[4(2n^3 + 6n^2 - 10n + 4) - (n-1)n^2(4n^2 - 12n + 2)]t^4 \beta^2}{4z^3} + \frac{(-6n^3 + 68n^2 - 20n - 24)t^4 \beta^3}{48z^2} \\ \vdots, \quad (16)$$

where  $z = i\omega_n + \mu^{(0)}$ . It is perhaps even more useful to extract a self-energy from this expansion. The Dyson-Mori self-energy (denoted simply with  $\Sigma_{\text{DM}} \rightarrow \Sigma$ ) can be deduced using  $\Sigma(z, k) = z - a_G[\epsilon_k + G(z, k)^{-1}]$ , where  $a_G = (1 - n/2)$  [3].

$$\Sigma^{(0)}(z, k) = 0 \\ \Sigma^{(1)}(z, k) = 0 \\ \Sigma^{(2)}(z, k) = \frac{t^2 \beta (6n-4)}{n-2} - \frac{(n^2-4n)t^2}{z} \\ \Sigma^{(3)}(z, k) = \frac{n^2 t^2 \epsilon_k \beta^2 (1+n^2-2n)}{2(n-2)} - \frac{(n-4)(n-2)nt^2 \epsilon_k}{8z^2} + \frac{nt^2 \epsilon_k \beta (n-1)}{z} \\ \Sigma^{(4)}(z, k) = \frac{t^4 \beta^3 (12+3n^3-34n^2+10n)}{12(n-2)} - \frac{3(n^4-8n^3+72n^2-80n)t^4}{4z^3} \\ + \frac{2t^4 \beta (2n^4-40n^3+65n^2-24n)}{(n-2)z^2} + \frac{t^4 \beta^2 (2n^6-12n^5+19n^4-19n^3+10n^2)}{(n-2)^2 z} \\ \vdots \quad (17)$$

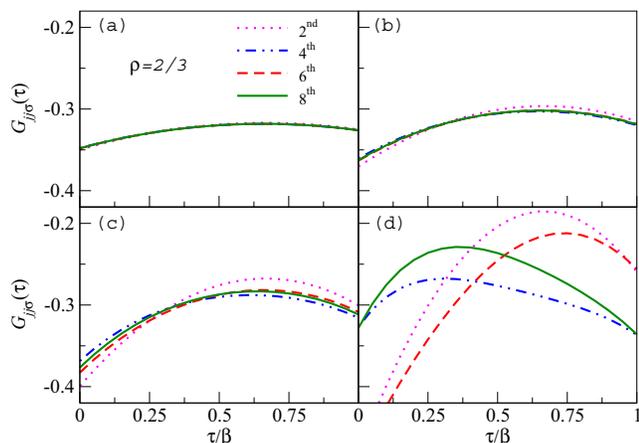


FIG. 2. (Color online) Local Green's function for a constant chemical potential  $\mu = 0$  vs imaginary time at (a)  $T = 3.0$ , (b)  $T = 2.0$ , (c)  $T = 1.5$ , and (d)  $T = 1.0$ .  $t = 1$  is the unit of energy throughout the paper.

### V. CONVERGENCE AND THE PADÉ APPROXIMATION

In Fig. 2, we show the local imaginary time Green's function for  $\rho = 2/3$ , corresponding to  $\mu = 0$ , at different temperatures. At  $T = 3.0$  (unless specified otherwise, we take  $t = 1$  as the unit of energy and work in units where  $k_B = 1$  throughout the paper), the series show very good convergence as expected in this high-temperature region. Note that odd terms in the series are zero for this local quantity. As we lower  $T$  to 2.0, there are some discrepancies between low orders, but the last two orders (6 and 8) still agree very well. This is no longer the case as we get closer to  $T = 1$ , below which the finite series is divergent by definition. This is because in the absence of any other energy scale in the system, an expansion in  $t$  can be viewed as an expansion in  $\beta$ . In other words,  $\beta^{m+1}$  always couples to  $t^m$  in the series for the Green's function. In Fig. 2(d), one can see large fluctuations between different orders already at  $T = 1.0$  and there is no clear picture from the bare results as to what the actual shape of the Green's function is.

To demonstrate the trends in the convergence of the series at other values of  $\mu$ , in Fig. 3, we show the equation of state at the same four temperatures as in Fig. 2. We also show the equation of state in the atomic limit ( $\rho$  vs  $\mu$ ). We find that the last two orders more or less agree with each other for all  $\mu$  at  $T \gtrsim 1.5$ . However, for  $T = 1$ , the convergence is lost in the vicinity of  $\mu = 0$ . This shows that the poor convergence of the local Green's function at this value of  $\mu$ , seen in Fig. 2(d), represents the worst case scenario. An important feature of the equation of state as observed in Fig. 3 is that even at these high temperatures, there are significant deviations of the many-body density from the density in the atomic limit near the extreme limits of  $n = 0$  and  $n = 1$ .

It is instructive now to study the temperature dependence of the density at a given  $\mu$ , and to find out how the region of convergence can be extended in temperature by the use of Padé approximations. In Fig. 4, we show the temperature dependence of the density for various positive and negative values of  $\mu$ . We show the direct sums as well as results after

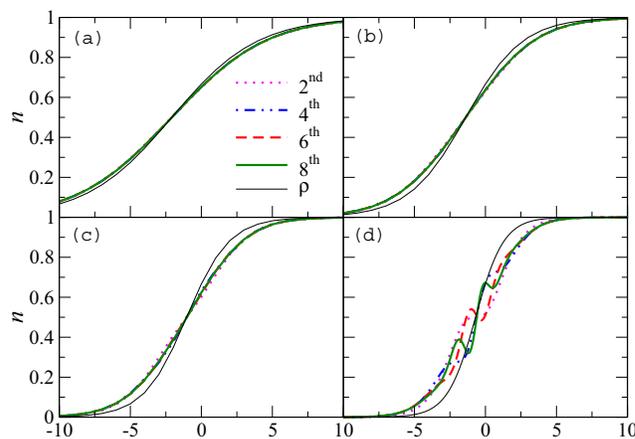


FIG. 3. (Color online) Density  $n$  as a function of the chemical potential at (a)  $T = 3.0$ , (b)  $T = 2.0$ , (c)  $T = 1.5$ , and (d)  $T = 1.0$ . Thin solid lines are the density in the atomic limit,  $\rho = \frac{2e^{\beta\mu}}{1+2e^{\beta\mu}}$ .

two different Padé approximations. The results in the atomic limit [ $\rho(T)$ ] are shown for  $\mu = 0$  and  $\pm 2.0$ . In the atomic limit, the system has two ground states depending on the sign of  $\mu$ . They correspond to  $\rho = 1$  and  $\rho = 0$  for positive and negative  $\mu$ , respectively. At exactly  $\mu = 0$ ,  $\rho$  is temperature independent at  $2/3$ . As one can see in Fig. 4, the real density for the many-body system has a qualitatively different behavior than  $\rho$  starting at relatively high temperatures. The temperature where  $n$  starts deviating from  $\rho$  due to correlations is around  $T \sim 2$  for  $\mu = -2$  and  $T \sim 5$  for  $\mu = 2$ . As expected, the density for  $\mu = 0$  falls below  $2/3$  for all  $T$ . To perform Padé approximation for  $n$  vs  $T$ , we first expand  $\rho$ , i.e., the

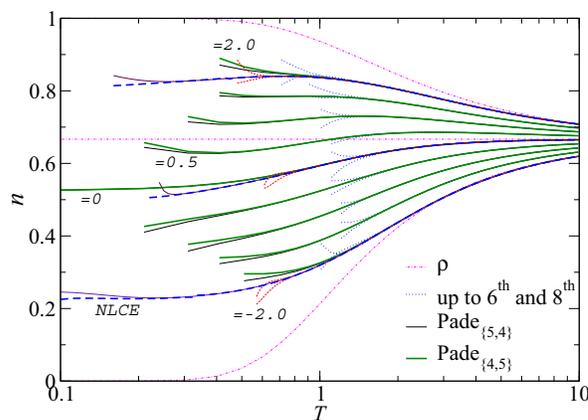


FIG. 4. (Color online) Average density  $n$  as a function of temperature for a range of  $\mu$  from  $-2.0$  to  $2.0$ , with the increment of  $0.5$ . The two indices of Padé in curly brackets indicate the order of the polynomials in the numerator and the denominator. From bottom to top, the dotted-dashed magenta lines are  $\rho$  for  $\mu = -2, 0$ , and  $2$ . We are also showing results from the NLCE for these three values of the chemical potential as thin dashed red lines (last two orders of the bare sums), and thick dashed blue and thin solid violet lines (after Wynn resummations with five and four cycles of improvement, respectively) [20].

zeroth-order term, in powers of  $\beta$  and then add the rest of the higher order terms from the series. Therefore, in the case of  $\mu = 0$ , where  $\rho$  is temperature independent, the odd powers of  $\beta$  in the series for  $n$  vanish and the two Padé approximants yield the same function, leading to  $n \sim 0.525$  for the ground state. Nevertheless, we cannot verify that this is the true value of the ground-state density of the system for  $\mu = 0$ .

The static properties of the model, such as the density, can in principle be obtained in higher orders by avoiding the relatively difficult calculation of the Green's function, and calculating only the free energy instead. However, for this purpose, we can also take advantage of the novel NLCE method that has been developed in recent years [20]. NLCE uses the same basis as high-temperature expansions, but calculates properties of finite clusters exactly, as opposed to perturbatively, using full diagonalization techniques. As a result, the convergence region of the NLCE is typically extended to lower temperatures in comparison to high-temperature expansions with the same number of terms.

In Fig. 4, we show results from the NLCE for the  $t$ - $J$  model with  $J = 0$  for up to the 11th order in the site expansion, where contributions of all clusters with up to 11 sites are considered, for  $\mu = 0$  and  $\pm 2.0$ . By comparing the direct sums in NLCE (thin dashed red lines represent the last two orders) with those from our series, we find that while we have perfect agreement between NLCE and the converged bare sums in the series, the Padé approximants overestimate the value of  $n$  in all cases at temperatures lower than one. The convergence of the NLCE results at low temperatures can be further improved using numerical resummations. Here, we show those obtained from the Wynn algorithm [20] by thin solid violet and thick dashed blue lines. Remarkably, the convergence is extended to  $T \sim 0.2$  for  $\mu = -2.0$ , and  $T \sim 0.3$  for  $\mu = 0$  and  $2.0$ . The results for  $\mu = 0$  show that the ground-state density is likely less than 0.525.

In Fig. 5, we plot the chemical potential of the system as a function of temperature for various fixed densities by inverting functions such as those seen in Fig. 4. Here, the dotted dashed lines represent the zeroth-order chemical potential  $\mu^{(0)}$  for a fixed density. They all approach zero as  $T \rightarrow 0$  since they correspond to the atomic limit. The results from the series and the NLCE suggest a different behavior starting at relatively high temperatures for the correlated system, except for the density near 0.5, where the linearity of the chemical potential, and the coincidence with the results from the atomic limit, is extended to low temperatures. This is consistent with the  $\mu = 0$  curve in Fig. 3 approaching  $n \sim 0.5$  at low temperatures. On the other hand, in the low density Fermi liquid regime, the low-temperature chemical potential is expected to be proportional to  $T^2$ . We find that the resumed NLCE results for  $n = 0.1$  agree with this behavior as they provide a reasonable fit to the function  $A + BT^2$ , as shown by a light blue (light gray) line in Fig. 5.

Another feature seen in the plots of chemical potential at fixed density, with potentially important implications for the state of the system, is the change in sign of the slope of  $\mu$  vs  $T$  at low temperatures. Recent theories of thermopower of correlated systems identify the Kelvin formula for thermopower

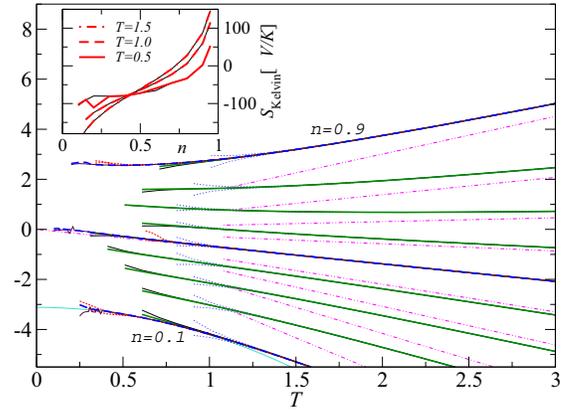


FIG. 5. (Color online) Chemical potential  $\mu$  at fixed density vs temperature for densities from  $n = 0.1$  to  $n = 0.9$  (from bottom to top with the increment of  $\Delta n = 0.1$ ). The lines are the same as in Fig. 4, except that the dotted-dashed magenta lines are the zeroth order of the chemical potential in the atomic limit, i.e.,  $\mu^{(0)} = T \log \frac{n}{2(1-n)}$ , and that thin solid lines are Padé<sub>(6,3)</sub>. Here, we show the NLCE results for  $n = 0.1, 0.5$ , and  $0.9$ . The light blue (light gray) solid line is the fit of the low-temperature NLCE results for  $n = 0.1$  after resummation to  $A + BT^2$  with  $A = -3.12$  and  $B = -1.10$ . The inset shows the Kelvin thermopower,  $S_{\text{Kelvin}}$ , from NLCE as defined in Eq. (18), in units of microvolts per degree Kelvin vs density. At each temperature, the two lines correspond to different Wynn resummations.

[27,28] by the expression,

$$S_{\text{Kelvin}} = \frac{-1}{q_e} \left( \frac{\partial \mu}{\partial T} \right)_{N,V} = \frac{1}{q_e} \left( \frac{\partial S}{\partial N} \right)_{T,V}, \quad (18)$$

where  $q_e = -|e|$  is the electron charge,  $S$  the entropy, and a Maxwell relation is employed in the second identity. This formula captures the considerations of Kelvin's famous paper on reciprocity in 1854 [29], within a contemporary setting. As explained in Refs. [27,28], this expression represents the “thermodynamic” contribution to the true thermopower in addition to the dynamical contributions, that are assumed small in many correlated systems and neglected here. We see from this expression that a flat chemical potential in temperature implies a maximum in entropy at the corresponding density, and locates a density where the thermopower changes sign (from electronlike to holelike), as often seen in correlated systems. From Fig. 5, we observe that  $\frac{\partial \mu}{\partial T} > 0$  and hence the Kelvin thermopower is positive for densities close to half filling, whereas near the empty band things are reversed and we get electronlike thermopower. The change in sign seems to arise at a density  $n$  between 0.7 and 0.9, somewhat greater than the value  $n = \frac{2}{3}$  from the naive atomic limit. A detailed discussion of the thermopower, and the related Hall constant in cuprates and in the two-dimensional  $t$ - $J$  model can be found in Refs. [28,30].

In Fig. 6, we show the analog of the quasiparticle fraction defined in the Matsubara frequency space as

$$Z_0(k) = \left[ 1 - \frac{\text{Im} \Sigma(\omega_0, k)}{\omega_0} \right]^{-1}, \quad (19)$$

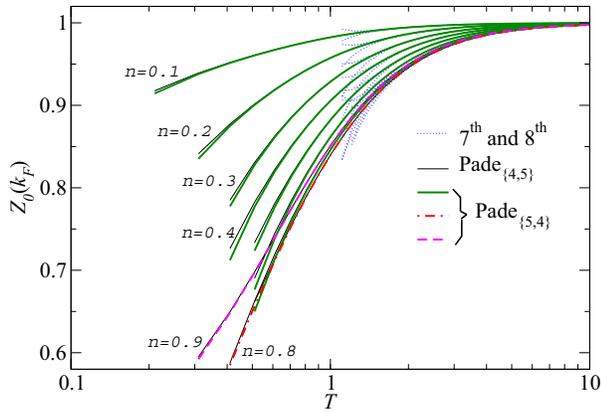


FIG. 6. (Color online) The quasiparticle fraction, defined in the Matsubara frequency space, at the nodal Fermi surface of the corresponding free Fermi gas, Eq. (19), after Padé approximation vs temperature for different values of density. At temperatures below one, the quasiparticle fraction initially decreases with increasing the density before increasing again for  $n > 0.7$ . The green thick solid lines are for  $n = 0.1 \dots 0.7$  from top to bottom.

where  $\omega_0 = \pi T$  is the lowest Matsubara frequency, as a function of temperature at various densities. We choose the momentum  $k$  to be the nodal Fermi vector of a free Fermi gas with the same density ( $k_F$ ). Previous studies based on the ECFL [2], or high-temperature expansions [31], suggest that this model possesses a Fermi surface coinciding with that of the free Fermi gas. The quantity in Eq. (19) will be equal to the actual quasiparticle fraction deduced from the self-energy in the real frequency axis,  $Z(k) = [1 - \frac{\partial \Sigma(\omega, k)}{\partial \omega}|_{\omega \rightarrow 0}]^{-1}$ , in the limit  $T \rightarrow 0$ . Therefore, the lowest temperatures we have access to may not be low enough to provide us with useful insight as to how the ground-state value of this quantity may vary with density. However, already at  $T \sim 0.5$ , Padé approximants offer an unexpected insight. We find that  $Z_0(k)$  decreases monotonically by increasing the density for  $n < 0.8$ , then increases as  $n$  increases to 0.9. Interestingly, the onset of this change of behavior coincides with that of the change of sign in the thermopower discussed earlier. As  $n \rightarrow 1$ , we do expect the true ground-state value of  $Z(k)$  to vanish, therefore this nonmonotonic dependence is presumably an artifact resulting from the finite  $T$  definition employed.

In Fig. 7, we show the momentum occupation number,  $m_k = \langle c_{k\sigma}^\dagger c_{k\sigma} \rangle$ , versus  $k$  at  $T = 0.77$  for different total densities. Features of this quantity at much lower temperatures were discussed in Ref. [5] for the  $t$ - $J$  model. However, the value of the density in the latter study was limited to  $n \lesssim 0.75$ . Here, we find that even at high temperatures, as the density approaches half filling, there is a huge redistribution of occupations in comparison to the free Fermi gas, as evidenced by the difference in  $m_k$  for  $n = 0.9$  between the two cases as seen in Fig. 7.

In a recent publication [14], the first moments of the electronic spectral functions of this model were studied using the same series expansion. It was shown that a modified first moment, (the “greater” moment) can better capture the location of the spectral peak at higher densities than the symmetric first moment. More information about the spectral properties of electrons in this model can be gathered from

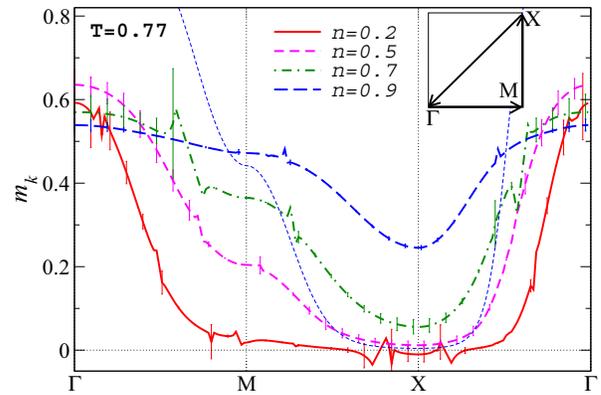


FIG. 7. (Color online) Momentum distribution function at  $T = 0.77$  for  $n = 0.2, 0.5, 0.7$ , and  $0.9$  vs momentum, as obtained from the average of the two Padé approximations ( $\{4,5\}$  and  $\{5,4\}$ ), around the irreducible wedge of the Brillouin zone as shown in the inset. Vertical lines show the difference between the two Padé approximants. The thin dashed line is the momentum occupation number of a free Fermi gas for  $n = 0.9$  at the same temperature.

higher order moments, also accessible through the series. In Fig. 8, we show the width of the quasiparticle peak, or the inverse lifetime, defined as

$$\Gamma^{-1}(k) = \sqrt{\varepsilon_2^>(k) - [\varepsilon_1^>(k)]^2}, \quad (20)$$

where  $\varepsilon_1^>(k)$  and  $\varepsilon_2^>(k)$  are the first and second greater moments, respectively, obtained from the series as described in Eq. (7) of Ref. [14]. Since the spectral function is largely skewed at higher densities [4], the width generally grows as the density increases.

## VI. SPECTRAL FUNCTIONS

We next turn to a study of the spectral functions  $\rho_G(k, \omega)$ , denoted by  $A(\omega, k)$  in standard photoemission studies. This can be found from the usual relation  $\rho_G(\omega, k) \equiv -\frac{1}{\pi} \text{Im} G(\omega + \mu^{(0)} + i\eta, k)$ , and requires a knowledge of the Greens function

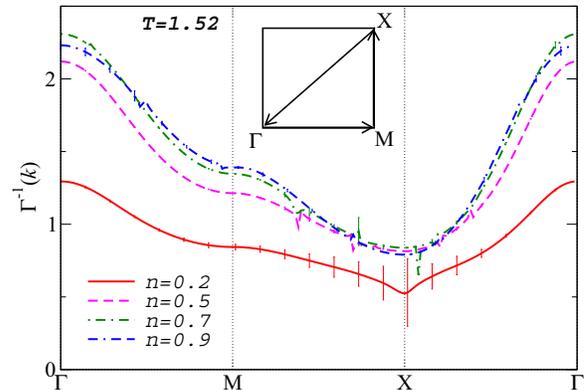


FIG. 8. (Color online) Inverse lifetime, defined in Eq. (20), at  $T = 1.52$  and for  $n = 0.2, 0.5, 0.7$ , and  $0.9$  vs momentum around the irreducible wedge of the Brillouin zone shown in the inset.  $t = 1$  sets the unit of energy. Lines are the same as in Fig. 7.

for complex frequencies. To extract spectral functions, we represent our Green's function as a continued fraction, which, when Taylor expanded to eighth order, reproduces Eq. (16). That is, we write  $G$  as (see Ref. [15] for the notation)

$$G(z, k) = \frac{a_G}{z + b_1 -} \frac{a_1}{z + b_2 -} \frac{a_2}{z + b_3 -} \frac{a_3}{z + b_4 -} \frac{a_4}{z}, \quad (21)$$

where  $a_l > 0$  and  $b_l$  are real. As explained in Ref. [15] (see also [16]), these conditions ensure that the resulting spectral function obtained from analytic continuation is positive definite. The formulas for the  $a_l$  and  $b_l$  can be obtained by suitably combining the "raw" moments; this procedure is detailed in Ref. [32]. In the infinite- $U$  Hubbard model, we know *a priori* how many floors will be in the continued fraction representation of a Green's function series of a given order. This is because the constants  $b_l$  have units of energy (and must therefore to leading order go like  $t$ ), and the constants  $a_l$  have units of energy squared (and must therefore to leading order go like  $t^2$ ). Therefore, we know that Eq. (21) is the correct, i.e., maximal continued fraction form obtainable from our eighth-order series. This is an advantage over the case of the finite- $U$  Hubbard model (see Ref. [16]), where the presence of the energy scale  $U$  means that the number of floors necessary to represent a series of a given order must be determined empirically.

In Ref. [3], Shastry establishes the relationship between the continued fraction representation of the Green's function [Eq. (21)], and a representation in terms of an infinite sequence of self-energies with spectral densities  $\rho_\Sigma^{(n)}(\omega)$ , with  $n = 0, 1, \dots$ . For the standard self-energy we omit the superscript so that  $\rho_\Sigma^{(0)}(\omega) \equiv \rho_\Sigma(\omega)$ . This is a particularly convenient reformulation of the well-known Mori scheme [33] for relaxation processes, where Laplace transforms over time-dependent correlations are used. In particular, denoting  $\Sigma_\infty \equiv \lim_{z \rightarrow \infty} \Sigma(z)$ , and recalling that

$$G(z, k) = \frac{a_G}{i\omega + \mu - a_G \epsilon_k - \Sigma_\infty - \int dv \frac{\rho_\Sigma^{(1)}(v)}{i\omega - v}},$$

$b_1 = -a_G \epsilon_k - \Sigma_\infty$ , and the standard self-energy is expressed as

$$\int \frac{\rho_\Sigma^{(1)}(v - \mu^{(0)})}{z - v} dv = \frac{a_1}{z + b_2 -} \frac{a_2}{z + b_3 -} \frac{a_3}{z + b_4 -} \frac{a_4}{z}, \quad (22)$$

where  $\rho_\Sigma(\omega) \equiv -\frac{1}{\pi} \text{Im} \Sigma(i\omega_n \rightarrow \omega + i\eta)$ . Following [3], we identify the constant  $a_1 \equiv a_\Sigma \equiv \int \rho_\Sigma^{(1)}(v) dv$ ,  $b_2 \equiv -\Sigma_\infty^{(1)}$ , and

$$\int \frac{\rho_\Sigma^{(1)}(v - \mu^{(0)})}{z - v} dv = \frac{a_2}{z + b_3 -} \frac{a_3}{z + b_4 -} \frac{a_4}{z}. \quad (23)$$

For  $l > 1$ , one has the general formula,

$$a_l = a_{\Sigma^{(l-1)}}; \quad b_l = -\Sigma_\infty^{(l-1)}. \quad (24)$$

The Green's function of Eq. (21) will lead to a spectral function with a small number of well-separated poles and residues. To obtain a continuous shape for the spectral function, there are several alternatives. We initially follow the procedure of Tomita and Mashiyama (TM) [34], which is useful in the spin relaxation problems, but does not seem to have features of a fermionic self-energy function built into it. Nevertheless,

we try it out in view of its simplicity, and as it provides a counterpoint to our preferred method presented next. In the spirit of Ref. [34], we assume that

$$\rho_\Sigma(\omega - \mu^{(0)}) = A \exp[-\alpha^2(\omega - \omega_0)^2], \quad (25)$$

so that the coefficients  $A, \alpha, \omega_0$  are fixed using the moments, and higher moments are forced to be those of the Gaussian. Using Eq. (22), we can solve for  $A, \alpha$ , and  $\omega_0$  in terms of  $a_1, a_2$ , and  $b_2$ . It is also possible to obtain a continuous spectral function whose moments correctly reproduce all of the coefficients in Eq. (21) by making the Gaussian approximation for the second-level self-energy:

$$\rho_\Sigma^{(2)}(\omega - \mu^{(0)}) = A \exp[-\alpha^2(\omega - \omega_0)^2]. \quad (26)$$

Then, using the relation,

$$\int \frac{\rho_\Sigma^{(2)}(v - \mu^{(0)})}{z - v} dv = \frac{a_3}{z + b_4 -} \frac{a_4}{z + \dots}, \quad (27)$$

we can solve for  $A, \alpha$ , and  $\omega_0$  in terms of  $a_3, a_4$ , and  $b_4$ . However, as shown in Fig. 9 below, this is actually a worse approximation as it accentuates an unphysical sharp peak in the TM scheme spectral function.

An alternative scheme for obtaining continuous spectral functions makes use of our knowledge of the approximate form of the self-energy as  $(T, \omega) \rightarrow 0$  [7]:

$$\rho_\Sigma(\omega) = A(\omega^2 + \pi^2 T^2) \left(1 - \frac{\omega}{\Delta}\right) \exp\left[-\frac{\omega^2 + \pi^2 T^2}{\omega_c^2}\right]. \quad (28)$$

Here,  $(\omega^2 + \pi^2 T^2)$  is the standard Fermi-liquid form,  $\frac{1}{\Delta}$  provides the aforementioned particle-hole asymmetry, and the exponential extrapolates the low-energy answer to higher energies in a natural way [2]. Once again, we can solve for  $A, \Delta$ , and  $\omega_c$  in terms of  $a_1, a_2$ , and  $b_2$  by using Eq. (22).

We obtain the spectral function  $\rho_G(\omega, k)$  using both Eqs. (25) and (28) at  $T = 1.1$  for  $n = 0.7$  and  $n = 0.9$  and at various points along the irreducible wedge of the Brillouin

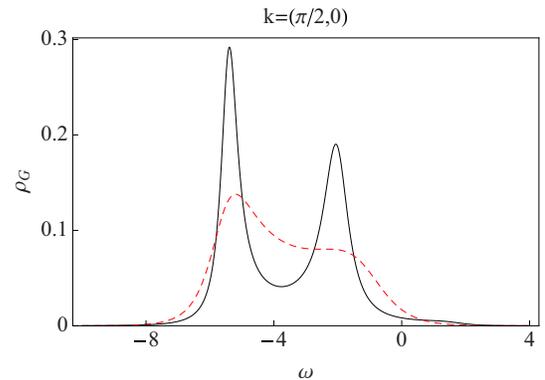


FIG. 9. (Color online) The spectral density for the physical Green's function vs  $\omega$  for  $T = 1.1$  and  $n = 0.9$ .  $t = 1$  sets the unit of energy. The red (dashed) curve is obtained from the TM scheme with the self-energy Eq. (25) and the black (solid) curve is obtained from the TM scheme with the second level self-energy [Eq. (26)]. The latter accentuates the unphysical secondary peak of the TM scheme spectral function.

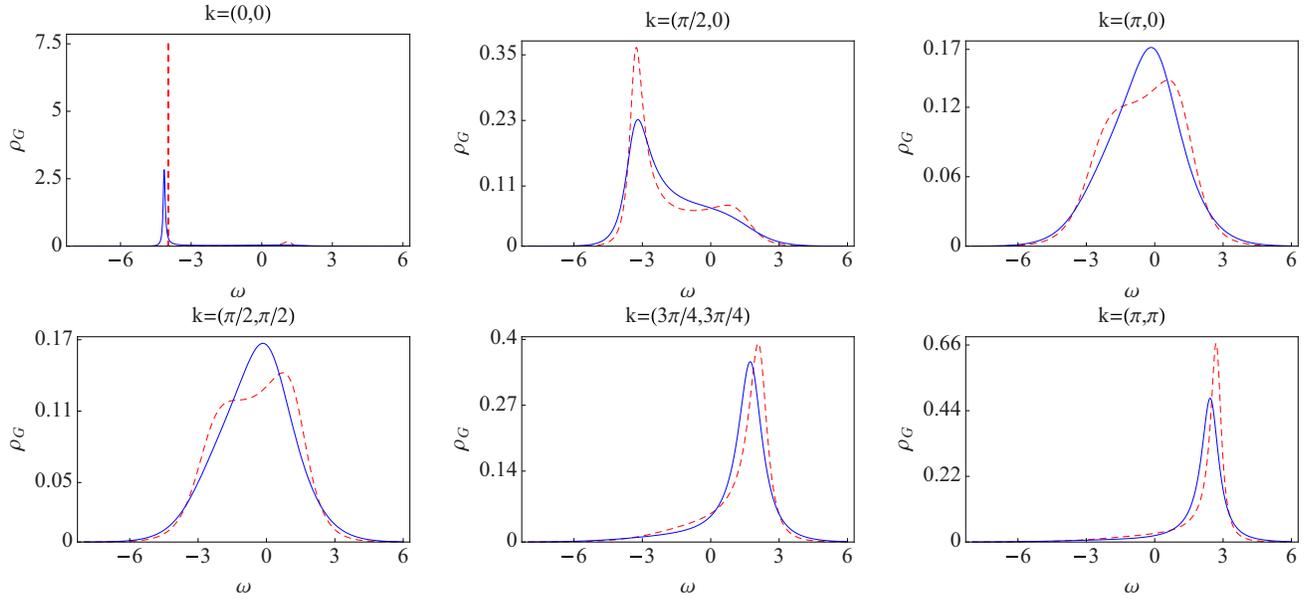


FIG. 10. (Color online) The spectral density for the physical Green's function vs  $\omega$  for  $T = 1.1$  and  $n = 0.7$ . The blue (solid) curve is obtained from the Fermi-liquid-type scheme [Eq. (28)] and the red (dashed) curve is obtained from the TM scheme [Eq. (25)]. The fairly sharp extra peaks obtained from the TM scheme, as compared to the Fermi-liquid scheme, seem to be physically unreasonable. We also note that the spectral functions from ECFL found numerically using the  $O(\lambda^2)$  scheme (see Fig. 3(f) of Ref. [14]) find rather broad peaks at high  $T$ .

zone. The spectral functions  $\rho_G(\omega, k)$  are plotted in Fig. 10 for  $n = 0.7$  and in Fig. 11 for  $n = 0.9$ .

## VII. SUMMARY

We present an implementation of the linked-cluster expansion for the Green's function of the infinite- $U$  Hubbard model on a computer, which is based on a formalism proposed by

Metzner [1]. Using efficient algorithms on parallel computers, we have carried out the expansion up to the eighth order in terms of the hopping amplitude, and obtained analytic results for the Green's function and the Dyson-Mori self-energy on the square lattice as a function of momentum and Matsubara frequency at a given fixed density. Since the lattice sums for graphs in this approach are evaluated independently of their time integrals and spin sums, our implementation paves the

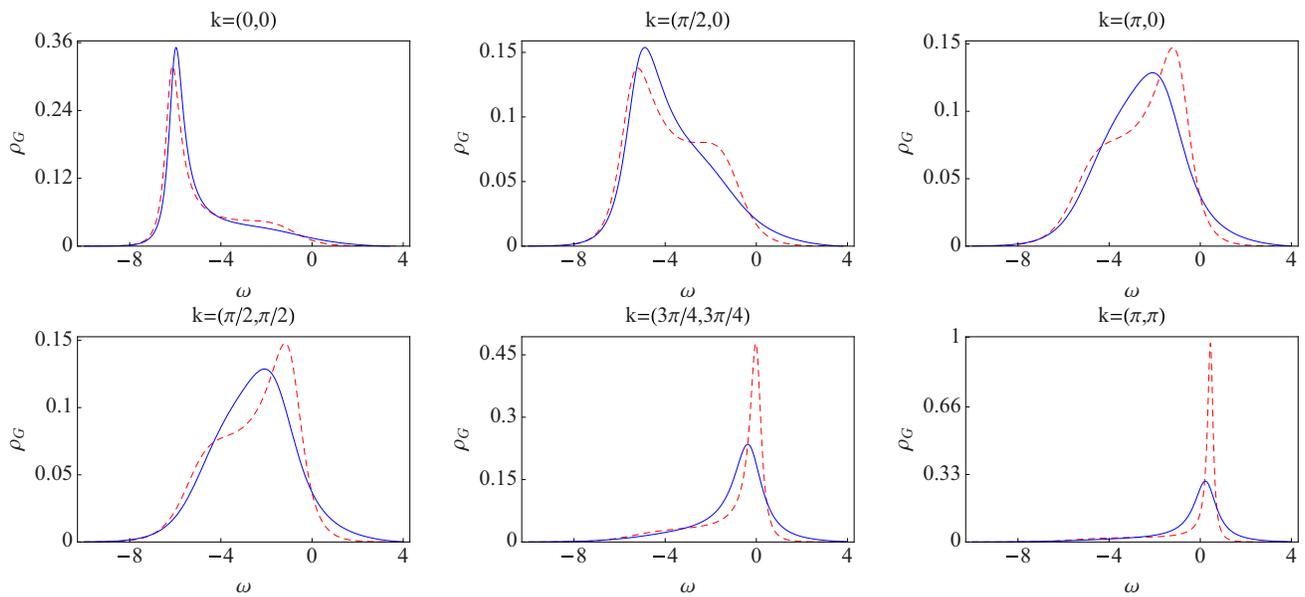


FIG. 11. (Color online) The spectral density for the physical Green's function vs  $\omega$  for  $T = 1.1$  and  $n = 0.9$ . Lines are the same as in Fig. 10.

way for obtaining similar results for other geometries and spatial dimensions.

To extend the region of convergence in temperature, we employ Padé approximations and study several static and dynamic quantities. The equation of state exhibits significant deviations from the atomic limit starting at relatively high temperatures and reveals interesting trends near  $n = 0.5$ , where we find that the chemical potential changes linearly with temperature and remains very close to the one in the atomic limit down to the lowest temperatures accessible to us. We also find that the change in sign of the derivative of  $\mu$  with respect to  $T$  at constant density, which is proportional to the thermopower in Kelvin's formula, takes place at increasingly higher densities due to correlations as the temperature is lowered. The momentum distribution function also shows significant deviations from free fermions, and becomes more uniform across the Brillouin zone as the correlations build up at higher densities. We further study dynamic quantities, such as the analog of the quasiparticle fraction in the Matsubara frequency space vs temperature, which shows a nonmonotonic dependence on density at low temperatures, and the lifetime of the quasiparticles at various densities, obtained in the series through the first two moments of the electronic spectral functions. To make contact with experiments and extend previous results for the spectral functions obtained within the ECFL or the dynamical mean-field theory, we calculate them here after transforming the Green's function series to continued fractions, or by employing certain forms for the spectral functions suggested by the ECFL theory. We present our results for densities close to half filling at several points in the momentum space.

To benchmark our results from the Padé approximations for the equation of state at temperatures lower than the hopping amplitude, where the direct sums in the series do not converge, and to shed more light on the state of the system at those temperatures, we also present results from the NLCE up to eleventh order for an equivalent model, i.e., the  $t$ - $J$  model with  $J = 0$ . We find perfect agreement between the direct sums from the two methods when they converge, and that at lower temperatures, the Padé approximants generally overestimate the density for a given chemical potential. The NLCE results after numerical resummations also help obtain the thermopower vs density at a temperature that is not otherwise accessible to the series even after the Padé approximations.

### ACKNOWLEDGMENT

This work was supported by DOE under Grant No. FG02-06ER46319 (B.S.S. and E.P.), and by NSF under Grant No. OCI-0904597 (E.K. and M.R.).

### APPENDIX A: RECURSIVE EXPANSION OF CUMULANTS

In the following, we combine the time and spin variables and denote them by their index only, i.e.,  $C_m^0(\tau_1\sigma_1, \dots, \tau_m\sigma_m | \tau'_1\sigma'_1, \dots, \tau'_m\sigma'_m) \rightarrow C_m^0(1, \dots, m | 1', \dots, m')$ . Cumulants are calculated by taking functional derivatives of a generating functional with respect to Grassmann variables [1,21], and can be expressed in terms of UGFs. We give explicit expressions for  $C_m^0$  through  $m = 3$ .

$$C_1^0(1|1') = G_1^0(1|1'), \quad C_2^0(1,2|1',2') = G_2^0(1,2|1',2') - G_1^0(1|1')G_1^0(2|2') + G_1^0(1|2')G_1^0(2|1'), \quad (\text{A1})$$

$$\begin{aligned} C_3^0(1,2,3|1',2',3') &= G_3^0(1,2,3|1',2',3') - C_2^0(1,2|1',2')G_1^0(3|3') + C_2^0(1,2|1',3')G_1^0(3|2') - C_2^0(1,2|2',3')G_1^0(3|1') \\ &\quad + C_2^0(1,3|1',2')G_1^0(2|3') + C_2^0(1,3|2',3')G_1^0(2|1') - C_2^0(1,3|1',3')G_1^0(2|2') \\ &\quad - C_2^0(2,3|1',2')G_1^0(1|3') - C_2^0(2,3|2',3')G_1^0(1|1') + C_2^0(2,3|1',3')G_1^0(1|2') \\ &\quad - G_1^0(1|1')G_1^0(2|2')G_1^0(3|3') + G_1^0(1|1')G_1^0(2|3')G_1^0(3|2') + G_1^0(1|2')G_1^0(2|1')G_1^0(3|3') \\ &\quad - G_1^0(1|2')G_1^0(2|3')G_1^0(3|1') + G_1^0(1|3')G_1^0(2|2')G_1^0(3|1') - G_1^0(1|3')G_1^0(2|1')G_1^0(3|2'). \end{aligned} \quad (\text{A2})$$

The rule for obtaining the expansion for  $C_m^0(1, \dots, m | 1', \dots, m') - G_m^0(1, \dots, m | 1', \dots, m')$  is as follows. Partition the unprimed integers  $1 \dots m$  into at least two sets. Each set in the partition corresponds to a cumulant, in which the unprimed numbers in the set are written in ascending order. The primed numbers  $1' \dots m'$  are then partitioned amongst the cumulants created by the unprimed number partitions, and are also written in ascending order. The sign of the term is (+) if the permutation to get from primed to unprimed numbers is odd, and (−) if it is even. The sign is due to the Grassmann variables in the generating functional, and is ultimately a consequence of the fermionic nature of the operators.  $C_3^0(1,2,3|1',2',3')$  can be expressed in terms of the UGFs by plugging Eq. (A1) into Eq. (A2). In general,  $C_m^0(1, \dots, m | 1', \dots, m')$  can be obtained in terms of UGFs of equal or lower orders by this recursive procedure.

### APPENDIX B: TIME INTEGRALS

In evaluating the time integrals, we use the following general result for the time integral of a product of step functions in terms of a series of ordered internal times  $\tau_i$ , over which the integrals are taken, and a fixed external time  $\tau$ :

$$\begin{aligned} &\int_0^\beta d\tau_n \int_0^\beta d\tau_{n-1} \dots \int_0^\beta d\tau_2 \int_0^\beta d\tau_1 \Theta(\tau_n - \tau_{n-1})\Theta(\tau_{n-1} - \tau_{n-2}) \dots \Theta(\tau_{m+1} - \tau)\Theta(\tau - \tau_m) \dots \Theta(\tau_3 - \tau_2)\Theta(\tau_2 - \tau_1) \\ &= \frac{\tau^m(\beta - \tau)^{n-m}}{m!(n-m)!}. \end{aligned} \quad (\text{B1})$$

- [1] W. Metzner, *Phys. Rev. B* **43**, 8549 (1991).
- [2] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011); **108**, 029702 (2012); *Phys. Rev. B* **87**, 125124 (2013).
- [3] B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011).
- [4] B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).
- [5] D. Hansen and B. S. Shastry, *Phys. Rev. B* **87**, 245101 (2013).
- [6] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, [arXiv:1307.3492](https://arxiv.org/abs/1307.3492); *Phys. Rev. B* **88**, 205108 (2013).
- [7] R. Zitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013).
- [8] E. Perepelitsky and B. S. Shastry, *Ann. Phys.* **338**, 283 (2013).
- [9] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [10] X. Deng, J. Mravlje, R. Zitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [11] G. Palsson and G. Kotliar, *Phys. Rev. Lett.* **80**, 4775 (1998).
- [12] L. F. Arsenault, B. S. Shastry, P. Sémon, and A.-M. S. Tremblay, *Phys. Rev. B* **87**, 035126 (2013).
- [13] G. H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011); K. Matsuyama and G. H. Gweon, *ibid.* **111**, 246401 (2013).
- [14] E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, and B. S. Shastry, *Phys. Rev. B* **87**, 161120 (2013).
- [15] J. A. Shohat and J. D. Tamarkin *The Problem of Moments* (American Mathematical Society, Providence, 1943).
- [16] S. Pairault, D. Senechal, and A.-M. S. Tremblay, *Eur. Phys. J. B* **16**, 85 (2000).
- [17] M. Plischke, *J. Stat. Phys.* **11**, 159 (1974); K. Kubo and M. Tada, *Progr. Theor. Phys.* **69**, 1345 (1983); **71**, 479 (1984); K. K. Pan and Y. L. Wang, *Phys. Rev. B* **43**, 3706 (1991); M. Bartkowiak and K. A. Chao, *ibid.* **46**, 9228 (1992); W. O. Putikka, M. U. Luchini, and R. R. P. Singh, *Phys. Rev. Lett.* **81**, 2966 (1998); V. W. Scarola, L. Pollet, J. Oitmaa, and M. Troyer, *ibid.* **102**, 135302 (2009); L. De Leo, J.-S. Bernier, C. Kollath, A. Georges, and V. W. Scarola, *Phys. Rev. A* **83**, 023606 (2011).
- [18] There have been a number of efforts to perform a strong-coupling expansion for the Green's function in a self-consistent manner, i.e., by resumming an infinite subclass of diagrams [16,19].
- [19] L. Craco and M. A. Gusmao, *Phys. Rev. B* **52**, 17135 (1995); **54**, 1629 (1996); L. Craco, *J. Phys.: Condens. Matter* **13**, 263 (2001); M. E. Foglio, T. Lobo, and M. S. Figueira, [arXiv:0903.0139](https://arxiv.org/abs/0903.0139); T. Lobo, M. S. Figueira, and M. E. Foglio, *Nanotechnology* **21**, 274007 (2010).
- [20] M. Rigol, T. Bryant, and R. R. P. Singh, *Phys. Rev. Lett.* **97**, 187202 (2006); *Phys. Rev. E* **75**, 061119 (2007).
- [21] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, New York, 1988).
- [22] A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, 1963).
- [23] This series has also been calculated by hand through fourth order in Ref. [24] using a completely independent method. The results of the two series match exactly.
- [24] E. Perepelitsky, [arXiv:1310.3797](https://arxiv.org/abs/1310.3797).
- [25] Note that when  $U = \infty$ , the creation and annihilation operators in Eq. (3) become Gutzwiller projected, and hence, follow different anticommutation relations than those valid for canonical fermion operators.
- [26] Higher orders for the square lattice or other geometries are available upon request.
- [27] M. R. Peterson and B. S. Shastry, *Phys. Rev. B* **82**, 195105 (2010).
- [28] B. S. Shastry, *Rep. Prog. Phys.* **72**, 016501 (2009).
- [29] W. Thomson (Lord Kelvin), Proc. R. Soc. Edinburgh 123, Collected Papers I, 23741 (1854).
- [30] A. Garg, B. S. Shastry, K. B. Dave, and P. Phillips, *New J. Phys.* **13**, 083032 (2011).
- [31] R. R. P. Singh and R. L. Glenister, *Phys. Rev. B* **46**, 14313 (1992).
- [32] M. Dupuis, *Progr. Theor. Phys.* **37**, 502 (1967).
- [33] H. Mori, *Progr. Theor. Phys.* **33**, 423 (1965).
- [34] K. Tomita and H. Mashiyama, *Progr. Theor. Phys.* **51**, 1312 (1974).



Contents lists available at ScienceDirect

## Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)

# Diagrammatic $\lambda$ series for extremely correlated Fermi liquids



Edward Perepelitsky\*, B. Sriram Shastry

Physics Department, University of California, Santa Cruz, CA 95064, USA

Centre de Physique Théorique, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex, France  
Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France

## ARTICLE INFO

## Article history:

Received 28 October 2014

Accepted 7 March 2015

Available online 14 March 2015

## Keywords:

 $\lambda$  expansion $t$ - $J$  model

Hubbard model

Extremely correlated Fermi liquid model

Strongly correlated electrons

## ABSTRACT

The recently developed theory of extremely correlated Fermi liquids (ECFL), applicable to models involving the physics of Gutzwiller projected electrons, shows considerable promise in understanding the phenomena displayed by the  $t$ - $J$  model. Its formal equations for the Greens function are reformulated by a new procedure that is intuitively close to that used in the usual Feynman–Dyson theory. We provide a systematic procedure by which one can draw diagrams for the  $\lambda$ -expansion of the ECFL introduced in Shastry (2011), where the parameter  $\lambda \in (0, 1)$  counts the order of the terms. In contrast to the Schwinger method originally used for this problem, we are able to write down the  $n$ th order diagrams ( $O(\lambda^n)$ ) directly with the appropriate coefficients, without enumerating *all* the previous order terms. This is a considerable advantage since it thereby enables the possible implementation of Monte Carlo methods to evaluate the  $\lambda$  series directly. The new procedure also provides a useful and intuitive alternative to the earlier methods.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

### 1.1. Motivation

The  $t$ - $J$  model is a model of fundamental importance in condensed matter physics, and is supposed to have the necessary ingredients to explain the physics of the high-temperature cuprates [1]. Its

\* Corresponding author at: Physics Department, University of California, Santa Cruz, CA 95064, USA.  
E-mail address: [eperepel@ucsc.edu](mailto:eperepel@ucsc.edu) (E. Perepelitsky).

Hamiltonian can be written in terms of the Hubbard  $X$  operators as [2]

$$H = - \sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i\sigma} X_i^{\sigma\sigma} + \frac{1}{2} \sum_{ij\sigma} J_{ij} X_i^{\sigma\sigma} + \frac{1}{4} \sum_{ij\sigma_1\sigma_2} J_{ij} \{X_i^{\sigma_1\sigma_2} X_j^{\sigma_2\sigma_1} - X_i^{\sigma_1\sigma_1} X_j^{\sigma_2\sigma_2}\}. \quad (1)$$

The operator  $X_i^{ab} = |a\rangle\langle b|$  takes the electron at site  $i$  from the state  $|b\rangle$  to the state  $|a\rangle$ , where  $|a\rangle$  and  $|b\rangle$  are one of the two occupied states  $|\uparrow\rangle, |\downarrow\rangle$ , or the unoccupied state  $|0\rangle$ . In terms of electron operators  $C, C^\dagger$ , and the Gutzwiller projection operator  $P_G$  that eliminates double occupancy, we may explicitly write  $X_i^{\sigma 0} = P_G C_{i\sigma}^\dagger P_G$ ,  $X_i^{0\sigma} = P_G C_{i\sigma} P_G$  and  $X_i^{\sigma\sigma'} = C_{i\sigma}^\dagger C_{i\sigma'} P_G$ . The key object of study for this model is the single-particle Green's function, given by the expression

$$\mathcal{G}_{\sigma_1\sigma_2}(i, f) = -\langle T_\tau X_i^{0\sigma_1}(\tau_i) X_f^{\sigma_2 0}(\tau_f) \rangle, \quad (2)$$

as well as higher order dynamical correlation functions. Several novel approaches for computing these objects have been tried in literature [3–8], but it has been found difficult to impose the Luttinger–Ward volume theorem in a consistent way, while providing a realistic description of both quasiparticle peaks and background terms in the spectral function.

The essential difficulties in computing these objects are (I) the non-canonical nature of the  $X$  operators, and hence the absence of the standard Wick's theorem, and (II) the lack of a convenient expansion parameter. In the recently developed *extremely correlated Fermi liquid theory* (ECFL) [9–11], Shastry proposed a formalism which successfully resolves both difficulties. This formalism is based on Schwinger's approach to field theory, which bypasses Wick's theorem, and is more generally applicable than the Feynman approach that is fundamentally based upon Wick's theorem. Building atop this powerful formalism, the ECFL theory consists of the following main ingredients:

- (1) The product ansatz, in which the physical Green's function  $\mathcal{G}[i, f]$  is written as a product of the auxiliary (Fermi-liquid type) Green's function  $\mathbf{g}[i, f]$ , and a caparison function  $\tilde{\mu}[i, f]$  (Eq. (9)). The former is a canonical, i.e. unprojected electron type Green's function, while the latter is a dynamical correction, which arises fundamentally from the removal of double occupancy from the Hilbert space. This addresses the difficulty (I) above.
- (2) The introduction of an expansion parameter  $\lambda \in (0, 1)$ , which continuously connects the  $t$ - $J$  model with the free Fermi gas, and enables the formulation of a systematic expansion. This parameter is related to the extent to which double occupancy is removed, and has a close parallel to the semiclassical expansion parameter  $\frac{1}{2S}$  arising in the expansion of spin  $S$  (angular momentum) operators in terms of canonical Bosons [11].

In addition the detailed calculations require certain crucial steps

- (3) The introduction of a particle–number sum rule for the auxiliary Green's function (Eq. (62)), fixing the number of auxiliary fermions to equal the number of physical fermions. This arises from requiring the charge of the particle to be unaffected by Gutzwiller projection, and is closely connected to the volume of the Fermi-surface of the physical fermions. In essence it ensures that the theory satisfies the Luttinger–Ward volume theorem [12,13].
- (4) The introduction of the second chemical potential  $u_0$ , which ensures that  $\mathbf{g}[i, f]$  and  $\tilde{\mu}[i, f]$  individually satisfy the shift invariance theorem [10], and together with the original chemical potential  $\mu$ , facilitates the fulfilling of the two particle–number sum rules.

In earlier work these ingredients are accomplished directly using the Schwinger equation of motion (EOM) for the  $t$ - $J$  model. In particular, the fundamental objects  $\mathbf{g}[i, f]$  and  $\tilde{\mu}[i, f]$  are defined through their respective equations of motion, and the expansion parameter  $\lambda$  is inserted directly into the equation of motion. The practical issue of computing objects to various orders in  $\lambda$  is also accomplished by iterating the EOM order by order. The technical details are given in Refs. [9,10], and are summarized in Section 2, facilitating a self contained presentation.

In recent papers, the  $O(\lambda^2)$  ECFL has been theoretically benchmarked using Dynamical Mean-Field Theory (DMFT) [14], Numerical Renormalization Group (NRG) calculations [15], and high-temperature series [16]. In all cases, the low order ECFL calculation compares remarkably well with these well established techniques. On the experimental side, a phenomenological version of ECFL which uses

simple Fermi-liquid expressions for the self-energies  $\Phi[i, f]$  and  $\Psi[i, f]$  (which are simply related  $\mathbf{g}[i, f]$  and  $\tilde{\mu}[i, f]$  respectively) was successful in explaining the anomalous lines shapes of Angle-Resolved Photoemission Spectroscopy (ARPES) experiments [17]. Encouraged by this, higher order terms e.g.  $O(\lambda^3)$  are of considerable interest in order to probe densities closer to the Mott limit than possible with the  $O(\lambda^2)$  theories, and in this context the present work is relevant. In this paper, we develop a diagrammatic  $\lambda$  expansion. This expansion allows one to calculate the Greens function and related objects to any order in  $\lambda$  by drawing diagrams. These diagrams are reminiscent of those in the Feynman series [18,19], although more complicated than the former. This extra complication stems from the non-canonical nature of the  $X$ -operators and the absence of Wick's theorem. The diagrammatic formulation of the  $\lambda$  series has the following advantages:

- It allows one to calculate the  $n$ th order contribution to any object by drawing diagrams directly for that order, without having to iterate the expressions from the previous orders. This not only allows for greater ease of computation of analytical expressions, but is also essential for powerful numerical series summation techniques, such as diagrammatic Monte Carlo [20]. Ultimately, it will allow the series to be evaluated to high orders in  $\lambda$ , whereas presently, only a second order calculation has been possible [21].
- It allows for the diagrammatic interpretation of the various objects in the theory such as the auxiliary Green's function  $\mathbf{g}[i, f]$  and the caparison factor  $\tilde{\mu}[i, f]$ . For example, one can see that the product ansatz (Eq. (9)) is a natural consequence of the structure of the  $\mathcal{G}[i, f]$  diagrams. In particular, it is necessitated by the extra complexity introduced into the diagrams (over those of the Feynman series) by the projection of the double occupancy.
- It allows one to visualize the structure of the diagrams to all orders in  $\lambda$ , therefore facilitating diagrammatic re-summations based on some physical principle.

## 1.2. Results

The main result of the paper is the formulation of diagrammatic rules to calculate the Green's function to any order in  $\lambda$ . More precisely, the rules state how to generate numerical representations (see Section 4.2), which are then converted into diagrams. A subset of these numerical representations (determined by a simple criterion) are in one-to-one correspondence with the standard Feynman diagrams. Therefore, the diagrams given here are a natural generalization of the Feynman diagrams. In this broader class of diagrams, we obtain a subset of numerical representations which are not in one-to-one correspondence with the resulting non-Feynman diagrams. In particular, two different numerical representations can lead to the same (non-Feynman) diagram. This occurs since in these non-Feynman diagrams, an interaction vertex can have more than two pairs of Green's function lines exiting and entering it (e.g. Fig. 40g). However, the contributions of both numerical representations must be kept. We also discuss the relationship between ECFL and a formalism using the high-temperature expansion for the  $t$ - $J$  model due to Zaitsev and Izyumov [7,8] in Section 8, and make some connections in the following.

We find that a certain subset of the  $\mathcal{G}[i, f]$  diagrams terminate with a self-energy insertion, rather than a single point, as in the case of the Feynman diagrams. This expresses the diagrammatic necessity for the factorization of  $\mathcal{G}$  into  $\mathbf{g}$  and  $\tilde{\mu}$ . These are in turn expressed in terms of the two self-energies  $\Phi$  and  $\Psi$ . It is interesting that within the Zaitsev-Izyumov [7,8] formalism, a two self-energy structure for the Green's function is necessary for the exact same reason. The fact that the two self-energy structure comes from three independent approaches, the  $\lambda$  expansion, the high-temperature expansion, and the factorization of the Schwinger EOM, shows that it is the correct representation of the Green's function for this model. In addition, as already reported in Ref. [11], the Dyson-Maleev approach developed by Harris, Kumar, Halperin and Hohenberg [22] also leads to a similar two self energy scheme in quantum spin systems, where again the algebra of the basic variables is non-canonical.

We derive diagrammatic rules for the constituent objects  $\mathbf{g}$ ,  $\tilde{\mu}$ ,  $\Phi$ , and  $\Psi$  from their definitions, starting from the Schwinger equations of motion. We avoid the use of dressed propagators (leading to skeleton terms), but rather expand various objects in powers of  $\lambda$  directly. The fact that these diagrammatic rules are consistent with those of  $\mathcal{G}$  and the product ansatz serves as an independent

proof of the rules given for  $\mathcal{G}$ . We find that  $\Phi$  consists of two independent pieces. The first can be obtained by adding a single interaction line to the terminal point of a  $\Psi$  diagram, while the second one is completely independent of  $\Psi$ . We denote the second piece by the letter  $\chi$ , which leads to the relation  $\Phi(\vec{k}, i\omega_k) = \epsilon_{\vec{k}} \Psi(\vec{k}, i\omega_k) + \chi(\vec{k}, i\omega_k)$  in momentum space. In a previous work by the same authors [23], we showed directly from the Schwinger equations of motion, that in the limit of infinite spatial dimensions,  $\Phi(\vec{k}, i\omega_k) = \epsilon_{\vec{k}} \Psi(i\omega_k) + \chi(i\omega_k)$ . Here, using the diagrammatic  $\lambda$  expansion, we show that this relationship continues to make sense in finite dimensions. In going from finite to infinite dimensions, we lose momentum dependence so that  $\Psi(\vec{k}, i\omega_k) \rightarrow \Psi(i\omega_k)$  and  $\chi(\vec{k}, i\omega_k) \rightarrow \chi(i\omega_k)$ . We also derive the Schwinger EOM defining the object  $\chi$  in finite dimensions.

We derive diagrammatic rules for the three point vertices  $\Lambda$  and  $\mathcal{U}$ , defined as functional derivatives of  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  (Eq. (11)). Diagrammatically, their relationship to  $\Phi$  and  $\Psi$  is seen to be consistent with the Schwinger equations of motion (Eq. (10)). We also derive a generalized Nozières relation for these vertices, which differs from the standard one for the three-point vertices of the Feynman diagrams. We introduce the concept of a skeleton diagram into our series. This enables us to make the rather subtle connection between our diagrammatic approach for the  $\lambda$  expansion, and the iterative one used previously. Finally, we use our diagrammatic approach to derive analytical expressions for the third order skeleton expansion of the objects  $\mathbf{g}$  and  $\tilde{\mu}$ , whereas previously only the second order expressions had been derived via iteration of the equations of motion.

### 1.3. Outline of the paper

In Section 2, we begin by reviewing the ECFL formalism from Refs. [9] and [10] in the simplified case of  $J = 0$ . In Section 3, we introduce the  $\lambda$  expansion diagrams in a heuristic way, drawing an analogy with the standard Feynman diagrams. In Section 4, we derive the rules for drawing and evaluating the bare diagrams for  $\mathcal{G}$  to each order in  $\lambda$ . We also draw and evaluate the first and second order bare diagrams for  $\mathcal{G}$ . In Section 5, we derive the diagrammatic rules for the constituent objects  $\mathbf{g}$ ,  $\tilde{\mu}$ ,  $\Phi$ ,  $\Psi$ ,  $\chi$ ,  $\gamma$ ,  $\Lambda$ , and  $\mathcal{U}$ . We also show how to evaluate diagrams in momentum space. We then introduce skeleton diagrams into the series, and complete the full circle by relating our diagrammatic approach to the  $\lambda$  expansion to the original iterative one reviewed in Section 2. In Section 6, we review the ECFL formalism [9,10] with  $J \neq 0$ , and introduce  $J$  into our diagrammatic series. In Section 7, we compute the skeleton expansion to third order in  $\lambda$  for the objects  $\mathbf{g}$  and  $\tilde{\mu}$ . We also discuss the high-frequency limit of  $\mathcal{G}$  to each order in the bare and skeleton expansions, as well as the “deviation” of the  $\lambda$  series from the Feynman series. Finally, in Section 8, we discuss the connection between the ECFL and the Zaitsev–Izyumov formalism for the high-temperature expansion of the  $t$ - $J$  model.

## 2. ECFL equations of motion and the $\lambda$ expansion

The Greens function is the fundamental object in this theory and is defined as usual by

$$\mathcal{G}_{\sigma_i, \sigma_f}[i, f] \equiv -\langle\langle X_i^{0\sigma_i} X_f^{\sigma_f 0} \rangle\rangle = -\frac{1}{Z[\mathcal{V}]} \text{tr} e^{-\beta H} T_\tau \left( e^{-\mathcal{A}} X_i^{0\sigma_i}(\tau_i) X_f^{\sigma_f 0}(\tau_f) \right), \quad (3)$$

where  $\mathcal{A} = \sum_j \int_0^\beta X_j^{\sigma\sigma'}(\tau') \mathcal{V}_j^{\sigma\sigma'}(\tau') d\tau'$ , is the additional term in the action due to the Bosonic source  $\mathcal{V}_i \equiv \mathcal{V}_i(\tau_i)$ , included in the partition functional  $Z[\mathcal{V}] = \text{tr} e^{-\beta H} T_\tau e^{-\mathcal{A}}$ . The angular brackets represent averages over the distribution in Eq. (3). The function  $\mathcal{G}$  satisfies the Schwinger equation of motion for the  $t$ - $J$  model as derived in Refs. [9], [10], and [2].

$$\begin{aligned} & \left\{ [(\partial_{\tau_i} - \mu)\delta[i, \mathbf{j}] - t[i, \mathbf{j}]] \delta_{\sigma_1, \sigma_j} + \mathcal{V}_i^{\sigma_1, \sigma_j} \delta[i, \mathbf{j}] \right\} \mathcal{G}_{\sigma_j, \sigma_2}[\mathbf{j}, f] \\ & = -\delta[i, f] \delta_{\sigma_1, \sigma_2} + \lambda \delta[i, f] \gamma_{\sigma_1, \sigma_2}[i] - \lambda t[i, \mathbf{j}] \gamma_{\sigma_1, \sigma_a}[i] \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] \\ & + \lambda t[i, \mathbf{j}] \sigma_1 \sigma_a \frac{\delta}{\delta \mathcal{V}_i^{\sigma_1, \sigma_a}} \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] + \frac{\lambda}{2} J[i, \mathbf{j}] \left( \gamma_{\sigma_1, \sigma_a}[\mathbf{j}] \mathcal{G}_{\sigma_a, \sigma_2}[i, f] - \sigma_1 \sigma_a \frac{\delta}{\delta \mathcal{V}_j^{\sigma_1, \sigma_a}} \mathcal{G}_{\sigma_a, \sigma_2}[i, f] \right), \quad (4) \end{aligned}$$

where the bold repeated indices are summed over. The functional derivative takes place at time  $(\tau_i^+)$ , and we have used the notation  $\delta[i, m] = \delta_{i,m}\delta(\tau_i - \tau_m)$ ,  $t[i, m] = t_{i,m}\delta(\tau_i - \tau_m)$ , and  $\gamma_{\sigma_1, \sigma_2}[i] = \sigma_1\sigma_2\mathcal{G}_{\bar{\sigma}_2, \bar{\sigma}_1}[i, i^+]$ .

We next outline how one obtains the above equation of motion from the definition Eq. (3). We take a time derivative  $\partial_{\tau_i}$  of Eq. (3), yielding several terms. We start with a simple contribution, namely the time derivative of the  $\theta(\tau_i - \tau_f)$ , which involves the anticommutator

$$\{X_i^{0\sigma_i}, X_j^{\sigma_j 0}\} = \delta_{ij} \left( \delta_{\sigma_i \sigma_j} - \lambda \sigma_i \sigma_j X_i^{\bar{\sigma}_i \bar{\sigma}_j} \right), \quad (5)$$

strictly speaking with  $\lambda = 1$ . We use the anticommutator, generalized as above by introducing the parameter  $\lambda \in [0, 1]$ , so that the result interpolates smoothly between the canonical value  $\lambda = 0$  and the fully Gutzwiller projected value  $\lambda = 1$ . This process is fundamental to obtaining the  $\lambda$  expansion. From this, we get  $\delta_{if}\delta(\tau_i - \tau_f) \left( \delta_{\sigma_i \sigma_f} - \lambda \sigma_i \sigma_f X_i^{\bar{\sigma}_i \bar{\sigma}_f}(\tau_i) \right)$ . This is expressed back in terms of the Greens function by writing  $(X_i^{\bar{\sigma}_i \bar{\sigma}_f}) \rightarrow \mathcal{G}_{\bar{\sigma}_f \bar{\sigma}_i}[i\tau_i, i\tau_i^+] = \sigma_i \sigma_f \gamma_{\sigma_i \sigma_f}[i]$ , and thus to the first two terms on the right hand side of Eq. (4).

Another contribution arises from the  $\tau_i$  dependence in the lower and upper limits of the time integrals in the expression

$$T_\tau e^{-\mathcal{A}_S} X_i^{0\sigma_i}(\tau_i) = e^{-\sum_j \int_{\tau_i}^{\beta} X_j^{\sigma \sigma'}(\tau') \mathcal{V}_j^{\sigma \sigma'}(\tau') d\tau'} X_i^{0\sigma_i}(\tau_i) e^{-\sum_j \int_0^{\tau_i} X_j^{\sigma \sigma'}(\tau') \mathcal{V}_j^{\sigma \sigma'}(\tau') d\tau'}, \quad (6)$$

involving the equal time commutator  $\sum_j \mathcal{V}_j^{\sigma \sigma'}(\tau_i) [X_j^{\sigma \sigma'}(\tau_i), X_i^{0\sigma_i}(\tau_i)] = \mathcal{V}_i^{\sigma_i \sigma_i'}(\tau_i) X_i^{0\sigma_i}(\tau_i)$ . This leads to the third term in the left hand side of Eq. (4).

The non trivial term is obtained when the  $\partial_{\tau_i} X_i^{0\sigma}(\tau_i)$  is evaluated from the Heisenberg equation of motion  $[H, X_i^{0\sigma}]$  and the fundamental anticommutator Eq. (5) yielding

$$[X_i^{0\sigma}, H] = -\mu X_i^{0\sigma} - t_{ij} X_j^{0\sigma_i} + \lambda \sum_{j\sigma_j} t_{ij} (\sigma_i \sigma_j) X_i^{\bar{\sigma}_i \bar{\sigma}_j} X_j^{0\sigma_j} - \frac{1}{2} \lambda \sum_{j \neq i} J_{ij} (\sigma_i \sigma_j) X_j^{\bar{\sigma}_i \bar{\sigma}_j} X_i^{0\sigma_j}. \quad (7)$$

Note that the  $J$  term has an almost identical structure to the  $t$  term, with  $i \leftrightarrow j$ . The term involving  $J$  actually does not come with the external  $\lambda$ , we introduce it so that the  $\lambda = 0$  limit is the Fermi gas. (This is permissible since we are finally interested in the limit  $\lambda = 1$ .) A higher order Greens function  $\langle\langle X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f) \rangle\rangle$  is generated by the third term and a similar one by the fourth term. These are re-expressed in terms of the Greens function by using the identity due to Schwinger

$$\begin{aligned} \langle\langle X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f) \rangle\rangle &= \langle\langle X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) \rangle\rangle \langle\langle X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f) \rangle\rangle \\ &- \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i)} \langle\langle X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f) \rangle\rangle. \end{aligned} \quad (8)$$

Using again  $\langle\langle X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) \rangle\rangle = \mathcal{G}_{\bar{\sigma}_j \bar{\sigma}_i}[i\tau_i, i\tau_i^+] = \sigma_i \sigma_j \gamma_{\sigma_i \sigma_j}[i]$ , we obtain the last four terms on the right hand side of equation Eq. (4). For ease of presentation we will initially set  $J \rightarrow 0$  and reinstate it at a later stage.

As demonstrated in Ref. [9], the electron Green's function  $\mathcal{G}[i, f]$  can be factored via the following product ansatz:

$$\mathcal{G}[i, f] = \mathbf{g}[i, \mathbf{j}] \cdot \tilde{\mu}[\mathbf{j}, f], \quad (9)$$

where  $\mathbf{g}[i, f]$  is the auxiliary Green's function,  $\tilde{\mu}[i, f]$  is the comparison factor, all objects have been represented as  $2 \times 2$  matrices in spin space, and matrix multiplication has been indicated by a dot.  $\mathbf{g}[i, f]$  and  $\tilde{\mu}[i, f]$  are defined by their respective Schwinger equations of motion.

$$\begin{aligned} \mathbf{g}^{-1}[i, m] &= (\mu - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] + t[i, m] (1 - \lambda \gamma[i]) - \lambda \Phi[i, m]. \\ \tilde{\mu}[i, m] &= (1 - \lambda \gamma[i]) \delta[i, m] + \lambda \Psi[i, m] \\ \Phi[i, m] &= -t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \Lambda_*[\mathbf{n}, m; i]; \quad \Psi[i, m] = -t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i]. \end{aligned} \quad (10)$$

These exact relations give the required objects  $\mathbf{g}$  and  $\tilde{\mu}$  in terms of the vertex functions. Here we also note that the local (in space and time) Green's function  $\gamma[i]$ , and the vertices  $\Lambda[n, m; i]$  and  $\mathcal{U}[n, m; i]$ , are defined as

$$\gamma[i] = \tilde{\mu}^{(k)}[\mathbf{n}, i^+].\mathbf{g}^{(k)}[i, \mathbf{n}]; \quad \Lambda[n, m; i] = -\frac{\delta}{\delta\mathcal{V}_i}\mathbf{g}^{-1}[n, m]; \quad \mathcal{U}[n, m; i] = \frac{\delta}{\delta\mathcal{V}_i}\tilde{\mu}[n, m], \quad (11)$$

where we have used the notation  $M_{\sigma_1, \sigma_2}^{(k)} = \sigma_1\sigma_2 M_{\bar{\sigma}_2, \bar{\sigma}_1}$  to denote the time reversed matrix  $M^{(k)}$  of an arbitrary matrix  $M$ . These exact relations give the vertex functions in terms of the objects  $\mathbf{g}$  and  $\tilde{\mu}$ . The vertices defined above ( $\Lambda$  and  $\mathcal{U}$ ) have four spin indices, those of the object being differentiated and those of the source. For example,  $\mathcal{U}_{\sigma_a\sigma_b}^{\sigma_1\sigma_2}[n, m; i] = \frac{\delta}{\delta\mathcal{V}_i^{\sigma_a\sigma_b}}\tilde{\mu}_{\sigma_1\sigma_2}[n, m]$ . In Eq. (10),  $\xi_{\sigma_a\sigma_b} = \sigma_a\sigma_b$ , and the  $*$  indicates that these spin indices should also be carried over (after being flipped) to the bottom indices of the vertex, which is also marked with a  $*$ . The top indices of the vertex are given by the usual matrix multiplication. An illustrative example is useful here:  $(\xi^*.\mathbf{g}[j, \mathbf{n}].\mathcal{U}_*[\mathbf{n}, m; i])_{\sigma_1\sigma_2} = \sigma_1\sigma_a \mathbf{g}_{\sigma_a, \sigma_b}[j, \mathbf{n}]\frac{\delta}{\delta\mathcal{V}_i^{\sigma_1\sigma_a}}\tilde{\mu}_{\sigma_b, \sigma_2}[\mathbf{n}, m]$ .

The  $\lambda$  expansion is obtained by expanding Eqs. (10) and (11) iteratively in the continuity parameter  $\lambda$ . The  $\lambda = 0$  limit of these equations is the free Fermi gas. Therefore, a direct expansion in  $\lambda$  will lead to a series in  $\lambda$  in which each term is made up of the hopping  $t_{ij}$  and the free Fermi gas Green's function  $\mathbf{g}^{(0)}[i, f]$ . As is the case in the Feynman series, this can be reorganized into a skeleton expansion in which only the skeleton graphs are kept and  $\mathbf{g}^{(0)}[i, f] \rightarrow \mathbf{g}[i, f]$ . However, one can also obtain the skeleton expansion directly by expanding Eqs. (10) and (11) in  $\lambda$ , but treating  $\mathbf{g}[i, f]$  as a zeroth order (i.e. unexpanded) object in the expansion. This expansion is carried out to second order in Ref. [9]. In doing this expansion, one must evaluate the functional derivative  $\frac{\delta\mathbf{g}}{\delta\mathcal{V}}$ . This is done with the help of the following useful formula which stems from the product rule for functional derivatives.

$$\frac{\delta\mathbf{g}[i, m]}{\delta\mathcal{V}_r} = \mathbf{g}[i, \mathbf{x}].\Lambda[\mathbf{x}, \mathbf{y}, r].\mathbf{g}[\mathbf{y}, m]. \quad (12)$$

Within the  $\lambda$  expansion, the LHS is evaluated to a certain order in  $\lambda$  by taking the vertex  $\Lambda$  on the RHS to be of that order in  $\lambda$ .

### 3. Heuristic discussion of $\lambda$ expansion diagrams

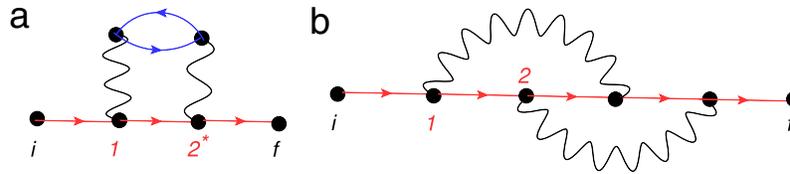
#### 3.1. Numerical representations of Feynman diagrams

Before deriving the precise rules for the  $\lambda$  expansion diagrams, it is useful to have a heuristic discussion in which we compare them to the more familiar Feynman diagrams [18,19]. To this end, we introduce numerical representations for the standard Feynman diagrams. These numerical representations will then be generalized to generate the  $\lambda$  expansion diagrams.

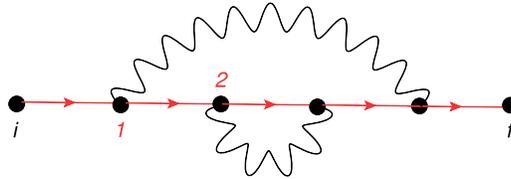
Consider any Feynman diagram for the Green's function  $\mathcal{G}[i, f]$  such as those displayed in Fig. 1. There is a unique path which runs between  $i$  and  $f$  which uses only Green's function lines, not counting the interaction lines. We denote this as the zeroth Fermi loop. It is drawn in red in Fig. 1. We number the interaction lines which connect to the zeroth Fermi loop in the order in which they appear in this loop. This list of numbers (along with  $f$ ) is placed in the top row of our numerical representation. In the case of both Figs. 1(a) and 1(b), it is

$$1 \ 2 \ f.$$

If the zeroth Fermi loop does not exhaust all of the Green's function lines in the diagram, such as in Fig. 1(a), we proceed to the first Fermi loop. To identify the first Fermi loop, we find the interaction vertex with the highest number which connects to the zeroth Fermi loop with only one of its two sides. In this case, this is the interaction vertex labeled 2. The other side has one incoming line and one outgoing line. There is a unique path in the diagram which connects these two lines using only Green's function lines, not interaction lines. This defines the first Fermi loop. It is drawn in blue in Fig. 1(a).



**Fig. 1.** Second order Feynman diagrams for  $\mathcal{G}[i, f]$ . The zeroth Fermi loop, which is the chain running from  $i$  to  $f$  is colored in red. In panel (a), the first Fermi loop is colored in blue. The numerical representation of the diagram in panel (a) is  $1\ 2^* f$ ;  $2^* : 0\ 1 f$ , while that of the diagram in panel (b) is  $1\ 2\ 1_2 f$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 2.** This Feynman diagram results from reversing the order of the subscripts in the numerical representation of the Feynman diagram in Fig. 1(b). Therefore, the numerical representation of this diagram is  $1\ 2\ 2_1 f$ .

Since the interaction vertex 2 spawned the first Fermi loop, it is starred in the top row of the representation. We also include a lower row for the first Fermi loop. Therefore, the numerical representation of Fig. 1(a) now reads.

$$\begin{array}{l} 1\ 2^* f \\ 2^* : 0\ 1 f. \end{array}$$

The second row, which represents the first Fermi loop, is labeled by  $2^*$ , since it was spawned by the second interaction vertex in the zeroth Fermi loop. The fact that only 0 and  $f$  are present in the second row tells us that there were no interaction vertices **introduced** in the first Fermi loop. That is to say there are no interaction vertices which connect to the first Fermi loop, but not to the previous ones (in this case the zeroth Fermi loop). Finally, after all of the Fermi loops have been recorded, all nonzero integers which are not starred indicate the position of one side of an interaction vertex in a Fermi loop. We record the position of the other side as a subscript. Therefore, the complete numerical representation of Fig. 1(a) is

$$\begin{array}{l} 1\ 2^* f \\ 2^* : 0\ 1 f. \end{array}$$

We can represent this in short as  $1\ 2^* f$ ;  $2^* : 0\ 1 f$ , where the semi colon indicates the next line. The complete numerical representation of Fig. 1(b) is

$$1\ 2\ 1_2 f.$$

Note that the order of appearance of the 1 and 2 as subscripts is important. Reversing them would yield the diagram in Fig. 2, which has the following numerical representation.

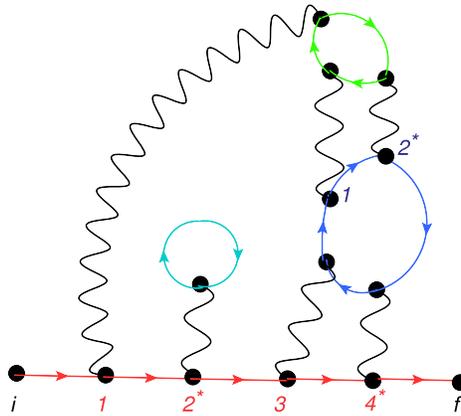
$$1\ 2\ 2_1 f.$$

We now consider the slightly more complicated diagram in Fig. 3 to illustrate the scope of this approach. We will now show how the numerical representation of this diagram is derived. We first identify the zeroth Fermi loop, which is drawn in red in Fig. 3. The top row now reads

$$1\ 2\ 3\ 4 f.$$

In this case, the vertex with the highest number which connects to this loop with only one side is 4. Hence, 4 spawns another Fermi loop, and gets a star in the top row.

$$1\ 2\ 3\ 4^* f.$$



**Fig. 3.** Sixth order Feynman diagram. The zeroth, first, second, and third Fermi loops are drawn in red, blue, green, and turquoise respectively. Interaction vertices introduced in a particular Fermi loop are numbered in the same color as that loop. An interaction vertex is starred if it spawns a new Fermi loop. The numerical representation for this diagram is  $1\ 2^* \ 3\ 4^* \ f$ ;  $4^* : 0\ 3\ 1\ 2^* \ f$ ;  $(4, 2)^* : 0\ (4,1)\ 1\ f$ ;  $2^* : 0\ f$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We identify this as the first Fermi loop. It is drawn in blue in Fig. 3. The numerical representation is modified to read

$$1\ 2\ 3\ 4^* \ f$$

$$4^* : 0\ 1\ 2\ f.$$

Considering only the interaction vertices introduced in the first Fermi loop, we now search for the one with the highest number which connects to the first Fermi loop with one side, but whose other side is **free**, that is to say that it does not connect to any of the Fermi loops introduced thus far (zeroth and first). This is the interaction vertex 2. Hence, it gets a star, and spawns the second Fermi loop, which is drawn in green. The numerical representation now reads

$$1\ 2\ 3\ 4^* \ f$$

$$4^* : 0\ 1\ 2^* \ f$$

$$(4, 2)^* : 0\ f.$$

Here, the ordered pair  $(4, 2)$  is used to distinguish the 2 in the first Fermi loop from the 2 in the zeroth Fermi loop, the latter being denoted simply as 2. The first number in the pair is 4 since the fourth interaction vertex in the zeroth Fermi loop spawned the first Fermi loop. Also note that no interaction vertices are introduced in the second Fermi loop, hence its row only has a 0 and an  $f$ . Therefore, we have arrived at the end of our first sequence of nested Fermi loops. We now take a step back in this sequence and return to the first Fermi loop. Considering only the interaction vertices introduced in the first loop with number less than 2, we search for the one with the highest number which connects with one side to the first Fermi loop, but whose other side is free (i.e. does not connect to the zeroth, first, or second Fermi loops). There is no such interaction vertex. Therefore, we take another step back in the sequence, and return to the zeroth Fermi loop. We find that the interaction vertex 2 connects to this loop with one side, but that the other side is free. Hence, 2 gets a star and spawns the fourth Fermi loop, which is drawn in turquoise. The numerical representation now reads

$$1\ 2^* \ 3\ 4^* \ f$$

$$4^* : 0\ 1\ 2^* \ f$$

$$(4, 2)^* : 0\ f$$

$$2^* : 0\ f.$$

Since there are no interaction vertices introduced in the fourth Fermi loop, we have arrived at the end of our second sequence of nested Fermi loops. We take a step back to the zeroth Fermi loop and find that there are no more interaction vertices introduced in this loop which have one side free. Since all of

the Fermi loops have been identified, as the final step, we must take the integers which are not starred, and place them in their final locations as subscripts. The complete numerical representation now reads

$$\begin{array}{l} 1 2^* 3 4^* f \\ 4^* : 0_3 1 2^* f \\ (4, 2)^* : 0_{(4,1)} 1 f \\ 2^* : 0 f. \end{array}$$

If we now wanted to formulate a set of rules for generating the numerical representations obtained from the Feynman diagrams, they would be the following.

- (1) Write a row of integers  $1 \dots m f$  where  $m \geq 1$ , e.g.

$$1 2 3 4 f.$$

- (2) Assign a star to any of the integers in the row ( $f$  does not count as an integer), e.g.

$$1 2^* 3 4^* f.$$

- (3) Every starred integer gives rise to a lower row. The  $i$ th lower row also consists of integers  $0 \dots m_i f$ , where  $m_i \geq 0$ , e.g.

$$\begin{array}{l} 1 2^* 3 4^* f \\ 4^* : 0 1 2 f \\ 2^* : 0 f. \end{array}$$

- (4) In the lower rows, assign a star to any of the integers excluding 0, e.g.

$$\begin{array}{l} 1 2^* 3 4^* f \\ 4^* : 0 1 2^* f \\ 2^* : 0 f. \end{array}$$

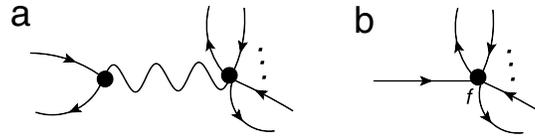
- (5) The integers starred in step 4 once again give rise to lower rows, etc. Continue this process until the last rows which you create have no starred integers, e.g.

$$\begin{array}{l} 1 2^* 3 4^* f \\ 4^* : 0 1 2^* f \\ (4, 2)^* : 0 f \\ 2^* : 0 f. \end{array}$$

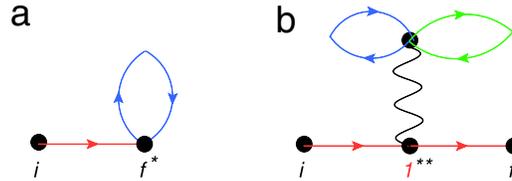
- (6) Label each integer with a tuple (an ordered list of numbers) which traces that integer back to the first row through the starred integers. For example, the 0 in the third row would be labeled  $(4, 2, 0)$ .
- (7) Between any 2 consecutive integers of a row (including 0's and  $f$ 's), one can place as subscripts an ordered list of tuples from the following set: all those corresponding to non-starred integers except 0 whose tuple can be obtained from the tuple of the smaller of the 2 consecutive integers in question, by taking the first  $k \leq l$  entries of this tuple (where  $l$  is the length of the tuple), and subtracting a non-negative integer from the last entry. For example, suppose that the two consecutive integers in question are the 2 and  $f$  of the second row. Then all tuples (corresponding to non-starred integers) eligible to be used as subscripts between them are:  $(4, 1)$ , 3, and 1. All non-starred integers (except 0's) must be used exactly once in this way, e.g.

$$\begin{array}{l} 1 2^* 3 4^* f \\ 4^* : 0_3 1 2^* f \\ (4, 2)^* : 0_{(4,1)} 1 f \\ 2^* : 0 f. \end{array}$$

If we think back to the order in which we generated Fermi loops (and hence the numerical representation) from a given Feynman diagram, we can see that it complies exactly with rule (7) stated above. Doing things in this way ensures that the mapping between Feynman diagrams and numerical representations is one-to-one.



**Fig. 4.** The 2 elements used for constructing the  $\lambda$  expansion diagrams.



**Fig. 5.** Two of the simplest non-Feynman diagrams in the  $\lambda$  expansion. A non-Feynman diagram occurs when the terminal point of a Fermi loop spawns another Fermi loop. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.2. Topologies of $\lambda$ expansion diagrams

The exact rules for drawing diagrams for the  $\lambda$  expansion, as defined in Section 2 will be derived in Section 4. There, it will be shown that the  $\lambda$  expansion diagrams are constructed from the 2 elements displayed in Fig. 4. The one in panel (a) is a generalization of the Feynman interaction vertex, in which one of the sides can have any number of pairs of incoming and outgoing lines rather than just one pair. The one in panel (b) is a generalization of the terminal point  $f$  in a Feynman diagram. In the case of the Feynman diagrams, it is a single point, while in the case of a  $\lambda$  expansion diagram, it is a single point along with any number of pairs of incoming and outgoing lines. These extra lines come from the second term on the RHS of Eq. (17). This term, which itself comes from the anti-commutator of the  $X$ -operators in Eq. (2), and which is absent in the EOM of canonical theories, allows a diagram to close in on itself in an iterative expansion of the EOM.

In Fig. 5, we have drawn two of the simplest non-Feynman diagrams which can be made from these elements. The one in panel (a) has the following numerical representation.

$$f^* : 0f.$$

The zeroth Fermi loop runs from the site  $i$  to the site  $f$  and is drawn in red. The site  $f$ , which is the terminal point of the zeroth Fermi loop spawns the first Fermi loop, drawn in blue. The one in panel (b) has the following numerical representation.

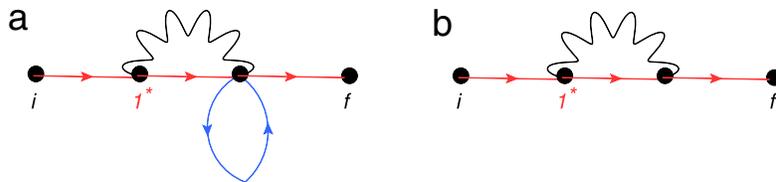
$$\begin{aligned} 1^{**}f \\ 1^{**} : 0f^* \\ (1, f)^* : 0f. \end{aligned}$$

Here, the interaction vertex 1, introduced in the zeroth Fermi loop (drawn in red) spawns the first Fermi loop (drawn in blue). Additionally, the terminal point of the first Fermi loop spawns the second Fermi loop (drawn in green).

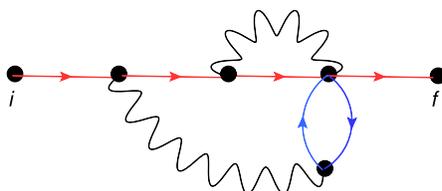
The diagrams drawn in Fig. 5 are both valid  $\lambda$  expansion diagrams. However, as shown below, the allowed topologies of  $\lambda$  expansion diagrams do not include all of the possible ways of combining the two elements in Fig. 4, but rather only a subset of these. To see which subset, consider the plausible diagram displayed in Fig. 6(a), which is not an allowed  $\lambda$  expansion diagram. This diagram is obtained from the Fock diagram in Fig. 6(b) by adding a Fermi loop to the latter. The numerical representation for the diagram in Fig. 6(b) is

$$1_1f.$$

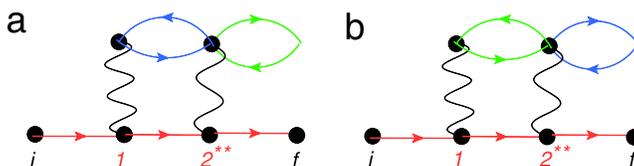
We see that the point from which the first Fermi loop emanates in Fig. 6(a) is represented by a subscript in Fig. 6(b). Alternatively, using the terminology introduced in Section 3.1, the first Fermi loop



**Fig. 6.** The diagram in panel (a) is not allowed in the  $\lambda$  expansion. This is because first Fermi loop emanates from a point which is represented by a subscript in the Fock diagram displayed in panel (b).



**Fig. 7.** A more elaborate version of the diagram in Fig. 6(a), which is also not allowed in the  $\lambda$  expansion.



**Fig. 8.** A demonstration that unlike Feynman diagrams,  $\lambda$  expansion diagrams are not in one-to-one correspondence with their numerical representations. In the diagram in panel (a), the interaction vertex 1 of the zeroth Fermi loop also connects to the first Fermi loop. In the diagram in panel (b), it connects to the second Fermi loop. The topologies of both diagrams, however, are identical. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is spawned by the interaction vertex 1 of the zeroth Fermi loop. However, the other side of this interaction vertex is not free, but rather connects to the zeroth Fermi loop itself. This is not allowed. In fact, we shall find that when there are more than one pair of lines connected to a single point of an interaction vertex, each pair must both start and terminate a Fermi loop at that point, as in Fig. 5(b). Another diagram which is not an allowed  $\lambda$  expansion diagram is drawn in Fig. 7.

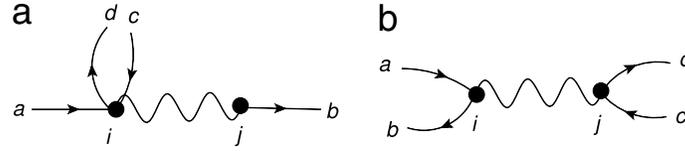
Another feature of the  $\lambda$  expansion diagrams is that they are not in one-to-one correspondence with their numerical representations. To see this, consider the diagrams drawn in Fig. 8. As usual, the zeroth, first, and second Fermi loops are drawn in red, blue, and green respectively. The diagram in (Fig. 8(a)) has the numerical representation

$$\begin{aligned} & 1 2^{**} f \\ 2^{**} : & 0_1 f^* \\ (2, f)^* : & 0 f. \end{aligned}$$

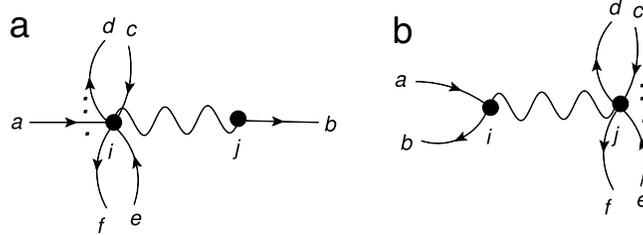
In words, this says that the interaction vertex 2 of the zeroth Fermi loop spawns the first Fermi loop. The interaction vertex 1 of the zeroth Fermi loop connects to the first Fermi loop. Finally, the terminal point of the first Fermi loop spawns the second Fermi loop. On the other hand, the diagram in (Fig. 8(b)) has the numerical representation

$$\begin{aligned} & 1 2^{**} f \\ 2^{**} : & 0 f^* \\ (2, f)^* : & 0_1 f. \end{aligned}$$

In this case, the interaction vertex 1 of the zeroth Fermi loop connects to the second Fermi loop rather than the first. We see that both of the above numerical representations lead to the same diagram, although they both have a contribution which must be accounted for.



**Fig. 9.**  $t$ -vertices in panel (a) versus  $J$ -vertices in panel (b) for the  $\lambda$  expansion diagrams which are also Feynman diagrams. The corresponding lines are marked with corresponding letters.



**Fig. 10.**  $t$ -vertices in panel (a) versus  $J$ -vertices in panel (b) for the more general  $\lambda$  expansion diagrams. The corresponding lines are marked with corresponding letters.

A final point to mention in this discussion of the  $\lambda$  expansion diagrams is that when drawing the diagrams in real space, the vertex appropriate for the  $t$ -interaction differs from the one appropriate for the  $J$ -interaction. While this is derived rigorously from the EOM below, one can understand it by examining the relevant terms in the Hamiltonian (Eq. (1)). First, we examine the  $t$ -term. Writing the  $X$  operators in terms of canonical creation and destruction operators, we obtain

$$-\sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} = -\sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger (1 - n_{i\bar{\sigma}}) c_{j\sigma} = -\sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger n_{i\bar{\sigma}} c_{j\sigma}. \quad (13)$$

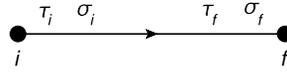
Here, we have used the non-Hermitean mapping described in Ref. [11]

$$X_i^{\sigma 0} \rightarrow c_{i\sigma}^\dagger (1 - n_{i\bar{\sigma}}); \quad X_j^{0\sigma} \rightarrow c_{j\sigma}. \quad (14)$$

As discussed in Ref. [11], it is permissible to drop the projection from the destruction operator, since if the system starts in the subspace of no double occupancy, the unprojected destruction operator cannot take it out of this subspace. The second term on the RHS of Eq. (13) can be represented with the interaction vertex drawn in Fig. 9(a). Next, we examine the  $J$  term. Since a spin flip operator or number operator cannot take the system out of the subspace of no double occupancy, the  $X$  operators in the  $J$  term can be replaced by their canonical counterparts. Therefore, written in terms of canonical operators, the  $J$  term looks like

$$\frac{1}{2} \sum_{ij\sigma} J_{ij} n_{i\sigma} + \frac{1}{4} \sum_{ij\sigma_1\sigma_2} \{ U_{ij} c_{i\sigma_1}^\dagger c_{i\sigma_2} c_{j\sigma_2}^\dagger c_{j\sigma_1} - n_{i\sigma_1} n_{j\sigma_2} \}. \quad (15)$$

The first term amounts to a shift in the chemical potential  $\mu$ , while the second one leads to the interaction vertex drawn in Fig. 9(b). The corresponding lines between Figs. 9(a) and 9(b) have been marked with corresponding letters. Throughout the text, we shall sometimes use the term “Feynman diagrams” to refer to the  $\lambda$  expansion diagrams formed solely from the interaction vertices in Fig. 9, and sometimes to refer to the usual Feynman diagrams [19,18]. It should be clear what we mean from the context. To obtain the more general  $\lambda$  expansion diagrams, one must use the vertices drawn in Fig. 10. Once again, the corresponding lines have been marked with corresponding letters between the  $t$ -vertices in panel (a) and the  $J$ -vertices in panel (b). A  $\lambda$  expansion diagram drawn in real space will of course have a mix of  $t$ -vertices and  $J$ -vertices. Luckily, when drawing the diagrams in momentum space, we can use only one type of vertex ( $t$  or  $J$ ). The details of this procedure are discussed in Section 6. To convert between “ $t$ -diagrams” and “ $J$ -diagrams”, we must rearrange every interaction vertex as indicated in Fig. 10. For example, the Hartree and Fock diagrams, when drawn



**Fig. 11.** The zeroth order contribution to the Green's function:  $\mathbf{g}^{(0)}[i, f]$ .

using  $t$ -vertices, appear as in Fig. 12(b) and Fig. 12(c) respectively. In this introductory section, we have used the more familiar  $J$ -vertices to construct our diagrams, while in the rest of the paper, we shall take the point of view of using the  $t$ -vertices. The counterparts of the diagrams drawn in Figs. 1(a), 1(b), 2, 5(b), and 8, are drawn in Figs. 17(g), 17(b), 17(c), 17(n), and 40(g) respectively.

To conclude this preliminary discussion, we point out that while it may be possible to define the  $\lambda$  diagrams as all inequivalent ways of combining the elements displayed in Fig. 4 with some topological constraints, this definition would not have much practical value. It also would not tell us how to evaluate the diagram once we had drawn it. On the other hand, the numerical representations of the  $\lambda$  expansion diagrams defined below are both easy to generate in a systematic manner, and easy to evaluate. In fact, one may argue that even for the standard Feynman diagrams, the definition in terms of the numerical representations presented in Section 3.1 is more useful than the usual one, since it gives a systematic way of generating, and a compact way of representing the diagrams.

#### 4. Bare diagrammatic $\lambda$ expansion for $\mathcal{G}[i, f]$ .

##### 4.1. Integral equation of motion and the first order $\lambda$ expansion.

As can be seen from Eq. (4), the parameter  $\lambda$  adiabatically connects the free Fermi gas at  $\lambda = 0$  with the fully projected model at  $\lambda = 1$ . Therefore, in the bare  $\lambda$  series for  $\mathcal{G}$ , to each order in  $\lambda$ ,  $\mathcal{G}[i, f]$  is expressed as a functional of the free Fermi gas,  $\mathbf{g}^{(0)}[i, f]$  and the hopping  $t_{ij}$ . In this section, we aim to derive a set of rules for drawing diagrams to compute the  $n$ th order contribution to the bare series for  $\mathcal{G}[i, f]$ . We do this by rewriting Eq. (4) as an integral equation, and then iterating this equation in  $\lambda$ . An analogous expansion is done for the first couple of orders of the Feynman series in Kadanoff and Baym in Ref. [24]. To this end, we rewrite Eq. (4) as

$$-\mathbf{g}_{\sigma_1, \sigma_j}^{-1(0)}[i, \mathbf{j}] \mathcal{G}_{\sigma_j, \sigma_2}[\mathbf{j}, f] = -\delta[i, f] \delta_{\sigma_1, \sigma_2} + \lambda \times \delta[i, f] \gamma_{\sigma_1, \sigma_2}[i] - \lambda \times t[i, \mathbf{j}] \gamma_{\sigma_1, \sigma_a}[i] \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] + \lambda \times t[i, \mathbf{j}] \sigma_1 \sigma_a \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_1, \bar{\sigma}_a}} \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f], \quad (16)$$

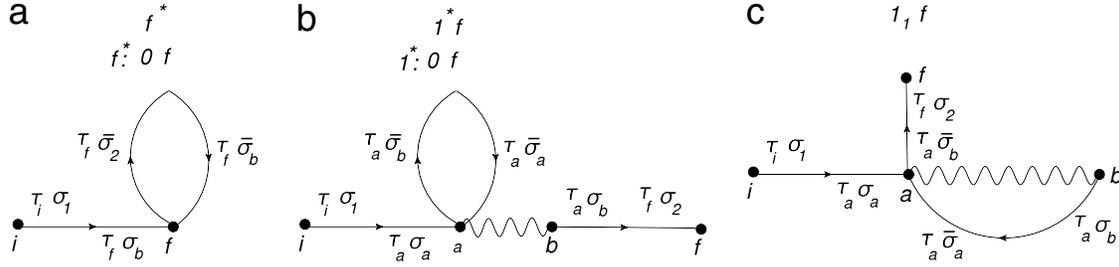
where  $\mathbf{g}^{-1(0)}[i, f]$ , the inverse of the free Fermi gas Green's function is obtained by setting  $\lambda = 0$  in Eq. (10). Rewriting Eq. (16), we obtain the following integral equation for  $\mathcal{G}[i, f]$ .

$$\mathcal{G}_{\sigma_1, \sigma_2}[i, f] = \mathbf{g}_{\sigma_1, \sigma_2}^{(0)}[i, f] - \lambda \mathbf{g}_{\sigma_1, \sigma_b}^{(0)}[i, f] \sigma_b \sigma_2 \mathcal{G}_{\bar{\sigma}_2, \bar{\sigma}_b}[f, f^+] - \lambda \times \mathbf{g}_{\sigma_1, \sigma_b}^{(0)}[i, \mathbf{k}] \left( -t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \mathcal{G}_{\bar{\sigma}_a, \bar{\sigma}_b}[\mathbf{k}, \mathbf{k}^+] \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] + t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}}^{\bar{\sigma}_b, \bar{\sigma}_a}} \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] \right), \quad (17)$$

This expression has considerable parallels to a similar expression for the (canonical) Hubbard model, with one exception, the second term on the RHS, (arising from the non-canonical nature of the  $X$ 's) has no counterpart in the canonical theory. If we drop this term, the series so generated is exactly the Feynman series.

We now proceed to draw the diagrams for the zeroth and first order contributions to  $\mathcal{G}$ . The zeroth order contribution to the Green's function, which is given by the free Fermi gas  $\mathbf{g}^{(0)}[i, f]$ , is represented by the diagram in Fig. 11.

To obtain the first order contribution to  $\mathcal{G}[i, f]$ , we plug  $\mathbf{g}^{(0)}[i, f]$  in for  $\mathcal{G}[i, f]$  in the RHS of Eq. (17). This leads to the three diagrams displayed in Fig. 12. The diagrams (a), (b), and (c) in Fig. 12 correspond to the three terms in the parenthesis on the RHS of Eq. (17) respectively. They correspond to the analytical expressions (a):  $-\lambda \sigma_b \sigma_2 \mathbf{g}_{\sigma_1 \sigma_b}^{(0)} [i, \tau_i, \tau_f] \mathbf{g}_{\bar{\sigma}_2 \bar{\sigma}_b}^{(0)} [\tau_f, \tau_f^+]$ ;



**Fig. 12.** The first order contribution to the Green's function:  $\mathcal{G}^{(1)}[i, f]$ . The diagrams in panels (a), (b), and (c) come from the first, second, and third terms on the RHS of Eq. (17), respectively.

(b):  $\lambda \sigma_a \sigma_b \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f]$ ; and (c):  $-\lambda \sigma_a \sigma_b \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{ba}^{(0)}[\tau_a, \tau_a^+] \mathbf{g}_{af}^{(0)}[\tau_a, \tau_f]$ . In drawing the diagram in Fig. 12(c), we have used the Schwinger identity

$$\frac{\delta \mathbf{g}_{\sigma_a, \sigma_b}^{(0)}[i, f]}{\delta \mathcal{V}_r^{\sigma_c \sigma_d}} = -\mathbf{g}_{\sigma_a, \sigma_x}^{(0)}[i, \mathbf{x}] \frac{\delta \mathbf{g}_{\sigma_x, \sigma_y}^{-1(0)}[\mathbf{x}, \mathbf{y}]}{\delta \mathcal{V}_r^{\sigma_c \sigma_d}} \mathbf{g}_{\sigma_y, \sigma_b}^{(0)}[\mathbf{y}, f] = \mathbf{g}_{\sigma_a, \sigma_c}^{(0)}[i, r] \mathbf{g}_{\sigma_d, \sigma_b}^{(0)}[r, f]. \quad (18)$$

In other words, the role of the functional derivative in the Eq. (17) is to pick a line in the diagram for  $\mathcal{G}_{\sigma_a \sigma_2}[\mathbf{j}, f]$ , and to split it into two lines, one entering the point  $\mathbf{k}$ , and the other one exiting it.

The reader would recognize that we bypassed the Wicks theorem, by utilizing instead the Schwinger identity Eq. (18).

#### 4.2. Rules for calculating the $n$ th order contribution

By plugging in the first and zeroth order diagrams into the RHS of Eq. (17), we can obtain the second order diagrams. Using this iterative process, we can obtain diagrams for  $\mathcal{G}$  to any order in  $\lambda$ . Moreover, by noticing the pattern in the iterative process, we can derive the rules for obtaining the  $n$ th order contribution to  $\mathcal{G}$  directly without calculating the lower order contributions. In the case of the Feynman diagrams, this is merely an alternate way of deriving the rules obtained from using Wick's theorem. However, in the present case, in which the standard Wick's theorem is not available, this derivation is essential in going from the EOM definition of the  $\lambda$  expansion introduced in Ref. [9] and the equivalent diagrammatic one developed here. We now present the diagrammatic rules for calculating the  $n$ th order contribution to  $\mathcal{G}$ .

- (1) Write a row of consecutive integers followed by the letter  $f$ , i.e.  $1 \dots m f$ , where  $m \geq 0$  (if  $m = 0$ , we simply write  $f$ ), e.g.

$$1 \ 2 \ 3 \ f.$$

- (2) Give any number of stars (including no stars) to each these integers (including  $f$ ), e.g.

$$1^{**} \ 2 \ 3 \ f^{*}.$$

- (3) Each integer (including  $f$ ) with  $p$  stars ( $p \geq 1$ ) gives rise to another row of integers which now starts with 0 (as opposed to 1), and which ends with an  $f$  with  $p - 1$  stars. Each new row can have any number of integers between the 0 and the  $f$ , each of which can have any number of stars, giving rise to further rows. 0 is not allowed to have any stars, e.g.

$$\begin{aligned} & 1^{**} \ 2 \ 3 \ f^{*} \\ 1^{**} : & 0 \ 1 \ 2^{***} \ f^{*} \\ (1, 2)^{***} : & 0 \ 1 \ f^{**} \\ (1, 2, f)^{**} : & 0 \ f^{*} \\ (1, 2, f, f)^{*} : & 0 \ 1 \ f \\ (1, f)^{*} : & 0 \ 1 \ f \\ f^{*} : & 0 \ 1 \ 2 \ f. \end{aligned}$$

Note that each integer in the above diagram is uniquely specified by a tuple which traces it back to the first row through the starred integers. For example, the number 1 in the fifth row corresponds to the tuple  $(1, 2, f, f, 1)$ .

- (4) Let  $\nu$  be the total number of integers without stars excluding 0's and  $f$ 's. Let  $s_f$  be the total number of stars on the  $f$  in the top row, and let  $s$  be the total number of stars excluding those on  $f$ 's. Then the order  $n$  must satisfy the relation  $n = \nu + s_f + s$ . In the above example,  $\nu = 8$ ,  $s_f = 1$ , and  $s = 5$ . Therefore this a 14th order diagram.
- (5) Between any 2 consecutive integers of a row (including 0's and  $f$ 's), one can place as subscripts an ordered list of tuples from the following set: All those corresponding to non-starred integers (except 0's and  $f$ 's) whose tuple can be obtained from the tuple of the smaller of the 2 consecutive integers in question, by taking the first  $k \leq l$  entries of this tuple (where  $l$  is the length of the tuple), and subtracting a non-negative integer from the last entry. We have taken  $f$ 's to be integers greater than all other integers in their respective rows. For example, suppose that the two consecutive integers in question are 1 and  $f$  in the fifth row of the above diagram. Then all integers eligible to be used as subscripts between them are:  $(1, 2, f, f, 1)$ ,  $(1, 2, 1)$ , and  $(1, 1)$ . All non-starred integers (except 0's and  $f$ 's) must be used exactly once in this way. e.g.

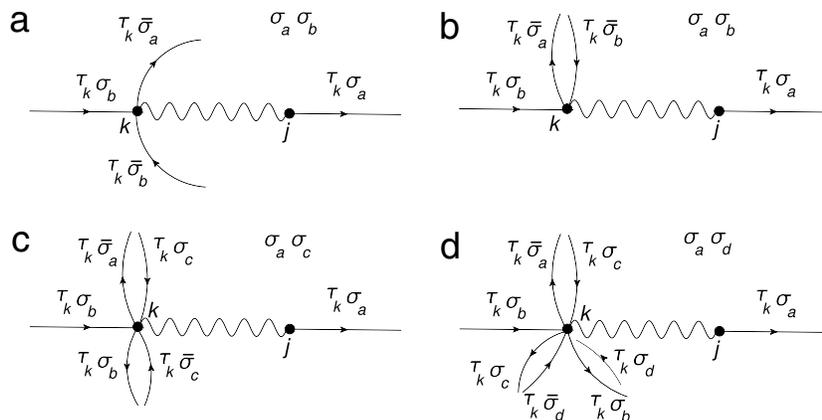
$$\begin{array}{l}
 1^{**} 2 3 {}_3 2 f^* \\
 1^{**} : 0 1 2^{***} f^* \\
 (1, 2)^{***} : 0 1 f^{**} \\
 (1, 2, f)^{**} : 0 {}_{(1,2,1)} f^* \\
 (1, 2, f, f)^* : 0 1 {}_{(1,2,f,f,1)} f \\
 (1, f)^* : 0 1 {}_{(1,f,1)} {}_{(1,1)} f \\
 f^* : 0 1 2 {}_{(f,1)} {}_{(f,2)} f.
 \end{array}$$

- (6) We use the numerical representation to draw the diagram in the following way. Each integer excluding 0's and  $f$ 's corresponds to an interaction vertex shown in Fig. 13. The interaction vertices displayed in panels (a), (b), (c), and (d) correspond to 0, 1, 2, and 3 stars respectively on the integer in question. On the top right of each panel, we indicate how the spins contribute to the sign of the diagram. Note that when two outgoing or two incoming lines share the same spin label, this spin contributes to the sign of the diagram, while when an outgoing and an incoming line share the same spin label, this spin does not contribute to the sign of the diagram. For example, in panel (d)  $\sigma_a$  and  $\sigma_d$  contribute to the sign while  $\sigma_b$  and  $\sigma_c$  do not.

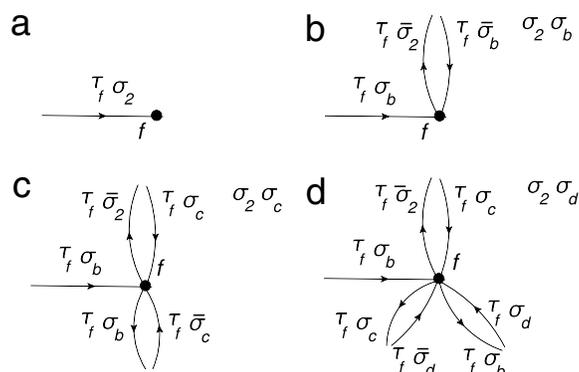
The  $f$  in the top row corresponds to a terminal point shown in Fig. 14. The terminal points displayed in panels (a), (b), (c), and (d) correspond to 0, 1, 2, and 3 stars respectively on the  $f$  in the top row. On the top right of each panel, we indicate how the spins contribute to the sign of the diagram. Note that the same general rule holds as in the case of the interaction vertices, except that now for the case of 1 or more stars, the spin  $\sigma_2$  also contributes to the sign of the diagram. For the case of one or more stars, one can obtain the terminal points in Fig. 14 from the interaction vertices in Fig. 13 by removing the interaction line and the Green's function line to the right of it, and making the substitution  $\sigma_a \rightarrow \sigma_2$ . The interaction vertices displayed in Fig. 13 and the terminal points displayed in Fig. 14 continue to follow the same pattern for greater than three stars.

To actually draw the diagram, let us momentarily ignore the subscripts in our numerical representation, and correspondingly the Green's function lines labeled by  $\bar{\sigma}_a$  and  $\bar{\sigma}_b$  in panel (a) of Fig. 13, (the case of 0 stars). Then the top row of the numerical representation corresponds to a chain of interaction lines connected to each other by Green's function lines running from the point  $i$  to the point  $f$ . The lower rows also correspond to a similar chain running from a single point back to itself. This is the point  $k$  (displayed in panels (b), (c), and (d) in Fig. 13) on the interaction vertex corresponding to the starred integer which gives rise to this lower row. Thus, the number of such chains beginning and ending at a point of a particular interaction vertex is equal to the number of stars on the starred integer which corresponds to this vertex. For the example given above, following this procedure yields the intermediate diagram displayed in Fig. 15.

Finally, to put the subscripts back into the diagram, we break each chain at any place where there are subscripts between two consecutive vertices of the chain, and pass the chain through the (non-starred) vertices indicated by the subscripts in the order in which they are written, after



**Fig. 13.** Interaction vertices appearing in the diagrams. Panels (a), (b), (c), and (d) correspond to 0, 1, 2, and 3 stars on the number representing the interaction vertex, respectively. Note that the lines are broken into pairs based on spin. A pair of two incoming or two outgoing lines share opposite spins, while a pair of one incoming and one outgoing line share the same spin. Moreover, in the case of the former, the spin contributes to the sign of the diagram, while in the case of the latter, it does not. The contribution to the sign is written in the top right of each panel.

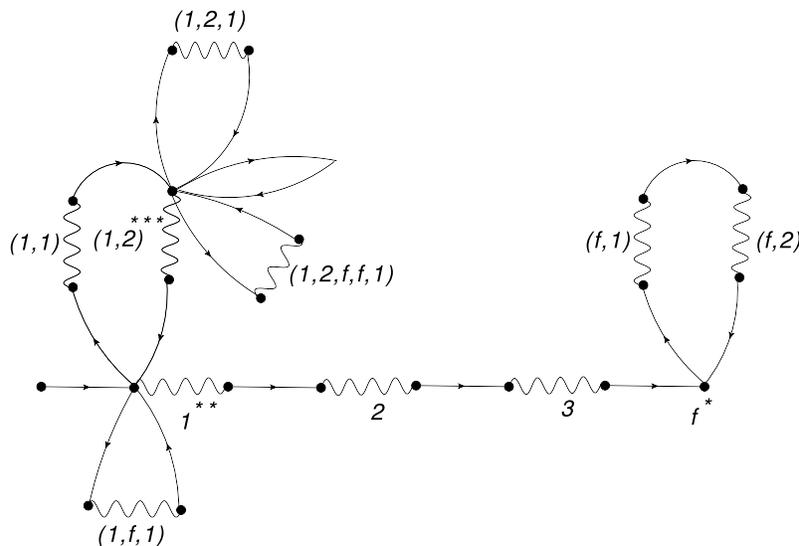


**Fig. 14.** Terminal point in the diagram corresponding to the  $f$  in the top row. Panels (a), (b), (c), and (d) correspond to 0, 1, 2, and 3 stars on the  $f$  in the top row, respectively. Same comments regarding spin apply as in Fig. 13. Note that in the case of one or more stars on the  $f$  in the top row, the line labeled by  $\bar{\sigma}_2$  is outgoing. This is compensated by the fact there are two more lines entering the point  $f$  than exiting it.

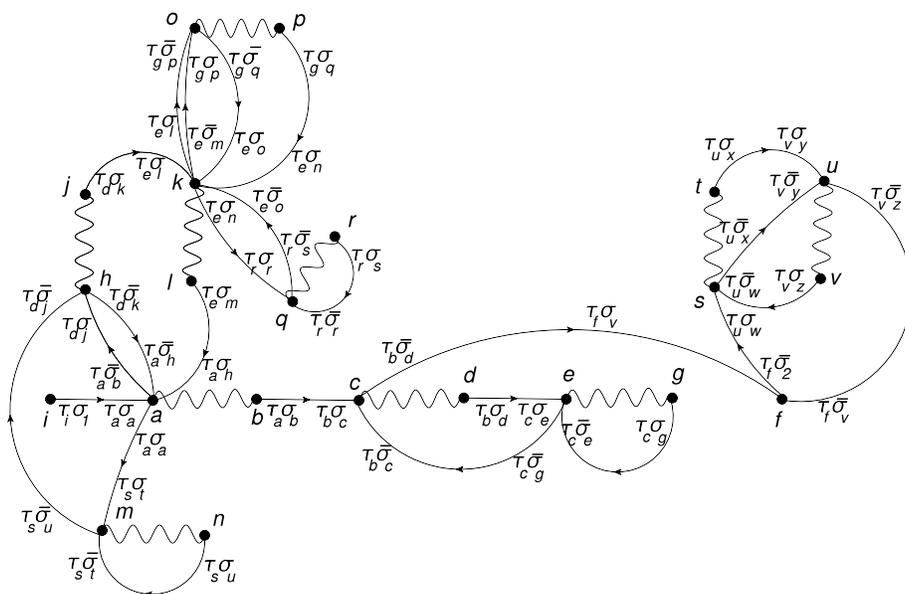
which it resumes its original course. This is accomplished with the help of the two Green's function lines labeled by  $\bar{\sigma}_a$  and  $\bar{\sigma}_b$  on the non-starred vertices, (displayed in panel (a) of Fig. 13), which were ignored in drawing the intermediate diagram in Fig. 15. The final diagram is displayed in Fig. 16.

Note that when drawing a vertex (or terminal point) with multiple stars, such as that displayed in Fig. 13(d), the lines  $\bar{\sigma}_a$  and  $\sigma_c$  (incoming) correspond to the row with 2 stars on its  $f$ , the lines  $\sigma_b$  (outgoing) and  $\sigma_d$  correspond to the row with 1 star on its  $f$ , and the lines  $\sigma_c$  (outgoing) and  $\bar{\sigma}_d$  correspond to the row with 0 stars on its  $f$ . Therefore, in Fig. 16, on the point  $k$  corresponding to the vertex (1, 2)\*\*\*, the lines  $\bar{\sigma}_m$  and  $\sigma_n$  (incoming) are part of the row (1, 2)\*\*\* (3rd row) in the numerical representation, the lines  $\sigma_l$  (outgoing) and  $\sigma_o$  are part of the row (1, 2,  $f$ )\*\* (4th row) in the numerical representation, and the lines  $\sigma_n$  (outgoing) and  $\bar{\sigma}_o$  are part of the row (1, 2,  $f$ ,  $f$ )\* (5th row) in the numerical representation.

- (7) Each solid line in the diagram contributes a non-interacting Green's function, each wavy line contributes a hopping matrix element. An equal-time Green's function is always taken to be  $\mathbf{g}^{(0)}(\tau, \tau^+)$ , i.e. the incoming (creation) line is given the greater time.
- (8) The total sign of the diagram is given by  $(-1)^n (-1)^s (-1)^{s_f - 1} \times (\text{sign from the spins})$ , where in the case of  $s_f = 0$ ,  $(-1)^{s_f - 1} \equiv 1$ , and the way in which the spins contribute to the sign is indicated



**Fig. 15.** Intermediate step in the process of drawing the diagram corresponding to the numerical representation in step 5 of the rules. All of the interaction vertices are drawn in. To complete the diagram, we must split some of the Green's function lines through the unused points in the interaction vertices in a manner indicated by the numerical representation.



**Fig. 16.** Diagram corresponding to the numerical representation in step 5 of the rules.

**Figs. 13 and 14.** Therefore, the diagram in **Fig. 16** has sign  $(-1)^{14}(-1)^5(-1)^0$  (sign from the spins)  $= -\sigma_b\sigma_h\sigma_c\sigma_d\sigma_e\sigma_g\sigma_v\sigma_2\sigma_w\sigma_x\sigma_y\sigma_z\sigma_t\sigma_u\sigma_j\sigma_k\sigma_m\sigma_o\sigma_p\sigma_q\sigma_r\sigma_s$ .

- (9) Sum over internal sites and spins, and integrate over internal times.

According to the above rules, the contribution of the diagram drawn in **Fig. 16** is

$$\begin{aligned}
 & -\sigma_b\sigma_h\sigma_c\sigma_d\sigma_e\sigma_g\sigma_v\sigma_2\sigma_w\sigma_x\sigma_y\sigma_z\sigma_t\sigma_u\sigma_j\sigma_k\sigma_m\sigma_o\sigma_p\sigma_q\sigma_r\sigma_s \\
 & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] \mathbf{t}_{ab} \mathbf{g}_{bc}^{(0)}[\tau_a, \tau_b] \mathbf{t}_{cd} \mathbf{g}_{de}^{(0)}[\tau_b, \tau_c] \mathbf{t}_{eg} \mathbf{g}_{ge}^{(0)}[\tau_c, \tau_c^+] \mathbf{g}_{ec}^{(0)}[\tau_c, \tau_b] \\
 & \mathbf{g}_{cf}^{(0)}[\tau_b, \tau_f] \mathbf{g}_{ah}^{(0)}[\tau_a, \tau_d] \mathbf{t}_{hj} \mathbf{g}_{jk}^{(0)}[\tau_d, \tau_l] \mathbf{t}_{kl} \mathbf{g}_{la}^{(0)}[\tau_l, \tau_a] \mathbf{g}_{ko}^{(0)}[\tau_e, \tau_g] \mathbf{t}_{op}
 \end{aligned}$$

$$\begin{aligned}
& \mathbf{g}_{\sigma_q \sigma_n}^{(0)} [\tau_g, \tau_e] \mathbf{g}_{\bar{\sigma}_m \sigma_p}^{(0)} [\tau_e, \tau_g] \mathbf{g}_{\bar{\sigma}_q \sigma_o}^{(0)} [\tau_g, \tau_e] \mathbf{g}_{\sigma_n \sigma_r}^{(0)} [\tau_e, \tau_r] t_{qr} \\
& \mathbf{g}_{\sigma_s \sigma_r}^{(0)} [\tau_r, \tau_r^+] \mathbf{g}_{\bar{\sigma}_s \sigma_o}^{(0)} [\tau_r, \tau_e] \mathbf{g}_{\sigma_a \sigma_t}^{(0)} [\tau_a, \tau_s] t_{mn} \mathbf{g}_{\sigma_u \sigma_t}^{(0)} [\tau_s, \tau_s^+] \mathbf{g}_{\bar{\sigma}_u \sigma_j}^{(0)} [\tau_s, \tau_d] \mathbf{g}_{\bar{\sigma}_k \sigma_h}^{(0)} [\tau_d, \tau_a] \\
& \mathbf{g}_{\bar{\sigma}_2 \sigma_v}^{(0)} [\tau_f, \tau_u] t_{st} \mathbf{g}_{\sigma_x \sigma_y}^{(0)} [\tau_u, \tau_v] t_{uv} \mathbf{g}_{\sigma_2 \sigma_w}^{(0)} [\tau_v, \tau_u] \mathbf{g}_{\bar{\sigma}_x \sigma_y}^{(0)} [\tau_u, \tau_v] \mathbf{g}_{\bar{\sigma}_z \sigma_v}^{(0)} [\tau_v, \tau_f].
\end{aligned}$$

Upon turning off the sources, the Green's functions become spin diagonal, i.e.  $\mathbf{g}_{\sigma_1 \sigma_2}^{(0)}[i, f] = \delta_{\sigma_1 \sigma_2} \mathbf{g}_{\uparrow \uparrow}^{(0)}[i, f] = \delta_{\sigma_1 \sigma_2} \mathbf{g}_{\downarrow \downarrow}^{(0)}[i, f] \equiv \delta_{\sigma_1 \sigma_2} \mathbf{g}^{(0)}[i, f]$ . This allows one to evaluate the spin sum and the sign of the above expression. A good way to evaluate the spin sum is to break the diagram into spin loops in the following manner. Recall that at each interaction vertex and at the terminal point, lines are paired according to spin. They share the same spin if one is incoming and the other is outgoing, and they have opposite spins if both lines of the pair are incoming or both are outgoing. Starting with the line exiting  $i$ , follow the path of Green's function lines created by the spin pairings until you reach the line labeled by  $\sigma_2$  (or  $\bar{\sigma}_2$  if  $f$  has one or more stars). These spins are all set by the value of  $\sigma_1 = \sigma_2$ , and therefore this is the zeroth spin loop. If not all of the lines have been used up by the zeroth loop, find a random line and follow the path created by the spin pairings to reach the line to which it is paired. This is the first spin loop, etc. Continue to do this until you have used up all of the lines in the diagram. Let  $F_s$  denote the number of spin loops in the diagram. Then, the spin sum is  $2^{F_s}$ . We emphasize that unlike the case of the standard Feynman diagrams, the spin loops of the  $\lambda$  expansion diagrams do not coincide with the Fermi loops (where each row of the numerical representation can be thought of as a Fermi loop).

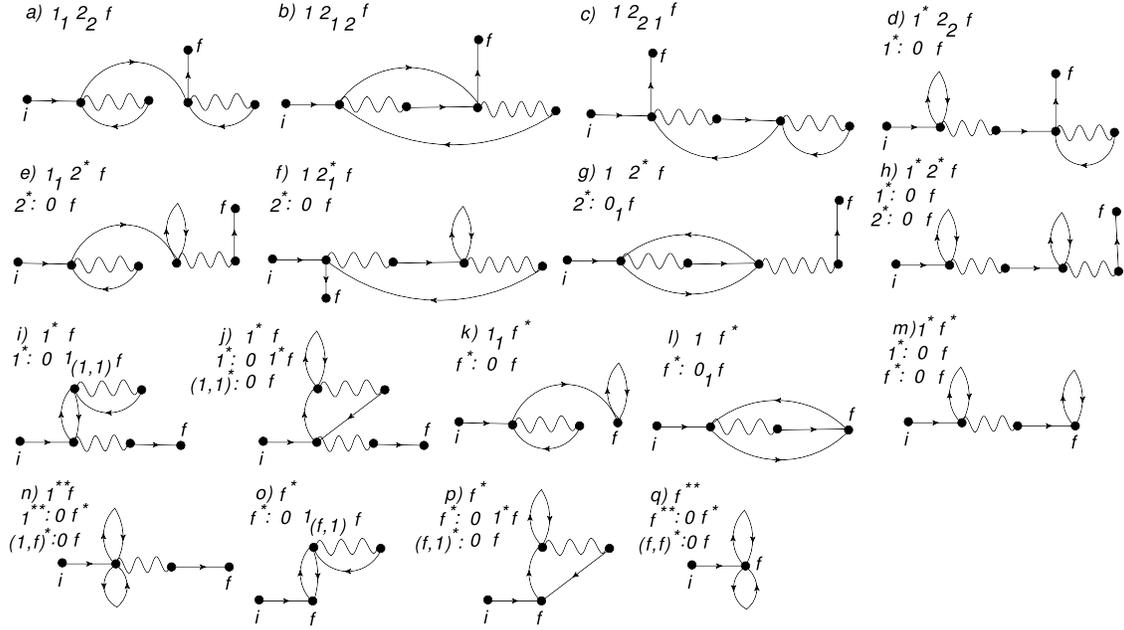
To determine the sign of the diagram, assign values to the spins in a manner consistent with the spin loops (i.e. the value of any one spin in the spin loop determines the values of all of them). Then, plug these values into the analytical expression for the diagram. It is important to note that the reason we can compute the spin sum and the sign independently, is that the choice we make for the values of the spins does not affect the sign of the diagram. To see this note that every spin loop consists of an even number of pairs that have either two incoming lines or two outgoing lines (since it has an equal number of each kind), and an arbitrary number which have one incoming line and one outgoing line. However, only the former contributes to the sign, while the latter does not (see Figs. 13 and 14.) Moreover, each pair contributes a distinct spin and appears in exactly one spin loop. Therefore, by flipping all of the spins in a spin loop, we flip an even number of spins, and therefore do not change the sign of the diagram. The only exception to this line of reasoning is the zeroth spin loop, in the case when the terminal point  $f$  has 1 or more stars (see Fig. 14). In this case, the zeroth spin loop must have one more pair where both lines are incoming than it has pairs where both lines are outgoing. This is due to the fact that in this case both the spins  $\sigma_1$  and  $\bar{\sigma}_2$  exit the sites  $i$  and  $f$  respectively. It is also consistent with the fact that the terminal point  $f$  now has one more pair with two incoming lines than two outgoing lines. Therefore, the spin pairs in the zeroth spin loop now contribute an odd number of spins. However, the spin  $\sigma_2$  from the zeroth spin loop now also appears explicitly in the sign. Therefore, flipping all of the spins in the zeroth spin loop once again does not change the sign of the diagram. In Fig. 16, we find

$$\begin{aligned}
(\sigma_1) &= (\sigma_d) = \sigma_t = \bar{\sigma}_u = \bar{\sigma}_j = \sigma_b = \sigma_c = \sigma_g = \bar{\sigma}_e = \bar{\sigma}_d = \sigma_v = \sigma_z = \bar{\sigma}_w = \sigma_2; \quad \sigma_x = \sigma_y; \\
\sigma_h &= \sigma_k = (\sigma_l) = \bar{\sigma}_p = \sigma_m; \quad \sigma_q = \bar{\sigma}_o = \bar{\sigma}_s = \sigma_r = (\sigma_n),
\end{aligned}$$

where the parenthesis indicates that the spin does not contribute to the sign of the diagram. Therefore,  $F_s = 3$ . The loops contribute  $(-1)^{5+0+1+2} = (-1)^8 = 1$  to the sign. Therefore, the final contribution of the diagram in Fig. 16 is

$$\begin{aligned}
& -8 \times \mathbf{g}_{i_a}^{(0)} [\tau_i, \tau_d] t_{ab} \mathbf{g}_{bc}^{(0)} [\tau_a, \tau_b] t_{cd} \mathbf{g}_{de}^{(0)} [\tau_b, \tau_c] t_{eg} \mathbf{g}_{ge}^{(0)} [\tau_c, \tau_c^+] \mathbf{g}_{ec}^{(0)} [\tau_c, \tau_b] \\
& \mathbf{g}_{cf}^{(0)} [\tau_b, \tau_f] \mathbf{g}_{ah}^{(0)} [\tau_a, \tau_d] t_{hj} \mathbf{g}_{jk}^{(0)} [\tau_d, \tau_l] t_{kl} \mathbf{g}_{la}^{(0)} [\tau_l, \tau_a] \mathbf{g}_{ko}^{(0)} [\tau_e, \tau_g] t_{op} \mathbf{g}_{pk}^{(0)} [\tau_g, \tau_e] \mathbf{g}_{ko}^{(0)} [\tau_e, \tau_g] \\
& \mathbf{g}_{ok}^{(0)} [\tau_g, \tau_e] \mathbf{g}_{kq}^{(0)} [\tau_e, \tau_r] t_{qr} \mathbf{g}_{rq}^{(0)} [\tau_r, \tau_r^+] \mathbf{g}_{qk}^{(0)} [\tau_r, \tau_e] \mathbf{g}_{am}^{(0)} [\tau_a, \tau_s] \\
& t_{mn} \mathbf{g}_{nm}^{(0)} [\tau_s, \tau_s^+] \mathbf{g}_{mh}^{(0)} [\tau_s, \tau_d] \mathbf{g}_{ha}^{(0)} [\tau_d, \tau_d] \mathbf{g}_{fs}^{(0)} [\tau_f, \tau_u] t_{st} \mathbf{g}_{tu}^{(0)} [\tau_u, \tau_v] t_{uv} \mathbf{g}_{vs}^{(0)} [\tau_v, \tau_u] \mathbf{g}_{su}^{(0)} [\tau_u, \tau_v] \mathbf{g}_{uf}^{(0)} [\tau_v, \tau_f],
\end{aligned}$$

where all sites and times other than  $i$  and  $f$ , and  $\tau_i$  and  $\tau_f$  are summed/integrated over.



**Fig. 17.** The second order diagrams contributing to the Green's function:  $\mathcal{G}^{(2)}[i, f]$ , and their corresponding numerical representations. Note that the diagrams (a) through (j) are the standard second order Feynman diagrams. The other diagrams are not.

#### 4.3. Second order contribution

Using the rules from Section 4.2, we draw the diagrams that contribute to  $\mathcal{G}[i, f]$  in second order in Fig. 17, and calculate their contributions below.

The contributions of these diagrams are

$$\begin{aligned}
 (a) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{ba}^{(0)}[\tau_a, \tau_a^+] \mathbf{g}_{ac}^{(0)}[\tau_a, \tau_b] t_{cd} \mathbf{g}_{dc}^{(0)}[\tau_b, \tau_b^+] \mathbf{g}_{cf}^{(0)}[\tau_b, \tau_f] \\
 (b) & - 2 \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bc}^{(0)}[\tau_a, \tau_c] t_{cd} \mathbf{g}_{da}^{(0)}[\tau_c, \tau_a] \mathbf{g}_{ac}^{(0)}[\tau_a, \tau_c] \mathbf{g}_{cf}^{(0)}[\tau_c, \tau_f] \\
 (c) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bc}^{(0)}[\tau_a, \tau_b] t_{cd} \mathbf{g}_{dc}^{(0)}[\tau_b, \tau_b^+] \mathbf{g}_{ca}^{(0)}[\tau_b, \tau_a] \mathbf{g}_{af}^{(0)}[\tau_a, \tau_f] \\
 (d) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] t_{ab} \mathbf{g}_{bc}^{(0)}[\tau_a, \tau_b] t_{cd} \mathbf{g}_{dc}^{(0)}[\tau_b, \tau_b^+] \mathbf{g}_{cf}^{(0)}[\tau_b, \tau_f] \\
 (e) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{ba}^{(0)}[\tau_a, \tau_a^+] \mathbf{g}_{ac}^{(0)}[\tau_a, \tau_b] \mathbf{g}_{cc}^{(0)}[\tau_b, \tau_b^+] t_{cd} \mathbf{g}_{df}^{(0)}[\tau_b, \tau_f] \\
 (f) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bc}^{(0)}[\tau_b, \tau_b^+] t_{cd} \mathbf{g}_{da}^{(0)}[\tau_b, \tau_a] \mathbf{g}_{af}^{(0)}[\tau_a, \tau_f] \\
 (g) & - 2 \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bc}^{(0)}[\tau_a, \tau_b] \mathbf{g}_{ca}^{(0)}[\tau_b, \tau_a] \mathbf{g}_{ac}^{(0)}[\tau_a, \tau_b] t_{cd} \mathbf{g}_{ia}^{(0)}[\tau_b, \tau_f] \\
 (h) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] t_{ab} \mathbf{g}_{bc}^{(0)}[\tau_a, \tau_b] \mathbf{g}_{cc}^{(0)}[\tau_b, \tau_b^+] t_{cd} \mathbf{g}_{df}^{(0)}[\tau_b, \tau_f] \\
 (i) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f] \mathbf{g}_{ac}^{(0)}[\tau_a, \tau_b] t_{cd} \mathbf{g}_{dc}^{(0)}[\tau_b, \tau_b^+] \mathbf{g}_{ca}^{(0)}[\tau_b, \tau_a] \\
 (j) & \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f] \mathbf{g}_{ac}^{(0)}[\tau_a, \tau_b] \mathbf{g}_{cc}^{(0)}[\tau_b, \tau_b^+] t_{cd} \mathbf{g}_{da}^{(0)}[\tau_b, \tau_a] \\
 (k) & - \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{ba}^{(0)}[\tau_a, \tau_a^+] \mathbf{g}_{af}^{(0)}[\tau_a, \tau_f] \mathbf{g}_{ff}^{(0)}[\tau_f, \tau_f^+] \\
 (l) & 2 \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f] \mathbf{g}_{fa}^{(0)}[\tau_f, \tau_a] \mathbf{g}_{af}^{(0)}[\tau_a, \tau_f] \\
 (m) & - \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f] \mathbf{g}_{ff}^{(0)}[\tau_f, \tau_f^+] \\
 (n) & - \mathbf{g}_{ia}^{(0)}[\tau_i, \tau_a] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f] \\
 (o) & - \mathbf{g}_{if}^{(0)}[\tau_i, \tau_f] \mathbf{g}_{fa}^{(0)}[\tau_f, \tau_a] t_{ab} \mathbf{g}_{ba}^{(0)}[\tau_a, \tau_a^+] \mathbf{g}_{af}^{(0)}[\tau_a, \tau_f]
 \end{aligned}$$

$$(p) - \mathbf{g}_{if}^{(0)}[\tau_i, \tau_f] \mathbf{g}_{fa}^{(0)}[\tau_f, \tau_a] \mathbf{g}_{aa}^{(0)}[\tau_a, \tau_a^+] t_{ab} \mathbf{g}_{bf}^{(0)}[\tau_a, \tau_f]$$

$$(q) \mathbf{g}_{if}^{(0)}[\tau_i, \tau_f] \mathbf{g}_{ff}^{(0)}[\tau_f, \tau_f^+] \mathbf{g}_{ff}^{(0)}[\tau_f, \tau_f^+].$$

## 5. Diagrammatic $\lambda$ expansion for constituent objects

### 5.1. Introduction of the two self-energies

We next consider the auxiliary Green's function  $\mathbf{g}[i, f]$ . Using Eq. (10) for  $\mathbf{g}^{-1}[i, f]$ , we can write the analog of Eq. (17) for  $\mathbf{g}[i, f]$ .

$$\mathbf{g}_{\sigma_1, \sigma_2}[i, f] = \mathbf{g}_{\sigma_1, \sigma_2}^{(0)}[i, f] - \lambda \mathbf{g}_{\sigma_1, \sigma_b}^{(0)}[i, \mathbf{k}] \left( -t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \mathcal{G}_{\bar{\sigma}_a, \bar{\sigma}_b}[\mathbf{k}, \mathbf{k}^+] \mathbf{g}_{\sigma_a, \sigma_2}[\mathbf{j}, f] + t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}}^{\bar{\sigma}_b, \bar{\sigma}_a}} \mathbf{g}_{\sigma_a, \sigma_2}[\mathbf{j}, f] \right), \quad (19)$$

Comparing the iterative expansion of  $\mathcal{G}[i, f]$  through Eq. (17) with that of  $\mathbf{g}[i, f]$  through Eq. (19), we see that the terms in the parenthesis are identical in both expansions. However, the second term on the RHS of Eq. (17), missing in Eq. (19), allows a Green's function diagram to close on itself in the iterative expansion, merging the initial point  $i$  and the terminal point  $f$ . Such a diagram must necessarily have more than one line connected to its terminal point, and therefore at least one star on the  $f$  in the top row. Therefore, the diagrams for  $\mathbf{g}[i, f]$  are the subset of the diagrams for  $\mathcal{G}[i, f]$  which have no stars on the  $f$  in the top row. In Fig. 17, these are diagrams (a) through (j), and diagram (n).

We see that in the diagrams for  $\mathbf{g}[i, f]$ , the terminal point labeled by  $f$  is connected to the rest of the diagram only by a single line. Therefore, it will be possible to describe these diagrams in terms of a Dyson equation, with a Dyson self-energy. This is not the case for the other diagrams in Fig. 17 (those which do have a star on the  $f$  in the top row), and these diagrams require the introduction of a second self-energy. We now proceed to define these two types of self-energies.

We shall denote the Dyson self-energy for  $\mathbf{g}[i, f]$  by  $\Sigma_a$ . As is the case in the Feynman diagrams, it is obtained from the diagrams for  $\mathbf{g}[i, f]$  by removing the external line coming in from the point  $i$ , and the one going out to the point  $f$ . If a diagram for  $\Sigma_a$  can be split into two pieces by cutting a single line, then it is reducible. Otherwise, it is irreducible. Denote the irreducible part of  $\Sigma_a$  by  $\Sigma_a^*$ .

Now consider those diagrams which do have a star on the  $f$  in the top row. The second self-energy,  $\Sigma_b$ , is obtained from these diagrams by removing the external line coming from the point  $i$ . Once again, if a diagram for  $\Sigma_b$  can be split into two pieces by cutting a single line, then it is reducible. Otherwise, it is irreducible. Denote the irreducible part of  $\Sigma_b$  by  $\Sigma_b^*$ .

From the diagrammatic structure of the series, it is clear that  $\mathcal{G}[i, f] = \mathbf{g}[i, f] + \mathbf{g}[i, \mathbf{j}] \Sigma_b^*[\mathbf{j}, f]$ . Comparing with Eq. (9), we see that  $\tilde{\mu}[i, f] = \delta[i, f] + \Sigma_b^*[i, f]$ . Also, from Dyson's equation, we know that  $\mathbf{g}^{-1}[i, f] = \mathbf{g}^{-1(0)}[i, f] - \Sigma_a^*[i, f]$ . We shall give an independent proof of these formulae starting from the equations of motion for  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  (Eq. (10)) in Section 5.2.

### 5.2. $\mathbf{g}^{-1}$ and $\tilde{\mu}$

We shall now prove, starting with the equations of motion in Eq. (10), the observations already made in Section 5.1, that

$$\tilde{\mu}[i, f] = \delta[i, f] + \Sigma_b^*[i, f]; \quad \mathbf{g}^{-1}[i, f] = \mathbf{g}^{-1(0)}[i, f] - \Sigma_a^*[i, f]. \quad (20)$$

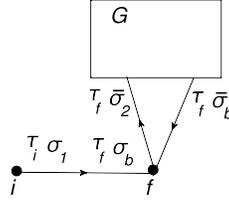
This is equivalent to showing that

$$\Sigma_b^*[i, f] = -\lambda \gamma[i] \delta[i, f] + \lambda \Psi[i, f]; \quad \Sigma_a^*[i, f] = \lambda \gamma[i] t[i, f] + \lambda \Phi[i, f]. \quad (21)$$

We rewrite the EOM for  $\tilde{\mu}[i, f]$  (Eq. (10)) in expanded form.

$$\tilde{\mu}_{\sigma_1 \sigma_2}[i, f] = (\delta_{\sigma_1 \sigma_2} - \lambda \sigma_1 \sigma_2 \mathcal{G}_{\bar{\sigma}_2 \bar{\sigma}_1}[i, i^+]) \delta[i, f] + \lambda \Psi_{\sigma_1 \sigma_2}[i, f]$$

$$\Psi_{\sigma_1 \sigma_2}[i, f] = -t[i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a, \sigma_b}[\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}_{\mathbf{i}}^{\sigma_1 \sigma_a}} \tilde{\mu}_{\sigma_b \sigma_2}[\mathbf{n}, f]. \quad (22)$$



**Fig. 18.** Schematic representation for a Green's function diagram with only the number  $f^{*...}$  in the top row. Upon removing the incoming external line, it contributes to  $\Sigma_{b1}^*[i, f]$ .

We now proceed to prove the first of Eq. (20) using induction in  $\lambda$ . The lowest order contribution to  $\Sigma_b^*[i, f]$  comes from diagram (a) in Fig. 12. Removing the incoming external line, we obtain  $\Sigma_{\sigma_1\sigma_2}^{*(1)}[i, f] = -\lambda\sigma_1\sigma_2\delta[i, f]\mathbf{g}_{\bar{\sigma}_2\bar{\sigma}_1}^{(0)}[\tau_f, \tau_f^+]$ . Using Eq. (22) to obtain the first order contribution to  $\tilde{\mu}[i, f]$ , we get  $\tilde{\mu}_{\sigma_1\sigma_2}^{(1)}[i, f] = -\lambda\sigma_1\sigma_2\mathbf{g}_{\bar{\sigma}_2, \bar{\sigma}_1}^{(0)}[i, i^+]\delta[i, f]$ . Clearly, these two are equal, and we have that  $\Sigma_{\sigma_1\sigma_2}^{*(1)}[i, f] = \tilde{\mu}_{\sigma_1\sigma_2}^{(1)}[i, f]$ .

Now consider the  $m$ th order contribution  $\Sigma_{\sigma_1\sigma_2}^{*(m)}[i, f]$ . This will be obtained from the corresponding  $m$ th order  $\mathcal{G}[i, f]$  diagram upon dropping the incoming external line. If in the numerical representation for this  $\mathcal{G}[i, f]$  diagram, there are no numbers other than  $f^{*...}$  in the top row (e.g. panels (o), (p), and (q) in Fig. 17), then the contribution of this diagram to  $\mathcal{G}[i, f]$  is  $\mathcal{G}_{\sigma_1\sigma_2}^{(m)}[i, f] = -\lambda\mathbf{g}_{\sigma_1\sigma_b}^{(0)}[i, f]\sigma_b\sigma_2\mathcal{G}_{\bar{\sigma}_2\bar{\sigma}_b}^{(m-1)}[f, f^+]$  (see Fig. 18). The resulting contribution to  $\Sigma_b^*$ , which we shall denote by  $\Sigma_{b1}^*$ , is

$$\Sigma_{\sigma_1\sigma_2}^{*(m)}[i, f] = -\lambda\sigma_1\sigma_2\mathcal{G}_{\bar{\sigma}_2\bar{\sigma}_1}^{(m-1)}[f, f^+]\delta[i, f]. \quad (23)$$

Alternatively, there are numbers other than  $f^{*...}$  in the top row of the numerical representation of the corresponding  $m$ th order  $\mathcal{G}[i, f]$  diagram. Then, the top row reads  $1 \dots f^{*...}$  (e.g. panels (k) through (m) of Fig. 17). In this case, we know that for the resulting  $\Sigma_b[i, f]$  diagram to be irreducible, i.e. for it to contribute to  $\Sigma_b^*[i, f]$ , the number 1 in the top row should not be starred. Therefore, we can represent the diagram schematically as in Fig. 19. This representation is obtained as follows. If we consider just the part of the diagram between the points  $j$  and  $f$ , we know that a line in this part of the diagram (denoted in Fig. 19 by the letter  $s$ ) is split by the point  $k$ . If we restore  $s$  by removing the lines labeled by  $\bar{\sigma}_a$  and  $\bar{\sigma}_b$  from the point  $k$ , then the part of the diagram running from  $j$  to  $f$  is a Green's function diagram which contributes to  $\mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \Sigma_b^*[\mathbf{n}, f]$  (since the  $f$  has at least one star on it). However, the line  $s$  cannot be contained in the  $\mathbf{g}[\mathbf{j}, \mathbf{n}]$  part of the diagram (represented in Fig. 19 by a double line), since then the resulting  $\Sigma_b$  (of the overall diagram) would be reducible. Therefore it must be contained in the  $\Sigma_b^*[\mathbf{n}, f]$  part of the diagram. The analytical expression for the diagram in Fig. 19 is

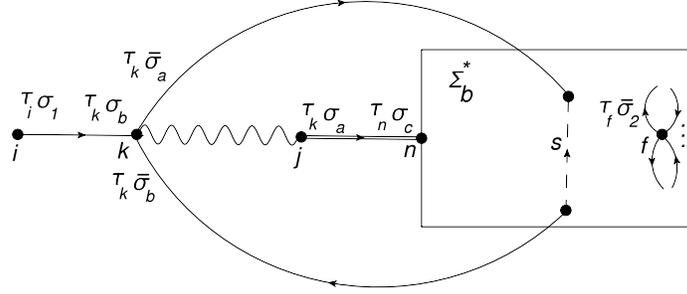
$$-\lambda\mathbf{g}_{\sigma_1\sigma_b}^{(0)}[i, \mathbf{k}]t[\mathbf{k}, \mathbf{j}]\sigma_b\sigma_a\mathbf{g}_{\sigma_a\sigma_c}[\mathbf{j}, \mathbf{n}]\frac{\delta}{\delta\mathcal{V}_{\mathbf{k}}^{\bar{\sigma}_b\bar{\sigma}_a}}\Sigma_{\sigma_c\sigma_2}^*[\mathbf{n}, f]. \quad (24)$$

Removing the incoming external line, and using the inductive hypothesis, we obtain the contribution of these types of diagrams to  $\Sigma_b^*[i, f]$ , which we shall denote as  $\Sigma_{b2}^*[i, f]$ .

$$\Sigma_{\sigma_1\sigma_2}^{*(m)}[i, f] = -\lambda t[i, \mathbf{j}]\sigma_1\sigma_a\mathbf{g}_{\sigma_a\sigma_b}^{(m_1)}[\mathbf{j}, \mathbf{n}]\frac{\delta}{\delta\mathcal{V}_i^{\bar{\sigma}_1\bar{\sigma}_a}}\tilde{\mu}_{\sigma_b\sigma_2}^{(m_2)}[\mathbf{n}, f], \quad (25)$$

where  $m = m_1 + m_2 + 1$ . Comparing Eq. (25) with Eq. (22), we see that

$$\Sigma_{\sigma_1\sigma_2}^*[i, f] = \lambda\Psi_{\sigma_1\sigma_2}[i, f]. \quad (26)$$



**Fig. 19.** Schematic representation for a Green's function diagram whose top row is  $1 \dots f^{* \dots *}$ . Upon removing the incoming external line, it contributes to  $\Sigma_{b2}^*[i, f]$ .

Combining Eqs. (23) and (26), we find that

$$\Sigma_{\sigma_1 \sigma_2}^* [i, f] = \Sigma_{\sigma_1 \sigma_2}^* [i, f] + \Sigma_{\sigma_1 \sigma_2}^* [i, f] = -\lambda \sigma_1 \sigma_2 \mathcal{G}_{\bar{\sigma}_2 \bar{\sigma}_1} [f, f^+] \delta [i, f] + \lambda \Psi_{\sigma_1 \sigma_2} [i, f]. \quad (27)$$

Therefore, comparing Eq. (27) with Eq. (22), we have shown the first of Eq. (20) to be true.

Now consider the EOM for  $\mathbf{g}^{-1}[i, f]$  (Eq. (10)) in expanded form.

$$\begin{aligned} \mathbf{g}_{\sigma_1 \sigma_2}^{-1} [i, f] &= \mathbf{g}_{\sigma_1 \sigma_2}^{-1(0)} [i, f] - \lambda t [i, f] \sigma_1 \sigma_2 \mathcal{G}_{\bar{\sigma}_2 \bar{\sigma}_1} [i, i^+] - \lambda \Phi_{\sigma_1 \sigma_2} [i, f] \\ \Phi_{\sigma_1 \sigma_2} [i, f] &= t [i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a, \sigma_b} [\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_1 \bar{\sigma}_a}} \mathbf{g}_{\sigma_b \sigma_2}^{-1} [\mathbf{n}, f]. \end{aligned} \quad (28)$$

Our goal is to prove the second of Eq. (20) using Eq. (28). To this end, we note that diagrams for  $\Sigma_a^*[i, f]$  can be split into four groups. Recall that a diagram for  $\Sigma_a[i, f]$  is obtained from a  $\mathbf{g}[i, f]$  diagram (or equivalently from a  $\mathcal{G}[i, f]$  diagram with no stars on the  $f$  in the top row) by removing the incoming and outgoing external lines. Consider a  $\mathbf{g}[i, f]$  diagrams whose numerical representation has the following property. There are no subscripts between the number immediately to the left of  $f$  in the top row (which we shall denote by  $c$ ) and  $f$  (e.g. panels (e), (g) through (j), and (n) of Fig. 17). This implies that  $c$  has at least one star, as otherwise  $c$  must be a subscript between  $c$  and  $f$ . Therefore, the top row looks like  $1 \dots c^{* \dots *} f$ . In the case that  $c = 1$  (e.g. panels (i), (j), and (n) of Fig. 17), these diagrams can be represented schematically as in Fig. 20. We denote the corresponding contribution to  $\Sigma_a^*$  by  $\Sigma_{a1}^*$ . If  $c > 1$  (e.g. panels (e), (g), and (h) of Fig. 17), then the diagrams can be represented schematically as in Fig. 21. We denote the corresponding contribution to  $\Sigma_a^*$  by  $\Sigma_{a2}^*$ . Comparing Fig. 18 with Fig. 20 and Fig. 19 with Fig. 21, and removing the external lines, we find that

$$\Sigma_{\sigma_1 \sigma_2}^* [i, f] = -\Sigma_{\sigma_1 \sigma_2}^* [i, \mathbf{j}] t [\mathbf{j}, f]; \quad \Sigma_{\sigma_1 \sigma_2}^* [i, f] = -\Sigma_{\sigma_1 \sigma_2}^* [i, \mathbf{j}] t [\mathbf{j}, f]. \quad (29)$$

Here, the minus comes from rule (8) of Section 4.2, where there is a minus sign discrepancy between the factors  $(-1)^{s_f-1}$  (applicable to  $f^{* \dots *}$  in  $\Sigma_b^*$ ) and  $(-1)^s$  (applicable to  $c^{* \dots *}$  in  $\Sigma_a^*$ ). Using Eqs. (23) and (26), we find that

$$-\Sigma_{\sigma_1 \sigma_2}^* [i, f] = -\lambda \sigma_1 \sigma_2 \mathcal{G}_{\bar{\sigma}_2 \bar{\sigma}_1} [i, i^+] t [i, f]; \quad -\Sigma_{\sigma_1 \sigma_2}^* [i, f] = \lambda \Psi_{\sigma_1 \sigma_2} [i, \mathbf{j}] t [\mathbf{j}, f]. \quad (30)$$

Motivated by this observation, we define a new object  $\chi_{\sigma_1 \sigma_2}[i, f]$  defined by the formula

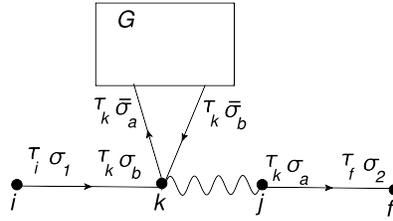
$$\Phi_{\sigma_1 \sigma_2} [i, f] = -\Psi_{\sigma_1 \sigma_2} [i, \mathbf{j}] t [\mathbf{j}, f] + \chi_{\sigma_1 \sigma_2} [i, f]. \quad (31)$$

Plugging this formula into Eq. (28), we obtain

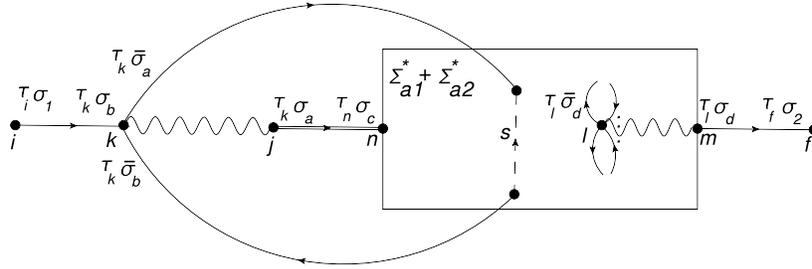
$$\mathbf{g}_{\sigma_1 \sigma_2}^{-1} [i, f] = \mathbf{g}_{\sigma_1 \sigma_2}^{-1(0)} [i, f] - \lambda t [i, f] \sigma_1 \sigma_2 \mathcal{G}_{\bar{\sigma}_2 \bar{\sigma}_1} [i, i^+] + \lambda \Psi_{\sigma_1 \sigma_2} [i, \mathbf{j}] t [\mathbf{j}, f] - \lambda \chi_{\sigma_1 \sigma_2} [i, f] \quad (32)$$

Plugging Eq. (32) into the equation for  $\Phi$  (Eq. (28)), we obtain

$$\begin{aligned} \Phi_{\sigma_1 \sigma_2} [i, f] &= -t [i, \mathbf{j}] \sigma_1 \bar{\sigma}_2 \mathbf{g}_{\bar{\sigma}_2 \bar{\sigma}_1} [\mathbf{j}, i] \delta [i, f] - \Psi_{\sigma_1 \sigma_2} [i, \mathbf{j}] t [\mathbf{j}, f] \\ &\quad - \lambda t [i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a \sigma_b} [\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_1 \bar{\sigma}_a}} \chi_{\sigma_b \sigma_2} [\mathbf{n}, f], \end{aligned} \quad (33)$$



**Fig. 20.** Schematic representation for Green's function diagram whose top row is  $1^* \dots^* f$ . Upon removing the incoming and outgoing external lines, it contributes to  $\Sigma_{a1}^*[i, f]$ . One can obtain  $\Sigma_{b1}^*[i, f]$  displayed in Fig. 18 by also removing the interaction line exiting the point  $k$ .



**Fig. 21.** Schematic representation for Green's function diagram whose top row is  $1 \dots c^* \dots^* f$ . Upon removing the outgoing and incoming external lines, it contributes to  $\Sigma_{a2}^*[i, f]$ . One can obtain  $\Sigma_{b2}^*[i, f]$  displayed in Fig. 19 by also removing the interaction line exiting the point  $l$ .

where we have used Eq. (22) to handle the second and third terms on the RHS of Eq. (32). Comparing Eq. (33) with Eq. (31), we obtain the following EOM for  $\chi_{\sigma_1\sigma_2}[i, f]$ .

$$\begin{aligned} \chi_{\sigma_1\sigma_2}[i, f] = & -t[i, \mathbf{j}]\sigma_1\bar{\sigma}_2\mathbf{g}_{\bar{\sigma}_2\bar{\sigma}_1}[\mathbf{j}, i]\delta[i, f] \\ & - \lambda t[i, \mathbf{j}]\sigma_1\sigma_a\mathbf{g}_{\sigma_a\sigma_b}[\mathbf{j}, \mathbf{n}]\frac{\delta}{\delta\mathcal{V}_i^{\bar{\sigma}_1\bar{\sigma}_a}}\chi_{\sigma_b\sigma_2}[\mathbf{n}, f]. \end{aligned} \quad (34)$$

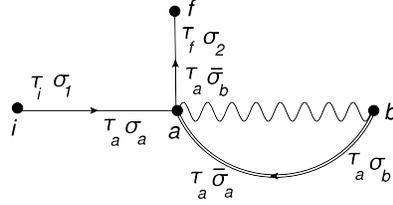
Comparing Eqs. (32) and (30), we see that  $-\Sigma_{\sigma_1\sigma_2}^{*a1}[i, f]$  and  $-\Sigma_{\sigma_1\sigma_2}^{*a2}[i, f]$  account for the second and third terms on the RHS of Eq. (32) respectively. Therefore, we now show that the remainder of the  $-\Sigma_a^*[i, f]$  diagrams account for the fourth term. To this end, we consider all  $\mathbf{g}[i, f]$  diagrams which do have a subscript between the number  $c$  (the number immediately to the left of  $f$  in the top row) and  $f$  in the top row. The top row now looks like  $1 \dots c \dots f$  (e.g. panels (a) through (d) and (f) of Fig. 17). Note that for the resulting  $\Sigma_a[i, f]$  diagram to be irreducible, the number 1 cannot have any stars. We further subdivide this group of  $\mathbf{g}[i, f]$  diagrams into 2 groups. In the first group, whose contribution to  $\Sigma_a^*[i, f]$  shall be denoted by  $\Sigma_{a3}^*[i, f]$ , the subscript immediately preceding  $f$  in the top row is 1. The top row for these diagrams looks like  $1 \dots c \dots 1f$  (e.g. panels (c) and (f) of Fig. 17). In the second group, whose contribution to  $\Sigma_a^*[i, f]$  shall be denoted by  $\Sigma_{a4}^*[i, f]$ , the subscript immediately preceding  $f$  in the top row is not 1. The top row for these diagrams looks like  $1 \dots c \dots df$ , where  $d \neq 1$  (e.g. panels (a), (b), and (d) of Fig. 17). Our goal is to show that  $\lambda\chi_{\sigma_1\sigma_2}[i, f] = \Sigma_{\sigma_1\sigma_2}^{*a3}[i, f] + \Sigma_{\sigma_1\sigma_2}^{*a4}[i, f]$ .

We do this by induction. The  $\mathbf{g}[i, f]$  diagrams contributing to  $\Sigma_{a3}^*[i, f]$  are shown in Fig. 22. The contribution of this diagram becomes

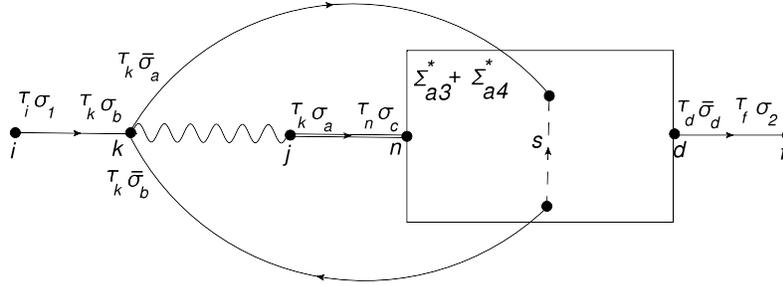
$$-\lambda\sigma_a\sigma_b\mathbf{g}_{\sigma_1\sigma_a}^{(0)ia}[\tau_i, \tau_a]t_{ab}\mathbf{g}_{\sigma_b\sigma_a}[\tau_a, \tau_a^+]\mathbf{g}_{\bar{\sigma}_b\bar{\sigma}_2}^{(0)af}[\tau_a, \tau_f]. \quad (35)$$

After removing the two external lines, we find that

$$\Sigma_{\sigma_1\sigma_2}^{*a3}[i, f] = -\lambda\sigma_1\bar{\sigma}_2t_{ib}\mathbf{g}_{\bar{\sigma}_2\bar{\sigma}_1}[\tau_i, \tau_i^+]\delta[i, f]. \quad (36)$$



**Fig. 22.** Schematic representation for Green's function diagram whose top row is  $1 \dots c \dots f$ . Upon removing the incoming and outgoing external lines, it contributes to  $\Sigma_{a3}^*[i, f]$ . This is the analog of the Fock diagram in the standard Feynman skeleton expansion.



**Fig. 23.** Schematic representation for Green's function diagram whose top row is  $1 \dots c \dots df$ , where  $d \neq 1$ . Upon removing the incoming and outgoing external lines, it contributes to  $\Sigma_{a4}^*[i, f]$ .

Thus,  $\Sigma_{a3}^*[i, f]$  is equal to the first term on the RHS of Eq. (34). Note that since this term contains the lowest order contribution to  $\chi[i, f]$ , this covers the base case of the induction. We now want to show that  $\Sigma_{a4}^*[i, f]$  equals the second term on the RHS of Eq. (34). The  $\mathbf{g}[i, f]$  diagrams contributing to  $\Sigma_{a4}^*[i, f]$  can be represented schematically as in Fig. 23. Here, the reasoning is similar to that which led to Fig. 19. If the line  $s$  were contained in  $\mathbf{g}_{jn}[\tau_k, \tau_n]$ , then the resulting  $\Sigma_a$  (of the overall diagram) would be reducible, while if  $s$  was the bare line  $\mathbf{g}_{df}^{(0)}[\tau_d, \tau_f]$ , the diagram would contribute to  $\Sigma_{a3}^*$  (see Fig. 22). The box can be either a  $\Sigma_{a3}^*$  insertion or a  $\Sigma_{a4}^*$  insertion, but cannot be a  $\Sigma_{a1}^*$  insertion or a  $\Sigma_{a2}^*$  insertion, since in this case the diagram would contribute to  $\Sigma_{a2}^*$  (see Fig. 21). The analytical contribution of Fig. 23 is

$$-\lambda \mathbf{g}_{\sigma_1 \sigma_b}^{(0)}[i, \mathbf{k}] t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \mathbf{g}_{\sigma_a \sigma_c}[\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}}^{\bar{\sigma}_b \bar{\sigma}_a}} \left( \Sigma_{\sigma_c \bar{\sigma}_d}^*[\mathbf{n}, \mathbf{d}] + \Sigma_{\sigma_c \bar{\sigma}_d}^*[\mathbf{n}, \mathbf{d}] \right) \mathbf{g}_{\bar{\sigma}_d \sigma_2}^{(0)}[\mathbf{d}, f]. \quad (37)$$

Dropping the external lines, and using the inductive hypothesis, we obtain

$$\Sigma_{\sigma_1 \sigma_2}^*[\mathbf{n}, f] = -\lambda^2 t[i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a \sigma_b}[\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_1 \bar{\sigma}_a}} \chi_{\sigma_b \sigma_2}[\mathbf{n}, f]. \quad (38)$$

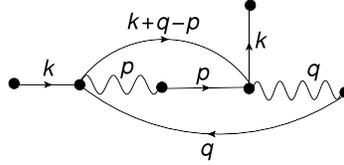
Combining this with Eq. (36) and comparing with Eq. (34), we find that

$$\Sigma_{\sigma_1 \sigma_2}^*[\mathbf{n}, f] + \Sigma_{\sigma_1 \sigma_2}^*[\mathbf{n}, f] = \lambda \chi_{\sigma_1 \sigma_2}[\mathbf{n}, f]. \quad (39)$$

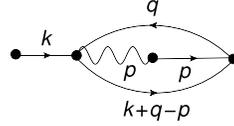
Using Eqs. (39), (30), and (32), we have proven the second of Eqs. (20).

### 5.3. Diagrams in momentum space

Upon turning off the sources, all objects become translationally-invariant in both space and time. We define the Fourier transform of all objects with two external points (e.g.  $\mathcal{G}[i, f]$ ), denoted below



**Fig. 24.** Momentum space representation of diagram for  $\mathbf{g}(k)$  from Fig. 17(b). Upon removing the incoming and outgoing external lines, it contributes to  $\chi(k)$ .



**Fig. 25.** Momentum space representation of diagram for  $\mathcal{G}(k)$  from Fig. 17(l). Upon removing the incoming external line, it contributes to  $\Psi(k)$ .

by the generic symbol  $Q[i, f]$ , as

$$Q[i, f] = \frac{1}{N_s \beta} \sum_k e^{ik(i-f)} Q(k), \quad (40)$$

where  $N_s$  is the number of sites on the lattice,  $\beta$  is the inverse temperature,  $k \equiv (\vec{k}, i\omega_k)$ , and  $k(i-f) \equiv \vec{k} \cdot (\vec{R}_i - \vec{R}_f) - \omega_k(\tau_i - \tau_f)$ . For the rest of the paper, we shall not write the explicit factor  $\frac{1}{N_s \beta}$  that goes along with each momentum sum. To obtain the momentum space contribution of a given  $\mathbf{g}(k)$  diagram, we assign momentum  $k$  to the outgoing and incoming external lines, and sum over the momenta of the internal lines, in such a way that momentum is conserved at each point in the diagram. We also associate with each Green's function line the factor  $\mathbf{g}^{(0)}(q)$ , where  $q$  is the momentum label of that line, and with each interaction line the factor  $-\epsilon_q$ , where  $q$  is the momentum label of that interaction line, and  $t[i, f] \equiv -\sum_q e^{iq(i-f)} \epsilon_q$ . The other rules are the same as in the coordinate space evaluation. For example, consider the diagram in panel (b) of Fig. 17, whose momentum space labels are displayed in Fig. 24. The momentum space contribution of this diagram is

$$-2 \mathbf{g}^{(0)}(k) \epsilon_p \mathbf{g}^{(0)}(p) \epsilon_q \mathbf{g}^{(0)}(q) \mathbf{g}^{(0)}(k+q-p) \mathbf{g}^{(0)}(k) \quad (41)$$

where a sum over the internal momenta  $p$  and  $q$  is implied. Upon removing the external lines, we obtain the following contribution to  $\Sigma_a^*(k)$ , or equivalently to  $\chi(k)$ :

$$-2 \epsilon_p \mathbf{g}^{(0)}(p) \epsilon_q \mathbf{g}^{(0)}(q) \mathbf{g}^{(0)}(k+q-p). \quad (42)$$

Additionally, consider the diagram for  $\mathcal{G}(k)$  displayed in panel (l) of Fig. 17, whose momentum space labels are displayed in Fig. 25. The incoming external line carries momentum  $k$  into the diagram, while the terminal point absorbs this momentum without transferring it to an outgoing external line. The momentum space contribution of this diagram is

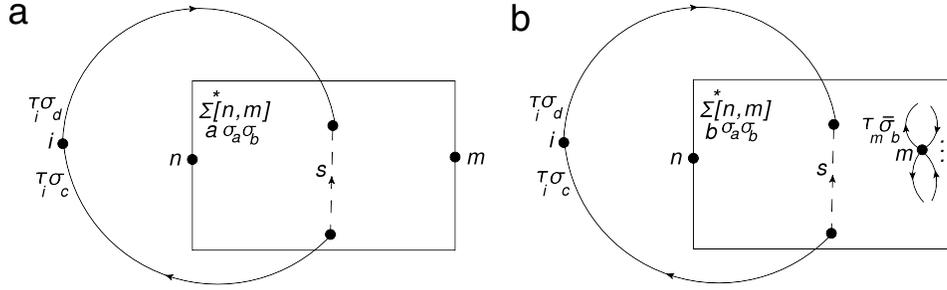
$$-2 \mathbf{g}^{(0)}(k) \epsilon_p \mathbf{g}^{(0)}(p) \mathbf{g}^{(0)}(q) \mathbf{g}^{(0)}(k+q-p) \quad (43)$$

Upon removing the incoming external line, we obtain the following contribution to  $\Sigma_b^*(k)$ , or equivalently to  $\Psi(k)$ :

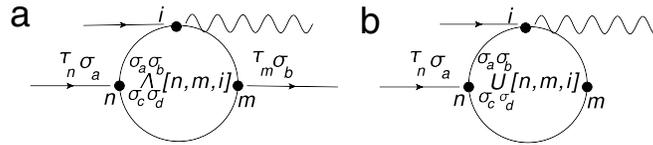
$$-2 \epsilon_p \mathbf{g}^{(0)}(p) \mathbf{g}^{(0)}(q) \mathbf{g}^{(0)}(k+q-p). \quad (44)$$

#### 5.4. The vertices $\Lambda$ and $\mathcal{U}$

In Section 5.2, we showed that our diagrammatic series is consistent with the ECFL EOM, Eqs. (10) and (11). We rewrite them here for convenience.



**Fig. 26.** Schematic diagram for the vertices.  $\Lambda_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$  is displayed in panel (a) while  $\mathcal{U}_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$  is displayed in panel (b).



**Fig. 27.** Schematic diagram for the vertices.  $\Lambda_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$  is displayed in panel (a) while  $\mathcal{U}_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$  is displayed in panel (b).

$$\begin{aligned} \mathbf{g}^{-1}[i, m] &= (\mu - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] + t[i, m] (1 - \lambda \gamma[i]) - \lambda \Phi[i, m], \\ \tilde{\mu}[i, m] &= (1 - \lambda \gamma[i]) \delta[i, m] + \lambda \Psi[i, m] \\ \Phi[i, m] &= -t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \Lambda_*[\mathbf{n}, m; i]; \quad \Psi[i, m] = -t[i, \mathbf{j}] \xi^* \cdot \mathbf{g}[\mathbf{j}, \mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i]. \end{aligned} \quad (45)$$

$$\gamma[i] = \tilde{\mu}^{(k)}[\mathbf{n}, i^+] \cdot \mathbf{g}^{(k)}[i, \mathbf{n}]; \quad \Lambda[n, m; i] = -\frac{\delta}{\delta \mathcal{V}_i} \mathbf{g}^{-1}[n, m]; \quad \mathcal{U}[n, m; i] = \frac{\delta}{\delta \mathcal{V}_i} \tilde{\mu}[n, m]. \quad (46)$$

We now examine the vertices  $\Lambda_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i] \equiv -\frac{\delta}{\delta \mathcal{V}_i} \mathbf{g}_{\sigma_a \sigma_b}^{-1}[n, m]$  and  $\mathcal{U}_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i] \equiv \frac{\delta}{\delta \mathcal{V}_i} \tilde{\mu}_{\sigma_a \sigma_b}[n, m]$  in more detail. The zeroth order vertices, also called the bare vertices, are given by

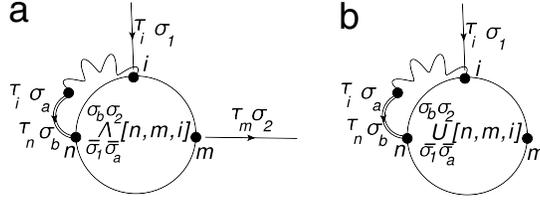
$$\Lambda_{\sigma_c \sigma_d}^{(0) \sigma_a \sigma_b}[n, m; i] = \delta[n, m] \delta[n, i] \delta_{\sigma_a \sigma_c} \delta_{\sigma_b \sigma_d}; \quad \mathcal{U}_{\sigma_c \sigma_d}^{(0) \sigma_a \sigma_b}[n, m; i] = 0. \quad (47)$$

The higher order terms contributing to  $\Lambda_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$  arise from splitting a line in  $\Sigma_{\sigma_a \sigma_b}^* [n, m]$  through the point  $i$ . The higher order terms contributing to  $\mathcal{U}_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$  arise from splitting a line in  $\Sigma_{\sigma_a \sigma_b}^* [n, m]$  through the point  $i$ . These terms can be represented schematically as in Fig. 26.

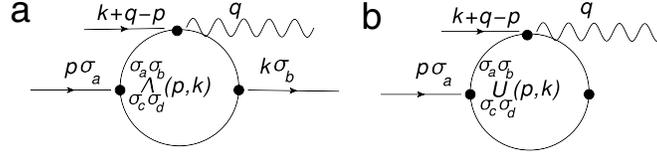
From Fig. 26(a), we see that in  $\Lambda_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$ , the external points  $n$  and  $m$  accommodate an incoming and outgoing external Green's function line, respectively, while the external point  $i$  accommodates an incoming external Green's function line and an external interaction line (Compare with Fig. 13(a)). In Fig. 26(b), we see that in  $\mathcal{U}_{\sigma_c \sigma_d}^{\sigma_a \sigma_b}[n, m; i]$ , the external point  $n$  accommodates an incoming external Green's function line, while the external point  $i$  accommodates an incoming external Green's function line and an external interaction line. However, the external point  $m$  is the terminal point and does not accommodate any external lines. Therefore, the vertices are represented schematically as in Fig. 27. In the case of the bare vertex  $\Lambda^{(0)}[n, m, i]$ , the diagram in Fig. 27(a) collapses onto a single point, which corresponds to the point  $k$  in Fig. 13(a).

In Eq. (45), the self-energies  $\Phi$  and  $\Psi$  are expressed in terms of the vertices  $\Lambda$  and  $\mathcal{U}$  respectively. These relationships can be expressed diagrammatically as in Fig. 28.

We now turn the sources off, so that we can represent the vertices in momentum space, as in Fig. 29. In the case of  $\Lambda(p, k)$ , the external lines carry a total of zero momentum out of the vertex. In the case of  $\mathcal{U}(p, k)$ , the terminal point (the one with no external lines coming in or out) absorbs momentum  $k$ , and therefore the remainder of the external lines have to bring momentum  $k$  into the



**Fig. 28.** Schematic diagram for the self-energies in terms of the vertices. In panel (a)  $\Phi_{\sigma_1\sigma_2}[i, m]$  is expressed in terms of  $\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[n, m; i]$  and in panel (b)  $\Psi_{\sigma_1\sigma_2}[i, m]$  is expressed in terms of  $\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[n, m; i]$ .



**Fig. 29.** Schematic diagram for the vertices in momentum space.  $\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}(p, k)$  is displayed in panel (a) while  $\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}(p, k)$  is displayed in panel (b).

vertex. Therefore, comparing Fig. 29 with Fig. 27, the Fourier transform of the three point vertices, denoted below by the generic symbol  $Q[n, m, i]$ , is:

$$Q[n, m, i] = \sum_{kp} e^{ipn} e^{-ikm} e^{i(k-p)i} Q(p, k) = \sum_{kp} e^{ip(n-i)} e^{ik(i-m)} Q(p, k). \quad (48)$$

Furthermore, there are only four non-zero spin configurations contributing to the vertex. These are  $Q^{(1)} \equiv Q_{\sigma\sigma\sigma}$ ,  $Q^{(2)} \equiv Q_{\sigma\sigma\bar{\sigma}}$ ,  $Q^{(3)} \equiv Q_{\sigma\bar{\sigma}\bar{\sigma}}$ , and  $Q^{(4)} \equiv Q_{\bar{\sigma}\bar{\sigma}\sigma}$ . These four spin configurations are related by the equation

$$Q^{(1)} - Q^{(2)} = Q^{(3)} + Q^{(4)}. \quad (49)$$

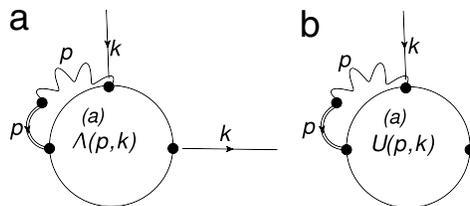
We shall now state the rules for computing the  $Q^{(i)}$  and derive Eq. (49). Recall that to obtain a diagram for  $\Lambda$  ( $\mathcal{U}$ ), we must split a line in the self-energy  $\Sigma_a^*$  ( $\Sigma_b^*$ ). This will give us an extra Green's function line in the diagram, and we must assign momenta to the external lines as indicated in Fig. 29, at the same time summing over the momenta of the internal lines in such a way as to conserve momentum at each point of the diagram. Also recall from Section 4.2 that the Green's function lines in the diagrams for  $\Sigma_a^*$  ( $\Sigma_b^*$ ) are partitioned into anywhere between 0 and  $F_s$  spin loops, where the zeroth loop contains the lines with the labels  $\sigma_1$  and  $\sigma_2$ . The spins carried by the Green's function lines in a single loop are allowed to alternate. However, the spin carried by each Green's function line in the loop is determined by that of any one of them (in the case of the zeroth loop it is the fixed spin  $\sigma_1$ ).

Now, in the case that the line split in going from  $\Sigma^* \rightarrow Q$  is from a loop which is not the zeroth loop, the resulting vertex diagram contributes only to  $Q^{(1)}$  and  $Q^{(2)}$  with a factor of  $\frac{1}{2}$  relative to the contribution of the original diagram to  $\Sigma^*$ . In the case that the line split in going from  $\Sigma^* \rightarrow Q$  is from the zeroth loop, the line split could either carry spin  $\sigma_1$  in the original  $\Sigma^*$  diagram or spin  $\bar{\sigma}_1$ . In the case of the former, the resulting vertex diagram contributes to both  $Q^{(1)}$  and  $Q^{(3)}$  with a factor of 1 relative to the contribution of the original diagram to  $\Sigma^*$ . In the case of the latter, the resulting vertex diagram contributes to  $Q^{(2)}$  with a factor of 1, and to  $Q^{(4)}$  with a factor of  $(-1)$ , relative to the contribution of the original diagram to  $\Sigma^*$ . Eq. (49) immediately follows. Note that in the Feynman diagrams, we have the simpler situation in which all of the Green's function lines in a single spin loop (also referred to as Fermi loop), carry the same spin [18]. Then, the very last case described above becomes impossible,  $Q^{(4)} \rightarrow 0$ , and Eq. (49) reduces to the standard Nozières relation  $Q^{(1)} - Q^{(2)} = Q^{(3)}$  [19].

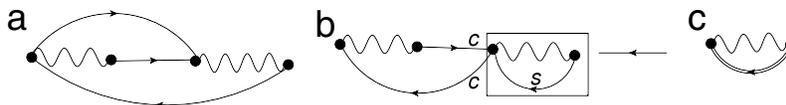
Following Ref. [9], we define  $Q^{(a)} \equiv Q^{(2)} - Q^{(3)}$ . Fourier transforming Eq. (45), we obtain:

$$\Phi(k) = \sum_p \epsilon_p \mathbf{g}(p) \Lambda^{(a)}(p, k); \quad \Psi(k) = \sum_p \epsilon_p \mathbf{g}(p) \mathcal{U}^{(a)}(p, k). \quad (50)$$

These relations are represented diagrammatically in Fig. 30.



**Fig. 30.** Schematic diagram for the self-energies in terms of the vertices. In panel (a)  $\Phi(k)$  is expressed in terms of  $\Lambda^{(a)}(p, k)$  and in panel (b)  $\Psi(k)$  is expressed in terms of  $\mathcal{U}^{(a)}(p, k)$ .



**Fig. 31.** Examples of skeleton and non-skeleton diagrams for the irreducible self-energy  $\Sigma_a^*$ . The diagram in panel (a) is a skeleton diagram. The diagram in panel (b) is not, since we can isolate a self-energy insertion by cutting the two lines labeled by  $c$ . The non-skeleton diagram in panel (b) can be obtained from the skeleton diagram in panel (c) by inserting the self-energy insertion enclosed by the box into the Green's function line. However, by splitting the line labeled by  $s$  in the diagram in panel (b), through another external point, we obtain a skeleton diagram for the vertex  $\Lambda$ .

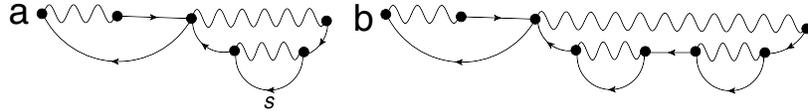
### 5.5. Skeleton diagrams

Consider the diagrammatic expansion for the irreducible self-energies that we have been using thus far, in which each diagram is composed of bare Green's function lines  $\mathbf{g}^{(0)}[i, f]$ , and hopping matrix elements  $t_{if}$ . We aim to reorganize this expansion in such a way that we only keep a subset of these diagrams, in which we replace each bare Green's function line  $\mathbf{g}^{(0)}[i, f]$ , by the full auxiliary Green's function  $\mathbf{g}[i, f]$ , thereby accounting for the diagrams which we discarded. We shall now define this subset of diagrams, which is referred to as the skeleton diagrams.

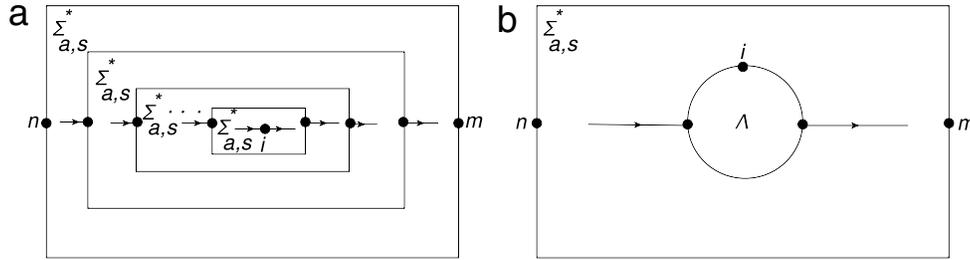
The skeleton diagrams are those diagrams in which one cannot separate a self-energy insertion  $\Sigma_a$  from the rest of the diagram by cutting two Green's function lines. For example, consider the  $\Sigma_a^*$  diagrams in Fig. 31 (the same considerations will apply to  $\Sigma_b^*$  diagrams). From left to right, these are the irreducible self-energies corresponding to the  $\mathbf{g}$  diagrams in Fig. 17(b), Fig. 17(c), and Fig. 12(c). We see that the  $\Sigma_a^*$  diagram in panel (b) of Fig. 31 is a non-skeleton diagram, since by cutting the two Green's function lines labeled by the letter  $c$ , we isolate the  $\Sigma_a$  self-energy insertion enclosed in the box. In contrast, the  $\Sigma_a^*$  diagram in panel (a) of Fig. 31 is a skeleton diagram, since it is impossible to isolate a  $\Sigma_a$  insertion by cutting two Green's function lines. Finally, the diagram in panel (c) of Fig. 31 is also a skeleton diagram. Furthermore, we see that by placing the self-energy insertion enclosed in the box into the Green's function line of the diagram in Fig. 31(c), we reproduce the diagram in Fig. 31(b). Since a full auxiliary Green's function line consists of an arbitrary self-energy insertion  $\Sigma_a$  surrounded by two bare Green's function lines  $\mathbf{g}^{(0)}$ , we see that the whole series is reproduced by keeping only the skeleton diagrams and making the substitution  $\mathbf{g}^{(0)}[i, f] \rightarrow \mathbf{g}[i, f]$ .

Now, consider the vertices  $\Lambda[n, m; i]$  and  $\mathcal{U}[n, m; i]$ . Recall from Fig. 26, that these correspond to splitting a Green's function line through the point  $i$  in  $\Sigma_a^*[n, m]$  and  $\Sigma_b^*[n, m]$ , respectively. How do we obtain the skeleton diagrams for the vertices? A naive guess would be that we do so by splitting a Green's function line in the skeleton diagrams for the irreducible self-energies. However, this is only partially correct. To see this, consider again the non-skeleton  $\Sigma_a^*$  diagram in panel (b) of Fig. 31. If we choose to split either of the two lines labeled by  $c$ , then we leave the self-energy insertion surrounded by the box intact, and the resulting diagram for  $\Lambda$  is a non-skeleton diagram. However, if we split the Green's function line labeled by  $s$ , this breaks up this self-energy insertion, and leads to a skeleton diagram for  $\Lambda$ .

Taking this reasoning a step further, consider the diagram for  $\Sigma_a^*$  in panel (a) of Fig. 32. This diagram can be obtained from the diagram in Fig. 31(b) by inserting the self-energy insertion enclosed by the box into the line labeled by  $s$  in Fig. 31(b). Once again, if we split any line other than the one labeled by  $s$  in Fig. 32(a), the resulting diagram for  $\Lambda$  will be a non-skeleton diagram, while if we split the line



**Fig. 32.** In this case both  $\Sigma_a^*$  diagrams displayed in panels (a) and (b) are non-skeleton diagrams. However, the diagram in panel (a) contains only irreducible self-energy insertions, while the one in panel (b) contains a reducible self-energy insertion. One can obtain a skeleton diagram for the vertex  $\Lambda$  only by splitting the line labeled by  $s$  in the diagram in panel (a). It is impossible to obtain a skeleton diagram for the vertex  $\Lambda$  from the diagram in panel (b) regardless of which line we split.



**Fig. 33.** Panel (a) demonstrates the general procedure for obtaining a skeleton diagram for the vertex  $\Lambda$  from a  $\Sigma_a^*$  diagram consisting of a sequence of  $\Sigma_a^*$  skeleton diagrams. The original  $\Sigma_a^*$  diagram is itself a skeleton diagram only if there is only one skeleton diagram in the sequence, i.e. the  $\Sigma_a^*$  diagram in question. If we remove the outermost box in panel (a), we are still left with a general skeleton diagram for the vertex  $\Lambda$ . Therefore, to obtain a skeleton diagram for  $\Lambda$ , one must insert the full vertex  $\Lambda$  into a green's function line of a skeleton diagram for  $\Sigma_a^*$ . This is displayed in panel (b). In the case that the original  $\Sigma_a^*$  diagram in panel (a) is itself a skeleton diagram (i.e. there is only one skeleton diagram in the sequence), the  $\Lambda$  vertex in panel (b) is a bare vertex.

labeled by  $s$ , the resulting diagram for  $\Lambda$  will be a skeleton diagram. Meanwhile, for the  $\Sigma_a^*$  diagram in Fig. 32(b), obtained from the diagram in Fig. 31(b) by putting a reducible self-energy insertion into the line labeled by  $s$  in Fig. 31(b), it is not possible to split any line in such a way that the resulting diagram for  $\Lambda$  will be a skeleton diagram.

Therefore, we see that to construct the skeleton diagrams for  $\Lambda[n, m; i]$  ( $\mathcal{U}[n, m; i]$ ), we have to use the following procedure. Take a skeleton diagram for  $\Sigma_a^*[n, m]$  ( $\Sigma_b^*[n, m]$ ), and insert into at most one line of this diagram, a skeleton diagram for  $\Sigma_a^*$ . Then, insert into at most line of that diagram, a skeleton diagram for  $\Sigma_a^*$ , and so on. This produces a sequence of skeleton diagrams for the irreducible self-energies. Then, in the last skeleton diagram of the sequence, split a single Green's function line through the point  $i$ . This procedure is represented schematically (for the case of  $\Lambda[n, m; i]$ ) in Fig. 33(a).

Now consider the part of Fig. 33(a) enclosed by the second box (counting from the very outer box). This is itself a skeleton diagram for the vertex  $\Lambda[w, v; i]$ , where  $w$  and  $v$  are internal variables. Therefore, we see that one can obtain the skeleton expansion for  $\Lambda[n, m; i]$  ( $\mathcal{U}[n, m; i]$ ) from the skeleton expansion for  $\Sigma_a^*[n, m]$  ( $\Sigma_b^*[n, m]$ ) by replacing in each skeleton diagram for  $\Sigma_a^*[n, m]$  ( $\Sigma_b^*[n, m]$ ), a single Green's function line  $\mathbf{g}[x, y]$ , with  $\mathbf{g}[x, \mathbf{w}].\Lambda[\mathbf{w}, \mathbf{v}; i].\mathbf{g}[\mathbf{v}, y]$ , where  $\Lambda[\mathbf{w}, \mathbf{v}; i]$  is the full vertex. This is represented schematically in Fig. 33(b). The case in which there is only one box in Fig. 33(a) corresponds to plugging in the bare vertex into Fig. 33(b).

We now have three skeleton expansions. The first is the original skeleton expansion of the self-energies in terms of the auxiliary Green's function.

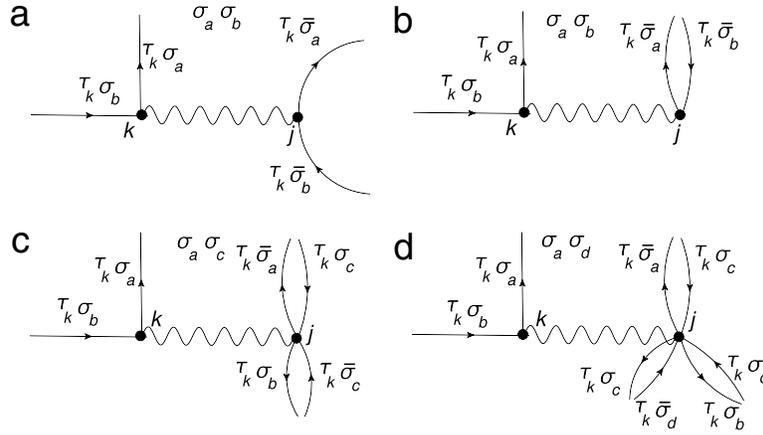
$$\mathbf{g}^{-1} \equiv \mathbf{g}^{-1}[\mathbf{g}]; \quad \tilde{\mu} \equiv \tilde{\mu}[\mathbf{g}], \quad (51)$$

The second is the skeleton expansion for the vertices in terms of the auxiliary Green's function. This is the skeleton expansion represented in Fig. 33(a).

$$\Lambda \equiv \Lambda[\mathbf{g}]; \quad \mathcal{U} \equiv \mathcal{U}[\mathbf{g}]. \quad (52)$$

The third is the skeleton expansion for the vertices in terms of the auxiliary Green's function and the full vertex  $\Lambda$ . This is the skeleton expansion represented in Fig. 33(b).

$$\Lambda \equiv \Lambda[\mathbf{g}, \Lambda]; \quad \mathcal{U} \equiv \mathcal{U}[\mathbf{g}, \Lambda]. \quad (53)$$



**Fig. 34.** The  $J$ -vertices in the diagrams of the  $\lambda$  expansion. They are more reminiscent of the Feynman diagram vertices than the  $t$ -vertices displayed in Fig. 13. The two types of vertices can be obtained from each other by interchanging lines between the two points of the vertex.

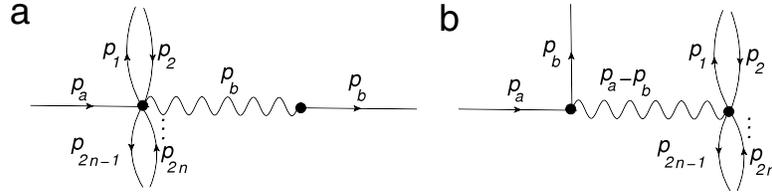
Using the diagrammatic rules developed here, we have access to all three of these skeleton expansions at any order. However, in the absence of these rules, we could derive the terms in these skeleton expansions by using Eqs. (51)–(53), and (45) in the following manner. Suppose that we have the skeleton expansions in Eq. (51)–(53) through  $m$ th order in  $\lambda$ . Then, plugging the  $m$ th order term of the skeleton expansion from Eq. (52) into Eq. (45) yields the  $m + 1$ st order term of the skeleton expansion in Eq. (51). Then, applying the rule  $\mathbf{g} \rightarrow \mathbf{g}\Lambda\mathbf{g}$  to the  $m + 1$ st order term of the skeleton expansion in Eq. (51), yields the  $m + 1$ st order contribution to the skeleton expansion in Eq. (53). Finally, plugging the  $k$ th order term of the skeleton expansion from Eq. (52) ( $0 \leq k \leq m$ ) into the  $m + 1 - k$ th term of the skeleton expansion from Eq. (53) yields the  $m + 1$ st order term of the skeleton expansion from Eq. (52), after which we can iterate the process again. This process starts at zeroth order by plugging the bare vertex into Eq. (45) and calculating the first order contribution to the skeleton expansion in Eq. (51), and so on. This is the approach used in the original ECFL papers [9, 10], and reviewed in Section 2. It reveals the power of the Schwinger approach in that it enables one to bypass the bare series and work directly with the skeleton expansion. However, the utility of the diagrams developed here is that they enable one to obtain the contribution of a given order directly, without iteration, and also to visualize all the higher order terms diagrammatically, therefore facilitating diagrammatic re-summations.

## 6. Putting $J$ back into the equations

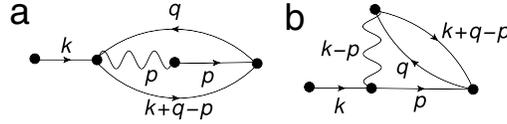
Let us rewrite Eq. (4) in the form of an integral equation as in Eq. (17), but this time keeping  $J$ .

$$\begin{aligned} \mathcal{G}_{\sigma_1, \sigma_2}[i, f] &= \mathbf{g}_{\sigma_1, \sigma_2}^{(0)}[i, f] - \lambda \mathbf{g}_{\sigma_1, \sigma_b}^{(0)}[i, f] \sigma_b \sigma_2 \mathcal{G}_{\bar{\sigma}_2, \bar{\sigma}_b}[f, f^+] \\ &\quad - \lambda \mathbf{g}_{\sigma_1, \sigma_b}^{(0)}[i, \mathbf{k}] \left( -t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \mathcal{G}_{\bar{\sigma}_a, \bar{\sigma}_b}[\mathbf{k}, \mathbf{k}^+] \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] + t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \frac{\delta}{\delta \mathcal{V}_{\mathbf{k}}^{\bar{\sigma}_b, \bar{\sigma}_a}} \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{j}, f] \right), \\ &\quad - \lambda \mathbf{g}_{\sigma_1, \sigma_b}^{(0)}[i, \mathbf{k}] \left( \frac{1}{2} J[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \mathcal{G}_{\bar{\sigma}_a, \bar{\sigma}_b}[\mathbf{j}, \mathbf{j}^+] \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{k}, f] - \frac{1}{2} J[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \frac{\delta}{\delta \mathcal{V}_{\mathbf{j}}^{\bar{\sigma}_b, \bar{\sigma}_a}} \mathcal{G}_{\sigma_a, \sigma_2}[\mathbf{k}, f] \right). \end{aligned} \quad (54)$$

The  $\lambda$ -expansion of Eq. (54) is given by the same set of rules as in Section 4.2, with the only difference being that now each vertex can be either a  $t$ -vertex or a  $J$ -vertex. Comparing the second and third lines on the RHS of Eq. (54), we see that the  $J$ -vertices can be obtained from the  $t$ -vertices in Fig. 13 by moving the line labeled by  $\sigma_a$  from the point  $j$  to the point  $k$ , and moving all lines but the one labeled by  $\sigma_b$  from the point  $k$  to the point  $j$ . The  $J$ -vertices are displayed in Fig. 34. They are more reminiscent of the standard Feynman diagram vertices.



**Fig. 35.** The  $t$  and  $J$  vertices are displayed in momentum space in panels (a) and (b) respectively. For each interaction vertex of a diagram, we can choose to use either one as long as we associate with it the factor  $\epsilon_{p_b} + \frac{1}{2}J_{p_a-p_b}$ .



**Fig. 36.** The  $g(k)$  diagrams drawn in panel (a) and panel (b) correspond to the same diagram. The one in panel (a) is drawn using a  $t$ -vertex, while the one in panel (b) is drawn using a  $J$ -vertex. In both cases, we associate the factor  $\epsilon_p + \frac{1}{2}J_{k-p}$  with the interaction vertex.

Now, let us compare an arbitrary  $t$ -vertex and an arbitrary  $J$ -vertex in momentum space. The  $t$ -vertex is shown in panel (a) of Fig. 35, while the  $J$ -vertex is shown in panel (b). Conserving momentum at each point of the  $t$  and  $J$  vertices yields the relation

$$p_a + \sum_{m=1}^n p_{2m} = p_b + \sum_{m=1}^n p_{2m-1}. \quad (55)$$

In Fig. 35(a), the interaction line contributes a factor of  $\epsilon_{p_b}$ , while in Fig. 35(b), the interaction line contributes a factor of  $\frac{1}{2}J_{p_a-p_b}$ . At first, it seems as though for each diagram with  $i$  interaction vertices, we must now draw  $2^i$  separate diagrams, since for each vertex we must decide whether it will be a  $t$ -vertex or a  $J$ -vertex. For example, consider the diagram in Fig. 25, also displayed in Fig. 36(a), in which the interaction vertex is a  $t$ -vertex. In Fig. 36(b), it is drawn with a  $J$ -vertex. However, we see that the Green's function lines in both diagrams have the same momentum labels, and the only difference is the momentum label of the interaction line. This is because the two things which determine the momentum labels of the Green's function lines are

- (1) the interconnections (via Green's function lines) between the interaction vertices (irrespective of where on these vertices these lines appear),
- (2) Eq. (55).

Since the  $J$ -vertex simply reshuffles the lines on the  $t$ -vertex, and Eq. (55) applies equally well to both types of vertices, both (1) and (2) are unaffected by the choice of  $t$  vertex vs.  $J$  vertex. Therefore, we can choose to use either the diagram in Fig. 36(a) or the diagram in Fig. 36(b) if we associate with the interaction line in each diagram the factor  $\epsilon_p + \frac{1}{2}J_{k-p}$ . In general, we can construct diagrams either from the vertices in Fig. 35(a) (as we have already been doing), or from the vertices in Fig. 35(b) (which would be more reminiscent of the Feynman diagrams), as long as we associate with each interaction vertex the factor  $\epsilon_{p_b} + \frac{1}{2}J_{p_a-p_b}$ .

The ECFL equations with  $J$  included are given as follows [9].

$$\begin{aligned} \mathbf{g}^{-1}[i, m] &= (\mu - \partial_{\tau_i} - \mathcal{V}_i) \delta[i, m] + t[i, m] (1 - \lambda\gamma[i]) + \frac{\lambda}{2} J[i, \mathbf{j}] \gamma[\mathbf{j}] \delta[i, m] - \lambda \Phi[i, m], \\ \tilde{\mu}[i, m] &= (1 - \lambda\gamma[i]) \delta[i, m] + \lambda \Psi[i, m], \\ \Phi[i, m] &= \mathbf{L}[i, \mathbf{n}] \cdot \mathbf{g}^{-1}[\mathbf{n}, m]; \quad \Psi[i, m] = -\mathbf{L}[i, \mathbf{n}] \cdot \tilde{\mu}[\mathbf{n}, m], \end{aligned} \quad (56)$$

where the operator  $\mathbf{L}$  is given by:

$$\mathbf{L}_{\sigma_1 \sigma_2}[i, m] = t[i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a \sigma_2}[\mathbf{j}, m] \frac{\delta}{\delta \mathcal{V}_i^{\sigma_1 \sigma_a}} - \frac{1}{2} J[i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a \sigma_2}[i, m] \frac{\delta}{\delta \mathcal{V}_j^{\sigma_1 \sigma_a}}. \quad (57)$$

Using the same decomposition as in Eq. (31), i.e.

$$\Phi[i, m] = -\Psi[i, \mathbf{j}]t[\mathbf{j}, m] + \chi[i, m], \quad (58)$$

we find that

$$\chi[i, m] = \mathbf{L}[i, \mathbf{n}].\mathbf{g}^{-1(0)}[\mathbf{n}, m] + \frac{\lambda}{2}J[m, \mathbf{k}]\mathbf{L}[i, m]\gamma[\mathbf{k}] - \lambda\mathbf{L}[i, \mathbf{n}]\chi[\mathbf{n}, m], \quad (59)$$

where

$$(\mathbf{L}[i, \mathbf{n}].\mathbf{g}^{-1(0)}[\mathbf{n}, m])_{\sigma_1\sigma_2} = t[i, \mathbf{j}]_{\sigma_1\sigma_2}\mathbf{g}_{\tilde{\sigma}_2\tilde{\sigma}_1}[\mathbf{j}, i]\delta[i, m] - \frac{1}{2}J[i, m]_{\sigma_1\sigma_2}\mathbf{g}_{\tilde{\sigma}_2\tilde{\sigma}_1}[i, m]. \quad (60)$$

Finally, we note that the equations associated with Fig. 30 now become

$$\Phi(k) = \sum_p \left( \epsilon_p + \frac{1}{2}J_{k-p} \right) \mathbf{g}(p)\Lambda^{(a)}(p, k); \quad \Psi(k) = \sum_p \left( \epsilon_p + \frac{1}{2}J_{k-p} \right) \mathbf{g}(p)\mathcal{U}^{(a)}(p, k). \quad (61)$$

## 7. Finite order calculations

### 7.1. Zeroth through third order calculation

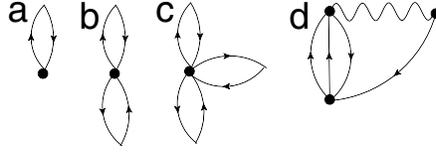
In this section, we compute the skeleton expansion for the objects  $\gamma$ ,  $\Psi$ , and  $\chi$  through second order in  $\lambda$  in momentum space. As can be seen from Eq. (63), this yields the skeleton expansion for  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  through third order in  $\lambda$ . Before proceeding with this computation, we follow Ref. [10] in introducing a second chemical potential  $u_0$  into the theory. As explained in Ref. [10], there is a so-called shift identity of the  $t$ - $J$  model, which states that adding an onsite term to the hopping affects  $\mathcal{G}$  only through a shift of the chemical potential  $\mu$ . However, the same is not true of the constituent factors  $\mathbf{g}$  and  $\tilde{\mu}$ , which will be affected by such a shift. To remedy this, in Ref. [10], the second chemical potential  $u_0$  is introduced directly into the definitions of  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  (Eq. (56)) through the formula  $t[i, j] \rightarrow t[i, j] + \frac{u_0}{2}\delta[i, j]$  in every term but the  $t[i, f]$  term in the equation for  $\mathbf{g}^{-1}[i, f]$ . Now, an onsite shift in the hopping affects  $\mathbf{g}$  and  $\tilde{\mu}$  only through a shift in the second chemical potential  $u_0$ . Moreover, the fact that  $\mathcal{G}$  will not be affected for any value of  $u_0$  (other than through a shift of the original chemical potential  $\mu$ ) is a consequence of the shift identity. Furthermore, the two chemical potentials  $\mu$  and  $u_0$  can now be used to satisfy the two sum rules

$$\sum_k \mathcal{G}(k) = \frac{n}{2}; \quad \sum_k \mathbf{g}(k) = \frac{n}{2}. \quad (62)$$

The first of these ensures the correct particle sum-rule for the physical electrons. The second one states that the auxiliary fermions must satisfy the same particle sum-rule as the physical ones. We can think of the Hubbard operator  $X_i^{0\sigma} = c_{i\sigma}(1 - n_{i\bar{\sigma}})$  as representing the physical fermions, and the canonical operator  $c_{i\sigma}$  as representing the auxiliary fermions. Since, the number operator is a charge neutral object, charge conservation implies that the physical and auxiliary fermions must satisfy the same particle sum-rule. As a consequence of this, the physical electrons have a Fermi-surface which complies with the Luttinger–Ward volume theorem (see Ref. [9] where these sum rules were originally introduced and their implications discussed).

We now proceed to present the diagrams and analytical expressions for  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  through third order in  $\lambda$ . Taking the Fourier transform of Eqs. (9) and (56), and using Eq. (58), we obtain

$$\begin{aligned} \mathbf{g}^{-1}(k) &= i\omega_k + \mu' - \left(\epsilon_k - \frac{u_0}{2}\right)\tilde{\mu}(k) - \lambda\chi(k), \\ \mu' &= \mu - \frac{u_0}{2} + \frac{\lambda}{2}\gamma J_0 \\ \tilde{\mu}(k) &= (1 - \lambda\gamma) + \lambda\Psi(k), \\ \mathcal{G}(k) &= \mathbf{g}(k)\tilde{\mu}(k), \end{aligned} \quad (63)$$



**Fig. 37.** Second order skeleton expansion for  $\gamma$ . Only the diagram in panel (a) is a standard Feynman diagram.  $\gamma^{(0)}$  is given by the diagram in panel (a),  $\gamma^{(1)}$  is given by the diagram in panel (b), and  $\gamma^{(2)}$  is given by diagrams in panels (c) and (d). We conserve momentum at each interaction vertex as indicated in Figs. 35 and 36.

where  $J_0$  is the zero-momentum component of the Fourier transform of  $J_{ij}$ . Our strategy is to compute the skeleton expansion for  $\gamma$ ,  $\Psi$ , and  $\chi$  through second order in  $\lambda$  (i.e.  $\gamma = \gamma^{(0)} + \gamma^{(1)} + \gamma^{(2)}$ , etc.) After plugging in the expressions from this skeleton expansion into Eq. (63), we must set  $\lambda = 1$ , and solve the resulting integral equations. The two Lagrange multipliers  $\mu$  and  $u_0$  are then determined by the sum rules in Eq. (62).

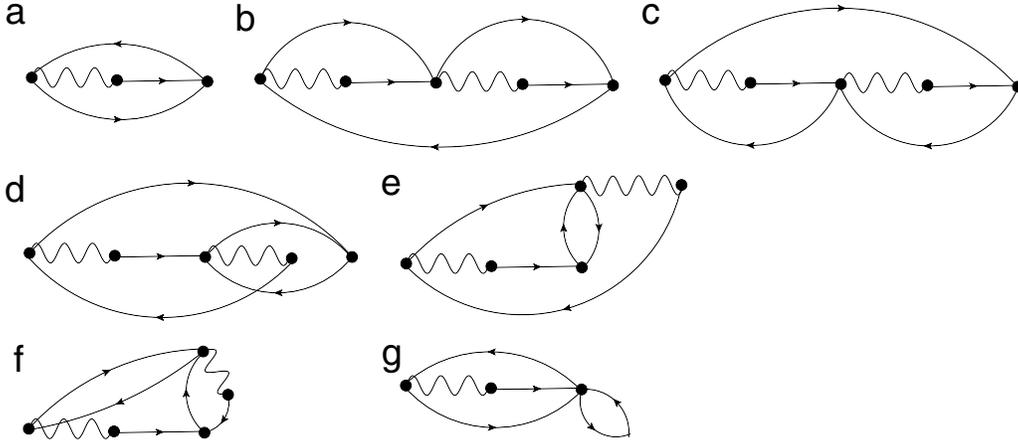
In Fig. 37, we have drawn the skeleton diagrams for  $\gamma$  (which is just a constant when the sources are off) through second order in  $\lambda$ . Therefore,  $\gamma$  is the sum of the following terms

$$\begin{aligned}
 & \text{(a)} \quad \frac{n}{2} \\
 & \text{(b)} \quad -\lambda \left(\frac{n}{2}\right)^2 \\
 & \text{(c)} \quad \lambda^2 \left(\frac{n}{2}\right)^3 \\
 & \text{(d)} \quad -2\lambda^2 \sum_{p,l,q} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(p+l-q) \left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q}\right). \tag{64}
 \end{aligned}$$

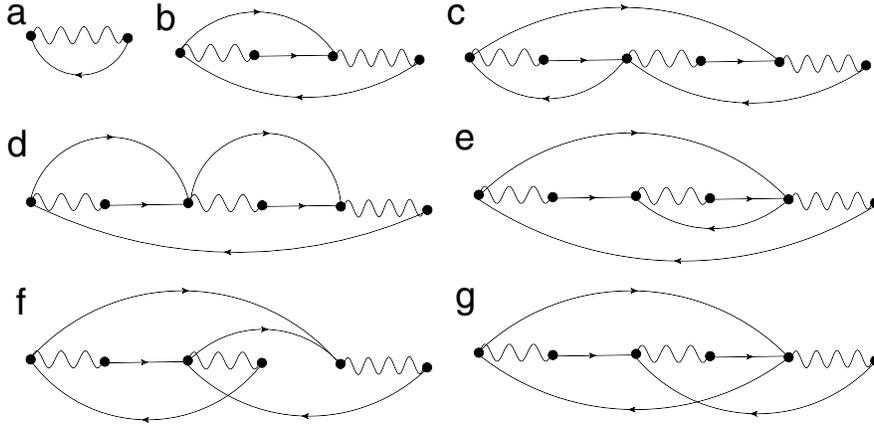
In Fig. 38, we have done the same for  $\Psi(k)$ . Therefore,  $\Psi(k)$  is the sum of the following terms

$$\begin{aligned}
 & \text{(a)} \quad -2\lambda \sum_{pq} \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \\
 & \text{(b)} \quad -4\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+q-l) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l}\right) \\
 & \text{(c)} \quad -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(q+l-p) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l}\right) \\
 & \text{(d)} \quad -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(p+l-q) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q}\right) \\
 & \text{(e)} \quad -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q}\right) \\
 & \text{(f)} \quad -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(l+p-k) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-k}\right) \\
 & \text{(g)} \quad \lambda^2 \frac{n}{2} \sum_{pq} \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p) \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right) \tag{65}
 \end{aligned}$$

The skeleton diagrams for  $\chi(k)$  have been split into two groups. Those drawn in Fig. 40, whose contribution will be denoted by  $\chi_B(k)$ , can be obtained from the  $\Psi(k)$  diagrams in Fig. 38 by attaching an interaction line to the terminal point of those diagrams. Due to the decomposition Eq. (58), this interaction line will contribute only a  $J$  term, but no  $\epsilon$  term, to the expression for  $\chi_B(k)$ . The rest of the  $\chi(k)$  diagrams, whose contribution will be denoted by  $\chi_A(k)$ , are drawn in Fig. 39. Then,



**Fig. 38.** Second order skeleton expansion for  $\Psi(k)$ . All diagrams but the one in panel (g) are standard Feynman diagrams (with one interaction line set to unity).  $\Psi^{(1)}$  is given by the diagram in panel (a), and  $\Psi^{(2)}$  is given by the diagrams in panels (b) through (g). We conserve momentum at each interaction vertex as indicated in Figs. 35 and 36.



**Fig. 39.** Second order skeleton expansion for  $\chi_A(k)$ . These diagrams are independent of those for  $\Psi(k)$ . All diagrams are standard Feynman diagrams. The diagram in panel (a) contributes to  $\chi^{(0)}$ , the diagram in panel (b) contributes to  $\chi^{(1)}$ , and the diagrams in panels (c) through (g) contribute to  $\chi^{(2)}$ . We conserve momentum at each interaction vertex as indicated in Figs. 35 and 36.

$\chi(k) = \chi_A(k) + \chi_B(k)$ , where  $\chi_A(k)$  is the sum of the terms in Eq. (66) and  $\chi_B(k)$  is the sum of the terms in Eq. (67).

$$\begin{aligned}
 (a) & - \sum_p \mathbf{g}(p) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2} J_{k-p} \right) \\
 (b) & - 2\lambda \sum_{pq} \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(k+q-p) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2} J_{k-p} \right) \left( \epsilon_q - \frac{u_0}{2} + \frac{1}{2} J_{p-q} \right) \\
 (c) & - \lambda^2 \sum_{pql} \mathbf{g}(p) \mathbf{g}(l) \mathbf{g}(q) \mathbf{g}(k+q-p) \mathbf{g}(l+q-p) \\
 & \quad \times \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2} J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2} J_{p-l} \right) \left( \epsilon_{l+q-p} - \frac{u_0}{2} + \frac{1}{2} J_{p-q} \right) \\
 (d) & - 4\lambda^2 \sum_{pql} \mathbf{g}(p) \mathbf{g}(l) \mathbf{g}(q) \mathbf{g}(k+q-p) \mathbf{g}(k+q-l)
 \end{aligned}$$

$$\begin{aligned}
& \times \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l} \right) \left( \epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q} \right) \\
(e) & - \lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p) \\
& \times \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l} \right) \left( \epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q} \right) \\
(f) & - \lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(p+l-q) \\
& \times \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{k+q-p-l} \right) \\
(g) & - \lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p) \\
& \times \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l} \right) \left( \epsilon_{k+l-p} - \frac{u_0}{2} + \frac{1}{2}J_{p-k} \right) \tag{66}
\end{aligned}$$

$$\begin{aligned}
(a) & - \lambda \sum_{pq} \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) J_{p-k} \\
(b) & - 2\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+q-l) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l} \right) J_{l-k} \\
(c) & - \frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(q+l-p) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l} \right) J_{l-k} \\
(d) & - \frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(p+l-q) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q} \right) J_{q-p} \\
(e) & - \frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q} \right) J_{p-k} \\
(f) & - \frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(l+p-k) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) \left( \epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-k} \right) J_{p-k} \\
(g) & \lambda^2 \frac{n}{4} \sum_{pq} \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p) \left( \epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p} \right) J_{p-k} \tag{67}
\end{aligned}$$

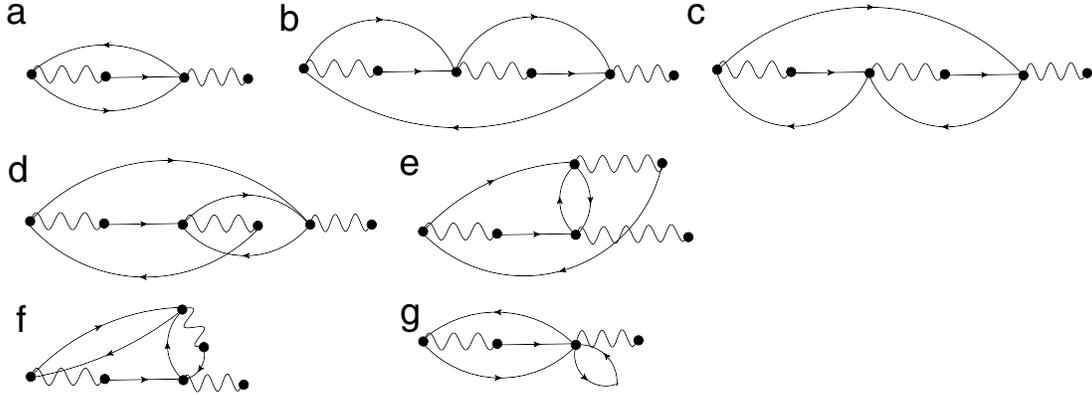
## 7.2. High frequency limit

We know from the anti-commutation relations for the Hubbard  $X$  operators, that the high frequency limit of the Green's function is  $\lim_{i\omega_k \rightarrow \infty} \mathcal{G}(k) = \frac{1-\frac{n}{2}}{i\omega_k}$ . From Eq. (63), we see that the high frequency limit of the Green's function can also be expressed as  $\lim_{i\omega_k \rightarrow \infty} \mathcal{G}(k) = \frac{1-\lambda\gamma}{i\omega_k}$ . Since  $\gamma = \sum_k \mathcal{G}(k) = \frac{n}{2}$  in the exact theory, after setting  $\lambda = 1$  the two expressions for the high frequency limit are equivalent.

From Eq. (63), we see that to obtain  $\mathbf{g}^{-1}(k)$  and  $\tilde{\mu}(k)$  to  $m$ th order in  $\lambda$ , we must calculate  $\gamma$ ,  $\Psi(k)$ , and  $\chi(k)$  to order  $m-1$ . If we are doing this using the bare expansion, then in order to satisfy the sum rules in Eq. (62) order by order, we must also expand the two chemical potentials  $\mu$  and  $u_0$  in  $\lambda$  [25].

$$\mu = \mu^{(0)} + \mu^{(1)} + \dots \quad u_0 = u_0^{(0)} + u_0^{(1)} + \dots, \tag{68}$$

where  $\mu^{(0)}$  is zeroth order in  $\lambda$ ,  $\mu^{(1)}$  is first order in  $\lambda$ , etc. Denoting  $\mathbf{g}$ ,  $\tilde{\mu}$ ,  $\gamma$ ,  $\Psi$ , and  $\chi$  by the generic symbol  $Q$ , and plugging the expansions from Eq. (68) into the bare expansion for  $Q^{(m)} = Q^{(m)}(\mu, u_0)$ ,



**Fig. 40.** Second order skeleton expansion for  $\chi_B(k)$ , which vanishes when  $J = 0$ . These diagrams can be obtained from those for  $\Psi(k)$  in Fig. 38 by adding an interaction line to the terminal point of those diagrams. However, this interaction line contributes only a factor of  $J$ , and not a factor of  $\epsilon$ . All diagrams but the one in panel (g) are standard Feynman diagrams. The diagram in panel (a) contributes to  $\chi^{(1)}$ , and the diagrams in panels (b) through (g) contribute to  $\chi^{(2)}$ . We conserve momentum at each interaction vertex as indicated in Figs. 35 and 36.

the latter is rearranged with the various orders being mixed due to the expansion of the chemical potentials. Then, we can solve for the various quantities  $\mu^{(0)}$ ,  $\mu^{(1)}$ , etc. such that in the rearranged series for  $\gamma^{(m)} = \gamma^{(m)}(n)$  and  $\mathbf{g}^{(m)} = \mathbf{g}^{(m)}(n)$ ,

$$\gamma^{(m)} = \delta_{m,0} \frac{n}{2}; \quad \sum_k \mathbf{g}^{(m)}(k) = \delta_{m,0} \frac{n}{2}. \quad (69)$$

Then, substituting the expression for  $\gamma^{(m)}$  back into Eq. (63), we see that only  $\mathcal{G}^{(0)}(k)$  and  $\mathcal{G}^{(1)}(k)$  contribute to the high-frequency limit of the Green's function, and that  $\lim_{i\omega_k \rightarrow \infty} \mathcal{G}(k) = \frac{1-\lambda\gamma}{i\omega_k} = \frac{1-\lambda\frac{n}{2}}{i\omega_k}$ .

In the skeleton expansion, the situation is different. In this case, after we set  $\lambda = 1$ , the diagrams from all orders in the skeleton expansion are mixed together on equal footing to generate one integral equation which together with the sum rules in Eq. (62) determines  $\mathbf{g}$ ,  $\mu$ , and  $u_0$ . The other objects are then obtained from these. In this case, if the skeleton expansions for  $\gamma$ ,  $\Psi(k)$ , and  $\chi(k)$  have been carried out to  $m - 1$ st order before being plugged into Eq. (63), then the sum rule Eq. (62) implies that (after setting  $\lambda = 1$ )  $\sum_{l=0}^m \gamma^{(l)} = \frac{n}{2}$ . However, from Eq. (63), the high frequency limit is given by  $\lim_{i\omega_k \rightarrow \infty} \mathcal{G}(k) = \frac{1-\sum_{l=0}^{m-1} \gamma^{(l)}}{i\omega_k}$ . Therefore, the error in the high frequency limit is equal to  $\gamma^{(m)}$ , and we have that

$$\lim_{i\omega_k \rightarrow \infty} \mathcal{G}(k) = \frac{1 - \frac{n}{2} + \gamma^{(m)}}{i\omega_k}. \quad (70)$$

This error vanishes as  $m \rightarrow \infty$ .

### 7.3. Analysis of the $\lambda$ expansion: Feynman type diagrams and non-Feynman diagrams.

The  $\lambda$  series for  $\mathcal{G}$  differs from the Feynman series for  $\mathcal{G}$  in two fundamental ways. The first is the presence of the term  $-\lambda\gamma + \lambda\Psi(k)$  in the numerator of  $\mathcal{G}(k)$ . In the Feynman series, this term is absent. To discuss the second one, let us identify  $\lambda\gamma$  with the Hartree term in the Feynman diagrams, and  $\lambda\Phi$  with all self-energy diagrams other than the Hartree term.  $\Psi$  forms a subset of  $\Phi$  (except for a missing interaction line which is not important for the present discussion), and hence all considerations which apply to  $\Phi$  will apply equally well to  $\Psi$ . Hence, the second important difference is that there are diagrams which contribute to  $\lambda\gamma$  which do not contribute the Hartree term of the Feynman series, and there are diagrams that contribute to  $\lambda\Phi$  which do not contribute to the other self-energy diagrams of the Feynman series.

From Fig. 37, we can see that the first order  $\lambda\gamma$  diagram is exactly the Hartree term of the Feynman series, while the others are all diagrams which do not contribute to the Hartree term of the Feynman series. However, from Figs. 39 and 40, we can see that the only diagram in the 3rd order skeleton expansion for  $\lambda\Phi$  which is not a Feynman diagram, is diagram (g) in Fig. 40 (Feynman diagrams are the same order in  $\lambda$  as they are in the interaction, while non-Feynman diagrams are not). Therefore, the deviation of  $\lambda\Phi$  and  $\lambda\Psi$  from the Feynman series grows rather slowly as compared with the growth of the series itself. Moreover, if we consider the fact that the infinite series for  $\gamma$  must sum to  $\frac{\rho}{2}$ , we see that to “leading order”, the only difference between the  $\lambda$  series and the Feynman series is the presence of the term  $-\lambda\gamma + \lambda\Psi(k)$  in the numerator of  $\mathcal{G}(k)$ . This leads us to the point of view taken in the phenomenological ECFL [9,26,17,14], in which  $\gamma \rightarrow \frac{\rho}{2}$ , and the self-energies  $\Psi(k)$  and  $\Phi(k)$  are given simple Fermi-liquid forms. Then, the main correction to Fermi-liquid behavior is not seen as coming from the self-energies themselves, but from the interplay between the numerator and denominator of the single-particle Green’s function.

## 8. Connection with Zaitsev–Izyumov formalism

The Zaitsev–Izyumov formalism [8,7] is a technique for doing an expansion in  $t$  and  $J$  around the atomic limit of the  $t$ – $J$  model (given by  $t \rightarrow 0$  and  $J \rightarrow 0$  in Eq. (1)). This can also be viewed as a high-temperature expansion since each factor of  $t$  and  $J$  must necessarily appear with a factor of  $\beta$ . The diagrams of this series give rise to the same two self-energy structure for the single-particle Green’s function as found in ECFL. In particular, Eq. (3.6) of Ref. [8] reads

$$\mathcal{G}_\sigma = \frac{\langle F^{\sigma 0} \rangle + \Delta_\sigma}{(G_\sigma^0)^{-1} - \Sigma_\sigma}. \quad (71)$$

We can make the identifications

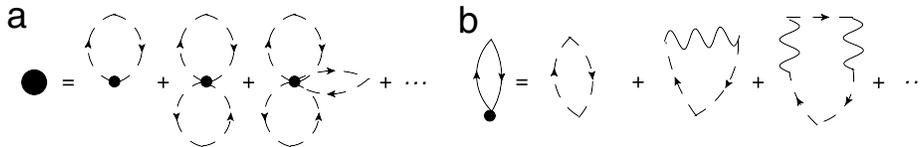
$$\langle F^{\sigma 0} \rangle \rightarrow 1 - \gamma; \quad \Delta_\sigma \rightarrow \Psi(k); \quad (G_\sigma^0)^{-1} \rightarrow \mathbf{g}^{-1(0)}; \quad \Sigma_\sigma \rightarrow -\epsilon_k \gamma + \Phi(k). \quad (72)$$

As is the case in the  $\lambda$  series, the fundamental object in the Zaitsev–Izyumov high-temperature series is the auxiliary Green’s function  $\mathbf{g}$ .

The main difference between the two series is the dimensionless expansion parameter. In the case of ECFL, it is the continuity parameter  $\lambda$ . In the case of the high-temperature series, it is  $\beta t$  and  $\beta J$ . To see this more explicitly, consider the simplest diagram in both series, which is the zeroth order diagram for  $\gamma$ . In ECFL, this is the diagram in Fig. 37(a). In Ref. [8], it is represented by a dot. The relationship between the two is shown in Fig. 41. In this figure, the dashed line indicates an atomic limit auxiliary Green’s function  $\mathbf{g}_{t \rightarrow 0, J \rightarrow 0}(i\omega_k) = \frac{1}{i\omega_k + \mu}$ . The big dot indicates the atomic limit value of  $\gamma$ , i.e.  $\gamma_{t \rightarrow 0, J \rightarrow 0} = \frac{\rho}{2}$ , where  $\rho = \frac{2e^{\beta\mu}}{1+2e^{\beta\mu}}$  is the atomic limit density. The wiggly line indicates a hopping  $\epsilon_k$ . Finally, the solid line indicates the bare auxiliary Green’s function  $\mathbf{g}^{(0)}(k) = \frac{1}{i\omega_k + \mu - \epsilon_k}$ . In panel (a), the zeroth order  $\gamma$  from the high-temperature series is expanded as an infinite series in  $\lambda$ . Here, each loop corresponds to  $\sum_{i\omega_k} \mathbf{g}_{t \rightarrow 0, J \rightarrow 0}(i\omega_k) = \frac{\frac{\rho}{2}}{1 - \frac{\rho}{2}}$ , and there is a minus sign between the successive terms of the series. Summing the geometric series, we find that  $\frac{\frac{\rho}{2}}{1 - \frac{\rho}{2}} \cdot \frac{1}{1 + \frac{\frac{\rho}{2}}{1 - \frac{\rho}{2}}} = \frac{\rho}{2}$ . In panel (b), the

zeroth order  $\gamma$  from the  $\lambda$  series is expanded as an infinite series in the hopping  $\epsilon_k$ . This gives the geometric series  $\sum_k \mathbf{g}^{(0)}(k) = \sum_k \sum_{n=0}^{\infty} \frac{\epsilon_k^n}{(i\omega_k + \mu)^{n+1}}$ . We see that to get from the high-temperature series to the  $\lambda$  series, one would have to break up all atomic limit objects into an infinite series in terms of  $\lambda$ , and replace every atomic limit auxiliary Green’s function with a bare propagating one.

We can summarize the fundamental difference between the two approaches as follows. In the case of zero magnetic field, the high-temperature series is an expansion around the atomic limit, i.e. an exponentially degenerate manifold of states, without giving preference to any one of them. In doing so, it is difficult to recover the adiabatic continuity aspect of physics relating to the Fermi-surface and the Luttinger–Ward volume theorem [19]. In contrast, ECFL builds the Fermi-surface into the  $\lambda$



**Fig. 41.** In panel (a), the zeroth order  $\gamma$  diagram from the high-temperature series (the big dot) is expanded as an infinite series in  $\lambda$ . The dashed lines indicate auxiliary Green's functions in the atomic limit  $\mathbf{g}_{t \rightarrow 0, J \rightarrow 0}(i\omega_k)$ . In panel (b), the zeroth order  $\gamma$  diagram from the  $\lambda$  series is expanded as an infinite series in the hopping. The solid line indicates a bare propagating auxiliary Green's function  $\mathbf{g}^{(0)}(k)$ , while the wavy line indicates the hopping  $\epsilon_k$ .

expansion at zeroth order, by expanding around the free Fermi gas and by maintaining continuity in  $\lambda$ . Finally, by enforcing that the number of auxiliary fermions equals the number of physical ones through the second chemical potential  $u_0$ , ECFL is able to satisfy the Luttinger–Ward volume theorem.

## 9. Conclusion

In conclusion, starting with the  $\lambda$  expansion as defined through iteration of the Schwinger EOM around the free Fermi gas [9,10], we derived a set of diagrammatic rules to calculate the  $n$ th order contribution to the physical Green's function  $\mathcal{G}$  in the  $t$ - $J$  model. The resulting diagrams suggested the need for two self-energies, which we denoted by  $\Sigma_a$  and  $\Sigma_b$ . Using the Schwinger equations of motion defining the ECFL objects,  $\mathbf{g}$ ,  $\tilde{\mu}$ ,  $\gamma$ ,  $\Phi$ , and  $\Psi$ , we derived diagrammatic rules for calculating these objects and found that they could be related simply to  $\Sigma_a^*$  and  $\Sigma_b^*$ , the irreducible parts of  $\Sigma_a$  and  $\Sigma_b$ . We also discovered diagrammatically that  $\Psi$  diagrams are simply a subset of the  $\Phi$  diagrams, with an interaction line missing. Denoting the remainder of the  $\Phi$  diagrams by the symbol  $\chi$ , this implied the expression  $\Phi(k) = \epsilon_k \Psi(k) + \chi(k)$ . We had already found this to be the case in the limit of infinite spatial dimensions with  $\chi(k) \rightarrow \chi(i\omega_k)$  and  $\Psi(k) \rightarrow \Psi(i\omega_k)$  in Ref. [23], and here we generalized it to finite dimensions. We also derived the Schwinger EOM for the object  $\chi$ . We derived diagrammatic rules for the three point vertices  $\Lambda$  and  $\mathcal{U}$ , defined as the functional derivatives of  $\mathbf{g}^{-1}$  and  $\tilde{\mu}$  respectively, with respect to the source. We derived a generalized Nozières relation for these vertices, which differs from the standard one for the Feynman diagrams. We then introduced skeleton diagrams into our series, thereby allowing us to make the connection with the iterative expansion of the Schwinger equations of motion (as done in Refs. [9] and [10]), which deals exclusively with skeleton diagrams.

We then derived the third order skeleton expansion for  $\mathbf{g}$  and  $\tilde{\mu}$ . Previously, this had been done only up to second order. We then discussed the error in the high-frequency limit incurred in the skeleton expansion carried out to any order in  $\lambda$ . We also discussed the “deviation” of the  $\lambda$  series from the Feynman series, thereby justifying on a qualitative level, the phenomenological ECFL [9,26], which has already been successful in explaining lines shapes found both from ARPES experiments [17], and from DMFT calculations [14]. Finally, we discussed the connection between ECFL and the Zaitsev–Izyumov high-temperature series. We found that while both formalisms dealt with the projection of double occupancy by introducing two self-energies, they had fundamentally different approaches to dealing with the problem of the Fermi-surface. While the high-temperature series is an expansion around a completely degenerate manifold of states, ECFL makes an adiabatic connection with the Fermi-surface and preserves the Luttinger–Ward volume theorem.

Our main motivation in deriving these diagrammatic rules is that they will allow the  $\lambda$  expansion to be evaluated to high orders using powerful numerical techniques such as diagrammatic Monte Carlo, and also that the intuition gained from the diagrams themselves could facilitate infinite resummations guided by some physical principles.

## Acknowledgments

This work was supported by DOE under Grant No. FG02-06ER46319. The authors thank Professor Antoine Georges, CPHT-Ecole Polytechnique and Collège de France, supported in part by the DARP/MURI-OLE program, for their warm hospitality, where this work was completed.

**References**

- [1] P.W. Anderson, *Science* 235 (1987) 1196; *The Theory of Superconductivity*, Princeton University Press, Princeton, NJ, 1997; arXiv:0709.0656 (unpublished).
- [2] B.S. Shastry, *Phys. Rev. B* 81 (2010) 045121.
- [3] Z. Wang, Y. Bang, G. Kotliar, *Phys. Rev. Lett.* 67 (1991) 2733–2736.
- [4] A. Greco, R. Zeyher, *Phys. Rev. B* 63 (2001) 064520.
- [5] A. Foussats, A. Greco, *Phys. Rev. B* 70 (2004) 205123.
- [6] R. Zeyher, A. Greco, *Phys. Rev. B* 80 (2009) 064519.
- [7] R.O. Zaitsev, *Sov. Phys - JETP* 43 (1976) 574.
- [8] Yu.A. Izyumov, B.M. Letfulov, *J. Phys.: Condens. Matter* 2 (1990) 8905–8923.
- [9] B.S. Shastry, *Phys. Rev. Lett.* 107 (2011) 056403.
- [10] B.S. Shastry, *Phys. Rev. B* 87 (2013) 125124.
- [11] B.S. Shastry, *Ann. Phys.* 343 (2014) 164.
- [12] J.M. Luttinger, J.C. Ward, *Phys. Rev.* 118 (1960) 1417.
- [13] J.M. Luttinger, *Phys. Rev.* 119 (1960) 1153.
- [14] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, B.S. Shastry, *Phys. Rev. B* 88 (2013) 235132.
- [15] B.S. Shastry, E. Perepelitsky, A.C. Hewson, *Phys. Rev. B* 88 (2013) 205108.
- [16] E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, B.S. Shastry, *Phys. Rev. B* 87 (2013) 161120.
- [17] G.H. Gweon, B.S. Shastry, G.D. Gu, *Phys. Rev. Lett.* 107 (2011) 056404; K. Matsuyama, G.H. Gweon, *Phys. Rev. Lett.* 111 (2013) 246401.
- [18] A.L. Fetter, J.D. Walecka, *Quantum Theory of Many-Particle Systems*, Dover Publications, Inc., Mineola, NY, 2003.
- [19] P. Nozières, *Theory of Interacting Fermi Systems*, W. A. Benjamin, Amsterdam, 1964.
- [20] N.V. Prokof'ev, B.V. Svistunov, I.S. Tupitsyn, *J. Exp. Theor. Phys.* 87 (2) (1998) 310–321; K.V. Houcke, E. Kozik, N. Prokof'ev, B. Svistunov, *Phys. Procedia* 6 (2010) 95–105.
- [21] D. Hansen, B.S. Shastry, *Phys. Rev. B* 87 (2013) 245101.
- [22] A.B. Harris, D. Kumar, B.I. Halperin, P.C. Hohenberg, *Phys. Rev. B* 3 (1971) 961.
- [23] E. Perepelitsky, B.S. Shastry, *Ann. Physics* 338 (2013) 283.
- [24] L.P. Kadanoff, G. Baym, *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problems*, Benjamin, NY, 1962.
- [25] W. Kohn, J.M. Luttinger, *Phys. Rev.* 118 (1960) 41.
- [26] B.S. Shastry, *Phys. Rev. B* 84 (2011) 165112.

# Low-energy physics of the $t$ - $J$ model in $d = \infty$ using extremely correlated Fermi liquid theory: Cutoff second-order equations

B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*

Edward Perepelitsky

*Centre de Physique Théorique, École Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France  
and Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*

(Received 26 May 2016; revised manuscript received 11 July 2016; published 28 July 2016)

We present the results for the low-energy properties of the infinite-dimensional  $t$ - $J$  model with  $J = 0$ , using  $O(\lambda^2)$  equations of the extremely correlated Fermi liquid formalism. The parameter  $\lambda \in [0, 1]$  is analogous to the inverse spin parameter  $1/(2S)$  in quantum magnets. The present analytical scheme allows us to approach the physically most interesting regime near the Mott insulating state  $n \lesssim 1$ . It overcomes the limitation to low densities  $n \lesssim 0.7$  of earlier calculations, by employing a variant of the skeleton graph expansion, and a high-frequency cutoff that is essential for maintaining the known high- $T$  entropy. The resulting quasiparticle weight  $Z$ , the low  $\omega, T$  self-energy, and the resistivity are reported. These are quite close at all densities to the exact numerical results of the  $U = \infty$  Hubbard model, obtained using the dynamical mean field theory. The present calculation offers the advantage of generalizing to finite  $T$  rather easily, and allows the visualization of the loss of coherence of Fermi liquid quasiparticles by raising  $T$ . The present scheme is generalizable to finite dimensions and a nonvanishing  $J$ .

DOI: [10.1103/PhysRevB.94.045138](https://doi.org/10.1103/PhysRevB.94.045138)

## I. INTRODUCTION

The fundamental importance of the  $t$ - $J$  model for understanding the physics of correlated matter, including high- $T_c$  superconductors, has been recognized for many years [1]. The  $t$ - $J$  model is a prototype of extreme correlations, incorporating the physics of (Gutzwiller) projection to the subspace of single occupancy. The added superexchange  $J$  provides the mechanism for quantum antiferromagnetism at half-filling, and upon hole doping, for superconductivity via singlet pairing [1]. This viewpoint has attracted much attention in the community. It has led to many approximate methods of calculation being applied to the  $t$ - $J$  model, in order to calculate experimentally measured variables. Despite intense effort in recent years, schemes for *controlled* calculations are rare since the model has well-known fundamental complexities that need to be overcome.

Motivated by this challenge, we have recently formulated the extremely correlated Fermi liquid (ECFL) theory [2,3] for tackling the  $t$ - $J$  and related  $U \rightarrow \infty$  type models. The ECFL theory deals with the  $t$ - $J$  model by viewing it as a noncanonical Fermi problem, and proceeds via a nonlinear representation of Gutzwiller projected fermions in terms of canonical fermions. It is pedagogically useful to draw a parallel [3] to the Dyson-Maleev representation of spins [4] used in quantum magnets. In this representation [4], the spins are hard core bosons, and are nonlinearly expressed in terms of the canonical bosons, namely, the spin waves. The ECFL methodology developed to date consists of successive approximations in the expansion parameter  $\lambda \in [0, 1]$ , playing a role analogous to the inverse spin parameter  $1/(2S)$  in quantum magnetism. This analogy is developed in [3], where parallels between the ECFL calculations and earlier calculations of the partition function and Green's functions of the spin problem are drawn. It is useful to note that the classical limit for spins  $1/S \rightarrow 0$  corresponds

to the limit of free-fermion limit  $\lambda \rightarrow 0$ . Continuity in  $\lambda$  leads to a protection of the Fermi surface volume for the interacting theory, i.e., the Luttinger-Ward volume theorem is obeyed. Low-order expansions can be performed analytically for the most part, and therefore have all the usual advantages of analytic approaches, such as explicit formulas for variables of interest and also flexibility for different situations. Several recent applications of the ECFL theory, mentioned below, show promise in terms of reproducing the salient features of exact numerical solutions of strong coupling models, wherever available [5,6]. The theory has also had success in reconciling extensive data on angle-resolved photoemission (ARPES) line shapes [7], including subtle features such as the low-energy kinks, and has made testable predictions on the asymmetry of line shapes [8].

In order to understand better the nature as well as limitations of a low-order expansion in  $\lambda$ , we have tested the solution against two important strongly correlated problems where the numerical renormalization group and related ideas provide exact numerical results. In [5], the asymmetric Anderson impurity problem, solved by Wilsonian renormalization numerical group methods [9–12] was used as one of the benchmarking models. Second, in [6], the  $d \rightarrow \infty$  Hubbard model at large  $U$ , solved numerically by the dynamical mean field theory (DMFT) method [13–39], was used as the benchmarking model. These benchmarking studies show that the ECFL approach is overall consistent with the exact solutions, with some caveats. There are indeed differences in detailed structures at higher energies [40]. However, the raw initial results seem both useful and reliable for obtaining the low-energy spectrum, and for a broad understanding of the occupied side of the spectral functions. We further found that the calculations are *very close* to the exact solutions, *provided we scale the frequencies* by the respective quasiparticle weights  $Z$  of the two theories.

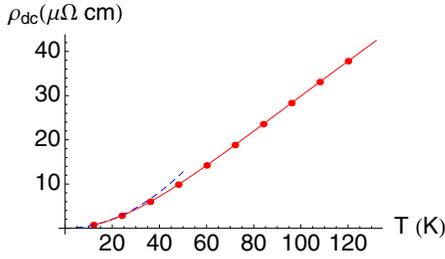


FIG. 1.  $\rho_{dc}$  on absolute scale vs  $T$  in Kelvin for particle density  $n = 0.85$ . We have used the estimates  $D = 12000$  K and  $\rho_0 = 258 \mu\Omega$  cm. The latter is obtained by using [57]  $\rho_0 \approx \frac{\hbar a_0 d}{e^2}$ , where we estimate  $a_0 d \approx 10^{-8}$  cm. The Fermi liquid behavior with quadratic resistivity in the blue dotted line breaks down above  $T_{FL} \approx 30$  K, and is followed by a regime of linear resistivity.

The version of the ECFL presented in [6] and the closely related [5] is therefore promising, but has the limitation of being confined to low density  $n \lesssim 0.7$ . In the most interesting density range  $n \lesssim 1$ , it falls short of being a “standalone theory” since the magnitude of the calculated  $Z$  is too large. One requires rescaling frequencies to compensate for the incorrect magnitude of  $Z$ , and thereby improve the agreement. It is therefore important to find ways to extend this analytical approach to cover the physically most interesting density regime  $0.7 \leq n \leq 1$ . A diagnostic objective of this paper is to identify the cause for the inaccurate  $Z$  in the earlier version, and to explore ways to overcome it. We have found it possible to do both. This paper presents an alternative scheme that can be pushed to high-particle densities as well. We show here that the resulting scheme gives satisfactory results for most of the interesting low  $\omega, T$  variables of the model.

Amongst the several variables of interest, the transport objects are the most difficult ones to compute reliably. The difficulty lies in their great sensitivity to the *lowest excitation energies*, and in the paucity of reliable tools to capture these. The limit of large dimensionality is helpful here since it has the great advantage of killing the vertex corrections [41]. Thus, a knowledge of the one-electron Green’s function can give us *the exact resistivity of a metal*, arising from inelastic mutual collisions of electrons. Despite the stated simplification, this calculation remains technically challenging. In important recent work, this calculation has been performed in [42,43] for the large- $U$  Hubbard model in infinite dimensions. The authors have produced exact resistivity results that are so rare in condensed matter systems. We can use them to benchmark our results for the resistivity at different densities and temperature. We report the results of this comparison in this paper. Figure 1 shows one of the main results of the calculation presented here, the details leading to it are described below.

In Sec. II, we summarize the second-order equations and introduce the various Green’s functions and self-energies needed. In Sec. III, we identify the conditions necessary for getting a satisfactory  $Z$  near half-filling. In Sec. IV, after summarizing the self-consistency loop, we give a prescription for modifying the earlier equations and give the new set. This requires using a slightly different skeleton graph expansion, where certain objects are evaluated exactly using the number sum rule. The ECFL theory has some intrinsic freedom in

choosing the details of the skeleton expansion, more so than in the standard Feynman graph based canonical models. That freedom can be usefully employed here. We find that it is also obligatory to introduce a high-energy cutoff, in order to recover the known high- $T$  entropy of the model. While the precise form of the cutoff is not uniquely given by theory, we found that several reasonable functional forms gave comparable results at low energies and low  $T$ , provided that the parameters were chosen to yield the *high- $T$  entropy*. This cutoff also eliminates weak tails in the spectral functions that otherwise extend to large negative (i.e., occupied) energies.

In Sec. V, we present results for the  $T$  and  $n$  variation of the chemical potential and the quasiparticle weight  $Z$ . We also present the  $\omega, T$ , and  $n$  variation of the self-energy and spectral functions, where the quasiparticles, the asymmetry of the spectral functions, and the thermal destruction of the quasiparticles are highlighted. In Sec. VI, we present results for the resistivity at low and intermediate  $T$  for various densities. In Sec. VII, we provide a summary and discuss the prospects for further work.

## II. SUMMARY OF SECOND-ORDER ECFL THEORY

Let us begin by recounting the exact formal expression for the Green’s function of the  $t$ - $J$  model. In the ECFL theory, this object is given exactly as

$$\mathcal{G}(k, i\omega_n) = \mathbf{g}(k, i\omega_n) \times \tilde{\mu}(k, i\omega_n), \quad (1)$$

a product of the auxiliary Green’s function  $\mathbf{g}$  and the “comparison” function [44] given in terms of a second self-energy  $\Psi(k, i\omega_n)$  and the particle density  $n$  as  $\tilde{\mu}(k, i\omega_n) = \{1 - n/2 + \Psi(k, i\omega_n)\}$ . The auxiliary Green’s function  $\mathbf{g}(k, i\omega_n)$  given by

$$\mathbf{g}(k, i\omega_n) = \frac{1}{i\omega_n + \mu - \{1 - n/2\} \varepsilon_k - \Phi(k, i\omega_n)}, \quad (2)$$

where  $\mu$  is the chemical potential and  $\varepsilon_k$  the band energy. In the infinite-dimensional limit, it is demonstrated in [45] that an exact simplification occurs with these equations, whereby the momentum dependence is given by

$$\Psi(k, i\omega_n) = \Psi(i\omega_n), \quad (3)$$

$$\Phi(k, i\omega_n) = \chi(i\omega_n) + \varepsilon_k \Psi(i\omega_n), \quad (4)$$

where both  $\Psi$  and  $\chi$  are functions of only the fermionic Matsubara frequency  $\omega_n = (2n + 1)\pi\beta$ , *but not the momentum  $k$* . These expressions can be used in Eq. (1) and upon using the analytic continuation  $i\omega_n \rightarrow \omega + i0^+$ , we may express the Green’s function in the standard Dyson representation

$$\mathcal{G}(k, \omega + i0^+) = \frac{1}{\omega + i0^+ + \mu - \varepsilon_k - \Sigma(\omega + i0^+)}, \quad (5)$$

where the Dyson self-energy is now manifestly momentum independent, and given by

$$\Sigma(\omega + i0^+) = \mu + \omega + \frac{\chi(\omega + i0^+) - \mu - \omega}{1 - \frac{n}{2} + \Psi(\omega + i0^+)}. \quad (6)$$

This result demonstrates the momentum independence of the Dyson self-energy of the  $t$ - $J$  model in infinite dimensions. It

is consistent with the analogous result for the Hubbard model at any  $U$  [13–15].

Within the ECFL theory we rely upon a systematic  $\lambda$  expansion to compute the two self-energies  $\Psi$  and  $\chi$ . This  $\lambda$  expansion is described in detail in [2,46,47], in brief the parameter  $\lambda$  lives in the range  $\in \{0, 1\}$ , and plays the role of the quantum parameter  $1/(2S)$  in the large spin expansions familiar in the theory of magnetism. A skeleton diagram method can be devised for expanding the self-energies  $\Psi$  and  $\chi$  in a formal power series in  $\lambda$ , with terms that are functionals of  $\mathbf{g}$  and the band energies  $\varepsilon_k$ . This expansion uses the full  $\mathbf{g}$  (rather than noninteracting propagators  $\mathbf{g}_0$ ) as fundamental units, or “atoms,” for the expansion. The procedure is in close analogy with the skeleton diagram methods used in many-body theory. Having the self-energies to a given order in  $\lambda$ , one now reconstructs the Green’s functions self-consistently, the scheme is to second order in the present case.

The explicit equations to second order are found to be

$$\mathcal{G}(k, i\omega_n) = \mathbf{g}(k, i\omega_n) \{a_G + \lambda \Psi(k, i\omega_n)\}, \quad (7)$$

$$\mathbf{g}^{-1}(k, i\omega_n) = i\omega_n + \boldsymbol{\mu}' - \lambda \chi(k, i\omega_n) - \{a_G + \lambda \Psi(k, i\omega_n)\} \left( \varepsilon_k - \frac{u_0}{2} \right), \quad (8)$$

with

$$a_G = 1 - \lambda \mathcal{G}(j, j^-) = 1 - \lambda \sum_k \mathcal{G}(k, i\omega_n) e^{i\omega_n 0^+}, \quad (9)$$

where  $\boldsymbol{\mu}' = \boldsymbol{\mu} - \frac{u_0}{2}$ . In Eq. (9), the middle (last) term is in space-time (wave-vector-frequency) variables, denoted respectively in the compact notation  $j \equiv (\vec{R}_j, \tau_j)$ ,  $k \equiv (\vec{k}, i\omega_n)$ , and denoting  $j^- \equiv (\vec{R}_j, \tau_j + i0^-)$ . The two self-energy functions  $\Psi$  and  $\chi$  are expanded formally in  $\lambda$  as  $\Psi = \Psi_{[0]} + \lambda \Psi_{[1]} + \dots$  and  $\chi = \chi_{[0]} + \lambda \chi_{[1]} + \dots$ . A systematic expansion in  $\lambda$  is available to third order in [46], from the low-order results [48] we find  $\Psi_{[0]} = 0$ ,  $\chi_{[0]} = -\sum_p \mathbf{g}(p)(\varepsilon_p - \frac{u_0}{2})$  and

$$\Psi_{[1]}(k) = -\sum_{pq} (\varepsilon_p + \varepsilon_q - u_0) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k), \quad (10)$$

$$\chi_{[1]}(k) = -\sum_{pq} \left( \varepsilon_{p+q-k} - \frac{u_0}{2} \right) (\varepsilon_p + \varepsilon_q - u_0) \times \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k). \quad (11)$$

In view of the explicit factors of  $\lambda$  in Eqs. (7) and (8), this leads to an  $O(\lambda^2)$  approximation for  $\mathcal{G}$ ; the recipe further requires that the parameter  $\lambda$  is set to unity before computing. Here,  $u_0$  denotes the second chemical potential. It enters the theory as a Hubbard-type term with a self-consistently determined coefficient  $u_0$ , as described in [47]. This chemical potential is essential in order to satisfy the shift invariance of the  $t$ - $J$  model order by order in  $\lambda$ , namely,  $t_{ij} \rightarrow t_{ij} + c \delta_{ij}$  with an arbitrary constant  $c$ . For instance, we see in Eq. (11) that a shift of the energies  $\varepsilon_k \rightarrow c + \varepsilon_k$  is rendered immaterial due to the structure of the terms, the constant  $c$  can be absorbed into  $u_0$ . The two chemical potentials  $\boldsymbol{\mu}$  and  $u_0$  are determined through the pair of sum rules on the auxiliary  $\mathbf{g}$  and the standard number

sum rule on  $\mathcal{G}$ :

$$\sum_k \mathbf{g}(k) e^{i\omega_n 0^+} = \frac{n}{2} = \sum_k \mathcal{G}(k) e^{i\omega_n 0^+}. \quad (12)$$

In dealing with Eq. (9), the composite nature of the  $\mathcal{G}$  on view in Eq. (1) offers a choice for implementing the skeleton expansion. Such a choice is absent in the more standard many-body problems. On the one hand, we could use the sum rule Eq. (12) for  $\mathcal{G}$  giving

$$a_G^{(I)} \rightarrow 1 - \lambda \frac{n}{2}, \quad (13)$$

reducing to the exact answer  $a_G^{\text{exact}} = 1 - \frac{n}{2}$  as  $\lambda \rightarrow 1$ .

Alternately, we could expand the  $\mathcal{G}$  in powers of  $\lambda$ , a procedure we followed in [6,49]. We expanded  $\mathcal{G}$  out to first order in  $\lambda$  from Eq. (7) since that already gives the required  $O(\lambda^2)$  correction. Thus, we set  $\mathcal{G} = \mathbf{g}(1 - \lambda n/2) + O(\lambda^2)$ , where the sum rule Eq. (12) was used for evaluating  $\sum_{k, \omega_n} \mathbf{g}(k)$ . As a result, we obtain the approximate result

$$a_G^{(II)} = 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4} + O(\lambda^3). \quad (14)$$

Setting  $\lambda \rightarrow 1$ , we thus get two alternate approximate skeleton versions of Eq. (7):

$$\mathcal{G}^{(I)}(k, i\omega_n) = \mathbf{g}^{(I)}(k, i\omega_n) \{1 - n/2 + \Psi(k, i\omega_n)\}, \quad (15)$$

$$\mathcal{G}^{(II)}(k, i\omega_n) = \mathbf{g}^{(II)}(k, i\omega_n) \times \{1 - n/2 + n^2/4 + \Psi(k, i\omega_n)\}, \quad (16)$$

where both expressions involve *the same* approximate  $\Psi$  given in Eq. (10), and the auxiliary  $\mathbf{g}^{(\dots)}$  is also adjusted to have the appropriate expression for  $a_G$  in Eq. (8). This dichotomous situation arises due to the composite nature of the physical  $\mathcal{G}$ , whereas in standard many-body problems the skeleton expansion is unique.

In [6] as well as [49] we employed Eq. (16) to compute the electron self-energy and spectral functions. It was argued that this expression should be valid for low-particle density  $n \lesssim 0.7$ . In [6], the results were compared with the numerically exact DMFT results for the same model. It was found that the self-energy is indeed close to the exact answer in the low-density limit. At the other end of *high densities*  $n \lesssim 1$ , it was found that the self-energy is also very close to the exact result, provided we scale the frequencies by the quasiparticle weight  $Z$  of that theory. This remarkable observation shows that in ECFL theory, the Dyson self-energy (6) found by compounding two simpler expressions  $\chi$  and  $\psi$ , has the correct functional form. Moreover, the unusual and important feature of particle-hole asymmetry, i.e., the presence of a strong  $\omega^3$  term in the  $\text{Im} \Sigma$ , comes about “naturally” within the scheme. This feature has been argued to be generic for strongly correlated systems, as argued in [8] and in the closely related [5] for the Anderson impurity model. The need for rescaling the frequency arises because the computed  $Z^{(II)}$  using the approximate version (16) overestimates this variable as  $n$  increases beyond the estimated limit of  $n \sim 0.7$ . We see in [6] (Fig. 16) that  $Z^{(II)}$  does not even vanish as  $n \rightarrow 1$ , as one expects in a Mott insulator.

Within the spirit of Eq. (16) one might expect that further approximations involving higher-order terms in  $\lambda$  will enhance the range of validity in density. Such a program is essentially numerically intensive since beyond second order, one needs to use other techniques, such as Monte Carlo generation and evaluation of diagrams [50–52]. We are currently performing these calculations, and have made formal progress towards this goal in [46], by enumerating the nontrivial diagrammatic rules in this model. The diagrams that we encounter include and go beyond Feynman diagrams, as necessitated by the lack of Wick's theorem in the noncanonical theory.

On the other hand, the analytical ease of the second-order theory offers considerable advantage relative to other contemporary methods. For low orders in  $\lambda$  most calculations can be done by hand, and the remaining computations are modest in scope. Analytical methods also have a much greater flexibility: they can be applied in lower dimensions as well. Further, the agreement with the other methods (DMFT [13–39], numerical renormalization group [5]) and also experiments on ARPES for the electron line shapes [7] is very good. In view of these positive factors, it appears to be useful to examine if the problem with the quasiparticle weight  $Z^{(11)}$  at  $n \lesssim 1$  can be understood and corrected, making other necessary approximations along the way. This is indeed the purpose of this paper; we will see below that the approximation (15) provides us with the correct direction for such an approach.

### III. SUM RULES NECESSARY FOR THE VANISHING OF $Z$ NEAR THE MOTT INSULATING STATE

Let us first understand the factors that make  $Z$  vanish as we approach the Mott insulating limit. For this purpose, it is useful to recall the local density of states of the Hubbard model for the case of a sufficiently large  $U$  (see [53] for a useful discussion). Here, we expect the formation and clear separation of characteristic lower and upper Hubbard bands, as indicated in the schematic Fig. 2. Specializing to  $T = 0$  for simplicity,

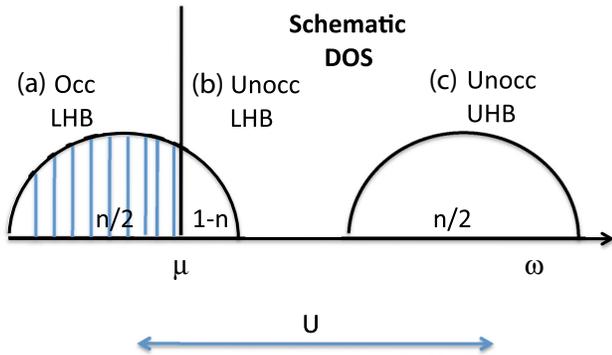


FIG. 2. A schematic depiction of the local spectral density of states  $\rho_{GLocal}(\omega)$  [popularly called  $A_{Local}(\omega)$ ] for the large- $U$  Hubbard model, where the correlation split Hubbard bands are clearly separated. It shows three regions: (a) occupied electronic states, (b) unoccupied lower Hubbard band states, and (c) unoccupied upper Hubbard band states, with their respective weights as in Eq. (18). The  $t$ - $J$  model sends the region (c) off to infinity with weights given in Eq. (19). The area in region (b) is exactly  $(1 - n)$ , and preserving this in an approximation is key to obtaining the correct low-energy scale.

we note that for the Hubbard model with  $n < 1$ , the spectral weight for the local  $\rho_G(\omega)$  of the physical electron satisfies the unitary sum rule  $\int d\omega \rho_G(\omega) = 1$ . We use a notation where a sum over  $\vec{k}$  is implied for unlabeled functions (without the  $\vec{k}$  argument), e.g.,  $\rho_G(\omega) \equiv \sum_{\vec{k}} \rho_G(\vec{k}, \omega)$ . The local Green's function itself is given by

$$G(\omega + i0^+) = \int dv \frac{\rho_G(v)}{\omega - v + i0^+}, \quad (17)$$

and so the  $\omega \rightarrow \infty$  asymptotic behavior is determined by this sum rule as  $G(\omega) \rightarrow \frac{\int dv \rho_G(v)}{\omega} = \frac{1}{\omega}$ . This can be partitioned into three sum rules as depicted in Fig. 2:

$$\begin{aligned} \int_{-\infty}^0 d\omega \rho_G(\omega) &= n/2, & \int_0^{\Omega_*} d\omega \rho_G(\omega) &= 1 - n, \\ \int_{\Omega_*}^{\infty} d\omega \rho_G(\omega) &= n/2, \end{aligned} \quad (18)$$

where  $\Omega_*$  is an energy scale denoting the upper end of the lower Hubbard band and hence is  $\sim O(W)$ ; it is well defined provided  $U \gg W$ . As stated, these three integrals add up to 1, ensuring that a full electron is captured. On the other hand, the  $t$ - $J$  model spectral function  $\rho_G(\omega)$  satisfies

$$\int_{-\infty}^0 d\omega \rho_G(\omega) = n/2, \quad \int_0^{\infty} d\omega \rho_G(\omega) = 1 - n, \quad (19)$$

where the upper Hubbard band (and  $\Omega_*$ ) is pushed off to  $+\infty$ , and thus the occupied and unoccupied portions add up to  $1 - n/2$ . This can be visualized clearly with the help of Fig. 2. This argument also determines the  $\omega \rightarrow \infty$  asymptotic form  $\lim_{\omega \rightarrow \infty} \mathcal{G}(\omega) \rightarrow \frac{1-n/2}{\omega}$ , and gives us a relation of importance to this study:

$$\left( \lim_{\omega \rightarrow \infty} \mathcal{G}(\omega) \rightarrow \frac{1-n/2}{\omega} \right) \leftrightarrow \left( \int_0^{\infty} d\omega \rho_G(\omega) = 1 - n \right). \quad (20)$$

To see its relevance, we note that as  $n \rightarrow 1$ , the chemical potential increases towards the top of the lower Hubbard band. This implies that the unoccupied portion of the lower Hubbard band shrinks to zero. Since roughly half of the quasiparticle's weight [54] resides in this shrinking energy domain of  $O(1 - n)$  times the bandwidth, the quasiparticle residue  $Z$  must vanish at least as fast as  $O(1 - n)$ .

We may now refer back to Eq. (16); since from the definitions (10) and (11) we can see that  $\lim_{\omega \rightarrow \infty} (\Psi(\omega), \chi(\omega)) \rightarrow 0$  and also  $\lim_{\omega \rightarrow \infty} \mathbf{g}(\omega) \rightarrow \frac{1}{\omega}$ , we combine these to obtain

$$\lim_{\omega \rightarrow \infty} \mathcal{G}^{(11)}(\omega) \rightarrow \frac{1 - n/2 + n^2/4}{\omega},$$

whereby the unoccupied region  $\int_0^{\infty} d\omega \rho_{G^{(11)}}(\omega) = 1 - n + n^2/4$ , in conflict with the condition (20) for a vanishing  $Z$ , as  $n \rightarrow 1$ .

Having thus identified this weakness of the approximation, we also see by the same argument that Eq. (15) would automatically give us a vanishing  $Z$ , as  $n \rightarrow 1$ ; the factors are now appropriate for the condition (20) to hold.

#### IV. CUTOFF SECOND-ORDER ECFL THEORY

Motivated by the above discussion, we now implement a skeleton graph expansion, where the basic atoms, or units, are still  $\mathbf{g}$ , but in static terms involving  $\mathcal{G}$ , such as in Eq. (9), we use the exact particle-number sum rule (12). This leads us to study the equations in Eq. (15).

##### A. Full set of self-consistent equations

For convenience and future reference, we summarize the full set of equations to be solved self-consistently. These are similar to the ones used in [6,49] with all the necessary changes for the present case made. The band density of states is taken as the semicircular expression  $D(\epsilon) = 2/(\pi D)\sqrt{1 - (\epsilon/D)^2}$ , and thus  $2D$  is the bare bandwidth. The complex frequency is denoted as  $z = \omega + i0^+$ , the local Green's function and its energy moments are defined by

$$\mathbf{g}^{-1}(\epsilon, z) = z + \mu' - (\epsilon - u_0/2) \left( 1 - \frac{n}{2} + \Psi_{[1]}(z) \right) - \chi_{[1]}(z), \quad (21)$$

$$\mathbf{g}_{Loc,m}(z) = \int d\epsilon D(\epsilon) \mathbf{g}(\epsilon, z) \epsilon^m = \int d\nu \frac{\rho_{\mathbf{g}L,m}(\nu)}{z - \nu}. \quad (22)$$

The chemical potential  $\mu'$  absorbs all constants such as  $\chi_{[0]}$ , leading to

$$\mu = \mu' + \frac{u_0}{2} \left( 1 + \frac{n}{2} \right) - \int d\omega f(\omega) \rho_{\mathbf{g}L,1}(\omega), \quad (23)$$

where  $f(\omega) = 1/(1 + \exp \beta\omega)$  is the Fermi function and we will need below  $\bar{f} = 1 - f$ . Equation (22) serves to introduce the spectral functions  $\rho_{\mathbf{g}L,m}(\nu)$ ; these are most often computed from the reversed relation

$$\rho_{\mathbf{g}L,m}(\omega) = -\frac{1}{\pi} \text{Im} \mathbf{g}_{Loc,m}(\omega + i0^+). \quad (24)$$

The physical Green's function is found from  $\mathcal{G}(\epsilon, z) = [1 - n/2 + \Psi(z)] \times \mathbf{g}(\epsilon, z)$ , and the Dyson self-energy from  $\Sigma(z) = z + \mu - \epsilon - \mathcal{G}^{-1}(\epsilon, z)$ . We define its local version  $\mathcal{G}_{Loc}(\omega)$  and its density through a band integration

$$\mathcal{G}_{Loc,m}(z) = \int d\epsilon D(\epsilon) \epsilon^m \mathcal{G}(\epsilon, z),$$

$$\rho_{\mathcal{G}L,m}(\omega) = -\frac{1}{\pi} \text{Im} \mathcal{G}_{Loc,m}(\omega + i0^+). \quad (25)$$

The physical momentum-integrated spectral function  $\rho_{\mathcal{G}L,0}$  is an object of central interest. It is also needed for the number sum rule below Eq. (30). The computation of  $\mathbf{g}$  requires the two complex self-energies  $\Psi, \chi$ . These can in turn be found from expressions involving the fundamental convolution:

$$\rho_{abc}^{(\mathcal{I})}(u) = \int_{u_1, u_2, u_3} \delta(u + u_3 - u_1 - u_2) \{ f(u_1) f(u_2) \bar{f}(u_3) + \bar{f}(u_1) \bar{f}(u_2) f(u_3) \} \rho_{\mathbf{g}L,a}(u_1) \rho_{\mathbf{g}L,b}(u_2) \rho_{\mathbf{g}L,c}(u_3), \quad (26)$$

where the right-hand side is conveniently computed from the local densities  $\rho_{\mathbf{g}L,a}$ , by using fast Fourier transforms.

This density is required for  $(a,b,c) = 0,1$ , and determines the complex function

$$\mathcal{I}_{abc}(z) = \mathcal{P} \int d\nu \frac{\rho_{abc}^{(\mathcal{I})}(\nu)}{z - \nu}. \quad (27)$$

From this object, the two self-energies can be found as the combinations

$$\Psi_{[1]}(z) = 2\mathcal{I}_{010}(z) - u_0 \mathcal{I}_{000}(z),$$

$$\chi_{[1]}(z) = 2\mathcal{I}_{011}(z) - u_0 [\mathcal{I}_{010}(z) + \mathcal{I}_{001}(z)] + \frac{u_0^2}{2} \mathcal{I}_{000}(z). \quad (28)$$

In summary, we can compute  $\mathbf{g}$  in terms of  $\chi, \Psi$  from Eq. (21). Having done so, we compute  $\chi, \Psi$  in terms of the  $\mathbf{g}$  from Eq. (28), thus defining the second part of the loop. The two chemical potentials  $\mu$  and  $u_0$  are found from Eq. (23) and the two particle-number sum rules:

$$\int d\omega f(\omega) \rho_{\mathbf{g}L,0}(\omega) = \frac{n}{2}, \quad (29)$$

$$\int d\omega \bar{f}(\omega) \rho_{\mathbf{g}L,0}(\omega) = \frac{n}{2}, \quad (30)$$

thereby all variables can be self-consistently calculated through a simple iterative scheme. The only inputs are the density of particles  $n$  and the temperature  $T$ .

##### B. Considerations of high density $n \rightarrow 1$ at low $T$ , and the entropy at high $T$

Before discussing the results, we note an important constraint that arises when we study the theory at high temperatures. We need to make sure that the number of states after the Gutzwiller projection has the correct value; this requires that the chemical potential has the correct asymptotic value at high  $T$ . When  $T \gg t, J$  the chemical potential grows linearly with  $T$ . From simple considerations of the atomic limit  $t = 0 = J$ , one can calculate the partition function exactly, from this one finds

$$\mu \sim k_B T \ln \{ n/[2(1-n)] \}, \quad (31)$$

where  $N_s$  and  $n = N/N_s$  are the number of sites and the density, respectively. This linear growth with  $T$  with the correct coefficient also ensures that the entropy near the Mott limit is correctly reproduced at high  $T$ . Upon using the Maxwell relation  $(\partial S/\partial N)/T = -(\partial \mu/\partial T)_N$ , and the initial condition  $S(n \rightarrow 0) = 0$ , we find

$$S \sim -k_B N_s \{ n \ln n/2 + (1-n) \ln(1-n) \}, \quad (32)$$

a well-known result. We must therefore also ensure that the approximation satisfies this condition (31), in order to obtain the correct entropy at high  $T$ .

Upon solving Eqs. (21)–(30) at high densities  $n \gtrsim 0.8$  as  $T \rightarrow 0$ , or high  $T \gg D$  with moderated densities  $n \gtrsim 0.7$ , we find that in each case the spectral function tends to flatten out on the occupied side, extending in range to  $\omega \ll -D$  with little weight in the tails. For the high- $T$  case, a second consequence is that the computed slope  $d\mu/dT$  begins to depart from Eq. (31). The flattening is consequence of the

growth of  $u_0$  which also increases linearly with  $T$ , becoming larger than the bandwidth  $2D$ , as seen in Fig. 5. This growth enhances the coefficients in the self-energies Eq. (28) and pushes one into a strong  $u_0$  regime, unless we impose some cutoff. In the  $T \rightarrow 0$  limit, the exact numerical results for spectral functions from DMFT [6] do confirm the expectation of a compact support for the spectral function, and hence the observed growth is artificial.

### C. Cutoff scheme with a Tukey window

We saw above that two physically distinct regimes involving different types of physics, namely, the high- $T$  regime at any density and the high-density regime at low  $T$ , share the common problem of growing tails of the spectral function.

In order to control this unphysical growth in both cases, we need to impose an appropriate high-energy cutoff. Higher-order terms in the  $\lambda$  expansion are expected to eliminate this growth in a systematic way, without needing an extra prescription. A detailed analysis of the cutoff issue within the  $\lambda$  expansion is underway currently, and we expect to present the details in a forthcoming paper. However, at the level of the lowest-order approximations, it seems that we do need to impose an extra cutoff, thereby introducing one more approximation. A rough estimate of the cutoff can be made by observing that the self-energy calculated by using the bare  $\mathbf{g}_0$  (setting  $\chi \rightarrow 0$  and  $\Psi \rightarrow 0$ ) in Eq. (26) would give the spectral weights a width of maximum range  $\pm 3D$ ; by setting  $u_1 = D$ ,  $u_2 = D$ ,  $u_3 = -D$ , we satisfy one of the Fermi combinations with  $u \sim 3D$ . By flipping signs we can reach  $u = -3D$ , thus, a range of frequencies  $-3D \leq \omega \leq 3D$  is feasible. The region near  $|\omega| \sim 3D$  would then be in the tails of the function. In a skeleton expansion, on the other hand, with increasing interaction strength  $u_0$ , we have the possibility of a runaway growth since under first iteration, the computed  $\rho_{\mathbf{g}L}$  can now extend to  $\pm 3D$  as compared to the range  $\pm D$  of the bare density, and so forth. Hence, one plausible strategy would be to limit the growth of the auxiliary spectral functions to a range  $\pm c_0$ , with  $c_0 \sim 2D$ , with the physical spectral functions possibly extending somewhat beyond this. Since two very different regimes, that of high  $T$  and high density are involved, we can test the additional approximations self-consistently, and thereby avoid unduly biasing the results.

It appears reasonable to choose the high-energy cutoff by requiring that we obtain the known high- $T$  slope and therefore the high- $T$  entropy (32) at all densities. While it might be possible to obtain the exact entropy by adjusting the cutoffs at each density separately, we content ourselves by finding a reasonable global fit instead, i.e., one set of *density-independent cutoffs* yielding the roughly correct entropy at relevant densities. The high- $T$  entropy is estimated at  $T \lesssim 1$ . It should be noted that  $T \sim 1$  is not always in the high- $T$  limit, especially for the tricky region close to  $n \sim 2/3$  where we know that  $d\mu/dT$  vanishes at high  $T$  from Eq. (31), hence, it is expedient to limit the high- $T$  region to  $T \lesssim 1$ . Having chosen such a cutoff, one can then explore the other physically interesting domain, and study the spectral functions at low  $T$  in the energy range  $|\omega| \lesssim D$ . This is a low-energy scale compared to the cutoffs, but already a very high-energy scale, in comparison to the physically interesting regimes  $|\omega| \lesssim \frac{D}{3}$

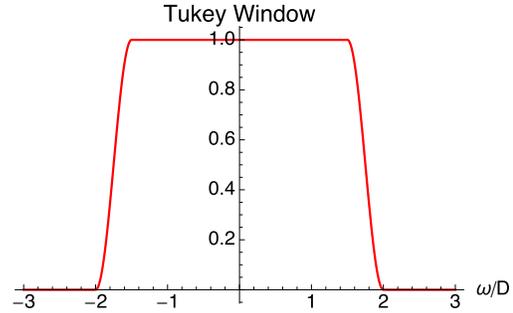


FIG. 3. Multiplication through the Tukey window  $\mathbf{W}_T(\omega)$  [Eq. (34)] is used for providing a cutoff in our scheme (33). It is applied only to the auxiliary local Green's function  $\rho_{\mathbf{g}L,m}(\omega)$ , while the physical spectral functions  $\rho_{\mathbf{g}L}(\omega)$  are unconstrained, apart from an overall window  $|\omega| \leq 5D$  used for numerical purposes. In this work, the upper cutoff used is  $\Omega_c^{(+)} = 2D$ , and the lower cutoff  $\Omega_c^{(-)} = 1.5D$ .

or even lower. We find below that the low- $T$  spectra indeed are better behaved with the cutoff. The low-energy results presented here are quite insensitive to the details of the choice for the cutoff, and hence one might be reasonably confident that the answers are not unduly biased by the choice made.

The method employed for imposing the high-energy cutoff was arrived at after some experimentation. We multiply the local spectral function (24) by a Tukey window function used in data filtering:

$$\hat{\rho}_{\mathbf{g}L,m}(\omega) = \frac{1}{\mathcal{V}} \rho_{\mathbf{g}L,m}(\omega) \mathbf{W}_T(\omega), \quad (33)$$

where the constant  $\mathcal{V}$  is found from the normalization condition  $\int \hat{\rho}_{\mathbf{g}L,0}(\omega) d\omega = 1$ . Here, the smooth Tukey window  $\mathbf{W}_T(\omega)$  is unity over the physically interesting, i.e., feature-rich frequency domain  $|\omega| \leq \Omega_c^{(-)}$ , where it starts falling off smoothly, and vanishing beyond the high-frequency cutoff  $|\omega| = \Omega_c^{(+)}$ . It is defined as a piecewise function (see Fig. 3)

$$\begin{aligned} \mathbf{W}_T(\omega) &= 1 \quad \text{for } \Omega_c^{(-)} \geq |\omega| \\ &= \frac{1}{2} \left( 1 + \sin \left\{ \pi/2 \frac{\Omega_c^{(+)} + \Omega_c^{(-)} - 2|\omega|}{\Omega_c^{(+)} - \Omega_c^{(-)}} \right\} \right) \\ &\quad \text{for } \Omega_c^{(+)} \geq |\omega| \geq \Omega_c^{(-)} \\ &= 0 \quad \text{for } |\omega| > \Omega_c^{(+)}. \end{aligned} \quad (34)$$

This procedure involves a *single rescaling*: after computing the local spectral functions  $\rho_{\mathbf{g}L,m}$  (with  $m = 0, 1$ ) from the self-energies as in Eq. (24), we multiply with  $\mathbf{W}_T$  and rescale as in Eq. (33) before sending the result back into the self-energy calculation in Eq. (28). Note that the prescription (33) involves the auxiliary local Green's function  $\mathbf{g}_{L,m}$  which is the basic building block in the theory. The cutoff is imposed *only* on  $\rho_{\mathbf{g}}$  in Eq. (24), and the other spectral functions are then computed by the unchanged equations (21)–(30).

We chose the parameters  $\Omega_c^{(+)} = 2D$ , and the lower cutoff  $\Omega_c^{(-)} = 1.5D$  after some experimentation. This choice of the cutoffs is in accord with the discussion above where we concluded  $c_0 \sim 2D$ . With this cutoff and rescaled auxiliary Green's function, the physical spectral function  $\rho_{\mathbf{g}}$  is computed as per the rules without any further assumptions. It typically

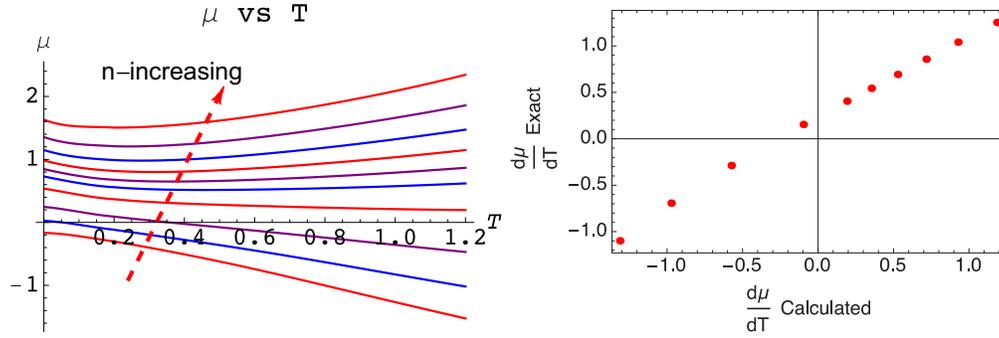


FIG. 4. Left: The chemical potential at particle densities  $n = 0.4, 0.5, 0.6, 0.7, 0.75, 0.775, 0.8, 0.825, 0.85, 0.875$  increasing from bottom to top. We confine ourselves to the limited regime  $T \leq 1.2$  since higher  $T$  requires further adjustment of the cutoffs. Note that within this regime, the  $\mu(T)$  curve turns around at a density around  $n \sim 0.7$ . For lower densities,  $\mu$  decreases monotonically with increasing  $T$ , whereas at higher densities we have a shallow minimum followed by a regime of rising  $\mu$ . This change of behavior is expected from Eq. (31), and has important physical consequence of changing the sign of the Kelvin thermopower for correlated matter [55]. Right: The slope  $d\mu/dT$  is calculated from the  $\mu(T)$  curves at  $T = 1$ , and contrasted with the exact values from Eq. (31). The points are taken from the same set of particle densities  $n$  as the figure on left, increasing from left to right. Since there is yet some curvature in the figures at left when  $T = 1$ , our procedure provides only a rough estimate. We note that these are in fair correspondence, especially at low hole density (see top right quadrant).

does extend to about  $4.5D$  or  $5D$  on the occupied side, but not beyond this scale. For numerical purposes, we also use an upper cutoff for the physical spectral function range as  $\sim 5D$ , this energy corresponds to  $\Omega_*$  in Fig. 2.

## V. RESULTS FOR CHEMICAL POTENTIAL, QUASIPARTICLE WEIGHT, SELF-ENERGY, AND SPECTRAL FUNCTIONS

### A. Chemical potential and quasiparticle weight $Z$

With the chosen cutoff, we examine the chemical potential as a function of density and  $T$  in Fig. 4. We observe in the left panel of Fig. 4 that the chosen cutoff provides a reasonable description of the  $\mu$  versus  $T$  curves at different densities. These exhibit an upturn between  $n = 0.6$  and  $0.7$  in the  $T$  domain that is computationally reliable within this scheme. The right panel of Fig. 4 shows that the slope  $d\mu/dT$  is also in reasonable agreement with the exact answer for this slope, apart from some error near the difficult regime of  $n \sim \frac{2}{3}$ . Here, we know from Eq. (31) that the slope is zero at high enough  $T$  and this causes problems of convergence.

We examine the various pieces adding up to the chemical potential in the right panel of Fig. 5. These curves also show that the Mott-Hubbard physics of the upturn of  $\mu(T)$  is enforced by the  $u_0$  term, it is thus quite crucial within this formalism. We also note that calculations without the cutoff lead to much larger values of  $u_0$ .

Overall, it seems that the results for  $\mu$  are quite reasonable in the hole-rich region  $n \geq 0.75$  (i.e.,  $\delta \leq 0.25$ ) with the global choice made, i.e., without requiring a fine tuning of the cutoffs with the density. We therefore proceed to use this for computing the spectral functions, and other physically interesting variables, also evaluated in the complementary low- $T$  region.

Turning to the main objective of this work of calculating the correct energy scale near the Mott limit, we display the computed  $Z$  versus the hole density  $\delta = 1 - n$  in the left panel of Fig. 6. It is interesting that the values obtained are significantly better than those reported in [6], we now find

$Z$  vanishes as  $\delta \rightarrow 0$ . The solid line gives the numerically exactly determined  $Z$  from DMFT, which is extremely well fit by  $Z \sim \delta^{1.39}$ . This latter behavior is noteworthy in that it vanishes faster than linear in  $\delta$ . The ‘‘mean field descriptions’’ involving slave auxiliary particles as well as the Brinkman-Rice theory [56] of the correlated metallic state give a linear  $Z \propto \delta$ . Therefore, this result indicates the need to account for fluctuations beyond the mean field description. It is interesting that the present calculation also gives a nonlinear behavior, with a slightly larger exponent than 1.39. We plan to return to a closer analytical study of this interesting result, obtained from the numerics of our solution.

### B. Self-energy and spectral functions at low $T$

We have also studied the quasiparticle decay rate at  $T \sim 0$ , defined for  $|\omega| \leq ZD$  through a Fermi liquid form with the expected particle-hole asymmetric correction [8]

$$-\text{Im}\Sigma(\omega) = \frac{\omega^2}{\Omega_0} \left(1 - \frac{\omega}{\Delta}\right), \quad (35)$$

whereby introducing two energies:  $\Omega_0$ , which determines the magnitude of  $\text{Im}\Sigma$  and  $\Delta$  the asymmetry scale. In [6] and also in [5], it was pointed out that  $\Omega_0$  varies like  $Z^2$  near the Mott insulating limit, leading to a scaling of the Green’s function frequency with  $Z$  at low energies. In this work, the  $\Omega_0$  is computed by averaging  $\text{Im}\Sigma(\omega)$  in the domain  $|\omega| \leq ZD$ . In the bottom right panel of Fig. 7, we show the variation of  $\Omega_0$  versus  $Z^2$  and in the inset with  $\delta^2$ . Since we have seen nonlinear corrections in  $Z$  as seen in Fig. 6, these two plots seem to support more closely the scaling of  $\Omega_0$  with  $Z^2$ , rather than  $\delta^2$  at the lowest  $\delta$ . It seems possible to improve the agreement by choosing a density-dependent cutoff, however, the global cutoff already achieves fair agreement.

In the top left panel of Fig. 7, we plot  $-\text{Im}\Sigma$  versus  $\omega/Z$  at different densities. As already noted in [6], these curves fall on top of each other quite well. The curves also exhibit particle-hole asymmetry as noted before [2,8]. This is exhibited by decomposing the  $\text{Im}\Sigma$  into symmetric and

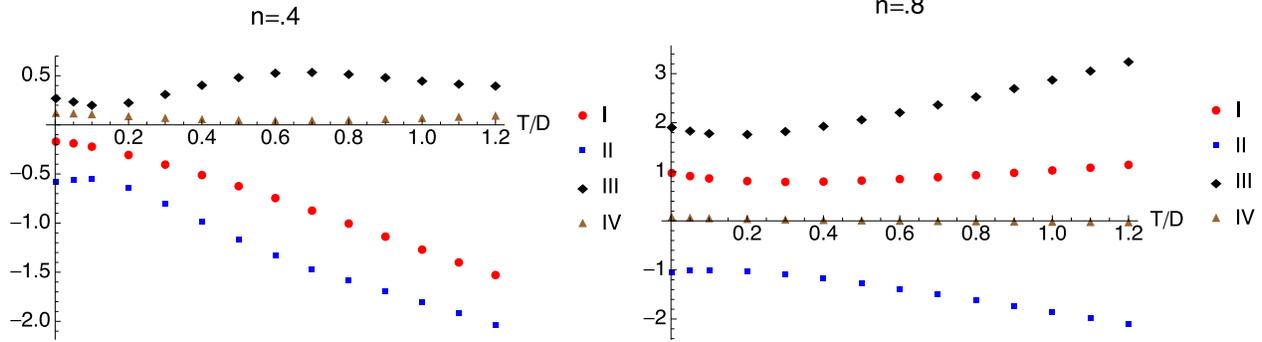


FIG. 5. The  $T$  dependence of the chemical potential  $\mu$  and its three additive contributions from Eq. (23) at two densities. The physical chemical potential  $\mu$  (I, red), the auxiliary part:  $\mu'$  (II, blue), the  $u_0$  contribution:  $(1/2 + n/4)u_0$  (III, purple), and the small part from the integral  $-\int f\rho_{gL,1}$  (IV, magenta). The observed upturn in  $\mu$  at high  $T$  for  $n = 0.8$ , reflecting the physics of Mott holes near half-filling, is predominantly due to the upturn of the second chemical potential  $u_0$ . Its growth, in turn, causes the numerical issues requiring the implementation of a cutoff in this work.

antisymmetric components in the the top right and bottom left panels. The antisymmetric part can be analyzed to read off the energy scale  $\Delta$  in Eq. (35). We find that  $\Delta$  is proportional to  $Z$  again, but with a weak density-dependent correction:

$$\Delta(\delta) = Z(\delta)\{3.38 - 15.6\delta + 27.1\delta^2\}. \quad (36)$$

The region beyond the straight line is captured on average, by extending Eq. (35) to

$$-\text{Im}\Sigma(\omega) = \frac{\omega^2}{\Omega_0} \left( 1 - \frac{\omega}{\Delta\sqrt{1 + 2\omega^2/Z^2}} \right). \quad (37)$$

This expression is potentially useful for phenomenological extensions of the theory.

In Figs. 8 and 9, we display the raw unscaled spectral functions and the imaginary part of the self-energy for various physical parameters. In Fig. 8, the low- $T$  spectra are shown at different densities. Note that the significant range of  $\omega$  where the spectral functions and self-energy vary, shrinks rapidly

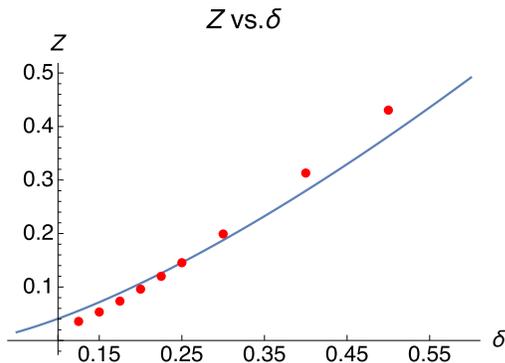


FIG. 6. The computed quasiparticle weight  $Z$  (dots) versus the hole density  $\delta = 1 - n$ , compared with the exact numerical results from DMFT ([6] solid curve), which fits very well to the formula  $Z \sim \delta^{1.39}$ . We see that the present scheme accounts well the suppression of  $Z$  near  $\delta \sim 0$ , even reproducing nonlinear vanishing near the Mott limit seen in [6]. This nonlinear feature goes beyond the predictions of both slave-boson mean field and Brinkman-Rice theory [56], and signifies an important correction to the mean field behavior.

with increasing  $n$ ; this is indirectly a reflection of variation of the  $Z$  with density in Fig. 6 since the scale of variation of  $\Sigma$  is set by  $Z$ . We also note that the spectral asymmetry in  $\text{Im}\Sigma$  is very clearly visible here.

### C. Temperature variation of the self-energy and spectral functions

In Fig. 9, we display the  $T$  dependence of the spectral function and the self-energy. One of the advantages of our computational scheme is the ease with which  $T$  variation can be computed. We are thus able to obtain easily the crossover from a coherent (extremely correlated) Fermi liquid regime at low  $T$  to an incoherent nondegenerate correlated state. The spectral function peaks rapidly broaden and shift as the temperature is increased. We also note that the Fermi coherence, signaled by a small magnitude of  $\text{Im}\Sigma$  at small  $\omega$ , is rapidly lost on heating, leading to a flat and structureless function. A comparison of the curves at  $n = 0.85$  and  $0.875$  shows that in this range of densities, where the  $Z$  is already very small, the effective Fermi temperature is also diminished since the same (small) variation of  $T$  produces a relatively large change in the damping.

## VI. TEMPERATURE DEPENDENCE OF RESISTIVITY AND RELATED QUANTITIES

Perhaps the single most important characterization of a theory is via the resistivity. It is a notoriously hard object to calculate reliably, and yet one that is most sensitive to the lowest-energy excitations of the system. Since we have argued that the present version of ECFL captures the low-energy excitations of the electron, it is useful to examine its results for resistivity for the  $t$ - $J$  model in infinite dimensions, or equivalently the  $U = \infty$  Hubbard model. The resistivity has been calculated numerically from DMFT quite recently in two papers [42,43], and hence it is of interest to see how our analytical calculation compares with these exact results.

We start with the Kubo expression for resistivity, with the vertex correction thrown out, thanks to the simplification

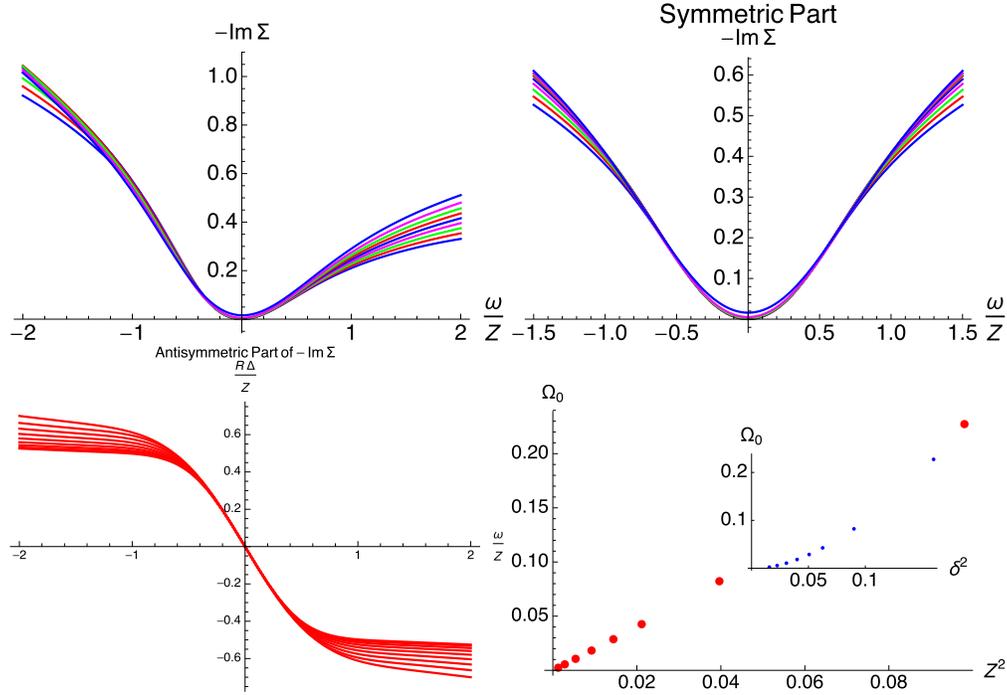


FIG. 7. Top left:  $-\text{Im}\Sigma(\omega)$  versus  $\omega/Z$  at several densities  $n = 0.7, 0.725, 0.75, 0.775, 0.8, 0.825, 0.85, 0.875, 0.9$  from bottom to top. We see that the frequency dependence scales well with  $Z$ , with better behavior on the occupied side  $\omega \leq 0$ . Top right: The symmetrized function  $[-\Sigma''(\omega) - \Sigma''(-\omega)]/2$  exhibits the quadratic behavior at  $\omega \sim 0$  expected from a Fermi liquid. Bottom left: The antisymmetric part is defined as  $R = [\Sigma''(\omega) - \Sigma''(-\omega)]/[\Sigma''(\omega) + \Sigma''(-\omega)]$ , so that if we assume Eq. (35) then  $R = -\omega/\Delta$ . We show the computed  $R$  multiplied by  $\Delta/Z$  at the above densities versus  $\omega/Z$ , with  $n = 0.9$  at the top and  $n = 0.7$  at the bottom for  $\omega \leq 0$ . These collapse to a straight line with slope  $-1$  in the range  $|\omega| \leq Z$ , provided we allow for an additional mild density dependence of the ratio  $\Delta/Z$ , as in Eq. (36). Bottom right: The energy scale  $\Omega_0$  [Eq. (35)] determining the magnitude of the  $\text{Im}\Sigma$  at  $T = 0$  is shown versus  $Z^2$ , and in the inset versus the hole density  $\delta^2$ . Here,  $\Omega_0$  is seen to scale better with  $Z^2$  rather than with  $\delta^2$ .

arising from  $d \rightarrow \infty$ :

$$\sigma_{DC} = \frac{2\pi\hbar e^2}{V} \sum_k (v_k^x)^2 \int d\omega (-\partial f/\partial\omega) \rho_G^2(\epsilon_k, \omega), \quad (38)$$

where the band velocity is given as  $\hbar v_k^x = \partial\epsilon_k/\partial k_x$ . We wrap the velocity into a useful function

$$\begin{aligned} \Phi(\epsilon) &= \frac{1}{a_0} \frac{1}{N_s} \sum_k \delta(\epsilon - \epsilon_k) (v_k^x)^2 / a_0^2 \\ &= \frac{1}{a_0} D(\epsilon) \left\langle \frac{(v_k^x)^2}{a_0^2} \right\rangle_{\epsilon_k=\epsilon}, \end{aligned} \quad (39)$$

where  $a_0$  is the lattice constant in the hypercubic lattice, and  $N_s$  the number of sites and we use the Bethe lattice semicircular density of states  $D(\epsilon) = \frac{2}{\pi D} \sqrt{1 - \epsilon^2/D^2}$ . Deng *et al.* [42,57] calculate that

$$\frac{\Phi(\epsilon)}{\Phi(0)} = \Theta(1 - \epsilon^2/D^2) \sqrt[3/2]{1 - \epsilon^2/D^2}, \quad (40)$$

where  $\Phi(0)$  is absorbed into a constant  $\sigma_0 = e^2\hbar\Phi(0)/D$ , which is identified with the Ioffe-Regel-Mott conductivity. With this choice of the vertex we obtain

$$\sigma_{DC} = \sigma_0 2\pi D \iint d\epsilon d\omega (-\partial f/\partial\omega) \left( \frac{\Phi(\epsilon)}{\Phi(0)} \right) \rho_G^2(\epsilon, \omega). \quad (41)$$

We write the (inverse) Green's function at real  $\omega$  as

$$\mathcal{G}_{\pm}^{-1}(\epsilon, \omega) = A(\omega) - \epsilon \pm iB(\omega), \quad (42)$$

where the retarded case corresponds to  $\mathcal{G}_+$ , and

$$\begin{aligned} A(\omega, T) &= \omega + \mu - \text{Re} \Sigma(\omega, T), \\ B(\omega, T) &= \pi \rho_{\Sigma}(\omega, T) = -\text{Im} \Sigma(\omega, T), \end{aligned} \quad (43)$$

and  $\Sigma$  is the Dyson self-energy. Setting  $D = 1$  and using the identities  $\rho_G = i/(2\pi)(\mathcal{G}_- - \mathcal{G}_+)$  and  $\mathcal{G}_{\pm}^2 = \partial_{\epsilon}\mathcal{G}_{\pm}$ , and further integrating by parts over  $\epsilon$  we obtain

$$\begin{aligned} \sigma &= \sigma_0 \int d\omega (-\partial f/\partial\omega) \xi(\omega), \\ \xi(\omega) &= \frac{1}{2\pi} \int d\epsilon \left\{ \frac{i}{B} (\mathcal{G}_+ - \mathcal{G}_-) \frac{\Phi(\epsilon)}{\Phi(0)} + (\mathcal{G}_+ + \mathcal{G}_-) \frac{\Phi'(\epsilon)}{\Phi(0)} \right\}. \end{aligned} \quad (44)$$

Using the explicit form of  $\Phi$  and  $\mathcal{G}_{\pm}$ , we reexpress  $\xi$  exactly as

$$\xi(\omega) = \frac{1}{\pi} \int_{-1}^1 d\epsilon \sqrt{1 - \epsilon^2} \frac{1 - 3\epsilon A + 2\epsilon^2}{B^2 + (A - \epsilon)^2}. \quad (45)$$

The evaluation of this integral is straightforward, and leads to a cumbersome result. A simple answer for the leading behavior when  $B \ll 1$  can be found, provided  $(A - \epsilon)$  goes through zero in the interval of integration. Since we will see that  $|A| \ll 1$  for

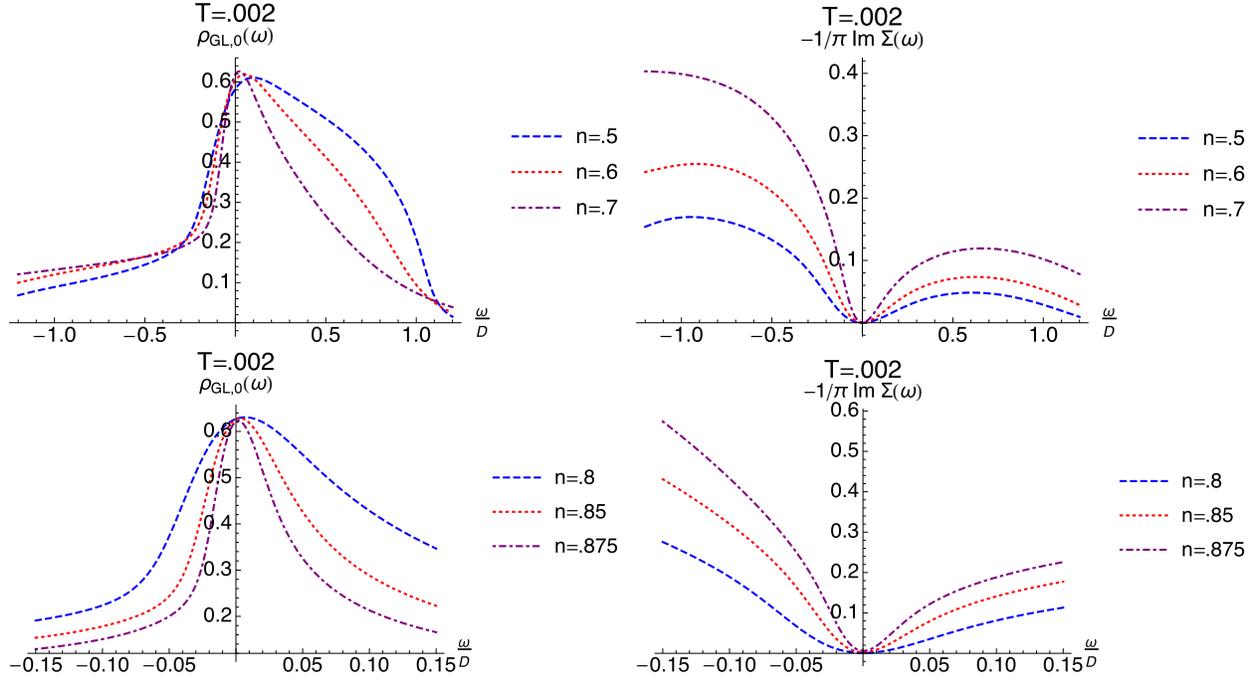


FIG. 8. The two figures on the left display the physical local spectral function  $\rho_{GL,0} = -\frac{1}{\pi} \text{Im} \mathcal{G}_{Loc,0}(\omega + i0^+)$  from Eq. (25), and the two figures on the right show the Dyson self-energy  $-\frac{1}{\pi} \text{Im} \Sigma(\omega)$ , plotted against the frequency  $\omega/D$ . The figures are at low  $T$  for the six indicated values of the density, and display a region that is somewhat greater than the one, where it is expected to be reliable  $|\omega| \lesssim ZD$ . One sees a correlation between the quasiparticle weight  $Z$  (Fig. 6) and the scale of variation of the decay rate. Densities  $n > 0.875$  have larger errors in  $Z$  compared to the exact DMFT results (see Fig. 6), and therefore not shown. However, it is easy to picture them at low  $\omega$ , using the observation that scaling  $\omega$  with  $Z$  collapses  $\Sigma''$ .

all temperatures and frequencies of interest ( $\omega \sim 0, \frac{T}{D} \lesssim 0.3$ ), this will always be the case. We may write  $\epsilon = A + B \tan(\theta)$ , retain the leading terms for small  $B$ , and set  $B \rightarrow 0$  in the remainder. With this, we obtain the asymptotic approximation

$$\lim_{B \ll 1} \xi(\omega) \sim \frac{[1 - A^2(\omega)]^{3/2}}{B(\omega)} \Theta[1 - A^2(\omega)]. \quad (46)$$

In Fig. 10, we use Eq. (46) to plot  $\frac{\rho_{dc}}{\rho_0}$  versus  $\frac{T}{D}$  for  $0.75 \leq n \leq 0.85$ , where  $\rho_0 = \frac{1}{\sigma_0}$ . These resistivity curves have both the same shape and the same scale as those found through DMFT (see Fig. (1a) of Ref. [42]). We find a Fermi liquid regime [ $\frac{\rho_{dc}}{\rho_0} \propto (\frac{T}{D})^2$ ] for  $0 < T < T_{FL}$ , where  $T_{FL} = (cD) \times Z(T=0)$ , and  $c \approx 0.05$ . Furthermore,  $\frac{\rho_{dc}}{\rho_0}$  is a function of  $\frac{T}{DZ(T=0)}$  for  $T \lesssim 2T_{FL}$  [Fig. 10(c)]. An important scale emphasized in DMFT studies [42,43] is the Brinkman-Rice scale ( $T_{BR} = D\delta$ ), which is the renormalized bandwidth of the quasiparticles rather than the quasiparticle weight. Since  $Z(T=0) \propto \delta^\alpha$ , with  $\alpha > 1$ , the Fermi liquid scale is contained within the Brinkman-Rice scale, and is smaller than the latter by some power of  $\delta$ . As  $T$  is increased above  $T_{FL}$ , the Fermi liquid regime is followed by a linear regime for  $T_{FL} < T \lesssim 0.01D$ . In Fig. 10(a), the Fermi liquid regime is tracked using the blue dashed parabola, while the linear regime is tracked using the magenta dashed line. Finally, this linear regime connects continuously to a second linear regime, existing for  $T \gtrsim 0.07D$  [displayed in Fig. 10(b)].

We now analyze more closely the low-temperature regime ( $T \lesssim 0.01D$ ). For this range of temperatures, the Sommerfeld

expansion can be applied to Eq. (44). To leading order [ $-\partial f/\partial \omega = \delta(\omega)$ ], and using Eq. (46), this gives

$$\rho_{DC} \sim \rho_0 \frac{-\text{Im} \Sigma(0, T)}{[1 - \{\mu - \text{Re} \Sigma(0, T)\}^2]^{3/2}}. \quad (47)$$

The constituent objects  $-\text{Im} \Sigma(0, T)$  and  $A(0, T)$  are plotted along with  $Z(T)$  in the relevant temperature range in Fig. 11. We first examine  $A(0, T) = \mu - \text{Re} \Sigma(0, T)$ , displayed in Fig. 11(c). For  $T_{FL} \lesssim T \lesssim 0.01D$ , it is linear, as tracked by the dashed blue line. We also notice that  $A^2(0, T) \ll 1$ , and can therefore be neglected in Eq. (47). Equation (47) then implies that the resistivity is proportional to  $[-\text{Im} \Sigma(0, T)]$  in this low-temperature range. Accordingly, in Fig. 11(a), we see that  $[-\text{Im} \Sigma(0, T)]$  is quadratic for  $T \lesssim T_{FL}$  (tracked by the blue dashed parabola) and linear for  $T_{FL} \lesssim T \lesssim 0.01D$  (tracked by the magenta dashed line). Finally, in Fig. 11(b), we see that  $Z(T)$  is approximately constant for  $T \lesssim T_{FL}$ , and grows linearly for  $T_{FL} \lesssim T \lesssim 0.01D$ , with a slope on the order of the bandwidth (tracked by the magenta dashed line). The blue dashed curve tracks the functional form discussed below, which approximates  $Z(T)$  very well for  $T \gtrsim T_{FL}$ . As emphasized in [43], the temperature dependence of  $[-\text{Im} \Sigma(0, T)]$  and  $Z(T)$  lead to a quasiparticle scattering rate, defined as  $[-\text{Im} \Sigma(0, T)] \times Z(T)$ , which is quadratic well above  $T_{FL}$ .

In Fig. 12, we plot the temperature dependence of these objects in a broader temperature range. In Fig. 12(c), the blue dashed line indicates the presence of a second linear regime in  $A(T)$  (with a slope slightly smaller than the first), meeting the latter at a kink at  $T \approx 0.01D$ . Figure 12(a) shows that for

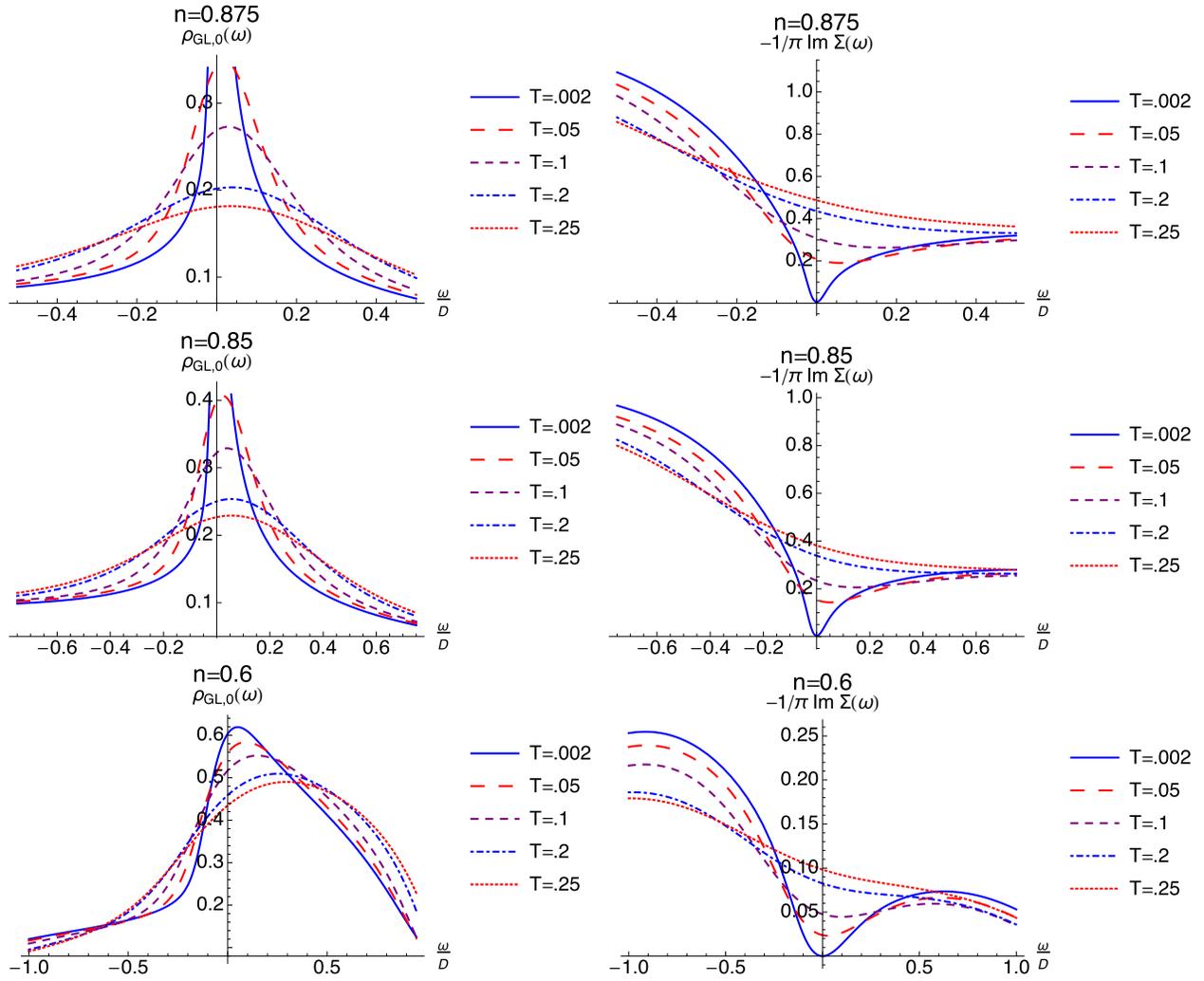


FIG. 9. The temperature variation with the frequency  $\omega/D$ , of the spectral function  $\rho_{GL,0}$  on the left and the Dyson self-energy  $-\frac{1}{\pi}\text{Im}\Sigma$  on the right, at density  $n = 0.875$  (top),  $n = 0.85$  (middle), and at  $n = 0.6$  (bottom). With increasing  $T$  we note the rapid broadening and shifting of  $\rho_{GL,0}$ . Here,  $-\frac{1}{\pi}\text{Im}\Sigma$  displays a rapid destruction of the coherent Fermi liquid behavior observed at the lowest  $T$ , by the filling up of the minimum at  $\omega = 0$ . Comparing the top two sets shows that at the lowest hole density, a small change in  $T$  has a large effect, due to the low effective Fermi temperature. We also observe here, as well as in Fig. 8, that  $-\frac{1}{\pi}\text{Im}\Sigma$  has a strong asymmetric correction to the quadratic  $\omega$  dependence of the standard Fermi liquid, as highlighted in the bottom left panel of Fig. 7. This is in accord with one of the basic analytical predictions of the ECFL theory, and also is found in the DMFT results.

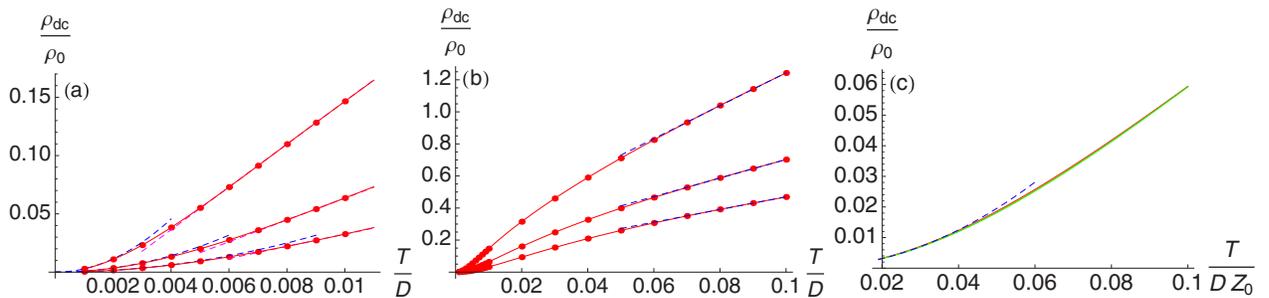


FIG. 10. (a), (b)  $\frac{\rho_{dc}}{\rho_0}$  vs  $\frac{T}{D}$  for  $n = 0.75, 0.8, 0.85$  from bottom to top. In (a), the blue dashed parabola tracks the FL regime  $0 < T < T_{FL}$  where  $\frac{\rho_{dc}}{\rho_0} \propto (\frac{T}{D})^2$ . The magenta dashed line tracks the first linear regime  $T_{FL} < T \lesssim 0.01D$ . In (b), the blue dashed line tracks the second linear regime  $T \gtrsim 0.07D$ . (c)  $\frac{\rho_{dc}}{\rho_0}$  vs  $\frac{T}{DZ(T=0)}$  for  $n = 0.75, 0.8, 0.85$  (red, orange, green). The blue dashed parabola tracks the Fermi liquid regime, demonstrating that  $T_{FL} = (c D) \times Z(T = 0)$ , with  $c \approx 0.05$ , and that  $\frac{\rho_{dc}}{\rho_0}$  is a function of  $\frac{T}{DZ(T=0)}$  for  $T \lesssim 2T_{FL}$ .

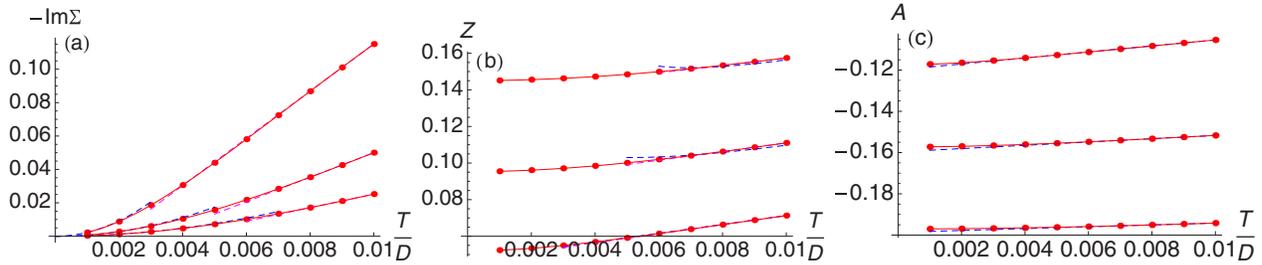


FIG. 11. (a)  $[-\text{Im}\Sigma(0, T)]$  vs  $\frac{T}{D}$  for  $n = 0.75, 0.8, 0.85$  from bottom to top.  $[-\text{Im}\Sigma(0, T)]$  is quadratic for  $T \lesssim T_{\text{FL}}$  (tracked by the blue dashed parabola) and linear for  $T_{\text{FL}} \lesssim T \lesssim 0.01D$  (tracked by the magenta dashed line). (b)  $Z(T)$  vs  $\frac{T}{D}$  for  $n = 0.75, 0.8, 0.85$  from top to bottom.  $Z(T)$  is approximately constant for  $T \lesssim T_{\text{FL}}$ , and grows linearly for  $T_{\text{FL}} \lesssim T \lesssim 0.01D$ , with a slope on the order of the bandwidth (tracked by the magenta dashed line). The blue dashed curve is the fit to the functional form  $Z(T) = \sqrt{\frac{1+aT+bT^2}{c+dT}}$  using a broader range of temperatures than the one shown here [Fig. 12(b)]. This form works well for  $T \gtrsim T_{\text{FL}}$ . (c)  $A(0, T) = \mu(T) - \text{Re} \Sigma(0, T)$  vs  $\frac{T}{D}$  for  $n = 0.75, 0.8, 0.85$  from bottom to top. For  $T_{\text{FL}} \lesssim T \lesssim 0.01D$ , it is linear, as tracked by the dashed blue line.

$T > 0.01D$ ,  $[-\text{Im}\Sigma(0, T)]$  continues to grow, until it finally begins to saturate at higher temperatures. Finally, in Fig. 12(b), we fit  $Z(T)$  to the functional form  $Z(T) = \sqrt{\frac{1+aT+bT^2}{c+dT}}$ , tracked by the blue dashed curve. This form works well for  $T \gtrsim T_{\text{FL}}$ . For  $T \lesssim 0.01D$ , it reproduces the behavior shown in Fig. 11(b), while for  $T \gtrsim 0.01D$ , it is consistent with the behavior  $Z(T) \propto \sqrt{T}$ . Therefore,  $Z^2(T)$  is linear in  $T$  over a very wide temperature range starting with  $T \approx 0.01D$ .

## VII. CONCLUSIONS

In this work, we have presented an analytical calculation of properties of the  $t$ - $J$  model in infinite dimensions, and shown that it provides a quantitative description of variables known from exact numerical work in [42,43]. The results include the quasiparticle weight, the self-energies, and spectral functions with particle-hole asymmetry that have been argued to be characteristic of very strong correlation [6,8]. Finally, we also give a good account of the temperature variation of resistivity. Results with the present technique at high  $T$  are less reliable and are not presented. In the low- to intermediate- $T$  results reported here, we reproduce the main features of the exact DMFT calculations, including a narrow regime with quadratic  $T$  dependence followed by two distinct linear  $T$ -dependent regimes. We are further able to identify the origin of these regimes in terms of the parameters of the theory.

The  $t$ - $J$  model studied here contains two essential ingredients of strong correlations: the physics of Gutzwiller projection to the subspace of single occupancy, and the physics of the superexchange. The first is captured in the present scheme, while the second is lost since we limit the study to infinite dimension for the purpose of benchmarking against known exact results. The scheme by itself has no intrinsic limitations to the case studied, and is generalizable to finite dimensions as well as finite superexchange. Thus, it may be expected to yield interesting results in lower dimensions, including transitions between different broken-symmetry states. Such calculations are currently underway.

In this work, we have discussed the characteristics of the resulting ECFL state. The state reported here is Fermi liquid like, but only so at a surprisingly low temperature. Upon minimal warming, this state devolves into one exhibiting linear resistivity. Our calculation yields a reduction of the effective Fermi temperature, due to extreme correlations, that far exceeds the expectations [58] based on a simple estimate  $T_F^{\text{eff}} \sim \delta T_F$ .

Within the terms of its limitations of  $d \rightarrow \infty$  and  $J = 0$ , this work provides useful insights. At the density  $n \sim 0.85$  relevant for cuprate superconductors, we obtain a state displaying linear resistivity for  $T$  beyond  $\sim 45$  K as seen in Fig. 1. A similar onset of linearity occurs at a slightly higher  $T$  within DMFT, the difference is due to our  $Z$  (from Fig. 6) being about half of the exact value. If we imagine that the effects of reduced dimensionality and nonzero  $J$  can stabilize

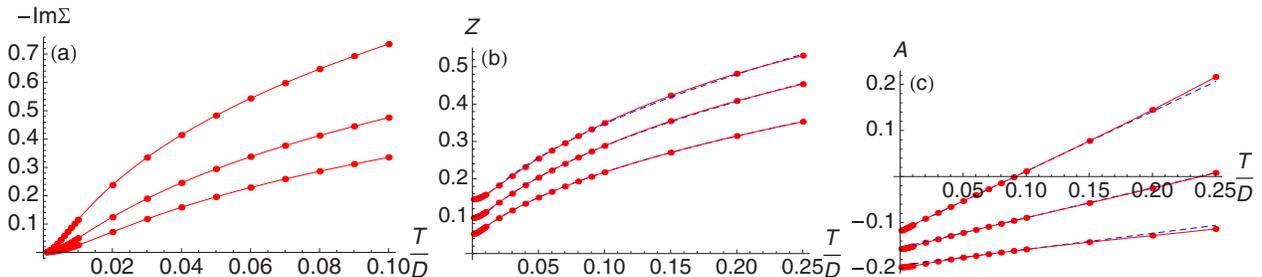


FIG. 12. Same plots as in Fig. 11 over a broader range of temperatures. (a)  $[-\text{Im}\Sigma(0, T)]$  continues to grow as  $T$  is increased beyond  $0.01D$ , until it finally begins to saturate at higher temperatures. (b) The blue dashed curve is the fit to the functional form  $Z(T) = \sqrt{\frac{1+aT+bT^2}{c+dT}}$ , which works well for  $T \gtrsim T_{\text{FL}}$ . For  $T \gtrsim 0.01D$ ,  $Z(T) \propto \sqrt{T}$ . (c) The blue dashed line tracks the second linear regime in  $A(T)$  (with a slope slightly smaller than the first) for  $T \gtrsim 0.01D$ .

this smaller onset scale, then the possibility of observing the asymptotic  $T^2$  resistivity of a Fermi liquid would become remote. Thus, the quadratic behavior, so essential for making a *formal distinction* between Fermi liquids and the elusive non-Fermi liquids [59], could be rendered unobservable in practice as well as divested of any essential difference.

## ACKNOWLEDGMENTS

We thank A. Georges for useful discussions about resistivity in DMFT. The work at UCSC was supported by the U. S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award No. DE-FG02-06ER46319.

- 
- [1] P. W. Anderson, *Science* **235**, 1196 (1987).
- [2] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011); **108**, 029702 (2012).
- [3] B. S. Shastry, *Ann. Phys. (NY)* **343**, 164 (2014).
- [4] F. J. Dyson, *Phys. Rev.* **102**, 1230 (1956); S. V. Maleev, *Zh. Eksp. Teor. Fiz.* **33**, 1010 (1957) [*Sov. Phys.-JETP* **6**, 776 (1958)]; A. B. Harris, D. Kumar, B. I. Halperin, and P. C. Hohenberg, *Phys. Rev. B* **3**, 961 (1971).
- [5] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *Phys. Rev. B* **88**, 205108 (2013).
- [6] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013).
- [7] G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).
- [8] B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).
- [9] K. G. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975).
- [10] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, *Phys. Rev. B* **21**, 1044 (1980).
- [11] T. A. Costi, J. Kroha, and P. Wölfle, *Phys. Rev. B* **53**, 1850 (1996).
- [12] T. A. Costi and A. C. Hewson, *J. Phys. C* **5**, L361 (1993).
- [13] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [14] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989).
- [15] Y. Kuramoto and T. Watanabe, *Physica B (Amsterdam)* **148**, 246 (1987); O. Sakai and Y. Kuramoto, *Solid State Commun.* **89**, 307 (1994).
- [16] D. Nicoletti, O. Limaj, P. Calvani, G. Rohringer, A. Toschi, G. Sangiovanni, M. Capone, K. Held, S. Ono, Yoichi Ando, and S. Lupi, *Phys. Rev. Lett.* **105**, 077002 (2010).
- [17] G. R. Boyd, V. Zlatic, and J. K. Freericks, *Phys. Rev. B* **91**, 075118 (2015).
- [18] R. Bulla, *Phys. Rev. Lett.* **83**, 136 (1999).
- [19] R. Chitra and G. Kotliar, *Phys. Rev. Lett.* **83**, 2386 (1999).
- [20] Marcelo J. Rozenberg, R. Chitra, and Gabriel Kotliar, *Phys. Rev. Lett.* **83**, 3498 (1999).
- [21] K. Held and D. Vollhardt, *Phys. Rev. Lett.* **84**, 5168 (2000).
- [22] Massimo Capone, Michele Fabrizio, Claudio Castellani, and Erio Tosatti, *Phys. Rev. Lett.* **93**, 047001 (2004).
- [23] D. J. Garcia, Karen Hallberg, and Marcelo J. Rozenberg, *Phys. Rev. Lett.* **93**, 246403 (2004).
- [24] Krzysztof Byczuk, Walter Hofstetter, and Dieter Vollhardt, *Phys. Rev. Lett.* **94**, 056404 (2005).
- [25] M. Keller, W. Metzner, and U. Schollwock, *Phys. Rev. Lett.* **86**, 4612 (2001).
- [26] A. I. Lichtenstein, M. I. Katsnelson, and G. Kotliar, *Phys. Rev. Lett.* **87**, 067205 (2001).
- [27] Gabriel Kotliar, Sergej Y. Savrasov, Gunnar Pálsson, and Giulio Biroli, *Phys. Rev. Lett.* **87**, 186401 (2001).
- [28] A. Chattopadhyay, S. Das Sarma, and A. J. Millis, *Phys. Rev. Lett.* **87**, 227202 (2001).
- [29] Antoine Georges, Rahul Siddharthan, and Serge Florens, *Phys. Rev. Lett.* **87**, 277203 (2001).
- [30] B. Kyung and A.-M. S. Tremblay, *Phys. Rev. Lett.* **97**, 046402 (2006).
- [31] Junya Otsuki, Hiroaki Kusunose, and Yoshio Kuramoto, *Phys. Rev. Lett.* **102**, 017202 (2009).
- [32] Michael Potthoff, *Phys. Rev. B* **64**, 165114 (2001).
- [33] K.-S. Chen, Z. Y. Meng, S.-X. Yang, T. Pruschke, J. Moreno, and M. Jarrell, *Phys. Rev. B* **88**, 245110 (2013).
- [34] M. H. Hettler, M. Mukherjee, M. Jarrell, and H. R. Krishna-murthy, *Phys. Rev. B* **61**, 12739 (2000).
- [35] M. Jarrell and H. R. Krishnamurthy, *Phys. Rev. B* **63**, 125102 (2001).
- [36] Luigi Amico and Vittorio Penna, *Phys. Rev. Lett.* **80**, 2189 (1998).
- [37] Marcus Fleck, Alexander I. Liechtenstein, Andrzej M. Oleś, Lars Hedin, and Vladimir I. Anisimov, *Phys. Rev. Lett.* **80**, 2393 (1998).
- [38] Gunnar Pálsson and Gabriel Kotliar, *Phys. Rev. Lett.* **80**, 4775 (1998).
- [39] A. J. Millis, Jun Hu, and S. Das Sarma, *Phys. Rev. Lett.* **82**, 2354 (1999).
- [40] For example, in [6] we comment on the insufficient height of the peak in the imaginary part of self-energy on the occupied side, and the absent “nipple” in the ECFL spectrum. We also note that for a few cases, the high-energy  $\omega \gg ZD$  (actually for  $\omega > 1.5D$ ) *unoccupied states* show unphysical negative weights. Their occasional appearance seems difficult to avoid completely at the current preliminary stage of development in this unconventional theory. In the authors’ view, a mitigating factor is that such errors arise at very high excitation energies, a regime that is not easily probed, at least within ARPES. This regime is also irrelevant for the number sum rule, where only occupied states are counted. We should therefore be careful in relying on the results for high-energy unoccupied states beyond say  $\omega \geq D$ .
- [41] A. Khurana, *Phys. Rev. Lett.* **64**, 1990 (1990).
- [42] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [43] W. Xu, K. Haule, and G. Kotliar, *Phys. Rev. Lett.* **111**, 036401 (2013).
- [44] The term *caparison*, defined in [www.merriam-webster.com/dictionary](http://www.merriam-webster.com/dictionary) as *rich clothing* or *adornment*, is used to stress the idea that the physical Green’s function  $\mathcal{G}$  acquires a second dressing (renormalization) on top of the

standard dressing provided the Dyson-type self-energy of the auxiliary Green's function  $\mathbf{g}$ , namely, by  $\Phi$ .

- [45] E. Perepelitsky and B. S. Shastry, *Ann. Phys. (NY)* **338**, 283 (2013).
- [46] E. Perepelitsky and B. S. Shastry, *Ann. Phys. (NY)* **357**, 1 (2015).
- [47] B. S. Shastry, *Phys. Rev. B* **87**, 125124 (2013).
- [48] These equations can be found in [46] [Eqs. (56)–(61)] by setting the exchange  $J \rightarrow 0$ . In particular, Eq. (65a) gives the quoted  $\Psi_{[1]}$ , Eq. (66a) gives the quoted  $\chi_{[0]}$ , and Eqs. (66b) and (67a) give the quoted  $\chi_{[1]}$ . We stress that in this version, the sum rule for  $\gamma$  is used exactly and hence we retain only the first term (64a) and abandon the higher-order  $\lambda$  expansion terms [Eqs. (64b)–(64d), etc.].
- [49] D. Hansen and B. S. Shastry, *Phys. Rev. B* **87**, 245101 (2013).
- [50] N. Prokof'ev and B. Svistunov, *Phys. Rev. Lett.* **99**, 250201 (2007).
- [51] E. Kozik, K. Van Houcke, E. Gull, L. Pollet, N. Prokof'ev, B. Svistunov and M. Troyer, *Europhys. Lett.* **90**, 10004 (2010).
- [52] N. V. Prokofiev, B. V. Svistunov, and I. S. Tupitsyn, *Phys. Lett. A* **238**, 253 (1993).
- [53] W.-C. Lee and P. W. Phillips, *Phys. Rev. B* **84**, 115101 (2011).
- [54] The other half is clearly in the occupied region  $\omega < 0$ .
- [55] M. R. Peterson and B. S. Shastry, *Phys. Rev. B* **82**, 195105 (2010); B. S. Shastry, *Rep. Prog. Phys.* **72**, 016501 (2009).
- [56] W. F. Brinkman and T. M. Rice, *Phys. Rev. B* **2**, 4302 (1970).
- [57] The value  $\Phi(0)$  can be calculated as

$$\Phi(0) = \frac{D^3}{2\pi(a_0 d \hbar^2)} \frac{\varepsilon_F}{(D^2 - \varepsilon_F^2)^{3/2}},$$

and yields an estimate of the IRM conductivity scale

$$\sigma_0 = \left( \frac{e^2}{h} \frac{1}{a_0 d} \right) \frac{D^2 \varepsilon_F}{(D^2 - \varepsilon_F^2)^{3/2}}.$$

The proportionality to inverse dimension is expected since the current flows in one of  $d$  directions.

- [58] Using  $C \equiv k_B \frac{T}{T_F}$  for a band of noninteracting electrons, where  $C$  is the specific heat, we find that  $T_F = \frac{3}{4\pi} \frac{D}{k_B}$ . With the half-bandwidth  $D \sim 12\,000$  K (1 eV) a naive estimate using  $n = 0.85$  and a crude formula  $T_F^{\text{eff}} = \delta T_F$  would yield  $T_F^{\text{eff}} \sim 430$  K. In DMFT [42], this scale is reduced to 90 K, while in the present theory, it is reduced even further down to 30 K.
- [59] R. B. Laughlin, *Phys. Rev. Lett.* **112**, 017004 (2014).

## Origin of kinks in the energy dispersion of strongly correlated matter

Kazue Matsuyama,<sup>1</sup> Edward Perepelitsky,<sup>2,3</sup> and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>Physics Department, University of California, Santa Cruz, California 95064, USA

<sup>2</sup>Centre de Physique Théorique, École Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France

<sup>3</sup>Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France

(Received 22 January 2017; published 20 April 2017)

We investigate the origin of ubiquitous low-energy kinks found in angle-resolved photoemission experiments in a variety of correlated matter. Such kinks are unexpected from weakly interacting electrons and hence identifying their origin should lead to fundamental insights in strongly correlated matter. We devise a protocol for extracting the kink momentum and energy from the experimental data which relies solely on the two asymptotic tangents of each dispersion curve, away from the feature itself. It is thereby insensitive to the different shapes of the kinks as seen in experiments. The body of available data are then analyzed using this method. We proceed to discuss two alternate theoretical explanations of the origin of the kinks. Some theoretical proposals invoke local bosonic excitations (Einstein phonons or other modes with spin or charge character), located exactly at the energy of observed kinks, leading to a *momentum-independent* self-energy of the electrons. A recent alternate is the theory of extremely correlated Fermi liquids (ECFL). This theory predicts kinks in the dispersion arising from a *momentum-dependent* self-energy of correlated electrons. We present the essential results from both classes of theories, and identify experimental features that can help distinguish between the two mechanisms. The ECFL theory is found to be consistent with currently available data on kinks in the nodal direction of cuprate superconductors, but conclusive tests require higher-resolution energy distribution curve data.

DOI: [10.1103/PhysRevB.95.165435](https://doi.org/10.1103/PhysRevB.95.165435)

### I. INTRODUCTION

High-precision measurements of electronic spectral dispersions have been possible in recent years, thanks to the impressive enhancement of the experimental resolution in the angle-resolved photoemission spectroscopy (ARPES). This technique measures the single-electron spectral function  $A(\vec{k}, \omega)$  multiplied by the Fermi occupation function; it can be scanned at either fixed  $\vec{k}$  as a function of  $\omega$  or at fixed  $\omega$  as a function of  $\vec{k}$ . These scans produce, respectively, the energy distribution curves (EDCs) and momentum distribution curves (MDCs). The line shapes in both these scans are of fundamental interest since they provide a direct picture of the quasiparticle and background components of interacting Fermi systems, and thus unravel the roles of various interactions that are at play in strongly correlated Fermi systems. The dispersion relation of the electrons can be studied through the location of the peaks of  $A(\vec{k}, \omega)$  in constant  $\omega$  or constant  $\vec{k}$  scans.

Recent experimental studies have displayed a surprising ubiquity of *kinks* in the dispersion of strongly correlated matter at low energies  $\sim 50$ – $100$  meV. The kinks are bending-type anomalies (see Fig. 1) of the simple  $\omega = v_F(\vec{k} - \vec{k}_F)$ , i.e., linear energy versus momentum dispersion that is expected near  $\vec{k}_F$  from band theory. The special significance of kinks lies in the fact that their existence *must* signal a *departure* from band theory. This departure could be either due to electron-electron interactions or to interaction of the electrons with other bosonic degrees of freedom. Either of them are therefore significant enough to leave a direct and observable fingerprint in the spectrum. The goal of this work is to elucidate the origin of the observed kinks, and therefore to throw light on the dominant interactions that might presumably lead to high- $T_c$  superconductivity.

The purpose of this paper is multifold: We (i) survey the occurrence of the kinks in a variety of correlated systems of

current interest, (ii) provide a robust protocol for characterizing the kinks which is insensitive to the detailed shape of the kink, (iii) discuss how these kinks arise in two classes of theories, one based on coupling to a bosonic mode and the other to strong correlations, and (iv) identify testable predictions that ARPES experiments can use to distinguish between these.

The 15 systems reporting kinks are listed in Table I: these include (1) most high- $T_c$  cuprates in the (nodal) direction  $\langle 11 \rangle$  at various levels of doping from insulating to normal metallic states in the phase diagram [1,2], (2) charge density wave systems, (3) cobaltates, and (4) ferromagnetic iron surfaces. The kinks lose their sharpness as temperature is raised [2–4], and appear to evolve smoothly between the  $d$ -wave superconducting state and the normal state.

The kinks above  $T_c$  are smoothed out as one moves away from nodal direction [5]. Recent experiments [6] resolve this movement of the kinks more finely into two subfeatures. Most of the studies in Table I focus on MDC kinks; the EDC kinks data are available for only eight systems so far. Bosonic modes have been reported in six systems using different probes such as inelastic x rays or magnetic scattering, with either charge (phonons, plasmons) or spin (magnetic) character, while the remaining nine systems do not report such modes. A few theoretical studies of the kinks have implicated the observed low-energy modes via electron-boson-type calculations; we summarize this calculation in the Supplemental Material (SM) [7]. We find, in agreement with earlier studies, that the boson coupling mechanism yields kinks in the MDC dispersion, provided the electron-boson coupling is taken to be sufficiently large. In addition, we find in all cases studied this mechanism also predicts a jump in the EDC dispersion. It also predicts an extra peak in the spectral function pinned to the kink energy after the wave vector crosses the kink. These two features are experimentally testable and differ from the predictions of the correlations mechanism discussed next.

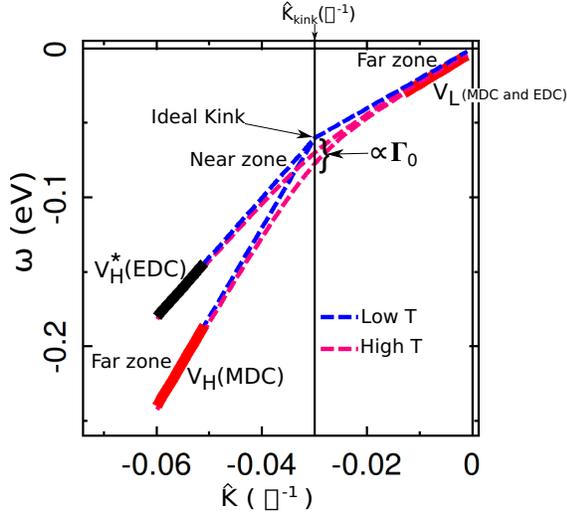


FIG. 1. A schematic MDC and EDC spectrum displaying typical features of experiments discussed below. Here,  $\hat{k} = (\vec{k} - \vec{k}_F) \cdot \vec{\nabla}_{\varepsilon_{k_F}} / |\vec{\nabla}_{\varepsilon_{k_F}}|$  is the momentum component normal to the Fermi surface, and we label EDC variables with a star. [The sketch uses parameters  $V_L = 2 \text{ eV \AA}$ ,  $V_H = 6 \text{ eV \AA}$ ,  $r = 1.5$ ,  $\hat{k}_{\text{kink}} = -0.03 \text{ \AA}^{-1}$ ,  $\Delta_0 = 0.03 \text{ eV}$ , and  $\Gamma_0 = 0.01 \text{ eV}$  in Eqs. (3) and (4).] The tangents in the *far zones* identify the asymptotic velocities  $V_L < V_H$  and  $V_L^* < V_H^*$  that characterize the MDC and EDC spectra. The intersection of the extrapolated MDC tangents fixes the kink momentum  $\hat{k}_{\text{kink}}$  and the ideal energy  $E_{\text{kink}}^{\text{ideal}}$ . The dispersion is rounded with raising  $T$ , as in the lower (red) curve. We *define* the MDC kink energy  $E_{\text{kink}}^{\text{MDC}}$  as  $E(\hat{k}_{\text{kink}})$ , i.e., the binding energy measured *at* the kink momentum, and similarly the EDC kink energy. In all cases,  $V_L = V_L^*$ . A testable consequence of the ECFL theory is that  $V_H^*$  is fixed in terms of the two MDC velocities by a strikingly simple relation:  $V_H^* = \frac{3V_H - V_L}{V_H + V_L} \times V_L$  [see Eq. (10)]. This easily testable prediction is tried against experimental data in Fig. 2 where both EDC and MDC data are available. In contrast, the electron-boson theory predicts a jump in the EDC dispersion at the kink energy, followed by  $V_H^* = V_H$ . Note that the difference between the EDC (MDC) kink energy,  $E_{\text{kink}}^{\text{EDC}} = E_{\text{kink}}^{\text{ideal}} - \Gamma_0$  and  $E_{\text{kink}}^{\text{MDC}} = E_{\text{kink}}^{\text{ideal}} - \Gamma_0 \sqrt{\frac{r}{2-r}}$ , and the ideal kink energy is equal (proportional) to  $\Gamma_0$ .

Since kinks are also observed in cases where no obvious bosonic mode is visible, it is important to explore alternate mechanisms that give rise to such features. In this context, we note that a recent theoretical work using the extremely strongly correlated Fermi liquid (ECFL) theory [8,9] calculates the dispersion using low-momentum and frequency expansions of the constituent self-energies. This calculation [9] shows that both EDC and MDC energy dispersions display qualitatively similar kinks, in particular, there is no jump in either dispersion. In essence, this work implies that a purely electronic mechanism with a strong momentum dependence of the Dyson self-energy results in kink-type anomalies. In terms of parameter counting, the calculation is *overdetermined*; it can be represented in terms of four parameters which can be fixed from a subset of measurements. With this determination one can then predict many other measurables and testable relations between these, as we show below. We show below that the various predictions are reasonably satisfied in one

case (of OPT Bi2212 below), while in other cases, there are insufficient experimental data to test the theories.

The ECFL theory incorporates strong Gutzwiller-type correlation effects into the electron dynamics [7]. It produces line shapes that are in close correspondence to experimental results for the high- $T_c$  systems [11,12]. The presence of a low-energy kink in the theoretical dispersion was already noted in Ref. [11]; this work substantially elaborates that observation. In order to understand the origin of a low-energy scale in the ECFL theory, it is useful to recall the predicted cubic correction to Fermi liquid self-energy  $\text{Im} \Sigma(\vec{k}_F, \omega) \sim \omega^2 (1 - \frac{\omega}{\Delta_0})$  from equations (SM-42) and (8) and (9). Here,  $\Delta_0$  is an emergent low-energy scale; it is related to the correlation-induced reduction of the quasiparticle weight  $Z$ . It reveals itself most clearly in the observed particle-hole asymmetry of the spectral functions, and therefore can be estimated independently from spectral *line-shape* analysis. A related and similar low value of the effective Fermi temperature is found in recent studies of the resistivity [10]. Here and in our earlier studies it is coincidentally found that  $\Delta_0 \sim 20\text{--}50 \text{ meV}$ , i.e., it is also roughly the energy scale of the kinks when the bandwidth is a few eV.

## II. ARPES SPECTRAL DISPERSIONS, KINKS, AND A PROTOCOL FOR DATA ANALYSIS

### A. Summary of variables in the theory

A few common features of spectral dispersions found in experiments are summarized in Fig. 1. The schematic figure shows a region of low spectral velocity near the Fermi level followed by a region of steeper velocity; these are separated by a bend in the dispersion, namely, the kink. While the kink itself has a somewhat variable shape in different experiments, the “far zone” is much better defined and is usually independent of the temperature; we denote the velocities in the far zones  $V_L, V_H$  for the MDC dispersion and the EDC dispersion counterparts by  $V_L^*, V_H^*$ . In terms of the normal component of the momentum measured from the Fermi surface

$$\hat{k} = (\vec{k} - \vec{k}_F) \cdot \vec{\nabla}_{\varepsilon_{k_F}} / |\vec{\nabla}_{\varepsilon_{k_F}}|, \quad (1)$$

the kink momentum  $\hat{k}_{\text{kink}}$  is uniquely defined by extrapolating the two asymptotic tangents, and the binding energy at this momentum defines the ideal kink energy  $E_{\text{kink}}^{\text{ideal}}$  [see Eq. (7)], which serves as a useful reference energy.

Our picture is that all lines of temperature-varying MDC dispersion curves in near zone converges into one line in the far zone in Fig. 1. We find that both the low and high velocities are independent of the temperature while depending on the doping levels. Lastly, the new laser ARPES data reveal that we need low-temperature dispersion data to determine  $V_L$  because temperature effect strongly influences the spectrum near the Fermi level.

We first define the important ratio parameter  $r$  ( $1 \leq r \leq 2$ ) from the MDC dispersion velocities as

$$r = \frac{2V_H}{V_H + V_L}. \quad (2)$$

The EDC dispersion relation  $E^*(\hat{k})$  locates the maximum of the spectral function  $A(\vec{k}, \omega)$  in  $\omega$  at constant  $\hat{k}$ , while the MDC

TABLE I. Comprehensive survey for ARPES kinks.

Name of the compounds	Above $T_c$		Below $T_c$		Local bosonic mode		
	MDC	EDC	MDC	EDC	Charge	Spin	Not reported
LSCO	✓[3,13]		✓[1,3,13,14]	✓[15]	✓[16–18]	✓[19]	
Bi2201	✓[3,5,13,20,21]	✓[23]	✓[5,21]		✓[24]		
Bi2212	✓[2–5,13,25,26]	✓[4]	✓[2–5,13,25,26]		✓[27]	✓[28,29]	
Bi2223	✓[5,30]		✓[5,30,31]				✓
YBCO			✓[32]		✓[33,34]	✓[35–38]	
Hg1201			✓[39]		✓[40]	✓[41–43]	
F0234			✓[44]				✓
CCOC			✓[45]				✓
LSMO			✓[46]	✓[46]			✓
2H-TaSe2 (CDW)			✓[47]		✓[48]		
Iron (110) surface			✓[49]				✓
BiBaCo1			✓[50] 5 K	✓[50] 5 K			✓
BiBaCo2			✓[50] 5 K	✓[50] 5 K			✓
BiBaCo	✓[50] 200 K	✓[50] 200 K					✓
NaCoO			✓[50] 5 K	✓[50] 5 K			✓

dispersion and  $E(\hat{k})$  locates the maximum  $\hat{k}$  at a fixed energy  $\omega$ . These are found from the ECFL theory (see SM [7] and Ref. [9]) as

$$E^*(\hat{k}) = (r V_L \hat{k} + \Delta_0 - \sqrt{\Gamma_0^2 + Q^2}), \quad (3)$$

$$E(\hat{k}) = \frac{1}{2-r} (V_L \hat{k} + \Delta_0 - \sqrt{r(2-r)\Gamma_0^2 + Q^2}), \quad (4)$$

where we introduced an energy parameter related to  $r$ ,  $V_L$  and  $\hat{k}_{\text{kink}}$

$$\Delta_0 = \hat{k}_{\text{kink}} V_L (1-r), \quad (5)$$

and a momentum-type variable  $Q = (r-1)V_L(\hat{k} - \hat{k}_{\text{kink}})$ . The variable  $\Gamma_0$  is temperaturelike

$$\Gamma_0 = \eta + \pi \{\pi k_B T\}^2 / \Omega_\Phi; \quad (6)$$

$\eta$  is an elastic scattering parameter dependent upon the incident photon energy, and  $\eta$  is very small for laser ARPES experiments and can be neglected to a first approximation. Here,  $\Omega_\Phi$  is a self-energy decay constant explained further in the SM [7]. The ideal kink energy  $V_L \hat{k}_{\text{kink}}$  can be expressed in terms of  $\Delta_0$  scale as

$$E_{\text{kink}}^{\text{ideal}} = -\frac{1}{r-1} \Delta_0. \quad (7)$$

It is important to note that these dispersion relation equations (3) and (4) are different from the standard dispersion relations  $E_{\text{FLT}}(\hat{k}) = E_{\text{FLT}}^*(\hat{k}) = V_H \hat{k}$ , which follow in the simplest Fermi liquid theory (FLT) near the Fermi energy  $A_{\text{FLT}}(\vec{k}, \omega) = \frac{1}{\pi} \frac{\Gamma_0}{(\omega - V_H \hat{k})^2 + \Gamma_0^2}$ . The FLT dispersions are identical in EDCs and MDCs, and are independent of the temperaturelike variable  $\Gamma_0$ , and do not show kinks. On the other hand, Eqs. (3) and (4) do have kinks, as we show below, and the temperaturelike variable  $\Gamma_0$  plays a significant role in the dispersion. At  $\Gamma_0 = 0$  one has an ideal spectrum, where the kinks are sharpest. When  $\Gamma_0 \neq 0$ , due to either finite temperature or finite damping  $\eta$ , related to the energy of the incoming photon, the kinks are rounded.

A few consequences of Eqs. (3) and (4) can be noted for the purpose of an experimental determination of the Fermi momentum. The chemical potential is usually fixed by referencing an external metallic contact and is unambiguous. Experimentally, the Fermi momentum is usually found from the MDC, as the momentum where the spectral function is maximum with energy fixed at the chemical potential, i.e.,  $\omega = 0$ . This corresponds to the generally wrong expectation that  $E(\hat{k}_{\text{peak}}) = 0$  implies  $\hat{k}_{\text{peak}} = 0$ . When  $\Gamma_0 \geq 0$ , from Eq. (4) we see that the condition  $E(\hat{k}_{\text{peak}}) = 0$  gives  $\hat{k}_{\text{peak}} = \frac{\sqrt{\Delta_0^2 + r^2 \Gamma_0^2 - \Delta_0}}{r V_L}$ , a positive number that equals zero only in the ideal case  $\Gamma_0 = 0$ . Thus, there is an apparent enlargement of the Fermi surface due to a finite  $\Gamma_0$  that needs to be corrected. By the same token, at the true (Luttinger theorem related) Fermi momentum  $\hat{k} = 0$ , the MDC energy  $E(0) = \frac{\Delta_0 - \sqrt{\Delta_0^2 + r(2-r)\Gamma_0^2}}{2-r}$ , a negative number when  $\Gamma_0 \neq 0$ . In recent laser ARPES Bi2201 data [[21], panel (a) in Fig. 4], we see that  $E(\hat{k}_{\text{peak}})$  vanishes at increasing  $\hat{k}_{\text{peak}}$  as  $T$  is raised, as predicted in our calculation. Recent laser ARPES experiment on OPT Bi2212 compounds reports a similar temperature dependence of momentum of MDC dispersion at the Fermi level in Ref. [26], strongly supporting our picture of its origin.

Similarly, the EDC peak at the true Luttinger theorem related Fermi surface  $\hat{k} = 0$  is nonzero. We find  $E^*(0) = (\Delta_0 - \sqrt{\Delta_0^2 + \Gamma_0^2}) \leq 0$ . Clearly,  $E^*(0)$  is negative unless  $\Gamma_0 = 0$ , i.e., it is generically red-shifted. If we are tempted to identify the Fermi momentum from the condition  $E^*(\hat{k}_{\text{peak}}^*) = 0$ , a similar cautionary remark is needed. The condition  $E^*(\hat{k}_{\text{peak}}^*) = 0$  gives  $\hat{k}_{\text{peak}}^* = \frac{\sqrt{\Delta_0^2 + (2r-1)\Gamma_0^2 - \Delta_0}}{(2r-1)V_L}$ , again a positive number as in the MDC case, and thus a slightly different enlargement of the apparent Fermi surface.

The above comments illustrate the difficulty of finding the correct Fermi surface when  $\Gamma_0$  is non-negligible, as in the case of synchrotron ARPES with substantial values  $\Gamma_0 \gtrsim 50$  meV. On the other hand, the laser ARPES studies have a much smaller  $\eta \lesssim 10$  meV, where our analysis can be tested by varying the temperature and the consequent change of the

spectrum. In the following, we analyze the data from the Bi2201 system where the laser data are available at various  $T$ , and allows us to test the above in detail. Our analysis below of two other synchrotron data, the OPT Bi2212 has  $10 \leq \eta \leq 40$  meV, while the low- $T$  LSCO data are assumed to be in the limit of  $\eta = 0$  because of the lack of high-temperature dispersion data.

The spectral function at low frequencies close to  $\vec{k}_F$  is also obtainable from these parameters; the relevant formula is noted below. In terms of  $\xi$

$$\xi = \frac{1}{\Delta_0}(\omega - r V_L \hat{k}) \quad (8)$$

the spectral function is

$$A(\vec{k}, \omega) = \frac{z_0}{\pi} \frac{\Gamma_0}{(\omega - V_L \hat{k})^2 + \Gamma_0^2} \left\{ 1 - \frac{\xi}{\sqrt{1 + c_a \xi^2}} \right\}. \quad (9)$$

Here,  $z_0$  is the quasiparticle weight and  $c_a \sim 5.4$  (see SM [7]). We should keep in mind that these expressions follow from a low-energy expansion, and is limited to small  $\hat{k}$  and  $\omega$ ; in practical terms the dimensionless variable  $|\xi| \lesssim 4$ , so that  $\omega$  (or  $\hat{k}$ ) is bounded by the kink energy (or momentum), as defined below. Finally, we note a strikingly simple relation that relates the high velocity of the EDC spectrum to the two velocities  $V_H$  and  $V_L$  in the MDC dispersion defined in Fig. 1:

$$V_H^* = \frac{3V_H - V_L}{V_H + V_L} \times V_L. \quad (10)$$

The origin of this simple but *key formula* lies in the fact that the entire ECFL spectrum is determined in terms of a few parameters, and therefore one should expect inter-relationships of this kind on general grounds. The details are provided in the Supplemental Material [7] Eq. (SM-36).

### III. OPT BI2212 ARPES DISPERSION DATA

In the well-studied case of optimally doped Bi2212 (BSCCO) superconductors, the kink has been observed in both EDC and MDC. We summarize the ECFL fit parameters in Table II obtained from literature [4]. We also display the predicted energy and high velocity of the EDC dispersion. The velocity ratio  $V_H/V_H^* \sim 1.3$  in this case is quite large and measurable. In this case, the EDC dispersion has fortunately

already been measured, allowing us to test the prediction. From Table II we see that the energy of the EDC kink and its velocity are close to the predictions.

In Fig. 2(a), we plot the predicted EDC dispersion using the parameters extracted from the MDC dispersion in Fig. 2(b), and compare with the ARPES data measured [4]. It is interesting that the predicted slope of the EDC dispersion from Eq. (10) is close to the measured one. Indeed, the measured EDC dispersion is close to that expected from the ECFL theory. To probe further, in Fig. 2(c) we compare the theoretical EDC line shape (solid blue line) given by the same parameters through Eq. (9) with the ARPES line shape measured at high temperature [4]. Figure 2(d) compares the theoretical MDC curve with the data. The theoretical curves are from the low-energy expansion and hence are chopped at the high end, corresponding to roughly  $|\xi|_{\max} \sim \frac{r V_L \hat{k}_{\text{kink}}}{\Delta_0}$  for MDC and  $|\xi|_{\max} \sim \frac{E_{\text{kink}}^{\text{ideal}}}{\Delta_0}$  for the EDC. With this cutoff, the momentum is less than the kink momentum and the energy is less than the kink energy. We used  $\Gamma_0 = 40$  meV since it provides a rough fit for both EDC and MDC spectral functions.

This value is somewhat larger than the bound  $\sim 10$  meV given in Table II; a smaller value leads to narrower lines but with the same shape. In rigorous terms, the same  $\Gamma_0$  must fit the dispersion and also the spectral functions. Our fit, requiring a different  $\Gamma_0$ , is not ideal in that sense. However, the resolution of the available data is somewhat rough, and should improve with the newer experimental setups that have become available. We thus expect that higher-resolution data with laser ARPES should provide an interesting challenge to this theory. We also stress that from Eq. (9), the MDC line shapes look more symmetric than the EDC line shapes at low energies. While many experimental results do show rather symmetric MDCs, there are well-known exceptions. For instance, MDCs asymmetry has indeed been reported for nearly optimally doped Hg1201 ( $T_c = 95$  K) at binding energy very close to the Fermi level,  $\omega \sim -5$  meV and  $\omega \sim -18$  meV in Fig. 5 in Ref. [39]. Note that the  $\omega = 0$  MDC plot of the spectral function  $A(k, \omega)$  from Eq. (9) locates the peak momentum  $\hat{k}_{\text{peak}} > 0$ , i.e., slightly to the right of the physical Fermi momentum  $\vec{k}_F$ , and we consider this implies that the *experimental* Fermi momentum determination is subject to such a correction, whenever the spectral function Eq. (9) has a momentum-dependent caparison factor (see caption in Fig. 2).

TABLE II. Parameter table for ARPES kink analysis for OPT Bi2212 [4] in Fig. 2 presents three essential parameters:  $V_L$ ,  $V_H$ , and  $\hat{k}_{\text{kink}}$ . From the high- and low-temperature MDC dispersions, we measured  $\Gamma_0 \lesssim 10$  meV in Fig. 2(b). With the measured experimental parameters and determining the velocity ratio  $r$  in Eq. (2), we are able to estimate the finite-temperature kink energy for EDC and MDC dispersions by  $E_{\text{kink}}^{\text{EDC}} = E_{\text{kink}}^{\text{ideal}} - \Gamma_0$  and  $E_{\text{kink}}^{\text{MDC}} = E_{\text{kink}}^{\text{ideal}} - \Gamma_0 \sqrt{\frac{r}{2-r}}$  and predict  $V_H^*$  by  $V_H^* = \frac{3V_H - V_L}{V_H + V_L} \times V_L$  in Eq. (10). The uncertainties for calculated variables were determined by error propagation, and the uncertainties for experimental variables were given by the half of the instrumental resolution.

MDCs			EDCs					
OPT Bi2212 ARPES data			$E_{\text{kink}}^{\text{MDC}}$ (meV)		$E_{\text{kink}}^{\text{EDC}}$ (meV)		$V_H^*$ (eV Å)	
$V_L$ (eV Å)	$V_H$ (eV Å)	$\hat{k}_{\text{kink}}$ (Å <sup>-1</sup> )	Calculated	Measured	Calculated	Measured	Predicted	Measured
$1.47 \pm 0.07$	$3.3 \pm 0.3$	$-0.037 \pm 0.005$	$67 \pm 21$	$67 \pm 8$	$63 \pm 21$	$65 \pm 8$	$2.60 \pm 0.56$	$2.1 \pm 1.1$

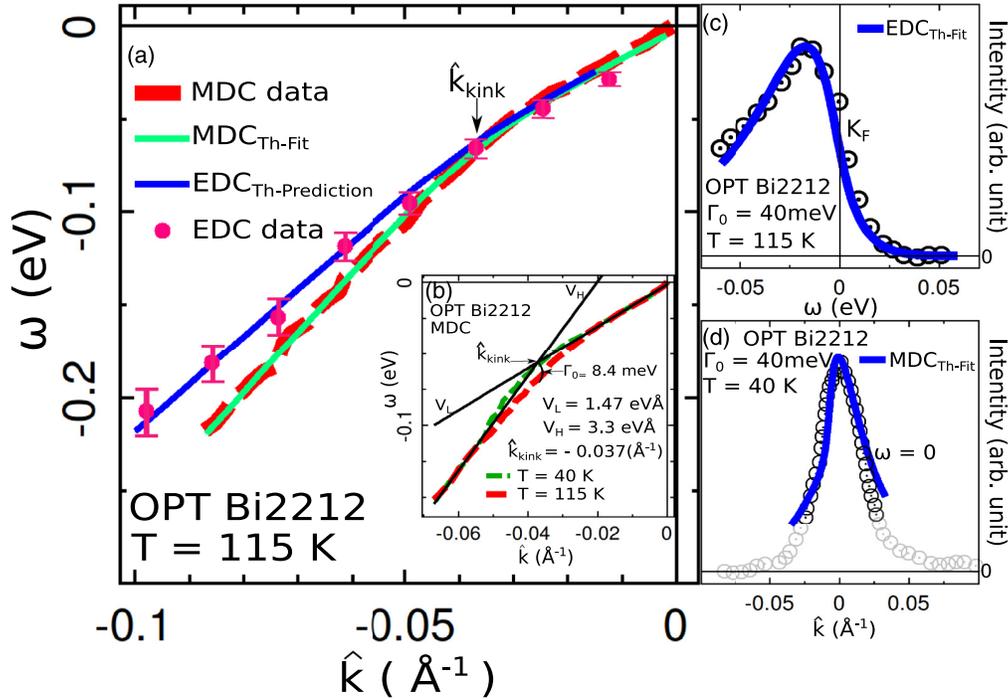


FIG. 2. ARPES kinks data for OPT Bi2212 from Ref. [4] compared to theoretical ECFL curves (solid lines) using parameters listed in Table II. (a) The predicted EDC spectrum (blue) from Eq. (3) versus the experimental EDC data (magenta symbols) at  $T = 115$  K. For reference we also show the MDC data (red dashed curve) and the corresponding ECFL fit (green solid curve). (b) Experimental MDC spectra at 40 K (below  $T_c$  in green dashed line) and 115 K (above  $T_c$  in red dashed line) yield common asymptotes shown in black lines from the far zone. These determine the parameters displayed in Table II. (c) At low energy  $\pm 60$  meV, the EDCs spectral function (blue solid line) from Eq. (9) is contrasted with the corresponding ARPES data from [4]. (d) At  $\omega = 0$  we compare the MDCs spectral function (blue solid line) from Eq. (9) with the corresponding ARPES data from Ref. [4]. The range of validity for the theoretical expansion is  $\pm \hat{k}_{\text{kink}}(0.037 \text{ \AA}^{-1})$ , the data points in the range are shown in black circle symbols, while the light gray circle symbols are outside this range. The peak position of the theoretical curve has been shifted to left by  $0.007 \text{ \AA}^{-1}$ , a bit less than the instrumental resolution. A similar shift is made in Fig. 3(l). For analogous reasons, the EDC peak in  $A(k, \omega)$  at  $k_F$  is shifted to the left, i.e.,  $E^*(0) \leq 0$ . A small shift to the right is made in Fig. 3(k), in order to compensate for this effect. These shift effects are within the resolution with present setups, but should be interesting to look for in future generation experiments since they give useful insights into the energy momentum dependence of the spectral function.

#### IV. LSCO LOW-TEMPERATURE DATA

Here, we analyze the LSCO data at low temperature (20 K) and at various doping levels ranging from the insulator ( $x = 0.03$ ) to normal metal ( $x = 0.3$ ) from Ref. [1]. The parameters are listed in Table III, where we observe that the velocity  $V_L$  is roughly independent of  $x$ , and has a somewhat larger magnitude to that in OPT Bi2212 in Table II. The kink momentum decreases with decreasing  $x$ , roughly as  $\hat{k}_{\text{kink}} = -(0.37x - 0.77x^2) \text{ \AA}^{-1}$ , and the kink energies of EDC and MDC dispersions are essentially identical. In the region beyond the kink, the prediction for  $V_H^*$  is interesting since it differs measurably from the MDC velocity  $V_H$ . We find the ratio  $V_H/V_H^* \sim 1.02\text{--}1.5$  is quite spread out at different doping.

Our analysis becomes unreliable as lower doping level  $x < 0.075$  in Figs. 3(h)–3(j), where the dispersion kink is no longer a simple bending kink, an extra curving tendency begins to appear. To put this in context, recall that the line shape of LSCO becomes extremely broad at small  $x$  [14], and so the peak position of the spectral function becomes more uncertain than at higher energy.

We should point out that in Fig. 3(k) the spectral function has been shifted to right by 4 meV for a better fit. This shifting is consistent with our argument that the Fermi momentum determination has a possible small error of in order  $0.006 \text{ \AA}^{-1}$ , arising from the  $\hat{k}$ -dependent caparison factor, and hence the peak position has an uncertainty  $V_L \times 0.006 \sim 10$  meV.

#### V. BI2201 LASER ARPES DATA

In this section, we present our analysis of the high-resolution laser ARPES data of the single-layered compounds Bi2201, at various different doping levels taken from a recent study in Ref. [21]. In earlier studies of this compound using synchrotron emitted high-energy photons, as also LSCO [3], the ARPES kinks were observed to have only a weak temperature dependence [5]. However, the new high-resolution laser ARPES data enables us to observe clear and significant temperature dependence of the ARPES kinks; it is comparable to that of the double-layered Bi2212 compounds. In fact, we find that the new data of Bi2201 compounds in Ref. [21] seem to provide a textbook example of our ECFL kink analysis.

TABLE III. Data table for ARPES kink analysis for OPT LSCO ( $T = 20$  K) [1] in Fig. 3. We were unable to reliably estimate  $\Gamma_0$  here due to the lack of data at high temperature, and hence set it at zero. The uncertainties for measured values were given by half of the instrumental resolution (10 meV,  $\sim 0.005 \text{ \AA}^{-1}$ ). The uncertainties for the calculated values were determined by error propagation.

$x$ (doping level)	MDCs			EDCs					
	LSCO low-temperature ARPES data			$E_{\text{kink}}^{\text{MDC}}$ (meV)		$E_{\text{kink}}^{\text{EDC}}$ (meV)		$V_H^*$ (eV $\text{\AA}$ )	
	$V_L$ (eV $\text{\AA}$ )	$V_H$ (eV $\text{\AA}$ )	$\hat{k}_{\text{kink}}$ ( $\text{\AA}^{-1}$ )	Calculated	Measured	Calculated	Measured	Calculated	Measured
0.3	$2.4 \pm 0.2$	$3.0 \pm 0.3$	$-0.047 \pm 0.005$	$113 \pm 29$	$110 \pm 10$	$113 \pm 29$		$2.93 \pm 0.45$	
0.22	$2.0 \pm 0.1$	$3.6 \pm 0.2$	$-0.042 \pm 0.005$	$84 \pm 18$	$85 \pm 10$	$84 \pm 18$		$3.14 \pm 0.35$	
0.18	$1.7 \pm 0.3$	$4.5 \pm 0.6$	$-0.040 \pm 0.005$	$68 \pm 43$	$72 \pm 10$	$68 \pm 43$		$3.2 \pm 1.2$	
0.15	$1.75 \pm 0.07$	$4.3 \pm 0.1$	$-0.037 \pm 0.005$	$65 \pm 11$	$64 \pm 10$	$65 \pm 11$		$3.23 \pm 0.20$	
0.12	$2.0 \pm 0.3$	$3.7 \pm 0.5$	$-0.029 \pm 0.005$	$58 \pm 28$	$55 \pm 10$	$58 \pm 28$		$3.19 \pm 0.89$	
0.1	$1.8 \pm 0.2$	$5.0 \pm 0.7$	$-0.035 \pm 0.005$	$63 \pm 44$	$64 \pm 10$	$63 \pm 44$		$3.5 \pm 1.4$	
0.075	$1.9 \pm 0.2$	$5.6 \pm 0.8$	$-0.026 \pm 0.005$	$49 \pm 37$	$51 \pm 10$	$49 \pm 37$		$3.8 \pm 1.7$	
0.063	$1.8 \pm 0.3$	$6.0 \pm 0.5$	$-0.022 \pm 0.005$	$40 \pm 21$	$43 \pm 10$	$40 \pm 21$		$3.7 \pm 1.1$	
0.05	$1.7 \pm 0.2$	$5.7 \pm 0.6$	$-0.023 \pm 0.005$	$39 \pm 25$	$41 \pm 10$	$39 \pm 25$		$3.5 \pm 1.3$	
0.03	$2.0 \pm 0.3$	$6.1 \pm 0.4$	$-0.016 \pm 0.005$	$32 \pm 15$	$32 \pm 10$	$32 \pm 15$		$4.02 \pm 0.85$	

In Table IV, we list the kink parameters corresponding to different doping levels of Bi2201 and tabulate the kink parameters. The entries are in correspondence to the panels in Fig. 4. In Figs. 4(a)–4(f), we depict the measured MDC dispersion and the predicted EDC dispersions at different doping levels. The latter are found from Eq. (3) using the variables in Table IV. Figures 4(g) and 4(h) of OPT Bi2201 are especially interesting. Combining the low  $T = 15$  K dispersion data and the finite- $T$  value of  $\Gamma_0$ , found from the depression of the kink energy  $E_{\text{kink}}^{\text{MDC}} = E_{\text{kink}}^{\text{ideal}} - \Gamma_0 \sqrt{\frac{r}{2-r}}$ , we can reconstruct the entire MDC dispersion at a finite  $T$ . This may be compared with the measured finite- $T$  MDC data, thus checking the validity of the formalism. This exercise is carried

out at  $T = 200$  K in Fig. 4(g) and  $T = 100$  K in Fig. 4(h), where we find a remarkably good fit in all details. In Figs. 4(g) and 4(h) we show the actual momentum (rather than  $\hat{k}$ ) to facilitate a comparison with data. Figure 4(g) especially clearly shows that  $E(\hat{k})$  vanishes at a  $\hat{k}$  that is different from 0. The shift corresponds to  $\sim 0.01 \text{ \AA}^{-1}$ . We have commented above that this apparent expansion of the Fermi surface with  $T$  is due to the nontrivial physics underlying Eq. (4) lying beyond the simple minded FLT.

Figure 4(i) plots the temperature dependence of  $\Gamma_0$  in Fig. 4(a) in Ref. [21]. The measured  $\Gamma_0$  curve is fitted with Eq. (6), and we estimate  $\eta = 5.3 \pm 2$  meV and  $\Omega_\Phi = 410 \pm 100$  meV.

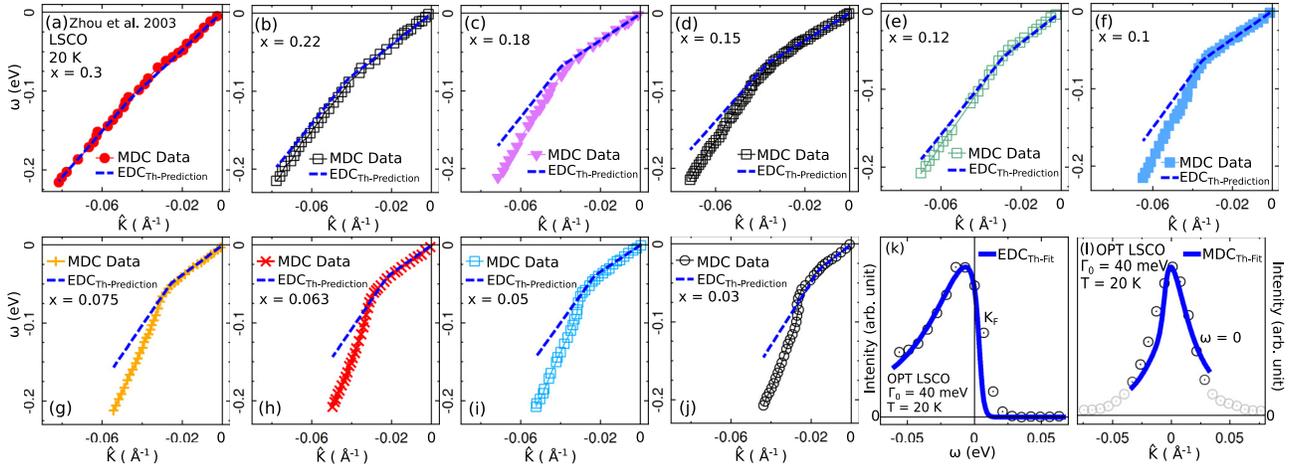


FIG. 3. ARPES kinks data for LSCO data [1] compared to theoretical ECFL curves (solid lines) using parameters listed in Table III. The doping level  $x$  varies between (normal metal)  $0.3 \leq x \leq 0.03$  (insulator) in (a)–(j). Each panel shows MDC nodal dispersion data (symbols), whose uncertainties are  $\pm 10$  meV. The blue dashed line is the theoretical prediction for EDC dispersion by Eq. (3). (k) We compare the spectral line shape for EDCs at  $k_F$  from Eq. (9) (blue solid line) in the range  $\pm E_{\text{kink}}^{\text{ideal}} \sim 65$  meV with the corresponding ARPES data (black circles) [12]. (l) At  $\omega = 0$  we compare the MDCs spectral function (blue solid line) from Eq. (9) with the corresponding ARPES data from Ref. [12]. The range of validity for the theoretical expansion is  $\pm \hat{k}_{\text{kink}}$  ( $0.037 \text{ \AA}^{-1}$ ), the data points in the range are shown in black circle symbols, while the light grey circle symbols are outside this range. The peak position of the theoretical curve MDC has been shifted to left by  $0.006 \text{ \AA}^{-1}$ .

TABLE IV. Parameter table for ARPES kink analysis for laser ARPES data of Bi2201 at various different doping levels [21] in Fig. 4. From  $0.1 < x < 0.16$ , we measured  $\Gamma_0 \sim 0$ . For  $x = 0.23$  and  $0.26$ , we measured  $\Gamma_0 \lesssim 17$  meV. For  $x = 0.16$  data, we report variables for high-temperature kinks data 200 K (g) and 100 K (h) in Fig. 4, and  $\Gamma_0$  values for 200 K and 100 K data are in corresponding panels (g) and (h) in Fig. 4. The uncertainties for the calculated parameters were determined by error propagation, and the uncertainties for the experimental parameters were given by half of the instrumental resolution.

$x$ (doping level)	MDCs			EDCs					
	Bi2201 laser ARPES data			$E_{\text{kink}}^{\text{MDC}}$ (meV)		$E_{\text{kink}}^{\text{EDC}}$ (meV)		$V_H^*$ (eV $\text{\AA}$ )	
	$V_L$ (eV $\text{\AA}$ )	$V_H$ (eV $\text{\AA}$ )	$k_{\text{kink}}$ ( $\text{\AA}^{-1}$ )	Calculated	Measured	Calculated	Measured	Calculated	Measured
0.1	$1.47 \pm 0.12$	$4.7 \pm 0.3$	$-0.022 \pm 0.002$	$32 \pm 3$	$37 \pm 0.5$	$32 \pm 6$		$3.0 \pm 0.3$	
0.11	$1.34 \pm 0.06$	$2.78 \pm 0.06$	$-0.021 \pm 0.002$	$28 \pm 1$	$28 \pm 0.5$	$28 \pm 4$		$2.28 \pm 0.12$	
0.13	$1.37 \pm 0.07$	$2.71 \pm 0.18$	$-0.025 \pm 0.002$	$38 \pm 3$	$39 \pm 0.5$	$37 \pm 5$		$2.27 \pm 0.17$	
0.16	$1.5 \pm 0.1$	$3.5 \pm 0.2$	$-0.026 \pm 0.002$	$39 \pm 3$	$43 \pm 0.5$	$39 \pm 6$		$2.7 \pm 0.2$	
0.23	$2.1 \pm 0.11$	$5.4 \pm 0.3$	$-0.036 \pm 0.002$	$98 \pm 6$	$97 \pm 0.5$	$89 \pm 10$		$3.9 \pm 0.3$	
0.26	$2.17 \pm 0.16$	$4.8 \pm 0.4$	$-0.045 \pm 0.002$	$123 \pm 11$	$122 \pm 0.5$	$114 \pm 18$		$3.8 \pm 0.4$	
0.16 (200 K)	$1.61 \pm 0.18$	$3.5 \pm 0.3$	$0.364 \pm 0.002$	$87 \pm 11$	$89 \pm 0.5$	$75 \pm 11$		$2.8 \pm 0.4$	
0.16 (100 K)	$1.61 \pm 0.18$	$3.5 \pm 0.3$	$0.364 \pm 0.002$	$69 \pm 11$	$70 \pm 0.5$	$62 \pm 11$		$2.8 \pm 0.4$	

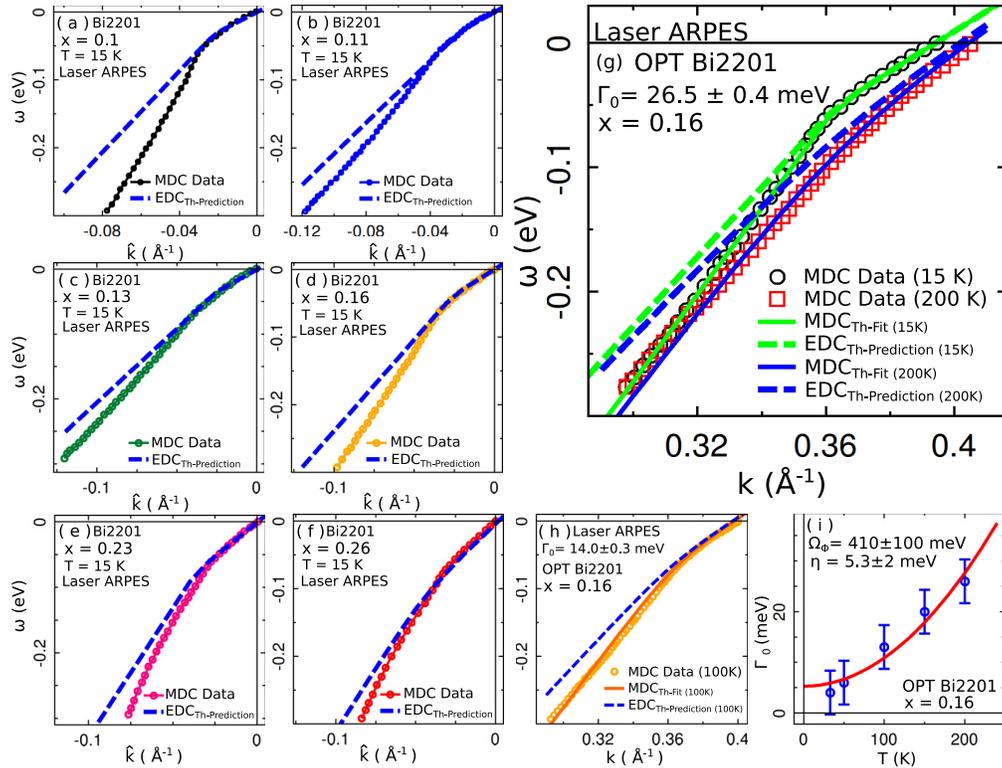


FIG. 4. ARPES kink analysis for laser ARPES data of Bi2201 at various different doping levels in Ref. [21]. (a)–(f) We predict EDC dispersions (blue dashed lines) using Eq. (3) for various different doping levels of Bi2201 laser ARPES data. (g), (h) First in (g), we present ECFL MDC fit (green solid line) for low-temperature (15 K) laser ARPES dispersion data of OPT Bi2201 from in Fig. 4(a) in Ref. [21] (black circles) and predict low-temperature EDC dispersion (green dashed line). Next, in (g) and (h), we predict high-temperature EDC (blue dashed lines) dispersions (g) 200 K and (h) 100 K for laser ARPES data of OPT Bi2201 [Fig. 4(a) in Ref. [21]], and show the MDC dispersion fits for two temperatures also, blue solid line for 200-K data (red squares) in (g) and brown solid line for 100-K data (yellow circles) in (h). We estimate  $\Gamma_0$  from measuring the difference between the ideal kink energy and the MDC kink energy. In order to compare with experiments, the  $x$ -axis representation in (g) and (h) is given by the physical  $k$  (rather than the momentum difference  $\hat{k}$ ). In (g), the MDC dispersion fit (blue solid line) of 200 K vanishes at  $k = 0.404 \pm 0.002 \text{ \AA}^{-1}$ , very close to the measured  $k = 0.405 \pm 0.002 \text{ \AA}^{-1}$  of the MDC dispersion data at 200 K. Similarly, in (h) the MDC dispersion fit (brown solid line) at 100 K vanishes at  $k = 0.398 \pm 0.002 \text{ \AA}^{-1}$ , close to the measured  $k = 0.4 \pm 0.002 \text{ \AA}^{-1}$  of the MDC dispersion data at 100 K. Note that the true Fermi momentum as estimated from the low- $T$  (15-K) data is  $k = 0.394 \pm 0.002 \text{ \AA}^{-1}$ , so that the deviations are bigger than the momentum resolution  $\Delta k \sim 0.004 \text{ \AA}^{-1}$ . (i) We plot the temperature dependence of  $\Gamma_0$  in Fig. 4(a) in Ref. [21]. Here, the temperature dependence data of  $\Gamma_0$  are fitted with Eq. (6), and  $\eta$  is determined  $5.3 \pm 2$  meV and  $\Omega_\phi = 410 \pm 100$  meV.

## VI. CONCLUSION

The main goal of this work is to understand the physical origin of kinks in the dispersion seen in ARPES studies of a wide class of systems. For this purpose we have listed 15 systems of topical interest where ARPES kink data are available. Our focus is on the nodal direction data since the largest volume is available here. We have devised a useful protocol to extract kink parameters from data, where the asymptotic tangents of the kink are used. Using this protocol we have analyzed in detail three families of systems: two synchrotron and one laser ARPES data of cuprate superconductors. The main parameters of the kinks are the energy, momentum, and the dispersion velocities in EDC and MDC scans; these provide a quantitative data set for testing various theoretical proposals for explaining kinks.

We have outlined two competing theories for the origin of kinks, and highlighted their distinctive predictions. One is the electron-boson model, where an Einstein mode of either spin or charge origin couples to the electrons, resulting in a momentum-independent self-energy. This theory gives rise to kinks in the electron dispersion. The other theory is the strong or extreme correlation theory, where the interactions lead to a momentum-dependent self-energy in two dimensions. This theory also gives rise to kinks in the electron dispersion. We expect that other contemporary theories of strong correlations, such as the cellular dynamical mean field theory (CDMFT) [22] method would give comparable results to those of the ECFL theory presented here, which provides the extra convenience of simple analytical expressions.

The predictions of the two theories differ significantly and in experimentally testable ways. Let us summarize the proposed tests.

*The boson-mode theory* [7] predicts the following:

- (1) A kink in the continuous MDC dispersion, located at the energy of the localized mode.
- (2) A momentum-independent peak at the kink energy, in the spectral function versus energy curve.
- (3) A jump discontinuity (rather than a kink) in the EDC dispersion.
- (4) The EDC and MDC velocities are identical both *above* and *below* the kink energy.

*The (extremely) strongly correlated Fermi liquid theory* [7] predicts the following:

- (1) A kink in the continuous MDC dispersion, located at a (calculable) *emergent* energy.
- (2) No peak in the spectral function at the kink energy.
- (3) A kink (rather than a jump discontinuity) in the continuous EDC dispersion.
- (4) The EDC and MDC velocities are identical *above* the kink energy.
- (5) *Below* the kink energy, the EDC velocity is determined by the two MDC velocities through a simple relation.

It is remarkable that a knowledge of the two MDC dispersions ( $V_H$  and  $V_L$ ) suffices to predict the EDC dispersion below the kink  $V_H^*$ , through the relation  $V_H^* = \frac{3V_H - V_L}{V_H + V_L} \times V_L$  [see Eq. (10)].

Thus, the parameters obtained from the MDC dispersion enable us to reconstruct the spectral function at low momentum and energy, in both MDC and EDC scans. We have carried out this exercise in three cases above.

It is thus clear that EDC dispersions hold the key to distinguishing between the two competing theories. EDC dispersion data are sparse but exist, the work on OPT Bi2212 from Ref. [4] shown in Fig. 2 presents both EDC and MDC dispersions at 115 K. Its resolution is presumably not optimal since it was an early experiment. Nevertheless, we can use it to make a first pass at comparing the two theories. This data set plotted in Fig. 2 shows that the EDC dispersion is continuous, i.e., has no jump. Further, the EDC higher velocity  $V_H^*$  is close to that predicted by the ECFL analysis. The measured spectral function in EDC, overlooking the noise, seems not to have any immovable feature at  $E_{\text{kink}}$ . Thus, all three characteristics noted above appear to be consistent with the ECFL predictions rather than the bosonic mode theory predictions. It is roughly fit by the low-energy parametrized curves as well, where the MDC is seen to be more symmetric than the EDC cuts.

As noted in Table I, the above case OPT Bi2212 is particularly interesting. Low-energy bosonic modes have been observed in neutron scattering [28,29] and in momentum-resolved electron energy loss experiments [27]. In Ref. [27] an MDC dispersion is presented using parameters taken from the bosonic data. This leads to a rather detailed model, and is shown to provide a reasonable fit to the MDC dispersion and the observed kink, but the important EDC dispersion is not displayed.

While we focused attention on dispersion kinks in the nodal direction in this work, the ECFL theory is also valid for other directions; it has a momentum dependence in the self-energy both normal to the Fermi surface and also along the tangent. The ECFL theory applied to the *d*-wave superconducting state in the *t*-*J* model is expected to lead to further interesting results in the future. For now, we note that the observed nodal direction spectra are essentially unchanged at  $T_c$ , which makes the nodal direction particularly interesting.

In conclusion, we have presented a current summary of the physics of the kinks in dispersion of cuprate high- $T_c$  superconductors, and given a set of measurements that can distinguish between competing theories. We believe that there is urgent need for further high-resolution EDC data, and also *T*-dependent scans to explore the rounding of kinks. Using such data one should be able to check the predictions of the theory more thoroughly, and thereby obtain definitive understanding of the origin of low-energy ARPES kinks of strongly correlated matter.

## ACKNOWLEDGMENTS

We thank A. Georges and J. Hancock for stimulating discussions. The work at UCSC was supported by the U. S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award No. DE-FG02-06ER46319.

- [1] X.-J. Zhou, T. Yoshida, A. Lanzara, P. V. Bogdanov, S. A. Kellar, K. M. Shen, W. L. Yang, F. Ronning, T. Sasagawa, T. Kakeshita, T. Noda, H. Eisaki, S. Uchida, C. T. Lin, F. Zhou, J. W. Xiong, W. X. Ti, Z. X. Zhao, A. Fujimori, Z. Hussain, and Z.-X. Shen, *Nature (London)* **423**, 398 (2003).
- [2] P. D. Johnson, T. Valla, A. V. Fedorov, Z. Yusof, B. O. Wells, Q. Li, A. R. Moodenbaugh, G. D. Gu, N. Koshizuka, C. Kendziora, Sha Jian, and D. G. Hinks, *Phys. Rev. Lett.* **87**, 177007 (2001).
- [3] A. Lanzara, P. V. Bogdanov, X. J. Zhou, S. A. Kellar, D. L. Feng, E. D. Lu, T. Yoshida, H. Eisaki, A. Fujimori, K. Kishio, J.-I. Shimoyama, T. Noda, S. Uchida, Z. Hussain, and Z.-X. Shen, *Nature (London)* **412**, 510 (2001).
- [4] A. Kaminski, M. Randeria, J. C. Campuzano, M. R. Norman, H. Fretwell, J. Mesot, T. Sato, T. Takahashi, and K. Kadowaki, *Phys. Rev. Lett.* **86**, 1070 (2001).
- [5] T. Sato, H. Matsui, T. Takahashi, H. Ding, H.-B. Yang, S.-C. Wang, T. Fujii, T. Watanabe, A. Matsuda, T. Terashima, and K. Kadowaki, *Phys. Rev. Lett.* **91**, 157003 (2003).
- [6] J. He, W. Zhang, J. M. Bok, D. Mou, L. Zhao, Y. Peng, S. He, G. Liu, X. Dong, J. Zhang, J. S. Wen, Z. J. Xu, G. D. Gu, X. Wang, Q. Peng, Z. Wang, S. Zhang, F. Yang, C. Chen, Z. Xu, H.-Y. Choi, C. M. Varma, and X. J. Zhou, *Phys. Rev. Lett.* **111**, 107005 (2013).
- [7] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.95.165435> for details of the doping dependence of the fit parameters, predictions of the electron-boson coupling model for kinks, predictions of the ECFL theory for kinks.
- [8] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).
- [9] B. S. Shastry, *Ann. Phys. (NY)* **343**, 164 (2014); **373**, 717 (2016).
- [10] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016).
- [11] G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).
- [12] K. Matsuyama and G.-H. Gweon, *Phys. Rev. Lett.* **111**, 246401 (2013).
- [13] D. R. Garcia and A. Lanzara, *Advances in Condensed Matter Physics* **2010**, 807412 (2010).
- [14] T. Yoshida, X. J. Zhou, D. H. Lu, S. Komiya, Y. Ando, H. Eisaki, T. Kakeshita, S. Uchida, Z. Hussain, Z.-X. Shen, and A. Fujimori, *J. Phys. Condens. Matter* **19**, 125209 (2007).
- [15] A. S. Mishchenko, N. Nagaosa, K. M. Shen, Z.-X. Shen, X. J. Zhou, and T. P. Devereaux, *Europhys. Lett.* **95**, 57007 (2011).
- [16] R. J. McQueeney, Y. Petrov, T. Egami, M. Yethiraj, G. Shirane, and Y. Endoh, *Phys. Rev. Lett.* **82**, 628 (1999).
- [17] L. Pintschovius and M. Braden, *Phys. Rev. B* **60**, R15039(R) (1999).
- [18] T. Fukuda, J. Mizuki, K. Ikeuchi, K. Yamada, A. Q. R. Baron, and S. Tsutsui, *Phys. Rev. B* **71**, 060501(R) (2005).
- [19] B. Vignolle, S. M. Hayden, D. F. McMorrow, H. M. Rønnow, B. Lake, C. D. Frost, and T. G. Perring, *Nat. Phys.* **3**, 163 (2007).
- [20] K. Yang, B. P. Xie, D. W. Shen, J. F. Zhao, H. W. Ou, J. Wei, S. Wang, Y. H. Wang, D. H. Lu, R. H. He, M. Arita, S. Qiao, A. Ino, H. Namatame, M. Taniguchi, F. Q. Xu, N. Kaneko, H. Eisaki, and D. L. Feng, *Phys. Rev. B* **73**, 144507 (2006).
- [21] Y. Y. Peng, J.-Q. Meng, L. Zhao, Y. Liu, J.-F. He, G.-D. Liu, X.-L. Dong, S.-L. He, J. Zhang, C.-T. Chen, Z.-Y. Xu, and X. J. Zhou, *Chin. Phys. Lett.* **30**, 067402 (2013).
- [22] M. H. Hettler, A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, *Phys. Rev. B* **58**, R7475 (1998); G. Kotliar, S. Y. Savrasov, G. Palsson, and G. Biroli, *Phys. Rev. Lett.* **87**, 186401 (2001).
- [23] W. Meevasana, X. J. Zhou, S. Sahrakorpi, W. S. Lee, W. L. Yang, K. Tanaka, N. Mannella, T. Yoshida, D. H. Lu, Y. L. Chen, R. H. He, Hsin Lin, S. Komiya, Y. Ando, F. Zhou, W. X. Ti, J. W. Xiong, Z. X. Zhao, T. Sasagawa, T. Kakeshita, K. Fujita, S. Uchida, H. Eisaki, A. Fujimori, Z. Hussain, R. S. Markiewicz, A. Bansil, N. Nagaosa, J. Zaanen, T. P. Devereaux, and Z.-X. Shen, *Phys. Rev. B* **75**, 174506 (2007).
- [24] J. Graf, M. d'Astuto, C. Jozwiak, D. R. Garcia, N. L. Saini, M. Krisch, K. Ikeuchi, A. Q. R. Baron, H. Eisaki, and A. Lanzara, *Phys. Rev. Lett.* **100**, 227002 (2008).
- [25] P. V. Bogdanov, A. Lanzara, S. A. Kellar, X. J. Zhou, E. D. Lu, W. J. Zheng, G. Gu, J.-I. Shimoyama, K. Kishio, H. Ikeda, R. Yoshizaki, Z. Hussain, and Z.-X. Shen, *Phys. Rev. Lett.* **85**, 2581 (2000).
- [26] W. Zhang, G. Liu, L. Zhao, H. Liu, J. Meng, X. Dong, W. Lu, J. S. Wen, Z. J. Xu, G. D. Gu, T. Sasagawa, G. Wang, Y. Zhu, H. Zhang, Y. Zhou, X. Wang, Z. Zhao, C. Chen, Z. Xu, and X. J. Zhou, *Phys. Rev. Lett.* **100**, 107002 (2008).
- [27] S. Vig, A. Kogar, V. Mishra, L. Venema, M. S. Rak, A. A. Husain, P. D. Johnson, G. D. Gu, E. Fradkin, M. R. Norman, and P. Abbamonte, [arXiv:1509.04230](https://arxiv.org/abs/1509.04230).
- [28] H. F. Fong, P. Bourges, Y. Sidis, L. P. Regnault, A. Ivanov, G. D. Gu, N. Koshizuka, and B. Keimer, *Nature (London)* **398**, 588 (1999).
- [29] H. He, Y. Sidis, P. Bourges, G. D. Gu, A. Ivanov, N. Koshizuka, B. Liang, C. T. Lin, L. P. Regnault, E. Schoenher, and B. Keimer, *Phys. Rev. Lett.* **86**, 1610 (2001).
- [30] S. Ideta, K. Takashima, M. Hashimoto, T. Yoshida, A. Fujimori, M. Kubota, K. Ono, K. Kojima, and S. Uchida, *J. Phys.: Conf. Ser.* **108**, 1 (2008).
- [31] S. Ideta, T. Yoshida, M. Hashimoto, A. Fujimori, H. Anzai, A. Ino, M. Arita, H. Namatame, M. Taniguchi, K. Takashima, K. M. Kojima, and S. Uchida, *J. Phys. Conf. Ser.* **428**, 012039 (2013).
- [32] S. V. Borisenko, A. A. Kordyuk, V. Zabolotnyy, J. Geck, D. Inosov, A. Koitzsch, J. Fink, M. Knupfer, B. Büchner, V. Hinkov, C. T. Lin, B. Keimer, T. Wolf, S. G. Chiuzbăian, L. Patthey, and R. Follath, *Phys. Rev. Lett.* **96**, 117004 (2006).
- [33] W. Reichardt, N. Pyka, L. Pintschovius, B. Hennion, and G. Collin, *Phys. C (Amsterdam)* **162**, 464 (1989).
- [34] L. Pintschovius, D. Reznik, and W. Reichardt, *Phys. Rev. B* **69**, 214506 (2004).
- [35] J. Rossat-Mignod, L. P. Regnault, C. Vettier, P. Bourges, P. Burllet, J. Bossy, J. Y. Henry, and G. Lapertot, *Phys. C (Amsterdam)* **185**, 86 (1991).
- [36] H. A. Mook, M. Yethiraj, G. Aeppli, T. E. Mason, and T. Armstrong, *Phys. Rev. Lett.* **70**, 3490 (1993).
- [37] P. Dai, M. Yethiraj, H. A. Mook, T. B. Lindemer, and F. Doğan, *Phys. Rev. Lett.* **77**, 5425 (1996).
- [38] P. Dai, H. A. Mook, R. D. Hunt, F. Doğan, *Phys. Rev. B* **63**, 054525 (2001).
- [39] I. M. Vishik, N. Barišić, M. K. Chan, Y. Li, D. D. Xia, G. Yu, X. Zhao, W. S. Lee, W. Meevasana, T. P. Devereaux, M. Greven, and Z.-X. Shen, *Phys. Rev. B* **89**, 195141 (2014).
- [40] M. d'Astuto, A. Mirone, P. Giura, D. Colson, A. Forget, and M. Krisch, *J. Phys. Condens. Matter* **15**, 8827 (2003).
- [41] Y. Li, V. Balédent, G. Yu, N. Barišić, K. Hradil, R. A. Mole, Y. Sidis, P. Steffens, X. Zhao, P. Bourges, and M. Greven, *Nature* **468**, 283 (2010).

- [42] Y. Li, G. Yu, M. K. Chan, V. Balédent, Y. Li, N. Barišić, X. Zhao, K. Hradil, R. A. Mole, Y. Sidis, P. Steffens, P. Bourges, and M. Greven, *Nat. Phys.* **8**, 404 (2012).
- [43] M. K. Chan, C. J. Dorow, L. Mangin-Thro, Y. Tang, Y. Ge, M. J. Veit, G. Yu, X. Zhao, A. D. Christianson, J. T. Park, Y. Sidis, P. Steffens, D. L. Abernathy, P. Bourges, and M. Greven, *Nat. Commun.* **7**, 10819 (2016).
- [44] Y. Chen, A. Iyo, W. Yang, A. Ino, M. Arita, S. Johnston, H. Eisaki, H. Namatame, M. Taniguchi, T. P. Devereaux, Z. Hussain, and Z.-X. Shen, *Phys. Rev. Lett.* **103**, 036403 (2009).
- [45] F. Ronning, T. Sasagawa, Y. Kohsaka, K. M. Shen, A. Damascelli, C. Kim, T. Yoshida, N. P. Armitage, D. H. Lu, D. L. Feng, L. L. Miller, H. Takagi, and Z.-X. Shen, *Phys. Rev. B* **67**, 165101 (2003).
- [46] N. Mannella, W. L. Yang, X. J. Zhou, H. Zheng, J. F. Mitchell, J. Zaanen, T. P. Devereaux, N. Nagaosa, Z. Hussain, and Z.-X. Shen, *Nature* **438**, 474 (2005).
- [47] T. Valla, A. V. Fedorov, P. D. Johnson, J. Xue, K. E. Smith, and F. J. DiSalvo, *Phys. Rev. Lett.* **85**, 4759 (2000).
- [48] G. Brusdeylins, C. Heimlich, J. G. Skofronick, J. P. Toennis, R. Vollmer, G. Benedek, and L. Miglio, *Phys. Rev. B* **41**, 5707 (1990).
- [49] J. Schäfer, D. Schrupp, E. Rotenberg, K. Rossnagel, H. Koh, P. Blaha, and R. Claessen, *Phys. Rev. Lett.* **92**, 097205 (2004).
- [50] V. Brouet, A. Nicolaou, M. Zacchigna, A. Taleb-Ibrahimi, P. Le Fèvre, and F. Bertran, *J. Electron Spectrosc. Relat. Phenom.* **185**, 146 (2012).

## Supplementary Information: Origin of Kinks in Energy Dispersion of Strongly Correlated Matter

Kazue Matsuyama,<sup>1</sup> Edward Perepelitsky,<sup>2,3</sup> and B Sriram Shastry<sup>1</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, CA 95064*

<sup>2</sup>*Centre de Physique Théorique, École Polytechnique,  
CNRS, Université Paris-Saclay, 91128 Palaiseau, France*

<sup>3</sup>*Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*

(Dated: April 20, 2017)

In this supplemental note we provide (I) some details of the doping dependence of the fit parameters (II) detailed predictions of the electron-Boson coupling model for kinks and (III) detailed predictions of the extremely strong correlation theory for kinks. In the main paper, we have discussed alternate mechanisms for generating the low-energy kink observed in ARPES. Although both mechanisms are capable of generating similar MDC dispersions, they produce EDCs and EDC dispersions which are distinct from one another in several clearly identifiable ways. These differences, detailed below, can be used to distinguish between the two mechanisms using ARPES, especially as higher resolution data becomes available in the future.

### A. Fixing the parameters

The independent parameters in the ECFL expressions for the kink can be taken as  $V_H, V_L, \hat{k}_{kink}$  and  $\Gamma_0$ . These can be fixed with four measurements as we indicate below. While the first three can be measured with precision, the variable  $\Gamma_0$  depends on the temperature and is also quite sensitive to the various experimental conditions including the incident photon energy, thus making it less precisely known than the others; we will perforce be content with rough estimates of this variable. The remaining parameters can be calculated using equation (MS-2) and equation (MS-5) etc. As mentioned above, the theory is overdetermined, in terms of these four parameters, the theory predicts a number of other quantities: a) the dispersion curves for both EDCs and MDCs, b) the location of both EDC and MDC kinks at finite temperature, and c) the spectral functions near the Fermi level (up to roughly the kink energy). Below we present an analysis of the ARPES data of Bi2212, LSCO and Bi2201 taken from literature, where we give the details of the fits and the predicted EDC velocities for future experiments.

The asymptotic velocities  $V_H, V_L$  determine the ratio  $r$  from equation (MS-2). The energy  $\Delta_0$  and the ideal kink energy are determined from equations (MS-5, SI-28). As discussed in Fig. 1  $E_{kink}^{MDC}$  is found by measuring the dispersion at the kink wave vector  $E(\hat{k}_{kink})$ , and similarly the EDC kink energy  $E_{kink}^{EDC}$  is found from  $E^*(\hat{k}_{kink})$ . For understanding the finite temperature data, the theory provides temperature dependent correction terms for

the two spectra, determined by the parameter  $\Gamma_0$ ,

$$E_{kink}^{EDC} = E_{kink}^{ideal} - \Gamma_0, \quad (\text{SI-1})$$

$$E_{kink}^{MDC} = E_{kink}^{ideal} - \Gamma_0 \sqrt{\frac{r}{2-r}}. \quad (\text{SI-2})$$

Since  $\Gamma_0$  determines the non-zero T (or  $\eta$ ) correction, we estimate from the difference between low and high temperature MDC dispersion curves

$$\Gamma_0 = \Delta E_{kink} = \sqrt{\frac{2-r}{r}} (E_{kink}^{ideal} - E_{kink}^{MDC}). \quad (\text{SI-3})$$

Clearly uncertainties in  $\Gamma_0$  are governed by those in the MDC dispersion at the kink momentum.

As noted in Fig. 1, the ECFL theory predicts a kink, rather than a jump in the EDC spectrum, quite analogous to that in the MDC dispersion, but with a different velocity on the steeper side, i.e.  $V_H^* \neq V_H$ . In fact the theory provides an experimentally testable expression relating the two,  $V_H^*$  is expressed quite simply in terms of measurable experimental variables,

$$V_H^* = \frac{3V_H - V_L}{V_H + V_L} \times V_L. \quad (\text{SI-4})$$

As mentioned in the introduction the Boson-mode coupled theories predict a jump in the EDC spectrum at the kink energy. The velocity beyond the jump is the same in EDC and MDC, i.e.  $V_H^* = V_H$ , in contrast to Eq. (SI-4). This velocity is reported in only a few cases, and provides a ready test of the ECFL theory.

The theory also predicts  $V_L = V_L^*$ , which is satisfied by inspection in all reported cases and is common to the Boson-mode theory. We use this protocol to analyze the experiments on three well studied families of high  $T_c$  materials next.

### B. Fit parameters

#### I. $\Delta_0$ FOR LSCO DATA IN THE MAIN TEXT

For the LSCO data discussed in the main text, we quoted the ECFL theory parameters, velocity ratio  $r$ , the ideal kink energy  $E_{kink}^{ideal}$  and the small energy parameter  $\Delta_0$ , in Eqs (2,7,5) (see also Eq. (SI-28)). In Fig. 1, we display the doping dependence of these parameters

$x = 1 - n$ . The size of the data point represents the uncertainty for each data points. While  $r$  and  $\Delta_0$  stay

almost constant, the ideal kink energy decreases linearly with increasing  $x$ .

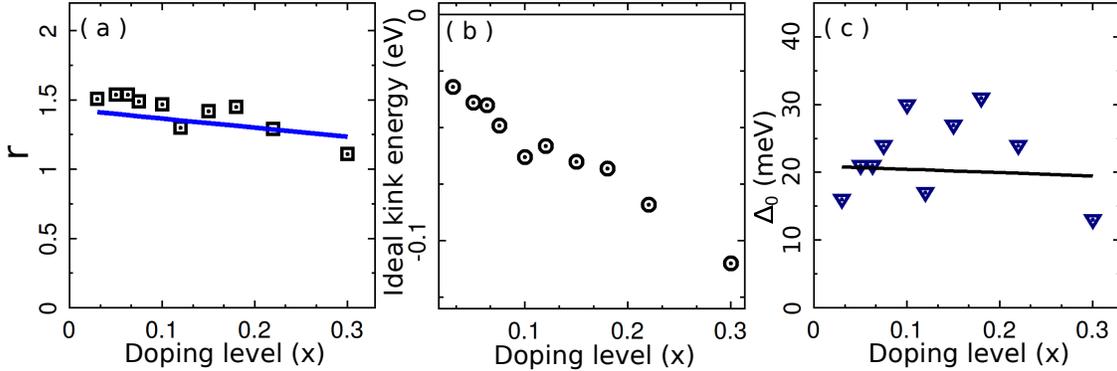


FIG. 1: ( a ) The ratio of low and high velocities,  $r$ , as a function of doping levels, and ( b ) ideal kink energy, ( c ) ECFL energy parameter  $\Delta_0$  as a function of doping levels for LSCO data in the main text.

## II. ELECTRON-BOSON COUPLING THEORY OF KINKS

The electron Boson mechanism suggested in Ref. (1) and others<sup>2,3</sup>, is the coupling of the electrons to Bosonic modes (such as phonons), located at the kink energy. To illustrate the basic idea, we first consider free electrons coupled to an Einstein phonon mode of energy  $\omega_0 = .08$  eV<sup>2,3</sup>, with coupling constant  $g$ . In this case, the spectral function is expressed in terms of a momentum independent self-energy  $\Sigma(\omega)$ , as

$$A(\vec{k}, \omega) = -\frac{1}{\pi} \frac{\Im m \Sigma(\omega)}{(\omega - \xi_k - \Re e \Sigma(\omega))^2 + (\Im m \Sigma(\omega))^2}, \quad (\text{SI-5})$$

where  $\xi_k \equiv \varepsilon_k - \mu$ ,  $\varepsilon_k$  is the bare dispersion, and  $\mu$  is the chemical potential. The real and imaginary parts of the self-energy due to the electron-phonon interactions are given by the well known formulas:<sup>4,5</sup>

$$\begin{aligned} \Im m \Sigma(\omega) &= -\pi g^2 \sum_{\pm} N(\omega + \mu \pm \omega_0) \times \\ &\quad [f^{\mp}(\omega \pm \omega_0) + n(\omega_0)], \\ \Re e \Sigma(\omega) &= -\frac{1}{\pi} \int d\nu \frac{\Im m \Sigma(\nu)}{\omega - \nu}, \end{aligned} \quad (\text{SI-6})$$

where  $f^-(\nu) \equiv f(\nu)$ ,  $f^+(\nu) \equiv \bar{f}(\nu) \equiv 1 - f(\nu)$ ,  $f(\nu)$  and  $n(\nu)$  are the Fermi and Bose distribution functions respectively, and  $N(E) \equiv \frac{1}{N_s} \sum_k \delta(E - \varepsilon_k)$  is the local density of states for the free electrons. Since the relevant frequency range for the self-energy is  $|\omega| \sim \omega_0$ ,

and  $\omega_0 \ll W$ , where  $W$  is the bandwidth, we neglect the frequency dependence in the density of states, i.e.  $N(\omega + \mu \pm \omega_0) \approx N(\mu) \approx N(\varepsilon_f)$ , where  $\varepsilon_f$  is the Fermi energy. Furthermore, the strength of the electron-phonon coupling is given by the dimensionless parameter<sup>6</sup>  $\lambda \equiv \frac{2N(\varepsilon_f)g^2}{\omega_0}$ . Therefore, the imaginary part of the self-energy is expressed directly in terms of  $\lambda$  as

$$\Im m \Sigma(\omega) = -\frac{\pi \lambda \omega_0}{2} \sum_{\pm} [f^{\mp}(\omega \pm \omega_0) + n(\omega_0)]. \quad (\text{SI-7})$$

We initially choose a typical intermediate strength value of  $\lambda = 0.5$ . We also add a small broadening  $\eta = .01$  eV to the imaginary part of the self-energy. In Fig. (2), we display  $\omega - \Re e \Sigma(\omega)$  and  $-\Im m \Sigma(\omega)$  vs.  $\omega$  (left panel), the EDC and MDC dispersions (middle panel), as well as the EDCs at several representative momenta (right panel) at  $T = 10$  K. The EDC and MDC dispersions as well as the EDCs can be understood directly from the real and imaginary parts of the self-energy using Eq. (SI-5). From Eq. (SI-5), the the MDC at fixed  $\omega$  is a Lorentzian of width  $-\Im m \Sigma(\omega)$  and peak position  $\xi^*(\omega) = \omega - \Re e \Sigma(\omega)$ <sup>1</sup>. Therefore, the MDC dispersion is obtained by inverting  $\xi^*(\omega)$  to obtain  $E(\xi)$ . Since  $\omega - \Re e \Sigma(\omega)$  is not one-to-one,  $E(\xi)$  is a multi-valued function.

To understand the EDC dispersion, we first examine the EDC curves in the right panel of Fig. (2). The momentum  $\xi$  associated with each curve is given by the location of the corresponding horizontal dashed line along the vertical axis in the left panel. The EDC at each momentum has two distinguishable features, a peak followed

by a hump. In the left panel, the red and green dots indicate the location of the peak and hump, respectively, at each momentum, as determined directly from the EDC.

We partition the EDCs into three distinct momentum regions,  $|\xi| < |\xi_1|$ ,  $|\xi_1| < |\xi| < |\xi_2|$ , and  $|\xi| > |\xi_2|$ , where the momenta  $\xi_1$  and  $\xi_2$  (the low-energy kink momentum) are denoted by the dashed vertical lines in the middle panel of Fig. (2). In the first region,  $|\xi| < |\xi_1|$ , the peak location,  $E_p^*$ , disperses according to the equation  $\xi = E_p^* - \Re e \Sigma(E_p^*)$ , while the hump location,  $E_h^*$ , remains at a fixed frequency, displayed by the horizontal dashed

line in the middle panel. In addition, there is a sharp dip between the peak and the hump which is pinned to the phonon frequency,  $-\omega_0$ . Since  $\Im m \Sigma(E_p^*)$  is constant throughout this region, the height of the peak does not change. On the other hand, since  $|E_h^* - \xi - \Re e \Sigma(E_h^*)|$  decreases as  $|\xi|$  is increased (and of course  $\Im m \Sigma(E_h^*)$  is constant), the hump height grows as  $|\xi|$  approaches  $|\xi_1|$ . Nevertheless, since the peak height remains greater than the hump height throughout this region (as will be shown below), the EDC dispersion is given by  $E^* = E_p^*$ .

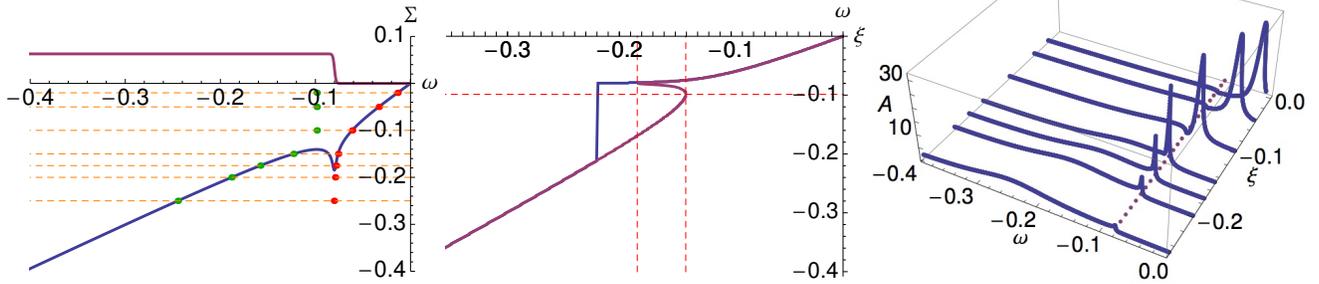


FIG. 2: Results for free electrons coupled to an Einstein phonon mode of frequency  $\omega_0 = .08$  eV, with coupling strength  $\lambda = 0.5$ , at  $T = 10$  K. **Right** panel: The EDCs at several representative momenta, the variable  $\xi = v_f(k - k_F) = (1 + \lambda)V_L(k - k_F)$  here and in later figures. The dashed line indicates the phonon frequency,  $\omega = -\omega_0$ . Each EDC has two well-defined features, a peak followed by a hump (separated by a sharp dip for low momentum EDCs). **Middle** panel: The MDC dispersion (magenta) has no jump while the EDC dispersion (blue) shows a jump. The two vertical dashed lines partition momentum space into three regions. The horizontal dashed line indicates the location of the hump in the EDCs in the first (low-momentum) region. In the first two regions, the EDC dispersion follows the MDC dispersion (closest to zero frequency), while in the third (high momentum) region, it stays pinned to the phonon frequency over a large range of momentum, until it discontinuously jumps back down to the MDC dispersion. Note that  $V_H = V_H^*$ . **Left** panel:  $\omega - \Re e \Sigma(\omega)$  and  $-\Im m \Sigma(\omega)$  vs.  $\omega$ . The horizontal dashed lines indicate the momenta associated with the corresponding EDCs in the right panel. The red dots indicate the locations of the peaks, and the green dots indicate the locations of the humps, as determined directly from each EDC.

In the second region,  $|\xi_1| < |\xi| < |\xi_2|$ , both  $E_p^*$  and  $E_h^*$  disperse according to the equation  $\xi = E_{p,h}^* - \Re e \Sigma(E_{p,h}^*)$ ,  $E_p^*$  being the root closest to, and  $E_h^*$  being the root farthest from, zero frequency. Since  $\Im m \Sigma(E_p^*)$  continues to remain constant and has the same value as in the first region, so does the height of the peak. Moreover, since  $\Im m \Sigma(E_h^*)$  remains constant as well, the height of the hump remains the one which it reached at  $\xi = \xi_1$ . Finally, since  $|\Im m \Sigma(E_h^*)| > |\Im m \Sigma(E_p^*)|$ , the peak height is greater than the hump height, and therefore  $E^* = E_p^*$ .

In the third region,  $|\xi| > |\xi_2|$ ,  $E_p^*$  is pinned to the phonon frequency  $-\omega_0$ , while  $E_h^*$  continues to disperse according to the equation  $\xi = E_h^* - \Re e \Sigma(E_h^*)$ . Since  $\Im m \Sigma(E_h^*)$  continues to have the same value as in the second region, so does the height of the hump. Meanwhile, the peak height decreases, since  $|E_p^* - \xi - \Re e \Sigma(E_p^*)|$

increases as  $|\xi|$  is increased. Although initially  $E^* = E_p^* = -\omega_0$ , eventually, after  $|\xi|$  has been sufficiently increased, the peak height falls below the hump height, and  $E^* = E_h^*$ . Accordingly, in the middle panel, we see that in first two regions, the EDC dispersion follows the MDC dispersion,  $E^* = E$  (closest to zero frequency). However, in the third region,  $E^*$  stays fixed at  $-\omega_0$ , until at sufficiently high momentum, it jumps back down to the MDC dispersion. Since the MDC and EDC dispersions coincide for large momentum, the velocities  $V_H$  and  $V_H^*$  are equal. We take these three features, a discontinuous jump in the EDC dispersion, a peak pinned to the phonon frequency in the EDC over a prolonged range of momentum, and the equality  $V_H = V_H^*$ , to be signatures of electron-Boson coupling in ARPES experiments. Similar calculations to the one above can be found in<sup>1,2</sup>, with analogous results.

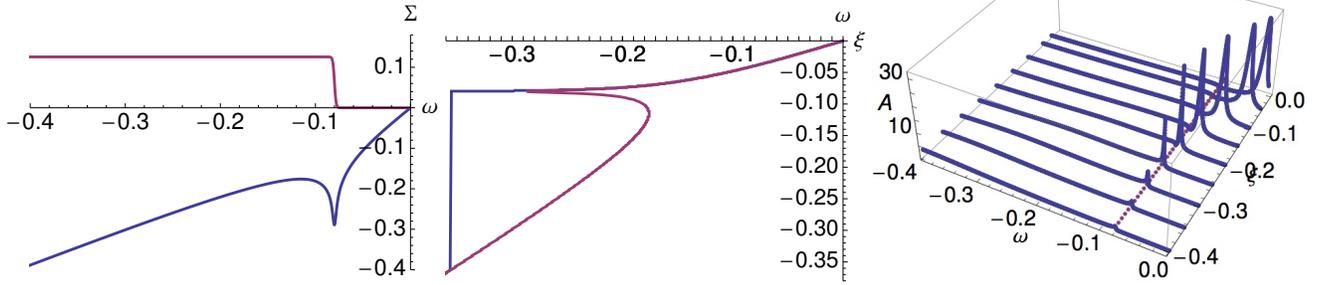


FIG. 3: To explore the effects of raising  $\lambda$ , we set  $\lambda = 1$  while leaving all other parameters unchanged from Fig. (2). As a result, the kink momentum in the MDC dispersion becomes bigger, the hump in the EDCs is suppressed, the EDC dispersion stays pinned to the phonon frequency over a larger range of momentum, and the magnitude of the jump in the EDC dispersion grows.

To examine the effects of raising  $\lambda$ , we set  $\lambda = 1$  leaving all other parameters unchanged, and plot the corresponding results in Fig. (3). This causes several noticeable changes to the results in Fig. (2). 1) The kink in the real part of the self-energy becomes sharper, which leads to a larger kink momentum,  $\xi_2$ , in the MDC dispersion. 2)  $-\Im m\Sigma(E_h^*)$  becomes bigger, causing the height of the hump to go down. 3) As a direct consequence of 2), the range over which the EDC dispersion stays pinned to the phonon frequency becomes more prolonged in momentum space, and therefore the magnitude of the jump in

the EDC dispersion also becomes bigger.

Setting  $T \rightarrow 0$  in Eq. (SI-7), and plugging it into Eq. (SI-6), we find that to linear order in  $\omega \ll \omega_0$ ,  $\Re\Sigma(\omega) = -\lambda\omega$ . Therefore,  $\lambda = \frac{v_f}{V_L} - 1$  (see also<sup>7</sup>). According to the normal state data ( $T = 115$  K) from<sup>3,8,9</sup> (since  $T \ll \omega_0$ , this zero temperature formula still applies),  $V_L = 1.47\text{eV}\text{\AA}$  and  $v_f = 2.7\text{eV}\text{\AA}$ , yielding  $\lambda = 0.84$ . In principle, one might argue for the larger value of  $v_f \sim 5.4\text{ eV}\text{\AA}$  from the ARPES observed width of the band<sup>16</sup>, leading to  $\lambda \sim 2.67$ , a very high value indeed.

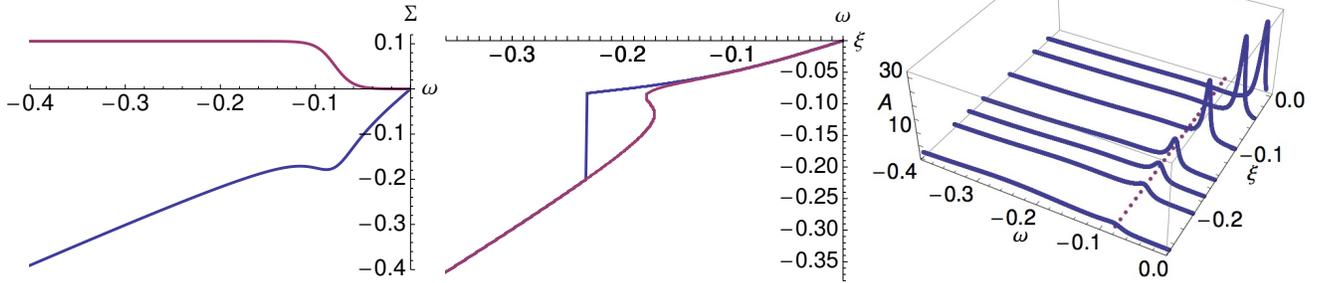


FIG. 4: We now use the experimentally relevant values of  $\lambda = 0.84$  and  $T = 115$  K. The curves retain the same qualitative features as in Fig. (2), which are less sharp in the present case due to the higher value of  $T$ .

However, we will assume, with several authors of the Boson-coupling models, that the smaller estimate is overall more reasonable. Using these experimentally relevant values, in Fig. (4), we plot  $\omega - \Re\Sigma(\omega)$  and  $-\Im m\Sigma(\omega)$  vs.

$\omega$  (left panel), as well as the MDC and EDC dispersions (middle panel), and the EDCs at several representative momenta (right panel). Due to the higher value of  $T$ , the self-energy curves have been rounded out somewhat

as compared to Fig. (2), but retain the same features. We see that the EDC dispersion once again follows the MDC dispersion (closest to zero frequency) in the first two momentum regions, until it (nearly) flattens out in the third region, where the peak is pinned to the phonon frequency,  $-\omega_0$ , in the corresponding EDCs. As the momentum is

increased such that the height of this peak shrinks below the height of the hump, the EDC dispersion jumps discontinuously down from the phonon frequency, to the MDC dispersion. Consequently, we see that the velocities of the MDC and EDC dispersion coincide above the kink; i.e.  $V_H = V_H^*$ .

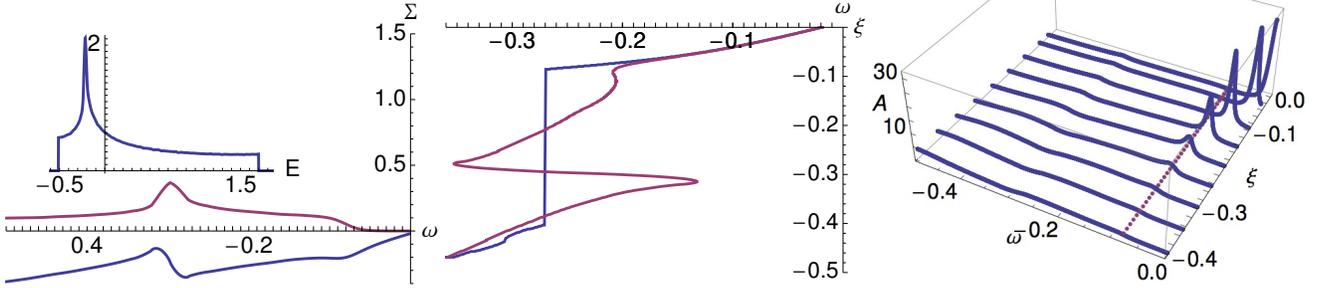


FIG. 5: We explore the effects of using the full frequency-dependence of the density of states in Eq. (SI-6), with  $\lambda = 0.84$  and  $T = 115$  K. Due to the functional form of the density of states (displayed as an inset in the left panel), the MDC dispersion acquires two additional branches which yield large frequency values. Below the low-energy kink momentum, the EDC dispersion follows the lowest-frequency branch of the MDC dispersion. Above the low-energy kink momentum, the EDC dispersion initially stays pinned to the phonon frequency, until it discontinuously jumps onto the highest-frequency branch of the MDC dispersion ( $V_H = V_H^*$ ). A noticeable hump also develops at high-frequencies, in the corresponding EDCs.

We now examine how these results are affected by retaining the full frequency-dependence of the density of states in Eq. (SI-6). Just as was done in<sup>3</sup>, we use the dispersion  $\text{tb2}$  from<sup>9</sup>. In this case,  $\varepsilon_f = 0$  and  $N(\varepsilon_f) = 0.61 \text{ eV}^{-1}$ . Retaining the same values of  $T = 115$  K and  $\lambda = 0.84$ , we set  $g = 0.23 \text{ eV}$  in Eq. (SI-6). We also set  $\mu \approx \varepsilon_f = 0$ . In Fig. (5), we plot  $\omega - \Re e \Sigma(\omega)$  and  $-\Im m \Sigma(\omega)$  vs.  $\omega$  (left panel), as well as the MDC and EDC dispersions (middle panel), and the EDCs at several representative momenta (right panel). Due to the functional form of the density of states (see the inset of the left panel), the MDC dispersion acquires two additional branches which yield large frequency values. In the first two momentum regions (below the low-energy kink momentum), the EDC dispersion follows the lowest-frequency branch of the MDC dispersion. As the momentum increases into the third region (above the low-energy kink momentum), the peak stays pinned to the phonon frequency in the corresponding EDCs. Moreover, since  $|\Im m \Sigma(E(\xi))| \gg |\Im m \Sigma(-\omega_0)|$ , where  $E(\xi)$  can be any branch of the MDC dispersion, the EDC dispersion stays pinned to the phonon frequency as well. As the momentum is increased further and the height of the peak decreases sufficiently, the EDC dispersion jumps discontinuously onto the highest-frequency branch of the MDC dispersion, since this is the one with the smallest value of  $|\Im m \Sigma(E(\xi))|$ , and hence  $V_H = V_H^*$ .

This small value of  $|\Im m \Sigma(E(\xi))|$  leads to a noticeable hump at high-frequencies in the corresponding EDCs.

Thus far, we have considered only free electrons coupled to a Boson mode. We now include electron-electron correlations. Following<sup>10</sup>, we assume that

$$\Im m \Sigma_{el-el}(\omega) = -\frac{(\tau^2 + \omega^2)}{\Omega_0} \exp\left[-\frac{(\tau^2 + \omega^2)}{\nu_0^2}\right] - \eta, \quad (\text{SI-8})$$

where  $\Sigma_{el-el}(\omega)$  is the self-energy due only to electron-electron correlations,  $\tau \equiv \pi k_B T$ ,  $T = 115$  K,  $\Omega_0 = .14 \text{ eV}$ ,  $\nu_0 = .5 \text{ eV}$ , and we set  $\eta = .01 \text{ eV}$ . This phenomenological form for  $\Im m \Sigma_{el-el}(\omega)$  reproduces the correct Fermi-liquid behavior at low frequencies, and extrapolates to high frequencies in a reasonable way. Furthermore, we assume a flat band for  $\varepsilon_k$  of bandwidth  $W$ , i.e.  $N(E) = \frac{1}{W} \Theta(\frac{W}{2} - |E|)$ , and set  $\mu \approx \varepsilon_f = 0$ . Retaining the same values of  $N(\varepsilon_f) = 0.61 \text{ eV}^{-1}$  and  $\lambda = 0.84$  as before, yields the values  $W = 1.64 \text{ eV}$  and  $g = 0.23 \text{ eV}$ . The self-energy is now given by the sum  $\Sigma(\omega) = \Sigma_{el-el}(\omega) + \Sigma_{el-ph}(\omega)$ , where the imaginary part of the latter term is

$$\Im m \Sigma_{el-ph}(\omega) = -\pi g^2 \sum_{\pm} A_{el-el,loc}(\omega \pm \omega_0) \times [f^{\mp}(\omega \pm \omega_0) + n(\omega_0)], \quad (\text{SI-9})$$

while the real part is as usual given by applying the Hilbert transform to Eq. (SI-9). Here,  $A_{\text{el-el,loc}}(\omega) = \frac{1}{N_s} \sum_{\vec{k}} A_{\text{el-el}}(\vec{k}, \omega)$ , where  $A_{\text{el-el}}(\vec{k}, \omega)$  is given by Eq. (SI-5) with the substitution  $\Sigma(\omega) \rightarrow \Sigma_{\text{el-el}}(\omega)$ . Eq. (SI-5) continues to express  $A(\vec{k}, \omega)$  in terms of  $\Sigma(\vec{k}, \omega)$ , where both objects now include electron-electron and electron-phonon correlations.

In Fig. (6), we plot  $\omega - \Re e \Sigma(\omega)$  and  $-\Im m \Sigma(\omega)$  vs.  $\omega$  (left panel), as well as the MDC and EDC dispersions (middle panel), and the EDCs at several representative momenta (right panel), from this calculation. Due to the specific form of the self-energy,  $\Sigma_{\text{el-el}}(\omega)$  (both  $-\Im m \Sigma_{\text{el-el}}(\omega)$  and  $A_{\text{el-el,loc}}(\omega)$ ) are displayed as an inset in the left panel), the highest-frequency branch of the MDC dispersion yields very large values of the frequency. Just as in the cases considered above, for momentum  $|\xi|$  below the low-energy kink momentum, the EDC dispersion follows the lowest-frequency branch of the MDC dispersion,  $E_l(\xi)$ . As the momentum  $|\xi|$  is increased above the low-energy kink momentum, the rapid increase in

$|\Im m \Sigma(E_l(\xi))|$  causes the peak in the EDC as well as the EDC dispersion to stay pinned to the phonon frequency. As the momentum is increased further,  $|\Im m \Sigma(E_h(\xi))|$  becomes comparable to  $|\Im m \Sigma(-\omega_0)|$ , where  $E_h(\xi)$  is the highest-frequency branch of the MDC dispersion. At this point, the EDC dispersion jumps discontinuously from the phonon frequency onto the highest-frequency branch of the MDC dispersion, and hence  $V_H = V_H^*$ . This is also reflected in the corresponding EDCs, which acquire a hump at high-frequencies.

In conclusion, we find that in all of the above cases of electrons interacting with a Boson mode, the EDCs are characterized by three signatures: (1) a peak pinned to the Boson-frequency over a large range of momentum, (2) the EDC dispersion jumps discontinuously from the Boson-frequency onto (the highest-frequency branch of) the MDC dispersion, and (3)  $V_H = V_H^*$ . These three features are jointly present for most parameters explored, and may be viewed as the signatures of kinks produced by this mechanism.

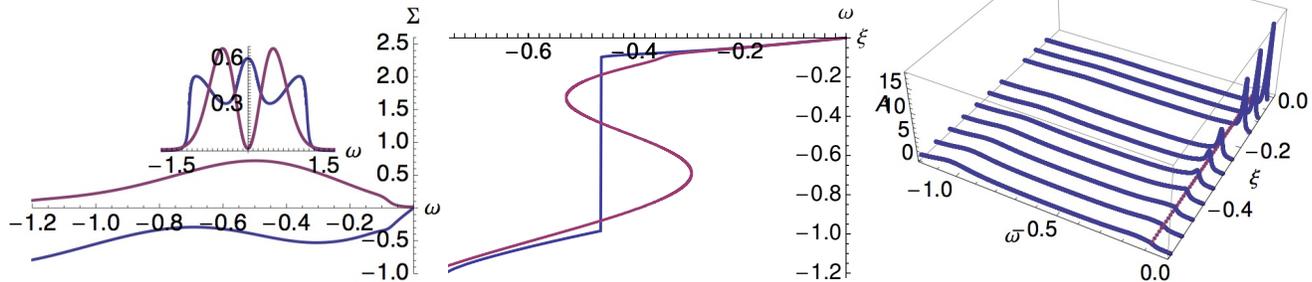


FIG. 6: We explore the effects of Fermi-liquid-like electron-electron correlations (Eq. (SI-8)), with  $\lambda = 0.84$  and  $T = 115$  K. Due to the functional form of the self-energy,  $\Sigma_{\text{el-el}}(\omega)$  (both  $-\Im m \Sigma_{\text{el-el}}(\omega)$  and  $A_{\text{el-el,loc}}(\omega)$ ) are displayed as an inset in the left panel), the highest-frequency branch of the MDC dispersion yields very large values of the frequency. Below the low-energy kink momentum, the EDC dispersion follows the lowest-frequency branch of the MDC dispersion. Above the low-energy kink momentum, the EDC dispersion initially stays pinned to the phonon frequency, until it discontinuously jumps onto the highest-frequency branch of the MDC dispersion ( $V_H = V_H^*$ ). This is also reflected in the corresponding EDCs, which acquire a hump at high-frequencies.

### III. EXTREMELY CORRELATED FERMILIQUID THEORY OF KINKS

In this section we present the theoretical details of the ECFL calculation of kinks. We first show the results of a low energy and momentum expansion of the ECFL Greens function in terms of a few parameters. Earlier studies<sup>11-13</sup> show that the two self energies  $\Phi, \Psi$  of the ECFL theory are to a large extent similar to the self energies of a standard intermediate coupling Fermi liquid, and yet due to their specific combination that oc-

curs in Eq. (SI-10) and Eq. (SI-12) end up providing a non trivial resulting theory. Indeed in Ref. (11) a similar low energy expansion in high dimensions, was tested successfully against the numerical results of the Dynamical Mean Field Theory (DMFT). It should be noted that the DMFT theory is designed for high dimensions, where the momentum dependence of the Dyson self energy and  $\Psi$  self energy of the ECFL theory drops out. In this section we allow for momentum dependence of both self energies in the ECFL formalism, this is in-fact the only distinction between the present expansion and that in Ref. (11).

We see below that this momentum dependence is essential for describing the low energy kinks in the occupied part of the ARPES spectrum.

### A. Low energy expansion of the ECFL theory

We start with the ECFL Greens function  $\mathcal{G}$  expressed in terms of the auxiliary Greens function  $\mathbf{g}$  and the caparison function  $\tilde{\mu}$  Ref. (14) and Ref. (15), we write

$$\mathcal{G}(\vec{k}, i\omega) = \mathbf{g}(\vec{k}, i\omega) \times \tilde{\mu}(\vec{k}, i\omega), \quad (\text{SI-10})$$

and with the latter expressed in terms of the two self energies  $\Phi(\vec{k}, i\omega_n), \Psi(\vec{k}, i\omega_n)$  as:

$$\tilde{\mu}(\vec{k}, i\omega_n) = 1 - \frac{n}{2} + \Psi(\vec{k}, i\omega_n) \quad (\text{SI-11})$$

$$\mathbf{g}^{-1}(\vec{k}, i\omega_n) = i\omega_n + \boldsymbol{\mu} - (1 - \frac{n}{2})\varepsilon_k - \Phi(\vec{k}, i\omega_n), \quad (\text{SI-12})$$

where  $n$  is the electron number per site,  $\omega_n = (2n+1)\pi/\beta$  the Matsubara frequency, which we analytically continue  $i\omega \rightarrow \omega + i0^+$ . Let us define  $\hat{k}$  as the *normal deviation* from the Fermi surface i.e.  $\hat{k} = (\vec{k} - \vec{k}_F) \cdot \vec{\nabla}_{\varepsilon_{k_F}} / |\vec{\nabla}_{\varepsilon_{k_F}}|$ . Our first objective is to Taylor expand these equations for small  $\omega$  and  $\hat{k}$ , as explained above. We carry out a low frequency expansion as follows:

$$1 - \frac{n}{2} + \Psi(\vec{k}, \omega) = \alpha_0 + c_\Psi(\omega + \nu_\Psi \hat{k} v_f) + i\mathcal{R}/\gamma_\Psi + \mathcal{O}(\omega^3), \quad (\text{SI-13})$$

where the frequently occurring Fermi liquid function  $\mathcal{R} = \pi\{\omega^2 + (\pi k_B T)^2\}$ ,  $v_f = (\partial_k \varepsilon_k)_{k_F}$  is the *bare* Fermi velocity, and the four parameters  $\alpha_0, c_\Psi, \nu_\Psi, \gamma_\Psi$  are coefficients in the Taylor expansion having suitable dimensions. Similarly we expand the auxiliary Greens function

$$\mathbf{g}^{-1}(k, \omega) = (1 + c_\Phi) \left( \omega - \nu_\Phi \hat{k} v_f + i\mathcal{R}/\Omega_\Phi + \mathcal{O}(\omega^3) \right), \quad (\text{SI-14})$$

where we have added another three coefficients in the Taylor expansion  $c_\Phi, \nu_\Phi, \Omega_\Phi$ .

To carry out this reduction we first trade the two parameters  $c_\Psi, \gamma_\Psi$  in favor of parameters  $\Omega_\Psi$  and  $s$  by defining  $c_\Psi = \frac{\alpha_0}{\Omega_\Psi}$  and  $\gamma_\Psi = \frac{s\Omega_\Psi}{c_\Psi}$ , where the dimensionless parameter  $0 \leq s \leq 1$ . With these expansions and the quasiparticle weight determined in terms of the expansion parameters as  $Z = \frac{\alpha_0}{1+c_\Phi}$ , we find

$$\mathcal{G} = \frac{Z}{\Omega_\Psi} \frac{\Omega_\Psi + \omega + \nu_\Psi \hat{k} v_f + i\mathcal{R}/(s\Omega_\Phi)}{\omega - \nu_\Phi \hat{k} v_f + i\mathcal{R}/\Omega_\Phi}. \quad (\text{SI-15})$$

Using  $A(\hat{k}, \omega) = -\frac{1}{\pi} \Im m \mathcal{G}$  we find the spectral function

$$A(\hat{k}, \omega) = \frac{Z}{\pi} \frac{\frac{\mathcal{R}}{\Omega_\Phi}}{(\omega - \nu_\Phi \hat{k} v_f)^2 + (\frac{\mathcal{R}}{\Omega_\Phi})^2} \times \tilde{\mu}_c(\hat{k}, \omega) \quad (\text{SI-16})$$

Here the caparison *factor*, (not to be confused with the caparison *function* in Eq. (SI-10)), is found as

$$\begin{aligned} \tilde{\mu}_c(\hat{k}, \omega) &= 1 - \xi(\hat{k}, \omega) \\ \xi(\hat{k}, \omega) &= \frac{1}{\Delta_0} (\omega - \nu_0 \hat{k} v_f) \end{aligned} \quad (\text{SI-17})$$

In Eq. (SI-17) we have introduced two composite parameters

$$\Delta_0 = \frac{s}{1-s} \Omega_\Psi, \quad \text{and} \quad \nu_0 = \frac{1}{1-s} \nu_\Phi + \frac{s}{1-s} \nu_\Psi \quad (\text{SI-18})$$

This procedure eliminates the *three* old parameters  $s, \Omega_\Psi$  and  $\nu_\Psi$  in favor of the *two* emergent energy scale  $\Delta_0$  and velocity  $\nu_0$ .

It is interesting to count the reduction in the number of free parameters from the starting value of seven in Eq. (SI-13) and Eq. (SI-14). Already in Eq. (SI-15) we have a reduction to six, since the quasiparticle weight  $Z$  combines two of the original parameters. Since Eq. (SI-18) subsumes three parameters into two, the spectral function in Eq. (SI-16) contains only five parameters: the two velocities  $\nu_0 v_f, \nu_\Phi v_f$ , and the two energies  $\Omega_\Phi, \Delta_0$ , in addition to the overall scale factor  $Z$ .

We will see below that the parameters that are measurable from energy dispersions are best expressed in terms of certain combinations of the velocities. In order to make the connection with the experiments close, we will redefine the two velocities in terms of an important dispersion velocity at the lowest energies  $V_L$  and a dimensionless ratio  $r$ , on using the definitions:

$$\begin{aligned} \nu_\Phi v_f &= V_L \\ \nu_0 v_f &= r \times V_L. \end{aligned} \quad (\text{SI-19})$$

In order to account for the difference between laser ARPES and synchrotron AREPS having different incident photon energies, we will make two phenomenological modifications in Eq. (SI-16) following Ref. (16)

$$\mathcal{R}(\omega)/\Omega_\Phi \rightarrow \mathcal{R}(0)/\Omega_\Phi = \pi\{\pi k_B T\}^2/\Omega_\Phi + \eta \equiv \mathbb{I} \quad (\text{SI-20})$$

where  $\eta$  represents an elastic energy from impurity scattering, dependent upon the energy of the incident photon in the ARPES experiments. In the spirit of a low energy expansion  $\mathcal{R}$  is evaluated at  $\omega = 0$ . Thus  $\Gamma_0$  is a T dependent constant, which subsumes the two parameters  $\eta$  and  $\Omega_\Phi$ , and thus the total parameter count is still five. Secondly for extension to higher energies, we “renormalize” the parameter  $\xi$  in Eq. (SI-17) according to a recently discussed prescription following from a theoretical calculation Ref. (17) as  $\tilde{\mu}_c \rightarrow \{1 - \frac{\xi}{\sqrt{1+c_a\xi^2}}\}$ , where  $c_a \sim 5.4$  near optimum doping  $\delta \sim 0.15$  as estimated recently. This correction ensures that the caparison factor exhibits the correct linear behavior for small  $\xi$ , and remains positive definite at high energies. Thus we write the spectral

function in terms of the new variables as

$$A(\vec{k}, \omega) = \frac{Z}{\pi} \frac{\Gamma_0}{(\omega - V_L \hat{k})^2 + \Gamma_0^2} \times \left\{ 1 - \frac{\xi}{\sqrt{1 + c_a \xi^2}} \right\}, \quad (\text{SI-21})$$

with  $\xi = \frac{1}{\Delta_0}(\omega - r V_L \hat{k})$ . We should keep in mind that these expressions follow from a low energy expansion, and is limited to small  $\hat{k}$  and  $\omega$ , so that the dimensionless variable  $|\xi|_{max} \sim \mathcal{O}(1)$ . Microscopic calculations of all these parameters is possible in the ECFL theory. One important parameter is the energy scale  $\Delta_0$  which is found to be much reduced from the band width, due to extremely strong correlations. A related energy is the effective Fermi liquid temperature scale where the  $T^2$  dependence of the resistivity gives way to a linear dependence. This scale is estimated in the limit of large dimensions from Ref. (17) to be as low as 45 K near optimum doping, i.e. much reduced from naive expectations.

For the present purposes we take a different track, we note that the ARPES fits are overdetermined, so that we can determine the few parameters of the low energy theory from a fairly small subset of measurements. The five final (composite) parameters defining the spectral function Eq. (SI-21) are  $Z, V_L, r, \Delta_0, \Gamma_0$ , where  $c_a \sim 5.4$ . Of these  $Z$  is multiplicative, it is only needed for getting the absolute scale of the spectral function, and  $c_a$  does not play a significant role near zero energy, it is required only at high energies. Thus the spectra relevant to EDC and MDC will require only *four* parameters  $V_L, r, \Delta_0, \Gamma_0$ . These suffice to determine the low energy theory and thus to make a large number of predictions; i.e. implying non trivial relationships amongst observables. Many of the predictions rely only on the overall structure of the theory and not its details.

## B. The EDC and MDC dispersion relations and kinks

Starting from Eq. (SI-21), we can compute the energy dispersions for MDC (varying  $\hat{k}$  while keeping  $\omega$  fixed) and the EDC spectra (varying  $\omega$  while keeping  $\hat{k}$  fixed). In terms of a momentum type variable

$$Q(\hat{k}) = \Delta_0 + (r-1)\hat{k} V_L \quad (\text{SI-22})$$

we can locate the peaks of Eq. (SI-21) using elementary calculus since  $c_a$  only plays a role at high energies, we set  $c_a \rightarrow 0$  when performing the extremization and find the MDC dispersion

$$E(k) = \frac{1}{2-r} \left( \hat{k} V_L + \Delta_0 - \sqrt{r(2-r)\Gamma_0^2 + Q^2} \right), \quad (\text{SI-23})$$

and the EDC dispersion

$$E^*(k) = \left( r \hat{k} V_L + \Delta_0 - \sqrt{\Gamma_0^2 + Q^2} \right). \quad (\text{SI-24})$$

Using these two dispersions and expanding them in different regimes, we can extract all the parameters of the kinks.

### 1. Kink momentum

As explained in the main paper, when we set  $T = 0 = \eta$  so that  $\Gamma_0 = 0$ , both the EDC and MDC dispersions contain an ideal kink at the kink momentum. Therefore, using Eqs. (SI-23) and (SI-24), the condition  $Q = 0$  locates the kink momentum for both dispersions:

$$\hat{k}_{kink} = \frac{\Delta_0}{(1-r)V_L}, \quad (\text{SI-25})$$

it corresponds to occupied momenta, i.e.  $\hat{k}_{kink} v_f < 0$ , provided that  $r > 1$ . We thus can express  $\Delta_0 = \hat{k}_{kink} V_L (1-r)$ , enabling us to usefully rewrite

$$Q = (r-1)V_L(\hat{k} - \hat{k}_{kink}) = \Delta_0 \left\{ 1 - \frac{\hat{k}}{\hat{k}_{kink}} \right\}. \quad (\text{SI-26})$$

As required by the ideal kink,  $Q$  changes sign at the kink momentum,

$$\text{sign}(Q) = \text{sign}(\hat{k} - \hat{k}_{kink}). \quad (\text{SI-27})$$

### 2. Ideal Kink energies: $T=0$

Using Eq. (SI-23) and Eq. (SI-24), in conjunction with Eq. (SI-25), the ideal kink energy is the same for both dispersions, and is given by

$$E_{kink}^{ideal} = -\frac{1}{r-1}\Delta_0. \quad (\text{SI-28})$$

We can also usefully estimate this ideal kink energy from the asymptotic velocities in the far zone, as explained in the main paper.

### 3. The non-ideal i.e. $T > 0$ kink energy

The EDC and MDC kink energies for the non-ideal case can be viewed in a couple of ways. We have argued in the main paper that these are best defined by fixing the momentum  $\hat{k} = \hat{k}_{kink}$  and reading off the energy at this value. This is an unambiguous method independent of the detailed shape of the kink, since it only requires knowledge of  $\hat{k}_{kink}$ , which can be found from an asymptotic measurement as we have argued in the main paper. We can put  $Q = 0$  and  $\hat{k} \rightarrow \hat{k}_{kink}$  in Eq. (SI-24) and Eq. (SI-23) and read off the kink energies:

$$E_{kink}^{EDC} = E_{kink}^{ideal} - \Gamma_0, \quad (\text{SI-29})$$

$$E_{kink}^{MDC} = E_{kink}^{ideal} - \Gamma_0 \sqrt{\frac{r}{2-r}}. \quad (\text{SI-30})$$

We observe that the MDC kink energy is real provided  $2 \geq r \geq 1$ . Note also that at  $T = 0$  and  $\eta = 0$ , the two energies both reduce to the ideal kink energy.

#### 4. The ideal energy dispersions

At  $T = 0$  or for  $|Q| \gg \Gamma_0$ , the two dispersions Eq. (SI-24) and Eq. (SI-23) become:

$$E^*(k) \sim \left[ r - (r-1) \text{sign}(\hat{k} - \hat{k}_{kink}) \right] \hat{k} V_L + 2\Delta_0 \Theta(\hat{k}_{kink} - \hat{k}) \quad (\text{SI-31})$$

and

$$E(k) \sim \frac{1}{2-r} \left[ 1 - (r-1) \text{sign}(\hat{k} - \hat{k}_{kink}) \right] \hat{k} V_L + \frac{2\Delta_0}{2-r} \Theta(\hat{k}_{kink} - \hat{k}). \quad (\text{SI-32})$$

The velocities in the asymptotic regime  $|\hat{k}| \gg \hat{k}_{kink}$  can be found from the slopes of these, and are therefore temperature-independent. For  $\hat{k} \gg \hat{k}_{kink}$  we get the ‘‘low’’ velocities

$$\begin{aligned} \frac{dE(k)}{d\hat{k}} &= V_L \\ \frac{dE^*(k)}{d\hat{k}} &= V_L^* = V_L \end{aligned} \quad (\text{SI-33})$$

and thus the EDC and MDC velocities are identical. For  $\hat{k} \ll \hat{k}_{kink}$  we get the ‘‘high’’ velocities

$$V_H = \frac{dE(k)}{d\hat{k}} = \frac{r}{2-r} V_L, \quad (\text{SI-34})$$

$$V_H^* = \frac{dE^*(k)}{d\hat{k}} = (2r-1)V_L. \quad (\text{SI-35})$$

We may cast Eq. (SI-35) into an interesting form

$$V_H^* = \left\{ \frac{3V_H - V_L}{V_H + V_L} \right\} V_L, \quad (\text{SI-36})$$

it is significant since the EDC spectrum velocity is exactly determined in terms of the two MDC spectrum velocities. It is also a testable result, we show elsewhere in the paper how this compares with known data. Note that the four independent parameters  $V_L, r, \Delta_0, \Gamma_0$  alluded to in the discussion below Eq. (SI-21), can be determined from the directly measurable parameters  $V_L, V_H, \hat{k}_{kink}, \Gamma_0$  (SI-34, SI-25, SI-3). Therefore, either set of parameters gives complete knowledge of the EDC and MDC dispersions, as well as the spectral function (up to an overall scale).

#### 5. Near Zone: Corrections to Energy dispersion due to finite T.

In the regime dominated by finite T and effects of  $\eta$  the elastic scattering parameter, we can also perform an

expansion in the limit when  $|Q| \ll \Gamma_0$ , using Eq. (SI-23) and Eq. (SI-24). The the first few terms are

$$E(k) = \frac{\Delta_0}{1-r} - \sqrt{\frac{r}{2-r}} \Gamma_0 + \frac{V_L}{2-r} (\hat{k} - \hat{k}_{kink}) - \frac{(1-r)^2}{2\sqrt{r(2-r)^3}} \frac{V_L^2}{\Gamma_0} (\hat{k} - \hat{k}_{kink})^2 + \dots \quad (\text{SI-37})$$

Similarly for the EDC dispersion

$$E^*(k) = \frac{\Delta_0}{1-r} - \Gamma_0 + rV_L (\hat{k} - \hat{k}_{kink}) - \frac{(1-r)^2}{2} \frac{V_L^2}{\Gamma_0} (\hat{k} - \hat{k}_{kink})^2 + \dots \quad (\text{SI-38})$$

These formulas display a shift in the energies due to  $\Gamma_0$  and also a  $\Gamma_0$  dependent curvature. Since the regime of this expansion,  $|Q| < \Gamma_0$  is different from that of the expansion in Eq. (SI-35) and Eq. (SI-33), we note that velocities are different as well. Thus one must be careful about specifying the regime for using the velocity formulae.

Let us note that in this regime  $|Q| < \Gamma_0$  the two dispersions differ, with the EDC higher.

$$E^*(k) - E(k) = \left\{ \sqrt{\frac{r}{2-r}} - 1 \right\} \Gamma_0 - \frac{(1-r)^2}{2-r} V_L (\hat{k} - \hat{k}_{kink}) + \dots \quad (\text{SI-39})$$

This equation gives a prescription for estimating  $\Gamma_0$  in cases where the other parameters are known. Alternatively in the MDC dispersion we expect to see a curvature only near the location of the kink, this is sufficient to fix  $\Gamma_0$ : from Eq. (SI-37)

$$\frac{d^2 E(k)}{d\hat{k}^2} = -\frac{(r-1)^2}{\sqrt{r(2-r)^3}} \frac{V_L^2}{\Gamma_0}. \quad (\text{SI-40})$$

The curvature  $\frac{d^2 E(k)}{d\hat{k}^2}$  can be estimated from the experimental data to provide an estimate of  $\Gamma_0$ .

### C. The Dyson self energy

For completeness we present the low energy expansion of the Dyson self energy, which gives rise to the spectral function in Eq. (SI-21). We may define the Dyson self energy from

$$\Sigma_D = \omega + \mu - \varepsilon_k - \mathcal{G}^{-1} \quad (\text{SI-41})$$

Using Eq. (SI-15) we obtain

$$\Im m \Sigma_D = -\frac{1}{Z} \frac{\mathcal{R}}{\Omega_\Phi} \frac{1 - \frac{1}{\Delta_0} (\omega - \nu_0 \hat{k} v_f)}{\{1 + (\omega + \nu_\Psi \hat{k} v_f)/\Omega_\Psi\}^2 + \frac{\mathcal{R}^2}{s^2 \Omega_\Phi^2 \Omega_\Psi^2}} \quad (\text{SI-42})$$

The corresponding real part is given by

$$\begin{aligned} \Re \Sigma_D &= \boldsymbol{\mu} - \boldsymbol{\mu}_0 + \omega - \hat{k} v_f \\ &- \frac{1}{Z} \frac{(\omega - \nu_\Phi \hat{k} v_f) + \frac{1}{\Omega_\Psi} q_2}{\{1 + (\omega + \nu_\Psi \hat{k} v_f)/\Omega_\Psi\}^2 + \frac{\mathcal{R}^2}{s^2 \Omega_\Phi^2 \Omega_\Psi^2}} \\ q_2 &= (\omega + \nu_\Psi \hat{k} v_f)(\omega - \nu_\Phi \hat{k} v_f) + \frac{\mathcal{R}^2}{s \Omega_\Phi^2}. \end{aligned} \quad (\text{SI-43})$$

The  $q_2$  term is quadratic (or higher) in the small variables  $\omega$ ,  $\hat{k} v_f$ , however these small terms are needed if we want to reproduce exactly Eq. (SI-16).

### 1. Useful identities and some Fermi Liquid parameters.

We list a few useful identities relating the various parameters

$$\begin{aligned} \Omega_\Psi &= \frac{1-s}{s} \Delta_0, \\ s &= \frac{\Delta_0}{\Delta_0 + \Omega_\Psi} \\ \nu_0 &= \frac{\nu_\Phi + s \nu_\Psi}{1-s} = r \nu_\Phi \\ \nu_\Psi &= \frac{r-1-rs}{s} \nu_\Phi \\ r-1 &= \frac{\Delta_0}{\Omega_\Psi} \left(1 + \frac{\nu_\Psi}{\nu_\Phi}\right) \end{aligned} \quad (\text{SI-44})$$

Let us note the Fermi liquid renormalizations from Eq. (SI-41)

$$\begin{aligned} \left. \frac{d\Sigma_D}{d\hat{k}} \right|_{FS} &= \left( \frac{V_L}{Z} - v_f \right) \\ \left. \frac{d\Sigma_D}{d\omega} \right|_{FS} &= \left( 1 - \frac{1}{Z} \right) \end{aligned} \quad (\text{SI-45})$$

Therefore we write the Fermi liquid mass enhancement that determines the heat capacity as:

$$\frac{m}{m^*} = Z \left\{ 1 + \frac{1}{v_f} \left. \frac{d\Sigma_D}{d\hat{k}} \right|_{FS} \right\} = V_L/v_f = \nu_\Phi. \quad (\text{SI-46})$$

Thus  $\nu_\Phi$  is the inverse mass enhancement factor, obtainable from the ratio of the heat capacity and the bare density of states. In this model we note that  $\nu_\Phi$  is not obliged to vanish as  $Z$  near the half filled limit  $n \rightarrow 1$ , but may be a finite number of  $O(1)$ . This is unlike the Brinkman Rice ‘‘heavy metal’’ type behavior  $m/m^* \propto Z$ , which is prototypical of theories with a momentum independent self energy.

Finally we note that the condition for the kink to occur is, we recall,  $r > 1$ . From Eq. (SI-44) we see that this requires a finite  $\Omega_\Psi$  (so that  $1 > s > 0$ ). We also need  $\Delta_0 > 0$  and  $\left(1 + \frac{\nu_\Psi}{\nu_\Phi}\right) > 0$ .

- 
- [SI-1] T. Cuk, D. H. Lu, X. J. Zhou, Z.-X. Shen, T. P. Devereaux, and N. Nagaosa, *Physica Status Solidi (b)* **242** (2005).
- [SI-2] J. He, W. Zhang, J. M. Bok, D. Mou, L. Zhao, Y. Peng, S. He, G. Liu, X. Dong, J. Zhang, J. S. Wen, Z. J. Xu, G. D. Gu, X. Wang, Q. Peng, Z. Wang, S. Zhang, F. Yang, C. Chen, Z. Xu, H.-Y. Choi, C. M. Varma, and X. J. Zhou, *Phys. Rev. Lett.* **111**, 107005 (2013).
- [SI-3] S. Vig, A. Kogar, V. Mishra, L. Venema, M. S. Rak, A. A. Husain, P. D. Johnson, G. D. Gu, E. Fradkin, M. R. Norman, and P. Abbamonte, arXiv preprint arXiv:1509.04230 (2015). (unpublished)
- [SI-4] A.B. Migdal, *Sov. Phys. JETP* **7**, 996 (1958).
- [SI-5] S. Engelsberg and J. R. Schrieffer, *Physical Review* **131**, 993 (1963).
- [SI-6] O. Gunnarsson, V. Meden, and K. Schnhammer, *Phys. Rev. B* **50**, 10462 (1994).
- [SI-7] A. V. Chubukov and M. R. Norman, *Phys. Rev. B* **70**, 174505 (2004).
- [SI-8] A. Kaminski, M. Randeria, J. C. Campuzano, M. R. Norman, H. Fretwell, J. Mesot, T. Sato, T. Takahashi, and K. Kadowaki, *Phys. Rev. Lett.* **86**, 1070 (2001).
- [SI-9] M. R. Norman, *Phys. Rev. B* **75**, 184514 (2007).
- [SI-10] B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011).
- [SI-11] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013).
- [SI-12] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *Phys. Rev. B* **88**, 205108 (2013).
- [SI-13] D. Hansen and B. S. Shastry, *Phys. Rev. B* **87**, 245101 (2013).
- [SI-14] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).
- [SI-15] B. S. Shastry, *Ann. Phys.* **343**, 164 (2014); (Erratum), *Ann. Phys.* **373**, 717 (2016).
- [SI-16] G-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).
- [SI-17] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016).

## Strange metal from Gutzwiller correlations in infinite dimensions

Wenxin Ding,<sup>1</sup> Rok Žitko,<sup>2,3</sup> Peizhi Mai,<sup>1</sup> Edward Perepelitsky,<sup>1</sup> and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, California 95060, USA*

<sup>2</sup>*Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

<sup>3</sup>*Faculty for Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia*

(Received 7 March 2017; published 18 August 2017)

Recent progress in *extremely correlated Fermi liquid theory (ECFL)* and the *dynamical mean field theory (DMFT)* enables us to accurately compute in the  $d \rightarrow \infty$  limit the resistivity of the  $t - J$  model after setting  $J \rightarrow 0$ . This is also the  $U = \infty$  Hubbard model. Since  $J$  is set to zero, our study isolates the *dynamical effects* of the single occupation constraint enforced by the projection operator originally introduced by Gutzwiller. We study three densities  $n = .75, .8, .85$  that correspond to a range between the overdoped and optimally doped Mott insulating state. We delineate four distinct regimes separated by three crossovers, which are characterized by different behaviors of the resistivity  $\rho$ . We find at the lowest temperature  $T$  a *Gutzwiller correlated Fermi liquid* regime with  $\rho \propto T^2$  extending up to an effective Fermi temperature that is dramatically suppressed from the noninteracting value by the proximity to half filling,  $n \sim 1$ . This is followed by a *Gutzwiller correlated strange metal* regime with  $\rho \propto (T - T_0)$ , i.e., a linear resistivity extrapolating back to  $\rho = 0$  at a positive  $T_0$ . At a higher temperature scale this crosses over into the *bad metal* regime with  $\rho \propto (T + T_1)$ , i.e., a linear resistivity extrapolating back to a finite resistivity at  $T = 0$  and passing through the Ioffe-Regel-Mott value where the mean free path is a few lattice constants. This regime finally gives way to the *high T metal* regime, where we find  $\rho \propto T$ , i.e., a linear resistivity extrapolating back to zero at  $T = 0$ . The present work emphasizes the first two, i.e., the two lowest temperature regimes, where the availability of an analytical ECFL theory is of help in identifying the changes in related variables entering the resistivity formula that accompanies the onset of linear resistivity, and the numerically exact DMFT helps to validate the results. We also examine thermodynamical variables such as the magnetic susceptibility, compressibility, heat capacity, and entropy and correlate changes in these with the change in resistivity. This exercise casts valuable light on the nature of charge and spin correlations in the Gutzwiller correlated strange metal regime, which has features in common with the physically relevant strange metal phase seen in strongly correlated matter.

DOI: [10.1103/PhysRevB.96.054114](https://doi.org/10.1103/PhysRevB.96.054114)

### I. INTRODUCTION

The resistivity due to mutual collisions of electrons at low temperatures reveals the lowest energy scale physics of charge excitations in metallic systems and therefore is very important. While it is fairly straightforward to measure experimentally, it is also one of the most difficult quantities to calculate theoretically, especially if electron-electron interactions are strong. Motivated by the unexpected behavior of resistivity and other variables in cuprate superconductors and related two-dimensional experimental systems, some works have postulated that the Fermi liquid theory—originally developed and justified for weakly interacting systems—would break down. In its place a zoo of non-Fermi liquids have been postulated, without necessarily having a rigorous theoretical underpinning. On the other hand the analytical framework of the extremely correlated Fermi liquid theory (ECFL) [1] and the well established dynamical mean field theory (DMFT) [2] give a different type of result, where the strong interactions compress the regime of Fermi-liquid type variation to a very small temperature and frequency scale. This Fermi-liquid regime is succeeded by a variety of regimes that display unusual non-Fermi-liquid dependences on frequency and temperature. The main goal of this work is to elucidate and characterize the different regimes that arise in the ECFL and DMFT theories and to provide a quantitative comparison between the qualitatively similar results of these two theories,

as applied to the infinite-dimensional Hubbard model, with the Hubbard charge repulsion parameter  $U$  taken to infinity,  $U \rightarrow \infty$ .

In earlier work [3] we have compared the ECFL and DMFT results for the zero-temperature spectral functions, finding an encouraging similarity. On scaling the frequency with the respective quasiparticle weights  $Z$  of the two theories the agreement is even close to quantitative. In the present work we undertake the more ambitious comparison of the resistivity and thermodynamic variables at finite temperatures.

In both the ECFL theory and the DMFT, the strong interactions cause the quasiparticles of the lowest temperature Fermi liquid to become fragile, i.e., the resulting quasiparticle weight  $Z$  is very small,  $Z \ll 1$ . This is also arguably the relevant regime in contemporary materials such as cuprate superconductors, and hence interest in this problem is very high.

In the problem studied here, namely  $U \rightarrow \infty$  and  $d \rightarrow \infty$ , the DMFT theory is formally exact. Further, the possibility of computing the resistivity from the sole knowledge of the single-particle Green's function is enabled by the vanishing of vertex corrections [4]. Despite these simplifications, obtaining *reliable* results for the resistivity is technically formidable due to the requirement of an impurity solver providing accurate and reliable results for the self-energy function  $\Sigma$  on the real frequency axis for both very low and very high temperatures. This problem has only recently been solved in Ref. [5],

almost 25 years after the formulation of the DMFT theory. The resistivity of the Hubbard model is now known for all densities and all values of  $U$ , including  $U = \infty$ . This is a set of exact results for the resistivity in interacting metallic systems resulting from inelastic scattering and therefore represents an important advance in the field. The DMFT results [5,6] offer a unique opportunity to test a variety of techniques and approximate methods for computing this variable. The ECFL formalism, on the other hand, is in its early stages of development and several technical innovations are ongoing so as to enable reliable calculations in the challenging regimes of the density  $n \lesssim 1$  [3,7].

Lastly, in a recent work [8] our group has published a voluminous high-temperature study using series expansion techniques adapted for very strong correlations, thus extending our understanding of the resistivity to the full range of temperatures. This study is on the same model as the present work and extends the results of Ref. [5] to much higher temperatures. In these studies the effect of the superexchange  $J$  is absent due to the  $U = \infty$  limit, and therefore there is no superconducting regime that one might expect from a  $t - J$  model in finite dimensions. By taking the limit of infinite  $U$  we have also banished the static superexchange that the DMFT includes for finite  $U$  [9–17]. However, these studies *do* capture the notoriously difficult nonperturbative local Gutzwiller correlation effects on the resistivity quantitatively. It seems fair to say that our understanding of the strong correlation problem has advanced significantly with these recent works.

In summary, at the lowest temperatures these earlier studies [5–8] found a Fermi-liquid type resistivity with  $\rho \propto T^2$ . This regime extends only up to  $T_{FL}(\delta)$ , a Fermi-liquid temperature scale dependent on the hole density ( $\delta \equiv 1 - n$ ). We shall term this the *Gutzwiller correlated Fermi liquid* (GCFL) regime. This regime is followed by three distinguishable regimes with linear in  $T$  resistivity having different slopes and intercepts, which are separated by crossovers; a *Gutzwiller correlated strange metal* (GCSM) followed by a “bad metal” and finally a “high-T metal” regime, as discussed below (see Fig. 1). The nomenclature stresses that these regimes originate purely from Gutzwiller correlations (i.e., double occupancy avoidance). In particular the regimes have no dependence upon the superexchange energy  $J$  or other energy scales which might be additionally involved in producing the related strange metal found in cuprates [18,19].

In order to understand the low-temperature regimes, we would like to throw light on the factors that lead to extraordinarily low values of the Fermi temperature  $T_{FL}(\delta)$  that are found. We also wish to provide a detailed understanding of the behavior of constituent variables that lead to a linear resistivity in the GCSM regime, starting at this low temperature. Here the ECFL theory provides us with a great advantage since it is largely analytical, and one can inspect the various constituents in detail. It is also interesting to seek a possible causal relationship between the linear temperature dependence of  $\rho$  in the GCSM regime and the nature of incipient order (either spin or charge) that might be present. For this purpose, it is useful to compute, by using the techniques of Refs. [5,7], the entropy and heat capacity, the magnetic susceptibilities and compressibility. For completeness we also study the thermoelectric transport, as well as a few dynamical quantities

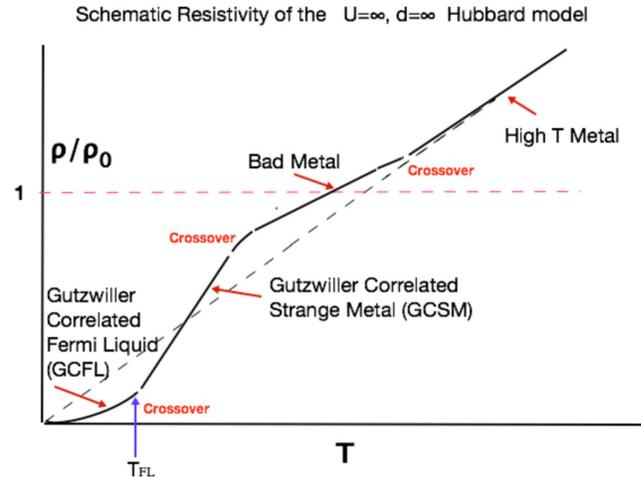


FIG. 1. A schematic view of the different regimes of temperature dependent resistivities found in the calculations of Refs. [5–8]. The various temperature scales are schematic. At the lowest  $T$  we have a Gutzwiller-correlated-Fermi liquid regime (GCFL) with  $\rho \propto T^2$ . This quadratic variation terminates at a characteristic Fermi temperature  $T_{FL}(\delta)$ , which is found to be surprisingly small relative to  $T_{BR} = \delta D$ , the Brinkman-Rice temperature scale ( $2D$  is the bandwidth). Upon warming we reach the Gutzwiller-correlated-strange metal (GCSM) regime, which is the main focus of this work. This gives way at higher  $T$  to the so-called bad-metal regime with a resistivity that increases linearly beyond the Ioffe-Regel-Mott value  $\rho_0$  characteristic of disordered metals. The temperature scale of this regime is  $T_{BR}$  discussed above. Finally at the highest  $T$  we reach the high  $T$  regime with  $\rho \propto T$  that can be extrapolated back to pass through the origin. We thus find a total of four regimes separated by three crossovers. It should be noted that in both theories considered here, the approximate range of the temperatures scales are  $T_{FL} \sim 0.004\text{--}0.01D$ , and the crossover to the bad-metal regime occurs at  $T \sim 0.04\text{--}0.06D$  for the densities considered ( $n = 0.75$  to  $n = 0.85$ ).

such as the self energy of the electrons. In a following paper we present other dynamical variables such as the optical conductivity. These quantities provide a complete picture of the metallic states having various temperature dependences sketched in Fig. 1.

The lowest temperature *Gutzwiller-correlated Fermi liquid* (GCFL) with  $\rho \propto T^2$  shows enhancements of certain static susceptibilities that are similar to those of the normal state of liquid  $^3\text{He}$ . The *almost localized Fermi liquid theory* (ALFL) of these enhancements is discussed by Vollhardt, Wölfle, and Anderson in Refs. [24,25] on the basis of Gutzwiller’s wave function and its approximation to the Hubbard model, where the variation of the Landau parameters with density at fixed (large)  $U$  is considered. In particular Ref. [25] studies the enhancements of Fermi liquid parameters leading to enhanced effective mass  $m^*/m$ , magnetic susceptibility  $\chi_{\text{spin}}/\chi_{\text{spin}}^0$ , and the bulk modulus (i.e., the inverse compressibility). Within the ALFL all three stated enhancements are proportional to the inverse of  $Z$  in that theory as well as in  $^3\text{He}$ . We check below the extent to which this is true in the GCFL regime, to see how it compares with the predictions of the ALFL theory, and find that the behavior of the compressibility is somewhat different.

Upon warming we reach the GCSM regime with a linear temperature dependence of the resistivity  $\rho$ . This regime is interesting since it is reminiscent of the *strange metal* regime in the cuprate phase diagrams [18]. It is remarkable that this linear resistivity regime extends to very low  $T$ , essentially the  $T_{\text{FL}}(\delta)$ , and one wants to know if this behavior is causally linked to a change in entropy, i.e., to disordering. We aim at correlating the GCSM regime with the extent of short ranged spin or charge order in this regime. These should be reflected in the heat capacity and the entropy gain. By computing these variables, we show that upon warming from  $T = 0$  substantial entropy is released as we reach  $T_{\text{FL}}$ . However in the entire GCSM regime the magnetic susceptibility is Pauli like, i.e., with an approximately  $T$  independent behavior, and hence spin entropy should be unchanged. From a high- $T$  expansion and on various general grounds, it is known that it changes into a Curie-Weiss type behavior at the onset of the bad-metal regime.

The GCSM regime is followed by other subtly different  $T$  dependences as described in Sec. III A which are obtained in the bad-metal regime and the high-temperature regime. The density dependences of the various crossover scales give important insight into the physics of the resistivity. With one exception, all calculations reported here are performed using both ECFL and single-site DMFT methods. Using the two methods is very important since it gives us the opportunity to benchmark the mostly analytical and relatively new ECFL technique with the established and largely numerical DMFT method. The magnetic susceptibility is available only from DMFT, and our presentation below seems to be the most extensive result for this subtle variable reported to date [11,26,27].

The plan of the paper is as follows. In Sec. II we first make some further technical remarks about the methods. In Sec. III A we describe the various  $T$  dependences of the resistivity which serve to define the GCFL and GCSM regimes and also point to the higher  $T$  bad-metal and high- $T$  regimes. In Sec. III B we compare the chemical potential and compressibility. In Sec. III C we discuss the frequently made bubble approximation for the charge and spin susceptibilities and show that the bubble susceptibility is exactly expressible as an integral of the energy derivative of momentum distribution function in  $d = \infty$ . We also note that it is a good approximation to the exact result for the charge susceptibility but not so for the spin susceptibility. In Sec. III D we illustrate the self energy and local density of states from the two methods and find that within ECFL the quasiparticles tend to have somewhat smaller  $Z$  at the highest densities, as compared to DMFT. This causes a few other differences described later. In Sec. III E we examine further  $T$  dependent properties, the heat capacity and entropy. Section III F discusses the magnetic susceptibility  $\chi$  from the DMFT calculations and lists some of the technical difficulties that prevent its evaluation in the ECFL theory. In Sec. III G we discuss the thermoelectric transport coefficients, the Seebeck coefficient, and the Lorenz number as well as the thermoelectric efficiency. In Sec. IV we discuss the salient features of our results.

## II. METHODS

In ECFL we have thus far used an expansion in the parameter  $\lambda$ , which plays a role analogous to the quantum

parameter  $\frac{1}{2S}$  in quantum theories of magnetism, where  $S$  is magnitude of the spin. In the first DMFT-ECFL comparison paper [3], we used the second order terms in an expansion in  $\lambda$ . This approximation led to a quantitatively reliable answer for the quasiparticle weight  $Z$  at low temperature only in the overdoped regime  $n \lesssim .75$  but to a nonvanishing value of  $Z$  for  $n \rightarrow 1$ . In the more recent paper [7] this problem was addressed using the exact, rather than the  $\lambda^2$ , version of the hole number sum rule, together with a cutoff for the tails of the spectral function at very high energies. This procedure extends the validity of the second order terms to higher density  $n \lesssim 0.85$  so that the  $Z$  values at low  $T$  tend to zero as the insulating state is approached and are comparable to, if somewhat smaller than, the DMFT results. Due to this improvement, we found that the resistivity is now on the same scale and exhibits very similar crossover features as the results in Refs. [5,6], as detailed below. In this work we report the comparison between the  $T$  dependent resistivity and other thermodynamic variables found from this cutoff scheme [28] and the exact results from DMFT. We use the Bethe lattice semicircular density of states  $D(\epsilon) = \frac{2}{\pi D} \sqrt{1 - \epsilon^2/D^2}$  in both theories.

The ECFL scheme used here has been described in detail in Ref. [7] and consists of using the  $\mathcal{O}(\lambda^2)$  expansion with the full number sum rule and the Tukey window used to cut off the spectral width at very high energies. The DMFT scheme has been described in detail in Ref. [3]. The NRG calculations [29,30] in this work were performed with the discretization parameter  $\Lambda = 2$ , using the discretization scheme from Refs. [31,32] with  $N_z = 16$  interleaved discretization grids. The truncation cutoff was set at  $10\omega_N$ , where  $\omega_N$  is the characteristic energy scale at the  $N$ th step of the iteration. We used charge conservation and spin SU(2) symmetries. The spectral functions were computed with the full-density-matrix algorithm [33] and broadened with a log-Gaussian kernel with  $\alpha = 0.05$ , followed by a Gaussian kernel with  $\sigma = 0.3T$ . The occupancy was controlled using the Broyden method [34]. The self-energy was computed through the ratio of correlators,  $\langle\langle n_{\bar{\sigma}} d_{\sigma}; d_{\sigma}^{\dagger} \rangle\rangle / \langle\langle d_{\sigma}; d_{\sigma}^{\dagger} \rangle\rangle$  [35], corrected by the term  $-w_{\text{UHB}} / \langle\langle d_{\sigma}; d_{\sigma}^{\dagger} \rangle\rangle$ , where  $w_{\text{UHB}}$  is the spectral weight of the upper Hubbard peak which was outside the NRG energy window (we redid some calculations using the standard approach that explicitly includes the UHB in the energy window, using a very large but finite value of  $U$ ; we found excellent agreement between the two computational schemes).

## III. RESULTS

In this work we consider the temperature region  $T \leq 0.02D$ , which covers the range up to 200 K if we assume  $D \sim 10000$  K, i.e.  $\mathcal{O}(1)$  eV. Here  $D$  is the half bandwidth. We study three densities (number of electrons per site)  $n = 0.75, 0.8, 0.85$ . These are typical of the overdoped and optimally doped cuprates.

### A. DC resistivity

We begin with a summary of the results for the resistivity which form the bedrock for this study. The findings in Refs. [5–7] are extended in Ref. [8] to higher temperatures, and from these we have a fairly complete understanding of

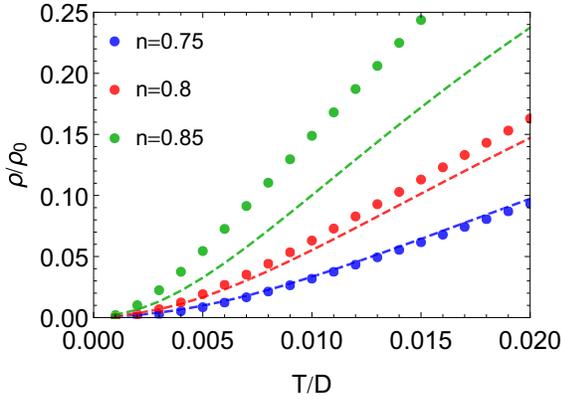


FIG. 2. Comparison of the resistivity computed using the ECFL (symbols) and the DMFT (dashed).  $\sigma_0 = 1/\rho_0$  is the Ioffe-Regel-Mott conductivity. As  $n$  gets closer to unity, the ECFL scheme employed systematically underestimates  $Z$  relative to the exact DMFT values (see Fig. 6 of Ref. [7]). This lowers the effective Fermi temperature  $T_{\text{FL}}$  and simultaneously enhances the magnitude of  $\rho$  for  $T > T_{\text{FL}}$ , a feature that is prominently visible above. It should be possible to improve the quantitative agreement between the two theories in the future [28].

the behavior of  $\rho$  at *essentially all*  $T$ . A cartoon of these is sketched in Fig. 1. The resistivity exhibits a variety of dependences on  $T$  upon warming from the absolute zero: (i) the Gutzwiller correlated Fermi liquid (GCFL) regime with a quadratic  $T$  dependence  $\rho \propto T^2$  up to a (hole) density-dependent Fermi-liquid temperature  $T_{\text{FL}}(\delta)$  ( $\delta = 1 - n$ ); (ii) the Gutzwiller-correlated-strange metal (GCSM) regime with a linear  $T$  dependence  $\rho \propto T + \text{constant}$  (constant  $< 0$ ), (iii) a “knee” connecting to the bad-metal (BM) regime with again a linear  $T$  dependence  $\rho \propto T + \text{constant}$  (constant  $> 0$ ). This regime is so named since the  $\rho$  crosses the fiduciary Ioffe-Regel-Mott maximal resistance  $\rho_0$  at temperature on the order of the Brinkman-Rice energy  $\delta D$ , followed by (iv) a crossover to a high-temperature regime again with linear  $T$  dependence  $\rho = AT$ , devoid of an offset so that the line extrapolates back to pass through the origin.

In Fig. 2 we present the resistivity in the GCFL and GCSM regimes. It is striking that the GC strange metal has a robust linear  $T$  resistivity over a wide  $T$  scale. The linear resistivity begins at  $T_{\text{FL}}(\delta)$  which can be driven to low values,  $\sim 45$  K (see Ref. [7]), by the Gutzwiller correlations alone, even though the bandwidth is of  $\mathcal{O}(2)$  eV. We emphasize that this unexpectedly drastic scale reduction yielding  $T_{\text{FL}} \ll ZD \ll \delta D$  requires a “hard” calculation for justification and can hardly be argued from general principles. The slight difference in the  $T_{\text{FL}}(\delta)$  between the two theories is due to the somewhat different  $Z(\delta)$  found in the two theories, for example Fig. 6 in Ref. [7] shows that the ECFL gives a smaller  $Z$  than the DMFT [28]. We also note that using the standard value for  $\rho_0 \sim 300 \mu\Omega$  cm, the Ioffe-Regel-Mott resistivity [36], the absolute scale of the resistivity computed in these approaches is quite similar to that found in the experiments. For example, Fig. 1 in Ref. [7] compares well on an absolute scale with the well-known linear resistivity result of S. Martin *et al.* in Ref. [37] on Bi2212, where the superconducting phase cuts off the region  $T \leq 80$  K.

Building on the analysis of Refs. [5–7], we derive a closed form expression for the resistivity in terms of the chemical potential and the real and imaginary parts of the single-particle self-energy on the Fermi surface [Eq. (7)]. We begin with the formula (Eq. (41) in Ref. [7]) for the conductivity on the infinite-dimensional Bethe lattice:

$$\sigma = 2\pi D \sigma_0 \int d\omega \int d\epsilon \left( -\frac{\partial f}{\partial \omega} \right) \phi(\epsilon) \rho_G^2(\epsilon, \omega), \quad (1)$$

where  $\sigma_0 = e^2 \hbar \Phi(0)/D$  ( $\Phi$  is defined in Eq. (39) of Ref. [7]),  $\sigma_0 = 1/\rho_0$ , and the transport function  $\phi(\epsilon) = \Phi(\epsilon)/\Phi(0)$  is given explicitly in Eq. (40) of Ref. [7] as  $\phi(\epsilon) = \Theta(1 - \frac{\epsilon^2}{D^2}) \times (1 - \frac{\epsilon^2}{D^2})^{\frac{3}{2}}$ . The single-particle spectral function is

$$\rho_G(\epsilon, \omega) = \frac{1}{\pi} \frac{B(\omega)}{[A(\omega) - \epsilon]^2 + B^2(\omega)}, \quad (2)$$

where  $A(\omega) \equiv \omega + \mu - \Re e \Sigma(\omega)$ ,  $B(\omega) = -\Im m \Sigma(\omega)$ , and all objects depend implicitly on the temperature  $T$ . At low temperatures and frequencies  $B(\omega) \ll D$ , so that Eq. (1) simplifies to

$$\sigma = \sigma_0 \int d\omega \left( -\frac{\partial f}{\partial \omega} \right) \frac{\phi[A(\omega)]}{B(\omega)}. \quad (3)$$

Following [6], we perform a small-frequency expansion

$$\phi[A(\omega)] = \phi[A(0)] + \dots; \quad B(\omega) = B_0 + B_2 \omega^2 + \dots \quad (4)$$

The linear order term in  $B(\omega)$  as well as all higher order terms in  $B(\omega)$  and  $\phi[A(\omega)]$  make negligible contributions to the conductivity in the temperature range considered and are therefore neglected. The integral may be evaluated analytically and yields

$$\sigma = \frac{\sigma_0 \phi[A(0)]}{2\pi T \sqrt{B_2 B_0}} \psi_1 \left( \frac{1}{2} + \frac{1}{2\pi T} \sqrt{\frac{B_0}{B_2}} \right), \quad (5)$$

where  $\psi_1(z)$  is the polygamma function, related to the digamma function,  $\Psi(z)$ , through  $\psi_1(z) \equiv \frac{d}{dz} \Psi(z)$  [38]. The ratio  $\frac{B_0}{B_2 \pi^2 T^2}$  is weakly dependent on temperature and may be replaced by its zero-temperature limit, see Fig. 3(b). In order to find this limiting value, consider the GCFL regime where

$$B_0 = B_2 \pi^2 T^2 \quad (\text{GCFL}). \quad (6)$$

Substituting Eq. (6) into Eq. (5) and eliminating  $B_2$ , we finally obtain the simple formula

$$\rho = \frac{12 \rho_0}{\pi^2 \phi[\bar{\mu} - \Re e \bar{\Sigma}(0)]} \times B_0, \quad (7)$$

where we have used that  $\psi_1(1) = \frac{\pi^2}{6}$ . Here, we denote the zero-temperature limit of any variable  $Q$  as  $\bar{Q}$  and have used that  $\phi[A(0)]$  is practically temperature independent [Fig. 3(c)]. Hence, the resistivity is proportional to the imaginary part of the self-energy on the Fermi surface. Moreover, the proportionality constant is very weakly density dependent (since this is true of  $\phi[\bar{A}(0)]$ ). Equation (7) can be obtained from Eq. (47) in Ref. [7] by multiplying the RHS of the latter by the constant  $\frac{12}{\pi^2}$  and setting  $T \rightarrow 0$  in the denominator. The latter equation is obtained by retaining the leading order

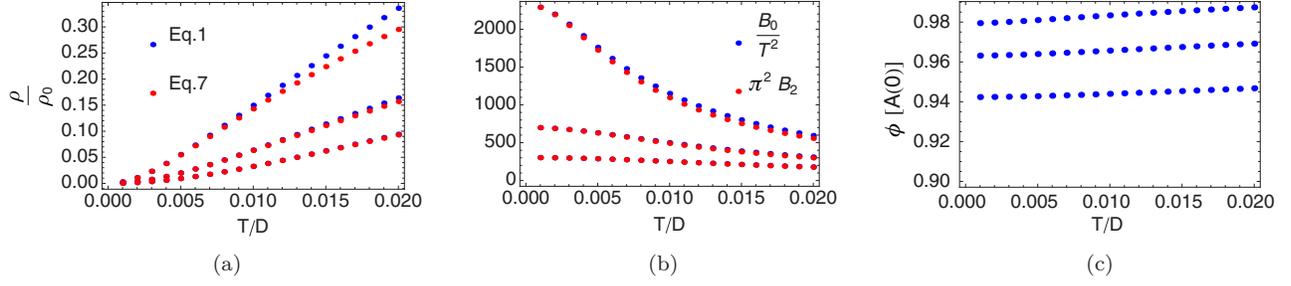


FIG. 3. ECFL calculation of the resistivity and related objects. Panel (a): The resistivity as a function of the temperature using the exact formula, Eq. (1), compared with the approximation, Eq. (7), for  $n = 0.75, 0.8, 0.85$  (bottom to top). Equation (7) is an excellent approximation at all densities for all temperatures. Panel (b): Parameters resulting from a low-frequency expansion of the imaginary part of the self-energy in the vicinity of the Fermi surface, plotted as a function of temperature, for  $n = 0.75, 0.8, 0.85$  (bottom to top).  $B_0$  is the self-energy on the Fermi surface, while  $B_2$  is the quadratic-frequency term. The ratio  $\frac{B_2 \pi^2 T^2}{B_0} \rightarrow 1$  as  $T \rightarrow 0$  and is approximately constant as a function of temperature. Panel (c):  $\phi[A(0)] = \phi[\mu - \Re \Sigma(0)]$ , plotted as a function of the temperature, for  $n = 0.75, 0.8, 0.85$  (bottom to top).  $\phi[A(0)]$  is practically independent of temperature and has very weak density dependence.

term in the Sommerfeld expansion of Eq. (3). In Fig. 3(a), we plot the resistivity as a function of the temperature, using both Eqs. (1) and (7), in the ECFL scheme. We find that Eq. (7) is an excellent approximation at all densities and temperatures considered, i.e., it holds in both the GCFL and GCSM regimes.

In the GCFL regime, substituting Eq. (6) into Eq. (7), and using the fact that  $B_2$  is approximately constant, we find that

$$\rho = \frac{12 \bar{B}_2 \rho_0}{\phi[A(0)]} \times T^2 \quad (GCFL). \quad (8)$$

From Fig. 7 of Ref. [7], we know that  $\bar{B}_2 \propto \frac{1}{Z^2}$ , where  $Z$  is the quasiparticle weight on the Fermi surface. Therefore, Eq. (8) implies that  $\rho \propto \frac{T^2}{Z^2}$  in the GCFL regime.

In Fig. 4, we plot the exact resistivity, together with the approximation Eq. (7), both obtained using the DMFT calculation [corresponding to Fig. 3(a) in the case of ECFL]. Once again, we find that Eq. (7) is an excellent approximation at all densities and temperatures considered, i.e., it holds in both the GCFL and GCSM regimes.

Finally, we note that the important effective Fermi temperature  $T_{FL}$  can be estimated as the temperature at which the resistivity deviates from its low-temperature quadratic behavior. We find at the three densities considered, the

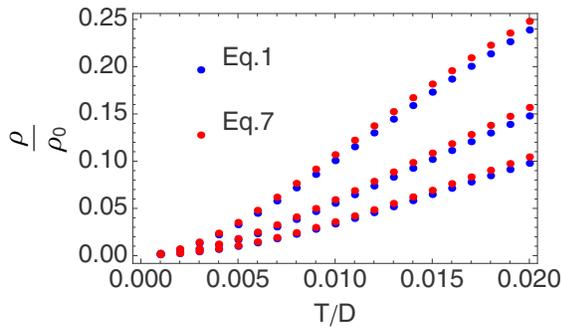


FIG. 4. The exact resistivity [Eq. (1)] compared with the approximation Eq. (7), using the DMFT calculation for  $n = 0.75, 0.8, 0.85$  (bottom to top). Equation (7) is an excellent approximation at all densities for all temperatures. [See Fig. 3(a) for the corresponding figure in ECFL.]

so-determined effective Fermi temperature for ECFL is, in agreement with Ref. [7], given by  $T_{FL} \sim .05 \bar{Z} D$ . In the case of DMFT, we also find  $T_{FL} \sim .05 \bar{Z} D$ , where a slightly higher value of  $\bar{Z}$  results in a slightly higher value of  $T_{FL}$ , as compared to ECFL.

## B. Chemical potential and compressibility

The chemical potential in the ECFL theory is found from the self-consistency condition of the Green's function. The compressibility  $\kappa = n^{-2} \partial n / \partial \mu$  is determined by numerical differentiation. The derivative is computed using the finite difference formula  $\partial n / \partial \mu = [(n + \delta n) - n] / [\mu(n + \delta n) - \mu(n)]$  with  $\delta n = 0.001$ . In the DMFT we used larger  $\delta n = 0.01$  and we performed two full DMFT runs for fillings  $n$  and  $n + \delta n$ .

We see that the chemical potentials (Fig. 5) match well apart from a constant shift [39]. The results obtained using

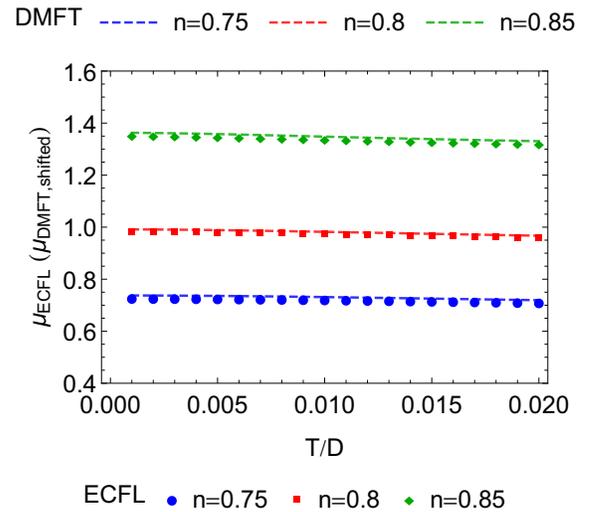


FIG. 5. Chemical potentials at  $n = 0.75, 0.8, 0.85$  for ECFL (symbols) and DMFT (dashed lines). The DMFT results are shifted by a density-dependent constant. After the shift, the chemical potentials almost coincide.

two different impurity solvers (NRG and CT-HYB QMC) in the DMFT are in agreement, thus the difference is not related to some technical issue in the NRG but is an actual discrepancy between DMFT and ECFL.

In our earlier work on the single impurity Anderson model [40], using a scheme that is an adaptation of that in Ref. [3], we studied the single impurity energy, which is a close analog of the chemical potential in the present problem. There we found that the location of the impurity energy found from the second order ECFL equations matched very closely the impurity energy found in the NRG (see Table 1 in Ref. [40]). In view of that excellent agreement, the current discrepancy on the absolute scale of the chemical potential between the DMFT results (also from NRG) and the present second order scheme is somewhat unexpected. It would appear that the different hole number sum rule and the cutoff scheme used here relative to the scheme in Refs. [3,40] influences this variable and needs to be investigated more closely in the future.

We note that the compressibilities (Fig. 6) are also roughly similar, and both theories show a suppression relative to the free fermion theory. The free fermion theory shows a slight monotonic decrease of the compressibility with  $T$ . In the GCFL and GCSM regimes, the ECFL compressibility shows an increase with  $T$ , followed by a slight fall with  $T$  in the bad metal regime. In Fig. 6(b), we show that in the ECFL theory  $Z/\kappa$  is a constant within numerical errors ( $\sim \pm 3.4\%$ ) at  $T = 0.001D$ . This is not the case in the DMFT, where  $Z$  is proportional to  $\delta$ , while  $\kappa$  behaves approximately as  $\kappa \propto \delta^{0.2}$  close to the doping-driven Mott transition [3]. In the GCFL regime, if we assume that the limit  $n \rightarrow 1$  follows the almost localized Fermi liquid theory [24,25], we should expect the compressibility to scale with  $Z$ . This is in accord with the results of ECFL Fig. 6 panel (b) but not with the DMFT.

### C. Bubble susceptibility

The knowledge of the Green's functions and the numerically determined exact compressibility and magnetic susceptibility  $\chi_{\text{spin}}$  [see below Sec. III F] enable us to check a popular assumption of retaining only the bubble graphs and throwing away the vertex correction for these quantities. We write the charge susceptibility  $\chi_c = dn/d\mu$  as

$$\begin{aligned} \chi_c &= \frac{1}{\beta N_s} \frac{d}{d\mu} \sum_{k, \omega_n, \sigma} e^{i\omega_n 0^+} G_\sigma(k, i\omega_n) \\ &= -\frac{1}{\beta N_s} \sum_{k, \omega_n, \sigma} G_\sigma^2(k, i\omega_n) \left\{ 1 - \frac{d}{d\mu} \Sigma_\sigma(k, i\omega_n) \right\} \end{aligned} \quad (9)$$

and similarly for  $\chi_{\text{spin}}$  by replacing  $\frac{d}{d\mu} \rightarrow \frac{d}{dB}$ , where  $B$  is the magnetic field. The vertex corrections thus correspond to the  $\mu$  or  $B$  derivatives of the *self energy*. Approximating this by dropping the derivative of the self energy, we get  $\chi_c \sim \chi_{\text{spin}} \sim \chi_{\text{Bubble}}$  where

$$\chi_{\text{Bubble}} = -\frac{1}{\beta N_s} \sum_{k, \omega_n, \sigma} G_\sigma^2(k, i\omega_n). \quad (10)$$

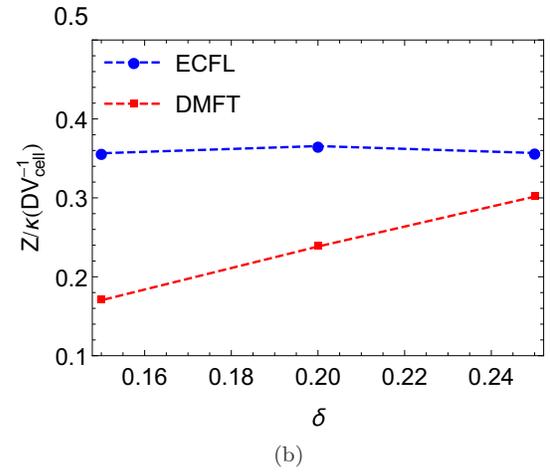
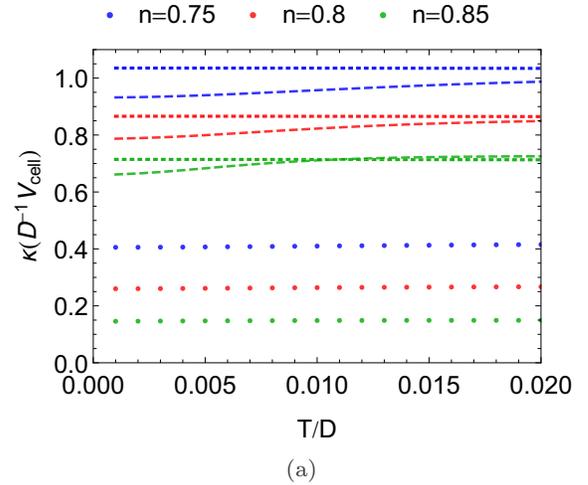


FIG. 6. (a) Compressibility  $\kappa = n^{-2} \partial n / \partial \mu$  of ECFL (symbols), DMFT (dashed lines), and free fermions (dotted lines). The DMFT results give a systematically higher value of compressibility than the ECFL theory. (b)  $Z/\kappa$  for the lowest temperature  $T = 0.001D$  at the three densities considered for ECFL (blue) and DMFT (red). The ECFL result for the compressibility is proportional to the quasiparticle weight  $Z$ , unlike the DMFT result which displays some variation. The difference in compressibility between the two theories seems related to the density dependent shift in chemical potentials noted in Fig. 5.

As usual we can convert the sum to a contour integral using the pole structure of the Fermi function  $f(\omega)$  and write

$$\begin{aligned} \chi_{\text{Bubble}} &= \frac{2}{N_s} \sum_k \int_{\Gamma} \frac{d\omega}{2\pi i} f(\omega) G^2(k, \omega) \\ &= \frac{2}{\pi N_s} \sum_k \int d\omega f(\omega) \Im m G^2(k, \omega + i0^+), \end{aligned} \quad (11)$$

where  $\Gamma$  is a closed contour encircling the imaginary axis in a counterclockwise fashion, and we rotated the axis to a pair of lines parallel to the real axis to obtain the final line. Using the standard definition of the spectral function  $\rho_G(k, \omega) = -\frac{1}{\pi} \Im m G(k, \omega + i0^+)$  we may write  $\Im m G^2(k, \omega + i0^+) = (-2\pi) \Re e G(k, \omega) \rho_G(k, \omega)$  to express  $\chi_{\text{Bubble}} = -\frac{4}{N_s} \sum_k \int d\omega f(\omega) \Re e G(k, \omega) \rho_G(k, \omega)$ . In the limit

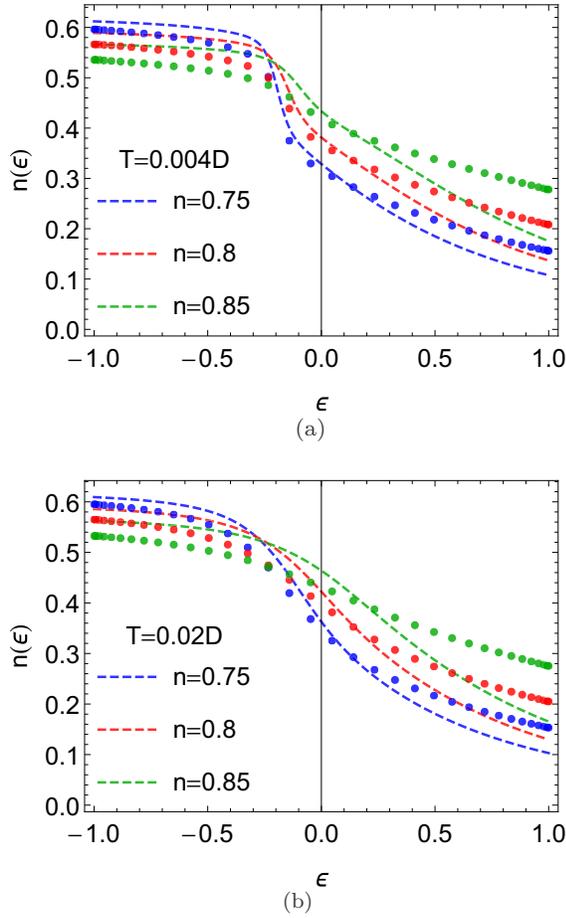


FIG. 7. The momentum distribution curves at three densities  $n = .75, .8, .85$  (top to bottom at  $\epsilon = -1$ ) at  $T = .004 D$  [panel (a)] and  $T = .02 D$  [panel (b)]. The ECFL curves are solid symbols and the DMFT curves are dashed lines.

$d \rightarrow \infty$  the Dyson self energy is independent of  $k$ , and therefore we can write  $\Im m G^2(\epsilon, \omega + i0^+) = \Im m \frac{d}{d\epsilon} G(\epsilon, \omega + i0^+) = -\pi \frac{d}{d\epsilon} \rho_G(\epsilon, \omega)$ , where we exchanged the two operations in the last line. Using the definition of the single particle momentum distribution function  $n_k \rightarrow n(\epsilon) \equiv \int d\omega f(\omega) \rho_G(\epsilon, \omega)$  we can perform the  $\omega$  integration in Eq. (11) and get a compact relation valid in high dimensions:

$$\chi_{\text{Bubble}} = -2 \int d\epsilon \mathcal{D}(\epsilon) \frac{d}{d\epsilon} n(\epsilon). \quad (12)$$

Here  $\mathcal{D}(\epsilon) = \frac{2}{\pi D} \sqrt{1 - \epsilon^2/D^2}$  is the band density of states per site per spin, and  $D$  is the half bandwidth. For noninteracting electrons the function  $n(\epsilon)$  is a constant with a unit jump at  $\epsilon_F$ , and we recover the standard result  $\chi^0 = 2\mathcal{D}(\epsilon_F)$ .

In the correlated problem, the jump at the Fermi energy is  $Z_k$  by Migdal's theorem, and so its contribution to  $\chi_{\text{Bubble}}$  is  $Z_k$ . The background also contributes to the integral in Eq. (12), and it is important to understand its behavior as  $n \rightarrow 1$ . In Fig. 7 we display the momentum distribution at the three densities considered at two temperatures. We note that the entire variation of the monotonic function  $n(\epsilon)$  is on the scale of  $\delta$ ; it settles down to a flat function  $n(\epsilon) = 0.5$  at  $n = 1^-$  and for small departures from half filling, the occupied (unoccupied)

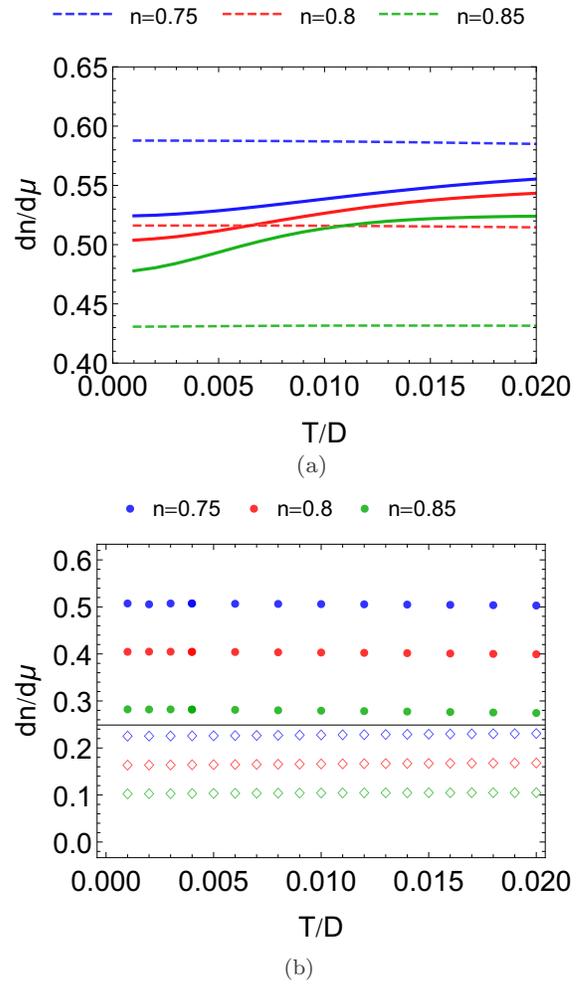


FIG. 8. The charge susceptibilities  $\chi_c = dn/d\mu$ , which are related to compressibility  $\kappa$  as  $\chi_c = n^2 \kappa$ . The numerically exact values versus bubble estimates [Eq. (12)] in panel (a) DMFT (full and dashed lines) and in panel (b) from ECFL (empty diamonds and solid circles).

region is enhanced (depleted) by an area that is proportional to  $\delta = 1 - n$ . Thus we see that as  $n \rightarrow 1$ , the background contribution is at most as large as  $\delta$ , and thus  $\chi_{\text{Bubble}}$  is a suitably weighted average of  $\delta$  and  $Z$ . In the density regimes we are considering, the  $\delta$  variation of  $Z$  is close to  $\delta^{1.39}$  rather than  $\delta$  (see discussion in Ref. [3]), and hence this balance can only be determined by a numerical evaluation. From Eq. (12) we can evaluate  $\chi_{\text{Bubble}}$ , and the results are shown from both theories at the three densities  $\delta = .25, .2, .15$  in Fig. 8. Within ECFL it appears that  $\chi_{\text{Bubble}}$  is dominated by the Migdal jump contribution; the spacing between the three relatively constant lines increases at lower  $\delta$ . Within DMFT the situation appears to be reversed and  $\chi_{\text{Bubble}}$  seems to scale with  $\delta$ . In Fig. 6 we see that the DMFT results for  $Z/\kappa$  have a distinct positive slope relative to the ECFL results, and this is consistent with the above discussed differences in the computed  $\chi_{\text{Bubble}}$  as well.

#### D. Self-energy and local density of states

In this section we study the imaginary part of the self energy  $\rho_\Sigma(\omega) = -\frac{1}{\pi} \Im m \Sigma(\omega)$  and the (local) spectral

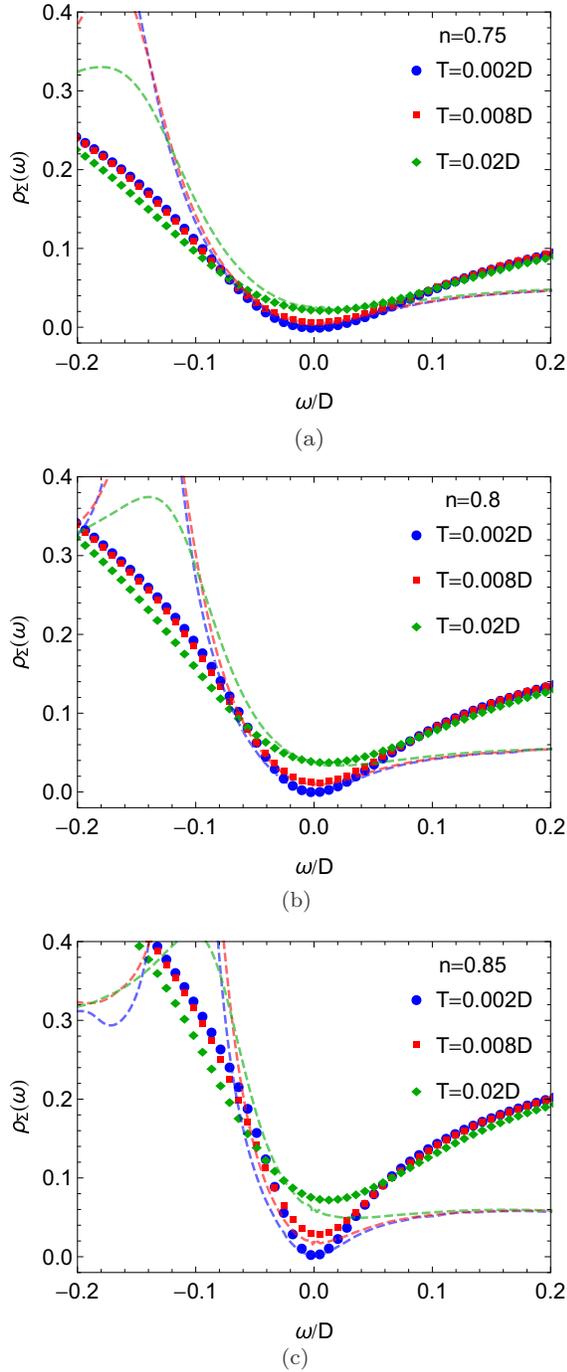


FIG. 9. Single particle decay rates, i.e., the spectral functions of self-energy [ $\rho_{\Sigma}(\omega) = -\pi^{-1} \Im m \Sigma(\omega)$ ] of ECFL (symbols) and DMFT (dashed lines) for a range of temperatures.

function integrated over the band energies  $\rho_G^{\text{loc}}(\omega) = -\frac{1}{\pi} \Im m \int d\epsilon \mathcal{D}(\epsilon) G(\epsilon, \omega)$ . The results of the two theories, including the magnitudes and their variation, are very close at low energies. The ECFL self-energy misses a maximum in  $\rho_{\Sigma}(\omega)$  found in DMFT between  $\omega \sim -0.1D$  and  $\omega \sim -0.2D$ , see Fig. 9. This feature was already noted in Ref. [3] and it is expected to influence the results of various quantities, such as the optical conductivity and dynamical Hall constant, but

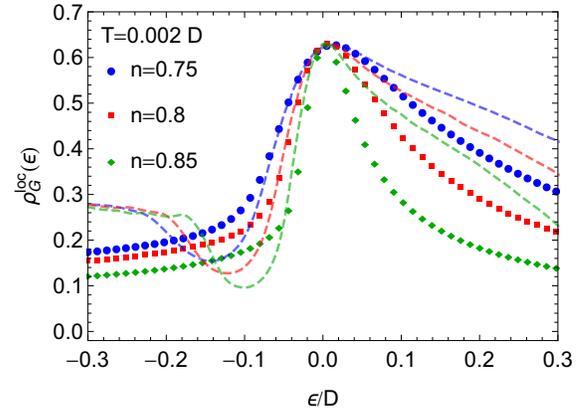


FIG. 10. Local density of states  $\rho_G^{\text{loc}}(\epsilon)$  of ECFL (symbols) and DMFT (dashed lines) at  $T = 0.002D$ .

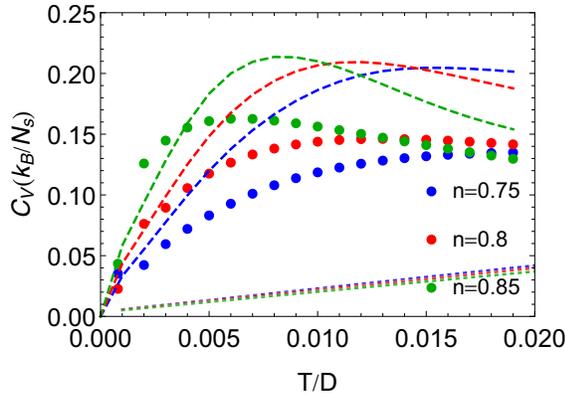
only at a fairly large energy. The imaginary part of the self energy in both theories shows a significant  $\omega^3$  type (i.e., odd in frequency) correction to the simple-minded expectation of a  $\omega^2$  behavior from Fermi liquid theory. This type of a skew has been argued in Ref. [41] to be responsible for the unusual and distinctive spectral functions in real materials such as the cuprates.

The local spectral functions of the two theories, shown in Fig. 10, are similar. They exhibit a sharpening of the maximum as  $n$  increases. Let us note that this object is relevant for angle integrated photoemission studies as well as STM studies, where one would also have to correct for the one electron density of states showing structure beyond that in the present theory.

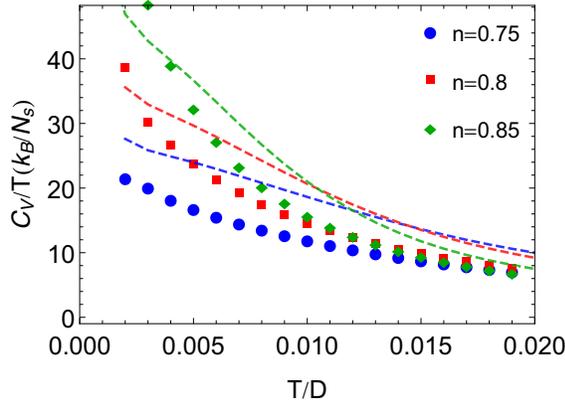
### E. Entropy and heat capacity

The heat capacity is computed in the ECFL theory by numerically differentiating the internal energy as  $C_V = \partial E_K / \partial T$  on a fine  $T$  grid. From its numerical integration  $\int_0^T dT' C_V(T') / T'$  we find the entropy. A similar procedure is used in the DMFT: The kinetic energies were computed on an equally spaced temperature grid (step size  $\Delta T = 10^{-3}D$ ), numerically differentiated, smoothed using a Gaussian filter to obtain the heat capacity  $C_V$ , then interpolated using second-order polynomials, and finally integrated to obtain the entropy.

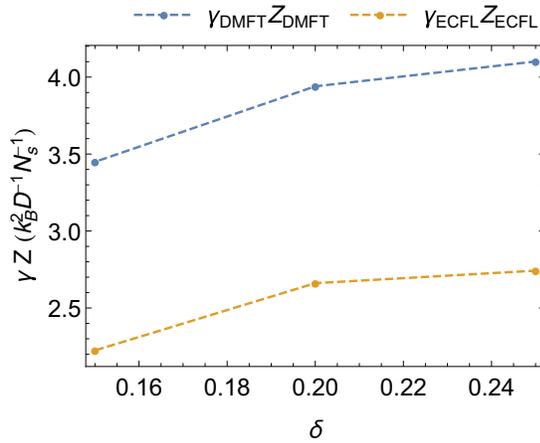
The heat capacity  $C_V$  is displayed in Fig. 11(a). We note that  $C_V$  has a Schottky peak near  $T \sim T_{\text{FL}}$  which becomes sharper as the density increases. At lower densities ( $n = 0.7, 0.75$ ), a linear- $T$  behavior is resolved, as we expect for a Fermi liquid. In Fig. 11(b) we display  $C_V/T$ , from which we see that for densities closer to half-filling ( $n = 0.8, 0.85$ ), the linear behavior of heat capacity is not clearly resolved due to the small  $T_{\text{FL}}$  scale, and also due to increasing numerical uncertainties near half filling. Consequently, we find  $C_V/T$  appears to be growing as  $T$  decreases, instead of saturating. In Fig. 11(c) we show the product of the heat-capacity slope  $\gamma = C_V(T)/T$  and the quasiparticle residue  $Z$  at a low  $T$  corresponding to the GCFL regime. This product is expected to be a constant for localized Fermi liquids [24]. At  $\delta = 0.15$ , we see however some variation in both ECFL and DMFT



(a)



(b)



(c)

FIG. 11. (a) Specific heat computed from the kinetic energy by differentiation as  $C_V = \partial E_K / \partial T$  for ECFL (symbols), DMFT (dashed lines), and free fermions (dotted lines). For  $n = 0.8$  and  $n = 0.85$  the heat capacity shows a gentle maximum at a characteristic  $T$ . (b) The ratio  $C_V/T$  versus  $T$  of ECFL (symbols) and DMFT (dashed lines). Taking the ratio with  $T$  wipes out the maximum seen in (a). (c)  $\gamma \times Z$  at  $T = 0.001D$ .

results. For higher hole densities  $\delta = 0.2, 0.25$ , it is indeed almost a constant.

In Fig. 12 we plot the entropy of the two theories, which give very similar results, and that of the free Fermi gas

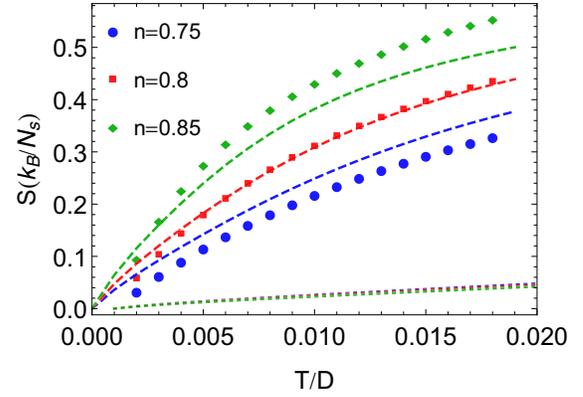


FIG. 12. The entropy versus  $T$  computed as  $\int_0^T dT' C_V(T')/T'$  for ECFL (symbols), DMFT (dashed lines), and free fermions (dotted lines).

with a much lower entropy recovery at these temperatures. It is revealing to compare the heat capacity curve at  $n = 0.8$  in Fig. 11(a), with the resistivity results in Fig. 2 at the same densities. Both theories show a broad maximum in the heat capacity near the corresponding Fermi liquid temperature  $T_{FL}(\delta)$ ; this is the temperature where the GCFL quadratic behavior of resistivity gives way to a linear behavior of the GCSM. At this temperature the entropy per site [see Fig. 12] is  $\sim 0.2 k_B$ , compared to the high  $T$  ( $T = \infty$ ) value of  $1.0119 k_B$ , obtained from  $S_{ideal} \equiv S_{T=\infty} = k_B \{n \log 2 - n \log n - (1-n) \log(1-n)\}$ . This corresponds to about 20% release of the entropy. For comparison, the Fermi gas on the Bethe lattice releases much less, about 1–2% entropy at a comparable  $T/D$ . At lower particle densities  $n = 0.8, 0.75$  we again see that a  $\sim 15$ –20% release of the entropy occurs at the corresponding Fermi liquid temperature  $T_{FL}(\delta)$ , however the heat capacity has a more rounded behavior.

In order to explore this further, in Fig. 13 we display the resistivity and the entropy recovery on the same  $T$  scale. We may thus take as a rule of thumb that at  $T_{FL}$ , the GCFL entropy release is  $\sim 15$ –20% relative to the maximum. This implies a substantial loss of coherence relative to the Fermi gas, i.e., the disordering of either the configurational (i.e., charge) degrees

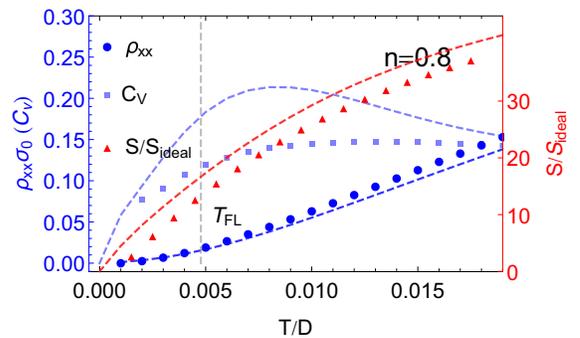


FIG. 13. Resistivity (blue circle), specific heat (light blue square), and entropy (red triangle) as percentage of the ideal entropy at infinite temperature  $S_{ideal}$ . The (Schottky) peak in the heat capacity is close to  $T_{FL}$ , the onset point of the linear- $T$  resistivity, or the end of the crossover region.

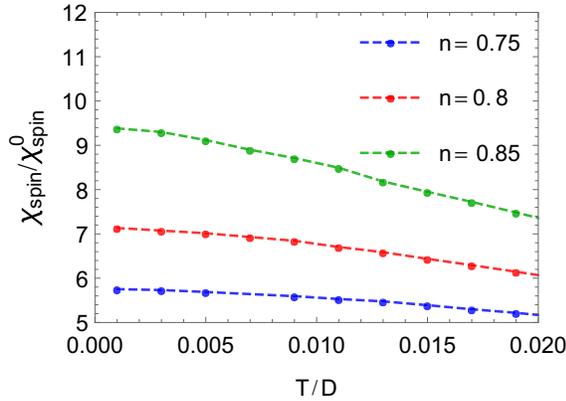


FIG. 14. Magnetic susceptibility (DMFT results). We note that the Stoner enhancement grows as  $\delta \rightarrow 0$  and its  $T$  dependence is Pauli like but with a somewhat enhanced  $T$  dependence at higher  $n$ . The crossover to linear resistivity occurs (see Fig. 2) at fairly low  $T \lesssim .005D$  at these densities but has no reflection on the variation of  $\chi_{\text{spin}}$ . We may thus infer that spin disordering is not relevant to the linear resistivity seen here.

of freedom or to the spins. Below we study the magnetic susceptibility to explore which of these is responsible. We find that the spins are largely unaffected when we go through  $T_{\text{FL}}$ , thereby implicating the charge degrees of freedom.

### F. Magnetic susceptibility

The uniform magnetic susceptibility close to the Mott transition,  $n \gtrsim 0.75$ , is one of the more difficult variables to compute reliably by any technique, since it is highly enhanced by Stoner factors  $\chi_{\text{spin}}/\chi_{\text{spin}}^0 \sim 10$ . In the ECFL theory we found the numerical precision required for computing the susceptibility hard to achieve with the scheme outlined in Ref. [7]. Although the local spectral functions for either spin are confined to a compact region in frequency, it is their difference that is needed for the susceptibility. This difference is numerically very small and smeared over a large frequency range making it very difficult to control. The magnetic susceptibility  $\chi$  is a sensitive variable also within the DMFT using the NRG as the impurity solver, in particular away from half filling at low temperatures, thus it is seldom studied using this approach (see, however Refs. [11,26] for some very early DMFT results, and Ref. [27] for a more recent study using the DMFT(NRG) of the half-filled Hubbard model in magnetic field at  $T = 0$ ). With some effort we have found it possible to estimate its temperature dependence. We used the method of finite field [2,42] with  $H = 10^{-4}D \ll T$ , which is small enough for the system to remain well inside the linear response regime but sufficiently large to be little affected by numerical noise. As a further test, we redid some calculations for  $H = 10^{-3}D$ , finding good consistency of the results.

In Fig. 14 we present the DMFT Stoner enhancement of the susceptibility  $\chi_{\text{spin}}/\chi_{\text{spin}}^0$  as a function of  $T$ . Here the spin susceptibility is denoted by  $\chi_{\text{spin}}$  and for the noninteracting band case it is given by  $\chi_{\text{spin}}^0 = 2\mu_B^2 D(\varepsilon_F)$ , where  $D$  is the band density of states per spin per site defined earlier. The scale of the Stoner enhancement is rather large,  $\sim 10$ . We find that

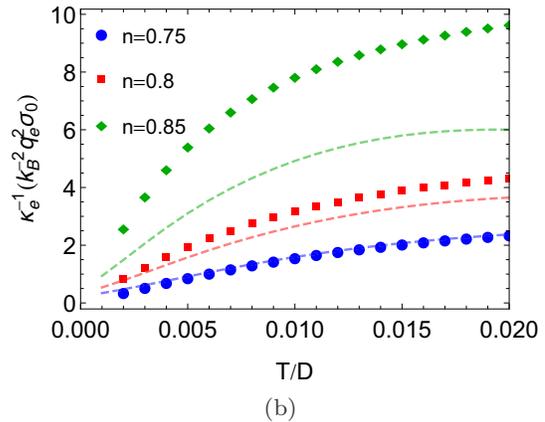
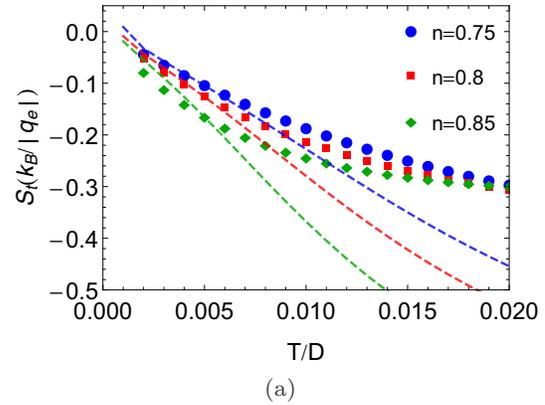


FIG. 15. (a) Thermopower of ECFL (symbols) and DMFT (dashed lines). Both amplitudes and temperature derivatives are similar for  $T \lesssim .005$  but depart at higher  $T$ . (b) Electronic thermal resistivity  $\kappa_e^{-1}$  of ECFL (symbols) and DMFT (dashed lines).

the  $T \rightarrow 0$  value is roughly consistent with  $1/Z$ , as expected for an almost localized Fermi liquid [24].

It is interesting that the Stoner factor and hence  $\chi_{\text{spin}}$  is Pauli-like in the temperature range studied here, i.e., the GCFL and the GCSM regimes. It does not reflect the change in the resistivity behavior from quadratic to linear. Thus the magnetic contribution to the entropy change in Fig. 11 is very small, and we must infer that the GCSM regime continues to have a quenched spin entropy, as in the Fermi liquid. It would appear, by inference, that the entropy released at  $T_{\text{FL}}$  is charge related and the crossover from the Fermi liquid to the GCSM may be viewed as partial charge disordering. This is to be contrasted to the crossover from GCSM to the higher temperature bad metal regime, where the spin degrees of freedom do become partially unscreened [43,44].

### G. Thermoelectric transport

For completeness we present the results for the thermopower  $S_t$ , the electronic thermal conductivity  $\kappa_e$ , and the Lorenz number  $L$ , as well as the thermoelectric figure of merit in Figs. 15 and 16. We record the expressions following from standard transport theory [45]; the thermopower  $S_t$  and electronic thermal conductivity  $\kappa_e$  are expressed in terms of three Onsager transport coefficients  $L_{11}$ ,  $L_{12} = L_{21}$ , and  $L_{22}$  as

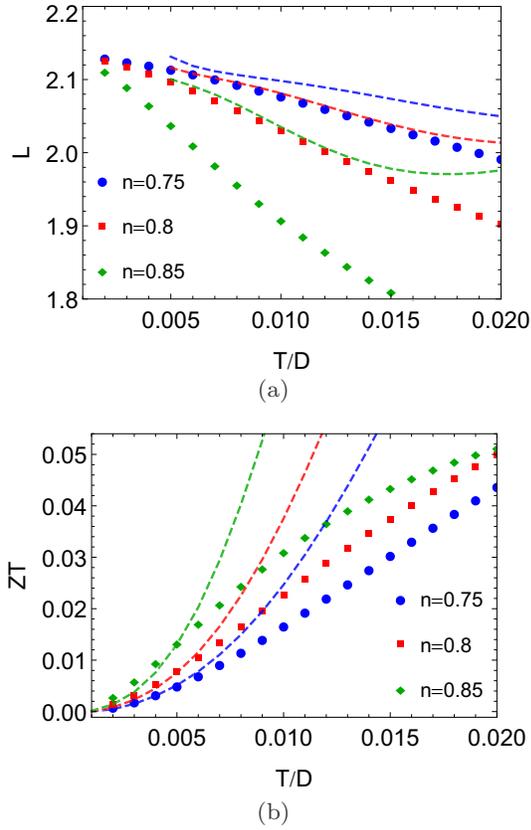


FIG. 16. (a) Lorenz number of ECFL (symbols) and DMFT (dashed lines). The Lorenz number saturates to a constant ( $\simeq 2.1$ ) which is typically expected for a Fermi liquid at low temperatures. (b) Figure of merit for ECFL (symbols) and DMFT (dashed lines). The low values of  $ZT$  found here are typical of normal metals.

follows:

$$\sigma_{xx} = e^2 L_{11}, \quad (13)$$

$$S_t = -\frac{k_B}{|e|T} \frac{L_{12}}{L_{11}}, \quad (14)$$

$$\kappa_e = \frac{k_B^2}{T} \left( L_{22} - \frac{L_{12}^2}{L_{11}} \right). \quad (15)$$

In infinite dimensions, these can be found in a straightforward way from the spectral functions due to vanishing vertex corrections:

$$L_{ij} = \frac{\sigma_0}{e^2} \int d\omega (-f'(\omega)) \omega^{i+j-2} \int d\epsilon \Phi_{xx}(\epsilon) A^2(\omega, \epsilon). \quad (16)$$

The Lorenz number is

$$L = \frac{e^2}{k_B^2} \frac{\kappa_e}{\sigma_{xx} T}, \quad (17)$$

and the electronic thermoelectric figure of merit

$$ZT = T \sigma_{xx} S_t^2 / \kappa_e \quad (18)$$

In the usual Fermi liquid theory, the electronic thermal conductivity  $\kappa_e \sim T^{-1}$  and the thermopower  $S_t \sim \gamma T$ . The

classic Lorenz number for a gas of particles with constant relaxation time is  $L_0 = \pi^2/3$  when we set  $k_B = |e| = 1$ , while for Fermi liquid one expects  $L_{FL} = L_0/1.54 \approx 2.13$  [46]. In previous DMFT studies [45,47–49], thermal transport coefficients were studied focusing on the very high temperature regime of the bad metals. While our results qualitatively agree with the previous studies, the crossover of thermal transport coefficients from GCFL to GCSM in the low- $T$  regime (relative to the very high- $T$  bad metal regime) are resolved. Both the thermopower and thermal resistivity of ECFL change slope near  $T_{FL}$ . In DMFT calculation, only the thermal resistivity shows similar crossover behavior, while the thermopower seems to be insensitive to the crossover from GCFL to GCSM. The Lorenz number of both ECFL and DMFT converges to  $L \simeq 2.1$  in the low- $T$  limit, as expected for a Fermi-liquid ground state. The low values of  $ZT$ , shown in Fig. 16(b), are typical of normal metals.

#### IV. CONCLUSIONS

This work achieves two goals. On one hand, we explored the low-temperature transport regimes of lattice fermions with the constraint of no double occupancy (Gutzwiller projection) in the limit of infinite dimensions. We focus on the temperature range where the Fermi-liquid quadratic resistivity gives way to the first  $T$  linear regime that we dubbed Gutzwiller correlated strange metal; this crossover occurs on the temperature scale which is much lower compared with the bandwidth (and the Brinkman-Rice scale) but which actually corresponds to the experimentally most relevant range of order 100 K. On the other hand, this work had a further methodological goal of comparing the results for a number of transport, spectroscopic, and thermodynamic quantities obtained using the mostly analytical extremely correlated Fermi liquid (ECFL) theory and the accurate numerical results from the dynamical mean field theory (DMFT) approach based on the numerical renormalization group as the impurity solver. We found that at the crossover temperature scale both techniques indicate a change of behavior in most of the quantities we investigated. The two methods have generally good agreement, which improves upon lowering either the temperature or the density.

The origin of the crossover in the resistivity has been tracked down to the temperature dependence of  $-\Im m \Sigma(0, T)$ , the imaginary part of the self-energy on the Fermi surface, which starts to deviate from its low-temperature asymptotic behavior on the scale  $T_{FL}$  (Fermi-liquid temperature). This low-energy scale is produced by purely local Gutzwiller correlation effects, i.e., it is a direct consequence of the constraint of no double occupancy of the lattice sites. We managed to show that  $\rho(T) \propto -\Im m \Sigma(0, T)$  [Eq. (7)], which accounts well for the  $\rho(T)$  dependence in the (GCFL)-Fermi-liquid and (GCSM)-strange metal regimes. As a result, we are able to explain the temperature dependence of the resistivity in terms of the temperature dependence of the imaginary part of the self-energy on the Fermi surface.

The charge compressibility of the DMFT theory at infinite  $U$  is seen to differ somewhat from that of the ECFL and also from the almost localized Fermi liquid. Developments in ECFL are underway in order to resolve the difference from DMFT. The compressibility shows a kink on the scale of  $T_{FL}$  and the

heat capacity has a weak peak. The magnetic susceptibility, however, shows no change across this crossover. The crossover hence seems to be related to the charge degrees of freedom, while the spin entropy is quenched in both Fermi liquid and strange metal regimes. It thus seems that the GCSM regime has a highly unusual composition, with some disordering of the charges, presumably in anticipation of the incipient Mott insulating state, without the participation of the spins.

In a following paper, Ref. [50], we present results for the dynamical Hall constant and Hall angle indicating that the two-relaxation-time behavior in transport properties observed in a number of cuprates emerges upon entering the GCSM regime. Finally we note a recent paper, Ref. [51], where the results of a two-dimensional version of the equations studied here are presented.

## ACKNOWLEDGMENTS

The work at UCSC was supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award No. DE-FG02-06ER46319. R.Ž. acknowledges the financial support from the Slovenian Research Agency (Research Core Funding No. P1-0044 and Project No. J1-7259). We thank Patrick Lee for a stimulating discussion and Dieter Vollhardt for a helpful correspondence. This work used the Extreme Science and Engineering Discovery Environment (XSEDE [52] TG-DMR160144), which is supported by National Science Foundation Grant No. ACI-1053575, and the UCSC supercomputer Hyades, which is supported by National Science Foundation (Award No. AST-1229745) and UCSC.

- 
- [1] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011); **108**, 029702 (2012); *Ann. Phys. (NY)* **343**, 164 (2014).
- [2] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989); A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [3] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013).
- [4] A. Khurana, *Phys. Rev. Lett.* **64**, 1990 (1990).
- [5] X. Y. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [6] W. Xu, K. Haule, and G. Kotliar, *Phys. Rev. Lett.* **111**, 036401 (2013).
- [7] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016).
- [8] E. Perepelitsky, A. Galatas, J. Mravlje, R. Žitko, E. Khatami, B. Sriram Shastry, and A. Georges, *Phys. Rev. B* **94**, 235115 (2016).
- [9] E. Müller-Hartmann, *Z. Phys. B* **74**, 507 (1989).
- [10] M. Jarrell, *Phys. Rev. Lett.* **69**, 168 (1992).
- [11] M. Jarrell and Th. Pruschke, *Phys. Rev. B* **49**, 1458 (1994).
- [12] R. Zitzler, Th. Pruschke, and R. Bulla, *Eur. Phys. J. B* **27**, 473 (2002).
- [13] T. Pruschke and R. Zitzler, *J. Phys.: Condens. Matter* **15**, 7867 (2003).
- [14] G. Sangiovanni, A. Toschi, E. Koch, K. Held, M. Capone, C. Castellani, O. Gunnarsson, S.-K. Mo, J. W. Allen, H.-D. Kim, A. Sekiyama, A. Yamasaki, S. Suga, and P. Metcalf, *Phys. Rev. B* **73**, 205121 (2006).
- [15] R. Peters and T. Pruschke, *Phys. Rev. B* **76**, 245101 (2007).
- [16] M. Fleck, A. I. Liechtenstein, A. M. Oleś, L. Hedin, and V. I. Anisimov, *Phys. Rev. Lett.* **80**, 2393 (1998).
- [17] R. Strack and D. Vollhardt, *Phys. Rev. B* **46**, 13852 (1992).
- [18] Y. Ando, S. Komiya, K. Segawa, S. Ono, and Y. Kurita, *Phys. Rev. Lett.* **93**, 267001 (2004).
- [19] In the Gutzwiller approximation [20,21], double occupancy is projected out of the Gutzwiller wave function, while spin correlations are accounted for by purely combinatorial means. In Refs. [22,23], this approximation is derived diagrammatically and shown to be exact in the limit of infinite spatial dimensions. In lower spatial dimensions, spin correlations play a greater role, and hence the full Gutzwiller wave function, with finite double occupancy, must be used.
- [20] M. C. Gutzwiller, *Phys. Rev.* **134**, A923 (1964).
- [21] M. C. Gutzwiller, *Phys. Rev.* **137**, A1726 (1965).
- [22] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **59**, 121 (1987).
- [23] W. Metzner and D. Vollhardt, *Phys. Rev. B* **37**, 7382 (1988).
- [24] D. Vollhardt, *Rev. Mod. Phys.* **56**, 99 (1984).
- [25] D. Vollhardt, P. Wolfle, and P. W. Anderson, *Phys. Rev. B* **35**, 6703 (1987).
- [26] H. Kajueter, G. Kotliar, and G. Moeller, *Phys. Rev. B* **53**, 16214 (1996).
- [27] J. Bauer, *Eur. Phys. J. B* **68**, 201 (2009).
- [28] We are currently developing a few promising methods for yielding a closer quantitative agreement between the ECFL and DMFT for both quasiparticle residue  $Z$  and the resistivity at *all* densities.
- [29] K. G. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975).
- [30] R. Bulla, T. Costi, and Th. Pruschke, *Rev. Mod. Phys.* **80**, 395 (2008).
- [31] R. Žitko and Th. Pruschke, *Phys. Rev. B* **79**, 085106 (2009).
- [32] R. Žitko, *Comput. Phys. Commun.* **180**, 1271 (2009).
- [33] A. Weichselbaum and J. von Delft, *Phys. Rev. Lett.* **99**, 076402 (2007).
- [34] R. Žitko, *Phys. Rev. B* **80**, 125125 (2009).
- [35] R. Bulla, A. C. Hewson, and Th. Pruschke, *J. Phys.: Condens. Matter* **10**, 8365 (1998).
- [36] The Ioffe-Regel-Mott resistivity is theoretically the largest value of resistance in an impure metal, attained when the mean free path is comparable to the lattice constant. Experimentally it has become standard to identify it with the largest resistivity attained in metallic systems exhibiting the phenomenon of resistivity saturation. This is seen in some strongly electron phonon coupled systems, such as the A-15 materials (e.g., Mo<sub>3</sub>Ge, Nb<sub>3</sub>Sb, or Nb<sub>3</sub>Sn), where the  $T$  linear rise of the phonon scattering resistance rolls over and saturates. The suggested magnitude is from; M. Gurvitch, *Phys. Rev. B* **24**, 7404 (1981); Z. Fisk and G. W. Webb, *Phys. Rev. Lett.* **36**, 1084 (1976).

- [37] S. Martin, A. T. Fiory, R. M. Fleming, L. F. Schneemeyer, and J. V. Waszczak, *Phys. Rev. Lett.* **60**, 2194 (1988).
- [38] G. Bevilacqua, [arXiv:1303.6206](https://arxiv.org/abs/1303.6206).
- [39] The chemical potentials at the lowest  $T$  for the three densities in Fig. 5 are 0.72742, 0.98132, 1.35263 for ECFL and 0.43986, 0.53668, 0.63789 for DMFT.
- [40] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *Phys. Rev. B* **88**, 205108 (2013).
- [41] B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012); G.-H. Gweon, B. S. Shastry, and G. D. Gu, *ibid.* **107**, 056404 (2011).
- [42] M. J. Rozenberg, G. Kotliar, and X. Y. Zhang, *Phys. Rev. B* **49**, 10181 (1994).
- [43] N. Dasari, N. S. Vidhyadhiraja, M. Jarrell, and R. H. McKenzie, *Phys. Rev. B* **95**, 165105 (2017).
- [44] J. Kokalj and R. H. McKenzie, *Phys. Rev. Lett.* **110**, 206402 (2013).
- [45] B. S. Shastry, *Rep. Prog. Phys.* **72**, 016501 (2009); L.-F. Arsenault, B. S. Shastry, P. Sémon, and A.-M. S. Tremblay, *Phys. Rev. B* **87**, 035126 (2013).
- [46] C. Herring, *Phys. Rev. Lett.* **19**, 167 (1967); **19**, 684 (1967); L. V. Pourovskii, J. Mravlje, A. Georges, S. I. Simak, and I. A. Abrikosov, [arXiv:1603.02287](https://arxiv.org/abs/1603.02287).
- [47] J. Merino and R. H. McKenzie, *Phys. Rev. B* **61**, 7996 (2000).
- [48] V. Zlatić and J. K. Freericks, *Phys. Rev. Lett.* **109**, 266601 (2012).
- [49] V. Zlatić and G. R. Boyd, and J. K. Freericks, *Phys. Rev. B* **89**, 155101 (2014).
- [50] A Strange Metal from Gutzwiller correlations in infinite dimensions II: Transverse Transport, Optical Response and Rise of Two Relaxation Rates, W. Ding, R. Žitko, and B. Sriram Shastry, [arXiv:1705.01914](https://arxiv.org/abs/1705.01914).
- [51] B. Sriram Shastry and P. Mai, [arXiv:1703.08142](https://arxiv.org/abs/1703.08142).
- [52] J. Town *et al.*, *Comput. Sci. Eng.* **66**, 62 (2014).



## PAPER

Extremely correlated Fermi liquid theory of the  $t$ - $J$  model in 2 dimensions: low energy properties

B Sriram Shastry and Peizhi Mai

Physics Department, University of California, Santa Cruz, CA 95064, United States of America

E-mail: [sriram@physics.ucsc.edu](mailto:sriram@physics.ucsc.edu)**Keywords:** strong correlations, 2-dimensional  $t$ - $J$  model, temperature dependent resistivity, cotangent Hall angle, extremely correlated Fermi liquid

## RECEIVED

24 August 2017

## REVISED

20 October 2017

## ACCEPTED FOR PUBLICATION

17 November 2017

## PUBLISHED

22 January 2018

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Abstract**

Low energy properties of the metallic state of the two-dimensional  $t$ - $J$  model are presented for second neighbor hopping with hole-doping ( $t' \leq 0$ ) and electron-doping ( $t' > 0$ ), with various super-exchange energy  $J$ . We use a closed set of equations for the Greens functions obtained from the *extremely correlated Fermi liquid theory*. These equations reproduce the known low energies features of the large  $U$  Hubbard model in infinite dimensions. The density and temperature dependent quasiparticle weight, decay rate and the peak spectral heights over the Brillouin zone are calculated. We also calculate the resistivity, Hall conductivity, Hall number and cotangent Hall angle. The spectral features display high thermal sensitivity at modest  $T$  for density  $n \gtrsim 0.8$ , implying a suppression of the effective Fermi-liquid temperature by two orders of magnitude relative to the bare bandwidth. The cotangent Hall angle exhibits a  $T^2$  behavior at low  $T$ , followed by an interesting kink at higher  $T$ . The Hall number exhibits strong renormalization due to correlations. Flipping the sign of  $t'$  changes the curvature of the resistivity versus  $T$  curves between convex and concave. Our results provide a natural route for understanding the observed difference in the temperature dependent resistivity of strongly correlated electron-doped and hole-doped matter.

**1. Introduction**

The  $t$ - $J$  model in 2-dimensions (2d) has been argued to be of fundamental importance for understanding strongly correlated matter, including the high  $T_c$  superconductors [1, 2]. Due to the difficulties inherent in the strong coupling problem, very few techniques are available for extracting its low temperature physics. Towards this end we have recently developed the *extremely correlated Fermi liquid* (ECFL) theory [3, 4]. It is an analytical method for treating very strong correlations of lattice Fermions, employing Schwinger's technique of functional differential equations together with several important added ingredients. While further details can be found in [3, 4], a brief summary of the main idea behind the ECFL theory seems appropriate. We consider the Hubbard model with a large interaction  $U \rightarrow \infty$ , and hence the name of the theory. A well known expansion in the inverse powers of  $U$  leads to the  $t$ - $J$  model (defined below [2]). Taking the infinite  $U$  limit forces one to abandon the conventional Feynman diagram based perturbation theory in  $U$ , and to make a fresh start. The ECFL theory starts with the graded Lie-algebra of the Gutzwiller projected, i.e. infinite- $U$  limit Fermi operators equations (2), (3). This leads to an exact functional differential equation for the Greens functions, known as the Schwinger equation of motion equation (18) or (22). In this equation, a parameter  $\lambda$  is introduced;  $\lambda$  is bounded in the range  $\in [0, 1]$  and represents the evolution from the free Fermi limit. We then use a systematic expansion in the parameter  $\lambda$ , for solving the Schwinger equations perturbatively in  $\lambda$ . In this scheme we start with the uncorrelated Fermi gas at  $\lambda = 0$  and end up at the fully correlated projected Fermion problem at  $\lambda = 1$ . The scheme thus represents a generalization of the usual perturbation theory for canonical Fermionic models, in order to handle a non-canonical Fermionic problem such as the  $t$ - $J$  model. The context of interacting Bosons provides a useful parallel. In the well known problem of representing spin  $S$  variables in terms of canonical Bosons, one uses the expansion parameter  $\frac{1}{2S}$  with a similar range  $\frac{1}{2S} \in [0, 1]$ . We may think of  $\lambda$  as being

analogous to the parameter  $\frac{1}{25}$  as shown in [4]. The introduction of the parameter  $\lambda$  and the  $\lambda$ -expansion scheme thus enabled are among the main technical advances introduced in the ECFL theory.

This approach has been recently benchmarked [5–7] against the numerically exact results from the single impurity Anderson model, and the  $d = \infty$  Hubbard model from dynamical mean field theory (DMFT) [8, 9]. These tests provide quantitative support to our general scheme described below, especially for low energy response. Our scheme has no specific limitation to  $d = \infty$ , and is expected to be reasonably accurate in any dimension  $d > 1$ , including 2 dimensions, a case of great experimental importance due to the High  $T_c$  cuprate materials. It is applied here to probe the metallic state of the  $t$ - $J$  model in 2d. We present results for the electron self energy, the spectral functions, the resistivity the Hall constant and the Hall angle at various temperatures and electron density  $n = N/N_s$  (number of electrons per site). We also frequently use the notation of hole density  $\delta = (1 - n)$  (in addition to  $n$ ), following the convention used in several experimental studies of doped Mott systems.

We explore various values of the parameters of the  $t$ - $J$  model, including the second neighbor hopping, which turns out to play a very important role in determining the effective Fermi liquid (FL) temperature scale. We investigate the resistivity due to mutual collisions of electrons at low temperatures, and its dependence on the parameters of the model. We pay special attention to the resistivity since this easily measured—but notoriously hard to calculate object, reveals the lowest energy scale physics of charge excitations in metallic systems, and therefore is of central importance.

## 2. Methods

In this section we summarize the equations used in the present calculation, together with the arguments leading to them—further details may be found in earlier papers on this theory [3, 5–7]. In section 2.1 the model is defined and the exact Schwinger–Dyson equations of motion (EOM) are written out. In section 2.2 the  $\lambda$  parameter is introduced and the exact factorization of the Greens function into an auxiliary Greens function and a caparison function are noted. In section 2.3 we summarize the shift identities of the  $t$ - $J$  model. The shift transformation is a simple and yet important invariance of the  $t$ - $J$  model leading to important constraints on possible approximations. Within the  $\lambda$  expansion, this invariance obligates the introduction of a second chemical potential  $u_0$ , which is then treated as a Lagrange multiplier to be fixed through sum-rules. In section 2.4 we collect the equations of the second order theory. In section 2.5 we summarize the rationale for a high energy cutoff of the equations given in section 2.4.

### 2.1. The $t$ - $J$ model preliminaries

The  $t$ - $J$  model is a two component Fermi system on a lattice, defined on the restricted subspace of three local states, obtained by excluding all doubly occupied configurations. The allowed states at a single site are  $|a\rangle$  with  $a = 0, \uparrow, \downarrow$ , and the double occupancy state  $|\uparrow\downarrow\rangle$  is removed by the (Gutzwiller) projection operator  $P_G = \prod_i (1 - n_{i\uparrow}n_{i\downarrow})$ . We use the Hubbard operators  $X_i^{a,b} = |a\rangle\langle b|$ , which are expressible in terms of the usual Fermions  $C_{i\sigma}, C_{i\sigma}^\dagger$  and the Gutzwiller projector  $P_G$  as:

$$X_i^{\sigma 0} = P_G C_{i\sigma}^\dagger P_G; \quad X_i^{0\sigma} = P_G C_{i\sigma} P_G; \quad X_i^{\sigma\sigma'} = P_G C_{i\sigma}^\dagger C_{i\sigma'} P_G. \quad (1)$$

These obey the anti-commutation relations

$$\{X_i^{0\sigma_i}, X_j^{\sigma_j 0}\} = \delta_{i,j} (\delta_{\sigma_i, \sigma_j} - \sigma_i \sigma_j X_i^{\sigma_i, \sigma_j}) \quad (2)$$

and the commutators

$$[X_i^{0\sigma_i}, X_j^{\sigma_j \sigma_k}] = \delta_{ij} \delta_{\sigma_i, \sigma_j} X_i^{0\sigma_k}; \quad [X_i^{\sigma_i 0}, X_j^{\sigma_j \sigma_k}] = -\delta_{ij} \delta_{\sigma_i, \sigma_k} X_i^{\sigma_i 0}. \quad (3)$$

The Hamiltonian of the general  $t$ - $J$  model  $H_{tJ}$  is

$$H_{tJ} = H_t + H_J, \quad (4)$$

$$H_t = -\sum_{ij} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_i X_i^{\sigma\sigma}; \quad H_J = \frac{1}{2} \sum_{ij} J_{ij} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} X_i^{\sigma\sigma} X_j^{\sigma'\sigma'} \right),$$

where we sum over repeated spin indices. Here  $\mu$  is the chemical potential and the spin is given in terms of the Fermions and the Pauli matrices  $\vec{\tau}$  as usual  $\vec{S} = \frac{1}{2} X_i^{\sigma 0} \vec{\tau}_{\sigma\sigma'} X_i^{0\sigma'}$ . We will restrict in the following to nearest neighbor exchange  $J$ , and first ( $t$ ) and second neighbor ( $t'$ ) hopping on a square lattice.

For the purpose of computing the Green's functions we add Schwinger sources to the Hamiltonian; the commuting (Bosonic) potential  $\mathcal{V}$  couples to the charge as well as spin density. These sources serve to generate compact Schwinger EOM, and are set to zero at the end. The zero source equations are usually termed as the Schwinger–Dyson equations. In that limit we recover spatial and temporal translation invariance of the Greens

function. Explicitly we write

$$\hat{\mathcal{A}}_S = \sum_i \int_0^\beta \hat{\mathcal{A}}_S(i, \tau) d\tau; \quad \hat{\mathcal{A}}_S(i, \tau) = \mathcal{V}_i^{\sigma' \sigma}(\tau) X_i^{\sigma' \sigma}(\tau) \quad (5)$$

and all time dependences are as in  $Q(\tau) = e^{\tau H_{ij}} Q e^{-\tau H_{ij}}$ . The generating functional of Green's functions of the  $t$ - $J$  model is

$$Z[\mathcal{V}] \equiv \text{Tr}_{ij} e^{-\beta H_{ij}} T_\tau(e^{-\hat{\mathcal{A}}_S}). \quad (6)$$

It reduces to the standard partition function on turning off the indicated source terms the Green's functions for positive times  $0 \leq \tau_j \leq \beta$ , are defined as usual:

$$\mathcal{G}_{\sigma\sigma'}(i\tau_i, f\tau_f) = -\langle T_\tau(e^{-\hat{\mathcal{A}}_S} X_i^{0\sigma}(\tau_i) X_f^{\sigma'0}(\tau_f)) \rangle, \quad (7)$$

where for an arbitrary  $\mathcal{Q}$  we define

$$\langle \mathcal{Q} \rangle \equiv \frac{1}{Z} \text{Tr}_{ij} e^{-\beta H_{ij}} T_\tau(e^{-\hat{\mathcal{A}}_S} \mathcal{Q}). \quad (8)$$

We note that  $n_\sigma$ , the number of particles per site, is determined from the number sum rule:

$$n_\sigma = \mathcal{G}_{\sigma\sigma}(i\tau^-, i\tau) \quad (9)$$

and  $\mu$  the chemical potential is fixed by this constraint. By taking the time derivative of equation (7) we see that the Green's function satisfies the EOM

$$\partial_{\tau_i} \mathcal{G}_{\sigma_i \sigma_j}(i, f) = -\delta(\tau_i - \tau_f) \delta_{ij} (1 - \gamma_{\sigma_i \sigma_j}(i\tau_i)) - \langle T_\tau(e^{-\hat{\mathcal{A}}_S} [H_{ij} + \hat{\mathcal{A}}_S(i, \tau_i), X_i^{0\sigma_i}(\tau_i)] X_f^{\sigma_j 0}(\tau_f)) \rangle, \quad (10)$$

where the local Green's function is defined as

$$\gamma_{\sigma_a \sigma_b}(i\tau_i) = \sigma_a \sigma_b \mathcal{G}_{\bar{\sigma}_b \bar{\sigma}_a}(i\tau_i^-, i\tau_i), \quad (11)$$

with the notation

$$\bar{\sigma}_i = -\sigma_i. \quad (12)$$

Using the Hamiltonian equation (4) and canonical relations equations (2), (3) we find

$$[H_{ij}, X_i^{0\sigma_i}] = \sum_j t_{ij} X_j^{0\sigma_i} + \mu X_i^{0\sigma_i} - \sum_{j\sigma_j} t_{ij} (\sigma_i \sigma_j) X_i^{\bar{\sigma}_i \bar{\sigma}_j} X_j^{0\sigma_j} + \frac{1}{2} \sum_{j \neq i} J_{ij} (\sigma_i \sigma_j) X_j^{\bar{\sigma}_i \bar{\sigma}_j} X_i^{0\sigma_j} \quad (13)$$

and

$$[\hat{\mathcal{A}}_S(i\tau_i), X_i^{0\sigma_i}] = -\mathcal{V}_i^{\sigma_i \sigma_i} X_i^{0\sigma_i}. \quad (14)$$

Substituting into equation (10) and using the free Fermi gas Green's function:

$$\mathbf{g}_{0, \sigma_i \sigma_j}^{-1}(i\tau_i, j\tau_j) = \{\delta_{\sigma_i \sigma_j} [\delta_{ij} (\mu - \partial_{\tau_i}) + t_{ij}] - \delta_{ij} \mathcal{V}_i^{\sigma_i \sigma_j}(\tau_i)\} \delta(\tau_i - \tau_j), \quad (15)$$

we obtain

$$\begin{aligned} \mathbf{g}_{0, \sigma_i \sigma_j}^{-1}(i\tau_i, j\tau_j) \mathcal{G}_{\sigma_i \sigma_j}(j\tau_j, f\tau_f) &= \delta(\tau_i - \tau_f) \delta_{ij} (1 - \gamma_{\sigma_i \sigma_j}(i\tau_i)) \\ &- \sum_{j\sigma_j} t_{ij} (\sigma_i \sigma_j) \langle T_\tau(X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) X_j^{0\sigma_j}(\tau_i) X_f^{\sigma_j 0}(\tau_f)) \rangle + \frac{1}{2} \sum_{k\sigma_k} J_{ik} (\sigma_i \sigma_j) \langle T_\tau(X_k^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i) X_i^{0\sigma_i}(\tau_i) X_f^{\sigma_j 0}(\tau_f)) \rangle. \end{aligned} \quad (16)$$

We next 'reduce' the higher order Green's function to a lower one using the identity (valid for any operator  $\mathcal{Q}$ ):

$$\langle T_\tau X_i^{\sigma\sigma'}(\tau) \mathcal{Q} \rangle = \langle T_\tau X_i^{\sigma\sigma'}(\tau) \rangle \langle T_\tau \mathcal{Q} \rangle - \frac{\delta}{\delta \mathcal{V}_i^{\sigma\sigma'}(\tau)} \langle T_\tau \mathcal{Q} \rangle \quad (17)$$

and rearranging terms we obtain the fundamental Schwinger EOM:

$$(\mathbf{g}_{0, \sigma_i \sigma_j}^{-1}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) - Y_{\sigma_i \sigma_j}(i\tau_i, j\tau_j)) \times \mathcal{G}_{\sigma_i \sigma_j}(j\tau_j, f\tau_f) = \delta_{ij} \delta(\tau_i - \tau_f) (\delta_{\sigma_i \sigma_j} - \gamma_{\sigma_i \sigma_j}(i\tau_i)), \quad (18)$$

where we defined the functional derivative operator at site  $i$  and time  $\tau_i$

$$D_{\sigma_i \sigma_j}(i\tau_i) = \sigma_i \sigma_j \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau_i)}, \quad (19)$$

the composite derivative operator

$$\hat{X}_{\sigma_i \sigma_j}(i\tau_i, j\tau_j) = \delta(\tau_i - \tau_j) \times \left( -t_{ij} D_{\sigma_i \sigma_j}(i\tau_i) + \delta_{ij} \sum_k \frac{1}{2} J_{ik} D_{\sigma_i \sigma_j}(k\tau_i) \right) \quad (20)$$

and corresponding  $Y_i$  as

$$Y_{i\sigma_i\sigma_j}(i\tau_i, j\tau_j) = -\delta(\tau_i - \tau_j) \times \left( -t_{ij}\gamma_{\sigma_i\sigma_j}(i\tau_i) + \delta_{ij} \sum_k \frac{1}{2} J_{ik}\gamma_{\sigma_i\sigma_j}(k\tau_i) \right). \quad (21)$$

By considering the spin, space and time variables as generalized matrix indices, we can symbolically write equation (18) as

$$(\mathbf{g}_0^{-1} - \hat{X} - Y_i) \cdot \mathcal{G} = \delta (\mathbf{1} - \gamma). \quad (22)$$

## 2.2. The $\lambda$ expansion and the auxiliary Greens function

The main task is to compute solutions of the Schwinger–Dyson equation, i.e. the functional differential equation (18) or (22). If symmetry-breaking, such as magnetism or superconductivity is ignored, then a liquid state ensues, where we would like the solution to connect continuously with the Fermi gas. For this purpose we seek guidance from standard Feynman–Dyson perturbation theory for canonical models. The repulsive Hubbard model is an ideal example, where the corresponding Schwinger–Dyson equation can be schematically written as:

$$(\mathbf{g}_0^{-1} - U\delta/\delta\mathcal{V} - UG) \cdot G = \delta \mathbf{1}. \quad (23)$$

Comparing with equation (22), we see that the left-hand sides are of the same form, but the right-hand sides differ, in equation (22) the local Greens function  $\gamma$  multiplies the delta function. In turn this extra term originates from the second (non canonical) term in the anti-commutator in equation (2), and is therefore the signature term of extremely strong correlations.

Within the Schwinger viewpoint of equation (23), we can view the skeleton graph perturbation theory (Feynman–Dyson) as an iterative scheme in  $U$ , i.e. using the  $n$ th order results to generate the  $(n + 1)$ th order terms by functional differentiation. In the ECFL theory the iterative scheme used is defined by generalizing equation (22) to

$$(\mathbf{g}_0^{-1} - \lambda\hat{X} - \lambda Y_i) \cdot \mathcal{G} = \delta (\mathbf{1} - \lambda\gamma). \quad (24)$$

The explicit solutions in the ECFL theory start from this basic equation. More explicitly, in equation (24) the exact equation (18) is generalized to include the  $\lambda$  parameter<sup>1</sup> by scaling  $\hat{X}_{\sigma_i\sigma_j}, Y_{i\sigma_i\sigma_j}, \gamma_{\sigma_i\sigma_j} \rightarrow \lambda\hat{X}_{\sigma_i\sigma_j}, \lambda Y_{i\sigma_i\sigma_j}, \lambda\gamma_{\sigma_i\sigma_j}$ . The starting point for the iteration is  $\lambda = 0$ , corresponding to the Fermi gas. As we iterate towards  $\lambda = 1$ , equation (24) reduces to the exact equation (22). The Gutzwiller projection is fully effective only at the end point of the iterative scheme  $\lambda = 1$ , while for intermediate values of  $\lambda$ , we have only a partial reduction of the number of doubly occupied sites. The role of  $U$  in equation (23) is roughly similar, at  $U = 0$  we have the Fermi gas, which evolves into an interacting theory with increasing  $U$ , giving us the Feynman–Dyson perturbation theory. The range of  $\lambda \in [0, 1]$  in equation (24) is bounded above, as opposed to that of  $U \in [0, \infty]$  in equation (23). Therefore the ECFL theory avoids dealing with a major headache of the canonical theory whenever a coupling constant becomes large. Recall that realistic interactions in correlated matter usually involve a large coupling parameter  $U$ . For this purpose one is forced to make hard-to-control approximations, such as summing specific classes of diagrams in different parameter ranges. The introduction of  $\lambda$  into the ECFL equations opens the possibility that a low order calculation might suffice to give accurate results at low excitation energies. This possibility is in-fact realized for important strong coupling problems as shown earlier [5].

We found in [3] that an efficient method for proceeding with the iterative scheme is to first perform a factorization of the Greens function into two parts. The first is an auxiliary Greens function  $\mathbf{g}$  satisfying a canonical equation, thus admitting a Dysonian expansion with its attendant advantage of summing a geometric series with every added term of the denominator. There remain some terms that cannot be pushed into the denominator, these are collected together as the caparison function  $\tilde{\mu}$ . In the matrix notation used above we first decompose the Greens function as:

$$\mathcal{G} = \mathbf{g} \cdot \tilde{\mu}, \quad (25)$$

this implies a product in the  $\vec{k}, \omega$  domain as written below in equation (32). The differential operator  $X$  in equation (24) is distributed over the two factors of equation (25) using the Leibniz product rule, as

$$X \cdot \mathbf{g} \cdot \tilde{\mu} = \overline{X \cdot \mathbf{g}} \cdot \tilde{\mu} + \overline{X \cdot \tilde{\mu}} \cdot \mathbf{g} \quad (26)$$

<sup>1</sup> In [4] we have noted an important generalization of these commutators to include a continuous parameter  $\lambda \in [0, 1]$ , thus defining the so called  $\lambda$  fermions. Using them one can systematically obtain the  $\lambda$  expansion encountered below from these relations directly. Here we stick to a simpler description with  $\lambda$  introduced by hand, in the EOM below.

where the contraction symbol  $\overline{\lambda}_a$  indicates the term being differentiated by the functional derivative terms in  $X$ , while the matrix indices follow the dots. Using  $\mathbf{g}^{-1} \cdot \mathbf{g} = \mathbb{1}$  equation (24) is now written as

$$(\mathbf{g}_0^{-1} - \lambda \overline{\lambda}_a \cdot \mathbf{g} \cdot \mathbf{g}^{-1} - \lambda Y_1) \cdot \mathbf{g} \cdot \tilde{\mu} = \delta (\mathbb{1} - \lambda \gamma) + \lambda \overline{\lambda}_a \cdot \mathbf{g} \cdot \tilde{\mu}. \quad (27)$$

This equation factors exactly into two equations upon insisting that  $\mathbf{g}$  has a canonical structure:

$$(\mathbf{g}_0^{-1} - \lambda \overline{\lambda}_a \cdot \mathbf{g} \cdot \mathbf{g}^{-1} - \lambda Y_1) = \mathbf{g}^{-1} \quad (28)$$

and

$$\tilde{\mu} = \delta (\mathbb{1} - \lambda \gamma) + \lambda \overline{\lambda}_a \cdot \mathbf{g} \cdot \tilde{\mu}. \quad (29)$$

We can then use  $\mathbf{g} \cdot \mathbf{g}^{-1} = \mathbb{1}$  to simplify the term  $\overline{\lambda}_a \cdot \mathbf{g} \cdot \mathbf{g}^{-1} = -\overline{\lambda}_a \cdot \mathbf{g} \cdot \mathbf{g}^{-1}$ , giving rise to a Dyson self-energy expressed in terms of a Dyson vertex function. The idea then is to iterate the pair of equations (28), (29) jointly in  $\lambda$ . Details of the skeleton expansion nature can be found in [3, 5, 6]. The main point to note is that while  $\mathbf{g}^{-1}$ ,  $\tilde{\mu}$  in equations (28) and (29) are expanded in powers of  $\lambda$ , the function  $\mathbf{g}$  is kept unexpanded as a basis term (or ‘atom’) of the skeleton expansion, temporarily ignoring its relationship as the inverse of  $\mathbf{g}^{-1}$ . The equal time value of the variable  $\gamma$  in equation (11) is taken from the exact sum-rule for  $\mathcal{G}$  in equation (9). The initial values at  $\lambda = 0$  are  $\mathbf{g} = \mathbf{g}_0$  and  $\tilde{\mu} = \mathbb{1}$ , and we must remember to use the product form equation (25) to determine the local Greens function  $\gamma$  in equation (11). We should note that when the source is turned off  $\mathcal{V} \rightarrow 0$  we recover space and time translation invariance so that equation (25) is simply  $\mathcal{G}(\vec{k}, i\omega_j) = \mathbf{g}(\vec{k}, i\omega_j) \cdot \tilde{\mu}(\vec{k}, i\omega_j)$ , with the Matsubara frequency  $\omega_j = (2j + 1)\pi k_B T$ . At low  $T$ , the leading singularities of  $\mathcal{G}$  are co-located with those of  $\mathbf{g}$ , provided the caparison function  $\tilde{\mu}$  is sufficiently smooth- this situation is realized in all studies done so far.

### 2.3. The shift identities and second chemical potential $u_0$

Before proceeding with the iterative scheme, it is important to discuss a simple but crucial symmetry of the  $t$ - $J$  model—the *shift invariance*, first noted in [10]. In an exact treatment shifting  $t_{ij} \rightarrow t_{ij} + c_t \delta_{ij}$  with  $c_t$  arbitrary, is easily seen to be innocuous, it merely adds to equation (4) a term  $-c_t \sum_{\sigma} \hat{N}_{\sigma}$  whereby the center of gravity of the band is displaced. (Here  $\hat{N}_{\sigma}$  is the number operator for electrons with spin  $\sigma$ .) However in situations such as the  $\lambda$  expansion, the Gutzwiller constraint is released at intermediate values, here it has the effect of adding terms derivable from a local (i.e. Hubbard type) interaction term.<sup>2</sup> To see this consider the fundamental commutator term  $[H_{ij}, X_i^{0\sigma}]$  in equation (13), here under the shift  $t_{ij} \rightarrow t_{ij} + c_t \delta_{ij}$ , the third term gives rise to an extra term  $c_t X_i^{\bar{\sigma}_i} X_i^{0\sigma}$ . This term vanishes only in a Gutzwiller projected state, the EOM by themselves do not eliminate it. Its appearance is tantamount to adding a Hubbard like term  $\frac{c_t}{2} \sum_i X_i^{\sigma\sigma} X_i^{\bar{\sigma}\bar{\sigma}}$  to the Hamiltonian  $H_{ij}$ . As argued in [10] we would like the EOM for the Greens functions to be explicitly invariant under the above shift of  $t_{ij}$  to *each order in  $\lambda$* . Enforcing this shift invariance to each order in the  $\lambda$  expansion plays an important ‘watchdog’ role on the  $\lambda$  expansion.

An efficient method to do so is to explicitly introduce an extra Lagrange multiplier  $u_0$  through a term  $\lambda u_0 \sum_i N_{i\uparrow} N_{i\downarrow}$  to the Hamiltonian equation (4). This amounts to replacing  $t_{ij} \rightarrow t_{ij} + \delta_{ij} \frac{u_0}{2}$  in all terms other than in the bare propagator  $\mathbf{g}_0$ . The  $u_0$  term makes no difference when  $\lambda$  is set at unity in the exact series, since double occupancy is excluded. In practice, we set  $\lambda = 1$  in equations that are truncated at various orders of  $\lambda$ , and the magnitude of  $u_0$  is fixed through a second constraint. We thus have two variables to fix, namely  $u_0$  and  $\mu$ . We also have two constraints, the number sum-rules  $n_{\sigma} = \mathbf{g}_{\sigma\sigma}(i\tau^-, i\tau)$ , and  $n_{\sigma} = \mathcal{G}_{\sigma\sigma}(i\tau^-, i\tau)$  (equation (9)). In the absence of a magnetic field the number densities  $n_{\sigma}$  reduce as  $n_{\sigma} \rightarrow \frac{n}{2}$ , where  $n$  is the number of particles per site.

After turning off the sources, in the momentum–frequency space we can further introducing two self energies  $\Psi(k, i\omega_j)$ , and  $\Phi(k, i\omega_j)$  with

$$\tilde{\mu}(\vec{k}, i\omega_j) = 1 - \lambda \frac{n}{2} + \lambda \Psi(\vec{k}, i\omega_j), \quad (30)$$

$$\mathbf{g}^{-1}(\vec{k}, i\omega_j) = \mathbf{g}_0^{(-1)}(\vec{k}, i\omega_j) + \lambda \left( \frac{n}{2} \varepsilon_k + \frac{n}{4} J_0 \right) - \lambda \Phi(\vec{k}, i\omega_j). \quad (31)$$

Here  $\varepsilon_k$  and  $J_k$  are the Fourier transforms of  $-t_{ij}$  and  $J_{ij}$ . In the right hand side of equation (31), the second and third terms arise respectively from the equal-time limit of  $\lambda Y_1$  and  $\lambda \overline{\lambda}_a \cdot \mathbf{g} \cdot \mathbf{g}^{-1}$  in equation (28) respectively. The two self energies  $\Phi$ ,  $\Psi$  are explicitly  $\lambda$  dependent, they vanish at infinite frequency for any  $\lambda$ . Thus we write

<sup>2</sup> Similarly we note that shifting  $J_{ij} \rightarrow J_{ij} + c_j \delta_{ij}$  with arbitrary  $c_j$  also adds a similar unphysical local interaction term, as discussed in greater detail in [10].

$$\mathcal{G}(k, i\omega_j) = \mathbf{g}(k, i\omega_j) \times \tilde{\mu}(k, i\omega_j). \quad (32)$$

The auxiliary Greens function satisfies a second sum-rule that is identical to equation (9), both may written in the Fourier domain:

$$(k_B T) \sum_{k,j} e^{i\omega_j 0^+} G_{\sigma\sigma}(k, i\omega_j) = n_\sigma; \text{ for both } G = \mathcal{G} \text{ and } \mathbf{g}. \quad (33)$$

Equation (25) can now be written explicitly in the non-Dysonian form proposed in [3]

$$\mathcal{G}(\vec{k}, i\omega_j) = \frac{1 - \lambda \frac{n}{2} + \lambda \Psi(\vec{k}, i\omega_j)}{\mathbf{g}_0^{(-1)}(\vec{k}, i\omega_j) + \lambda \frac{n}{2} \varepsilon_k + \lambda \frac{n}{4} J_0 - \lambda \Phi(\vec{k}, i\omega_j)}. \quad (34)$$

We observe that simple FL-type self energies  $\Psi$  and  $\Phi$  can, in the combination above, lead to highly asymmetric (in frequency) Dyson self energy  $\Sigma(k, \omega)$  obtainable from the Greens function through  $\Sigma = G_0^{-1} - G^{-1}$  [3, 5, 6, 10]. Finally we note that our calculations are performed in terms of spectral function obtainable from analytic continuation of the Matsubara frequencies into the upper complex half plane of frequencies:

$$\begin{aligned} \rho_{\mathcal{G}}(\vec{k}, \omega) &= -\frac{1}{\pi} \Im \mathcal{G}(\vec{k}, i\omega_j \rightarrow \omega + i0^+), \\ \mathcal{G}(\vec{k}, i\omega_j) &= \int \frac{\rho_{\mathcal{G}}(\vec{k}, \omega)}{i\omega_j - \omega} \end{aligned} \quad (35)$$

and similarly defined spectral functions for variables  $\mathbf{g}$ ,  $\Phi$ ,  $\Psi$  etc. Note that the physical spectral function  $\rho_{\mathcal{G}}(\vec{k}, \omega)$  is identical to  $A(\vec{k}, \omega)$ , a notation used in much of experimental literature.

#### 2.4. Summary of equations to second order in $\lambda$

In the following, we use the minimal second order equations [5–7] obtained by expanding equations (28) and (29) to second order in  $\lambda$ . The calculation is straightforward and a systematic derivation is detailed in [6], which is followed here. We use the abbreviation  $^3 k \equiv (\vec{k}, i\omega_n)$ , and also redefine  $\Phi(k) = \chi(k) + \varepsilon_k \Psi(k)$ , keeping in mind that one set of terms in  $\Phi$  have an external common factor of  $\varepsilon_k$  multiplied by all terms in  $\Psi$ . We next collect the answers below in terms of the two self energies  $\chi$ ,  $\Psi$

$$\mathbf{g}^{-1}(k) = i\omega_n + \boldsymbol{\mu} - \underline{\varepsilon}_k + \lambda \frac{1}{4} n J_0 - \varepsilon_k \left( -\lambda \frac{n}{2} + \lambda \Psi \right) - \lambda \chi(k), \quad (36)$$

the tag below the band energy  $\underline{\varepsilon}_k$  can be ignored after the next paragraph. We now expand  $\Psi$  and  $\chi$  from equations (28) and (29) in powers of  $\lambda$ . To the lowest two orders we find  $\Psi = \lambda \Psi_{[1]} + O(\lambda^2)$  and  $\chi = \chi_{[0]} + \lambda \chi_{[1]} + O(\lambda^2)$ , where  $\chi_{[0]} = -\sum_p \mathbf{g}_p (\varepsilon_p + \frac{1}{2} J_{k-p})$ .

The next step is to introduce  $u_0$  explicitly: we write  $\varepsilon_k \rightarrow \varepsilon'_k = \varepsilon_k - \frac{u_0}{2}$  in every occurrence of  $\varepsilon_k$ , except in the bare band energy term  $\underline{\varepsilon}_k$  in equation (36).

$$\mathbf{g}^{-1}(k) = i\omega_n + \boldsymbol{\mu} + \lambda \frac{1}{4} n J_0 - \frac{1}{2} u_0 - \tilde{\mu}(k) \varepsilon'_k - \lambda \chi_{[0]}(k) - \lambda^2 \chi_{[1]}(k). \quad (37)$$

Note that the shift with  $u_0$  also applies to the term  $\chi_{[0]}$ , it now reads  $\chi_{[0]} = -\sum_p \mathbf{g}_p (\varepsilon'_p + \frac{1}{2} J_{k-p})$ . We note the expressions for  $\chi_{[1]}$ ,  $\Psi_{[1]}$  from [6] equations (65)–(67):

$$\chi_{[1]}(k) = -\sum_{pq} \left( \varepsilon'_p + \varepsilon'_q + \frac{1}{2} (J_{k-p} + J_{k-q}) \right) \times (\varepsilon'_{p+q-k} + J_{q-k}) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k), \quad (38)$$

$$\Psi_{[1]}(k) = -\sum_{pq} (\varepsilon'_p + \varepsilon'_q + J_{k-p}) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k). \quad (39)$$

We now set  $\lambda = 1$  and record the final equations:

$$\mathbf{g}^{-1}(k) = i\omega_n + \left( \boldsymbol{\mu} + \frac{1}{4} n (J_0 - u_0) - \frac{1}{2} u_0 + \sum_p \mathbf{g}_p \varepsilon_p + \frac{J_k}{2} \sum_p \mathbf{g}_p \cos p_x \right) - \tilde{\mu}(k) \varepsilon'_k - \chi_{[1]}(k), \quad (40)$$

$$\tilde{\mu}(k) = 1 - \frac{n}{2} + \Psi_{[1]}(k), \quad (41)$$

where we used a nearest neighbor  $J_{ij}$  and cubic symmetry in the simplifications. We can verify that the above expressions obey the shift invariance: if we shift  $\varepsilon_k \rightarrow \varepsilon_k + c_0$ , the arbitrary constant  $c_0$  can be absorbed by shifting  $\boldsymbol{\mu} \rightarrow \boldsymbol{\mu} + c_0$  and  $u_0 \rightarrow u_0 + 2c_0$ , and is thus immaterial. The band energy is given explicitly as

<sup>3</sup> We denote  $k \equiv (\vec{k}, i\omega_n)$ ,  $\omega_n = (2n+1)\pi k_B T$  the Matsubara frequencies,  $N_s$  the number of sites and  $\sum_k \equiv \frac{k_B T}{N_s} \sum_{k_x, k_y, \omega_n} J_k$  is the Fourier transform of the exchange.

$\varepsilon_k = -2t(\cos(k_x a_0) + \cos(k_y a_0)) - 4t' \cos(k_x a_0) \cos(k_y a_0)$ , where  $t$  and  $t'$  are the first and second neighbor hopping amplitudes.

### 2.5. High energy cutoff scheme

The self consistent solution of the second order equations of equations (38)–(41) plus the number sum-rules, can be found numerically by discretizing the momentum and frequency variables on a suitable grid. This procedure can be carried out in a straightforward way for low  $T \lesssim t$  and high hole densities  $\delta \gtrsim 0.3$  (low particle densities  $n \lesssim 0.7$ ). At lower hole densities or at high temperature  $T \gg t$ , the equations run into convergence problems. The origin of this problem is the formation of weak and featureless tails of the spectral functions extending to quite high energies. These tails are known to be artificial, since they do not occur in the exact numerical solutions where available. Thus the second order theory seems insufficient in the regime of low hole densities  $\delta \lesssim 0.2$ , where much of the current interest lies. A technically rigorous resolution of the problem of weak tails seems possible. However it requires the non-trivial calculation of higher order terms in the  $\lambda$  expansion. Such higher order terms oscillate in sign and hence cancellations at high energies are expected.

In view of the substantial magnitude of the program of summing the  $\lambda$  series to high orders, it seems worthwhile to investigate simpler and physically motivated approximations for improving the lowest order scheme. It turns out that there are a few interesting alternatives in this direction. In [5] we showed one convenient way to handle the high energy tail problem practically, through the introduction of a high energy cutoff. The choice of an objective cutoff was rationalized by considering two physically different limits, that of high particle density  $n \rightarrow 1$  and the simpler high temperature limit, where related tails are found. The cutoff is chosen using the analytically available high  $T$  limit results and then applied to all densities and  $T$ .

The cutoff scheme of [5] is not rigorous, but enables us to extract meaningful results for low energy excitations from the second order  $\lambda$  equations, out to fairly low hole densities  $\delta \lesssim 0.2$ . It is benchmarked in the case of  $d = \infty$ , where the cutoff scheme quantitatively reproduces the spectral weights in the most interesting regime of low energies  $|\omega| \ll t$ , while erring somewhat at energies above the scale of quarter bandwidth. In [5, 7] the resulting physical quantities such as resistivity are shown to be in good correspondence to the exact results from DMFT. In view of this success we use a similar cutoff scheme for 2d below, with the expectation that the physics of the low energy excitations is captured. In the present 2d case we employ a single (re)-normalization the spectral function for each  $\vec{k}$  as

$$\hat{\rho}_{\mathbf{g}}(\vec{k}, \omega) = \frac{1}{\mathcal{N}_{\vec{k}}} W_T(\omega - \bar{\varepsilon}_{\vec{k}}) \rho_{\mathbf{g}}(\vec{k}, \omega), \quad (42)$$

where  $W_T$  is a smooth window (even) function shown in figure 3 [5]. It is centered on  $\bar{\varepsilon}_{\vec{k}}$ , the self-consistent location of the peak in  $\rho_{\mathbf{g}}(\vec{k}, \omega)$ , determined iteratively. It has width  $4D$ , where  $2D$  is the bandwidth ( $\sim 8t$  in this case). The constant  $\mathcal{N}_{\vec{k}}$  is fixed by the normalization condition  $\int \hat{\rho}_{\mathbf{g}}(\vec{k}, \omega) d\omega = 1$ . In the present case of 2d we can impose this cutoff window at each  $\vec{k}$  individually, so that only  $\vec{k}$  states very far from the chemical potential are affected by the cutoff.

The two chemical potentials  $\mu$  and  $u_0$  are determined through the number sum rules written in terms of the Fermi function  $f(\omega) = (1 + e^{\beta\omega})^{-1}$  and the spectral functions:

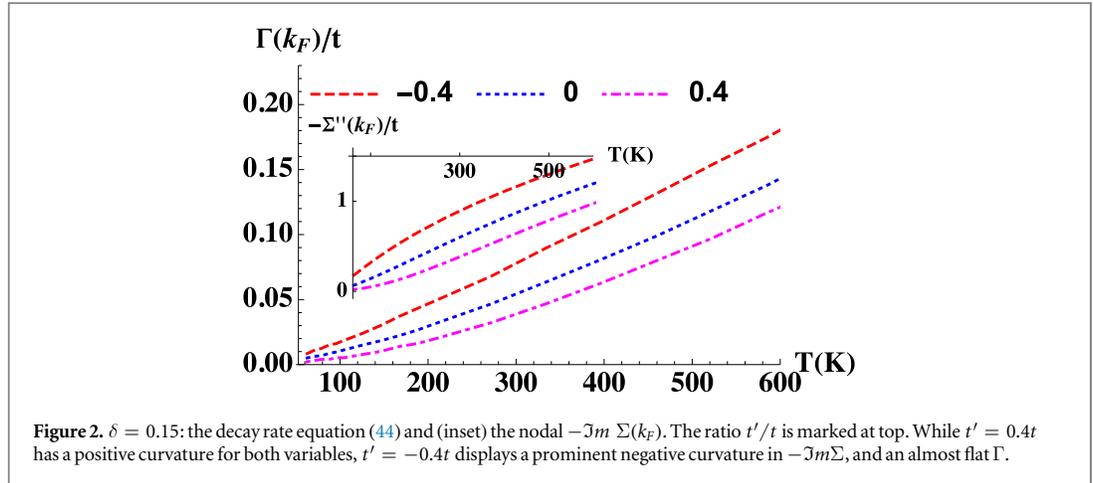
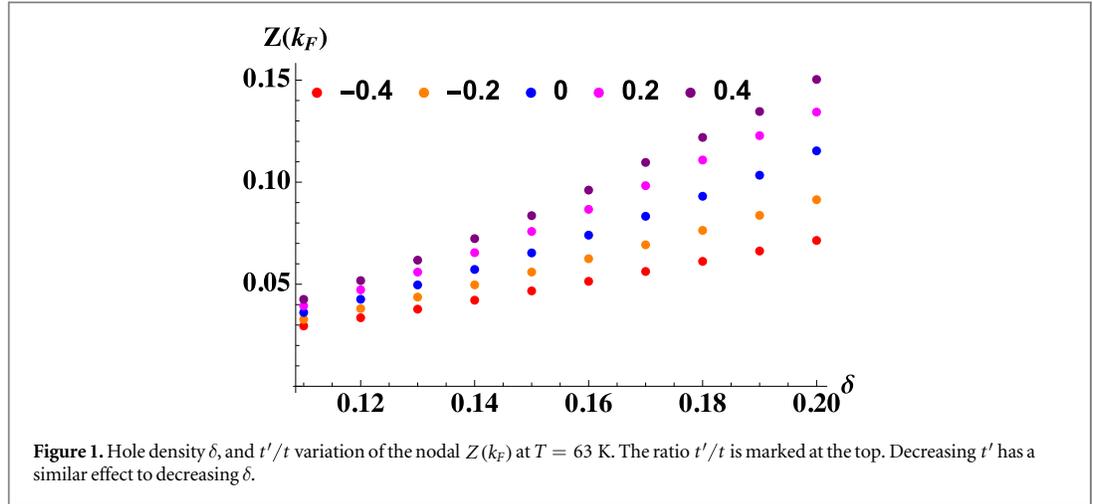
$$\sum_{\mathbf{k}} \int \hat{\rho}_{\mathbf{g}}(\mathbf{k}, \omega) f(\omega) d\omega = \frac{n}{2} = \sum_{\mathbf{k}} \int \rho_{\mathbf{G}}(\mathbf{k}, \omega) f(\omega) d\omega. \quad (43)$$

The set of equations equations (38)–(43) constitute the final set of equations to be computed. These are valid in any dimension, and reduce to the ones benchmarked in  $d = \infty$  after setting  $J \rightarrow 0$  [5, 7].

After analytically continuing  $i\omega_n \rightarrow \omega + i0^+$  we determine the spectral function of the interacting electron spectral function  $\rho_{\mathbf{G}}(\vec{k}, \omega) = -\frac{1}{\pi} \mathcal{I}m \mathcal{G}(\vec{k}, \omega)$ . The set of equations (1)–(5) was solved iteratively on  $L \times L$  lattices with  $L = 19, 37, 61$  and a frequency grid with  $N_{\omega} = 2^{14}, 2^{16}$  points. We find that  $L = 61$  produces the most accurate results at low temperature, while different  $L$  do not make a difference at high temperature. Also,  $N_{\omega} = 2^{14}, 2^{16}$  lead to the same result in the relevant range of parameters. Therefore, we only display the result computed at  $L = 61$  and  $N_{\omega} = 2^{14}$  in this paper. Other details are essentially the same as in our recent study of the  $d = \infty, J \rightarrow 0$  case in [5, 7].

## 3. Results

*Band Parameters:* The  $t$ - $J$  model is studied on the square lattice with hopping parameters  $t$  and  $t'$  for first and second neighbors. The hopping parameter  $t > 0$ , while  $t'$  is varied between  $-0.4t$  and  $0.4t$ , thereby changing



the Fermi surface (FS) from hole-like to electron-like. Parameters relevant to cuprate High  $T_c$  materials are summarized in [2, 11]<sup>4</sup>. Following [2] we assume  $t \sim 0.45$  eV, giving a bandwidth  $\sim 3.6$  eV.

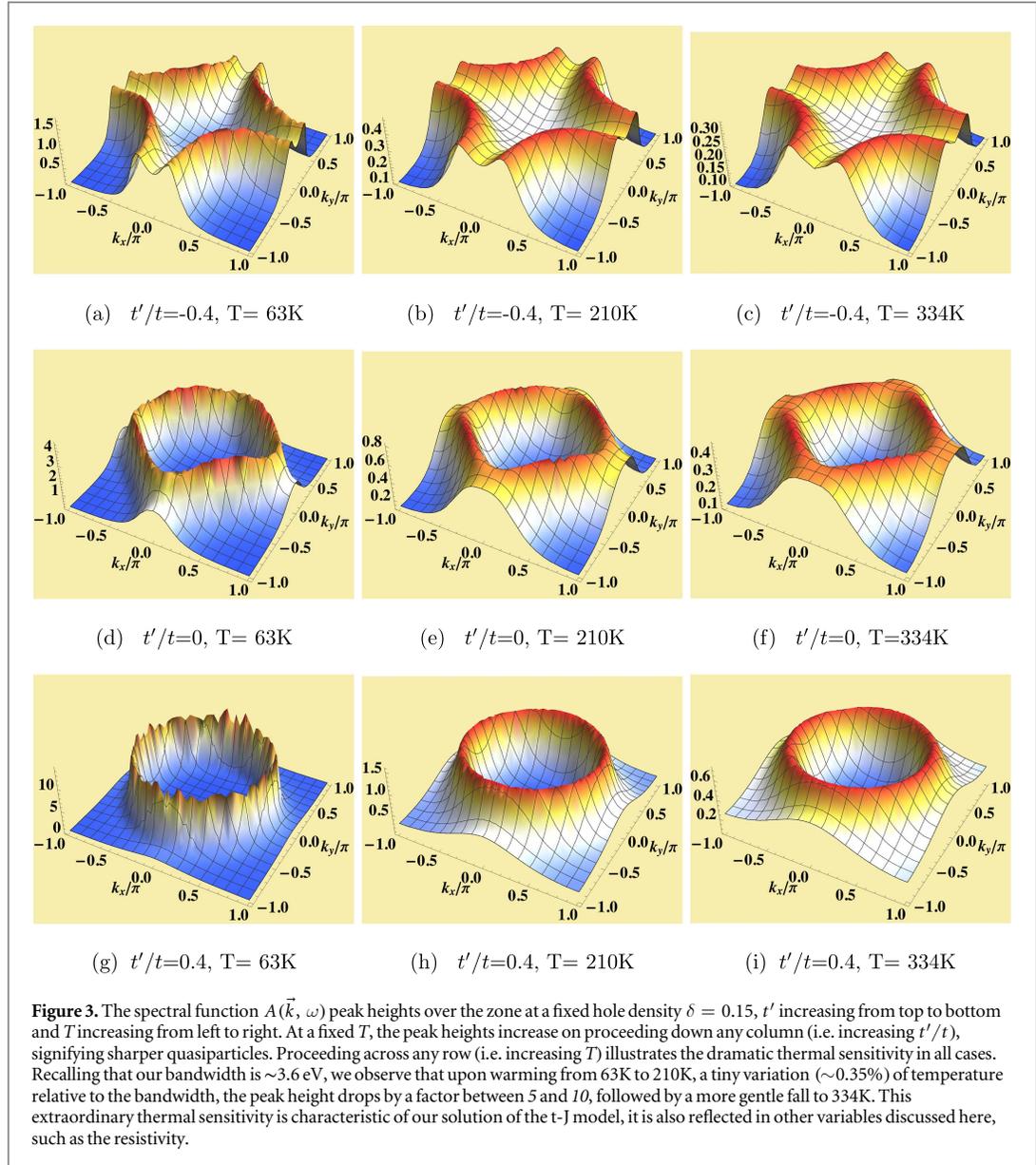
*Single-particle spectrum:* The quasiparticle energy  $E(\vec{k})$  and quasiparticle weight  $Z(\vec{k})$  are found from  $\mathcal{G}$  as usual [5]. In figure 1 we display the hole density  $\delta$  and  $t'/t$  dependence of the low temperature  $Z(k_F)$ , along the nodal (i.e.  $\langle 11 \rangle$ ) direction. The typical magnitudes of  $Z$  are comparable or lower than those reported in  $d = \infty$  [5]. A new and important feature is the strong sensitivity of  $Z(k_F)$  to the sign and magnitude of  $t'/t$ . Both decreasing  $t'$  (at fixed  $\delta$ ) and decreasing  $\delta$  (at fixed  $t'$ ) reduce  $Z$ . This feature is basic to understanding our main results. We next study the decay rate of the electrons

$$\Gamma(\vec{k}) = -Z(\vec{k}) \times \Im \Sigma(\vec{k}, E(\vec{k})), \quad (44)$$

found as the half-width at half-maximum of the spectral function  $\rho_G(\vec{k}, \omega)$  at fixed  $\vec{k}$ . We display the  $T$  variation of  $\Gamma$  and  $-\Im \Sigma$  at the Fermi surface for three representative values of  $t'/t$  in figure 2. Both variables display considerable variation with modest change of  $T$ . The case of  $t' > 0$  shows a distinct quadratic  $T$  dependence, but for  $t' \leq 0$  we note the strong reduction, or absence, of such a quadratic dependence. Below we note a closely parallel  $T$  and  $t'$  dependence of the resistivity.

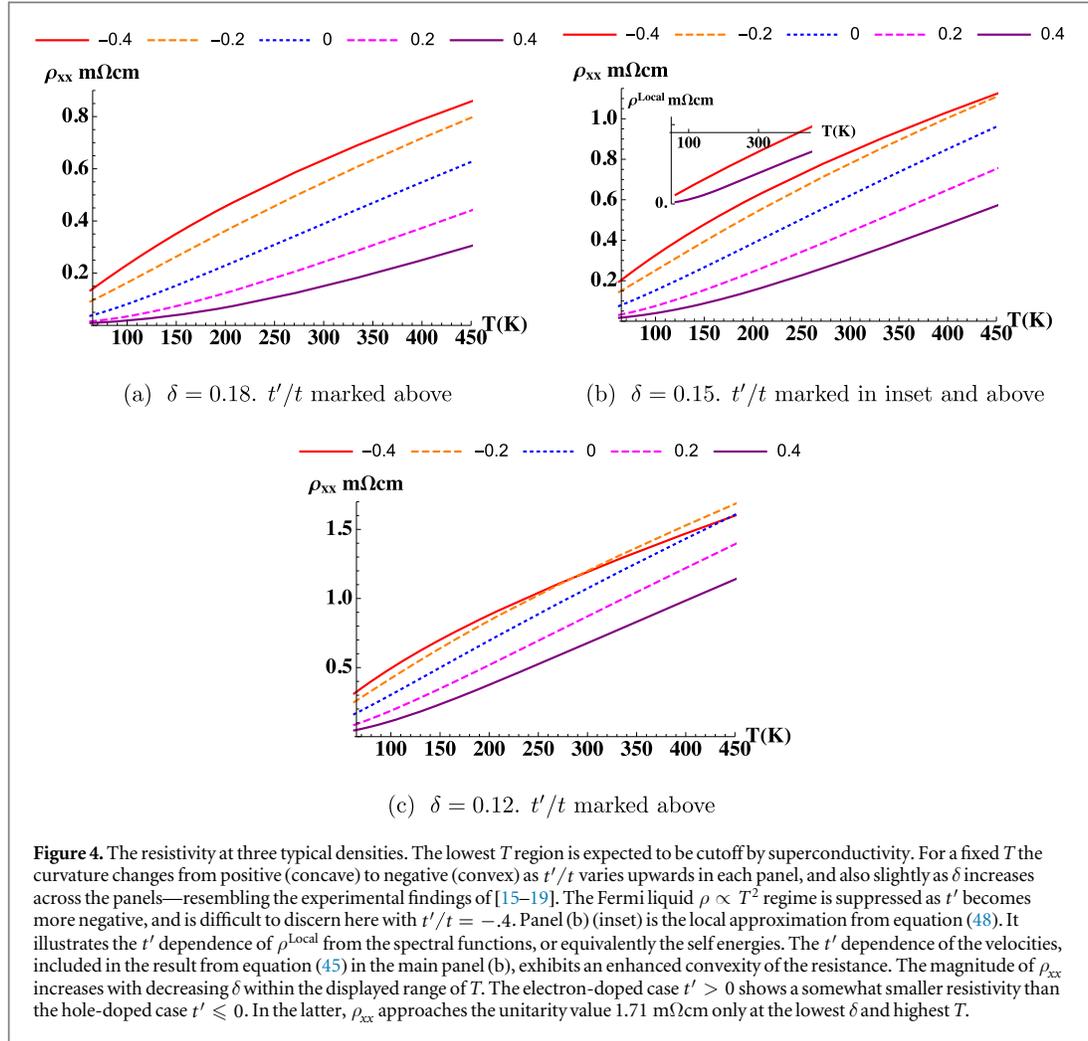
In figure 3 we display the photoemission accessible peak heights of the spectral function  $\{t \star \rho_G(\vec{k}, \omega)\}_{\max}$  over the BZ at three representative values of  $t'/t$ , at three temperatures  $T = 63, 210, 334$  K. The peaks locate the interacting Fermi surface and its thermal sensitivity. The Fermi surface closely tracks the non-interacting FS, changing from hole-like in panels (a)–(c) to strongly electron-like in panels (g)–(i). This implies that the momentum dependence of the Dyson self energy is mild. In contrast a strong momentum dependence would distort the Fermi surface shape significantly—while retaining the Luttinger–Ward volume. Several features are

<sup>4</sup>In high  $T_c$  systems [2] estimate  $t' \lesssim -0.27t$  for BSCCO, while for LSCO  $t' \sim -0.16t$ . NCCO is modeled with  $t' > 0$  after invoking a particle hole transformation. In this case we must flip the sign of the calculated  $R_H$  and  $\Theta_H$  to compare with data.



noteworthy. The peaks are higher in the nodal relative to the anti-nodal direction at low  $T$ . We observe the high sensitivity to warming, in going from  $T = 63$  K to  $T = 334$  K a small ( $\sim 0.7\%$ ) change in  $T$  relative to the bandwidth causes a five to fifteen-fold drop in the spectral peaks at the Fermi surface. This is correlated to the thermal variation of  $-\Im m \Sigma$  at the same set of  $t'$ , shown in the inset of figure 2, since the intensity at  $k_F$  is essentially the inverse of this object. Meanwhile the background spectral weight rises rapidly in all cases, to a roughly similar magnitude. The figure shows that at low  $T$  the curve with  $t' > 0$  has much higher peaks than  $t' \leq 0$ , giving the impression of weaker correlations. However the drop on warming is the largest in this case, which signifies another facet of strong correlations. The heights of the spectral peak and that of the background is predicted quantitatively in figure 3. Their ratio is straightforward to measure in angle resolved photo emission studies, and we suggest it should be interesting to closely study this ratio experimentally, as a check of the theoretical understanding of the temperature dependence of  $\Im m \Sigma(k_F)$  and the decay rate  $\Gamma(k_F)$ .

*Resistivity:* We now study the behavior of the resistivity from electron–electron scattering. We use the popular bubble approximation, factoring the current correlator as  $\langle J(t)J(0) \rangle \sim \sum_k v_k^2 \mathcal{G}^2(k)$ , where the bare current vertex is the velocity  $\tilde{v}_k^\alpha = \frac{\partial \epsilon_k}{\partial k_\alpha}$ . In tight binding theory  $v_k^\alpha$  oscillates in sign, resulting in a vanishing average over the Brillouin zone. This oscillation is expected to reduce magnitude of the vertex corrections [12]. For a 3d metal having well separated sheets in the  $c$  direction ( $c_0$  the separation of the sheets), with each sheet represented by the 2d  $t$ - $J$  model, the DC resistivity  $\rho_{xx}$  can be written in terms of the electron spectral function as follows. We define a dimensionless resistivity  $\tilde{\rho}_{xx}$  whose inverse is given by



$$\bar{\sigma}_{xx} = \langle \Upsilon(\vec{k}) (\hbar v_k^x)^2 / a_0^2 \rangle_k, \quad (45)$$

where  $\langle A \rangle_k \equiv \frac{1}{N_k} \sum_{\vec{k}} A(\vec{k})$ , while the momentum resolved relaxation scale is:

$$\Upsilon(\vec{k}) = (2\pi)^2 \int_{-\infty}^{\infty} d\omega (-\partial f / \partial \omega) \rho_G^2(\vec{k}, \omega), \quad (46)$$

and  $f \equiv 1/(1 + \exp \beta\omega)$  is the Fermi function. This object resembles the spectral peaks in figure 3, losing height and broadening rapidly with  $T$ . The physical 3d resistivity is given by  $\rho_{xx} = \rho_0 \times \bar{\rho}_{xx}$ , where  $\rho_0 \equiv c_0 \hbar / e^2$  ( $\sim 1.71$  mΩ cm) serves as the scale of resistivity<sup>5</sup>, and using the measured values of the lattice constants we can express our results in absolute units. For understanding the magnitude of the inelastic scattering it can be useful to convert the resistivity into the dimensionless parameter  $\langle k_F \rangle \lambda_m$  of an effective 2d continuum theory, where  $\lambda_m$  is the mean-free-path and where  $\langle k_F \rangle$  is an (angle averaged) effective Fermi momentum. We can use a relation argued for in [13, 14]

$$\langle k_F \rangle \lambda_m = \frac{\hbar c_0}{e^2 \rho_{xx}} = \frac{1}{\bar{\rho}_{xx}}. \quad (47)$$

In [13, 14] (see footnote 6) the authors note that in a metallic system this parameter is expected to be greater than unity, and its least value is  $\langle k_F \rangle \lambda_m = 1$  for the case of unitary (impurity) scattering. Thus we expect that  $\rho_{xx} \leq \rho_0$ , i.e.  $\bar{\rho}_{xx} \leq 1$  in a good metal. The Ioffe–Regel–Mott resistivity scale used in [5, 7, 8] provides a similar measure for quantifying the magnitudes of the resistivity found in strongly correlated metals. However we should keep in mind that both estimates suffer from ambiguities in defining a precise threshold value of resistivity, since factors of 2 (or of  $2\pi$ ) cannot be ruled out in equation (47).

<sup>5</sup> The numerics assume a bct unit cell ( $a, a, c$ ) with  $a = 3.79 \text{ \AA}$  and  $c = 13.29 \text{ \AA}$ . In the expression for  $\rho_0, c_0$  corresponds to the interlayer separation  $c_0 = c/2$ . In equations (50) and (51) we use  $v_0/|e| = .596 \times 10^{-3} \text{ cm}^3 \text{ C}^{-1}$  and  $\Phi_0/\Phi = 1440$  with  $B = 10 \text{ T}$ .

<sup>6</sup> See equation (6) in [13]. The origin of this formula is simple to understand, 3d conductivity is written in terms of the two dimensional density as  $\sigma = \frac{n_{2d} e^2 \tau}{c_0 m}$ , and writing  $n_{2d} = k_F / (2\pi)$  and  $\lambda_m = \tau \hbar k_F / m$  we obtain  $\sigma = e^2 / (\hbar c_0) k_F \lambda_m$ .

Figure 4 shows the temperature dependence of the resistance at three densities, and their strong variation with  $t'/t$ .  $J$  is taken as 900 K here, varying  $J$  between 0 and 1500 K makes almost no difference at these temperatures. We see that the scale of the resistivity for  $t' \leq 0$  exceeds the (approximately estimated) unitarity value 1.71 m $\Omega$  cm for the lowest  $\delta$  and the highest displayed  $T$ . On the other hand  $t' > 0$  shows a considerably smaller resistivity at most densities.

In all curves we see that the curvature changes from positive (for  $t' \geq 0$ ) to negative (for  $t' < 0$ ) at say 150 K. To understand the role of  $t'/t$  we note that the resistivity in equation (45) depends on  $t'/t$  through the velocity  $v_k^x$ , in addition to a dependence through the self energies equations (38), (39). To gauge their relative importance it is useful to examine a local approximation of equation (45) where the two functions are averaged separately over momentum:

$$\bar{\sigma}_{xx}^{\text{local}} = \langle \Upsilon(\vec{k}) \rangle_k \times \langle (\hbar v_k^x)^2 / a_0^2 \rangle_k. \quad (48)$$

The velocity squared average is independent of the sign of  $t'$ , therefore the local approximation, shown in figure 4, inset of panel (b), probes only the dependence through equations (38), (39). Comparing the inset and main figure in panel (b), we see that at  $t' = 0.4t$  both resistivity curves display a positive curvature. At  $t' = -0.4t$  we see that  $\rho^{\text{Local}}$  is essentially linear in  $T$ , while  $\rho_{xx}$  shows a negative curvature. The difference is therefore related to the velocity factors, which are very different effect between  $t' < 0$  and  $t' > 0$ . These cause the integrals to have very different thermal variation.

*Hall response:* Within the bubble scheme, we may also calculate the Hall conductivity [9, 20, 21] as

$\sigma_{xy} = -2\pi^2 / \rho_0 \times \left( \frac{\Phi}{\Phi_0} \right) \times \bar{\sigma}_{xy}$ , the dimensionless conductivity:

$$\bar{\sigma}_{xy} = \frac{4\pi^2}{3} \int_{-\infty}^{\infty} d\omega (-\partial f / \partial \omega) \langle \rho_G^3(k, \omega) \eta(k) \rangle_k \quad (49)$$

and  $\eta(k) = \frac{\hbar^2}{a_0^4} \left\{ (v_k^x)^2 \frac{\partial^2 \varepsilon_k}{\partial k_y^2} - (v_k^x v_k^y) \frac{\partial^2 \varepsilon_k}{\partial k_x \partial k_y} \right\}$ . Here  $\Phi = Ba_0^2$  is the flux (see footnote 5) and  $\Phi_0 = hc/(2|e|)$  is the flux quantum. In terms of these we can compute the Hall constant  $R_H$  and Hall angle  $\Theta_H$  from

$$c R_H = -\frac{4\pi^2 v_0}{|e|} \bar{\sigma}_{xy} \times \bar{\rho}_{xx}^2, \quad (50)$$

$$\cot(\Theta_H) = -\frac{1}{2\pi^2} \frac{\bar{\sigma}_{xx}}{\bar{\sigma}_{xy}} \times \frac{\Phi_0}{\Phi}, \quad (51)$$

with  $v_0 = (a_0^2 c_0)$  (see footnote 5). We also define the Hall number as

$$\begin{aligned} n_H &= -\frac{v_0}{|e|cR_H} \\ &= \frac{1}{4\pi^2} \frac{\bar{\sigma}_{xx}^2}{\bar{\sigma}_{xy}}, \end{aligned} \quad (52)$$

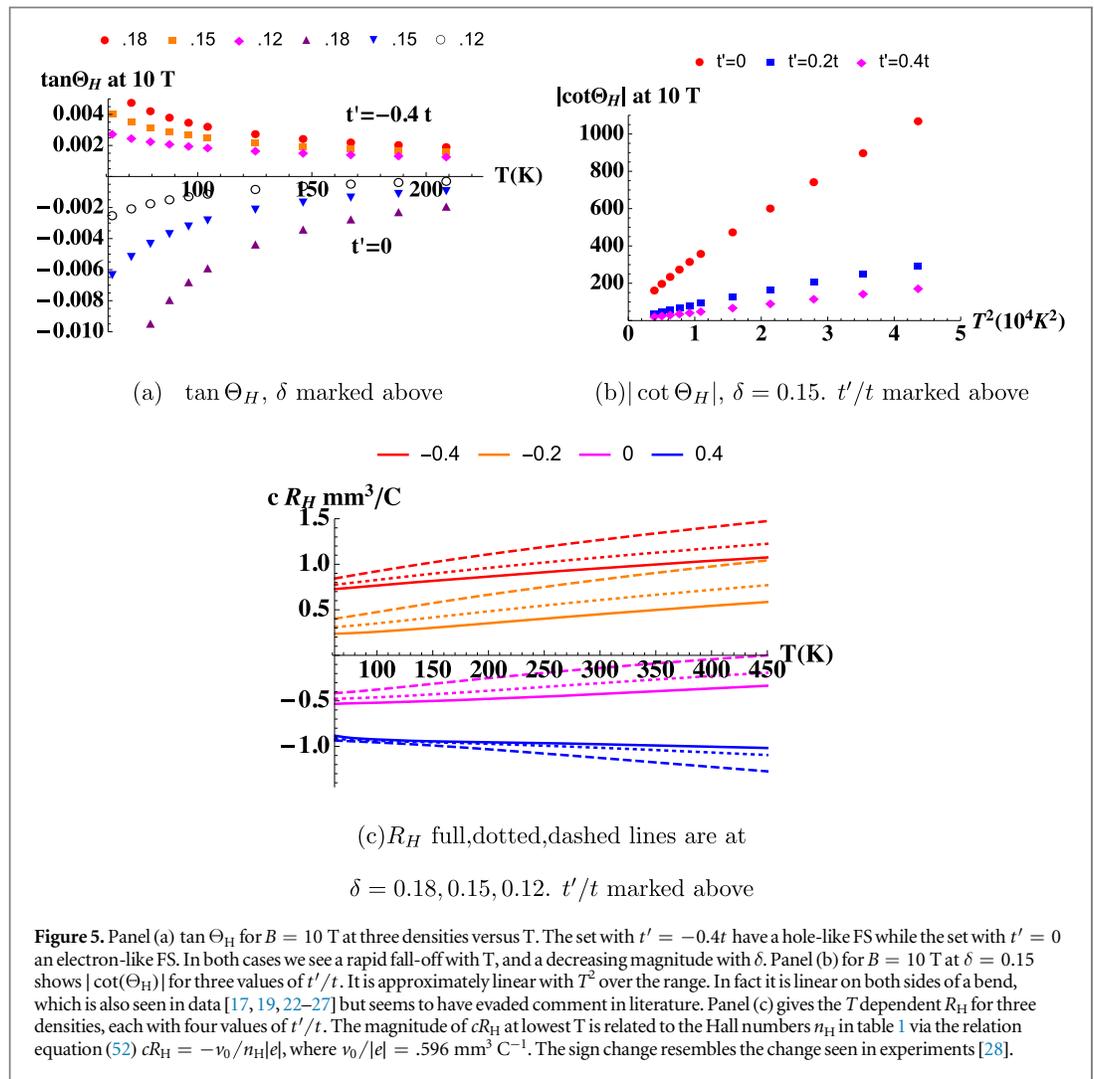
where the definition ensures that in the limit of a circular Fermi surface without interactions,  $n_H \rightarrow n$ , where  $n$  is the number of electrons per copper in the effective single band  $t$ - $J$  model (see footnote 8). The tight binding Fermi surface in the presence of  $t'$  is not circular, but rather resembles the surfaces shown in figure 3. Under these conditions we can evaluate the conductivities in equations (45) and (49), using an approximate Lorentzian spectral function peaked at the bare Fermi surface<sup>7</sup>: The resulting ‘bare’ Hall number  $n_H^0$  contains the corrections due to the curvature of the Fermi surface, including the change in sign in going from open to close surfaces as  $t'$  becomes negative. It is therefore helpful to compare our computed Hall number  $n_H$ , containing the effects of interactions and a complicated scattering rate  $\Gamma$ , with  $n_H^0$  containing only the band effects. This helps us to gauge the effects of interactions, left out in the formula for  $n_H^0$ . The computed  $n_H$ , bare Hall number  $n_H^0$  and their ratio  $n_H/n_H^0$  are shown in table 1. We see that strong correlation generally suppresses  $n_H$  from the bare value  $n_H^0$ , in some cases by as much as 40%. The  $n_H^0$  itself differs from  $n$  quite substantially, depending on doping, temperature and  $t'$ , and therefore one must exercise great care in extracting carrier densities from Hall numbers.

In figure 5 we display the computed Hall variables. In panel (a)  $\tan \Theta_H$  is shown for two values of  $t'/t$  displaying hole-like and electron-like behavior. A decrease in hole density reduces the magnitude in either case. In panel (b) we display the computed  $\cot(\Theta_H)$  versus  $T^2$  with three values of  $t'$  giving an electron-like FS. We note that  $\cot \Theta_H$  is approximately linear with  $T^2$  [22–24] and is strongly affected by the magnitude of  $t'$ . The two distinct  $\cot(\Theta_H) \propto T^2$  regimes seen in figure 5(b) are also seen in many experiments but seem to have evaded attention so far. In [24] it is noted that the bending temperature corresponds to a crossover from the Fermi liquid

<sup>7</sup> These definitions lead to intuitive results in simple cases. For 2-d electrons with  $\varepsilon_k = \hbar^2 k^2 / (2m)$ , and a Lorentzian  $\rho_G(k, \omega) = \frac{1}{\pi} \frac{\Gamma_0}{\Gamma_0^2 + (\omega + \varepsilon_F - \varepsilon_k)^2}$ , upon setting the width  $\Gamma_0 \rightarrow 0$ , we recover the Drude result  $\sigma_{xx} = nq_e^2 \tau / m$  and  $\bar{\sigma}_{xy} = n \left( \frac{\hbar^2}{2ma_0^2 \Gamma_0} \right)$  with  $\tau = \hbar / (2\Gamma_0)$ . Thus  $|e|R_H c / v_0 = -1/n$  in equation (50), and  $\cot(\Theta_H) = -1/(\omega_c \tau)$  in equation (51) where  $\omega_c \equiv |e|B/(mc)$ . If on the other hand we take  $\varepsilon_k$  from the Fourier transform of  $-t_{ij}$ , the limit  $\Gamma_0 \rightarrow 0$  yields the ‘bare’ Hall number  $n_H^0$ , incorporating the effects of a non-circular band structure.

**Table 1.** The Hall number  $n_H$  from equation (52), the bare Hall number  $n_H^0$ , and their ratio  $n_H/n_H^0$ .  $n_H$  is computed at lattice size  $61 \times 61$  at  $T = 63$  K and  $n_H^0$  is computed in the bare band with lattice size  $4000 \times 4000$  at  $T = 0$ . Note that the ratio  $n_H/n_H^0$  varies from .6 to .8. This substantial correction is due to strong correlations. Therefore the inverse problem of deducing the carrier density  $n$  from the Hall number  $n_H$  is quite complex. Finite temperature effects make this even more complicated, as seen in figure 5.

Electron density	Hall number	$t' = -0.4$	$t' = -0.2$	$t' = 0$	$t' = 0.2$	$t' = 0.4$
$n = 0.82$	$n_H$	-0.819	-2.514	1.119	0.679	0.675
	$n_H^0$	-1.163	-3.389	1.51	0.879	0.823
	$n_H/n_H^0$	0.704	0.742	0.741	0.773	0.82
$n = 0.85$	$n_H$	-.768	-1.918	1.249	0.67	0.65
	$n_H^0$	-1.137	-2.448	1.774	0.927	0.855
	$n_H/n_H^0$	0.676	0.783	0.704	0.722	0.76
$n = 0.88$	$n_H$	-.706	-1.479	1.436	0.67	0.637
	$n_H^0$	-1.109	-1.963	2.148	0.977	0.884
	$n_H/n_H^0$	0.637	0.754	0.669	0.686	0.721



to a strange metal, and is therefore of fundamental importance. In panel (c) we show the Hall constant  $R_H$  at three densities for representative values of  $t'/t$ . Its sign is electron-like for  $t' > 0$  and hole-like for  $t' \leq 0$ , tracking the change in topology of the Fermi surface in figure 2. The magnitude of  $R_H$  is substantially affected by changing  $t'$ . This is a strong correlation effect, and discourages envisaging any simple relationship between the Hall number and hole density.

## 4. Discussion

Using the recently developed second order equations of the ECFL theory in [5], we have presented results for the 2d  $t$ - $J$  model at low and intermediate temperatures. In keeping with our recent findings for the  $d = \infty$  solution of the same equations, we note that the quasiparticle weight  $Z(k_F)$  is non-zero, but remarkably small. This fragile FL therefore has an extremely low effective Fermi temperature, above which it displays characteristics of a Gutzwiller correlated strange metal, as listed in [5, 7], including a resistivity that is linear in  $T$ .

By varying  $t'$ , the second neighbor hopping at a fixed  $t$  and  $J$ , we found in figure 1 a remarkable variation of the quasiparticle weight  $Z(k_F)$  that is characteristic of the 2d square lattice, with no simple analog in  $d = \infty$ . We found  $t' < 0$  leads to a considerable reduction in its magnitude, while  $t' > 0$  leads to a larger value and thus a more robust FL. A direct calculation of the single particle spectral width  $\Gamma = -Z\Sigma''$  confirms this observation in figure 2, and when studied as a function of the temperature, shows a much larger magnitude, and hence broader spectral lines.

In figure 3 we present the Fermi surface, as found from the peaks of the spectral function. The shapes of the Fermi surfaces are quite close to those implied by the bare band parameters. This implies that the momentum dependence of the self energy, while non-zero, is fairly small. This also suggests that the vertex corrections, neglected here in 2d, may actually be quite small as well. Two key results concern the spectral heights over the Brillouin zone, and the resistivity as a function of  $T$  at various densities and  $t'$ . The spectral height is the peak value of  $\rho_G(\vec{k}, \omega)$  scanned over  $\omega$ , and equals the inverse of the least magnitude of  $\Im m \Sigma(\vec{k}, \omega)$ . In figure 3 we present both the  $T$  evolution (going horizontally) and the  $t'$  evolution (going vertically) of this important object visible in ARPES. We note that  $t' < 0$  model with a very small  $\Gamma$  also displays a rapid loss of coherence on warming. The quasiparticle peaks drop rapidly, while the valleys, representing the background spectral weight in photoemission, catch up with the peaks in magnitude. A similar variation happens for  $t' = 0$  but the drop of the peak heights is more pronounced. The case of  $t' > 0$  has the largest drop of peak heights, while its effective Fermi temperature is the largest of the three cases. It follows that the electron doped case has a more robust FL appearance for  $T$  lower than its Fermi scale. Our study provides absolute scale values for these observable heights, and it should be interesting to study these experimentally for comparison. Towards that objective we note that  $t' > 0$  maps to the electron doped High  $T_c$  superconductors, while  $t' \leq 0$  maps to the hole doped cases, as we may also deduce from the shapes of the Fermi surfaces in the above figure.

The other key result concerns the resistivity. We are able to calculate the longitudinal resistivity  $\rho_{xx}$  on a doubly absolute scale, both the magnitude of  $\rho_{xx}$  and that of  $T$  are given in physical units by using reasonable values for the basic parameters of the  $t$ - $J$  model and the lattice constants figure 4. We find essentially the experimentally observed scales for both axes, and there is room for further adjustments of bare scales if needed. The main finding is that as  $\delta$  is varied towards half filling, the regime of linear resistivity increases in the hole-like cases ( $t' \leq 0$ ) and the quadratic dependence regime shrinks to very low  $T$  scales, falling below the known superconducting transition temperatures. The other important finding is that the concavity (convexity) of resistance versus  $T$ , usually taken to denote a (non) FL behavior, requires an enlarged viewpoint; we find that the sign of  $t'$  flips the two cases. As an example, the case  $t' \leq 0$  has a pronounced convex regime at low  $T$ . This could be naively ascribed to a non-FL behavior, but in reality is a crossover range to the strange metal regime.

We also present results in figure 5 for the Hall constant and the Hall angle. These are calculated using simple versions of the Kubo formula, found by neglecting the vertex corrections, in the same spirit as the longitudinal resistivity. This approximation is as yet untested against exact results and hence requires some caution. We find that the Hall angle changes sign with  $t'$ . The  $\cot(\Theta_H)$  is found to be roughly linear with  $T^2$ , in agreement with the experimental situation. As noted above, the  $\cot(\Theta_H) \propto T^2$  regime is followed at higher  $T$  by a bend (or kink) corresponding to a crossover from the Fermi liquid to a strange metal, and is therefore important. A similar bend is also seen in many experiments [17, 19, 22–27], but seems to have evaded comment so far. In [24] this kink is discussed further and its connection to the crossover is explained.

In table 1, we present the Hall number  $n_H$  obtained from the Hall constant for various  $n$  and  $t'/t$ . It is compared with the corresponding  $n_H^0$  obtained from a non-interacting theory that incorporates the band-structure effects of changing curvature when  $t' < 0$ . We note that the  $n_H^0$  already captures the changes in sign due to varying  $t'$ . However, there are substantial quantitative corrections even at the lowest  $T$ , originating in strong correlations. This makes the inference of electron density  $n$  from  $n_H$  quite non-trivial; table 1 provides an estimate of the errors involved in this inversion.

## 5. Conclusions

In this work, we employed a recently developed scheme from the ECFL theory where the second order  $\lambda$  expansion terms are supplemented with a high energy cutoff. This scheme has been benchmarked in  $d = \infty$

against DMFT [5, 7] for computing transport and other low energy excitations, giving good agreement with exact numerical results. As detailed in [5] the magnitude of the quasiparticle weight  $Z$  is somewhat lower in this scheme as compared to the exact DMFT values for hole density  $\delta \lesssim 0.8$ . In this work *the same equations* are applied to the 2-d tJ model. While the close agreement found in the  $d = \infty$  case might not guarantee the accuracy of these results in  $d = 2$ , it is plausible that the variations of resistivity and Hall constant, induced by the magnitude and sign of  $t'$  found here, will persist in more exact (future) results. Hence it seems that we can draw some useful conclusions already, especially with regard to the difference between hole and electron doping.

We have shown a range of results for the 2d  $t$ - $J$  model, obtained by varying different parameters within our scheme. It is interesting that the magnitudes of various transport variables, presented here in physical units<sup>8</sup>, are roughly on the scale of reported measurements [15, 25, 16–19, 26, 27]. Although it is not our primary aim here to produce exact fits, we note that the agreement can be improved in many cases with suitable changes of the bare (band) parameters.

In the range of parameters considered here, a metallic state has been posited, and therefore the role of the exchange  $J$  is limited; we find very little variation of the transport quantities with a change in  $J$ . The transport parameter variation with density seems very similar to that found in  $d = \infty$  in [5, 7, 8] where a large variety of Gutzwiller correlated metallic states were shown to arise [7], with their origin in the  $U = \infty$  or Gutzwiller correlation rather than with  $J$ . The added feature in  $d = 2$  is the important role played by  $t'$ , as stressed here. We expect magnetic, superconducting and possibly other broken symmetry states at the lowest  $T$  and  $\delta$  to arise, largely due to the effect of  $J$ . Further work is necessary to find reliable calculational schemes for these broken symmetry states.

A few broad conclusions suggest themselves. The parameter  $t'/t$  plays a key role in determining the low-energy scales. In figure 1 we see that the quasiparticle weight  $Z$  has a large variation with  $t'$ . The origin of this sensitivity lies in the self energies in equations (39), (38), where combinations of the band energies  $\epsilon_k$  play the role of an effective interaction. Varying  $t'/t$  therefore changes the self-energies strongly, in contrast to the usual weak change via the band parameters in equation (37).

Our main findings are as follows. (I) The spectral functions are highly sensitive to thermal variation; in figure 3 we observe a five to fifteen fold drop in intensity with a variation of  $k_B T$  about 1/100th the bandwidth  $\sim 3.6$  eV. This is in severe conflict with expectations from conventional theories of metals. (II) We note from figure 4 that with  $t' \leq 0$ , a FL resistivity  $\rho \propto T^2$  is seen only at very low  $T$ . The very low  $T$  (FL) regime is followed by a 'strange metal' regime, also at low  $T$ , where we find a  $\rho$  versus  $T$  curve, with zero or negative curvature. This regime parallels the Gutzwiller-correlated strange metal regime reported in  $d = \infty$  [7], the negative curvature making it even stranger. (III) For the electron-doped case  $t' > 0$ , figure 4 shows that the curvature is positive and the FL regime extends to higher temperatures.

It is significant that the ECFL theory captures the diametrically opposite resistivity behaviors of hole doped [15–17] and electron doped materials [18, 19] within the same scheme, only differing in the sign of  $t'/t$ . The resistivity curvature mapping of [15] can also be viewed in terms of a variation of this ratio and the temperature, as in figure 4. In conclusion this work provides a sharp picture of the difference made by the second neighbor hopping  $t'$  in the presence of Gutzwiller correlations. It also yields quantitative results for several famously hard to compute variables in correlated matter, that are in rough agreement with a variety of experiments.

## Acknowledgments

We thank Edward Perepelitsky and Sergey Syzranov for helpful comments on the manuscript. The work at UCSC was supported by the US Department of Energy (DOE), Office of Science (BES) under Award # DE-FG02-06ER46319. Computations reported here used the XSEDE Environment [29] (TG-DMR160144) supported by National Science Foundation grant number ACI-1053575.

## References

- [1] Anderson P W 1987 *Science* **235** 1196
- [2] Ogata M and Fukuyama H 2008 *Rep. Prog. Phys.* **71** 036501
- [3] Shastry B S 2011 *Phys. Rev. Lett.* **107** 056403
- [4] Shastry B S 2014 *Ann. Phys.* **343** 164–99  
Shastry B S 2016 *Ann. Phys.* **373** 717–8 (erratum)
- [5] Shastry B S and Perepelitsky E 2016 *Phys. Rev. B* **94** 045138  
Žitko R, Hansen D, Perepelitsky E, Mravlje J, Georges A and Shastry B S 2013 *Phys. Rev. B* **88** 235132  
Shastry B S, Perepelitsky E and Hewson A C 2013 *Phys. Rev. B* **88** 205108

<sup>8</sup> From [5, 7] we may infer that the  $Z$  in the present calculation is a bit too low for  $t' = 0$  and  $.12 \leq \delta \leq .15$ . This is expected to result in overestimating  $\rho_{xx}$  by a factor  $\sim 3$  at  $T = 450$  K.

- [6] Perepelitsky E and Shastry B S 2015 *Ann. Phys.* **357** 1
- [7] Ding W, Žitko R, Mai P, Perepelitsky E and Shastry B S 2017 arXiv:1703.02206v2
- [8] Deng X Y, Mravlje J, Žitko R, Ferrero M, Kotliar G and Georges A 2013 *Phys. Rev. Lett.* **110** 086401
- [9] Xu W, Haule K and Kotliar G 2013 *Phys. Rev. Lett.* **111** 036401
- [10] Shastry B S 2013 *Phys. Rev. B* **87** 125124
- [11] Markiewicz R S, Sahrakorpi S, Lindroos M, Lin H and Bansil A 2005 *Phys. Rev. B* **72** 054519
- [12] Shastry B S and Shraiman B 1990 *Phys. Rev. Lett.* **65** 1068
- [13] Ando Y 2007 *High Tc Superconductors and Related Transition Metal Oxides* ed A Bussman-Holder and H Keller (Berlin: Springer)
- [14] Ando Y, Boebinger G S, Passner A, Kimura T and Kishio K 1995 *Phys. Rev. Lett.* **74** 3253
- [15] Ando Y, Komiya S, Segawa K, Ono S and Kurita Y 2004 *Phys. Rev. Lett.* **93** 267001
- [16] Martin S, Fiory A T, Fleming R M, Schneemeyer L F and Waszczak J V 1988 *Phys. Rev. B* **60** 2194
- [17] Takagi H, Ido T, Ishibashi S, Uota M, Uchida S and Tokura Y 1989 *Phys. Rev. B* **40** 2254
- [18] Onose Y, Taguchi Y, Ishizaka K and Tokura Y 2004 *Phys. Rev. B* **69** 024504
- [19] Li Y, Tabis W, Yu G, Barišić N and Greven M 2016 *Phys. Rev. Lett.* **117** 197001
- [20] Voruganti P, Golubentsev A and John S 1992 *Phys. Rev. B* **45** 13945  
Fukuyama H, Ebisawa H and Wada Y 1969 *Prog. Theor. Phys.* **42** 494  
Kohno H and Yamada K 1988 *Prog. Theor. Phys.* **80** 623
- [21] Arsenault L-F and Tremblay A M S 2013 *Phys. Rev. B* **88** 205109
- [22] Chien T R, Wang Z Z and Ong N P 1991 *Phys. Rev. Lett.* **67** 2088
- [23] Ong N P and Anderson P W 1997 *Phys. Rev. Lett.* **78** 977
- [24] Ding W, Žitko R and Shastry B S 2017 *Phys. Rev. B* **95** 115153
- [25] Ando Y, Kurita Y, Komiya S, Ono S and Segawa K 2004 *Phys. Rev. Lett.* **92** 197001
- [26] Hwang H Y, Batlogg B, Takagi H, Kao H L, Kwo J, Cava R J, Krajewski J J and Peck W F Jr. 1994 *Phys. Rev. Lett.* **72** 2636
- [27] Balakirev F F, Betts J B, Migliori A, Tsukada I, Ando Y and Boebinger G S 2009 *Phys. Rev. Lett.* **102** 017004
- [28] Takeda J, Nishikawa T and Sato M 1994 *Physica C* **231** 293
- [29] Town J *et al* 2014 XSEDE: accelerating scientific discovery *Comput. Sci. Eng.* **16** 62–74
- [30] Hansen D and Shastry B S 2013 *Phys. Rev. B* **87** 245101

## Strange metal from Gutzwiller correlations in infinite dimensions: Transverse transport, optical response, and rise of two relaxation rates

Wenxin Ding,<sup>1</sup> Rok Žitko,<sup>2,3</sup> and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, California 95060, USA*

<sup>2</sup>*Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

<sup>3</sup>*Faculty for Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia*

(Received 4 May 2017; revised manuscript received 8 September 2017; published 25 September 2017)

Using two approaches to strongly correlated systems, the extremely correlated Fermi liquid theory and the dynamical mean field theory, we compute the transverse transport coefficients, namely, the Hall constants  $R_H$  and Hall angles  $\theta_H$ , and the longitudinal and transverse optical response of the  $U = \infty$  Hubbard model in the limit of infinite dimensions. We focus on two successive low-temperature regimes, the Gutzwiller-correlated Fermi liquid (GCFL) and the Gutzwiller-correlated strange metal (GCSM). We find that the Hall angle  $\cot\theta_H$  is proportional to  $T^2$  in the GCFL regime, while upon warming into the GCSM regime it first passes through a downward bend and then continues as  $T^2$ . Equivalently,  $R_H$  is weakly temperature dependent in the GCFL regime, but becomes strongly temperature dependent in the GCSM regime. Drude peaks are found for both the longitudinal optical conductivity  $\sigma_{xx}(\omega)$  and the optical Hall angles  $\tan\theta_H(\omega)$  below certain characteristic energy scales. By comparing the relaxation rates extracted from fitting to the Drude formula, we find that in the GCFL regime there is a single relaxation rate controlling both longitudinal and transverse transport, while in the GCSM regime two different relaxation rates emerge. We trace the origin of this behavior to the dynamical particle-hole asymmetry of the Dyson self-energy, arguably a generic feature of doped Mott insulators.

DOI: [10.1103/PhysRevB.96.115153](https://doi.org/10.1103/PhysRevB.96.115153)

### I. INTRODUCTION

In a recent study [1] we have presented results for the longitudinal resistivity and low-temperature thermodynamics of the Hubbard model (with the repulsion parameter  $U = \infty$ ) in the infinite-dimensional limit. In this limit, we can obtain the complete single-particle Green's functions using two methods: the dynamic mean field theory (DMFT) [2–5] and the extremely correlated Fermi liquid (ECFL) theory [6,7], with some overlapping results and comparisons in Ref. [8]. These studies capture the nonperturbative local Gutzwiller correlation effects on the longitudinal resistivity  $\rho_{xx}$  quantitatively [4–6]. A recent study by our group addresses the physically relevant case of two dimensions [9], with important results for many variables discussed here.

The present work extends the study of Ref. [1], using the ECFL scheme of Ref. [6], to the case of the Hall conductivity  $\sigma_{xy}$  and the finite-frequency (i.e., optical) conductivities. One goal is to further test ECFL with the exact DMFT results for these quantities which are more challenging to calculate. More importantly, however, by combining the various calculated conductivities we are able to uncover the emergence of two different transport relaxation times. In cuprate superconductors, various authors [10–14] have commented on the different temperature ( $T$ ) dependence of the transport properties in the normal phase. The cotangent Hall angles, defined as the ratio of the longitudinal conductivity  $\sigma_{xx}$  and the Hall conductivity,  $\cot(\theta_H) = \sigma_{xx}/\sigma_{xy}$ , is close to quadratic as in conventional metals. Meanwhile, the longitudinal resistivity has unusual linear temperature dependence [15]. Understanding the ubiquitous  $T^2$  behavior of  $\cot(\theta_H)$  in spite of the unconventional temperature dependence of the longitudinal resistivity is therefore quite important.

In Ref. [1] we found that at the lowest temperatures the system is a *Gutzwiller-correlated Fermi liquid* (GCFL) with

$\rho_{xx} \propto T^2$ . Upon warming one finds a regime with linear temperature dependence of the resistivity  $\rho_{xx}$  [1], which is reminiscent of the *strange metal* regime in the cuprate phase diagrams [15]. It is termed the *Gutzwiller-correlated strange metal* (GCSM) regime [1]. Previous studies [4,5] established the GCFL and GCSM regimes using the longitudinal resistivity. Here we focus instead on the Hall constants  $R_H = \sigma_{xy}/\sigma_{xx}^2$  and the Hall angles [5], as well as on the optical conductivity [4] and optical Hall angles. In the GCFL regime, the primary excitations are coherent quasiparticles that survive the Gutzwiller correlation, and there is a single transport relaxation time, as one would expect for a conventional Fermi liquid. Upon warming up into the GCSM regime, the longitudinal and transverse optical scattering rates become different. It appears that the existence of two separate scattering times is a generic characteristic of the GCSM regime.

This work is organized as follows. First we summarize the Kubo formulas used to calculate the transport coefficients in Sec. II. We then revisit in Sec. III the familiar Boltzmann transport theory from which two separate relaxation times can be naturally derived. The results for the dc transport properties are presented in Sec. IV and those of optical conductivities in Sec. V. In Sec. VI we interpret the two scattering times found in the GCSM regime through the particle-hole asymmetry of dynamical properties (spectral function) of the system. In conclusion we discuss the implication of this work for strongly correlated matter.

### II. KUBO FORMULAS

The transport properties of correlated materials can be easily evaluated in the limit of infinite dimensions because the vertex corrections are absent [16]. For dimensions  $d > 3$ , the longitudinal conductivity  $\sigma_{xx}$  is straightforwardly generalized

as the electric field remains a  $d$ -dimensional vector. The generalization is less clear for the transverse conductivity and Hall constants, because the magnetic field is no longer a vector but rather a rank-2 tensor defined through the electromagnetic tensor. Nevertheless,  $\sigma_{xy}$  can still be defined through suitable current-current correlation functions.

The input to the transport calculation is the single-particle Green's function  $G(\omega, \mathbf{k})$ , calculated in the following within either ECFL or DMFT. The Kubo formulas can be written as [17,18]

$$\sigma_{xx} = 2\pi q_e^2 \sum_k \Phi_k^{xx} \int d\omega \left( -\frac{\partial f(\omega)}{\partial \omega} \right) \rho_G^2(\omega, \mathbf{k}), \quad (1)$$

$$\sigma_{xy}/B = \frac{4\pi^2 q_e^3}{3} \sum_k \Phi_k^{xy} \int d\omega \left( -\frac{\partial f(\omega)}{\partial \omega} \right) \rho_G^3(\omega, \mathbf{k}), \quad (2)$$

where  $\rho_G(\omega, \mathbf{k}) = -\text{Im} G(\omega, \mathbf{k})/\pi$  is the single-particle spectral function and  $q_e = -|e|$  is the electron charge.  $\Phi_k^{xx} = (\epsilon_k^x)^2$  and  $\Phi_k^{xy} = (\epsilon_k^y)^2 \epsilon_k^{xx} - \epsilon_k^y \epsilon_k^x \epsilon_k^{xy}$  are called transport functions, with  $\epsilon_k^\alpha = \partial \epsilon_k / \partial k_\alpha$  and  $\epsilon_k^{\alpha\beta} = \partial^2 \epsilon_k / \partial k_\alpha \partial k_\beta$ ,  $\epsilon_k$  being the energy dispersion. We set  $\hbar$  to 1.

It is more convenient to convert the multidimensional  $k$  sums into energy integrals:

$$\sigma_{xx} = \sigma_0 2\pi D \int d\epsilon \frac{\Phi^{xx}(\epsilon)}{\Phi^{xx}(0)} \int d\omega \left( -\frac{\partial f(\omega)}{\partial \omega} \right) \rho_G^2(\omega, \epsilon), \quad (3)$$

$$\begin{aligned} \sigma_{xy}/B &= \sigma_0 \frac{4\pi^2 D q_e}{3} \int d\epsilon \frac{\Phi^{xy}(\epsilon)}{\Phi^{xx}(0)} \\ &\times \int d\omega \left( -\frac{\partial f(\omega)}{\partial \omega} \right) \rho_G^3(\omega, \epsilon), \end{aligned} \quad (4)$$

where  $\Phi^{xx(xy)}(\epsilon) = \sum_k \Phi_k^{xx(xy)} \delta(\epsilon - \epsilon_k)$ ,  $\sigma_0 = q_e^2 \Phi^{xx}(0)/D$  is the Ioffe-Regel-Mott conductivity,  $D$  is the half-bandwidth, and  $\rho_G(\omega, \epsilon) = \rho_G(\omega, \mathbf{k})$  such that  $\epsilon = \epsilon_k$ . In  $d$  dimensions the transport functions on the Bethe lattice are [19]

$$\Phi^{xx}(\epsilon) = \frac{1}{3d} (D^2 - \epsilon^2) \rho_0(\epsilon), \quad (5)$$

$$\Phi^{xy}(\epsilon) = -\frac{1}{3d(d-1)} \epsilon (D^2 - \epsilon^2) \rho_0(\epsilon), \quad (6)$$

where  $\rho_0(\epsilon) = \frac{2}{\pi D^2} \sqrt{D^2 - \epsilon^2} \Theta(D - |\epsilon|)$  is the noninteracting density of states on the Bethe lattice and  $D$  is the half-bandwidth. Even though the transport function results indicate that  $\sigma$  vanishes as  $d \rightarrow \infty$ , we can redefine the conductivities in this limit as the sum of all components:  $\sigma_L = \sum_\alpha \sigma_{\alpha\alpha}$ ,  $\sigma_T = \sum_{\alpha \neq \beta} \text{sgn}[\alpha - \beta] \sigma_{\alpha\beta}$  with  $\alpha(\beta) = 1, 2, \dots, d$ . More importantly, the  $d$  dependence directly drops out when we compute the Hall constant  $R_H = \sigma_{xy}/\sigma_{xx}^2$ . For the rest of this work, we shall redefine  $\sigma_{xx}$  and  $\sigma_{xy}$  via  $\sigma_L$  and  $\sigma_T$  considering that all components of  $\sigma_{L(T)}$  are equal so that both the  $d$  factor and the constant factor drop out from the transport functions:

$$\sigma_{xx} = 3\sigma_L, \quad \Phi^{xx}(\epsilon) = (D^2 - \epsilon^2) \rho_0(\epsilon), \quad (7)$$

$$\sigma_{xy} = 3\sigma_T, \quad \Phi^{xy}(\epsilon) = -\epsilon (D^2 - \epsilon^2) \rho_0(\epsilon). \quad (8)$$

### III. TWO-RELAXATION-TIME BEHAVIOR IN THE BOLTZMANN THEORY

In Boltzmann theory, the transport properties can be obtained by solving for the distribution function in the presence of external fields from the Boltzmann equation [20]:

$$\frac{\partial \delta f}{\partial t} - \frac{q_e}{\hbar c} \mathbf{v} \times \mathbf{B} \cdot \frac{\partial \delta f}{\partial \mathbf{k}} + \mathbf{v} \cdot q_e \mathbf{E}(t) \left( -\frac{\partial f^0}{\partial \epsilon} \right) = \hat{L} \delta f, \quad (9)$$

where  $f$  is the full distribution function that needs to be solved,  $f^0$  is the Fermi-Dirac distribution function,  $\delta f = f - f^0$ , and  $\hat{L} \delta f$  represents the linearized collision integrals.

In the regime of linear response, we expand  $\delta f^{E,B}$  in powers of the external fields to second order as

$$\delta f^{E,B} = \delta f^{E,0} + B \delta f^{E,1}, \quad (10)$$

where  $\delta f^{E,0}$  is the solution in the absence of magnetic fields, and both  $\delta f^{E,0}$  and  $\delta f^{E,1}$  are linear in  $E$ . In the relaxation-time approximation (RTA) [21] we replace the collision integrals as  $\hat{L} \delta f \rightarrow -\delta f/\tau$  where  $\tau$  is assumed to be  $\mathbf{k}$  independent. However,  $\hat{L} \delta f^{E,0}$  and  $\hat{L} \delta f^{E,1}$  are in principle governed by different relaxation times, as pointed out by Anderson [10,14]. Writing

$$\hat{L} \delta f^{E,0} \rightarrow -\frac{\delta f^{E,0}}{\tau_{tr}}, \quad \hat{L} \delta f^{E,1} \rightarrow -\frac{\delta f^{E,1}}{\tau_H}, \quad (11)$$

we obtain

$$\sigma_{xx}(\omega) = \frac{\omega_p^2}{4\pi} \frac{\tau_{tr}}{1 - i\omega\tau_{tr}}, \quad (12)$$

$$\sigma_{xy}(\omega)/B = \frac{\omega_p^2 \omega_c / B}{4\pi^2} \frac{\tau_H}{1 - i\omega\tau_H} \frac{\tau_{tr}}{1 - i\omega\tau_{tr}}, \quad (13)$$

where

$$\frac{\omega_p^2}{4\pi} = \int \frac{d^d k}{(2\pi)^d} 2q_e^2 v_x^2 (-\partial_\epsilon f^0), \quad (14)$$

$$\frac{\omega_c}{B} = \omega_p^{-2} \int \frac{d^d k}{(2\pi)^d} 2q_e^3 (v_x^2 \partial_{k_y} v_y - v_x v_y \partial_{k_x} v_y) \partial_\epsilon f^0; \quad (15)$$

$v_a = \partial_{k_a} \epsilon(\mathbf{k})$  is the velocity in direction  $a$ ,  $\epsilon(\mathbf{k})$  is the energy dispersion of the electrons, and  $\mathbf{B} = \hat{z}B$ . Then the Hall angle is

$$\tan \theta_H(\omega) = \frac{\omega_c}{\pi} \frac{\tau_H}{1 - i\omega\tau_H}. \quad (16)$$

Therefore, the optical conductivities can be cast in the Boltzmann-RTA form as

$$\frac{\sigma_{xx}(0)}{\text{Re}[\sigma_{xx}(\omega)]} = 1 + \omega^2 \tau_{tr}^2, \quad (17)$$

$$\frac{\sigma_{xy}(0)/B}{\text{Re}[\sigma_{xy}(\omega)]/B} = 1 + \omega^2 (\tau_{tr}^2 + \tau_H^2) + \tau_{tr}^2 \tau_H^2 \omega^4, \quad (18)$$

$$\frac{\theta_H(0)}{\text{Re}[\theta_H(\omega)]} = 1 + \omega^2 \tau_H^2. \quad (19)$$

The dc and ac transport coefficients of a microscopic theory do not necessarily take the form of the Boltzmann RTA theory. In the rest of this work, we study both the dc and the real part

of the ac transport coefficients, and consider them as

$$\text{Re}[\sigma_{xx}(\omega)] = \frac{\sigma_{xx}(0)}{1 + \tau_r^2 \omega^2 + O(\omega^4)}, \quad (20)$$

$$\text{Re}[\tan \theta_H(\omega)/B] = \frac{\tan \theta_H(0)/B}{1 + \tau_H^2 \omega^2 + O(\omega^4)}. \quad (21)$$

The relaxation times  $\tau_r$  and  $\tau_H$  are extracted from the low-frequency part of  $\text{Re}[\sigma_{xx}(\omega)]$  and  $\text{Re}[\tan \theta_H(\omega)/B]$  by fitting to the above expressions. Although computing  $\text{Re}[\theta_H(\omega)]$  requires both real and imaginary parts of the optical conductivities, we can make the approximation  $\text{Re}[\theta_H(\omega)] \simeq \text{Re}[\sigma_{xy}(\omega)]/\text{Re}[\sigma_{xx}(\omega)]$  when  $\omega$  of concern is small. We expect  $\tau_r$  and  $\tau_H$  to have temperature and density dependence similar to those of  $\sigma_{xx}(0)$  and  $\tan \theta_H(0)/B$ .

#### IV. DC TRANSPORT

We now use the Kubo formulas to compute the transport coefficients within the ECFL and DMFT approaches. We plot the ECFL results as solid symbols and the DMFT results as dashed lines using the same color for each density unless specified otherwise. As we shall demonstrate, the agreement between the DMFT and ECFL results follows the same qualitative trend for all quantities considered: it is better at lower temperatures, lower frequencies, and lower density (higher hole doping).

We identify the GCFL and GCSM regimes, as well as the crossover scale  $T_{FL}$  separating them, from the  $T$  dependence of the longitudinal resistivity  $\rho_{xx}$ , shown in Fig. 1. Figure 1(a) shows resistivity for all densities and temperatures considered for ECFL (symbols) and DMFT (dashed lines) in this work. In Fig. 1(b), we use the resistivity of  $n = 0.8$  from ECFL as an example of how we determine  $T_{FL}$  with the help of the  $T^2$  fit at low temperatures (blue dashed line) and the linear- $T$  fit from the GCSM regime (purple dashed line). We identify the Fermi liquid temperature  $T_{FL}$  using the resistivity, rather than the more conventional thermodynamic measures, such as heat capacity. The latter variables do actually give rather similar values, but the resistivity seems most appropriate for this study. Our definition is that up to and below  $T_{FL}$ , the resistivity  $\rho_{xx} \sim T^2$ , while above  $T_{FL}$ ,  $\rho_{xx}$  displays a more complex set of  $T$  dependencies as outlined in Ref. [1]. The Fermi liquid temperature has been quantitatively estimated in Ref. [6]:

$$T_{FL} \simeq 0.05 \times DZ \simeq 0.05 \times D\delta^\alpha, \quad (22)$$

where  $\delta$  is the hole density  $\delta = 1 - n$ . The exponent  $\alpha \sim 1.39$  within DMFT [8]. The value of  $\alpha$  determined for ECFL within the current scheme [6] is somewhat larger than that from DMFT. As a consequence  $T_{FL}$  given by DMFT is slightly higher than that by ECFL (see [22] for exact numbers), as can also be seen in Fig. 1. Consequently as  $n$  increases, the ECFL curves for  $\rho_{xx}$  lie above those from DMFT.

##### A. Hall constant

In Fig. 2, we show  $R_H$  as a function of temperature at different densities for low temperatures  $T < 0.02D$  [Fig. 2(a)], and as a function of the hole density  $\delta = (1 - n)$  at  $T = 0.002D, 0.005D, 0.01D$  [Fig. 2(b)]. The Hall constant is

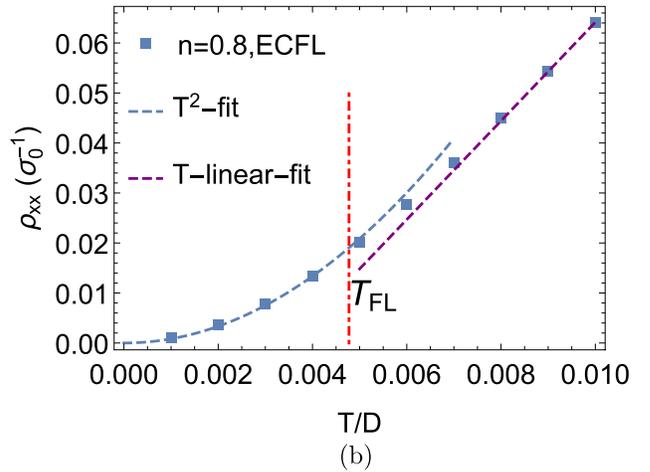
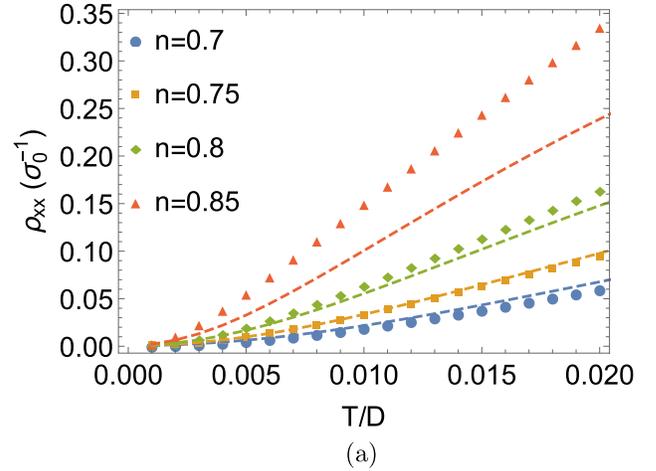


FIG. 1. Temperature dependence of the dc resistivity  $\rho_{xx}$  (a) of the  $U = \infty$  Hubbard model from DMFT (dashed lines) and ECFL (solid symbols) for a range of electron densities  $n$ . The horizontal axis corresponds to absolute temperatures. Using the resistivity of  $n = 0.8$  from ECFL (b) as an example, we show how we determine  $T_{FL}$  with the help of the  $T^2$  fit at low temperatures (blue dashed line) and the linear- $T$  fit from the GCSM regime (purple dashed line).

weakly temperature dependent for  $T \ll T_{FL}$ , but it starts to decrease upon warming, as seen in Fig. 2(a).

As a function of hole density  $\delta$  the Hall constants from the two theories agree quite well, and are roughly linear with  $\delta$ . The extrapolation to  $\delta \rightarrow 0$  is uncertain from the present data. One might be tempted to speculate that it vanishes, since the lattice density of states is particle-hole symmetric. This question deserves further study with different densities of states that break the particle-hole symmetry.

##### B. Cotangent of the Hall angle

The theoretical results for cotangent of the Hall angle ( $\cot \theta_H$ ) $B = (\sigma_{xx}/\sigma_{xy})B$  are shown as a function of  $T^2$  in Fig. 3(a). We see that in DMFT as well as ECFL, the  $\cot(\theta_H)$  is linear in  $T^2$  on both sides of a bend (or kink) temperature, which increases with increasing hole density  $\delta$ . However this kink is weaker in DMFT than in ECFL. This bending was already noted in Fig. 5(a) of Ref. [9], within the 2-d ECFL

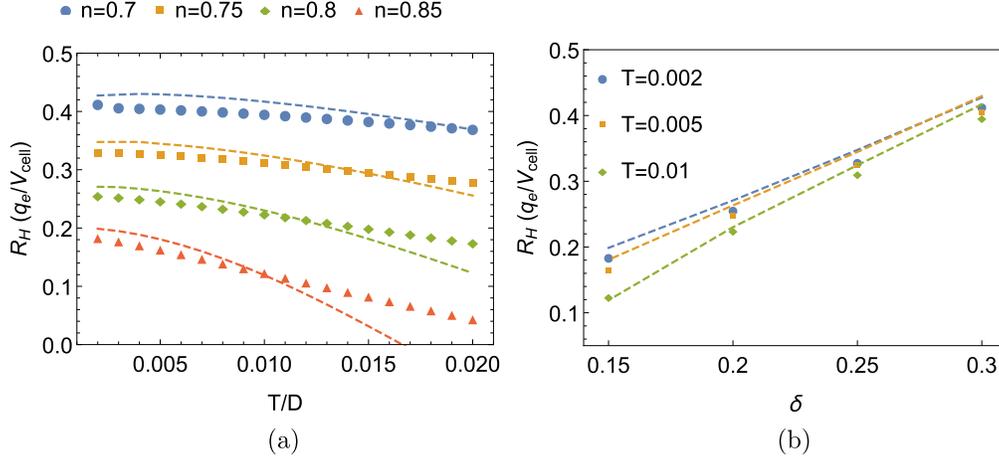


FIG. 2. Temperature dependence of the Hall constants  $R_H$  (a) and  $R_H$  at  $T = 0.002D, 0.005D, 0.01D$  as a function of the hole density  $\delta = 1 - n$  (b) for both DMFT (dashed lines) and ECFL (solid symbols).  $R_H$  is weakly  $T$  dependent below  $T_{FL}$  and develops stronger  $T$  dependence in the GCSM regime.  $R_H$  varies roughly linearly on  $\delta$  at all three temperatures shown in (b).

theory. We may thus infer that  $\cot(\theta_H)$  goes as  $Q_{FL}T^2$  in the Fermi liquid regime, passes through a slight downward bend, and continues as  $Q_{SM}T^2$  in the strange metal regimes, such that  $Q_{FL} > Q_{SM}$ . The difference,  $Q_{FL} - Q_{SM}$ , becomes smaller as  $\delta$  decreases.

In order to characterize this kink more precisely, we define the downward-bending regime by its onset temperature  $T_B^-$ , the crossing temperature of the two different  $T^2$  lines  $T_B$ , and its ending temperature  $T_B^+$ . The temperatures  $T_B^{-(+)}$  are determined by 5% deviation from the  $T^2$  fitting well below (above)  $T_{FL}$ , and  $T_B$  is well defined as the crossing point of the two  $T^2$  fittings. We illustrate the kink and the determination of  $T_B$ ,  $T_B^-$ , and  $T_B^+$  at  $n = 0.7$  for both ECFL in Fig. 3(b) and DMFT in Fig. 3(c). In Fig. 3(d), we show  $T_B$ ,  $T_B^-$ ,  $T_B^+$ , and  $T_{FL}$  obtained from ECFL as functions of  $\delta$ . We see that  $T_B^-$  is identical to  $T_{FL}$ , while  $T_B$  and  $T_B^+$  are  $T_{FL}$  plus some constants with weak  $\delta$  dependence. We plot  $\frac{\cot \theta_H}{\cot \theta_H(T=T_{FL})}$  as functions of  $(T/T_{FL})^2$  for ECFL in Fig. 3(e) and DMFT in Fig. 3(f) to show the systematic evolution of the kinks when the density is varied.

In Fig. 3(a) we note that with  $n \gtrsim 0.8$  the ECFL and DMFT curves separate out at modest  $T^2$ , unlike lower densities where the agreement is over a greater range. This may be ascribed to the limitations of the second-order scheme in ECFL used here, which underestimates  $Z$  at high densities.

### C. Kink in cotangent of the Hall angle

There has been much interest in the quadratic  $T$  dependence of  $\cot(\theta_H)$  in the literature [10,14]. Going beyond the much discussed low- $T$  quadratic correlation, we would like to point out that a bending anomaly, or kink, is seen in the plot of  $\cot(\theta_H)$  versus  $T^2$ , in almost all experiments. A kink is clearly seen in the experimental curves Fig. 2 of Ref. [10], Fig. 4 of Ref. [23], and Fig. 3(c) of Ref. [11]. This intriguing feature and its significance, noted here, seems to have escaped mention earlier.

From Fig. 3(c) of Ref. [11] we estimate  $T_B \simeq 100$  K, 80 K, 70 K for LSCO at  $\delta = 0.21, 0.17, 0.14$ , respec-

tively. These are comparable with the ECFL results  $T_B = 70$  K, 60 K, 40 K at  $\delta = 0.2, 0.175, 0.15$ , if we set  $D = 10^4$  K. The trend of  $T_B$  and the prefactor difference  $Q_{FL} - Q_{SM}$  also agree with what we find; i.e., both  $T_B$  and  $Q_{FL} - Q_{SM}$  decrease as  $\delta$  is lowered. An increase of  $Q_{SM}$  at even higher temperatures is also observed in Ref. [24], similarly to what we find in Fig. 3(a) above the GCSM regime. It is notable that the bending temperatures  $T_B$  in theory and in experiments are on a similar scale.

From the (shared) perspective of the ECFL and DMFT theories Ref. [1], the scale  $T_{FL}$  represents a crossover between the GCFL and GCSM regimes. In the present work, we have argued that the  $\cot(\theta_H)$  versus  $T^2$  curve further shows a bend at temperature  $T_B$ . This bending temperature scale  $T_B$  is related to the effective Fermi liquid scale  $T_{FL}$ . In Figs. 3(b)–3(d) this relationship is made clear through the identification of  $T_B^\pm$  and  $T_B$ . It should be possible to extract these scales from experiments following the same protocol. In view of our clear-cut prediction, it would be of considerable interest to experimentally explore the bending anomaly (i.e., kink) in  $\cot(\theta_H)$  versus  $T^2$  and to test the proposed correlation with  $T_{FL}$ .

## V. OPTICAL RESPONSE

### A. Optical conductivity and the longitudinal scattering rates $\Gamma_{tr}$

In Fig. 4 we show the optical conductivity  $\sigma_{xx}(\omega)$  as well as the quantity  $\sigma_{xx}(0)/\sigma_{xx}(\omega) - 1$ , which better presents the approach to the zero-frequency limit and is to be compared with the Boltzmann RTA form (Drude formula) in Eq. (17). We display plots obtained from both ECFL (symbols) and DMFT (dashed lines) for fixed  $n = 0.8$  and for three temperatures to show the generic behavior at  $T < T_{FL}$ ,  $T \simeq T_{FL}$ , and  $T > T_{FL}$ :  $T = 0.002D$  [Fig. 4(a)],  $T = 0.005D$  [Fig. 4(b)], and  $T = 0.01D$  [Fig. 4(c)]. ECFL results agree well with the exact solution of DMFT within this temperature range.

$\sigma_{xx}(\omega)$  shows a narrow Drude peak below  $T_{FL}$  which broadens as  $T$  increases and finally takes a form well approximated by a broad Lorentzian at  $T = 0.01D$ . Correspondingly,

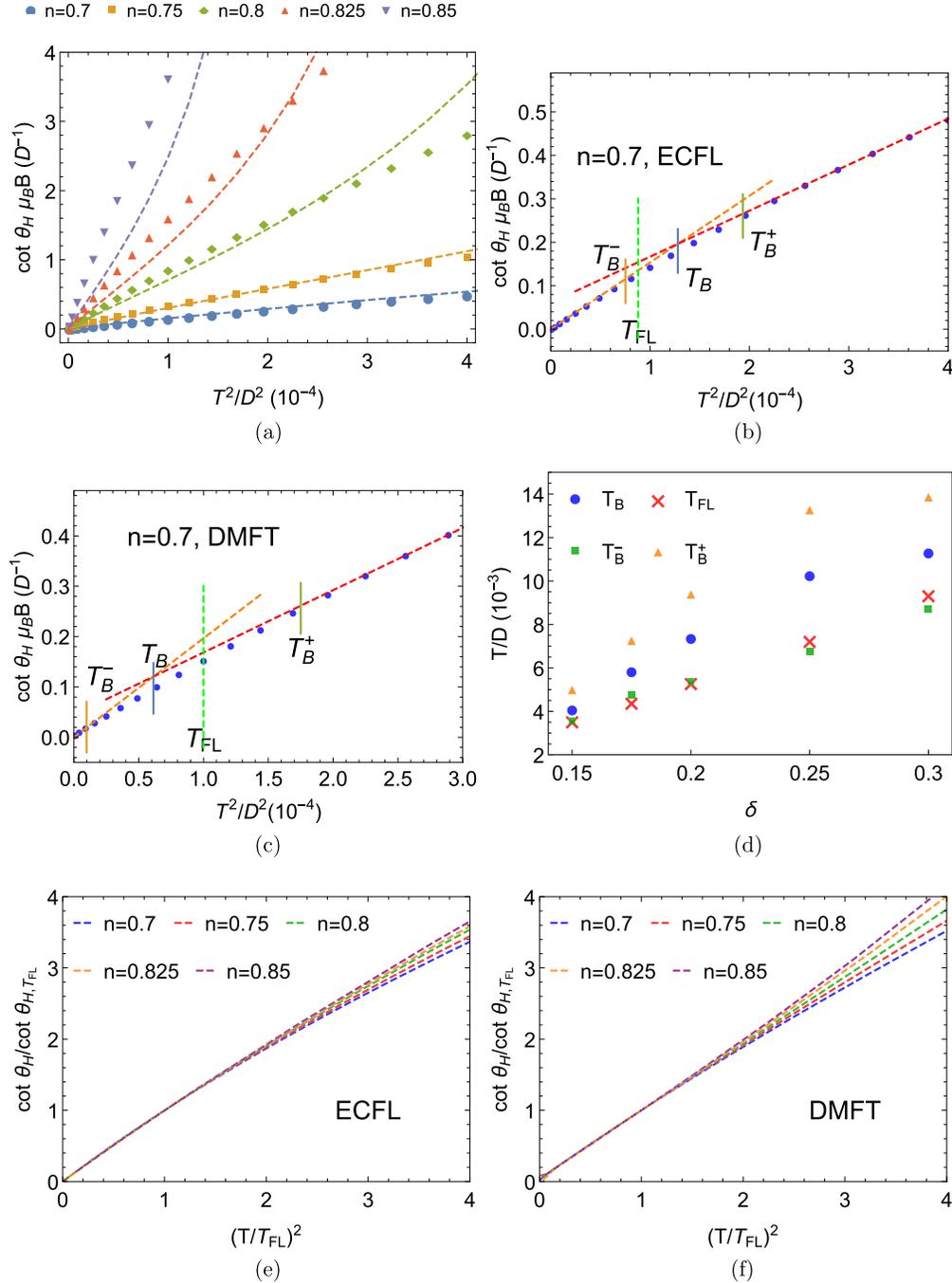


FIG. 3. Temperature dependence of the cotangent Hall angle  $\cot \theta_H B$  of both ECFL (symbols) and DMFT (dashed lines) shown as a function of  $T^2$  (a). The Hall angle  $\cot \theta_H B \propto T^2$  in the GCFL regime passes through a slight downward bend (i.e., a kink) and continues as  $T^2$  within the temperature range studied. The downward bending regime is characterized by its onset  $T_B^-$ , the crossing of the two different  $T^2$  lines  $T_B$ , and its ending  $T_B^+$ . We illustrate the kink and the determination of  $T_B$ ,  $T_B^-$ , and  $T_B^+$  at  $n = 0.7$  for both ECFL (b) and DMFT (c).  $T_B$ ,  $T_B^-$ , and  $T_B^+$  obtained from the ECFL are shown as a function of  $\delta$  in (d). We plot  $\frac{\cot \theta_H}{\cot \theta_H(T=T_{FL})}$  as a function of  $(T/T_{FL})^2$  for ECFL (e) and DMFT (f) to show the systematic evolution of the kinks when the density varies.

$[\sigma_{xx}(0)/\sigma_{xx}(\omega) - 1]$  is quadratic in frequency and can be fitted to  $\tau_{tr}^2 \omega^2$  to extract the relaxation time  $\tau_{tr}$ . The  $\omega^2$  regime has a width  $\propto \tau_{tr}^{-1}$ . The fitting is performed at very small frequencies well within this quadratic regime. At higher frequency,  $[\sigma_{xx}(0)/\sigma_{xx}(\omega) - 1]$  flattens out and creates a kneelike

feature in between. The flattening tendency decreases as  $T$  increases, and  $1/\sigma_{xx}(\omega)$  grows monotonically. This kneelike feature thus becomes smoother as  $T$  increases and eventually is lost for  $T > T_{FL}$ . This trend is illustrated in Fig. 4(d), where we normalize all curves of  $[\sigma_{xx}(0)/\sigma_{xx}(\omega) - 1]$  by their

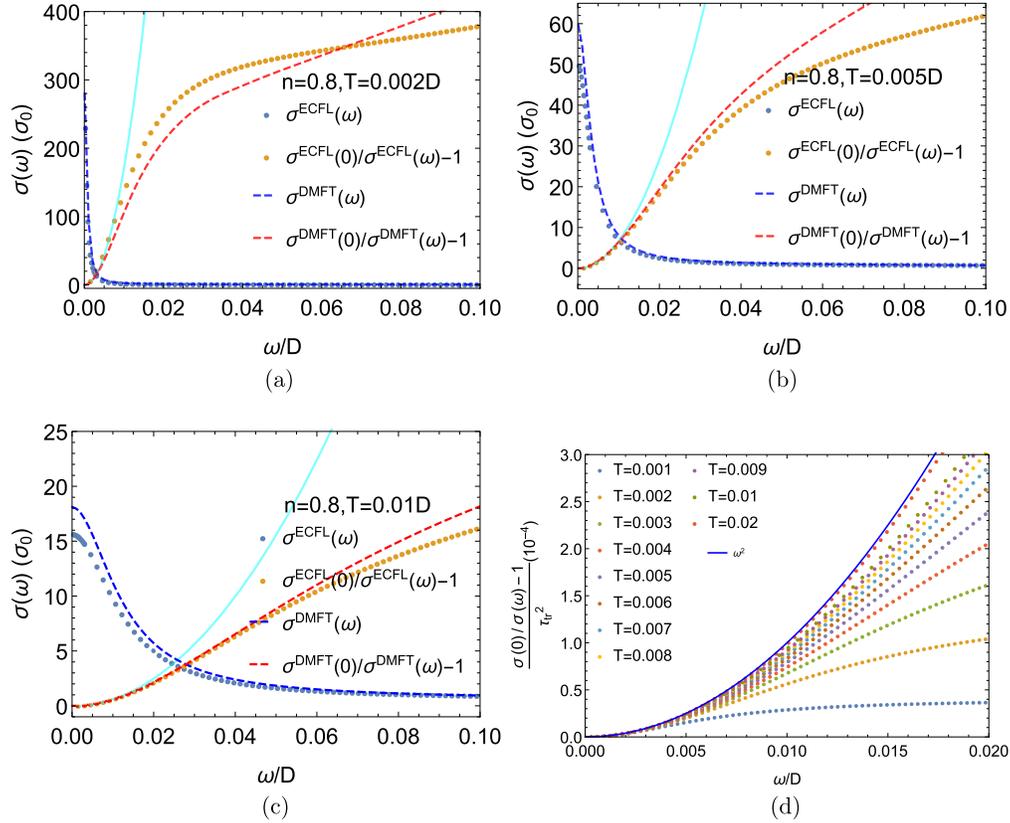


FIG. 4.  $\sigma_{xx}(\omega)$  and  $\sigma(0)/\sigma(\omega) - 1$  for  $n = 0.8$  at  $T = 0.002D$  (a),  $T = 0.005D$  (b), and  $T = 0.01D$  (c) for DMFT (dashed lines) and ECFL (solid symbols). The cyan solid lines are  $\omega^2$  fitting near  $\omega \rightarrow 0$ . In (d) we normalize  $\sigma(0)/\sigma(\omega) - 1$  curves computed from ECFL for various temperatures by  $\tau_{ir}^2$  with  $\tau_{ir}$  obtained from the fits at small frequencies to the Drude formula. The solid blue line is a  $\omega^2$  curve.

corresponding  $\tau_{ir}^2$ , while the  $\omega^2$  curve is shown as a solid blue line. All curves fall onto the  $\omega^2$  line at small frequencies, and peel off at a frequency which increases as  $T$  increases.

These scattering rates are shown as a function of temperature in Fig. 6(a). The scattering rate  $\Gamma$  has a temperature dependence similar to that of the resistivity, i.e., a quadratic- $T$  regime at low temperatures followed by a linear- $T$  regime at higher temperatures.

### B. Optical Hall angle and the transverse scattering rates $\Gamma_H$

In Fig. 5, we show the optical tangent Hall angle  $\tan \theta_H(\omega)$  and the quantity  $\tan \theta_H(0)/\tan \theta_H(\omega) - 1$ . We display plots obtained from both ECFL (symbols) and DMFT (dashed lines) for fixed  $n = 0.8$  and for three temperatures to show the generic behavior at  $T < T_{FL}$ ,  $T \simeq T_{FL}$ , and  $T > T_{FL}$ :  $T = 0.002D$  [Fig. 5(a)],  $T = 0.005D$  [Fig. 5(b)], and  $T = 0.01D$  [Fig. 5(c)]. The ECFL results agree well with those from DMFT within this temperature range.

Just like  $\sigma_{xx}(\omega)$ ,  $\tan \theta_H(\omega)$  possesses a narrow Drude peak below  $T_{FL}$  that broadens in a similar way with increasing temperature.  $[\tan \theta_H(0)/\tan \theta_H(\omega) - 1]$  is quadratic in frequency and we fit  $\tau_H^2 \omega^2$  to extract the transverse relaxation time  $\tau_H$ . The  $\omega^2$  regime, however, has a very narrow, weakly  $T$ -dependent width which is about  $0.003D$ . The relaxation time  $\tau_H$  is extracted by fitting within this very low frequency range. Above this energy a flattening behavior, similar to that

in the optical conductivity, takes place at low temperatures. At higher temperatures and lower hole density, a power-law behavior with an exponent that increases with  $T$  gradually replaces the flattening out behavior. Such a tendency is visible in Figs. 5(d) and 5(e), where all  $[\tan \theta_H(0)/\tan \theta_H(\omega) - 1]$  curves are normalized by their corresponding  $\tau_H^2$ .

In Fig. 6(b) we show  $\Gamma_H$  (defined as  $\Gamma_H \equiv \tau_H^{-1}$ ) for various densities and temperatures obtained from the Drude formula fitting. Their  $T$  dependence is quadratic for both GCFL and GCSM regimes.

### C. Emergence of two relaxation times

In Fig. 6(c), we show  $\Gamma_H/\Gamma_{ir}$  as a function of temperature. At all densities considered this ratio behaves differently for  $T$  below and above  $T_{FL}$ . Below  $T_{FL}$ , the ratio  $\Gamma_H/\Gamma_{ir} \simeq 0.5$  remains essentially constant, and hence the optical transport is dominated by a single scattering rate. Once  $T_{FL}$  is crossed, however,  $\Gamma_H/\Gamma_{ir}$  becomes strongly  $T$  dependent. This indicates that there are *two relaxation times* in the GCSM regime. This is possible since the quasiparticles are no longer well defined for  $T > T_{FL}$ , and different frequency regimes present in the spectral functions contribute differently to the two relaxation times. In Fig. 6(d), we plot  $\Gamma_H/\Gamma_{ir}$  versus the rescaled temperature  $T/T_{FL}$  to illustrate the clearly distinct behavior below and above  $T_{FL}$ .

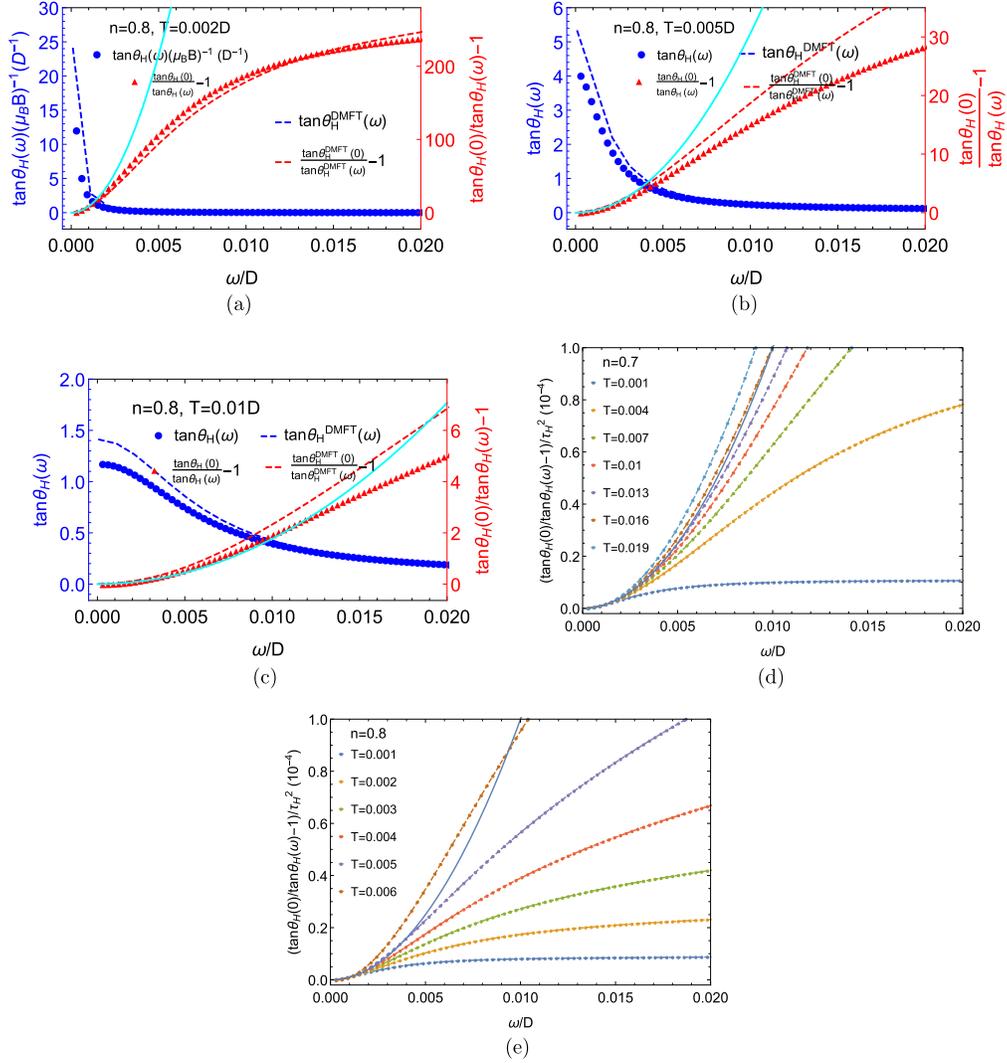


FIG. 5. Optical Hall angles  $\tan\theta_H(\omega)$  (blue) and  $\tan\theta_H(0)/\tan\theta_H(\omega) - 1$  (red) shown for  $n = 0.8$ ,  $T = 0.002D$  (a),  $T = 0.005D$  (b), and  $T = 0.01D$  (c) for DMFT (dashed lines) and ECFL (solid symbols). The cyan solid lines are  $\omega^2$  fitting near  $\omega \rightarrow 0$ .  $[\tan\theta_H(0)/\tan\theta_H(\omega) - 1]/\tau_H^2$  obtained from ECFL shown for  $n = 0.7$  (d) and  $n = 0.8$  (e). Drude peaks are found to be narrow (note the different horizontal axis scale compared to Fig. 3).

## VI. ANALYSIS

We begin by analyzing the exact formulas for the conductivities  $\sigma_{xx}, \sigma_{xy}$  of Eqs. (3) and (4), following Refs. [18] and [6] within ECFL theory where more analytic insight is available.

It has long been noted that the particle-hole asymmetry of the spectral function is one of the characteristic features of strongly correlated systems [25,28–33]. The dynamic particle-hole transformation is defined by simultaneously inverting the wave vector and energy in  $\rho_G(\mathbf{k}, \omega)$  relative to the chemical potential  $\mu$  as  $(\hat{\mathbf{k}}, \omega) \rightarrow -(\hat{\mathbf{k}}, \omega)$ , with  $\hat{\mathbf{k}} = \mathbf{k} - \mathbf{k}_F$  [25]. In the limit of  $d \rightarrow \infty$ , we ignore the  $\hat{\mathbf{k}}$  part of the transformation. Consequently, the dynamic particle-hole asymmetry solely stems from the asymmetry of the self-energy spectral function  $\rho_\Sigma(\omega, T) = -\text{Im} \Sigma(\omega, T)/\pi$ . Instead of analyzing  $\rho_G$ , we can

simply focus on  $\rho_\Sigma$  since

$$\rho_G = \frac{\rho_\Sigma}{(\omega + \mu - \epsilon - \text{Re} \Sigma)^2 + \pi^2 \rho_\Sigma^2} \quad (23)$$

$$= \frac{1}{\pi} \frac{B(\omega, T)}{[A(\omega, T) - \epsilon]^2 + B^2(\omega, T)}, \quad (24)$$

where

$$A(\omega, T) = \omega + \mu - \text{Re} \Sigma(\omega, T), \quad (25)$$

$$B(\omega, T) = \pi \rho_\Sigma(\omega, T) = -\text{Im} \Sigma(\omega, T). \quad (26)$$

Then we approximate the exact equations (3) and (4) by their asymptotic values at low enough  $T$ , following Ref. [6]. The idea is to first integrate over the band energy  $\epsilon$  viewing one

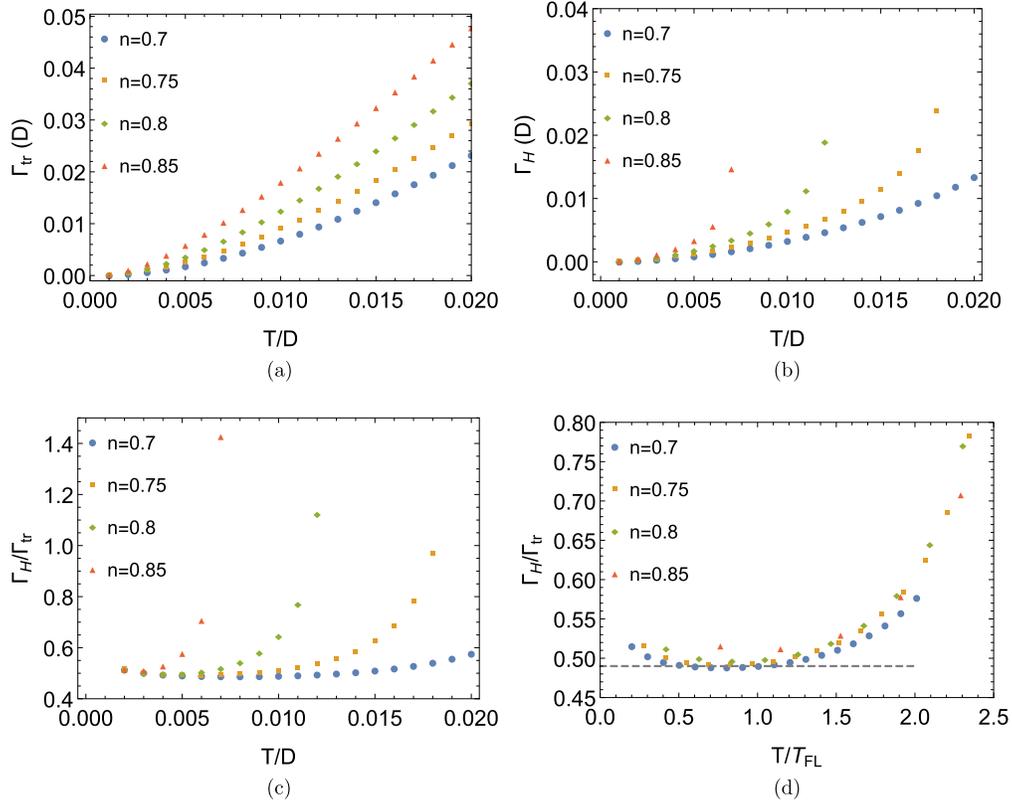


FIG. 6. Longitudinal relaxation rate  $\Gamma_{tr}$  extracted by fitting  $\sigma_{xx}(\omega)$  by the Drude formula (a), transverse relaxation time  $\Gamma_H$  extracted from  $\theta_H(\omega)$  (b), and their ratio  $\Gamma_H/\Gamma_{tr}$  as a function of  $T$  (c) and as a function of scaled temperature  $T/T_{FL}$  (d). All the relaxation rates are extracted from the ECFL optical response results.

of the powers of  $\rho_G$  as a  $\delta$  function constraining  $\epsilon \rightarrow A(\omega, T)$ . This gives

$$\sigma_{xx} = \frac{\sigma_0 D}{\Phi^{xx}[0]} \int d\omega (-f') \frac{\Phi^{xx}[A(\omega, T)]}{B(\omega, T)}, \quad (27)$$

$$\sigma_{xy} = \frac{\sigma_0 D q_e}{\Phi^{xx}[0]} \int d\omega (f') \times \left( \frac{\partial_\omega^2 \Phi^{xy}[A(\omega, T)]}{3} + \frac{\Phi^{xy}[A(\omega, T)]}{2[B(\omega, T)]^2} \right). \quad (28)$$

The first term in Eq. (28) turns out to be negligible compared to the second, and hence we will ignore it. Next, we track down the electronic properties that give rise to a second relaxation time using the above asymptotic expressions.

To the lowest order of approximation at low temperatures, we can make the substitution  $f'(\omega) \rightarrow -\delta(\omega)$  in Eq. (27) and (28), which gives

$$\cot \theta_{H,0}/B = \frac{2B(0, T)}{q_e A(0, T)}. \quad (29)$$

We show  $\cot \theta_{H,0}$  in Fig. 7. When plotted as a function of  $T^2$  as shown in the main panel of Fig. 7,  $\cot \theta_{H,0}$  (solid symbols) is in good agreement with the exact results (dashed lines) both qualitatively, i.e., showing a kinklike feature, and quantitatively except for relatively high temperatures and densities. However, when it is plotted as a function of  $T$  (inset of Fig. 7), we find that the “kink” is actually the crossover from a  $T^2$  behavior to a linear- $T$  behavior and  $\cot \theta_{H,0}$  follows

the  $T$  dependence of  $\rho_{xx}$ . The lowest order approximation is insufficient to capture and to understand the second  $T^2$  regime. Therefore, we pursue more accurate asymptotic expressions of Eqs. (27) and (28). Following Refs. [5] and [1], we do the

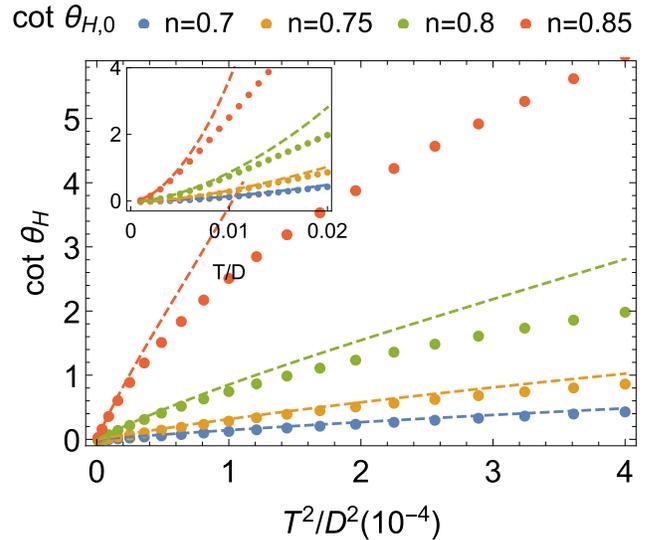


FIG. 7. Zeroth-order asymptotic cotangent Hall angle  $\cot \theta_{H,0}$  plotted as a function of  $T^2$  (main panel, symbols) compared with the exact results (dashed lines) and as a function of  $T$  (inset).

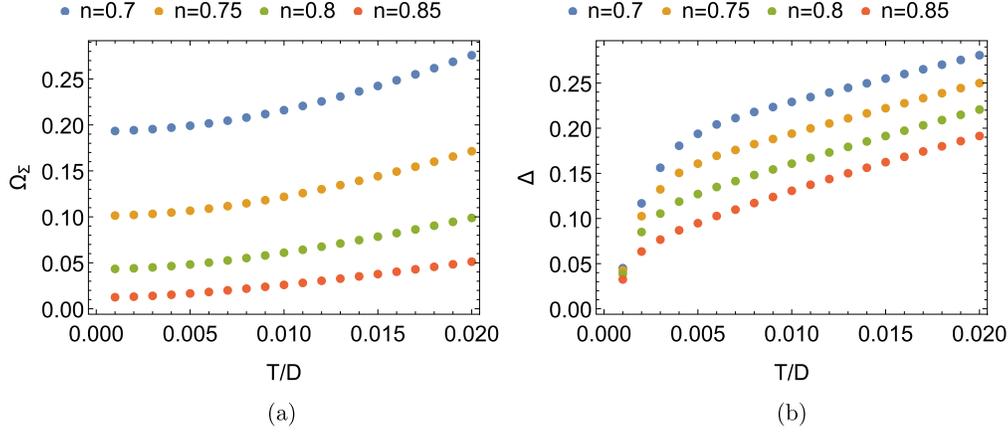


FIG. 8. Coefficients of the small frequency expansion of the ECFL Dyson self-energy  $\Omega_{\Sigma}$  (a) and  $\Delta$  (b) plotted as functions of temperature.

following small-frequency expansion:

$$\Phi^{xx(xy)}[A(\omega, T)] = \Phi^{xx(xy)}[A_0] + \Phi^{xx(xy)'}[A_0]A_1\omega + \dots, \quad (30)$$

$$B(\omega, T) = B_0 + B_1\omega + B_2\omega^2 + \dots, \quad (31)$$

where  $A_0$  and  $A_1$  are given by the expansion

$$A(\omega, T) = A_0 + A_1\omega + \dots. \quad (32)$$

Recall that  $A_1 = Z^{-1}$ ; it is therefore large. In order to provide further context to these coefficients  $B_n$  and to connect with earlier discussions of the self-energy, it is useful to recall a suggestive expression for the imaginary self-energy exhibiting particle-hole asymmetry at  $k_F$  at low  $\omega$  (e.g., see Eq. (28) in Ref. [8]):

$$-\text{Im} \Sigma(\omega, T) \sim \pi \frac{(\omega^2 + \pi^2 T^2)}{\Omega_{\Sigma}(T)} \left(1 - \frac{\omega}{\Delta}\right), \quad (33)$$

where  $\Omega_{\Sigma}$  behaves as  $\sim Z^2$  in the low- $T$  Fermi liquid regime. The scale  $\Delta$  breaks the particle-hole symmetry of the leading term.

The variation of  $\Omega_{\Sigma}$  and  $\Delta$  in the GCSM regime is illustrated in Fig. 8. Expanding this expression at low  $\omega$  we identify the coefficients  $B_0 = \pi \frac{\pi^2 T^2}{\Omega_{\Sigma}(T)}$ ,  $B_1 = -\frac{B_0}{\Delta}$ ,  $B_2 = \frac{\pi}{\Omega_{\Sigma}}$ , all of which are numerically verified to be valid for all temperatures we study in this work. The negative sign of  $B_1$  is easily understood.

Now we keep  $B(\omega, T)$  to  $O(\omega^2)$  and  $A(\omega, T)$  to  $O(\omega)$ , which are the lowest orders required to capture all important features of the exact results. Then Eqs. (27) and (28) can be simplified as

$$\sigma_{xx} \simeq \frac{\sigma_0 F_1^0}{D^2 B_0} (D^2 - A_0^2)^{3/2} \left(1 - \frac{3\pi^2 F_2^2}{F_1^0} \frac{T^2 A_0 A_1}{\Delta (D^2 - A_0^2)}\right), \quad (34)$$

$$\sigma_{xy}/B \simeq \frac{\sigma_0 q_e F_2^0}{2D^2 B_0^2} A_0 (D^2 - A_0^2)^{3/2} \times \left[1 + \frac{\pi^2 F_3^0}{F_2^0} \frac{T^2 A_1}{\Delta A_0} \left(1 - \frac{3A_0^2}{D^2 - A_0^2}\right)\right]. \quad (35)$$

The coefficients are defined as [26]

$$F_m^n = \frac{\pi}{4} \int_{-\infty}^{\infty} \frac{dx}{\cosh^2(\pi x/2)} \frac{x^n}{(1+x^2)^m}. \quad (36)$$

Using Eqs. (34) and (35) and note [26], we can write

$$\sigma_{xx} \simeq \sigma_{xx,0} (1 - \alpha_{xx}), \quad (37)$$

$$\sigma_{xy} \simeq \sigma_{xy,0} (1 - \alpha_{xy}), \quad (38)$$

with

$$\sigma_{xx,0} = \sigma_0 \frac{(D^2 - A_0^2)^{3/2}}{D^2} \frac{0.822467}{B_0}, \quad (39)$$

$$\sigma_{xy,0}/B = \sigma_0 q_e \frac{A_0 (D^2 - A_0^2)^{3/2}}{D^2} \frac{0.355874}{B_0^2}, \quad (40)$$

$$\alpha_{xx} = \frac{A_1 A_0}{D^2 - A_0^2} \frac{3.98598 \times T^2}{\Delta}, \quad (41)$$

$$\alpha_{xy} = -A_1 \left(\frac{1}{A_0} - \frac{3A_0}{D^2 - A_0^2}\right) \frac{2.12075 \times T^2}{\Delta}. \quad (42)$$

$\sigma_{xx,0}$  agrees with previous works [1,6].  $\alpha_{xx(xy)}$  are relative corrections due to  $\Delta$  and  $A_1$  comparing to  $\sigma_{xx(xy),0}$ . Numerical results of  $\alpha_{xx}$  and  $\alpha_{xy}$  are shown in Fig. 9(a). We find that  $|\alpha_{xx}|$  is less than 5% even at the highest temperature. However,  $\alpha_{xy}$  becomes  $O(1)$  in the GCSM regime. Therefore, we obtain the following asymptotic  $\tan \theta_H$  by omitting  $\alpha_{xx}$ :

$$\cot(\theta_H) \simeq \frac{\cot \theta_{H,0}}{(1 - \alpha_{xy})}, \quad (43)$$

$$\cot \theta_{H,0}/B = q_e \frac{B_0}{0.432691 A_0}. \quad (44)$$

We show  $\rho_{xx}$  and  $\cot(\theta_H)$  computed from the asymptotic expressions (37) and (38) in Fig. 9. The asymptotic values are denoted by crosses whereas the results of Eqs. (3) and (4) are denoted by solid circles. The numerical results of Eq. (43) recover the second  $T^2$  regime.

Therefore, we find that the  $\alpha_{xy}$  term due to the higher order terms of  $A(\omega, T)$  and  $B(\omega, T)$  gives rise to the second  $T^2$  regime of  $\cot(\theta_H)$ . Typically such correction is small, such as is the case of  $\alpha_{xx}$ . The significant difference between  $\alpha_{xx}$

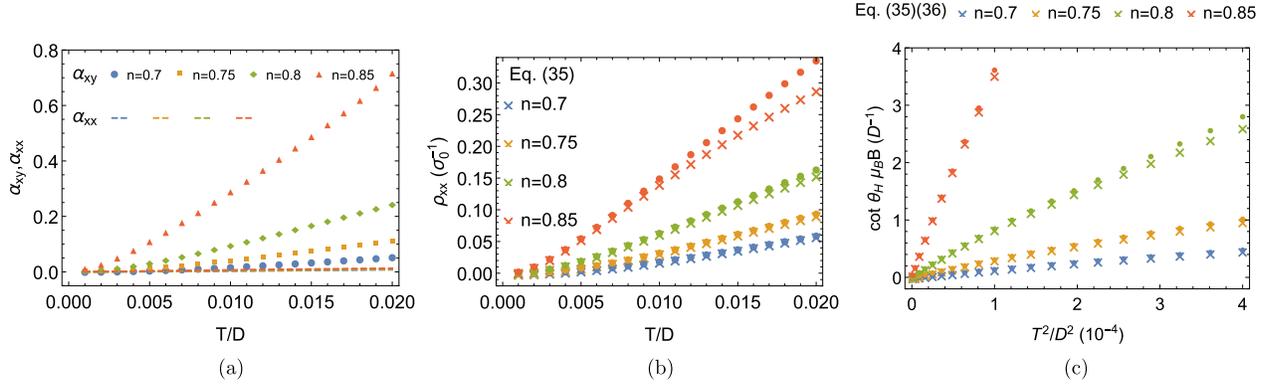


FIG. 9.  $\alpha_{xx}$  (dashed lines) and  $\alpha_{xy}$  (solid symbols) (a),  $\rho_{xx}$  (b),  $\cot(\theta_H)$  (c) computed from Eqs. (37) and (38) using ECFL results. The asymptotic values are denoted by crosses whereas the ECFL results of Eqs. (3) and (4) are denoted by solid circles.

and  $\alpha_{xy}$  is understood by examining Eqs. (41) and (42) more closely. Both  $\alpha_{xx}$  and  $\alpha_{xy}$  are  $\propto A_1 T^2 / \Delta$  with slightly different constant factors. Since  $A_0 \ll D$  and almost independent of  $T$ , we can ignore the  $3A_0(D^2 - A_0^2)^{-1}$  term of  $\alpha_{xy}$ . Hence the difference is mostly determined by a factor

$$\alpha_{xy}/\alpha_{xx} \sim A_0^{-2}, \quad (45)$$

which greatly enhances  $\alpha_{xy}$ .

In the GCFL regime,  $\alpha_{xy}$  is negligible, and the coefficient of the  $T^2$  behavior is

$$Q_{FL} = \frac{B \cot(\theta_H)}{T^2} \simeq \frac{\pi^3}{0.432691 \times q_e A_0 \Omega_\Sigma(T \rightarrow 0)}. \quad (46)$$

$\Omega_\Sigma(T)$  is almost a constant in the GCFL regime hence approximated by its zero-temperature value  $\Omega_\Sigma(T \rightarrow 0)$  [27]. In the GCSM regime, both  $\Omega_\Sigma$  and  $\Delta$  become linear in  $T$ :

$$\Omega_\Sigma(T) \simeq \Omega_0(T + T_\Omega), \quad (47)$$

$$\Delta(T) \simeq \Delta_0(T + T_\Delta), \quad (48)$$

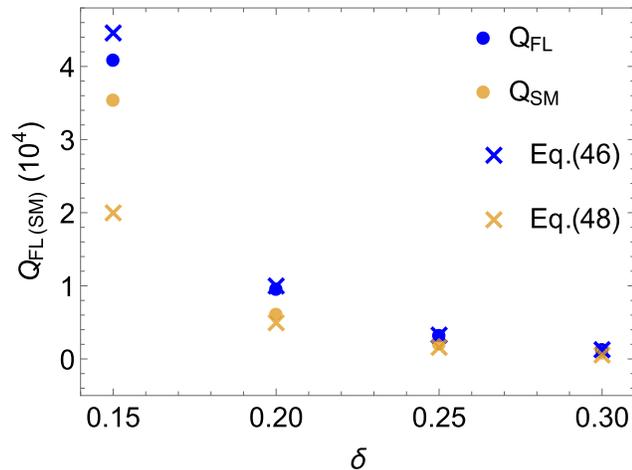


FIG. 10. Equations (46) and (49) (crosses) compared with  $Q_{FL}$  and  $Q_{SM}$  (solid circles) obtained by fitting the exact  $\cot(\theta_H)$ .

where  $\Omega_0(\Delta_0)$  and  $T_{\Omega(\Delta)}$  are fitting parameters [27]. By keeping only the constant term we obtain

$$Q_{SM} \simeq \frac{\pi^3}{0.432691 \times q_e A_0 \Omega_0(T_\Delta + T_\Omega)}. \quad (49)$$

We compare the actual  $Q_{FL}$  and  $Q_{SM}$  with Eqs. (46) and (49) in Fig. 10.

According to the above analysis, the second  $T^2$  behavior of  $\cot(\theta_H)$  is due to the combination of two things:

(1) The dynamic particle-hole antisymmetric component of  $\rho_\Sigma(\omega)$  characterized by the energy scale  $\Delta$ . Its contribution to transport becomes important when  $\pi T$  becomes comparable to  $\Delta$ .

(2) The particular form of the transverse transport function  $\Phi^{xy}(\epsilon)$  that causes  $\Phi^{xy}[A_0]/\Phi^{xy}[A_0] \propto A_0^{-1}$ . Without this factor,  $\alpha_{xy}$  would be negligible as  $\alpha_{xx}$ . This particular form of  $\Phi^{xy}(\epsilon)$  is due to the particle-hole symmetry of the bare band structure.

## VII. DISCUSSION

We have shown that Hall constants, Hall angles, optical conductivities, and optical Hall angles calculated by ECFL agree reasonably well with the DMFT results. The differences tend to increase at higher densities and higher temperatures as noted earlier [6].

We focused on the differences in the behavior above and below the Fermi liquid temperature scale  $T_{FL}$ , i.e., from the GCFL regime to the GCSM regime. Below  $T_{FL}$ , both  $\rho_{xx}$  and  $\cot(\theta_H) \propto T^2$ . Equivalently,  $R_H$  has very weak  $T$  dependence since  $R_H = \rho_{xx}/\cot(\theta_H)$ . When  $T > T_{FL}$ , however,  $\cot(\theta_H)$  passes through a slight downward bend and continues as  $T^2$  whereas  $\rho_{xx} \propto T$ . The significance of the downward bend is that it signals the crossover to the strange metal regime from the Fermi liquid regime.

We explored the long-standing two-scattering-rate problem by calculating both the optical conductivities and optical Hall angles, and the corresponding scattering rates. Below  $T_{FL}$ , both  $\sigma_{xx}(\omega)$  and  $\tan\theta_H(\omega)$  exhibit Drude peaks, which is a manifestation of transport dominated by quasiparticles. The corresponding scattering rates can be extracted by fitting to the Drude formula in the appropriate frequency range. Above  $T_{FL}$ , the Drude peak for  $\sigma_{xx}(\omega)$  becomes broadened; i.e.,

$\sigma_{xx}(0)/\sigma_{xx}(\omega) - 1 \sim \omega^2$  for an even larger range that keeps growing with increasing temperature. In this case, fitting to the Drude formula is still valid, and the scattering rate shows trends as a function of temperature similar to those of the dc resistivity. For  $\theta_H(\omega)$ , the Drude peak range is very narrow, but nonetheless persists for all temperatures that we study in this work. Similarly, the extracted scattering rate  $\Gamma_H$  shows trends as a function of temperature similar to those of the dc Hall angle. At lower dopings and higher temperatures, it seems possible that the Drude peaks of  $\theta_H(\omega)$  would disappear and the fractional power law would stretch down to nearly  $\omega = 0$ .

By comparing the two optical scattering rates through their ratio,  $\Gamma_H/\Gamma_{ir}$ , we clearly demonstrated that  $\Gamma_H$  and  $\Gamma_{ir}$  are equivalent below  $T_{FL}$ , but that they quickly become two distinguishable quantities when the system crosses over into the strange metal region.

By carefully examining the asymptotic expressions of  $\sigma_{xx}$  and  $\sigma_{xy}$  we established that the different temperature dependence of  $\cot(\theta_H)$  in the GCSM regime is governed by a correction caused by both the dynamical particle-hole

antisymmetric component of  $\rho_\Sigma(\omega)$  and the particle-hole symmetry of the bare band structure. This correction is turned on when  $T$  becomes comparable to  $\Delta$ , the characteristic energy scale of the antisymmetric components of  $\rho_\Sigma(\omega)$ .

It would be useful to examine the bend in  $\cot(\theta_H)$  more closely in experiments in cuprate materials, where such a feature is apparently widely prevalent but seems to have escaped comment so far. In particular, one would like to understand better whether the longitudinal resistivity and the cotangent Hall angle show simultaneous signatures of a crossover, as our theory predicts in this work.

## ACKNOWLEDGMENTS

The work at UCSC was supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award No. DE-FG02-06ER46319. R.Ž. acknowledges financial support from the Slovenian Research Agency (Research Core Funding No. P1-0044 and Project No. J1-7259).

- 
- [1] W. Ding, R. Žitko, P. Mai, E. Perepelitsky, and B. S. Shastry, *Phys. Rev. B* **96**, 054114 (2017).
- [2] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989).
- [3] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [4] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [5] W. Xu, K. Haule, and G. Kotliar, *Phys. Rev. Lett.* **111**, 036401 (2013).
- [6] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016).
- [7] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).
- [8] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013).
- [9] B. S. Shastry and P. Mai, [arXiv:1703.08142](https://arxiv.org/abs/1703.08142).
- [10] T. R. Chien, Z. Z. Wang, and N. P. Ong, *Phys. Rev. Lett.* **67**, 2088 (1991).
- [11] Y. Ando, Y. Kurita, S. Komiya, S. Ono, and K. Segawa, *Phys. Rev. Lett.* **92**, 197001 (2004).
- [12] F. F. Balakirev, J. B. Betts, A. Migliori, I. Tsukada, Y. Ando, and G. S. Boebinger, *Phys. Rev. Lett.* **102**, 017004 (2009).
- [13] J. Takeda, T. Nishikawa, and M. Sato, *Physica C* **231**, 293 (1994); see especially Fig. 4.
- [14] P. W. Anderson, *Phys. Rev. Lett.* **67**, 2092 (1991).
- [15] Y. Ando, S. Komiya, K. Segawa, S. Ono, and Y. Kurita, *Phys. Rev. Lett.* **93**, 267001 (2004).
- [16] A. Khurana, *Phys. Rev. Lett.* **64**, 1990 (1990).
- [17] T. Pruschke, D. L. Cox, and M. Jarrell, *Phys. Rev. B* **47**, 3553 (1993).
- [18] P. Voruganti, A. Golubentsev, and S. John, *Phys. Rev. B* **45**, 13945 (1992).
- [19] L.-F. Arsenault and A.-M. S. Tremblay, *Phys. Rev. B* **88**, 205109 (2013).
- [20] J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids* (Oxford University Press, 1960).
- [21] D. Feng and G. Jin, in *Introduction to Condensed Matter Physics* (World Scientific, 2005), pp. 199–229.
- [22] Here we give numerical values of  $T_{FL}$  of both ECFL and DMFT for  $n = 0.7, 0.75, 0.8, 0.85$ :
- |                  | 0.7     | 0.75    | 0.8     | 0.85    |
|------------------|---------|---------|---------|---------|
| $T_{FL,ECFL}(D)$ | 0.00994 | 0.00726 | 0.00478 | 0.00262 |
| $T_{FL,DMFT}(D)$ | 0.00937 | 0.00794 | 0.00601 | 0.00408 |
- [23] H. Y. Hwang, B. Batlogg, H. Takagi, H. L. Kao, J. Kwo, R. J. Cava, J. J. Krajewski, and W. F. Peck, Jr., *Phys. Rev. Lett.* **72**, 2636 (1994).
- [24] S. Ono, S. Komiya, and Y. Ando, *Phys. Rev. B* **75**, 024515 (2007).
- [25] B. S. Shastry, *Phys. Rev. Lett.* **109**, 067004 (2012).
- [26] Here we give numerical values of  $F_m^n$  here,  $F_1^0 = \pi^2/12 = 0.822467$ ,  $F_2^2 = \pi^2/24 - \zeta(3)/4 = 0.110719$ ,  $F_2^0 = \pi^2/24 + \zeta(3)/4 = 0.711748$ ,  $F_3^2 = \pi^2/96 - \pi^4/960 + \zeta(3)/16 = 0.0764691$ , where  $\zeta(3) = 1.20206\dots$  is the Reimann zeta function.
- [27] Here we give numerical values of  $\Omega_\Sigma(T \rightarrow 0)$ ,  $\Omega_0, \Delta, T_\Omega$ , and  $T_\Delta$  for  $n = 0.7, 0.75, 0.8, 0.85$ :
- |                                  | 0.7       | 0.75      | 0.8        | 0.85        |
|----------------------------------|-----------|-----------|------------|-------------|
| $\Omega_\Sigma(T \rightarrow 0)$ | 0.194326  | 0.10234   | 0.0443113  | 0.0135004   |
| $\Omega_0$                       | 5.99932   | 4.97066   | 3.79492    | 2.52377     |
| $\Delta$                         | 5.15313   | 5.58921   | 5.97819    | 6.05317     |
| $T_\Omega$                       | 0.0257944 | 0.0143857 | 0.00607066 | 0.000456418 |
| $T_\Delta$                       | 0.0346638 | 0.0248982 | 0.0171223  | 0.0118911   |

- [28] B. Revaz, J.-Y. Genoud, A. Junod, K. Neumaier, A. Erb, and E. Walker, *Phys. Rev. Lett.* **80**, 3364 (1998).
- [29] P. W. Anderson and N. P. Ong, *J. Phys. Chem. Solids* **67**, 1 (2006).
- [30] T. Hanaguri, C. Luplen, Y. Kohsaka, D.-H. Lee, M. Azuma, M. Takano, H. Takagi, and J. C. Davis, *Nature (London)* **430**, 1001 (2004).
- [31] A. N. Pasupathy, A. Pushp, K. K. Gomes, C. V. Parker, J. Wen, Z. Xu, G. Gu, S. Ono, Y. Ando, and A. Yazdani, *Science* **320**, 196 (2008).
- [32] P. A. Casey, J. D. Koralek, N. C. Plumb, D. S. Dessau, and P. W. Anderson, *Nat. Phys.* **4**, 210 (2008).
- [33] G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).

## $t$ - $t'$ - $J$ model in one dimension using extremely correlated Fermi-liquid theory and time-dependent density matrix renormalization group

Peizhi Mai,<sup>1,\*</sup> Steven R. White,<sup>2,†</sup> and B. Sriram Shastry<sup>1,‡</sup>

<sup>1</sup>Physics Department, University of California, Santa Cruz, California 95064, USA

<sup>2</sup>Department of Physics and Astronomy, University of California, Irvine, California 92717, USA



(Received 14 December 2017; revised manuscript received 24 February 2018; published 6 July 2018)

We study the one-dimensional  $t$ - $t'$ - $J$  model for generic couplings using two complementary theories, the *extremely correlated Fermi liquid theory* and *time dependent density matrix renormalization group* over a broad energy scale. The two methods provide a unique insight into the strong *momentum dependence* of the self-energy of this prototypical non-Fermi liquid, described at low energies as a Tomonaga-Luttinger liquid. We also demonstrate its intimate relationship to spin-charge separation, i.e., the splitting of Landau quasiparticles of higher dimensions into two constituents, driven by strong quantum fluctuations inherent in one dimension. The momentum distribution function, the spectral function, and the excitation dispersion of these two methods also compare well.

DOI: [10.1103/PhysRevB.98.035108](https://doi.org/10.1103/PhysRevB.98.035108)

### I. INTRODUCTION

In varying dimensions the  $t$ - $J$  model continues to attract attention owing to its relevance in cuprates and other important strongly interacting electronic systems. The model embodies very strong correlations, which lie outside the regime of validity of perturbation theory, and thus pose a challenging problem. Our main goal in this work is to obtain an understanding of the properties in one dimension (1D), *over a wide energy range*.

At low energies the bosonization technique has been widely applied to the (closely related) Hubbard model [1–5]. For large  $U$  several nonperturbative methods have been devised to study the  $t$ - $J$  model for general dimensions, including the study of finite clusters [6,7] and large- $N$  based slave particle mean-field theories [8]. In 1D we also have exact results using Bethe's ansatz [9–14] at special values of the parameters of the model, and also for long-ranged versions [15] of the  $t$ - $J$  model, using techniques developed in the Haldane-Shastry models. Photoemission experiments [16] have been carried out to study the spectral properties of several quasi-1D metals, relevant to the  $t$ - $J$  model.

To study a wider energy range, including the low to intermediate and high energy regimes, we employ and compare the results from two complementary techniques. In 1D, the density matrix renormalization group (DMRG) [17] provides nearly exact results for the ground state and can also be used for finite temperature and spectral properties. Ground state DMRG has been used to give the phase diagram of the  $t$ - $J$  model over a broad range of parameters in Ref. [18]. Here we study dynamics using the time dependent density matrix renormalization group (tDMRG). tDMRG [17,19] has been used to obtain virtually exact spectral functions for spin chains, but only a few times for doped Fermi systems. One such time was a tDMRG treatment of the  $t$ - $J$  model, obtaining spectral functions for the system at finite temperature [20]. In this

work we use tDMRG only at  $T = 0$ , but we have pushed much farther in terms of system size, accuracy, and frequency resolution than in Ref. [20]. This accuracy is needed to resolve the detailed features of the self-energy, which has not been done before with tDMRG.

The other technique used is the extremely correlated Fermi liquid (ECFL) theory [21]. This analytical theory, which can treat a large class of large  $U$  problems, including the  $t$ - $J$  model, uses Schwinger's functional differential equations for the electron Green's function. These equations are systematically expanded in a parameter  $\lambda \in [0, 1]$ , representing partial Gutzwiller projection. The  $\mathcal{O}(\lambda^2)$  theory leads to a closed set of coupled equations [21,22] for the Green's function. This treatment has been benchmarked in high dimensions and in 2D. In infinite dimensions, dynamical mean field theory (DMFT) [23] provides a solution to the Hubbard model, and ECFL has been benchmarked recently [24,25] against exact results from the single impurity Anderson model, and DMFT in  $d = \infty$  [26,27]. The limiting case  $U = \infty$  has been explored in detail in Ref. [28]. The agreement at low energies is good enough to yield accurate results for the low  $T$  resistivity, a highly sensitive variable. In 2D, ECFL has been applied recently to cuprate superconductors [29,30]. It is therefore interesting to see how well this scheme deals with the physics of 1D. The equations used here have the character of a skeleton graph series. We have checked that the second order skeleton graphs for the Hubbard model in 1D already displays characteristics of spin-charge separation and non-Fermi liquid spectral functions, while the nonskeleton, i.e., bare perturbation theory does not.

Understanding the extent of *momentum dependence* of the Dysonian self-energy  $\Sigma$  in various dimensions is one of the goals of the present paper. While the  $d = \infty$  models have a momentum *independent* self-energy, momentum dependence of  $\Sigma$  is inevitable in lower dimensions. However there is a scarcity of reliable information on its extent and location. In most published work, the self-energy in 1D is rarely presented [31], or even calculated, since standard solutions directly deal with the Green's function. In contrast we focus on unraveling the  $(\vec{k}, \omega)$  dependence of the Dysonian self-energy in 1D and comparing with its higher dimensional counterparts.

\*pemai@ucsc.edu

†srwhite@uci.edu

‡b.sriram.shastry@gmail.com

## II. OVERVIEW

In the present paper we solve the  $d = 1$   $t$ - $t'$ - $J$  model for generic parameters using *the same set of ECFL equations* as in higher dimensions. We calculate from the two theories the momentum distribution function, self-energy, spectral function, and excitation dispersion over a broad energy scale.

In the low  $k, \omega$  regime exhibiting non-Fermi liquid behavior, reasonable agreement is found between the two and the exact diagonalization (ED) data in the velocities of spinons and holons [6], as well as the Tomonaga-Luttinger liquid (TLL) theory in anomalous exponent [18]. Extending the  $\mathcal{O}(\lambda^2)$  ECFL equations to higher orders holds the promise of a better agreement. At higher energies, where few studies exist, the agreement between the two theories is quite good already. A valuable insight gained at low energies is the close relationship between a momentum dependent ridge in the  $\text{Im } \Sigma(k, \omega)$  and the spin-charge separation.

## III. MODEL AND PARAMETERS USED

The Hamiltonian of the 1D  $t$ - $t'$ - $J$  model is

$$H_{tJ} = -t \sum_{\langle ij \rangle} X_i^{\sigma 0} X_j^{0\sigma} - t' \sum_{\langle\langle ij \rangle\rangle} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_i X_i^{\sigma\sigma},$$

$$+ J \sum_{\langle ij \rangle} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} X_i^{\sigma\sigma} X_j^{\sigma'\sigma'} \right), \quad (1)$$

where repeated spin indices are summed,  $X_i^{\sigma 0} = P_G C_{i\sigma}^\dagger P_G$ ,  $X_i^{0\sigma} = P_G C_{i\sigma} P_G$ ,  $X_i^{\sigma\sigma'} = P_G C_{i\sigma}^\dagger C_{i\sigma'} P_G$  with  $P_G = \prod_i (1 - n_{i\uparrow} n_{i\downarrow})$  as the Gutzwiller projection operator.  $\langle ij \rangle$  and  $\langle\langle ij \rangle\rangle$  refers to summing over first and second neighbor pairs, respectively.

For this model [21,29] we compute the results from the two theories at density  $n = 0.7$ , second nearest neighbor hopping  $t'/t = 0, 0.2$  and  $J/t = 0.3, 0.6$ . We avoid the special cases of  $t' = 0 = J$  since this leads to a degenerate spectrum, with a charge sector that is isomorphic to the spinless Fermi gas. The ECFL results are shown at various  $T$  while the tDMRG results are at  $T = 0$  where most reliable calculations are possible.  $t = 1$  is the energy unit and will be neglected below.

The tDMRG methods used are very similar to those used in Ref. [32]. We start by obtaining the ground state  $|\psi\rangle$  using DMRG on a rather long but finite chain, with  $L = 400$ , and then apply  $\hat{c}_0$  or  $\hat{c}_0^\dagger$  to a site 0 near the center, forming  $|\psi(t=0)\rangle$ . We use a Trotter based time evolution algorithm, with fermionic swap gates to handle next-nearest neighbor terms. We specify a density matrix eigenvalue truncation cutoff of  $3 \times 10^{-8}$  during the evolution, subject to a constraint on the maximum number of states kept of  $m = 3000$ . (Results were checked by comparing to  $m = 2000$ .) We evolve out to a time  $t = 50$ . At  $t = 50$ , the normalization of  $|\psi(t)\rangle$  had decreased by a few percent, a small error affecting primarily the widths of any sharp peaks. The space and time dependent Green's function

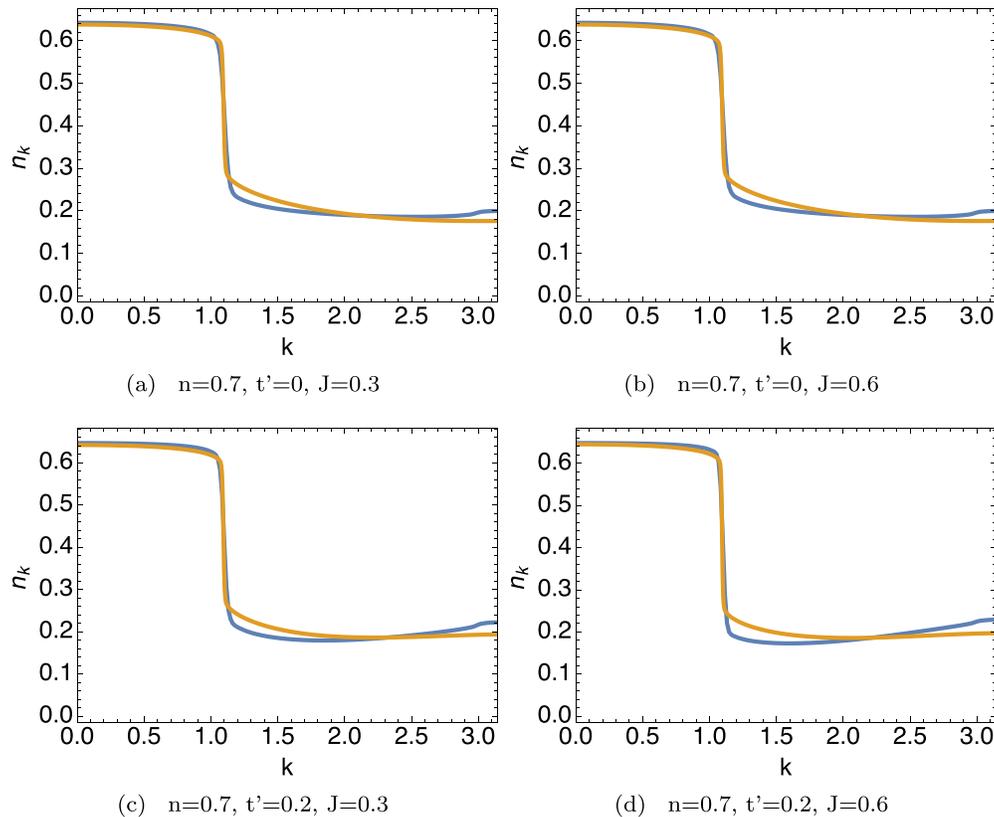


FIG. 1. Momentum distribution  $n_k$  for ECFL (yellow) at  $T = 0.005$  and tDMRG (blue) at  $T = 0$  with  $n = 0.7$ ,  $J = 0.3, 0.6$ , and  $t' = 0, 0.2$ . In all cases these two methods agree well especially in the occupied region and both give a power law singularity at  $k_F$ . The small discrepancy in the unoccupied region corresponds to the  $3k_F$  feature in the exact solutions discussed in Ref. [9]. This subtle singularity is missed by the  $\mathcal{O}(\lambda^2)$  equations.

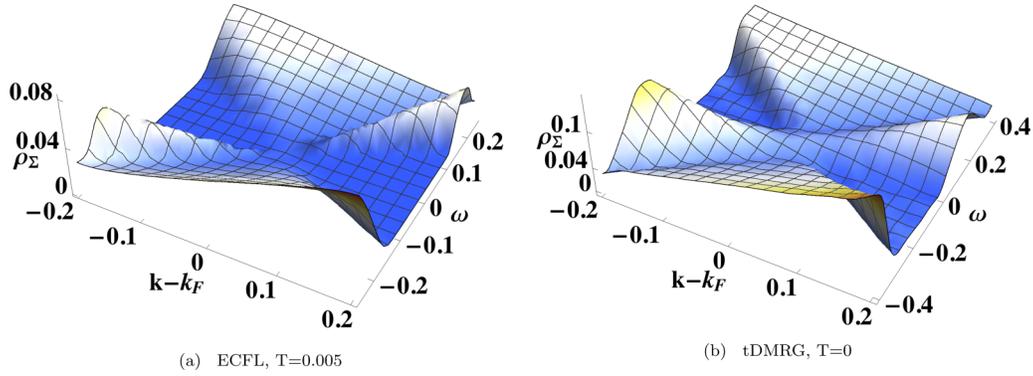


FIG. 2.  $n = 0.7$ ,  $J = 0.3$ ,  $t' = 0$ : Imaginary self-energy  $\rho_{\Sigma}(k, \omega)$  at low  $\omega$  and  $k - k_F$  from both methods. Both give a dominant  $(k, \omega)$  dependent ridge running from left to right, and a less prominent feature running from top-left to bottom-right. Both of them pass through the  $k = k_F, \omega = 0$  region. The dominant ridge is responsible for the appearance of the twin peaks structure in the spectral functions which represents the spin-charge separation. The peaks for  $k < k_F, \omega < 0$  are seen in the left half of the electronic spectral function in Fig. 6 panels (a) and (b), while the peaks for  $k > k_F, \omega > 0$  are seen in the right half of the same figures. As seen in Fig. 5 panel (c), the peak in the self-energy  $\rho_{\Sigma}$  directly leads to a dip in the electronic spectral function  $\rho_G$ , provided the real part is small.

is obtained by sandwiching  $\hat{c}_i$  or  $\hat{c}_i^{\dagger}$  between the ground state and  $|\psi(t)\rangle$  for all  $i$ . Linear prediction is used to extend the time dependent Green's function out to  $t = 100$ , after which the data is windowed and Fourier transformed. This calculation represents the most accurate and detailed study to date of the spectral properties of the model at  $T = 0$ .

#### IV. MOMENTUM DISTRIBUTION FUNCTION

In 1D  $t$ - $J$  model,  $n_k$  shows a power law singularity at  $k_F$  [2,5], a signature of the TLL, unlike a jump in higher dimensions as Fermi liquid behavior. This feature is observed from both methods in Fig. 1 for different  $t'$  and  $J$ . Due to the second order approximation, the weak  $3k_F$  singularity related to shadow band [9,12] is not observed in ECFL results. Besides this weak effect,  $n_k$  from both methods agrees well, especially in the occupied side, showing that ECFL describes the correct  $t'$  and  $J$  dependent behaviors.

#### V. SELF-ENERGY

Next we present the Dysonian self-energy in terms of its spectral function  $\rho_{\Sigma}$  defined as

$$\rho_{\Sigma}(k, \omega) = -\frac{1}{\pi} \text{Im} \Sigma(k, \omega). \quad (2)$$

It is derived separately from the Green's functions in ECFL and tDMRG methods. In tDMRG,  $\Sigma$  can be found from  $G$  by inverting the Dyson relation  $G^{-1} = G_0^{-1} - \Sigma$ . The ECFL theory produces two (non-Dysonian) self energies  $\Phi, \Psi$  [21], and the resulting  $G$  can again be inverted to find the standard Dysonian  $\Sigma$ . Both ECFL ( $T = 0.005$ ) and tDMRG ( $T = 0$ ) self-energies are shown in Fig. 2 for comparison.

In Fig. 2, the two theories have a similar pattern of  $k$  dependence, a dominant ridge running from left to right, and a less prominent feature running from top-left to bottom-right. They pass through the  $k = k_F, \omega = 0$  region. The ridge leads to the appearance of twin peaks in the spectral functions representing spin-charge separation. In the higher energy

region in Fig. 3, both theories agree well and are similar to their higher dimensional counterparts.

A powerful feature of ECFL theory is that it allows us to vary temperature without extra effort, at least in the low to intermediate temperature region. In Fig. 4,  $\rho_{\Sigma}$  at  $k_F$  is presented in several temperatures. The bump becomes higher with increasing temperature though no obvious change in larger scale [panel (b)]. This is expected because warming softens the peak height of spectral function at  $k_F$ , which is  $\rho_G(k_F, 0) = 1/(\pi^2 \rho_{\Sigma}(k_F, 0))$  in panel (c). The central peak height  $\rho_{\Sigma}(k_F, 0)$  scales as  $T^{\alpha}$  with  $\alpha \approx 1.1$ , as opposed to  $\alpha = 2$  expected for a Fermi liquid. Although  $T = 0.005$  is the lowest temperature in the current numerical scheme for second order ECFL due to the finite lattice size (up to  $L = 2417$  and  $N_{\omega} = 2^{17}$ ), we extrapolate the curve to  $T = 0$ .

The peak at  $k_F$  disappears at zero  $T$  and is replaced by a minimum at the origin corresponding to a singular peak in the spectral function, consistent with earlier studies [2,12]. The self-energy approaches zero as  $|\omega|^{\gamma}$ , where  $\gamma \approx 1.3$ . This behavior is difficult to observe in our present tDMRG implementation, because the finite time cutoff, leads to a broadening. The peak and its  $k$  dependence is recovered on moving away from  $k_F$ , causing spin-charge separated peaks at  $T = 0$ .

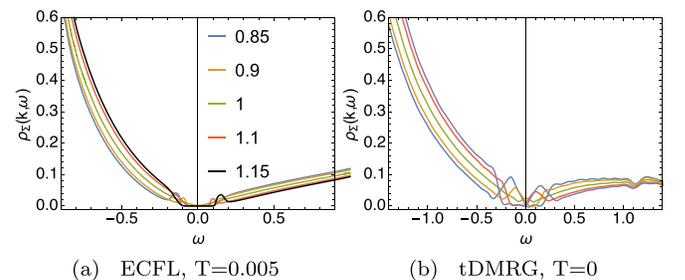


FIG. 3.  $n = 0.7, J = 0.3$ :  $\rho_{\Sigma}(k, \omega)$  vs  $\omega$  at marked  $k/k_F$ 's, from ECFL at  $T = 0.005$  (a) and tDMRG at  $T = 0$  (b) in a large scale. The two sets of results are similar on a broad energy scale, and are comparable to higher dimensional results. The low energy behavior is discussed below.

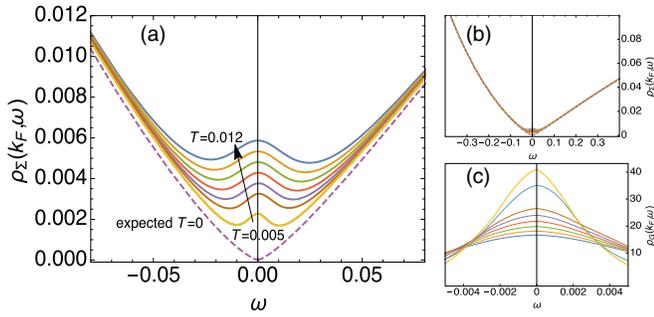


FIG. 4.  $\rho_\Sigma(k_F, \omega)$  from ECFL is shown in (a) for several  $T$  at  $J = 0.3, t' = 0$ . The central peak  $\rho_\Sigma(k_F, 0)$  scales as  $T^{1.1}$ , in contrast to Fermi liquid behavior  $T^2$ . Extrapolating to  $T = 0$  the double minimum structure disappears, leaving behind a  $\sim |\omega|^{1.3}$  dependence. (b) displays the self-energy in larger scale where changing  $T$  barely makes a difference. (c) shows the spectral function softened by warming.

## VI. SPECTRAL FUNCTION

We also compare the spectral functions from both methods. In Fig. 5 panels (a) and (b) both show a single peak at  $k_F$  and double peaks away from  $k_F$  representing spinons and holons,

respectively. Panel (c) puts together the spectral function away from  $k_F$  and different parts of its formula:

$$\rho_G(k, \omega) = \frac{\rho_\Sigma(k, \omega)}{[\omega + \mu - \varepsilon_k - \text{Re}\Sigma(k, \omega)]^2 + \pi^2 \rho_\Sigma^2(k, \omega)}. \quad (3)$$

It shows that  $\omega + \mu - \varepsilon_k - \text{Re}\Sigma(k, \omega)$  is very small in the frequency range that spans the two peaks and confirms that the visible twin peaks result from a peak in  $\rho_\Sigma$  in the middle. Thus the location of the ridge lies in the minimum between spinon and holon peaks in the spectral function in panels (a) and (b), and in fact the ridge causes the twin peaks. The exponents in panel (d) match reasonably with those from the TLL at  $J = 0.3, t' = 0$  and also at 0.6 (where  $\zeta' \sim -0.49$  versus  $\zeta' \sim -0.46$  from Ref. [18]). We take the Luttinger parameter  $K_\rho \approx 0.53$  at  $J = 0.3, t' = 0$  from Fig. 4 in Ref. [18]. Then we calculate  $\zeta = \gamma_\rho = (K_\rho + K_\rho^{-1} - 2)/8 \approx 0.05$  [1,4]. Therefore the anomalous exponent is  $\zeta' = \zeta - \frac{1}{2} = -0.45$ . The calculation is similar for  $J = 0.6$  with  $K_\rho \approx 0.56$  from Fig. 4. The tDMRG spectral function in panel (b) is too soft to extract the anomalous exponent, because its finite time cutoff leads to the broadening of spectral peaks in the low  $\omega$  region.

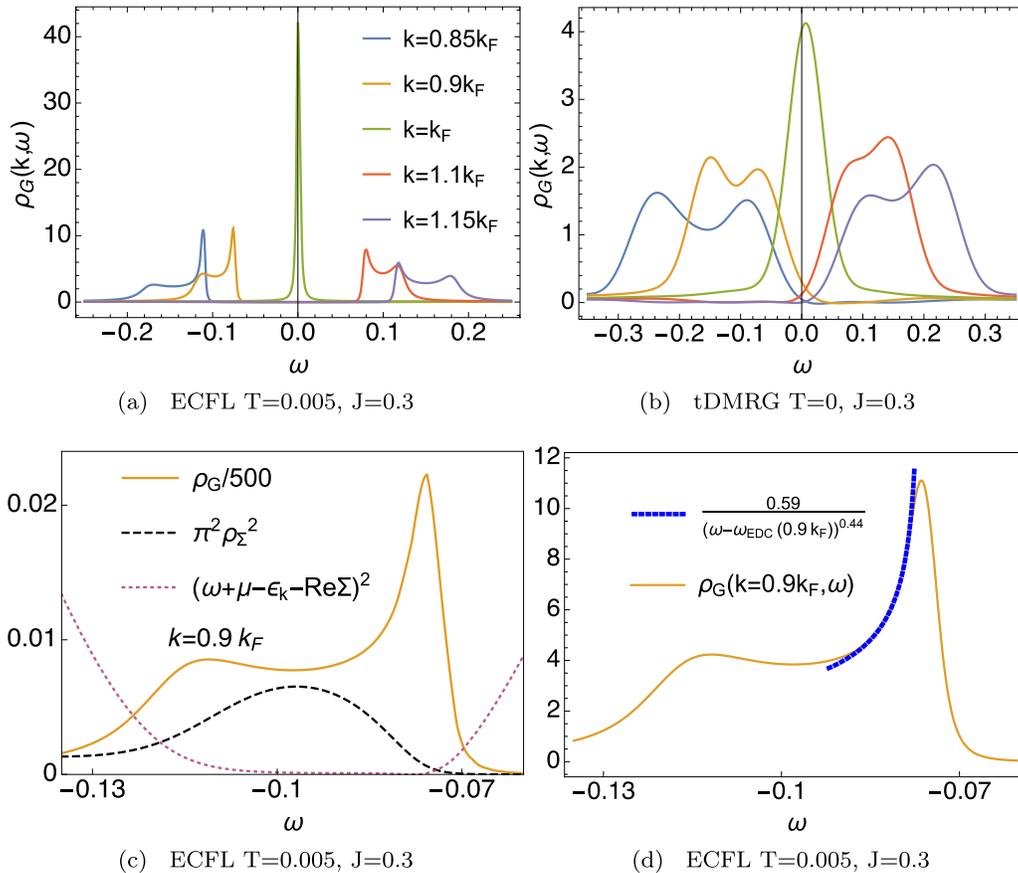


FIG. 5. Energy distribution curves (EDCs) at  $t' = 0, J = 0.3$ : (a) and (b) (same legends marking  $k/k_F$ ) displaying the spinon and the holon for  $k \neq k_F$ . Panel (c) at  $k = .9k_F$  shows that the peak in  $(\pi\rho_\Sigma)^2$  (dashed black) coincides with the dip in the spectral function  $\rho_G(\omega)$  (solid gold), while  $(\omega + \mu - \varepsilon_k - \text{Re}\Sigma)^2$  (magenta dots) is small everywhere. This implies that the twin peaks originate in the intervening peak of self-energy. Panel (d) also at  $k = .9k_F$  shows the fitting procedure for finding the anomalous exponent  $\zeta' \equiv \zeta - \frac{1}{2}$  for the spinon [1,4], we fit to  $.59(\omega - \omega_{\text{peak}})^{\zeta'}$  (dashed blue), the best fit value is  $\zeta' \sim -0.44$ , close to the TLL result  $-0.45$  [18].

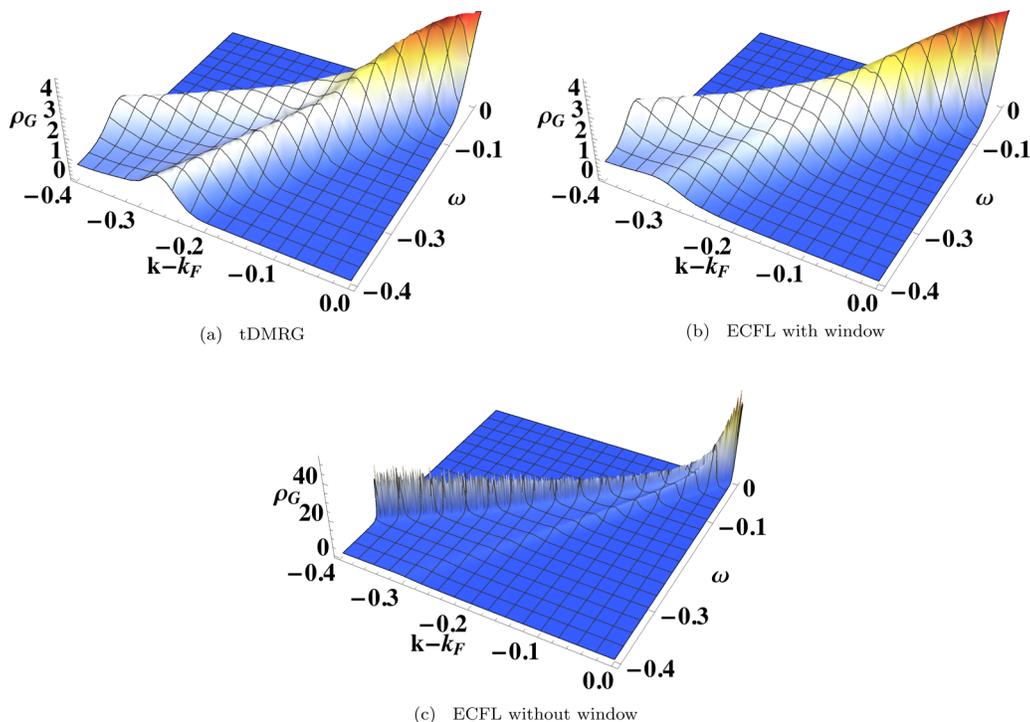


FIG. 6.  $J = 0.6, t' = 0$ . The spectral function of the tDMRG ( $T = 0$ ) with an intrinsic time window (a) and the ECFL ( $T = .005$ ) with (b) and without (c) a comparable time window. The introduction of a time window brings the two theories to the same scale. The central peak and the spinon peaks are of comparable height while the holon peak of ECFL is less prominent due to second order approximation.

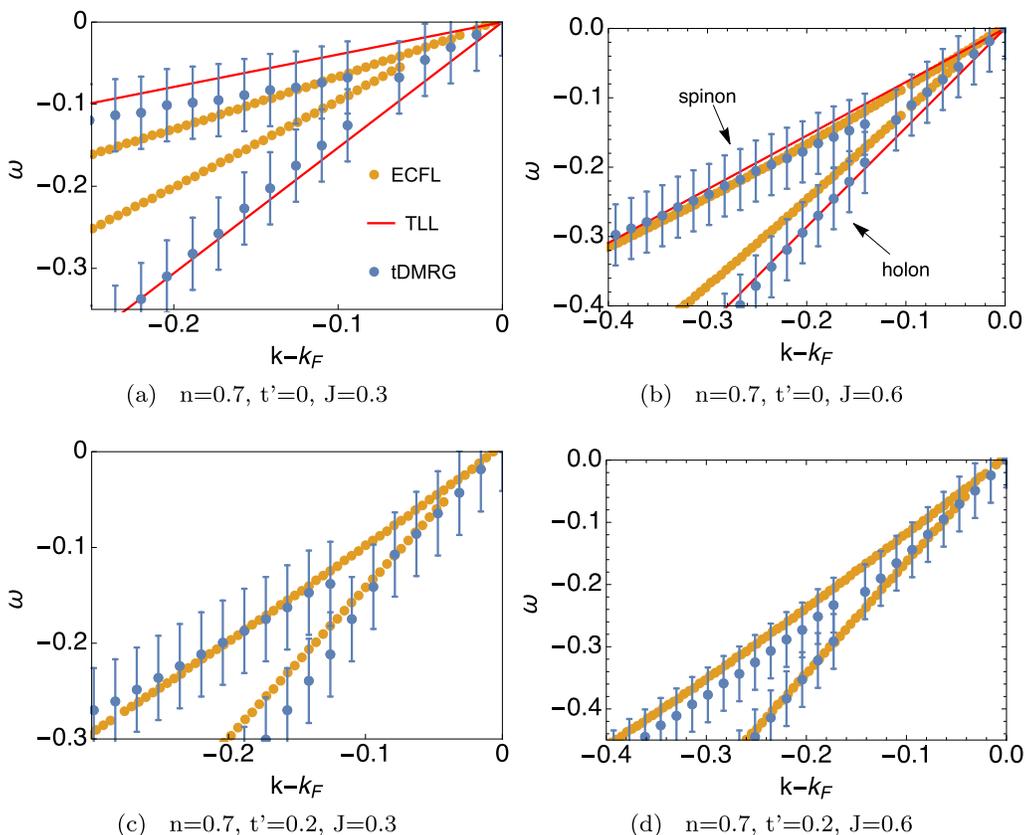


FIG. 7. Dispersion of excitations from both ECFL at  $T = 0.005$  (gold dots) and tDMRG at  $T = 0$  (blue dots), and the available ED data (red) [6]. The error bars in the tDMRG estimates are from the time window broadening. The tDMRG results are consistent with the ED results, while the ECFL holon dispersion deviates somewhat.

## Nonresonant Raman scattering in extremely correlated Fermi liquids

Peizhi Mai and B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*



(Received 24 May 2018; revised manuscript received 1 August 2018; published 4 September 2018)

We present theoretical results for the optical conductivity and the nonresonant Raman susceptibilities for three principal polarization geometries relevant to the square lattice. The susceptibilities are obtained using the recently developed extremely correlated Fermi liquid theory for the two-dimensional  $t$ - $t'$ - $J$  model, where  $t$  and  $t'$  are the nearest- and second-neighbor hopping. Our results sensitively depend on  $t$ ,  $t'$ . By studying this quartet of related dynamical susceptibilities, and their dependence on  $t$ ,  $t'$ , doping, and temperature, we provide a useful framework of interpreting and planning future Raman experiments on strongly correlated matter.

DOI: [10.1103/PhysRevB.98.115101](https://doi.org/10.1103/PhysRevB.98.115101)

### I. INTRODUCTION

Inelastic or Raman scattering of electrons by photons ( $e$ - $\gamma$ ) in strongly correlated systems is of considerable current interest. The scattering intensity, given by the Kramers-Heisenberg formula [1], consists of a resonant and a nonresonant piece. The nonresonant piece depends only on the energy transfer. In contrast, the resonant piece also depends on the incident energy, and it is the focus of this work. In typical weakly correlated metals, this contribution is confined to a small energy window of a few meV [2,3]. Raman scattering theory, if based solely on density fluctuations, would give a vanishing contribution as  $q \rightarrow 0$  due to the conservation law in that limit. The early works of Refs. [2,4] showed that nonparabolic bands lead to the coupling of light to a nonconserved operator (the stress tensor operators discussed below), rather than the density. These operators are exempt from conservation laws that govern the density, and therefore they can lead to nonresonant Raman scattering.

Recent experiments [5–17] in strongly correlated metallic systems such as high- $T_c$  superconductors have added further complexity to challenge to our understanding. It is found that the scattering is  $q$ -independent and extends over a much larger energy range  $O$  (eV), and it is also observed to have a complex  $T$  dependence [5–7,10,14]. To explain these, a systematic reformulation of light scattering in narrowband systems was developed in [18–23]. Shastry and Shraiman (SS) [18,19] developed a theory of Raman scattering in Mott-Hubbard systems using the Hubbard model, where nonparabolicity of bands is built in correctly, so that the conservation law concerns are taken care of. However, the large energy spread of the nonresonant signals remains unaccounted for. It cannot arise from quasiparticles in Fermi liquids, and hence SS argued that a large contribution from the incoherent background of the electron spectral function is required to explain the data (see, e.g., [5,6]). This qualitative argument is not fine enough to explain or predict differences in backgrounds in different geometries. The latter remains an unresolved problem, and it is the focus of the present work.

Progress toward a solution at the microscopic level has been slow since a suitable theory in two dimensions displaying such a phenomenon has been lacking so far. In this work, we apply the recently developed extremely correlated Fermi liquid theory (ECFL) [24,25] to calculate the Raman cross sections using the  $k$ -dependent bare vertices of Refs. [18,19]. This theory provides a framework for controlled calculations in the  $t$ - $J$  model, a prototypical model for very strong correlations, and a limiting case of the Hubbard model. The theory has been successfully benchmarked against essentially exact results in  $d = 0$  [26],  $d = 1$  [27], as well as  $d = \infty$  [28]. A recent application of the theory to the physically important case of  $d = 2$  in Refs. [29,30] gives detailed results for the spectral functions and the resistivity  $\rho$  in the  $t$ - $t'$ - $J$  model, with nearest- and second-neighboring hopping. The state obtained in ECFL at low hole densities has a very small quasiparticle weight  $Z \ll 1$ . A significant result is that the temperature dependence of resistivity is nonquadratic already at  $T \sim 100$  K for low hole doping.

In this work, we apply the solution found in Refs. [29,30] to compute the Raman scattering, in three standard polarization configuration channels  $A_{1g}$ ,  $B_{1g}$ ,  $B_{2g}$  defined below [31]. The results are applicable to either electron doping or hole-doped cuprates by choosing the sign of  $t'$ , and they may apply to other strongly correlated systems as well. Following SS, we also compare the Raman conductivities with the optical conductivity, and we shall focus on the quartet of these results on various values of material parameters.

The utility of comparing the optical conductivity with the Raman response requires a comment. SS [18,19] suggested that this comparison is useful, since these are exactly related in a limiting situation of  $d = \infty$ . Further, in  $d = 2, 3, \dots$  one often calculates the response within the bubble diagrams, where again these are related. In the bubble approximation, also used in the present work, one evaluates the current-current and related correlation functions by retaining only the lowest-order  $\chi_{JJ} \sim \sum_k (\gamma_k)^2 G(k)G(k)$  (i.e., bubble) terms with dressed Green's functions and suitable bare vertices  $\gamma$ . While this calculation misses a contribution due to the renormalization of one of the bare vertices  $\gamma \rightarrow \Gamma$ , it is hard to improve on

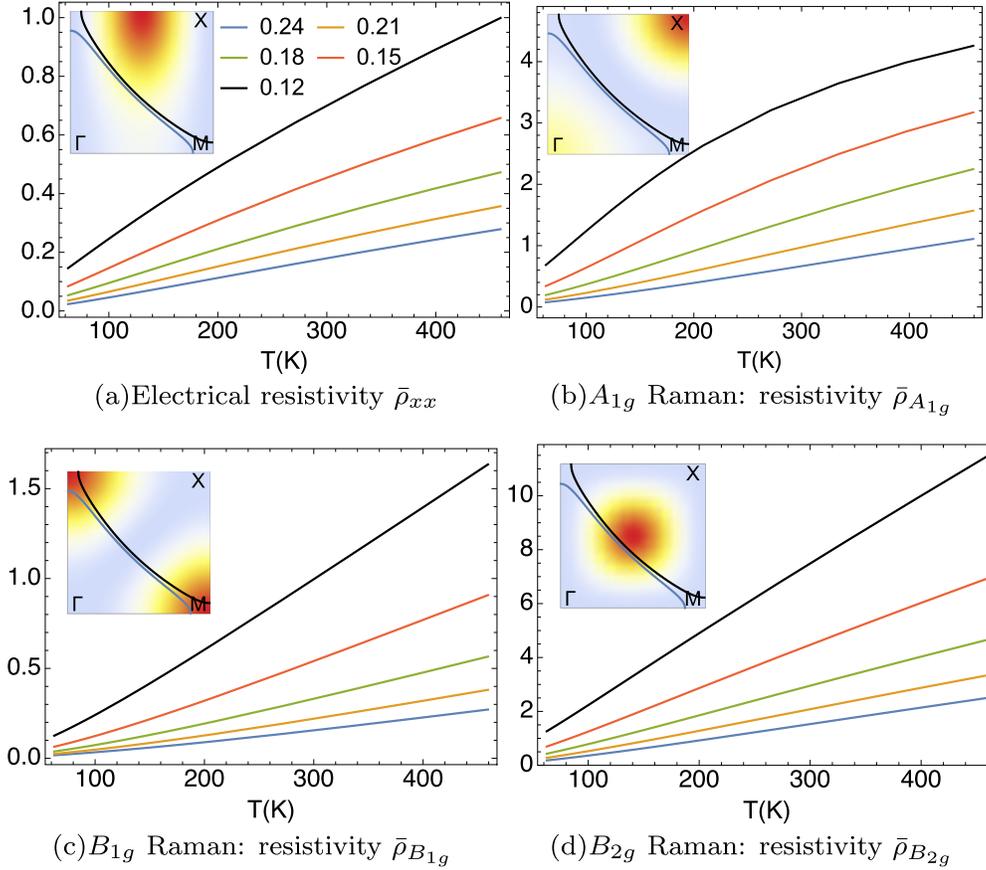


FIG. 1. Electrical and Raman resistivities from Eq. (5) at  $t' = -0.2$  with varying hole doping  $\delta$ , as marked. The  $T$  dependence of electrical resistivity and the  $A_{1g}$  resistivity are concave-down at small  $\delta$ , while the  $B_{1g}$  and  $B_{2g}$  resistivity are flat or concave-up. Inset: The displayed Fermi surfaces at  $\delta = 0.12, 0.24$  locate the maxima of  $\Upsilon(k, \omega)$ . The relevant squared vertices from Eq. (2) are shown as a heat map. The hot spots are movable by varying  $t'$  and  $t$ .

this already difficult calculation for strong correlations, since  $G$  is highly nontrivial. An exception is the special case of  $d \rightarrow \infty$ , where the vertex corrections vanish. Within the bubble scheme, the bare Raman and current vertices are different while everything else is the same. Therefore, one should be able to relate the two experimental results and explore the differences arising from the bare vertices. The ‘‘pseudo-identity’’ of the transport and Raman resistivities has been explored experimentally in Ref. [8] and finds some support. In this work, we use the correct bare vertices in the different geometries to explore the various Raman resistivities to refine the theory. These different bare vertices have a different dependence on the hopping parameters  $t, t'$ , and the calculations reflect these in specific and experimentally testable ways.

The neglect of vertex corrections also leads to a relationship between various Raman susceptibilities at finite  $\omega$ . In the experiments of Ref. [6], the same quartet of susceptibilities has been studied and found to have a roughly similar scale for their  $\omega$  dependence, although the curve shapes are distinct. On the theoretical side, one interesting aspect of the results of Refs. [29,30] is that the Fermi surface *shape* remains very close to that of the noninteracting tight-binding model, while of course conserving the area. Thus the Dyson self-energy is a weak function of  $\vec{k}$ , unlike the strong dependence in one

dimension [27]. This fact implies that the vertex corrections, while nonzero, are modest.

## II. THE RAMAN AND CURRENT VERTICES

We use the  $t$ - $t'$ - $J$  model with a tight-binding dispersion [29] on the square lattice  $\varepsilon(k) = -2t[\cos(k_x) + \cos(k_y)] - 4t' \cos(k_x) \cos(k_y)$ , and we set the lattice constant  $a_0 \rightarrow 1$ . The photons modulate the Peierls hopping factors as  $t_{ij} \rightarrow t_{ij} \exp\{ie/\hbar \int_i^j d\vec{r} \cdot \vec{A}\}$ , and the second-order expansion coefficients define the scattering operators. In this case, they are

$$\hat{\mathcal{J}}_{\alpha,q} = \sum_{k\sigma} \mathcal{J}_{\alpha}(k) C_{k+\frac{1}{2}q,\sigma}^{\dagger} C_{k-\frac{1}{2}q,\sigma}, \quad (1)$$

where  $\alpha$  is a composite index determined by the in-out polarizations of the photon. With that the vertices  $\mathcal{J}_{\alpha}$  for the three main Raman channels are

$$\begin{aligned} A_{1g} : \mathcal{J}_{A_{1g}}(k) &= 2t(\cos k_x + \cos k_y) + 4t' \cos k_x \cos k_y, \\ B_{1g} : \mathcal{J}_{B_{1g}}(k) &= 2t(\cos k_x - \cos k_y), \\ B_{2g} : \mathcal{J}_{B_{2g}}(k) &= -4t' \sin k_x \sin k_y, \\ xx : \mathcal{J}_{xx}(k) &= 2 \sin k_x (t + 2t' \cos k_y). \end{aligned} \quad (2)$$

The definition of  $\alpha = xx$  corresponds to the particle current along  $x$ . It integrates the charge current into the same scheme

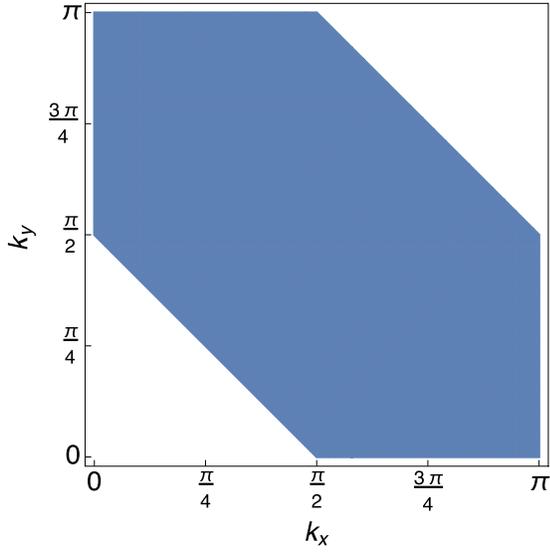


FIG. 2. Shaded region for estimating the average scale of the vertices.

as the Raman scattering. It is interesting that the  $B_{2g}$  vertex is independent of  $t$ , and is solely governed by  $t'$ . The vertex  $B_{1g}$  is complementary given its independence of  $t'$ . These geometries

sample different parts of  $k$  space in interesting ways due to their different  $\vec{k}$  dependences.

We next define the calculated variables, and we display the results for them from computations based on the spectral functions found in Refs. [29,30]. Results in the  $\omega = 0$  dc limit and also at finite  $\omega$  are shown. Finally, we discuss the results and their significance.

### III. RAMAN AND CHARGE SUSCEPTIBILITIES

We summarize the formulas for the (nonresonant) Raman susceptibility, and in the spirit of Refs. [18,19] we also define a *Raman conductivity and resistivity* in analogy as follows:

$$\chi_\alpha(q, z) = \sum_{nm} \frac{p_n - p_m}{\epsilon_m - \epsilon_n - z} |(\widehat{\mathcal{J}}_{\alpha,q})_{n,m}|^2, \quad (3)$$

where  $p_n$  is the probability of the state  $n$ . For visible light,  $qa_0 \ll 1$  and therefore we set  $q \rightarrow 0$ . The (nonresonant) Raman intensity  $\mathcal{I}_\alpha$  [1–3,18,19] and the Raman conductivities [18,19] are given by

$$\mathcal{I}_\alpha(0, \omega) = \frac{\chi_\alpha''(0, \omega)}{(1 - e^{-\beta\omega})}, \quad \sigma_\alpha(\omega) = \zeta_\alpha \frac{\chi_\alpha''(0, \omega)}{N_s \omega}, \quad (4)$$

with  $N_s$  the number of sites, and  $\zeta_{xx} = e^2$  accounting for the electric charge in the conductivity with all other  $\zeta_\alpha = 1$ . In the

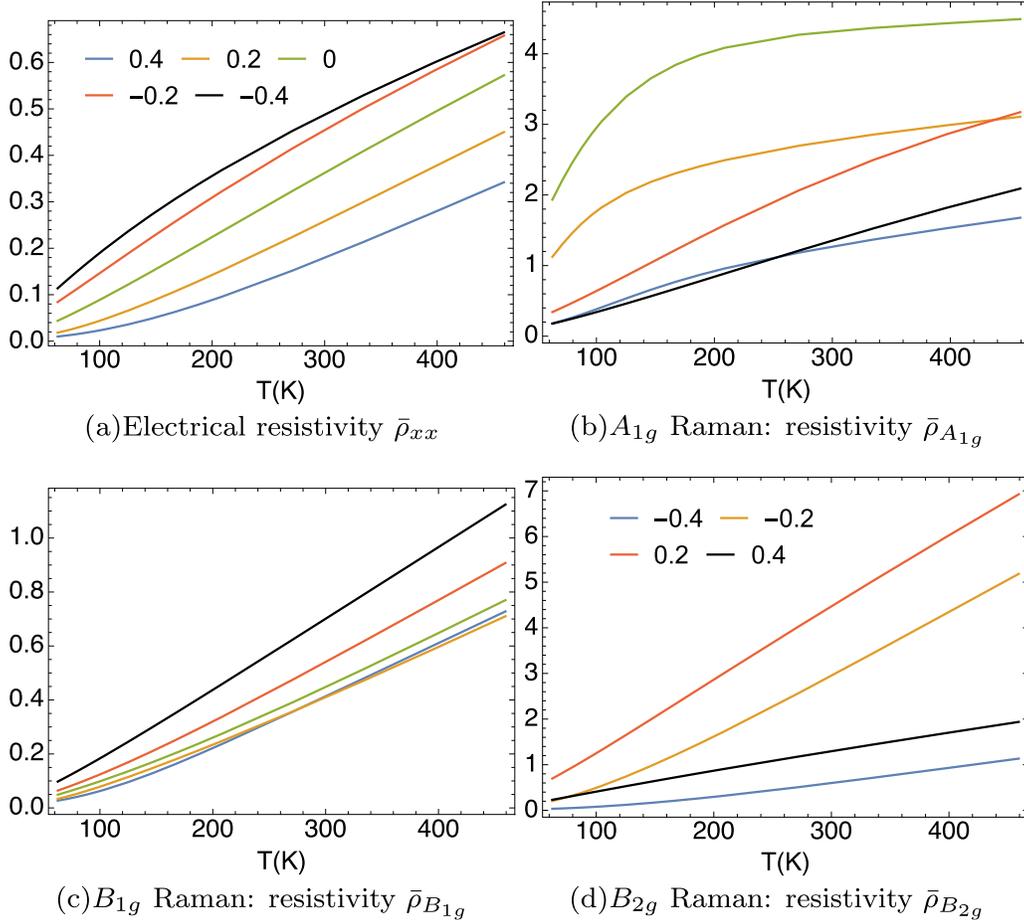


FIG. 3. Dimensionless  $\bar{\rho}_{xx}$  (taken from Ref. [29]),  $\bar{\rho}_{A_{1g}}$ ,  $\bar{\rho}_{B_{1g}}$ , and  $\bar{\rho}_{B_{2g}}$  at  $\delta = 0.15$  with varying second-neighbor hopping  $t'$ , as marked (same legend for all subfigures). Reference [8] displays data corresponding to the  $B_{1g}$  geometry.

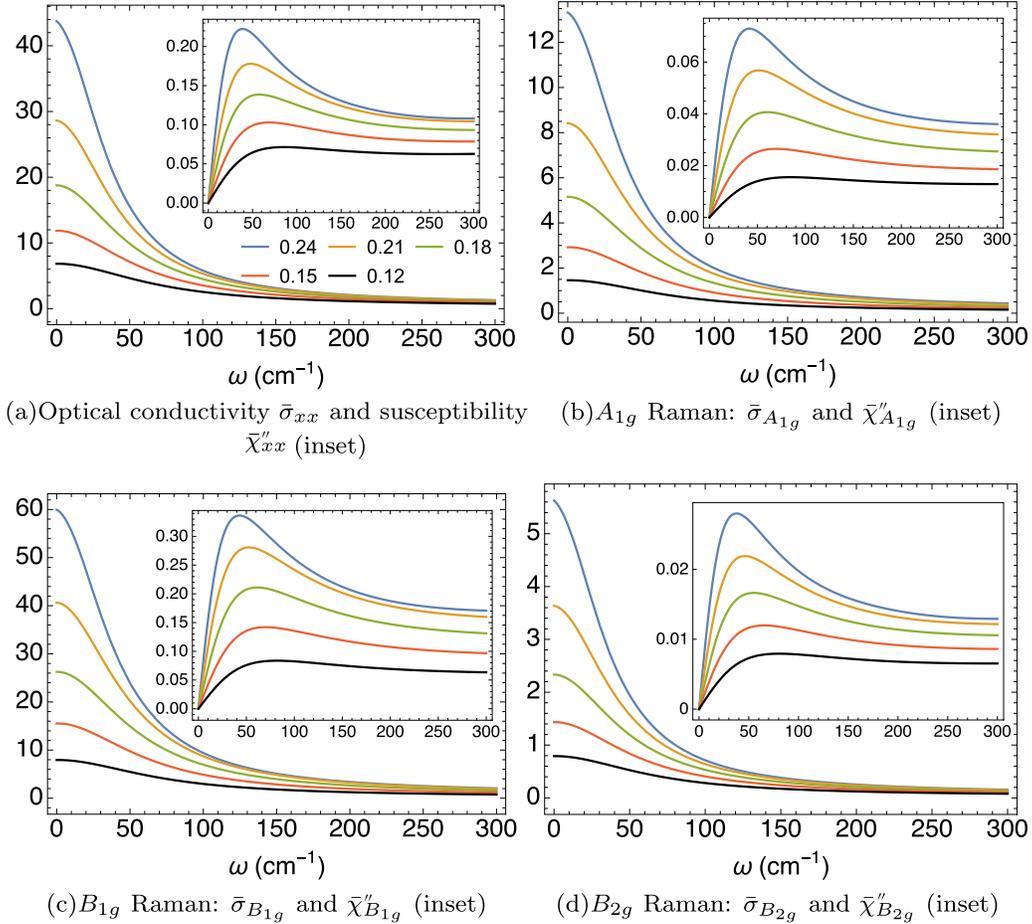


FIG. 4. Dynamical conductivities  $\bar{\sigma}_\alpha$  and susceptibilities  $\bar{\chi}''_\alpha$  (inset) for the hole-doped case  $t' = -0.2$ ,  $T = 63$  K at different  $\delta$ , as marked. In the experiments in Ref. [6] (Fig. 1), the same quartet of results is shown for LSCO. At the highest energy of over 1000 K, as in the data, the susceptibility shows no sign of dropping off.

dc limit we define the Raman resistivities

$$\rho_\alpha(0) = \frac{N_s}{\zeta_\alpha} \frac{k_B T}{\mathcal{I}_\alpha(0, 0)}, \quad (5)$$

where for  $\alpha = xx$ ,  $\rho_\alpha$  is the usual resistivity.

The ‘‘pseudo-identity,’’ a statement of universality relating electrical transport and the dc limit of Raman intensities noted by SS in Refs. [18, 19], is arrived at if we assume that  $\rho_\alpha$  has a similar  $T$  dependence for all  $\alpha$ :  $\mathcal{I}_\alpha(0, 0) \sim C_\alpha \frac{T}{\rho_{xx}(T)}$ , where  $C_\alpha$  is an  $\alpha$ -dependent constant. Thus  $\rho \sim T^\sigma$  behavior would give rise to  $T^{1-\sigma}$  behavior for the Raman intensity in all channels. We see in Fig. 1 that this suggestion is true for the  $A_{1g}$  resistivity at hole dopings, but it needs to be adjusted to the different  $k$ -dependent filters that make the  $B_{1g}$  and  $B_{2g}$  channels different from the others. Thus we limit the universality of the pseudo-identity in this work, and we quantify the effects of the bare vertices in the relationship between the members of the quartet of susceptibilities.

Proceeding further using the bubble scheme we get the imaginary part of the dimensionless susceptibility  $\bar{\chi}''_\alpha(0, \omega) \equiv \frac{c_0 \hbar}{N_s} \chi''_\alpha(0, \omega)$  as

$$\bar{\chi}''_\alpha(0, \omega) = \omega \langle \Upsilon(k, \omega) \mathcal{J}_\alpha^2(k) \rangle_k, \quad (6)$$

where  $c_0 \sim 6.64 \text{ \AA}$  is a typical interlayer separation [29]. The angular average is  $\langle A \rangle_k \equiv \frac{1}{N_s} \sum_k A(k)$  and the momentum resolved relaxation scale is

$$\begin{aligned} \Upsilon(k, \omega) &= \frac{4\pi^2}{\omega} \int_{-\infty}^{\infty} dy \rho_G(k, y) \rho_G(k, y + \omega) [f(y) - f(y + \omega)]. \end{aligned}$$

Here  $\rho_G(k, \omega)$  is the electron spectral function. With  $\rho_{1,\alpha} \equiv \frac{c_0 \hbar}{\zeta_\alpha}$ , the corresponding dimensionless conductivity  $\bar{\sigma}_\alpha(\omega) \equiv \rho_{1,\alpha} \times \sigma_\alpha(\omega)$  is given by

$$\bar{\sigma}_\alpha(\omega) = \langle \Upsilon(k, \omega) \mathcal{J}_\alpha^2(k) \rangle_k. \quad (7)$$

From Eqs. (6) and (7), we can see  $\bar{\chi}''_\alpha(0, \omega) = \omega * \bar{\sigma}_\alpha(\omega)$

#### IV. PARAMETER REGION

We explore how the variation of second-neighbor hopping  $t'$ , doping  $\delta$ , and temperature  $T$  affects the quartet of conductivities and susceptibilities in the normal state. We focus on optimal doping or slightly overdoped cases from electron-doped (positive  $t'$ ) to hole-doped (negative  $t'$ ) systems. Our temperature region starts from 63 K to a few hundred degrees Kelvin.

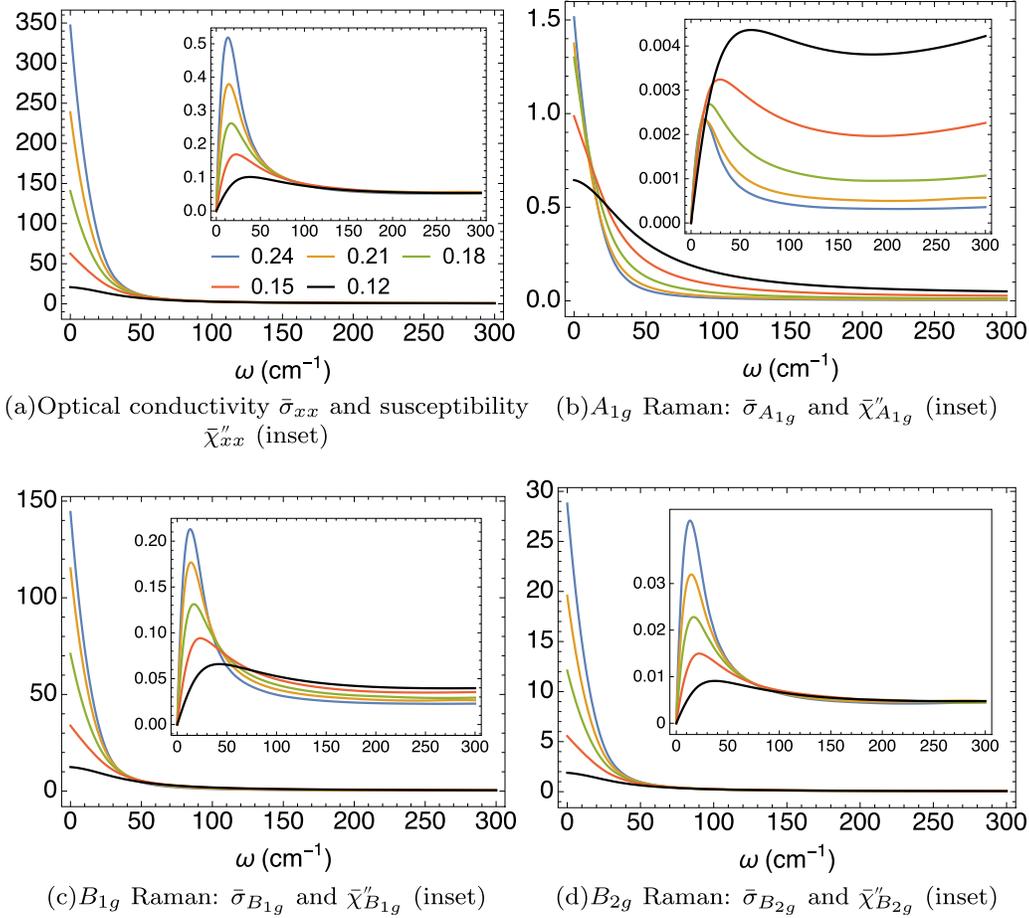


FIG. 5. Dynamical conductivities  $\bar{\sigma}_\alpha$  and susceptibilities  $\bar{\chi}''_\alpha$  (inset) for the electron-doped case  $t' = 0.2$ ,  $T = 63$  K at different  $\delta$ , as marked.

## V. DC LIMIT AND ELECTRICAL RESISTIVITY RESULTS:

Using the spectral function from the second-order ECFL theory, we calculate the dimensionless dc ( $\omega \rightarrow 0$ ) electrical and Raman conductivities  $\bar{\sigma}_\alpha$  from Eq. (7). The corresponding dimensionless resistivities are

$$\bar{\rho}_\alpha = \frac{1}{\bar{\sigma}_\alpha} = \frac{1}{\langle \Upsilon(k, 0) \mathcal{J}_\alpha^2(k) \rangle_k}. \quad (8)$$

The electrical resistivity in physical units is given by  $\rho_{xx} = \bar{\rho}_{xx} \times \rho_{1,xx}$ , with  $\rho_{1,xx} = c_0 \frac{h}{e^2} \sim 1.71$  m  $\Omega$  cm [29].

We calculate typical quantities for the three Raman geometries and the electrical conductivity from Eq. (2) as a set of quartets below. The comparison of the figures in each set is of interest, since the different functions in the bare vertices pick out different parts of the  $k$ -space. In this paper,  $t = 1$  serves as the energy unit; for the systems in mind, we estimate [29]  $t \sim 0.45$  eV.

In Fig. 1, we plot dc resistivity  $\bar{\rho}_{xx}$  and Raman resistivities in the dc limit  $\bar{\rho}_{A_{1g}}$ ,  $\bar{\rho}_{B_{1g}}$ ,  $\bar{\rho}_{B_{2g}}$  varying hole doping  $\delta$  and fixing  $t' = -0.2$ . The four figures have a roughly similar doping dependence, as suggested by the pseudo-identity. They all decrease when the doping increases, although the curvature changes more in  $\bar{\rho}_{xx}$  and  $\bar{\rho}_{A_{1g}}$  than the other two cases. This can be understood from Eq. (6) since they arise from the same kernel  $\Upsilon(k, 0)$  with different filters. The quasiparticle peak in  $\rho_G$ ,

contributing most to  $\Upsilon(k, 0)$ , is located along the Fermi surface and gets broadened when warming up. The inset shows the corresponding squared vertex  $\mathcal{J}_\alpha^2$  in the background and the Fermi surfaces at the lowest and highest dopings. The  $B_{1g}$  vertex vanishes along the line  $k_x = k_y$  while the  $B_{2g}$  vertices vanish near  $\{\pi, 0\}$  and  $\{0, \pi\}$  points. In our calculation, both  $B_{1g}$  and  $B_{2g}$  overlap well with the peak region of the spectral function, whereas  $A_{1g}$  and the resistivity do not. This results in the difference between the  $T$  dependence of them and the other two in Fig. 1. It would be of considerable interest to study this pattern of  $T$  dependences systematically in future Raman studies.

Although all  $\bar{\rho}_\alpha$  increase when reducing doping  $\delta$  approaching the half-filling limit due to the suppression of quasiparticles, their magnitudes at high temperature vary considerably as a result of different vertices filtering the contribution from  $\Upsilon(k, 0)$ . We can understand this scale difference by evaluating the average of vertices over the shaded region in Fig. 2. The shaded region covers the Fermi surface for all chosen  $\delta$  and  $t'$ , and therefore it contains the most significant contribution to  $\rho_\alpha$ .

At  $t' = -0.2$ ,  $\langle \mathcal{J}_{xx}^2 \rangle_s \approx 2.41$ ,  $\langle \mathcal{J}_{A_{1g}}^2 \rangle_s \approx 0.56$ ,  $\langle \mathcal{J}_{B_{1g}}^2 \rangle_s \approx 1.30$ ,  $\langle \mathcal{J}_{B_{2g}}^2 \rangle_s \approx 0.20$ , where  $\langle \rangle_s$  represents the  $k$  average over the shaded region. They not only explain the relation  $\bar{\rho}_{xx} < \bar{\rho}_{B_{1g}} < \bar{\rho}_{A_{1g}} < \bar{\rho}_{B_{2g}}$ , but they also capture the ratio among them rather closely at high enough  $T$ . The structure at low  $T$  is more

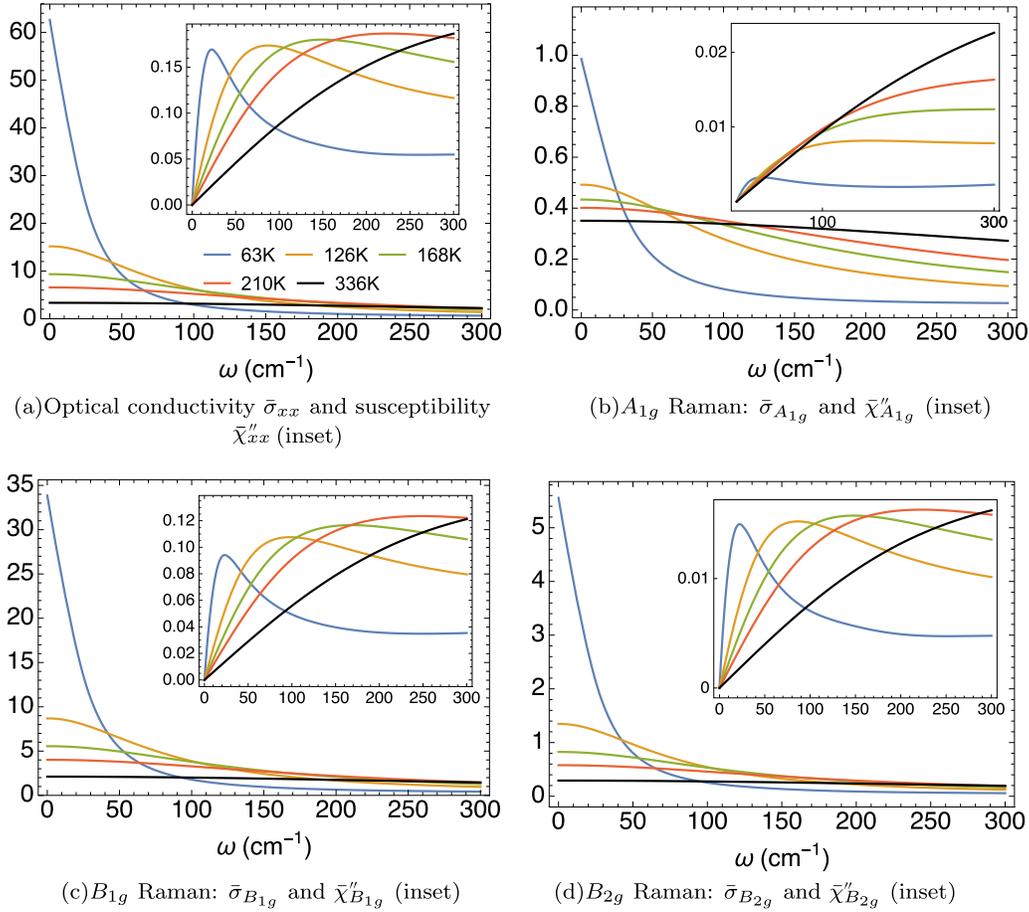


FIG. 6. Dynamical conductivities and (inset) susceptibilities for the electron-doped case with  $t' = 0.2$ ,  $\delta = 0.15$  for various  $T$ 's as marked. Part (d) with  $B_{2g}$  symmetry is comparable to the high-resolution experimental result in Fig. 2 of Ref. [10] at a comparable set of  $T$ 's. The theoretical curve reproduces well the quasielastic peaks and their  $T$  evolution.

subtle, and it carries information about the magnitude of  $t'$  that cannot be captured by the above high- $T$  argument.

Although all  $\bar{\rho}_\alpha$  increase as  $\delta$  decreases in general, their  $t'$  dependence can be rather different, as shown in Fig. 3.  $\bar{\rho}_{xx}$  and  $\bar{\rho}_{B_{1g}}$  decrease monotonically in general as  $t'$  increases from hole-doped (negative) to electron-doped (positive), while  $\bar{\rho}_{A_{1g}}$  and  $\bar{\rho}_{B_{2g}}$  decrease only as  $|t'|$  increases and their monotonicity with respect to  $t'$  changes upon sign change of  $t'$ . Another interesting observation is that  $\bar{\rho}_\alpha(t' = -0.2) > \bar{\rho}_\alpha(t' = 0.2)$  and  $\bar{\rho}_\alpha(t' = -0.4) > \bar{\rho}_\alpha(t' = 0.4)$  are generally true for  $\alpha = xx, B_{1g}$  and  $B_{2g}$ , but for the  $A_{1g}$  case,  $\bar{\rho}_\alpha(t' = -0.2) < \bar{\rho}_\alpha(t' = 0.2)$  in general and  $\bar{\rho}_\alpha(t' = -0.4) \approx \bar{\rho}_\alpha(t' = 0.4)$ .

In Eq. (8), the resistivities depend on  $t'$  through  $\Upsilon(k, 0)$  and  $\mathcal{J}_\alpha^2$ . To estimate their  $t'$  dependence, we can look at their average over the shaded region  $\langle \Upsilon(k, 0) \rangle_s$  and  $\langle \mathcal{J}_\alpha^2 \rangle_s$ . While  $\langle \Upsilon(k, 0) \rangle_s$  rises monotonically as  $t'$  increases,  $\langle \mathcal{J}_\alpha^2 \rangle_s$  ( $\alpha = xx, A_{1g}, B_{2g}$ ) is a quadratic function of  $t'$  that behaves differently at positive and negative  $t'$ , as shown in Eq. (2).

In the simplest  $B_{1g}$  case,  $\mathcal{J}_{B_{1g}}^2$  is independent of  $t'$ . Then  $t'$  only affects  $\bar{\rho}_{B_{1g}}$  through  $\Upsilon(k, 0)$  and therefore  $\bar{\rho}_{B_{1g}}$  increases almost monotonically as  $t'$  decreases (the crossing between  $t = 0.2$  and  $0.4$  is due to the fact that the change in Fermi surface geometry leads to a different filtering result when coupling to  $\mathcal{J}_{B_{1g}}^2$ ). In the charge-current case, the  $t'$  dependence

of  $\Upsilon(k, 0)$  still dominates since  $\bar{\rho}_{xx}$  behaves similarly to  $\bar{\rho}_{B_{1g}}$  and the contribution from  $\mathcal{J}_{xx}^2$  mostly modifies the curvature without affecting the relative scale.

The different behaviors in the other two cases indicate the quadratic  $t'$  dependence in  $\mathcal{J}_\alpha^2$  ( $\alpha = A_{1g}, B_{2g}$ ) becomes dominant. In the simpler  $B_{2g}$  case,  $\mathcal{J}_{B_{2g}}^2 \propto t'^2$  provides the dominant  $t'$  dependence in  $\bar{\rho}_{B_{2g}}$ , explaining  $\bar{\sigma}_{B_{2g}}(t' = 0) = 0$  and  $\bar{\rho}_{B_{2g}}(|t'| = 0.2) > \bar{\rho}_{B_{2g}}(|t'| = 0.4)$  regardless of the sign of  $t'$ . Similarly, due to the quadratic  $t'$  dependence of  $\mathcal{J}_{A_{1g}}^2$ ,  $\bar{\rho}_{A_{1g}}(t' = 0) > \bar{\rho}_{A_{1g}}(|t'| = 0.2) > \bar{\rho}_{A_{1g}}(|t'| = 0.4)$ .

Typically negative  $t'$  leads to stronger correlation and suppresses the quasiparticle peak [29] and hence for a certain  $|t'|$ ,  $\bar{\rho}_\alpha(t' < 0) > \bar{\rho}_\alpha(t' > 0)$  is generally true except for the  $A_{1g}$  case. In this exception, the negative linear  $t'$  term in  $\mathcal{J}_{A_{1g}}^2$  shifts the stationary point away from  $t' = 0$  and counters this effect from  $\Upsilon(k, 0)$  for small  $|t'|$  leading to  $\bar{\rho}_{A_{1g}}(t' = -0.2) < \bar{\rho}_{A_{1g}}(t' = 0.2)$  and  $\bar{\rho}_{A_{1g}}(t' = -0.4) \approx \bar{\rho}_{A_{1g}}(t' = 0.4)$ .

Besides,  $\bar{\rho}_{A_{1g}}$  shows rather different  $T$ -dependent behaviors between electron-doped  $t' \geq 0$  and hole-doped  $t' < 0$  cases. At negative  $t'$ ,  $\bar{\rho}_{A_{1g}}$  increases almost linearly with temperature. But at zero or positive  $t'$ ,  $\bar{\rho}_{A_{1g}}$  first increases sharply up to a certain temperature scale depending on  $t'$  and then crosses over to a region where the growth rate becomes much smaller.

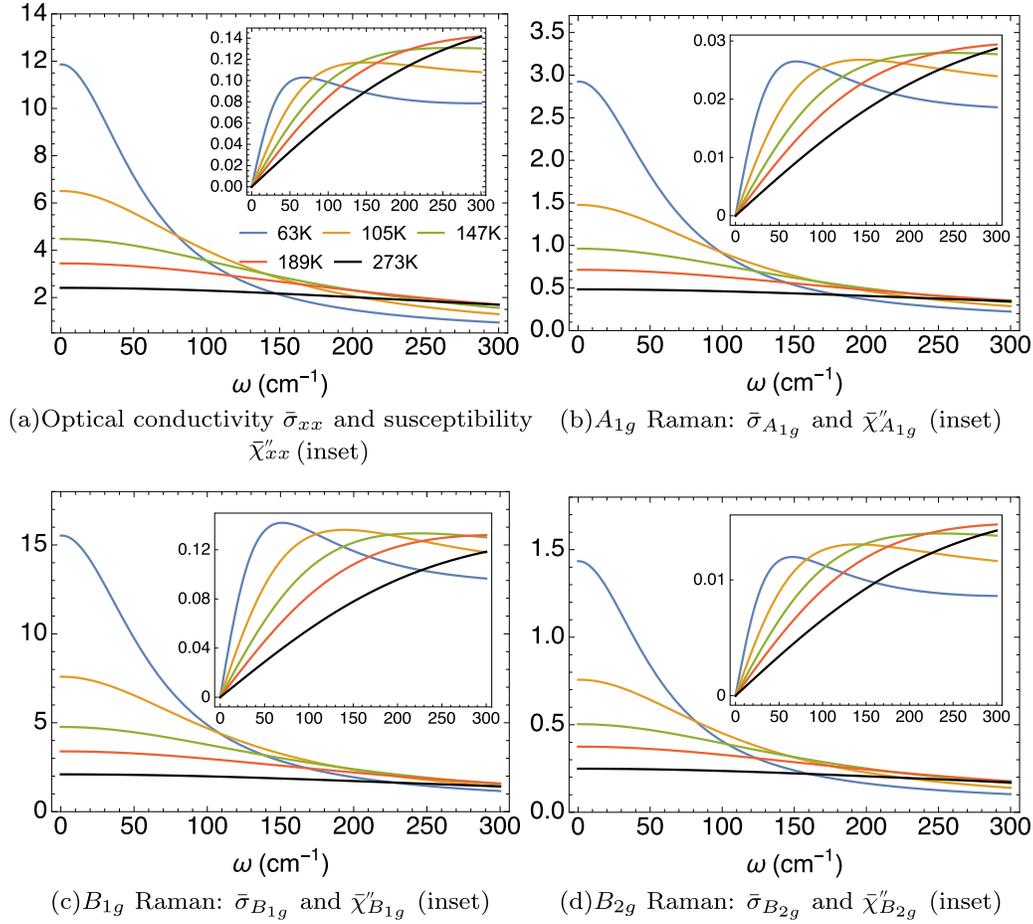


FIG. 7. Optical conductivity  $\bar{\sigma}_{xx}(\omega)$  and the Raman conductivities  $\bar{\sigma}_{A_{1g}}(\omega)$ ,  $\bar{\sigma}_{B_{1g}}(\omega)$ ,  $\bar{\sigma}_{B_{2g}}(\omega)$  at  $t' = -0.2$ ,  $\delta = 0.15$  and varying  $T$ , as marked (same legend for all subfigures). The corresponding dimensionless susceptibility is plotted in the inset with the same  $x$  axis. References [6,7,10] show data that correspond to these variables.

## VI. FINITE $\omega$ RESULTS

Next we present the  $\omega$ -dependent optical and Raman conductivities defined in Eq. (7). In Figs. 4 and 5, the set of four  $\omega$ -dependent conductivities are displayed for the hole-doped system at  $t' = -0.2$  and the electron-doped system at  $t' = 0.2$ , respectively, for a set of typical densities at low  $T$ . In the insets we display the corresponding imaginary part of susceptibility,

related through Eq. (6). In most cases, the quasielastic peak gets suppressed and shifts to higher frequency when reducing the carrier concentration. The only exception is  $\bar{\chi}''_{A_{1g}}$  at  $t' = 0.2$ . Its quasielastic peaks are considerably smaller than other geometries due to the fluctuation in the specific vertex, and they get higher and broader as doping increases.

In Fig. 6 we focus on the electron-doped case of varying  $T$  at  $t' = 0.2$ ,  $\delta = 0.15$ , where high-quality experimental results

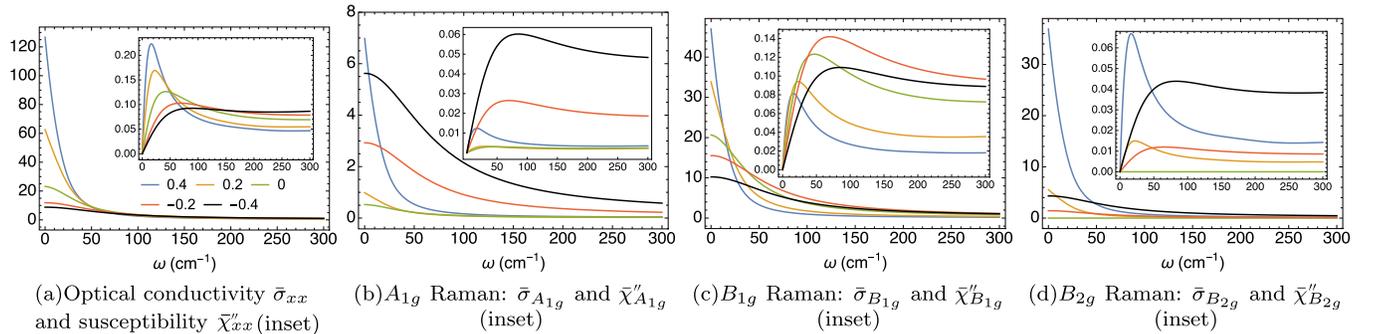


FIG. 8. Optical conductivity  $\bar{\sigma}_{xx}(\omega)$  and the Raman conductivities  $\bar{\sigma}_{A_{1g}}(\omega)$ ,  $\bar{\sigma}_{B_{1g}}(\omega)$ ,  $\bar{\sigma}_{B_{2g}}(\omega)$  at  $\delta = 0.15$ ,  $T = 63$  K, and varying  $t'$ , as marked (same legend for all subfigures). The corresponding dimensionless susceptibility is plotted in the inset with the same  $x$  axis. References [6,7,10] show data that correspond to these variables.

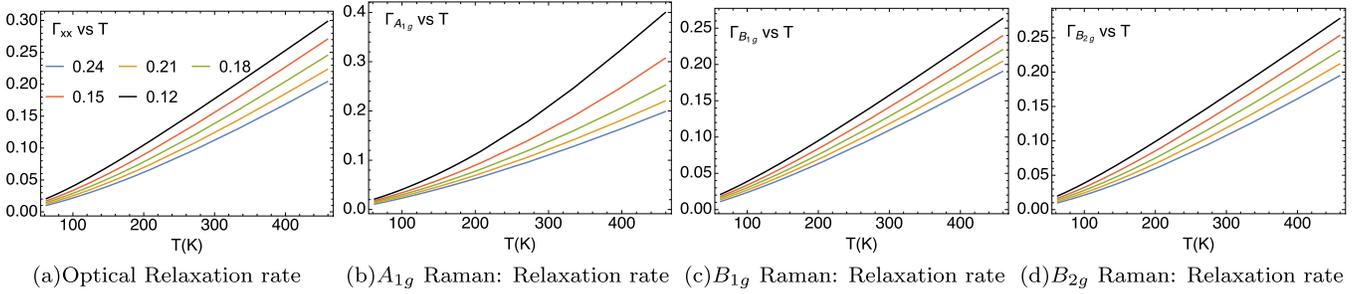


FIG. 9. Relaxation rates (half-widths at half-maximum) of  $\sigma_\alpha(\omega)$  in units of  $t$ , at  $t' = -0.2$  at various marked  $\delta$ . The optical rate shows less convexity than the corresponding dc resistivity of Ref. [29]. The rates in (a,b) and (c,d) have similar orders of magnitude, for reasons discussed in Fig. 1.

are available for the  $B_{2g}$  Raman channel in Ref. [10]; see particularly Fig. 2. We evaluate the susceptibility at  $T$  values corresponding to those in this experiment. There is a fair similarity between the theoretical curve [panel (d)] and the experiment. In particular, the theoretical curve reproduces the quasielastic peak and its  $T$  evolution. The other three panels in Fig. 6 are our theoretical predictions, and they are equally amenable to experimental verification.

In the  $xx$ ,  $B_{1g}$ ,  $B_{2g}$  geometries, the quasielastic peaks in susceptibility get slightly higher and quite broader upon warming. The  $A_{1g}$  case is different. Its quasielastic peaks are much less obvious (too broad) except for the lowest temperature, and the peak magnitude is rather sensitive to temperature increase.

We also vary  $T$  at hole doping  $t' = -0.2$  in Fig. 7. Comparing with the electron-doped case in Fig. 6, we note that the hole-doped optical and Raman objects share a greater similarity in shape dependence on  $T$  if we ignore the scale difference. As  $T$  increases, the quasiparticle peaks get softened, and hence it generally suppresses the conductivities as well as the quasielastic peak in susceptibilities.

For completeness, the  $t'$  variation in  $\bar{\sigma}_\alpha(\omega)$  and  $\bar{\chi}''_\alpha(\omega)$  is plotted in Fig. 8, and it looks rather different among various geometries. This can be understood as arising from the competition among various factors. We have a quadratic  $t'$  dependence in the squared vertices, and a monotonic  $t'$  dependence in the magnitude and geometry of  $\Upsilon(k, \omega)$ . The  $t'$  dependence of the shape of  $\bar{\sigma}_\alpha$  has more commonality. Another interesting observation is that, unlike the dc case when  $\bar{\sigma}_{xx}$  and  $\bar{\sigma}_{B_{1g}}$  are similarly affected by  $t'$ , at finite frequency their behaviors depend on  $t'$  rather differently. This difference is more obviously observed in terms of  $\bar{\chi}''$ .

From the optical and Raman conductivities  $\bar{\sigma}_\alpha$  we can extract a frequency scale  $\Gamma_\alpha$ , as the half-width at half-maximum,

in units of  $t$ . These are plotted against  $T$  in Fig. 9 for varying  $\delta$  and Fig. 10 for varying  $t'$ . It is remarkable that despite a bare bandwidth of  $\sim 3.6$  eV, these frequency scales appear close to linear in  $T$  down to very low  $T$ . This is closely related to the observation in Ref. [29] that the resistivity departs from a  $T^2$  behavior at extraordinarily low  $T$ 's, i.e., the effective Fermi temperatures are suppressed from the bare values by two or more orders of magnitude. Although the magnitude of the optical and Raman conductivities differs a lot, their relaxation rates describing the shape turn out to be much closer, as a result of a similar  $T$ -dependent line shape of the spectral function [29] in the normal state.

## VII. CONCLUSION AND DISCUSSION

We have presented calculations of the electrical and Raman resistivities in the dc limit, the optical conductivity, the Raman susceptibilities, and related objects based on the second-order ECFL theory in Ref. [29]. We computed the susceptibilities (using the leading-order approximation) with the shown results. Experiments on different geometries can test and put some bound on this hypothesis of weak vertex corrections for the Raman operators. This is clearly of theoretical importance, since going beyond the bubble graphs brings in a formidable level of complexity.

The ECFL theory leads to a very small quasiparticle weight  $Z$  and a large background extending over the bandwidth, and it has a very small effective Fermi temperature leading to an interesting  $T$  dependence of the resistivity, as discussed in [29]. The line shape of the calculated Raman susceptibility is close to that for the case of electron-doped NCCO [10] in terms of the  $T$  and  $\omega$  dependences, and therefore it is promising. Our calculation also gives the Raman susceptibility in two other

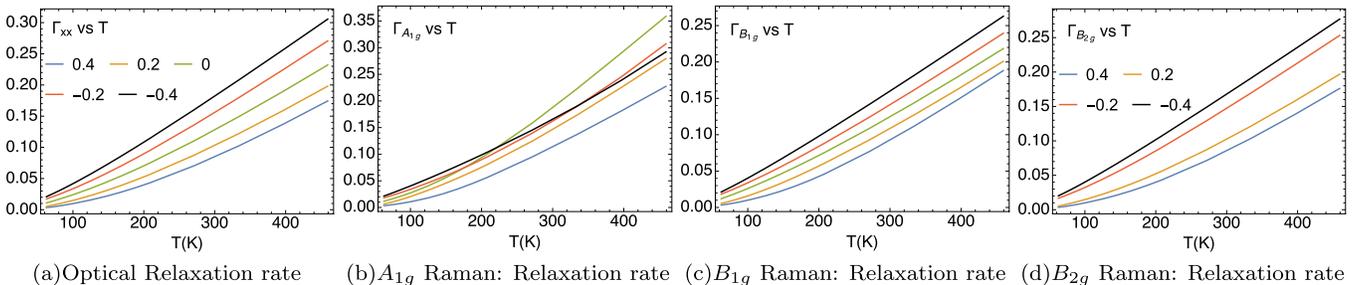


FIG. 10. The half-width at half-maximum for optical conductivity and Raman conductivities at  $\delta = 0.15$  and varying  $t'$ , as marked.

geometries, and this prediction can be checked against future experiments that are quite feasible. We note that the data [6] from Sugai *et al.* for this quartet of variables in the case of LSCO seem to be roughly consistent with our results, and a more detailed comparison is planned.

The focus on the  $T$  dependence in the  $\omega \rightarrow 0$  limit, i.e., on resistivities, can be quite a fruitful goal for future experiments, since this limit gets rid of all excitations and measures the “pure background.” It is an important exercise since the different geometries probe different combinations of  $t$ ,  $t'$  as they occur in the bare vertices [Eq. (2), as stressed above. We are predicting that the Raman resistivity in each channel can be found from the intensity at low  $T$ , and broadly speaking it is similar to resistivity. In further detail, it is predicted to be (a) channel-specific and (b)  $t'/t$ -dependent. These clear-cut predictions can be tested in future experiments.

Finally, although such a measurement is not commonly done, a systematic measurement of the ratios of the scattering

cross sections in different geometries should be feasible. This measurement, and a comparison between the quartet of susceptibilities presented here, can be profitably compared with recent theories of strongly correlated systems to yield material parameters. Most importantly, it can yield physical insights into the mechanism underlying the broad nonresonant Raman signals that have remained quite mysterious so far.

#### ACKNOWLEDGMENTS

We thank Tom Devereaux, Lance Cooper, and Girsh Blumberg for helpful discussions. The computation was done on the comet in XSEDE [32] (TG-DMR170044) supported by National Science Foundation Grant No. ACI-1053575. The work at UCSC was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, under Award No. DE-FG02-06ER46319.

- 
- [1] H. A. Kramers and W. Heisenberg, *Z. Phys.* **31**, 681 (1925).  
 [2] P. A. Wolff, *Phys. Rev.* **171**, 436 (1968).  
 [3] P. A. Wolff and P. Platzmann, *Solid State Physics* (Academic Press, New York, 1973), Suppl. 13; W. Hayes and R. H. Loudon, *Scattering of Light by Crystals* (Wiley, New York, 1978).  
 [4] A. A. Abrikosov and V. M. Genkin, *Sov. Phys. JETP* **38**, 417 (1974).  
 [5] S. Sugai, S. I. Shamoto, and M. Sato, *Phys. Rev. B* **38**, 6436 (1988).  
 [6] S. Sugai, Y. Takayanagi, N. Hayamizu, T. Muroi, J. Nohara, R. Shiozaki, K. Okazaki, and K. Takenaka, *Physica C* **470**, S97 (2010).  
 [7] S. Sugai, J. Nohara, R. Shiozaki, T. Muroi, Y. Takayanagi, N. Hayamizu, K. Takenaka, and K. Okazaki, *J. Phys.: Condens. Matter* **25**, 415701 (2013).  
 [8] R. Hackl, L. Tassinari, F. Venturini, C. Hartinger, A. Erb, N. Kikugawa, and T. Fujita, in *Advances in Solid State Physics*, edited by B. Kramer (Springer-Verlag, Berlin, Heidelberg, 2005), Vol. 45, pp. 227–238.  
 [9] M. M. Qazilbash, A. Koitzsch, B. S. Dennis, A. Gozar, H. Balci, C. A. Kendziora, R. L. Greene, and G. Blumberg, *Phys. Rev. B* **72**, 214510 (2005).  
 [10] A. Koitzsch, G. Blumberg, A. Gozar, B. S. Dennis, P. Fournier, and R. L. Greene, *Phys. Rev. B* **67**, 184522 (2003).  
 [11] C. Sauer and G. Blumberg, *Phys. Rev. B* **82**, 014525 (2010).  
 [12] M. V. Klein and S. B. Dierker, *Phys. Rev. B* **29**, 4976 (1984).  
 [13] M. V. Klein, S. L. Cooper, A. L. Kotz, R. Liu, D. Reznik, F. Slakey, W. C. Lee, and D. M. Ginsberg, *Physica C* **185**, 72 (1991); G. Blumberg and M. V. Klein, *J. Low Temp. Phys.* **117**, 1001 (1999).  
 [14] F. Slakey, S. L. Cooper, M. V. Klein, J. P. Rice, and D. M. Ginsberg, *Phys. Rev. B* **39**, 2781 (1989).  
 [15] S. L. Cooper, D. Reznik, A. Kotz, M. A. Karlow, R. Liu, M. V. Klein, W. C. Lee, J. Giapintzakis, D. M. Ginsberg, B. W. Veal, and A. P. Paulikas, *Phys. Rev. B* **47**, 8233 (1993).  
 [16] S. L. Cooper, F. Slakey, M. V. Klein, J. P. Rice, E. D. Bukowski, and D. M. Ginsberg, *J. Opt. Soc. Am. B* **6**, 436 (1989).  
 [17] S. L. Cooper, M. V. Klein, B. G. Pazol, J. P. Rice, and D. M. Ginsberg, *Phys. Rev. B* **37**, 5920 (1988).  
 [18] B. S. Shastry and B. I. Shraiman, *Phys. Rev. Lett.* **65**, 1068 (1990).  
 [19] B. S. Shastry and B. I. Shraiman, *Int. J. Mod. Phys. B* **5**, 365 (1991).  
 [20] J. K. Freericks and T. P. Devereaux, *Phys. Rev. B* **64**, 125110 (2001); J. K. Freericks, T. P. Devereaux, M. Moraghebi, and S. L. Cooper, *Phys. Rev. Lett.* **94**, 216401 (2005).  
 [21] T. P. Devereaux and R. Hackl, *Rev. Mod. Phys.* **79**, 175 (2007).  
 [22] L. de Medici, A. Georges, and G. Kotliar, *Phys. Rev. B* **77**, 245128 (2008).  
 [23] J. Kosztin and A. Zawadowski, *Solid State Commun.* **78**, 1029 (1991).  
 [24] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).  
 [25] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016).  
 [26] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *Phys. Rev. B* **88**, 205108 (2013).  
 [27] P. Mai, S. R. White, and B. S. Shastry, *Phys. Rev. B* **98**, 035108 (2018).  
 [28] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013); E. Perepelitsky and B. S. Shastry, *Ann. Phys. (N.Y.)* **338**, 283 (2013).  
 [29] B. S. Shastry and P. Mai, *New J. Phys.* **20**, 013027 (2018).  
 [30] P. Mai and B. S. Shastry, *arXiv:1808.09788* (2018).  
 [31] Parabolic bands leads to an exact cancellation, whereby the  $A_{1g}$  scattering is unobservable. References [3,4,18,19] show that the cancellation is inoperative with nonparabolic bands, or with disorder. Experimentally, Fig. 2 of Ref. [6] shows an  $A_{1g}$  contribution similar in scale to the other two geometries.  
 [32] J. Towns *et al.*, *Comput. Sci. Eng.* **16**, 62 (2014).

## Reversal of particle-hole scattering-rate asymmetry in the Anderson impurity model

R. Žitko,<sup>1,2</sup> H. R. Krishnamurthy,<sup>3</sup> and B. Sriram Shastry<sup>4</sup>

<sup>1</sup>*Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

<sup>2</sup>*Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia*

<sup>3</sup>*Department of Physics, Center for Condensed Matter Theory, Indian Institute of Science, Bengaluru 560012, India*

<sup>4</sup>*Physics Department, University of California, Santa Cruz, California 95064, USA*



(Received 30 July 2018; revised manuscript received 10 October 2018; published 29 October 2018)

We study the particle-hole asymmetry of the scattering rate in strongly correlated electron systems by examining the cubic  $\omega^3$  and  $\omega T^2$  terms in the imaginary part of the self-energy of the Anderson impurity model. We show that the sign is opposite in the weak-coupling and strong-coupling limits, explaining the differences found in theoretical approaches taking the respective limits as the starting points. The sign change in fact precisely delineates the crossover between the weak- and strong-correlation regimes of the model. For weak interaction  $U$  the sign reversal occurs for small values of the doping  $\delta = 1 - n$ , while for interaction of order  $U \approx 2\Gamma$ ,  $\Gamma$  being the hybridization strength, the crossover curve rapidly shifts to the large-doping range. This curve, based on the impurity dynamics, is genuinely different from other crossover curves defined through impurity thermodynamic and static properties.

DOI: [10.1103/PhysRevB.98.161121](https://doi.org/10.1103/PhysRevB.98.161121)

In contemporary strongly correlated quantum materials, such as the cuprate superconductors and sodium cobaltates, one finds that spectral line shapes from angle-resolved photoemission spectroscopy (ARPES) differ qualitatively from those in simple Fermi liquids. The origin of the difference has been traced to a large correlation-induced asymmetry in the imaginary part of the self-energy [1,2], which can be expanded as

$$\text{Im } \Sigma(\omega, T) = a(\omega^2 + \pi^2 T^2) + b\omega^3 + c\omega T^2 + \dots \quad (1)$$

For context, note that the usually quoted Fermi-liquid self-energy, namely, the first two terms in this expression, is *even* in  $\omega$ . While this is dominant at the lowest energies, the higher-order *odd* in  $\omega$  terms become important when their coefficients ( $b, c$ ) become sufficiently large. This is found to happen in the strong-correlation models, while in weakly correlated systems these coefficients are very small. The signs of ( $b, c$ ) are of particular importance: They determine whether particles or holes have the shorter lifetime. Since  $a < 0$ , if  $b < 0$ , the particlelike excitations scatter more strongly on the impurity (are more damped) than the holelike excitations with the same excitation energy (absolute value of  $\omega$ ), and vice versa for  $b > 0$ . Understanding the asymmetry of the self-energy is a problem of great current interest. The asymmetry of  $\text{Im } \Sigma$  is relevant to transport coefficients such as the thermopower, where the entropy and charge are carried by both particlelike excitations above the Fermi level and holelike excitations below it. However, in the low- $T$  thermopower there are other competing factors (asymmetry of the density of states, asymmetry of the quasiparticle velocities), hence the situation is not solely controlled by the sign of the scattering asymmetry.

The single-impurity Anderson model (SIAM) is a “laboratory example” of an exactly solvable many-body problem. It is simpler but has many similarities to the lattice many-body problems such as the Hubbard model. It is therefore a natural

place to understand the magnitude and signs of the asymmetric corrections to the lowest-order Fermi-liquid theory result mentioned above. The goal of this Rapid Communication is to explore this asymmetry by using the numerical renormalization group (NRG), and to contrast it with various approximate theories. We report a surprising result in this well-studied problem: We find a line in the  $U$ - $n$  plane where the asymmetry changes sign. Here,  $U$  is the interaction strength and  $n$  the impurity occupancy. Along this one-dimensional line in the  $U$ - $n$  plane, the particle-hole ( $p$ - $h$ ) symmetry of the scattering rate is exactly fulfilled up to the fifth- and higher-order terms. This change of sign demarcates the border between the qualitatively different regimes of weak and strong correlations. Indeed, this Rapid Communication was motivated by the puzzling observation that weak-coupling approaches (e.g., perturbation theory in the interaction strength  $U$ ) and strong-coupling techniques [e.g., the extremely correlated Fermi-liquid (ECFL) theory [2–4]] give opposite signs for the asymmetry, as illustrated in Fig. 1.

The SIAM is defined by the Hamiltonian

$$H = \epsilon_d(n_\uparrow + n_\downarrow) + U n_\uparrow n_\downarrow + \sum_{k\sigma} \epsilon_k c_{k\sigma}^\dagger c_{k\sigma} + \sum_{k\sigma} (t_k c_{k\sigma}^\dagger d_\sigma + \text{H.c.}), \quad (2)$$

where  $\epsilon_d$  is the impurity level,  $d_\sigma$  are impurity operators,  $n_\sigma = d_\sigma^\dagger d_\sigma$  and  $n = \langle n_\uparrow + n_\downarrow \rangle$ , and  $c_{k\sigma}$  are operators for conduction-band electrons with energy  $\epsilon_k$  that couple with the impurity with amplitude  $t_k$ . The hybridization strength is  $\Gamma = \pi \sum_k |t_k|^2 \delta(\omega - \epsilon_k)$ ; we will assume it to be a constant function in the domain  $-D < \omega < D$ . In the following we will make use of the Hartree-Fock parameter  $E_d = \epsilon_d + U \langle n_\sigma \rangle$  with  $E_d = 0$  corresponding to the  $p$ - $h$  symmetric  $n = 1$  case. We also define the dimensionless interaction  $u = U/\pi\Gamma$ . We will limit our consideration to  $n < 1$ , since the results for

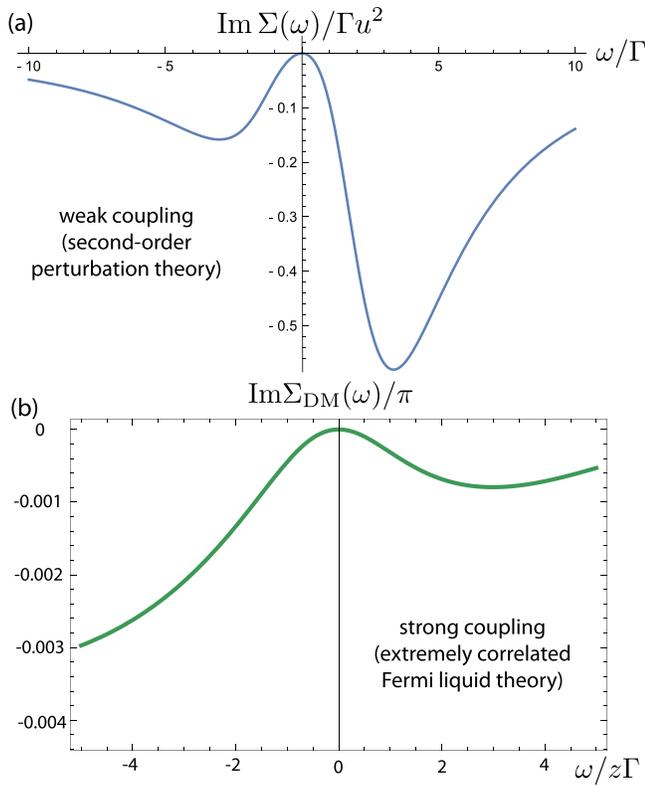


FIG. 1. Imaginary parts of the self-energy in (a) weak-coupling theory and in (b) strong-coupling theory have the opposite sign of the particle-hole asymmetry. The second-order perturbation theory result corresponds to  $E_d/\Gamma = 1$  in Fig. 2 of Ref. [5], and the extremely correlated Fermi-liquid theory result to  $n = 0.6$  in Fig. 5 of Ref. [4].

$n > 1$  can be obtained by the  $p$ - $h$  transformation  $d_\sigma \rightarrow d_\sigma^\dagger$ ,  $c_{k\sigma} \rightarrow -c_{k\sigma}^\dagger$ , which takes  $\omega$  to  $-\omega$ .

Second-order perturbation theory in  $U$  for the SIAM [5] [e.g., Fig. 1(a) for  $E_d/\Gamma = 1$ ] predicts  $b$  to be of constant sign as a function of  $n$  in the full domain  $0 < n < 1$ , with zero value at  $n = 1$ . Specifically,  $b < 0$ , i.e., the particlelike excitations scatter more strongly. At nonzero but low temperatures, the asymmetry of  $\text{Im } \Sigma(\omega)$  at low  $\omega$  will be controlled by the  $\omega T^2$  term. Due to conformal symmetry of the FL fixed point,  $c = b\pi^2$ , thus the sign of the asymmetry does not depend on  $T$  at low enough temperatures [6].

By extending the perturbative expansion to third order, we find that the third-order contributions tend to have the opposite sign of the second-order ones for small  $\omega$ , i.e., they reduce the scattering. Furthermore, the asymmetry of the third-order contribution is such that the  $\omega > 0$  part is dominant (for  $n < 1$ ), thus the third-order reduction in scattering is stronger for particlelike excitations [see Fig. 2(a)]. There is thus a competition between the second- and third-order contributions which may lead to a change in sign of the asymmetric terms. The contributions  $b^{(2)}$  and  $b^{(3)}$  to the coefficient  $b$  actually follow very similar qualitative  $E_d$  dependence [see Fig. 2(b)], except for the sign and evidently a different power of  $u = U/\pi\Gamma$ . Based on these results, the change of sign in  $b$  should occur for  $u \sim 1$ , i.e.,  $U \sim \pi\Gamma$ . Since the perturbation theory also breaks down at  $u \sim 1$ , we cannot make a precise

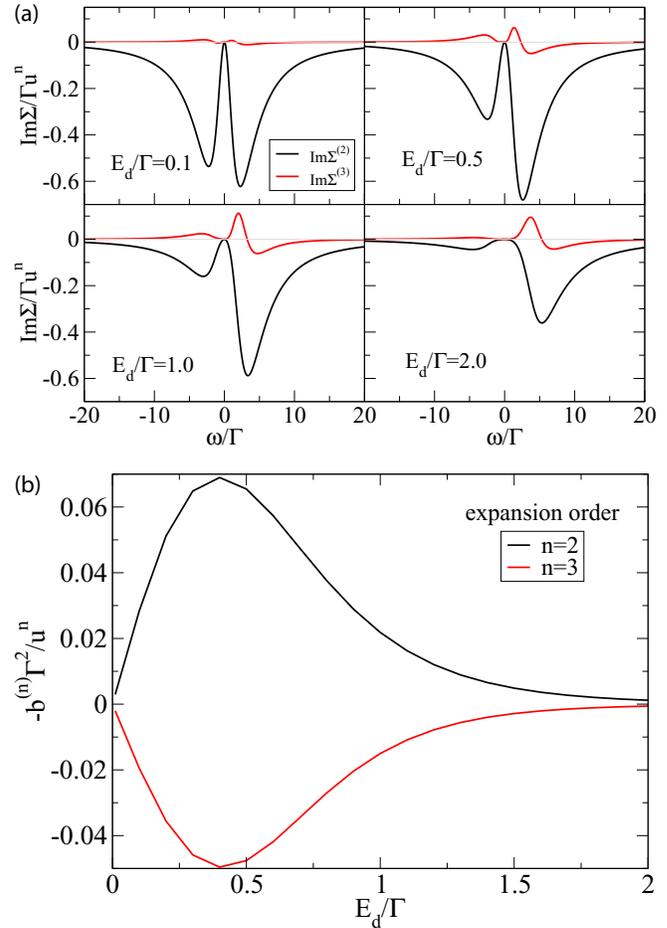


FIG. 2. (a) Second-order and third-order imaginary part of the self-energy for a range of the Hartree-Fock parameters  $E_d = \epsilon_d + U\langle n_\sigma \rangle$ . The y axis is scaled as  $1/(\Gamma u^n)$ , where  $n$  is the expansion order and  $u = U/\pi\Gamma$ .  $T = 0$ . (b) Coefficient  $b$  of the  $\omega^3$  term in  $\text{Im } \Sigma(\omega, T = 0)$  at second and third order in perturbation theory.

statement about the details of this sign change within the bare perturbative approach.

We therefore solved the impurity problem numerically using the numerical renormalization group (NRG) [7–12]. This nonperturbative approach is based on logarithmic discretization of the continuum, mapping onto a tight-binding chain with exponentially decreasing hopping constants, and iterative diagonalization of the resulting Hamiltonian [7,8,12]. Through various refinements over the years [11,13–20] the technique has developed into a powerful tool for computing the dynamical properties of impurity models. Comparisons with quantum Monte Carlo simulations indicate that the results of the NRG, when taken to full convergence, may be considered as essentially exact (up to very small systematic errors due to discretization and truncation). We performed the NRG calculations with a narrow broadening kernel by averaging over  $N_z = 32$  interleaved discretization grids with the discretization parameter  $\Lambda = 2$ , and increasing the truncation cutoff until convergence [20]: These steps reduced the oscillatory artifacts and allowed a reliable extraction of the cubic term in the self-energy function in the limit of small

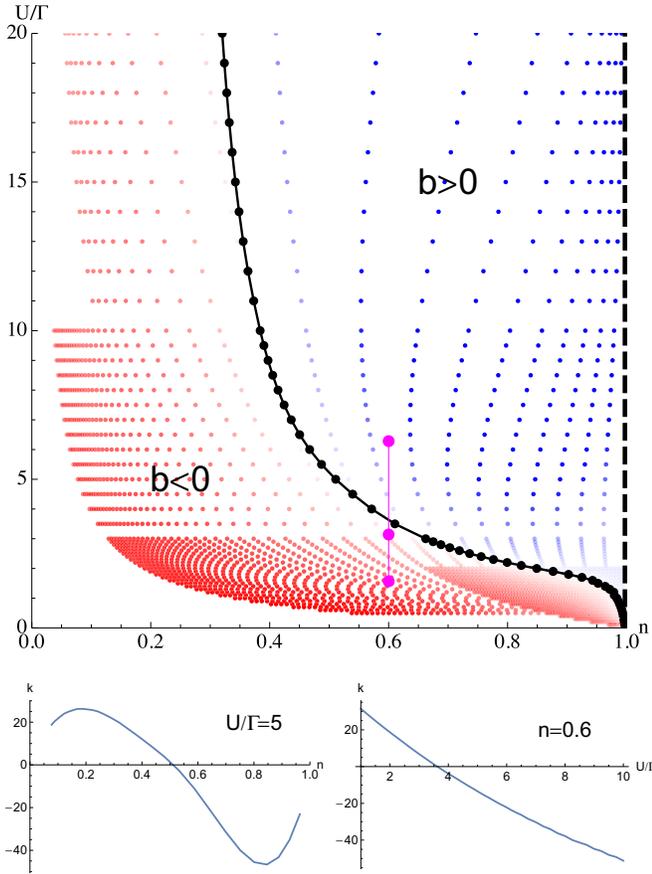


FIG. 3. Top: Sign (red negative, blue positive) and magnitude (color saturation) of the  $\omega^3$  term in  $\text{Im} \Sigma(\omega, T = 0)$ . The quantity shown is the coefficient  $k$  in the fit of the antisymmetrized and normalized combination  $[\text{Im} \Sigma(\omega) - \text{Im} \Sigma(-\omega)] / [\text{Im} \Sigma(\omega) + \text{Im} \Sigma(-\omega)]$  with the linear function  $k\omega$ . The fit is performed in an energy interval  $\omega \in [-\xi : \xi]$ ; here,  $\xi$  is the low-energy scale of the problem defined as the temperature where the impurity moment is screened (and is equivalent to the Kondo temperature in the Kondo regime of the model). Notice that  $k \approx b/a$  and that  $a < 0$ . The three magenta points joined by a line are considered in Fig. 4. The dashed line at  $n = 1$  indicates a further zero crossing of the coefficient  $k$  at the particle-hole symmetric point of the Hamiltonian itself. Bottom: Cross sections at constant interaction  $U/\Gamma = 5$  and constant occupancy  $n = 0.6$ .

$\omega$ . The results, shown in Fig. 3, reveal a change of sign of the coefficient  $b$  in  $\text{Im} \Sigma$  along a curve in the  $U$ - $n$  plane (black line in the figure). At low  $U \ll \Gamma$ , the sign reversal occurs close to half filling. For  $U$  of order  $\Gamma$ , the sign-change point rapidly moves away from half filling. At still higher  $U$ , the slope of the black curve in the  $U$ - $n$  plane redresses and becomes increasingly vertical for  $U \gg \pi\Gamma$ .

These results are fully consistent with our perturbative analysis. For very small  $U$ , the third-order term is negligibly small and the sign is constant in essentially the full  $0 < n < 1$  interval. In this regime, the curve separating the different signs in the  $U$ - $n$  plane is almost vertical and close to  $n = 1$ . At some value of  $U$  of order  $\pi\Gamma$ , the perturbation theory predicts that the third-order contribution will overtake the second-order contribution for most  $E_d$  at almost the same value of  $u$ ,

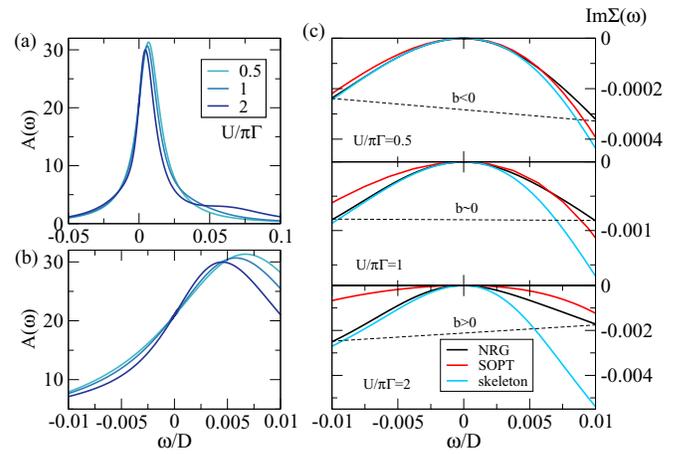


FIG. 4. Spectral function and imaginary part of self-energy for  $n = 0.6$  and three values of  $U/\pi\Gamma$  across the sign-reversal line (magenta line in Fig. 3). (a) Spectral functions computed using the NRG, and (b) closeups on the low- $\omega$  range. (c)  $\text{Im} \Sigma$  computed using the NRG and two weak-coupling approaches. Here,  $\Gamma = 0.01D$ ,  $T = 0$ .  $b$  for the NRG calculations has the sign of the slopes of the dashed lines in the figures.

and  $b$  will thus change in a wide  $n$  interval. Indeed, this seems to correspond to  $U \approx 2\Gamma$  where the curve in the  $U$ - $n$  plane abruptly changes slope and becomes almost horizontal. Around the same  $u$ , however, we enter the strong-coupling regime where the perturbation theory breaks down. Note that for large  $u = U/\Gamma$  the crossover (as  $n$  increases from 0) to the “strong-coupling” domain with  $b > 0$  already takes place in the mixed valent regime of the impurity model; the deep Kondo limit is confined to values of  $n$  close to 1.

We also performed the skeleton expansion to second order, which is a self-consistent calculation where the dressed Green’s function is used as the propagator in the second-order term of the self-energy. This corresponds to an infinite resummation of a certain class of diagrams, which for small  $U$  reduces to the bare perturbation theory in  $U$ . The results show that the coefficient  $b$  has the same sign for any value of  $U$  and  $n$ , i.e., the sign associated with the weak-coupling limit (see Fig. 4). This can also be shown analytically by invoking the Friedel sum rule [21]—for the skeleton expansion this leads to  $b/a \propto \sin(n\pi)/\Gamma$ , which makes  $b$  negative for all  $n$ . The absence of sign change seems to imply that the skeleton expansion is not able to describe the transition to the dynamics expected in the strong-coupling regime, presumably because its starting point is still the noninteracting limit. On the other hand, the extremely correlated Fermi-liquid theory produces the correct sign of the scattering-rate asymmetry because it is constructed as a strong-coupling approach by projecting out the double-occupancy from the outset.

Next, we compare the crossover curve with other quantities sensitive to the magnetic behavior of the impurity. The local-moment fraction  $f_{\text{LM}} = n - 2\langle n_{\uparrow}n_{\downarrow} \rangle$  is equal to the expectation value of the projection operator to the singly occupied impurity state (the state which carries the spin degree of freedom). It behaves as  $f_{\text{LM}} = n$  in the  $U \gg \Gamma$  limit, while the small- $U$  dependence is shown in Fig. 5. None of the contours

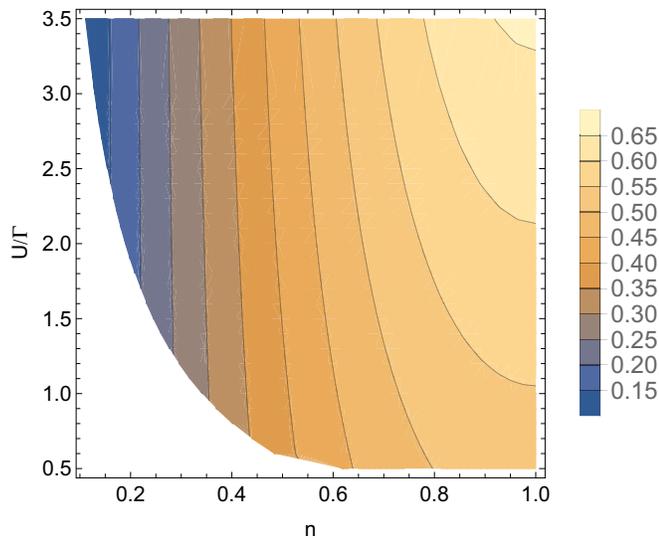


FIG. 5. Local-moment fraction  $f_{LM} = n - 2\langle n_{\uparrow}n_{\downarrow} \rangle$ .

(isocurves) resembles the sign-reversal curve from Fig. 3. Most notably, the contours in  $f_{LM}$  approach the  $n = 1$  line with zero slope and do not curve down. Other thermodynamic quantities, such as the characteristic low-energy scale of the problem defined by the temperature where the impurity entropy and the effective moment become small (i.e., the Kondo temperature in the Kondo regime), also correlate with the dependence of  $f_{LM}$  on  $U$  and  $n$  (not shown). The scattering asymmetry is thus not simply related to the degree of local-moment formation, but requires a calculation of dynamical properties.

We now discuss the relevance of these results to the low-temperature thermopower in correlated systems. Two cases need to be distinguished. In quantum dots that are directly described by the SIAM, the thermopower is determined by the asymmetry of the spectral function around the Fermi level, i.e., its Fermi-level slope [22–24]. The asymmetry of the particle-hole scattering is not important as such, as it only enters as one factor that affects the spectral-function slope. The situation is different for bulk systems described

by the Hubbard model that map within the dynamical mean-field theory (DMFT) approximation to a SIAM with a self-consistently defined hybridization function [25–31]. There, the thermopower is given by the “leading” term proportional to the Fermi-level slope of the transport function that is “corrected” by a term proportional to the coefficients of the cubic terms in  $\text{Im} \Sigma$ . This is actually an order 1 correction [29,30], which may become dominant close to half filling. Due to the DMFT self-consistency, this term has a complex dependence on specific details of the problem [shape of the noninteracting density of states (DOS), doping level, strength of the interaction compared to the critical  $U_{c2}$  of the Mott metal-to-insulator transition]. We plan to present a study of these in a separate publication.

In conclusion, we uncovered a crossover line in the phase diagram of the single-impurity Anderson model with a flat hybridization function which corresponds to a change in the scattering dynamics. On one side of this line the impurity behaves as a weakly renormalized resonant level, and on the other side as a magnetic impurity. On the weakly correlated side, the particlelike excitations scatter more strongly than the holelike excitations, while the opposite is the case on the strongly correlated side. This crossover might be directly observable in quantum dot experiments [32,33] provided that the spectral function can be measured in a sufficient energy window so that the reconstruction of the full Green’s function  $G(\omega)$  is possible via the Kramers-Kronig transformation: Assuming that the hybridization function is approximately constant close to the Fermi level, the particle-hole scattering asymmetry can be easily extracted from  $\text{Im}[G(\omega)^{-1}]$ . The predicted almost exact particle-hole symmetry of the scattering rate along the crossover line in the  $U$ - $n$  plane should make it an interesting feature to test for in experiments on magnetic adsorbates and quantum dots.

R.Ž. acknowledges the support of the Slovenian Research Agency (ARRS) under P1-0044 and J1-7259. The work at UCSC was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award No. DE-FG02-06ER46319. H.R.K. acknowledges support from the Department of Science and Technology (DST), India.

- [1] G.-H. Gweon, B. S. Shastry, and G. D. Gu, Extremely Correlated Fermi Liquid Description of Normal State ARPES in Cuprates, *Phys. Rev. Lett.* **107**, 056404 (2011).
- [2] B. S. Shastry, Dynamical Particle-Hole Asymmetry in Cuprate Superconductors, *Phys. Rev. Lett.* **109**, 067004 (2012).
- [3] B. S. Shastry, Extremely Correlated Fermi Liquids, *Phys. Rev. Lett.* **107**, 056403 (2011).
- [4] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, Extremely correlated Fermi liquid study of the  $U = \infty$  Anderson impurity model, *Phys. Rev. B* **88**, 205108 (2013).
- [5] B. Horvatić and V. Zlatić, Perturbation expansion for the asymmetric Anderson Hamiltonian, *Phys. Status Solidi* **99**, 251 (1980).
- [6] B. Horvatić and V. Zlatić, Perturbation expansion for the asymmetric Anderson Hamiltonian II. General asymmetry, *Phys. Status Solidi* **111**, 65 (1982).
- [7] K. G. Wilson, The renormalization group: Critical phenomena and the Kondo problem, *Rev. Mod. Phys.* **47**, 773 (1975).
- [8] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Renormalization-group approach to the Anderson model of dilute magnetic alloys. I. Static properties for the symmetric case, *Phys. Rev. B* **21**, 1003 (1980).
- [9] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Renormalization-group approach to the Anderson model of dilute magnetic alloys. II. Static properties for the asymmetric case, *Phys. Rev. B* **21**, 1044 (1980).
- [10] T. A. Costi and A. C. Hewson, Resistivity cross-over for the non-degenerate Anderson model, *Philos. Mag. B* **65**, 1165 (1992).
- [11] W. Hofstetter, Generalized Numerical Renormalization Group for Dynamical Quantities, *Phys. Rev. Lett.* **85**, 1508 (2000).

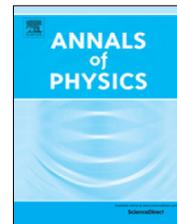
- [12] R. Bulla, T. Costi, and T. Pruschke, The numerical renormalization group method for quantum impurity systems, *Rev. Mod. Phys.* **80**, 395 (2008).
- [13] H. O. Frota and L. N. Oliveira, Photoemission spectroscopy for the spin-degenerate Anderson model, *Phys. Rev. B* **33**, 7871 (1986).
- [14] O. Sakai, Y. Shimizu, and T. Kasuya, Single-particle and magnetic excitation spectra of degenerate Anderson model with finite  $f$ - $f$  Coulomb interaction, *J. Phys. Soc. Jpn.* **58**, 3666 (1989).
- [15] T. A. Costi, A. C. Hewson, and V. Zlatić, Transport coefficients of the Anderson model via the numerical renormalization group, *J. Phys.: Condens. Matter* **6**, 2519 (1994).
- [16] R. Bulla, A. C. Hewson, and Th. Pruschke, Numerical renormalization group calculation for the self-energy of the impurity Anderson model, *J. Phys.: Condens. Matter* **10**, 8365 (1998).
- [17] V. L. Campo and L. N. Oliveira, Alternative discretization in the numerical renormalization group, *Phys. Rev. B* **72**, 104432 (2005).
- [18] R. Peters, T. Pruschke, and F. B. Anders, A numerical renormalization group approach to Green's functions for quantum impurity models, *Phys. Rev. B* **74**, 245114 (2006).
- [19] A. Weichselbaum and J. von Delft, Sum-Rule Conserving Spectral Functions from the Numerical Renormalization Group, *Phys. Rev. Lett.* **99**, 076402 (2007).
- [20] R. Žitko and T. Pruschke, Energy resolution and discretization artifacts in the numerical renormalization group, *Phys. Rev. B* **79**, 085106 (2009).
- [21] A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, UK, 1993).
- [22] R. Scheibner, H. Buhmann, D. Reuter, M. N. Kiselev, and L. W. Molenkamp, Thermopower of a Kondo Spin-Correlated Quantum Dot, *Phys. Rev. Lett.* **95**, 176602 (2005).
- [23] T. A. Costi and V. Zlatić, Thermoelectric transport through strongly correlated quantum dots, *Phys. Rev. B* **81**, 235127 (2010).
- [24] S. Andergassen, T. A. Costi, and V. Zlatić, Mechanism for large thermoelectric power in molecular quantum dots described by the negative- $u$  Anderson model, *Phys. Rev. B* **84**, 241107(R) (2011).
- [25] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, *Rev. Mod. Phys.* **68**, 13 (1996).
- [26] B. S. Shastry, Electrothermal transport coefficients at finite frequencies, *Rep. Prog. Phys.* **72**, 016501 (2009).
- [27] B. S. Shastry, in *New Materials for Thermoelectric Applications: Theory and Experiments*, edited by V. Zlatić and A. Hewson (Springer, Dordrecht, 2013), pp. 25–30.
- [28] K. Haule and G. Kotliar, Thermoelectrics near the Mott localization-delocalization transition, in *Properties and Applications of Thermoelectric Materials*, edited by V. Zlatić and A. C. Hewson (Springer, Dordrecht, 2009), p. 119.
- [29] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, How Bad Metals Turn Good: Spectroscopic Signatures of Resilient Quasiparticles, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [30] W. Xu, K. Haule, and G. Kotliar, Hidden Fermi Liquid, Scattering Rate Saturation, and Nernst Effect: A Dynamical Mean-Field Theory Perspective, *Phys. Rev. Lett.* **111**, 036401 (2013).
- [31] L.-F. Arsenault and A.-M. S. Tremblay, Transport functions for hypercubic and Bethe lattices, *Phys. Rev. B* **88**, 205109 (2013).
- [32] D. Goldhaber-Gordon, J. Göres, M. A. Kastner, H. Shtrikman, D. Mahalu, and U. Meirav, From the Kondo Regime to the Mixed-Valence Regime in a Single-Electron Transistor, *Phys. Rev. Lett.* **81**, 5225 (1998).
- [33] A. V. Kretinin, H. Shtrikman, D. Goldhaber-Gordon, M. Hanl, A. Weichselbaum, J. von Delft, T. Costi, and D. Mahalu, Spin-1/2 Kondo effect in an InAs nanowire quantum dot: Unitary limit, conductance scaling, and Zeeman splitting, *Phys. Rev. B* **84**, 245316 (2011).



ELSEVIER

Contents lists available at ScienceDirect

## Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)

## Fermi surface volume of interacting systems

B. Sriram Shastry

Physics Department, University of California, Santa Cruz, CA 95064, USA



## ARTICLE INFO

## Article history:

Received 18 March 2019

Accepted 22 March 2019

Available online 24 March 2019

## Keywords:

Fermi surface

*t*-*J* model

Strongly correlated matter

Superconductivity

## ABSTRACT

Three Fermion sumrules for interacting systems are derived at  $T = 0$ , involving the number expectation  $\bar{N}(\mu)$ , canonical chemical potentials  $\mu(m)$ , a logarithmic time derivative of the Greens function  $\gamma_{\vec{k}\sigma}$  and the static Greens function. In essence we establish at zero temperature the sumrules linking:

$$\bar{N}(\mu) \leftrightarrow \sum_m \Theta(\mu - \mu(m)) \leftrightarrow \sum_{\vec{k}, \sigma} \Theta(\gamma_{\vec{k}\sigma}) \leftrightarrow \sum_{\vec{k}, \sigma} \Theta(G_\sigma(\vec{k}, 0)).$$

Connecting them across leads to the Luttinger and Ward sumrule, originally proved perturbatively for Fermi liquids. Our sumrules are nonperturbative in character and valid in a considerably broader setting that additionally includes non-canonical Fermions and Tomonaga–Luttinger models. Generalizations are given for singlet-paired superconductors, where one of the sumrules requires a testable assumption of particle–hole symmetry at all couplings. The sumrules are found by requiring a continuous evolution from the Fermi gas, and by assuming a monotonic increase of  $\mu(m)$  with particle number  $m$ . At finite  $T$  a pseudo-Fermi surface, accessible to angle resolved photoemission, is defined using the zero crossings of the first frequency moment of a weighted spectral function.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

The Luttinger–Ward (LW) sumrule [1] for interacting electrons expresses the number of electrons in terms of the static limit of the imaginary frequency Greens function [2–4] for  $T \rightarrow 0$  as

$$\bar{N}(\mu) = \sum_{\vec{k}, \sigma} \Theta(G_\sigma(\vec{k}, \omega = 0|\mu)), \quad (1)$$

E-mail address: [sriram@physics.ucsc.edu](mailto:sriram@physics.ucsc.edu).<https://doi.org/10.1016/j.aop.2019.03.016>

0003-4916/© 2019 Elsevier Inc. All rights reserved.

with  $\Theta(x) = \frac{1}{2} (1 + \text{sgn}(x))$ . Since the static Greens function is negative outside the Fermi surface, its volume is fixed by the number of particles [1,4,5], independent of the magnitude of the interaction. This interaction independence is a fundamental result in Landau's theory of the Fermi liquid [6,7]. In condensed matter physics, field theory and statistical mechanics, the origin of this sumrule and its ramifications have been very influential [4,7–11]. It has continued to receive much attention to recent times [12–24], partly motivated by the search for novel phases of matter that might violate this sumrule. The present work provides a physically transparent derivation of the sumrule, and extends it in several directions. The extended version includes non-Fermi liquids, such as the 1-d Tomonaga–Luttinger model (TLM). It is also valid for non-canonical Fermions, such as  $U = \infty$  Gutzwiller projected electrons in the  $t$ - $J$  model, in treatments where continuity with the Fermi gas is maintained [25], but not otherwise [23,26]. Our extension also includes singlet pairing superconductors. These include the s-wave BCS–Gor'kov–Nambu case and d-wave cuprate superconductors. While analyticity in the coupling is lost in these extensions, they do evolve continuously from the non-interacting limit, which suffices for our purposes. Exotic superconductors, where one hollows out the  $k$ -space [27], provide an interesting counter-point where continuity with the gas limit is discarded; we need to exclude them here too.

Since the static Greens function entering Eq. (1) is not directly measurable, one needs to relate it to other directly visible signatures for using it. This work provides a new and experimentally accessible sum-rule Eq. (79), which is equivalent to Eq. (1) at  $T = 0$  in the cases considered. It also allows one to define a pseudo Fermi surface at any  $T$ . This surface carries useful information on the real part of the on-shell selfenergy.

### 1.1. Methods used

The technique used here is non-perturbative, it relies on *isothermal* continuity in some parameter  $\lambda$  connecting the interacting and non-interacting systems. Since this type of continuity has not been explicitly discussed in literature, a few words are in order. As some parameter in the Hamiltonian is varied, the variation is required to be isothermal, i.e. at each intermediate value of the parameter, the system is allowed to repopulate energy levels according to the thermal distribution. This is in contrast to adiabatic variations where the population of the energy levels is frozen at their starting values. By continuity we imply that the expectation of the energy and other observables change without discontinuity, i.e. we rule out first order transitions. Illustrating the distinction we note that the change of shape of the Fermi surface for anisotropic systems is allowed by isothermal continuity, but not by adiabatic continuity. Finally our method does not require analyticity in a coupling, isothermal continuity is sufficient for our purpose.

Our other main assumption is that the canonical ensemble (CE) chemical potentials  $\mu(m)$  increase monotonically with the particle number  $m$ , whereby the canonical free energy is a concave-up function of  $m$ . This is tantamount to ruling out phase separation. We argue in Section 2 that such a monotonic behavior could be regarded as a defining feature of repulsive interactions.

In each case covered by our argument, at non-zero  $T$  we construct an effective particle density  $n_{\text{eff}}(T)$ , and pseudo-Fermi surface, whose temperature variation reveals lowest lying characteristic energy scales in the system. The pseudo-Fermi surface has the potential to be studied using angle resolved photoemission (ARPES) technique, and hence is discussed in some detail in Section 7

### 1.2. Organization of the paper

The paper is organized as follows. I first establish in Section 2 a basic thermodynamic number sumrule for electrons with repulsive interactions;

$$\bar{N}(\mu) = \sum_{m=0}^{N_{\text{max}}-1} \Theta(\mu - \mu(m)), \quad (2)$$

where the CE chemical potential  $\mu(m) = F_{m+1} - F_m$  is the difference of the canonical free energies  $F$  with  $m+1$  and  $m$  particles. We will assume a hard-core set of particles, and therefore the maximum

number of particles is limited by  $N_{max}$ . In Section 3 I next introduce  $\gamma$ , the temporal log-derivative of the Greens function:

$$\gamma_{\vec{k}\sigma}(\mu, T) = \lim_{\tau=\beta/2} \partial_\tau \log G_\sigma(\vec{k}, \tau|\mu), \quad \beta = \frac{1}{k_B T}. \quad (3)$$

Setting  $\tau = \beta/2$  sandwiches each Fermionic operator of  $G$  symmetrically by factors that project all contributing states to the ground state as  $T \rightarrow 0$ . While the study of  $G_\sigma(\vec{k}, \frac{\beta}{2}|\mu)$  is popular in quantum Monte-Carlo studies [28], the log-derivative, playing a key role in this work, has not been discussed earlier. Its physical content at low  $T$ , as  $\mu$  minus a  $\vec{k}$ -weighted average over  $\mu(m)$  becomes clear later (see Eq. (32)). In Section 3 we make an important distinction between two ways of taking the zero temperature and thermodynamic limits, in **Limit-I** we take  $T \rightarrow 0$  first and  $L \rightarrow \infty$  later, while in **Limit-II** we take  $L \rightarrow \infty$  first and  $T \rightarrow 0$  later.

In Section 4 the  $T = 0$  limit is taken first (i.e. in the limit **Limit-I**), and shown to lead to the sumrule

$$\sum_{\vec{k}, \sigma} \Theta(\gamma_{\vec{k}\sigma}(\mu, 0)) = \sum_m \Theta(\mu - \mu(m)). \quad (4)$$

This is demonstrated for the Fermi liquid and also for the 1-d case of a Tomonaga–Luttinger model.

In Section 5 the  $L \rightarrow \infty$  limit is taken first (i.e. in the limit **Limit-II**), whereby we obtain a continuous frequency variable in terms of which a spectral function can be defined. Here the sumrule

$$\sum_{\vec{k}, \sigma} \Theta(\gamma_{\vec{k}\sigma}(\mu, 0^+)) = \sum_{\vec{k}, \sigma} \Theta(G_\sigma(\vec{k}, \omega = 0|\mu)), \quad (5)$$

is established for Fermi liquids in Section 5.1 and for 1-d TLL systems in Section 5.2.

Assuming unbroken symmetry, powerful theorems on the uniqueness of the ground state [29,30] are applicable, these allows us to equate the two zero temperature limits

$$\gamma_{\vec{k}\sigma}(\mu, 0) = \gamma_{\vec{k}\sigma}(\mu, 0^+). \quad (6)$$

Upon using Eqs. (2), (7), (4), (5) (or Eqs. (8), (75)) then imply the sumrule Eq. (1). In the infinite volume limit, the  $\vec{k}$  sums are replaced by integrals as usual.

In Section 6 a systematic development of the volume theorem for a singlet superconducting state is provided. This broken symmetry state not accessible by the methods of L–W. In Section 6.1 we study the canonical chemical potentials  $\mu_e(2m) \equiv \frac{1}{2}(F_{2m+2} - F_{2m})$  constrained to the even particle sector. The  $\mu_e(2m)$  are taken to be monotonically increasing in  $m$ , reflecting the inherent repulsion between pairs of electrons. In this ensemble we study the effects of adding or removing a particle and thence the Greens function, leading to the sumrule

$$\bar{N}_{SC}(\mu) = 2 \sum_{m=0}^{\frac{1}{2}N_{max}} \Theta(\mu - \mu_e(2m)), \quad (7)$$

which replaces Eq. (2) in the normal state.

In Section 6.2 the Greens function and  $\gamma_{\vec{k}}$  are studied at  $T = 0$ , (i.e. in the **Limit-I**) in the superconducting state, subject to the assumptions of particle–hole symmetry and of the repulsion between the Cooper pairs of electrons. Here one finds

$$\sum_{\vec{k}\sigma} \Theta(\gamma_{\vec{k}\sigma}(\mu, 0)) = 2 \sum_m \Theta(\mu - \mu_e(2m)) = \bar{N}(\mu), \quad (8)$$

a sumrule corresponding to Eqs. (4).

In Section 6.3 the Greens function and  $\gamma_{\vec{k}}$  are studied at  $L = \infty$ , (i.e. in the **Limit-II**) in the superconducting state. Here we use the Nambu–Go’rkov [31–35] formalism together with the formally exact quasiparticle representation [34] of the diagonal Greens function. This yields the sumrule Eq. (75), and completes the set of links giving the number sumrule Eq. (1). In summary the

sumrules corresponding to Eqs. (4), (5) for a superconductor are Eqs. (8), (75) in Sections 6.2 and 6.3.

In Section 7 details of the applications of the sumrules at finite  $T$  to angle resolved photoemission (ARPES) are given. The main finding is that one can use a first moment of the frequency with respect to the weight function

$$\mathcal{W}(\vec{k}, \omega, T) = W_0 \frac{A(\vec{k}, \omega)}{\cosh(\frac{1}{2}\beta\omega)} \quad (9)$$

where  $W_0$  is a normalization constant and  $A(\vec{k}, \omega)$  is the electronic spectral weight measured in experiments. It is denoted in the rest of the paper by the theoreticians favorite symbol  $\rho_G(k, \omega)$ . The first moment with respect to  $\mathcal{W}$  of the frequency  $\langle \omega \rangle_{\vec{k}}$  is found to be equal to  $-\gamma_{\vec{k}}(\mu, 0^+)$  at  $T = 0$ , and in view of the theorems proved here, can be used as a proxy for the inverse static Greens function. It can be found from photoemission at any  $T$ , and thereby permits us to define an observable pseudo-Fermi surface (PFS), which becomes the true Fermi surface (FS) at  $T \rightarrow 0$ . Section 7.1 examines the  $T$  dependence of the pseudo FS and notes that it can be used to unravel the often sensitive  $T$  dependence of the real part of selfenergy. In Section 7.2 the pseudo FS for a singlet superconducting state is discussed in some detail.

In Section 8 I summarize the paper and discuss the results.

### 1.3. The Hamiltonian

Consider a two component Fermion Hamiltonian

$$\mathcal{H} = \sum_{\vec{k}\sigma} \varepsilon(\vec{k}) C_{\vec{k}\sigma}^\dagger C_{\vec{k}\sigma} + U \times \text{interaction} - \mu \mathcal{N} \quad (10)$$

in the grand canonical ensemble (GCE), where  $\mathcal{N}$  is the number operator,  $\mu$  is the (running i.e. varying) chemical potential,  $\varepsilon(k)$  the energy dispersion. We take the interaction as a short-ranged Hubbard type interaction, possibly with a few further neighbor terms. The initial discussion assumes  $U > 0$ , and later we allow for pairing i.e.  $U < 0$ . We assume a finite lattice in  $d$ -dimensions with  $N_s = L^d$  sites ( $L$  the linear dimension) and take the limit of an infinite system at the end.

## 2. A number sumrule at $T = 0$

We derive a new and useful sumrule Eq. (2) for the electron number at  $T = 0$  for electrons with repulsive interactions. It is of thermodynamic origin and is based on an assumption of “good behavior” of the chemical potentials of repulsive finite systems. Let us define the common eigenstates of  $\mathcal{N}, \mathcal{H}$  as  $|m, a\rangle$  with eigenvalues  $m, E_a(m) - m\mu$  as the respective eigenvalues. In the canonical ensemble (CE)  $m$  particle sector, we will denote  $E_0(m)$  and  $F_m$  as the ground state energy and free energy  $F_m = -k_B T \log Z_m$ . We define the CE chemical potentials  $\mu(m)$  using

$$\mu(m) = F_{m+1} - F_m, \text{ for } 0 \leq m < N_{max}, \quad (11)$$

where  $T$  dependence is implied in all variables. The value of  $N_{max}$  is twice the number of sites for the prototypical spin- $\frac{1}{2}$  Hubbard model. The set of free energies  $F_m$  is conveniently extended by defining  $F_0$  and  $F_{N_{max}+1}$  satisfying the conditions  $F_0 > 2F_1 - F_2$  and  $F_{N_{max}+1} > 2F_{N_{max}} - F_{N_{max}-1}$  but are arbitrary otherwise. By inversion we obtain for  $m \geq 1$

$$F_m - F_0 = \mu(m-1) + \mu(m-2) + \dots + \mu(0). \quad (12)$$

Our essential assumption is that of a positive definite CE compressibility, i.e. a strictly concave-up free energy,

$$\begin{aligned} F_{m+1} + F_{m-1} - 2F_m &> 0, \\ \text{or } \mu(m) &> \mu(m-1). \end{aligned} \quad (13)$$

In a very large system, if we replace differences by derivatives, Eq. (13) becomes the more familiar condition of a positive physical compressibility. We can use it to order the CE chemical potentials as a monotonically increasing set

$$\mu(0) < \dots < \mu(j) < \dots < \mu(N_{max}). \quad (14)$$

From the interesting example of the Hubbard model on a buckyball cluster, we learn that this condition can be violated by ostensibly repulsive interactions [36], leading to phase separation and related phenomena. Therefore the ordering in Eq. (14) seem to us to be no more than a robust characterization of truly repulsive interactions.

We introduce a useful set of weight functions

$$\xi_n = e^{\beta(\mu - \mu(n))}. \quad (15)$$

Using these we may write  $p_\mu(m)$  the probability of finding  $m$  particles in the GCE. With  $p_\mu(m) \equiv \exp \beta(m\mu - F_m)/Z(\mu)$ , and the grand partition function  $Z(\mu) = \sum_m e^{\beta(\mu m - F_m)}$ , we obtain

$$p_\mu(m) = Z^{-1}(\mu) \xi_0 \xi_1 \dots \xi_{m-1} \quad (16)$$

$$Z(\mu) = 1 + \xi_0 + \xi_0 \xi_1 + \xi_0 \xi_1 \xi_2 + \xi_0 \xi_1 \xi_2 \xi_3 + \dots \quad (17)$$

The CE chemical potentials  $\mu(m)$  are computed at low  $T$  from the ground state energies  $E_0(m)$ .

When  $T \ll 2\pi \hbar v / (Lk_B)$  where  $v \sim v_F$  the band velocity, the free energies  $F_m$  can be replaced by the ground state energies  $F_m \rightarrow E_0(m)$ , and the canonical chemical potentials  $\mu(m)$  computed from the ground state energies  $E_0(m)$ . We note that

$$\lim_{T \rightarrow 0} \xi_j \rightarrow \begin{cases} \infty, & \text{if } \mu > \mu(j), \\ 1, & \text{if } \mu = \mu(j), \\ 0, & \text{if } \mu < \mu(j). \end{cases} \quad (18)$$

Let us consider the case when  $\mu$  is in the  $j$ th (open) interval  $\mathcal{I}_j$  defined as

$$\mathcal{I}_j = \{\mu \mid \mu(j-1) < \mu < \mu(j)\}. \quad (19)$$

When  $\mu \in \mathcal{I}_j$  at very low  $T$ , the  $j$  particle sector is occupied while  $j+1$  and higher sectors are unoccupied. To see this, when  $T \rightarrow 0$  we observe that  $\xi_0, \xi_1 \dots \xi_{j-1}$  grow while  $\xi_j, \xi_{j+1} \dots$  decrease towards zero. Therefore for  $\mu \in \mathcal{I}_j$ ,  $Z$  is dominated by a single term

$$Z = \xi_0 \xi_1 \dots \xi_{j-1} \times \mathcal{Y} \quad (20)$$

$$\mathcal{Y} = \left( 1 + \frac{1}{\xi_{j-1}} + \frac{1}{\xi_{j-1}\xi_{j-2}} + \dots + \xi_j + \xi_j \xi_{j+1} + \dots \right) \rightarrow 1,$$

and therefore

$$p_\mu(j) \rightarrow 1 \quad (21)$$

while the probabilities with lower and higher indices vanish:

$$\begin{aligned} p_\mu(j-r) &\rightarrow \frac{1}{\xi_{j-1}\xi_{j-2} \dots \xi_{j-r}} \sim 0 \\ p_\mu(j+r) &\rightarrow \xi_j \dots \xi_{j+r-1} \sim 0. \end{aligned} \quad (22)$$

Therefore at  $T = 0$  it follows that the system has  $j$ , and no more than  $j$  particles, i.e.

$$\lim_{T=0} p_\mu(j) = \Theta(\mu - \mu(j-1)) - \Theta(\mu - \mu(j)). \quad (23)$$

The number of particles can be found using Eq. (23) and  $\bar{N}(\mu) = \sum_{m=1}^{N_s} m p_\mu(m)$ . Shifting the sum in one of the terms and simplifying, we deduce the  $T = 0$  thermodynamic number sumrule Eq. (2).

Note the crucial role played by concavity of the free energy, it implies a  $1 \leftrightarrow 1$  relationship between  $m$  and  $\mu(m)$ . This rules out double bends  $\curvearrowright$ , i.e. a non monotonic relation which prevents inversion. The assumed monotonicity allows the relationship to be inverted, yielding  $m$  from  $\mu(m)$  uniquely and hence giving the sum-rule. In order to deal with degeneracies of  $\mu(m)$ , usually arising from discrete symmetries (spin, parity, rotation, ...), we relax the strictly increasing condition Eq. (13) to the weaker

$$\mu(m) \geq \mu(m-1), \quad (24)$$

we obtain a second form of the sumrule:

$$\bar{N}(\mu) = \sum'_m g_m \Theta(\mu - \mu(m)), \quad (25)$$

where  $g_m$  is the degeneracy of the particular  $\mu(m)$ , and the primed sum is over unequal  $\mu(m)$ 's.

### 3. Log-derivative of the Greens function

The log-derivative in Eq. (3) can be written as a ratio

$$\gamma_{\vec{k}\sigma}^- = \beta_{\vec{k}\sigma}^- / \alpha_{\vec{k}\sigma}^-, \quad (26)$$

where

$$\alpha_{\vec{k}\sigma}^-(\mu, T) = -G_\sigma(\vec{k}, \frac{\beta}{2} | \mu) \quad (27)$$

$$\beta_{\vec{k}\sigma}^-(\mu, T) = - \lim_{\tau \rightarrow \frac{1}{2}\beta} \partial_\tau G_\sigma(\vec{k}, \tau | \mu). \quad (28)$$

In terms of the convenient variable

$$\mathbf{f}(m, a, b) \equiv e^{\beta(\mu(m+\frac{1}{2}) - \frac{1}{2}(E_a(m) + E_b(m+1)))} / Z(\mu),$$

we find

$$\alpha_{\vec{k}\sigma}^-(\mu, T) = \sum_{m,a,b} \mathbf{f}(m, a, b) |\langle m, a | C_{\vec{k}\sigma}^- | m+1, b \rangle|^2, \quad (29)$$

and

$$\beta_{\vec{k}\sigma}^-(\mu, T) = \sum_{m,a,b} \mathbf{f}(m, a, b) (\mu + E_a(m) - E_b(m+1)) |\langle m, a | C_{\vec{k}\sigma}^- | m+1, b \rangle|^2 \quad (30)$$

These spectral representations imply that at low T *both initial and final states* are limited to their ground states in their respective number sectors.

We take the low temperature limit and the thermodynamic limit in two distinct ways, by comparing  $k_B T$  with an energy scale  $\Delta_E$  representing the excited state energy level separation in gapless systems:

$$\Delta_E \sim \frac{2\pi \hbar v}{L}, \quad (31)$$

where  $v \sim v_F$  the band velocity. We distinguish between two ways of taking the limit

- **Limit (I)**:  $\Delta_E > k_B T \gtrsim 0$ , or equivalently  $\{\frac{1}{L} \rightarrow 0, T \rightarrow 0\}$
- **Limit (II)**:  $k_B T > \Delta_E \gtrsim 0$ , or equivalently  $\{T \rightarrow 0, \frac{1}{L} \rightarrow 0\}$ .

The two limits can be taken with different sets of tools, **Limit(I)** leads to Eq. (4), and can be taken employing ideas and tools relevant to finite size systems, while **Limit(II)** leading to Eq. (5) allows the use of electronic spectral functions that are continuous functions of  $\omega$ . The results of [29,30] imply Eq. (6), i.e. that the two limits coincide asymptotically.

#### 4. Zero temperature Limit (I), i.e. $\{\frac{1}{L} \rightarrow 0, T \rightarrow 0\}$

In this section we consider Fermi liquids of TLM systems and take the  $T = 0$  limit in Eqs. (26), (29), (30). Upon taking the stated limit, we project the sum over the intermediate states  $a, b$  to the ground state, and write  $e^{\beta(\mu m - E_0(m))} \rightarrow p_\mu(m) \times Z(\mu)$ , whereby

$$\begin{aligned}\alpha_{\vec{k}\sigma}(\mu, 0) &= \sum_m \Phi_m(\vec{k}\sigma), \\ \beta_{\vec{k}\sigma}(\mu, 0) &= \sum_m \Phi_m(\vec{k}\sigma) \times \{\mu - \mu(m)\}, \\ \gamma_{\vec{k}\sigma}(\mu, 0) &= \mu - \sum_m \tilde{\Phi}_m(\vec{k}\sigma) \mu(m),\end{aligned}\quad (32)$$

the normalized weight function  $\tilde{\Phi}_m = \Phi_m / \sum_r \Phi_r$  is normalized to unity  $\sum_m \tilde{\Phi}_m = 1$ ) and its un-normalized counterpart

$$\Phi_m(\vec{k}\sigma) = p_\mu(m) e^{\frac{1}{2}(\beta - \mu(m))} Z_{\sigma N_s}(\vec{k}, m), \quad (33)$$

$$Z_{\sigma N_s}(\vec{k}, m) = |\langle m, 0 | C_{\vec{k}\sigma} | m+1, 0 \rangle|^2. \quad (34)$$

$Z_{\sigma N_s}$  is the ground state CE quasiparticle weight of a state with  $m$  particles in  $N_s$ . In Eq. (32) by writing  $\mu(m) \langle m, 0 | C_{\vec{k}\sigma} | m+1, 0 \rangle = \langle m, 0 | [C_{\vec{k}\sigma}, H] | m+1, 0 \rangle$  and evaluating the kinetic piece explicitly, we obtain

$$\gamma_{\vec{k}\sigma}(\mu, 0) = \mu - \varepsilon(\vec{k}) - \mathcal{M}(\vec{k}\sigma, \mu), \quad (35)$$

$$\mathcal{M}(\vec{k}\sigma, \mu) = \sum_m \frac{\tilde{\Phi}_m(\vec{k}\sigma)}{Z^{\frac{1}{2}}(\vec{k}, m)} \langle m, 0 | [C_{\vec{k}\sigma}, V] | m+1, 0 \rangle. \quad (36)$$

We require  $Z$  in Eq. (34) to be non-zero at finite  $N_s$ , although it could vanish as  $N_s \rightarrow \infty$ , in such a way that the normalized  $\tilde{\Phi}$  and  $\mathcal{M}$  involving the ratios of  $Z$ -like objects remain non-zero. Let us also observe that  $\mathcal{M}$  vanishes on turning off interactions. We comment on its relation to the conventional Dyson selfenergy below after Eq. (45).

We next argue that Eq. (32) implies Eq. (4) provided the interacting system is continuously connected to the gas limit. For the strictly monotonic case  $\mu(m) < \mu(m+1)$ , there is a  $1 \leftrightarrow 1$  map between the  $\vec{k}\sigma$  and the  $\mu(m)$ , extending the obvious map in the gas. Hence  $\tilde{\Phi}_m = \delta_{m, m_0}$  for some  $m_0$ , whereby  $\gamma_{\vec{k}\sigma} = \mu - \mu(m_0)$ . Summing over all  $\vec{k}\sigma$  leads to Eq. (4). This property of a sum over all  $\vec{k}\sigma$  recovering the sum over all  $\mu(m)$  follows from the completeness of Fermi operators with the  $\vec{k}\sigma$  labels, i.e. there is no state in the Hilbert space that is inaccessible by a combination of these operators.

We might relax strict monotonicity of  $\mu(m)$  and allow for the merging of a set of  $\mu(m)$  at different  $m$  with say  $\mu(m_0)$ . In this case  $\tilde{\Phi}_m$  are non-zero for the set of  $m$  with non vanishing matrix elements in  $Z_{\sigma N_s}(\vec{k}, m)$ . Summing over these  $m$ 's we again get  $\gamma_{\vec{k}\sigma} = \mu - \mu(m_0)$ . Further summing over all  $\vec{k}\sigma$  gives us back Eq. (4), with a suitable degeneracy factor, provided we use completeness of the sum. We verify completeness in the noninteracting case (including shell type degeneracies) by using the representation Eq. (35), with  $\mathcal{M} = 0$ . In an interacting theory this completeness requires invoking isothermal continuity.

#### 5. Zero temperature Limit (II), i.e. $\{T \rightarrow 0, \frac{1}{L} \rightarrow 0\}$

We now consider the log-derivative  $\gamma_{\vec{k}}(\mu, T)$  for Fermi liquids as well as 1-d TLM. We are interested in calculating the  $T \rightarrow 0^+$  limit of  $\gamma_{\vec{k}}(\mu, T)$  near its root.

In order to calculate  $\alpha$  (Eq. (27)) and  $\beta$  (Eq. (28)) (dropping the explicit spin label below) we use the spectral function representations for the time dependent  $G$  detailed in Appendix A. We start

with Eq. (A.4) where we put  $\tau = \frac{1}{2}\beta$  so that

$$\alpha_{\vec{k}}(\mu, T) = \int_{-\infty}^{\infty} \frac{d\omega}{2 \cosh(\beta\omega/2)} \rho_G(\vec{k}, \omega), \quad (37)$$

$$\beta_{\vec{k}}(\mu, T) = - \int_{-\infty}^{\infty} \frac{\omega d\omega}{2 \cosh(\beta\omega/2)} \rho_G(\vec{k}, \omega), \quad (38)$$

$$\gamma_{\vec{k}}(\mu, T) = \beta_{\vec{k}}/\alpha_{\vec{k}}. \quad (39)$$

### 5.1. Fermi liquids

The spectral function in a Fermi liquid can be expressed for low  $T$ ,  $|\omega| \ll T_F$  as a Lorentzian [4]

$$\rho_G(\vec{k}, \omega) \sim \frac{Z(\vec{k})}{\pi} \frac{\Gamma_k}{\Gamma_k^2 + (\omega - E(\vec{k}, T))^2}, \quad (40)$$

where the quasiparticle weight

$$Z^{-1}(\vec{k}) = 1 - \partial_{\omega} \Sigma(\vec{k}, \omega)|_{\omega \rightarrow 0}, \quad (41)$$

and the width of the peak  $\Gamma_k = -Z(\vec{k})\Sigma''(\vec{k}, 0, T)$ , these are implicitly functions of  $k, T, \mu$  etc. Note that  $\Gamma_k \sim T^2$  is the standard Fermi liquid result for this object. The quasiparticle energy is defined as usual from the root of the nonlinear equation

$$E(\vec{k}, T) = \varepsilon(\vec{k}) + \Sigma'(\vec{k}, E, T) - \mu(T), \quad (42)$$

and  $\Sigma'$  ( $\Sigma''$ ) denotes the real (imaginary) part of  $\Sigma$ . From Eq. (B.7) and Eq. (B.5) and using the convenient symbol

$$W(\vec{k}, T) = \frac{1}{2} + \frac{\Gamma_k + iE(\vec{k}, T)}{2\pi T} \quad (43)$$

we deduce that

$$\begin{aligned} \alpha_{\vec{k}}(\mu, T) &= \frac{Z(\vec{k})}{\pi} \Re e \xi(W) \\ \beta_{\vec{k}}(\mu, T) &= -E(\vec{k}, T)\alpha_{\vec{k}}(\mu, T) - \frac{Z(\vec{k})\Gamma_k}{\pi} \Im m \xi(W) \\ \gamma_{\vec{k}}(\mu, T) &= -E(\vec{k}, T) - \Gamma_k \frac{\Im m \xi(W)}{\Re e \xi(W)}. \end{aligned} \quad (44)$$

In the limit  $T \rightarrow 0^+$  the width  $\Gamma_k$  vanishes and we obtain

$$\gamma_{\vec{k}}(\mu, 0^+) = -E(\vec{k}, 0). \quad (45)$$

Comparing the relations Eqs. (42), (45) with Eqs. (35), (36), we observe that the variable  $\mathcal{M}$  is essentially the selfenergy  $\Sigma'$  from the perspective of **Limit-I**. A change in the sign of  $\gamma_{\vec{k}}$  therefore occurs at the zero of  $E(\vec{k}, 0)$ . Close to this root, i.e. with small  $E$  Eq. (42) and the Dyson expression for the Greens function give us

$$\Theta(\gamma_{\vec{k}}(\mu, 0^+)) = \Theta(G^{-1}(\vec{k}, \omega = 0|\mu)), \quad (46)$$

and by replacing the static Greens function  $G^{-1} \rightarrow G$  we obtain the sumrule Eq. (5). We return to these expressions later in Section 7, where we carry out a detailed analysis of the volume of the Fermi surface in connection with ARPES experiments.

### 5.2. Non-Fermi liquids in 1-d

In this section we apply our method to the case of Tomonaga–Luttinger systems. This is an extensively studied area where many methods for exact solution are available [37–39]. In these

systems the quasiparticle weight  $Z$  vanishes, in parallel to the discussion for **Limit (I)** after Eq. (36). We show below that as  $T \rightarrow 0$  in **Limit(II)**,  $\alpha_{\hat{k}}$  vanishes as well, and so does  $\beta_{\hat{k}}$  in such a way that  $\gamma_{\hat{k}}$  remains finite and switches sign at the true Fermi wave-vector.

For the canonical example of a spinless model, the spectral function is severely constrained by the Lorentz invariance of the theory and conformal invariance of the effective 2-d classical theory at finite  $T$  [37,38,40]. It can be expressed as a scaling function valid at low  $T$ ,  $\omega$  and  $\hat{k}$  defined as  $\hat{k} = k - \zeta k_*$  near the left ( $\zeta = -1$ ) and right ( $\zeta = 1$ ) Fermi points  $\mp k_*$  [37,39,40]

$$\rho_G(k, \omega) = \frac{1}{T^{\alpha_0}} \sum_{\zeta=\pm 1} \mathcal{F}\left(\frac{\omega - \zeta V \hat{k}}{T}\right), \quad (47)$$

where  $V$  the renormalized Fermion velocity is related to the bare Fermi velocity  $V_F$  by a non singular scaling factor, and we set  $k_B = 1$  in this section. Here and in the following we should retain only one of the two terms of the  $\zeta$  sum, where  $\hat{k}$  is small. Although we did not specify the value of  $k_*$  yet, it will turn out that  $k_* = k_F$  below, thanks to the sumrule. The exponent  $\alpha_0 < 1$ , both  $\alpha$ ,  $V$  depend on the interaction strength, and the positive definite scaling function is peaked at the origin. It satisfies  $\mathcal{F}(0) = 1$  and  $\mathcal{F}(\xi) \rightarrow 1/|\xi|^{\alpha_0}$  for  $|\xi| \gg 1$ . As  $T \rightarrow 0^+$  we obtain

$$\rho_G(k, \omega) \sim \sum_{\zeta=\pm 1} \frac{\mathcal{A}}{|\omega - \zeta V \hat{k}|^{\alpha_0}}, \quad (48)$$

with  $\mathcal{A} > 0$ . From the above spectral function and Eq. (A.5) we can calculate the Greens function near zero frequency close to the Fermi points with  $T \rightarrow 0^+$  as

$$G(k, 0|\mu) = -\mathcal{B} \frac{\zeta V \hat{k}}{|V \hat{k}|^{\alpha_0+1}}, \quad (49)$$

where  $\mathcal{B} > 0$ , and  $\zeta = \pm 1$  for the right and left Fermi points.

We next calculate Eqs. (37) and (38) using Eq. (47). The  $\cosh(\frac{1}{2}\beta\omega)$  factor in Eq. (37) cuts off frequencies with  $|\omega| > T$ , and if we restrict  $|V \hat{k}| \lesssim T$  as well, then the dimensionless argument of the scaling function  $\mathcal{F}$  in Eq. (47) is at most of  $\mathcal{O}(1)$ , and we get no contribution to the integrals from a regime where  $\mathcal{F}(\xi) \rightarrow 1/|\xi|^{\alpha_0}$ . We can therefore reasonably replace  $\mathcal{F}$  by a Lorentzian

$$\mathcal{F} \sim \frac{cT}{\pi} \frac{cT}{(cT)^2 + (\omega - \zeta V \hat{k})^2}, \quad (50)$$

where  $c$  is a positive constant. This enables the convenience of an explicit evaluation of the integrals. If needed it can be supplanted by a more lengthy and tedious argument that avoids this replacement, giving the same answer.

We therefore use the results Eq. (B.7) and Eqs. (B.8), (B.9) to explicitly perform the integrals and write down at low  $T$  the results when  $V \hat{k}$  is small;

$$\alpha_{\hat{k}}(\mu, T) = \frac{c}{\pi} T^{1-\alpha_0} \Re e \xi \left( \frac{1}{2} + \frac{cT + iV \hat{k}}{2\pi T} \right) \quad (51)$$

$$\gamma_{\hat{k}}(\mu, T) = -\zeta V \hat{k} - cT \frac{\Im m \xi \left( \frac{1}{2} + \frac{cT + i\zeta V \hat{k}}{2\pi T} \right)}{\Re e \xi \left( \frac{1}{2} + \frac{cT + iV \hat{k}}{2\pi T} \right)} \quad (52)$$

and  $\beta_{\hat{k}}(\mu, T) = \gamma_{\hat{k}}(\mu, T) \alpha_{\hat{k}}(\mu, T)$ . Here  $\zeta = \pm 1$  for the right and left Fermi points. Note that these equations closely resemble Eq. (44). At finite  $T$  both terms in Eq. (52) vanish when  $V \hat{k}$  vanishes. As  $T \rightarrow 0^+$  the second term in Eq. (52) drops out identically, and we get

$$\gamma_{\hat{k}}(\mu, 0^+) = -\zeta V \hat{k}. \quad (53)$$

Comparing with the static Greens function Eq. (49) we obtain

$$\lim_{T \rightarrow 0^+} \Theta(\gamma_{\vec{k}}(\mu, T)) = \Theta(G(k, 0|\mu)) \quad (54)$$

we get the sumrule Eq. (5).

We note that the vanishing of the quasiparticle weight  $Z$  in this model is reflected in the vanishing of  $\alpha_{\vec{k}}$  in Eq. (51) at  $T \rightarrow 0$ . Away from the Fermi points  $\beta_{\vec{k}}$  also vanishes but their ratio  $\gamma_{\vec{k}}$  in Eq. (52) is finite. It follows that

$$2 \sum_k \Theta(G(k, 0|\mu)) = 2 \sum_k \Theta(G_0(k, 0|\mu)) \quad (55)$$

since each side equals the number of particles and therefore equates the Fermi diameter of the interacting and non interacting theories. Therefore the unspecified  $k_*$  can now be identified with the bare Fermi momentum  $k_F$ .

## 6. Sumrules in the singlet superconducting state

The volume theorem can be generalized to singlet superconductors. Our work is inspired by an observation regarding Gor'kov's (diagonal) Greens function describing the superconducting state [32] (see Eq. (14))

$$G(\vec{k}, \omega) = \frac{1}{2} \frac{u^2(\vec{k})}{\omega - E(\vec{k}) + i0^+} + \frac{1}{2} \frac{v^2(\vec{k})}{\omega + E(\vec{k}) + i0^+}, \quad (56)$$

$$u^2(\vec{k}) = \frac{1}{2} \left( 1 + \frac{\varepsilon(\vec{k}) - \mu}{E(\vec{k})} \right) \text{ and } v^2(\vec{k}) = \frac{1}{2} \left( 1 - \frac{\varepsilon(\vec{k}) - \mu}{E(\vec{k})} \right). \quad (57)$$

Here  $\varepsilon(\vec{k})$  is the band energy,  $E(\vec{k}) = \sqrt{\varepsilon^2(\vec{k}) + \Delta^2(\vec{k})}$  the (positive) quasiparticle energy and  $\Delta(\vec{k})$  the gap function. It is remarkable to note that this expression contains in its innards, a precise encoding of the (submerged) normal state Fermi surface. Setting  $\omega = 0$  we find

$$G(\vec{k}, 0) = \frac{\mu - \varepsilon(\vec{k})}{E^2(\vec{k})}. \quad (58)$$

Therefore in system exhibiting superconductivity, the submerged normal state Fermi surface is revealed by a change in sign of  $G(\vec{k}, 0)$  occurring at

$$\varepsilon(\vec{k}) = \mu, \quad (59)$$

and at the root,

$$u(\vec{k}_F) = v(\vec{k}_F). \quad (60)$$

The latter condition expresses an emergent particle–hole symmetry on the Fermi surface of the weak coupling BCS solution. While the above relations are true at the mean-field (BCS) level of description, it is not clear if this encoding survives the effects of strong interactions, and further refinements of the theory. It is also not entirely clear as to how one might probe this encoding, since there is no known method for probing the static  $G$  directly. The first question is treated here with an affirmative answer for a fully gapless superconductor. For a partially or fully gapped case, it is subject to the survival of the particle–hole symmetry as at least an approximate symmetry for arbitrary coupling. The second is answered in Section 7, where we relate an observable first moment of frequency with a suitable weight function to this observation, thereby suggesting a potentially useful photoemission experiment.

The strategy used for the normal state is extended to the superconductors as follows. We first establish the thermodynamic sumrule Eq. (7), under the assumption that (Cooper) pairs of electrons exhibit mutual repulsion, when viewed as composite particles. We then take the  $T = 0$  limit to

obtain the sumrule Eq. (8). The main assumption here, used without a direct proof, is that of nearly particle–hole symmetric matrix elements for the interacting system, analogous to Eq. (60) for the free case. Finally we take the  $L \rightarrow \infty$  limit and using results from the Nambu–Gorkov formalism, obtain Eq. (75) and hence the sumrule Eq. (5). This completes the set of sumrules needed to establish Eq. (1) for the superconducting state as well.

### 6.1. Superconducting phase: Thermodynamic sumrule

We next study the thermodynamic sumrule for a superconducting state, using the canonical ensemble. This approach is familiar from the nuclear physics context [41,42] and has been recently applied in the context of mesoscopic superconductivity [43]. Our treatment initially assumes a finite gap such as s-wave BCS superconductors, and later generalized to include d-wave case relevant for cuprate superconductors. We define the CE chemical potentials remaining within the *even-canonical* ensemble [41,43]:

$$\mu_e(2m) \equiv \frac{1}{2}(F_{2m+2} - F_{2m}), \quad (61)$$

and require the monotonic property  $\mu_e(2m+2) > \mu_e(2m)$  leading to an ordering

$$\mu_e(0) < \mu_e(2) < \dots < \mu_e(2j) \dots < \mu_e(N_{max}). \quad (62)$$

This clearly implies a concavity condition on the free energies  $2F_{2n} < F_{2n-2} + F_{2n+2}$ , arising from represents repulsion between pairs, so that further clusters of electrons are forbidden, i.e. *the pairing stops with pairs*. This results in a homogeneous many-body eigenstate of pairs, qualitatively similar and continuously connected to a gas of (repulsive) Bosons as envisaged in Ref. [44,45]. Assuming the ordering Eq. (62) we may repeat the discussion leading to Eq. (2), yielding Eq. (7), the number sum-rule for paired superconductors at  $T = 0$ , with the extra factor of 2 from skipping odd fillings.

### 6.2. Superconducting phase: $T = 0$ sumrule

We now consider the low T log-derivative Greens function as in Eq. (3) with  $F_m \sim E_0(m)$ . For this we also need the odd sector energies  $F_{2n+1}$ , these are expressed in terms of the even energies and a gap function  $\Delta$ [41–43,46–54]:

$$E_0(2n+1) = \frac{1}{2}(E_0(2n) + E_0(2n+2)) + \Delta(2n+1). \quad (63)$$

Here  $\Delta$  playing the role of the BCS gap is assumed non-zero initially. It is interpretable as the energy of an unpaired electron in an otherwise paired state. It is essentially the lowest energy Bogoliubov–Valatin [55] quasiparticle in the CE.

We consider the **Limit(I)** of  $\alpha, \beta$  of Eqs. (11), (12), (13), (14) of the main paper. We proceed to calculate Eqs. (13), (14) by grouping pairs of terms  $\{2m, 2m+1\}$ , rewriting  $p_\mu$ 's using Eqs. (27) and (29) as  $p_\mu(2m) = e^{\beta(2m\mu - F_{2m})}/Z(\mu)$  and  $p_\mu(2m+1) = e^{-\beta\Delta(2m+1)}p_\mu^{\frac{1}{2}}(2m)p_\mu^{\frac{1}{2}}(2m+2)$ . We further define the matrix elements:

$$\begin{aligned} V_{k\sigma}^{ab}(2m+1) &= \langle 2m+1, a | C_{\bar{k}\sigma} | 2m+2, b \rangle, \\ U_{k\sigma}^{ab}(2m+1) &= \langle 2m, a | C_{\bar{k}\sigma} | 2m+1, b \rangle. \end{aligned} \quad (64)$$

The ground state  $|2m+1, 0\rangle$  has one unpaired quasiparticle (with 2-fold degeneracy), while  $|2m, 0\rangle$  and  $|2m+2, 0\rangle$  are the fully paired non-degenerate ground states. These matrix elements are therefore analogs of the familiar GCE coefficients  $(u_k, v_k) = \sqrt{\frac{1}{2}(1 \pm \frac{\xi_k}{E_k})}$  of the BCS–Gor'kov theory [32,55,56] noted in Eq. (57), with  $\xi = \varepsilon - \mu$ . Recall that  $\xi_k = 0$  at the Fermi momentum, therefore the relation  $u_{k_F} = v_{k_F}$  noted in Eq. (60) holds good in weak coupling [32,55,56]. This relation also underlies the Majorana zero modes discussed in [57], and is often viewed as expressing an emergent particle–hole symmetry. Following this we assume the more general ground states

matrix elements satisfy  $|U_{\vec{k}\sigma}^{00}| \sim |V_{\vec{k}\sigma}^{00}|$ , for the correct bridging momentum. For finite systems we require it to hold within a tolerance that is discussed below.

We closely follow the procedure in the Fermi liquid case, and express  $\alpha$ ,  $\beta$  in terms of the matrix elements  $U$ ,  $V$ .

$$\beta_{\vec{k}}(\mu, T) = \sum_m e^{\frac{1}{2}\beta(\mu - \mu_e(2m) - \Delta)} p_{\mu}^{\frac{1}{2}}(2m) p_{\mu}^{\frac{1}{2}}(2m+2) \mathcal{B}(m),$$

and

$$\alpha_{\vec{k}}(\mu, T) = \sum_m e^{\frac{1}{2}\beta(\mu - \mu_e(2m) - \Delta)} p_{\mu}^{\frac{1}{2}}(2m) p_{\mu}^{\frac{1}{2}}(2m+2) \mathcal{A}(m), \quad (65)$$

where

$$\mathcal{A}(m) = \{(U_{\vec{k}\sigma}^{00}(2m+1))^2 + e^{-\beta(\mu - \mu_e(2m))} (V_{\vec{k}\sigma}^{00}(2m+1))^2\}, \quad (66)$$

$$\mathcal{B}(m) = (\mu - \mu_e(2m))\mathcal{A}(m) + \Delta(2m+1) \{(U_{\vec{k}\sigma}^{00}(2m+1))^2 - e^{-\beta(\mu - \mu_e(2m))} (V_{\vec{k}\sigma}^{00}(2m+1))^2\}. \quad (67)$$

In computing  $\gamma_{\vec{k}}(\mu, T)$  as  $T \rightarrow 0$ , our calculation proceeds similar to the non-superconducting case but with the role of  $Z(\vec{k}, m)$  now played by the matrix elements  $U$ ,  $V$ . Assuming continuity from the Fermi gas via the weak coupling BCS–Gor’kov theory, the given wave vector  $\vec{k}$  picks out a single particle number  $m$  contributing to both  $\alpha$ ,  $\beta$ . In the gapless case for the given  $\vec{k}$ ,  $\Delta$  vanishes as an inverse power of  $L$ . Thus  $\mathcal{B}/\mathcal{A} = \mu - \mu_e(2m)$  with negligible corrections. If the gap is non-zero the ratio  $\mathcal{B}/\mathcal{A} = \mu - \mu_e(2m) + \Delta C$  where on dropping indices:

$$C = \frac{e^{\beta(\mu - \mu_e(2m))} U^2 - V^2}{e^{\beta(\mu - \mu_e(2m))} U^2 + V^2}.$$

We require the correction  $\Delta C$  to be small relative to the separation between  $\mu_e(2m)$  and  $\mu_e(2m+2)$ . If particle–hole symmetry were exactly true then  $U = V$ ,  $\Delta C = 0$  and the node in  $\gamma$  is situated exactly at  $\mu = \mu_e(2m)$  even if  $\Delta \neq 0$ . In practice an approximate equality between  $U$  and  $V$  suffices for this condition with a specified tolerance. If we require that

$$\frac{|U^2 - V^2|}{U^2 + V^2} < \frac{|\mu_e(2m \pm 2) - \mu_e(2m)|}{\Delta(2m+1)}, \quad (68)$$

the node in  $\gamma$  at  $\mu \sim \mu_e(2m)$  is essentially unshifted. Assuming this relation and summing over  $\vec{k}$ , it follows that

$$\sum_{\vec{k}\sigma} \Theta(\gamma_{\vec{k}\sigma}(\mu, 0)) = 2 \sum_m \Theta(\mu - \mu_e(2m)) = \bar{N}(\mu),$$

as noted in Eq. (8), where the factor of 2 comes from the equal contribution from  $\vec{k}\sigma$  and its time reversed partner  $-\vec{k}\bar{\sigma}$ .

### 6.3. Superconducting phase: $T \rightarrow 0^+$ sumrule using Nambu–Gor’kov formalism

We next take the limit  $\{T \rightarrow 0, \frac{1}{L} \rightarrow 0\}$  in the superconducting state. In order to go beyond the mean-field treatment in the Gor’kov’s paper [32], we use the formally exact Nambu formalism [31]. It contains all possible many body effects, including those neglected in mean field theory. We start with the Nambu–Gor’kov [31–33] four component theory where the selfenergy in the superconducting state is expanded as

$$\Sigma(\vec{k}, z) = (1 - Z_{\vec{k}}(z))z \mathbb{1} + \phi(\vec{k}, z) \tau_1 + \chi(\vec{k}, z) \tau_3, \quad (69)$$

with  $z = i\omega_n$  where the Nambu selfenergies  $Z, \chi, \phi$  are even functions of  $z$ . In this notation for superconductors  $Z \sim 1 - \partial_\omega \Sigma$ , i.e the inverse of the normal state convention where  $Z \sim (1 - \partial_\omega \Sigma)^{-1}$ . From this the matrix Greens function  $\mathbf{G}$  is written as

$$\mathbf{G}(\vec{k}, z|\mu) = \frac{zZ_{\vec{k}}(z)\mathbb{1} + \tau_3(\varepsilon(\vec{k}) - \mu + \chi_{\vec{k}}(z)) + \tau_1\phi_{\vec{k}}(z)}{z^2Z_{\vec{k}}^2(z) - E_{\vec{k}}^2(z)}. \quad (70)$$

We are only interested in the diagonal Greens function  $G_{11}$ , which we shall denote by  $G$  below. This is the component of the Greens function relevant to the volume theorem and also to photoemission studies. It can be found within the quasiparticle approximation by expanding Eq. (70) near the poles of the Greens function [31,33–35]. The poles of  $G(\vec{k}, \omega)$  are located at the Bogoliubov–Valatin (B–V) [55] quasiparticle energies  $\omega = \pm E_{r\vec{k}}$  where

$$E_{r\vec{k}} = \Re e(E_{\vec{k}}(\eta_{\vec{k}})/Z_{\vec{k}}(\eta_{\vec{k}})), \quad \text{with} \quad \eta_{\vec{k}} = E_{r\vec{k}} + i0^+, \quad (71)$$

and have a width

$$\Gamma_{\vec{k}} = Z_{\vec{k}}^{-1} \Im m\{\eta_{\vec{k}}Z_{\vec{k}}(\eta_{\vec{k}}) - \frac{1}{E_{\vec{k}}}(\tilde{\varepsilon}_{\vec{k}}\chi_{\vec{k}}(\eta_{\vec{k}}) + \phi_{\vec{k}}\phi_{\vec{k}}(\eta_{\vec{k}}))\},$$

expressed in terms of the following set of real constants (Eq. (2.25,2.26,2.27) of [34]).

$$\begin{aligned} \tilde{\varepsilon}_{\vec{k}} &= \varepsilon(\vec{k}) - \mu + \Re e \chi_{\vec{k}}(\eta_{\vec{k}}), \quad \phi_{\vec{k}} = \Re e \phi_{\vec{k}}(\eta_{\vec{k}}) \\ E_{\vec{k}} &= (\tilde{\varepsilon}_{\vec{k}}^2 + \phi_{\vec{k}}^2)^{\frac{1}{2}}, \quad Z_{\vec{k}} = \Re e Z_{\vec{k}}(\eta_{\vec{k}}). \end{aligned} \quad (72)$$

In the above expression  $\phi_{\vec{k}}$  plays the role of a gap function,  $\tilde{\varepsilon}_{\vec{k}}$  the dispersion of a gapless underlying Fermi liquid renormalized with selfenergy  $\chi_{\vec{k}}$ , and  $E_{\vec{k}}$  is proportional to the quasiparticle energy  $E_{r\vec{k}}$ .

For energies close to the BV quasiparticle energies, the quasiparticle Greens function is given by the asymptotic expressions

$$G(\vec{k}, i\omega_n|\mu) \sim \sum_{\alpha=\pm 1} \left\{ \frac{1}{2} + \alpha \frac{\tilde{\varepsilon}_{\vec{k}}}{2E_{\vec{k}}} \right\} \frac{Z_{\vec{k}}^{-1}}{i\omega_n - \alpha E_{r\vec{k}} + i\Gamma_{\vec{k}}}, \quad (73)$$

$$\rho_G(\vec{k}, \omega) \sim \frac{1}{\pi} \sum_{\alpha=\pm 1} \left\{ \frac{1}{2} + \alpha \frac{\tilde{\varepsilon}_{\vec{k}}}{2E_{\vec{k}}} \right\} \frac{Z_{\vec{k}}^{-1}\Gamma_{\vec{k}}}{(\omega - \alpha E_{r\vec{k}})^2 + \Gamma_{\vec{k}}^2} \quad (74)$$

The spectral function has a similar status for superconducting systems as Eq. (40) for Fermi liquids; both expressions capture the various many-body renormalizations in terms of a few parameters.

We calculate  $\alpha_{\vec{k}}, \beta_{\vec{k}}$  from Eqs. (37), (38) using the spectral function Eq. (74) in greater detail below in Eq. (92) in Section 7.2. However as  $T \rightarrow 0^+$  it is known [34] that  $\Gamma_{\vec{k}} \rightarrow 0$ , i.e. one has sharp poles, and  $\rho_G$  is a sum over two delta functions. In this case we easily calculate

$$\alpha_{\vec{k}} \sim \frac{1}{Z_{\vec{k}} \cosh(\frac{1}{2}\beta E_{r\vec{k}})}, \quad \beta_{\vec{k}} \sim \frac{-\tilde{\varepsilon}_{\vec{k}}}{Z_{\vec{k}}^2 \cosh(\frac{1}{2}\beta E_{r\vec{k}})}$$

therefore  $\gamma_{\vec{k}}(\mu, T) \rightarrow -\frac{\tilde{\varepsilon}_{\vec{k}}}{Z_{\vec{k}}}$ . Now  $G(\vec{k}, 0|\mu) = -\frac{\tilde{\varepsilon}_{\vec{k}}}{E_{\vec{k}}^2}$  from Eq. (73), and therefore

$$\Theta(\gamma_{\vec{k}}(\mu, 0^+)) = \Theta\left(G(\vec{k}, 0|\mu)\right), \quad (75)$$

and therefore by summing over  $\vec{k}$  we obtain the sumrule Eq. (5). This result is argued to be valid for all flavors of singlet pairing, including the gapless d-wave case. We combine Eq. (75) or Eq. (5) with Eq. (8) and infer the sumrule Eq. (1) in the superconductor.

## 7. The pseudo Fermi surface at finite $T$

Extending the ground state sum-rule to finite  $T$ , we define a “pseudo-Fermi surface” and an effective density  $n_{\text{eff}}(T)$  from the changes in sign with  $\vec{k}$  of  $\gamma_{\vec{k}\sigma}(\mu, T)$ . These tend to the true Fermi surface and particle density when  $T \rightarrow 0$ , and can be extracted from experimental photoemission data as follows. In terms of a dipole matrix-element  $M$  and the Fermi function  $f(\omega) = (\exp \beta\omega + 1)^{-1}$ , the photoelectron intensity is given by  $\mathcal{I}(\vec{k}, \omega) = M(\vec{k})\rho_G(\vec{k}, \omega)f(\omega)$ . From Eqs. (3), (37), (38) it follows that  $\gamma$  is a suitably weighted first moment of frequency:

$$\gamma_{\vec{k}\sigma}(\mu, T) = -\langle \omega \rangle_{\vec{k}}, \quad (76)$$

where

$$\langle \omega \rangle_{\vec{k}} = \int d\omega \mathcal{I}(\vec{k}, \omega) e^{\frac{1}{2}\beta\omega} \omega / \int d\omega \mathcal{I}(\vec{k}, \omega) e^{\frac{1}{2}\beta\omega}, \quad (77)$$

$$= \int \rho_G(\vec{k}, \omega) \frac{\omega d\omega}{\cosh(\frac{1}{2}\beta\omega)} / \int \rho_G(\vec{k}, \omega) \frac{d\omega}{\cosh(\frac{1}{2}\beta\omega)}, \quad (78)$$

the two expressions Eqs. (77), (78) are equivalent since the  $\vec{k}$  dependent matrix element and other factors cancel out. This weight function was already mentioned in Eq. (9) in Section 1. In averaging over  $\omega$ , the weight factors provide exponential cutoffs for high  $|\omega|$ . By replacing  $\omega$  by  $\omega^m$  in Eq. (76), one can generate the  $m$ th moment  $\langle \omega^m \rangle_{\vec{k}}$ . This novel set of moments characterize the low energy excitations of the spectral function, unlike the moments without the  $T$  dependent weight functions, and seem promising for further study.

From  $\gamma$  we define the effective density

$$n_{\text{eff}}(T) = 1/N_s \sum_{\vec{k}\sigma} \Theta(\gamma_{\vec{k}\sigma}(\mu, T)). \quad (79)$$

We can now define the pseudo-Fermi surface at any  $T$ ; it is defined as the set of Fermi points  $\vec{k}$  satisfying

$$\langle \omega \rangle_{\vec{k}} = 0. \quad (80)$$

The sign changes of  $\gamma$  with  $\vec{k}$  occur on this surface, and  $n_{\text{eff}}(T)$  counts the number of particles within this surface from Eq. (79). It reduces to the true Fermi-surface at  $T = 0$ . We next discuss the content of this sum-rule at finite  $T$  for two important cases.

### 7.1. Finite $T$ volume sumrule: Fermi liquids

We note that Eqs. (76), (77) are identical to Eq. (39) in Section 6.3. Therefore for Fermi liquids at finite (but low)  $T$ , we can use the quasiparticle approximation for the spectral function Eq. (40), so that

$$\langle \omega \rangle_{\vec{k}} = E(\vec{k}, T) + \Gamma_k \frac{\Im m \xi \left( \frac{1}{2} + \frac{\Gamma_k + iE(\vec{k}, T)}{2\pi T} \right)}{\Re e \xi \left( \frac{1}{2} + \frac{\Gamma_k + iE(\vec{k}, T)}{2\pi T} \right)}, \quad (81)$$

following Eq. (44). In order to deduce the pseudo Fermi points, we observe that when  $E(\vec{k}, T)$  vanishes in Eq. (39), the imaginary part of  $\xi$  vanishes as well. Thus at any  $T$  the pseudo Fermi point is located by

$$E(\vec{k}, T) = 0, \quad (82)$$

where  $E(\vec{k}, T)$  is defined in Eq. (42). At  $T = 0$  it reduces to

$$E(\vec{k}_F, 0) = \varepsilon(\vec{k}_F) + \Sigma'(\vec{k}_F, 0, 0) - \mu(0) = 0, \quad (83)$$

where we set  $\vec{k} = \vec{k}_F$ , the corresponding  $T = 0$  Fermi momentum upon using the volume theorem. In Eq. (42) we expand the selfenergy at low  $\omega$  and write  $E$  as

$$Z^{-1}(\vec{k}, T)E(\vec{k}, T) = \varepsilon(\vec{k}) + \Sigma'(\vec{k}, 0, T) - \mu(T) \quad (84)$$

$$= \varepsilon(\vec{k}) + (\Sigma'(\vec{k}, 0, T) - \Sigma'(\vec{k}, 0, 0)) + \Sigma'(\vec{k}, 0, 0) + (\mu(0) - \mu(T)) - \mu(0) \quad (85)$$

The vanishing of the right hand side locates the pseudo FS. Using Eq. (83) we obtain

$$\begin{aligned} (\mu(T) - \mu(0)) &= (\varepsilon(\vec{k}) - \varepsilon(\vec{k}_F)) + (\Sigma'(\vec{k}, 0, T) - \Sigma'(\vec{k}, 0, 0)) \\ &\quad + (\Sigma'(\vec{k}, 0, 0) - \Sigma'(\vec{k}_F, 0, 0)). \end{aligned} \quad (86)$$

As expected this equation is satisfied identically by setting  $\vec{k} = \vec{k}_F$  at  $T = 0$ . At low  $T$  we perturb by expanding  $\vec{k}$  about  $\vec{k}_F$ ,

$$\vec{k} = \vec{k}_F + \delta\vec{k}, \quad (87)$$

and linearize in  $\delta\vec{k}$  to find

$$\delta\vec{k} \cdot \vec{V}_{\vec{k}_F} = (\mu(T) - \mu(0)) - (\Sigma'(\vec{k}_F, 0, T) - \Sigma'(\vec{k}_F, 0, 0)), \quad (88)$$

where  $\vec{V}_{\vec{k}} = \vec{\nabla}_{\vec{k}}[\varepsilon(\vec{k}) + \Sigma'(\vec{k}, 0, 0)]$  is related to the Fermi velocity. The variation  $\delta\vec{k}$  is normal to the true (i.e.  $T = 0$ ) FS, and can be determined from this relation. Proceeding further we may write the change in FS area with  $T$  as a line integral over the true FS in 2-d

$$\delta A(T) = \oint_{\text{FS}} dk_{\perp} \frac{(\mu(T) - \mu(0)) - (\Sigma'(\vec{k}_F, 0, T) - \Sigma'(\vec{k}_F, 0, 0))}{|\vec{V}_{\vec{k}}|}, \quad (89)$$

where  $dk_{\perp}$  is the wave-vector element tangential to the FS.

The effective density at  $T$  differs from the true particle density by the usual counting rules leading to

$$n_{\text{eff}}(T) - n = 2 \times \delta A(T)/(2\pi)^2. \quad (90)$$

The variation Eq. (89) is driven by the  $T$  dependent shifts of  $\mu(T)$  and of the real part of the selfenergy  $\Sigma'(\vec{k}_F, 0, T)$ . The shift of  $\mu$  with  $T$  is the smaller of the two, and can in principle be estimated experimentally. For example in ARPES the apparent change of excitation energy with  $T$  of some fixed ( $T$  independent) feature, such as a band edge can be used for this purpose. The variation  $\delta A(T)$  is amplified when the quasiparticle Fermi velocity is reduced from the bare one, as it often happens in strongly correlated matter. An example of the  $T$  dependence of  $n_{\text{eff}}$  in the  $t$ - $J$  model is shown in Ref. [58], where the variation with  $T$  is quite significant due to strong correlations, revealing emergent low-energy scales in the problem.

The expression Eq. (89) allows us to explore the  $T$  dependent shift of the real part of  $\Sigma$ . This object is of great interest in strongly correlated materials. In the strange metal regime of the  $d = \infty$  Hubbard model, it has been reported in Ref. [59] (Fig (12.c)) to have a strong  $T$  dependence, which in turn leads to a linear  $T$  resistivity

## 7.2. Finite $T$ volume sumrule: Superconductors

In parallel to the treatment of the normal case above, we calculate the first moment Eq. (78) in the superconducting phase, using the quasiparticle spectral function in Eq. (74), and the useful integrals noted in Eq. (B.7). Canceling common factors we write

$$\begin{aligned} \langle \omega \rangle_{\vec{k}} &= \mathcal{N}/\mathcal{D}, \\ \mathcal{N} &= \sum_{\alpha=\pm 1} \left\{ \frac{1}{2} + \alpha \frac{\tilde{\varepsilon}_{\vec{k}}}{2E_{\vec{k}}} \right\} \left[ \alpha E_{r\vec{k}} \Re e \xi \left( \frac{1}{2} + \frac{\Gamma_{\vec{k}} + i\alpha E_{r\vec{k}}}{2\pi T} \right) + \Gamma_{\vec{k}} \Im m \xi \left( \frac{1}{2} + \frac{\Gamma_{\vec{k}} + i\alpha E_{r\vec{k}}}{2\pi T} \right) \right] \\ \mathcal{D} &= \sum_{\alpha=\pm 1} \left\{ \frac{1}{2} + \alpha \frac{\tilde{\varepsilon}_{\vec{k}}}{2E_{\vec{k}}} \right\} \Re e \xi \left( \frac{1}{2} + \frac{\Gamma_{\vec{k}} + i\alpha E_{r\vec{k}}}{2\pi T} \right) \end{aligned} \quad (91)$$

We now use the properties of the  $\xi$  function Eqs. (B.8) and (B.9) to simplify Eq. (91). This gives the final formula for the first moment:

$$\langle \omega \rangle_{\vec{k}} = \frac{\tilde{\varepsilon}_{\vec{k}}}{E_{\vec{k}}} \left( E_{r\vec{k}} + \Gamma_{\vec{k}} \frac{\Im m \xi \left( \frac{1}{2} + \frac{\Gamma_{\vec{k}} + iE_{r\vec{k}}}{2\pi T} \right)}{\Re e \xi \left( \frac{1}{2} + \frac{\Gamma_{\vec{k}} + iE_{r\vec{k}}}{2\pi T} \right)} \right). \quad (92)$$

The  $\Gamma_{\vec{k}}$  term in the above expression is expected to be exponentially small for the gapped states and a power law for gapless singlet paired states. This expression resembles Eq. (81) for the Fermi liquid state with the energy dispersion  $\tilde{\varepsilon}_{\vec{k}}$  replacing the quasiparticle energy  $E(\vec{k}, T)$ . The vanishing of the first moment locates the pseudo-FS for the superconductor through the condition

$$\tilde{\varepsilon}_{\vec{k}} = \varepsilon(\vec{k}) - \mu + \Re e \chi_{\vec{k}}(E_{r\vec{k}}) \rightarrow 0, \quad (93)$$

which replaces the simple relation of Gor'kov's theory Eq. (59). This implies that the shift in the chemical potential from the noninteracting value due to pairing effects is exactly canceled by the selfenergy term  $\Re e \chi_{\vec{k}}(E_{r\vec{k}})$ . This cancellation is analogous to exact cancellation in the normal state.

Our treatment of the pseudo FS of the superconducting state has a few precedents. The closely related papers [60,61] discuss the surface formed by  $\vec{k}$  with  $G(\vec{k}, 0) = 0$ , using a phenomenological model of  $G$  for superconductors in strongly correlated cuprate materials. The model uses a “renormalized” mean field theory [62] for this calculation. This method incorporates effects of strong correlations through a rescaled version of the BCS effective Hamiltonian with density dependent scale factors. The area of the above surface in these works is found to be only approximately the number density, even at  $T = 0$ . Their results are in contrast to the findings of the present work, where the pseudo FS area must match the particle density exactly at  $T = 0$ . The discrepancy could be due to missing a cancellation between the shifts of the selfenergy and the chemical potential, or due to a lack of the (unproven) particle–hole symmetry at strong coupling. Experimental checks of the particle–hole symmetry, as suggested in this work would be of considerable interest.

In the present work we propose a new suggestion for probing the pseudo-Fermi surface for superconductors. It differs from the signatures proposed earlier [60,61], advocating either locating the maxima of the spectral weight, or the maxima of the gradient of the momentum distribution function  $|\nabla_{\vec{k}} n_{\vec{k}}|$ . Our proposal involves studying the first moment of the frequency  $\langle \omega \rangle_{\vec{k}}$ , defined in Eq. (78). Its vanishing as in Eq. (80) defines the pseudo FS. As explained above this moment can be constructed from the dynamical information in ARPES. For singlet superconductors such as the cuprates, the pseudo-Fermi surface is definable on both sides of the superconducting transition using the moment in both phases Eqs. (81), (92). Apart from (usually small)  $T$  dependent corrections, its area is the same in both phases, being related to the number of particles.

## 8. Discussion

Given the importance of the Fermi volume theorem, and the attendant complexities of deriving it, a fresh approach seems relevant. This work presents a non-perturbative derivation of the volume sum-rule Eq. (1) in a broad setting. We avoid using the traditional number sum-rule

$$\bar{N}(\mu) = \sum_{k\sigma} G_{\sigma}(\vec{k}, \tau = 0^- | \mu); \quad (94)$$

instead we use different ways to compute the zero  $T$  limit of the log-derivative  $\gamma_{\vec{k}}$ . This is a major departure from the L–W route, where the introduction of the Luttinger–Ward functional is an essential second step. This functional can only be defined in perturbation theory, and leads to difficulties for strong coupling problems, as explicitly demonstrated in recent work [63].

In 1-d Ref. [17] uses adiabatic evolution of the system with a magnetic flux parameter, to give a non-perturbative argument for the invariance of Fermi diameter. Ref. [16] extends this to arbitrary dimensions  $d > 1$  assuming that the system is a Fermi liquid. In contrast to the 1-d result, Ref. [16] also requires adiabatic evolution through a large number  $\mathcal{O}(L^{d-1})$  of level crossings, arising from a large accumulation of phase with increasing flux.

More generally the use of adiabatic theorem in gapless situations, particularly for  $d > 1$  is risky, and often requires extra symmetry for justification. A well known example is provided by a gapless metal for  $d > 1$ , with a varying interaction strength. When the symmetry is less than circular (or spherical),  $k$ -space redistribution always occurs upon varying the interaction constant. This results in a change of shape of the Fermi surface [7,64,65], implying that adiabaticity is violated.

We note a few points regarding perturbative arguments. The  $T = 0$  Brueckner–Gammel–Goldstone formalism [66] is based on the adiabatic theorem and uses the non-interacting Greens function  $G_0$  as the foundation for the perturbation expansion. Therefore the invariance of the Fermi volume, as well as its shape, are automatic byproducts, we get back what we initially put in. A critique of this method by Kohn and Luttinger [67] led to the L–W work. They used finite  $T$  perturbation theory instead, allowing for a  $k$ -space redistribution of occupied states [7,64,65]. However the problem of strong coupling remains. It is hard to see how the L–W method can be justified in strong coupling, recalling that it is predicated on the existence of the Luttinger–Ward functional, defined term by term in powers of the coupling. Recent work explicitly displays pathologies of the L–W functional in Hubbard type models at strong coupling [63].

The present work utilizes continuity, instead of perturbation, to bypass the strong coupling problem. Isothermal continuity breaks down at level crossing transitions with tuning, and is signaled by a jump in expectation values. Therefore the guarantor of isothermal continuity is the absence of jump discontinuities in expectation values. In summary we may assume isothermal continuity within a continuously connected phase of matter, thus requiring the absence of first order quantum transitions. As our example of the superconductor shows, the isothermal argument works through the normal to superconducting transition, where the dependence on coupling is non-analytic (but continuous). Here the adiabatic methods seem to fail.

We use continuity in a parameter for linking the interacting system with the Fermi gas. The parameter used is most often, but not necessarily, the coupling constant in the Hamiltonian. In the case of the  $t$ - $J$  model with extreme coupling  $U = \infty$ , a more general interaction type parameter  $\lambda \in [0, 1]$  is invoked [25]. Continuous evolution with  $\lambda$  ensures the volume theorem for the  $t$ - $J$  model [25].

Our extension of the volume theorem to singlet superconductor is based on two assumptions. Firstly we assume that pairs of electrons act repulsively with respect to other pairs, thereby giving a monotonically increasing chemical potential  $\mu_e(m)$  in Eq. (62). This is certainly true in the BCS theory and in exactly solvable models [46,48] for superconductivity in finite size systems. It would break down if an as yet undiscovered glue were to result in say Cooper-quartets, instead of Cooper-pairs. The other main assumption is that of an approximately valid particle–hole symmetry Eq. (68) for the case of fully or partially gapped superconductors. This leads to  $U \sim V$  in the correlated superconductors, extending the known result Eq. (60) in the weak coupling BCS–Gor’kov case. This symmetry has been assumed to be true in other contexts, e.g. for the recently discussed Majorana Fermions [57]. For strongly coupled systems, this symmetry is hard to establish analytically. However numerical tests of the condition Eq. (68) involving ground-state to ground state matrix elements of the Fermi operators may be feasible for small systems using exact diagonalization, and are planned for future work. Finally since we establish a direct connection with observable variables, one could test the resulting sumrule experimentally in a variety of superconducting materials. The results would indicate if this symmetry holds good, and how widely, if so.

The present work leads to the notion of a pseudo-Fermi surface defined finite  $T$ . This surface is shown here to be accessible to ARPES studies from moments of the observed intensities. It seems well worth exploring this object and its  $T$  dependence experimentally to throw light on interesting issues in strongly correlated matter. For superconductors such measurements could complement information from the high magnetic field setups used to study the same submerged normal state Fermi surface by destroying the superconducting order using strong magnetic fields [68].

## Acknowledgments

I am grateful to P. W. Anderson, P. Coleman, B. Doucot, A. Georges, A. C. Hewson, H. R. Krishnamurthy, E. Perepelitsky, M. Randeria, and A. Tsvelik for helpful discussions on aspects of

this problem at various times. The work at UCSC was supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award No. DE-FG02-06ER46319.

### Appendix A. Spectral function and its relation to the Greens function

With  $-\beta < \tau \leq \beta$ , we recall the (Matsubara) imaginary time Greens function [2–4]

$$G_{\sigma}(k, \tau) = -\frac{1}{Z(\mu)} \text{Tr} e^{-\beta \mathcal{H}} \left( T_{\tau} C_{k\sigma}(\tau) C_{k\sigma}^{\dagger} \right), \quad (\text{A.1})$$

where the time dependence is  $Q(\tau) = e^{\tau \mathcal{H}} Q e^{-\tau \mathcal{H}}$ . Using the usual antiperiodicity  $G(\tau) = -G(\tau + \beta)$  we define the Fourier version as usual  $G(i\omega_n) = \frac{1}{2} \int_{-\beta}^{\beta} G(\tau) e^{i\omega_n \tau} d\tau$ . We may express  $G$  as

$$G(\vec{k}, i\omega_n) = \int d\omega \frac{\rho_G(\vec{k}, \omega)}{i\omega_n - \omega}, \quad (\text{A.2})$$

where the spectral function  $\rho_G(\vec{k}, \omega)$  can be conveniently found from the analytic continuations  $i\omega_n \rightarrow z$  followed by  $z \rightarrow \omega + i0^+$  as  $\rho_G(\vec{k}, \omega) = -\frac{1}{\pi} \Im m G(\vec{k}, \omega + i0^+)$ .

The spectral function has a further representation [4]

$$\begin{aligned} \rho_G(k, \omega) &= (1 + e^{-\beta\omega}) \sum_{n,m,a,b} p_{\mu}(n) |\langle n, a | C_{\vec{k}} | m, b \rangle|^2 \\ &\times \delta(\omega + E_a(n) - E_b(m) + \mu) e^{-\beta E_a(n) + \beta F_n}, \end{aligned} \quad (\text{A.3})$$

where  $F_n$  is the  $n$ -particle free energy.

In terms of  $\Delta_E \sim |E_a(n) - E_{a'}(n)|$ , i.e. a typical excitation energy at a fixed number for a finite system, we may distinguish between two regimes. At zero  $T$ , or more generally for  $\Delta_E/k_B \gtrsim T$ , the spectral function is a sum over separated delta functions and hence is very grainy. On the other hand provided  $(T, \omega) \gtrsim \Delta_E/k_B$ , the sum over the delta functions is taken over several states and hence the resulting spectral functions are smooth functions of  $\omega$ . This is therefore a complementary regime to the earlier one.

In terms of the spectral functions we may write the time dependent functions as

$$G(\vec{k}, \tau) = \int_{-\infty}^{\infty} d\omega \rho_G(\vec{k}, \omega) e^{-\tau\omega} (f(\omega)\theta(-\tau) - \bar{f}(\omega)\theta(\tau)) \quad (\text{A.4})$$

with the Fermi functions  $f(\omega) = \frac{1}{e^{\beta\omega} + 1}$  and  $\bar{f} = 1 - f$ . We will need the following relation for the real part of the Greens function

$$G(k, 0) = -\mathcal{P} \int \frac{d\omega}{\omega} \rho_G(k, \omega), \quad (\text{A.5})$$

where  $\mathcal{P}$  denotes the principal value.

### Appendix B. Some useful integrals arising in the sum-rule Eq. (78)

We outline the calculation of integrals that arise in Eq. (78):

$$\mathcal{J}_m = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \frac{(-\omega)^m}{\cosh(\frac{1}{2}\beta\omega)} \frac{\Gamma}{\Gamma^2 + (\omega - E)^2}, \quad (\text{B.1})$$

for real parameters  $\Gamma, E$  with  $m = 0, 1$ . A simple way to do these integrals is to use the Mittag-Leffler expansion

$$\frac{1}{\cosh(\pi z)} = \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n (n + \frac{1}{2})}{z^2 + (n + \frac{1}{2})^2}, \quad (\text{B.2})$$

so that we can integrate term by term using the simple result for convolution of two Lorentzians. This yields

$$\begin{aligned}\mathcal{J}_0 &= \frac{1}{\pi} \sum_{n=0}^{\infty} (-1)^n \frac{\Gamma/(2\pi T) + n + \frac{1}{2}}{E^2/(2\pi T)^2 + (\Gamma/(2\pi T) + n + \frac{1}{2})^2} \\ \mathcal{J}_1 &= -\frac{1}{\pi} \sum_{n=0}^{\infty} (-1)^n \frac{(n + \frac{1}{2})E}{E^2/(2\pi T)^2 + (\Gamma/(2\pi T) + n + \frac{1}{2})^2}.\end{aligned}\quad (\text{B.3})$$

These sums can be performed using the digamma function

$$\Psi(z) = \frac{d}{dz} \log \Gamma(z) = \lim_{M \rightarrow \infty} \left( \log M - \sum_{n=0}^M \frac{1}{z+n} \right).\quad (\text{B.4})$$

We define a meromorphic function  $\xi(z)$  via the alternating infinite sum

$$\xi(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{z+n} = \frac{1}{2} \left( \Psi\left(\frac{1}{2} + \frac{z}{2}\right) - \Psi\left(\frac{z}{2}\right) \right),\quad (\text{B.5})$$

In the complex  $z$  plane  $\xi(z)$  has a pole at the origin and at every negative integers, and is analytic everywhere else. Writing  $z = x + iy$  we record the useful corollaries

$$\begin{aligned}\Re \xi(x + iy) &= \sum_{n=0}^{\infty} (-1)^n \frac{x+n}{(x+n)^2 + y^2} \\ \Im \xi(x + iy) &= -\sum_{n=0}^{\infty} (-1)^n \frac{y}{(x+n)^2 + y^2}.\end{aligned}\quad (\text{B.6})$$

Using these we can perform the required summations in Eq. (B.3) analytically as

$$\begin{aligned}\mathcal{J}_0 &= \frac{1}{\pi} \Re \xi \left( \frac{1}{2} + \frac{\Gamma + iE}{2\pi T} \right) \\ \mathcal{J}_1 &= -\frac{E}{\pi} \Re \xi \left( \frac{1}{2} + \frac{\Gamma + iE}{2\pi T} \right) - \frac{\Gamma}{\pi} \Im \xi \left( \frac{1}{2} + \frac{\Gamma + iE}{2\pi T} \right).\end{aligned}\quad (\text{B.7})$$

From the series defining  $\xi(z)$  in Eq. (B.5), it is real for real  $z$ . Using the Schwarz reflection principle we deduce relations needed in the text; for  $\alpha = \pm 1$

$$\Re \xi \left( \frac{1}{2} + \frac{\Gamma + i\alpha E}{2\pi T} \right) = \Re \xi \left( \frac{1}{2} + \frac{\Gamma + iE}{2\pi T} \right)\quad (\text{B.8})$$

$$\Im \xi \left( \frac{1}{2} + \frac{\Gamma + i\alpha E}{2\pi T} \right) = \alpha \Im \xi \left( \frac{1}{2} + \frac{\Gamma + iE}{2\pi T} \right).\quad (\text{B.9})$$

## References

- [1] J.M. Luttinger, J.C. Ward, *Phys. Rev.* 118 (1960) 1417.
- [2] T. Matsubara, *Progr. Theoret. Phys.* 14 (1955) 351.
- [3] A. Abrikosov, L. Gor'kov, I.E. Dzyaloshinski, *Sov. Phys. J.E.T.P.* 36 (1959) 900.
- [4] A. Abrikosov, L. Gor'kov, I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, Englewood-Cliffs, NJ, 1963.
- [5] J.M. Luttinger, *Phys. Rev.* 119 (1960) 1153, (esp. Eqs. (31, 32, 35)).
- [6] L.D. Landau, *Sov. Phys. J.E.T.P.* 30 (1956) 1058, *Sov. Phys. J.E.T.P.* 32, 59.
- [7] P. Nozières, *Theory of Interacting Fermi Systems*, W. A. Benjamin, New York, 1964.
- [8] C. De Dominicis, *J. Math. Phys.* 3 (1964) 983;  
C. De Dominicis, P.C. Martin, *J. Math. Phys.* 5 (14) (1964) 31.
- [9] M. Dahmen, G. Jona-Lasinio, *Il Nuovo Cimento LII A* (1967) 807.
- [10] J.S. Langer, V. Ambegaokar, *Phys. Rev.* 121 (1961) 1090.
- [11] D.C. Langreth, *Phys. Rev.* 150 (1966) 516.

- [12] B.L. Altshuler, A.V. Chubukov, A. Dashevskii, A.M. Finkel'stein, D.K. Morr, E.P.L. 41 (1998) 401.
- [13] I. Dzyaloshinskii, *Phys. Rev. B* 68 (2013) 085113.
- [14] K.B. Blagoev, K.S. Bedell, *Phys. Rev. Lett.* 79 (1997) 1106.
- [15] K.B. Dave, P.W. Phillips, C.L. Kane, *Phys. Rev. Lett.* 110 (2013) 090403.
- [16] M. Oshikawa, *Phys. Rev. Lett.* 84 (2000) 3370.
- [17] M. Yamanaka, M. Oshikawa, I. Affleck, *Phys. Rev. Lett.* 79 (1997) 1110.
- [18] T. Senthil, S. Sachdev, M. Vojta, *Phys. Rev. Lett.* 90 (2003) 216403;  
T. Senthil, S. Sachdev, M. Vojta, *Physica B* 359 (2005) 9;  
T. Senthil, M. Vojta, S. Sachdev, *Phys. Rev. B* 69 (2004) 035111.
- [19] J. Kokalj, P. Prelovsek, *Eur. Phys. J. B* 63 (2008) 431; *Phys. Rev. B* 78 (2008) 153103.
- [20] B. Farid, A.M. Tselik, arXiv:0909.2886 (2009).
- [21] O.J. Curtin, Y. Nishikawa, A.C. Hewson, *J. Phys. Commun.* 2 (2018) 031001.
- [22] Y. Nishikawa, O.J. Curtin, A.C. Hewson, D.J.G. Crow, *Phys. Rev. B* 98 (2018) 104419.
- [23] E. Quinn, *Phys. Rev. B* 97 (2018) 115134.
- [24] K. Seki, S. Yunoki, *Phys. Rev. B* 96 (2017) 085124.
- [25] B.S. Shastry, *Phys. Rev. Lett.* 107 (2011) 056403, <http://physics.ucsc.edu/~sriram/papers/ECFL-Reprint-Collection.pdf>.
- [26] B.S. Shastry, *Phys. Rev. B* 81 (2010) 045121.
- [27] W.V. Liu, F. Wilczek, *Phys. Rev. Lett.* 90 (2003) 047002; *Phys. Rev. Lett.* 94 (2005) 017001.
- [28] N. Trivedi, M. Randeria, *Phys. Rev. Lett.* 75 (1995) 312;  
M. Randeria, N. Trivedi, A. Moreo, R.T. Scalettar, *Phys. Rev. Lett.* 69 (1992) 2001;  
N. Trivedi, R.T. Scalettar, M. Randeria, *Phys. Rev. B* 54 (1996) R3756.
- [29] L. van Hove, *Physica* 16 (1950) 137;  
C.N. Yang, T.D. Lee, *Phys. Rev.* 87 (1952) 404;  
M.E. Fisher, *Physics* 3 (1967) 255;  
Joel L. Lebowitz, International Symposium on Mathematical Problems in Theoretical Physics, Kyoto, Japan, 1975.
- [30] J.L. Lebowitz, E.H. Lieb, *Phys. Rev. Lett.* 22 (1969) 13;  
E.H. Lieb, J.L. Lebowitz, *Adv. Maths.* 9 (1972) 316;  
E.H. Lieb, R. Seiringer, *The Stability of Matter in Quantum Mechanics*, Cambridge University Press, 2010.
- [31] Y. Nambu, *Phys. Rev.* 117 (1960) 648.
- [32] L.P. Gor'kov, *Sov. Phys.—JETP* 7 (1958) 505.
- [33] J.R. Schrieffer, *Theory of Superconductivity*, Benjamin, New York, 1964.
- [34] L. Tewordt, *Phys. Rev.* 128 (1962) 12.
- [35] D.J. Scalapino, J.R. Schrieffer, J.W. Wilkins, *Phys. Rev.* 148 (1966) 263.
- [36] S. Kivelson, S. Chakravarty, *Europhys. Lett.* 16 (1991) 751;  
S.R. White, S. Chakravarty, M.P. Gelfand, S.A. Kivelson, *Phys. Rev. B* 45 (1992) 5062.
- [37] A.O. Gogolin, A.A. Nersesyan, A.M. Tselik, *Bosonization and Strongly Correlated Systems*, Cambridge University Press, Cambridge, 1998.
- [38] F.H.L. Essler, H. Frahm, F. Göhmann, A. Klümper, V.E. Korepin, *The One-Dimensional Hubbard Model*, Cambridge University Press, Cambridge, 2005.
- [39] T. Giamarchi, *Quantum Physics in One Dimension*, Clarendon Press, Oxford, 2004.
- [40] S. Eggert, A. Mattsson, J. Kinaret, cond-mat/9705157 (vide Eq. (15)).
- [41] A.B. Migdal, *Nuclear Phys.* 13 (1959) 655.
- [42] A. Bohr, B.R. Mottleson, D. Pines, *Phys. Rev.* 110 (1958) 936.
- [43] B. Janko, A. Smith, V. Ambegaokar, *Phys. Rev. B* 50 (1994) 1152;  
R.A. Smith, V. Ambegaokar, *Phys. Rev. Lett.* 77 (1996) 4962.
- [44] A.J. Leggett, *J. Phys. (Paris)* 41 (1980) C7–19.
- [45] P. Nozières, S. Schmitt-Rink, *J. Low. Temp. Phys.* 59 (1985) 195.
- [46] R.W. Richardson, *J. Math. Phys.* 18 (1977) 1802;  
R.W. Richardson, N. Sherman, *Nucl. Phys.* 52 (1964) 221; *Nucl. Phys.* 52 (1964) 253.
- [47] M.T. Tuominen, J.M. Hergenrother, T.S. Tighe, M. Tinkham, *Phys. Rev. Lett. B* 9 (1992) 1997;  
D.C. Ralph, C.T. Black, M. Tinkham, *Phys. Rev. Lett.* 74 (1995) 3241;  
C.T. Black, D.C. Ralph, M. Tinkham, *Phys. Rev. Lett.* 76 (1996) 688.
- [48] M. Gaudin, *La Fonction d'onde de Bethe*, Masson, Paris, 1983;  
M.C. Cambiaggio, A.M.F. Rivas, M. Saraceno, *Nucl. Phys. A* 424 (1997) 157.
- [49] J. von Delft, A.D. Zaikin, D.S. Golubev, W. Tichy, *Phys. Rev. Lett.* 77 (1996) 3189.
- [50] K.A. Matveev, A.I. Larkin, *Phys. Rev. Lett.* 78 (1997) 3749.
- [51] L. Amico, A. Osterloh, *Phys. Rev. Lett.* 88 (2002) 127003.
- [52] L. Amico, A. Di Lorenzo, A. Mastellone, A. Osterloh, R. Raimondi, *Ann. Phys.* 299 (2002) 228.
- [53] A. Faribault, P. Calabrese, J-S. Caux, *Phys. Rev. B* 77 (2008) 064503.
- [54] E.A. Yuzbashyan, A.A. Baytin, B.L. Altshuler, *Phys. Rev. B* 71 (2005) 094505; *Phys. Rev. B* 26 (2003) 214509.
- [55] N. Bogoliubov, *Sov. Phys.—JETP* 34 (1958) 73;  
J.G. Valatin, *Il Nuovo Cimento* 7 (1958) 883.
- [56] J. Bardeen, L.N. Cooper, J.R. Schrieffer, *Phys. Rev.* 108 (1957) 1175.
- [57] S.R. Elliott, M. Franz, *Rev. Modern Phys.* 87 (2015) 137.
- [58] P. Mai, B.S. Shastry, arXiv:1808.09788, *Phys. Rev. B*, 98, 2018, 205106. See especially Figures 9–12.
- [59] B.S. Shastry, E. Perepelitsky, arXiv:1605.08213, *Phys. Rev. B*, 94, 2016, 045138.

- [60] C. Gros, B. Edegger, V.N. Muthukumar, P.W. Anderson, P.N.A.S. 103 (2006) 14298.
- [61] R. Sensarma, M. Randeria, N. Trivedi, Phys. Rev. Lett. 98 (2007) 027004.
- [62] F.C. Zhang, C. Gros, T.M. Rice, H. Shiba, Superconductivity Sci. Tech. Vol. 1, 1988.
- [63] E. Kozik, M. Ferrero, A. Georges, Phys. Rev. Lett. 114 (2015) 156402.
- [64] I.J. and Pomeranchuk, Sov. Phys.—JETP 8 (1958) 361.
- [65] W. Metzner, D. Roher, S. Andergassen, Phys. Rev. Lett. 91 (2003) 066402.
- [66] K.A. Brueckner, J.L. Gammel, Phys. Rev. 109 (1958) 1023;  
J. Goldstone, Proc. Roy. Soc. (London) A 293 (1957) 267;  
M. Gell-Mann, F. Low, Phys. Rev. 84 (1951) 350.
- [67] W. Kohn, J.M. Luttinger, Phys. Rev. 118 (1960) 41.
- [68] G.S. Boebinger, Phys. Today 49 (1996) 36.

## Extremely correlated Fermi liquid of the $t$ - $J$ model in two dimensions

Peizhi Mai and B. Sriram Shastry

*Physics Department, University of California, Santa Cruz, California 95064, USA*



(Received 29 August 2018; published 5 November 2018)

We study the two-dimensional  $t$ - $J$  model with second-neighbor hopping parameter  $t'$  and in a broad range of doping  $\delta$  using a closed set of equations from the extremely correlated Fermi liquid theory. We obtain asymmetric energy distribution curves and symmetric momentum distribution curves of the spectral function, consistent with experimental data. We further explore the Fermi surface and local density of states for different parameter sets. Using the spectral function, we calculate the resistivity, Hall number, and spin susceptibility. The curvature change in the resistivity curves with varying  $\delta$  is presented and connected to intensity loss in angle-resolved photoemission spectroscopy experiments. We also discuss the role of the superexchange  $J$  in the spectral function and the resistivity in the optimal to overdoped density regimes.

DOI: [10.1103/PhysRevB.98.205106](https://doi.org/10.1103/PhysRevB.98.205106)

### I. INTRODUCTION

The  $t$ - $J$  model where extreme correlations are manifest plays a fundamentally important role in understanding the physics of correlated matter, including high- $T_c$  superconductors [1,2]. Despite the large progress [3–10] made in numerically solving the  $t$ - $J$  model and the related Hubbard model, very few analytical techniques are reliable to obtain the low-temperature physics in this model for a broad range of dopings due to its inherent difficulties including noncanonical algebra for Gutzwiller projected fermions and the lack of an obvious small parameter for perturbation expansion.

To tackle this challenge, we have recently developed the extremely correlated Fermi liquid (ECFL) theory [11,12]. It is a nonperturbative analytical theory employing Schwinger's functional differential equations of motion to deal with lattice fermions under extreme correlation  $U \rightarrow \infty$ . The ECFL theory uses a systematic expansion of a bounded parameter  $\lambda \in [0, 1]$ , analogous to the expansion parameter  $\frac{1}{2S}$  in the Dyson-Maleev representation of spins [13] via canonical bosons, and therefore provides a controlled calculation for the  $t$ - $J$  model. With recent advances in the theory [14], it is possible to represent the ECFL equations to any order in  $\lambda$  in terms of diagrams which are generalizations of the Feynman graphs, without having to consider previous orders.

The second-order  $O(\lambda^2)$  ECFL theory gives a closed set of equations for the Green's function and has been described in detail in Ref. [15]. It has been benchmarked successfully [16,17] against the exact results from the single-impurity Anderson model and the dynamical mean field theory (DMFT) [3,18–20], in the case of the infinite-dimensional large- $U$  Hubbard model. The benchmarking has also been carried out in the one-dimensional  $t$ - $J$  model, where  $k$ -dependent behavior is inevitable, against the density matrix renormalization group (DMRG) technique. ECFL and DMRG compare well [21] in describing the spin-charge separation in a Tomonaga-Luttinger liquid and the relevant strongly  $k$ -dependent self-energy.

Recently in Ref. [15], we have applied the second-order ECFL theory to studying the 2D  $t$ - $J$  model with a second-

neighbor hopping parameter  $t'$ . We calculated the spectral function peak, quasiparticle weight, and resistivity from hole doping ( $t' \leq 0$ ) to electron doping ( $t' > 0$ ). The high thermal sensitivity in the spectral function and small quasiparticle weight indicate a suppression of an effective Fermi liquid temperature scale. The curvature of resistivity vs  $T$  changes between concave and convex upon a sign change in  $t'$ , implying a change of the effective Fermi liquid temperature [17]. We also compute the optical conductivity and the nonresonant Raman susceptibilities in Ref. [22].

In the present work, we perform a more detailed study in the 2D  $t$ - $J$  model. Apart from the spectral function peak height, we compute the energy distribution curves (EDCs) and momentum distribution curves (MDCs) which are measured in angle-resolved photoemission spectroscopy (ARPES) [23]. For the first time from a microscopic theory, we obtain an asymmetric EDC line shape and a rather symmetric MDC line shape, which are consistent with experimental observation [23]. The self-energy is also calculated. It is independent of  $k$  in the infinite-dimensional limit [16] and has strong  $k$ -dependence in 1D [21]. In 2D our calculation gives a weakly  $k$ -dependent self-energy in the normal (metallic) state. For this reason, we expect the vertex correction to be modest. Then we compute the resistivity within the bubble scheme neglecting the vertex corrections. Unlike Ref. [15], here we focus on the doping dependence of resistivity vs  $T$  curves at different  $t'$ , corresponding to experimental observation [24]. Spin susceptibility and the NMR spin-lattice relaxation rate are also calculated with the ECFL Green's function and related to experiment [25,26]. At the end, we discuss the effect of the superexchange interaction and justify our choice of  $J$ .

This work is organized as follows: First we summarize the ECFL formalism to calculate the electron Green's function and introduce the parameter region in Sec. II. In Sec. III, we discuss the ECFL spectral properties, resistivity, Hall response, and spin susceptibility at a fixed typical superexchange  $J$ , as well as the effect of changing  $J$ . Section IV includes a conclusion and some remarks.

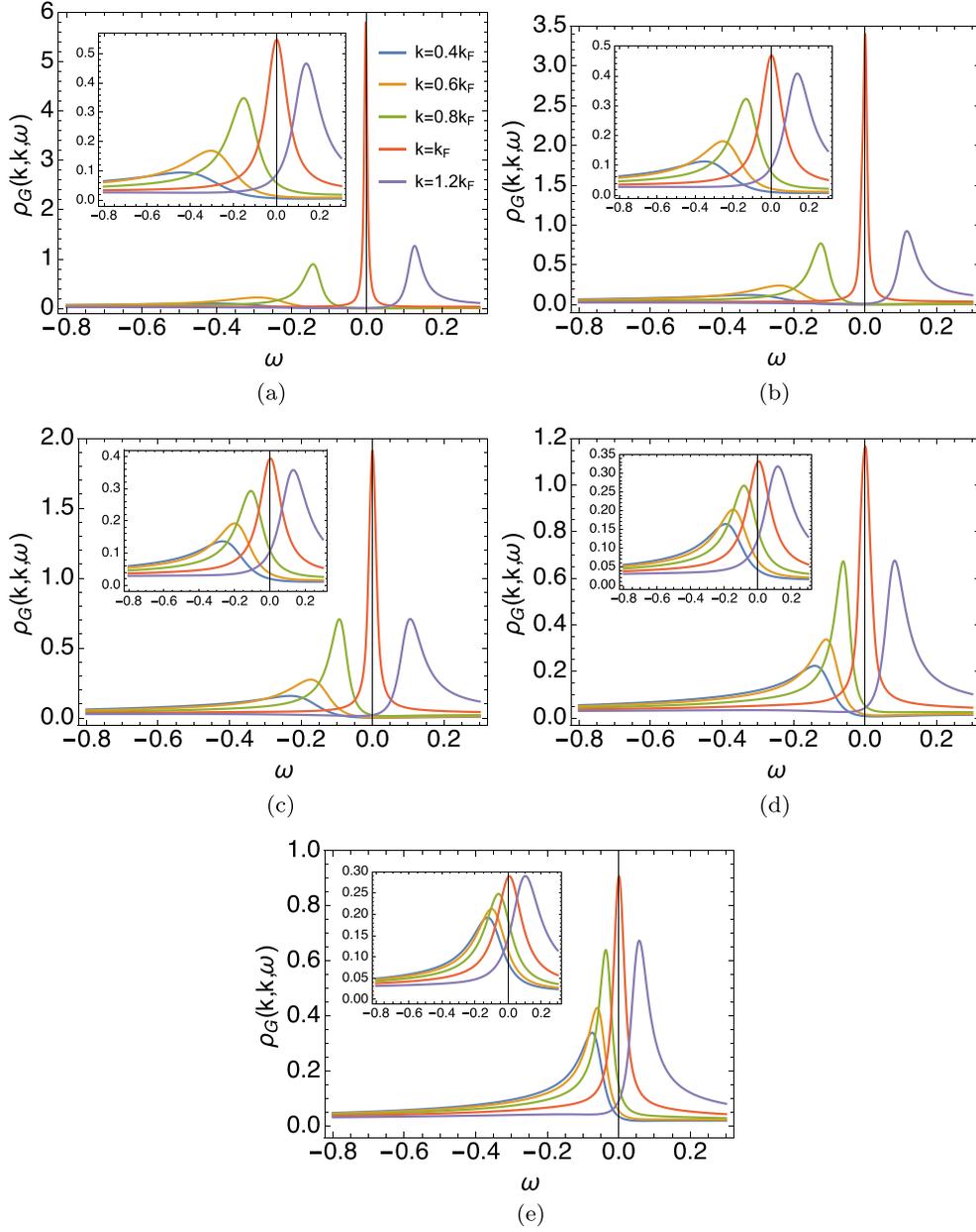


FIG. 1. EDC line shapes at different fixed values of momentum  $k$  in nodal direction ( $\Gamma \rightarrow X$ ). All figures including insets share the same legend. The parameters are set as  $\delta = 0.15$ ,  $T = 105$  K or 400 K (inset) and  $t'$  as specified. The line peak and width in the vicinity of the Fermi surface depends strongly on temperature. The peak magnitude at  $\omega = 0$  goes down as  $t'$  decreases due to stronger correlation. (a)  $t' = 0.4$ . (b)  $t' = 0.2$ . (c)  $t' = 0$ . (d)  $t' = -0.2$ . (e)  $t' = -0.4$ .

## II. METHOD AND PARAMETERS

### A. Summary of second-order ECFL theory

In the ECFL theory [11] the one-electron Green's function in momentum space is expressed as the product of an auxiliary Green's function  $\mathbf{g}$  and a caparison function  $\tilde{\mu}$ :

$$\mathcal{G}(k) = \mathbf{g}(k) \times \tilde{\mu}(k), \quad (1)$$

where  $k \equiv (\vec{k}, i\omega_n)$  and  $\omega_n = (2n + 1)\pi k_B T$  is the Matsubara frequency. Here  $\mathbf{g}(k)$  is a canonical fermion propagator vanishing as  $1/\omega$  as  $\omega \rightarrow \infty$ , and  $\tilde{\mu}(k)$  plays a role of adaptive spectral weight due to the noncanonical nature of the problem.

In the minimal version of second-order theory [16] including superexchange  $J$ , they can be written explicitly as

$$\tilde{\mu}(k) = 1 - \lambda \frac{n}{2} + \lambda \Psi(k), \quad (2)$$

$$\mathbf{g}^{-1}(k) = i\omega_n + \mu - \frac{u_0}{2} + \frac{\lambda}{4} n J_0 - \tilde{\mu}(k) \varepsilon'_k - \lambda \chi(k), \quad (3)$$

where  $\mu$  is the chemical potential, and  $\varepsilon'_k = \varepsilon_k - \frac{u_0}{2}$ . Here  $u_0$  is a Lagrange multiplier [27] guaranteeing the shift invariance of the  $t$ - $J$  model at every order of  $\lambda$ . To elaborate,  $u_0$  absorbs any arbitrary uniform shift of the band  $\varepsilon_k \rightarrow \varepsilon_k + c$ , a constant shift which should not change

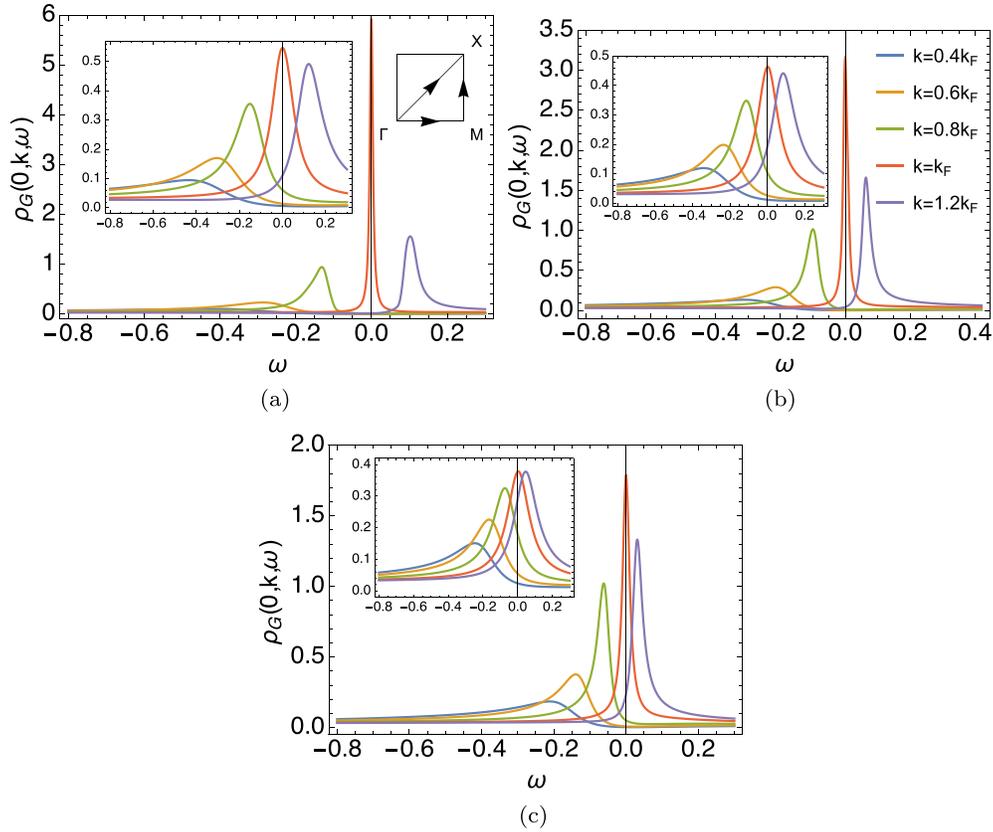


FIG. 2. EDC line shapes at different fixed values of momentum  $k$  in antinodal direction ( $\Gamma \rightarrow M$ ). All figures including insets share the same legend. The parameters are set as  $\delta = 0.15$ ,  $T = 105$  K or 400 K (inset), and  $t'$  as specified. The line peak and width in the vicinity of the Fermi surface depend strongly on temperature. The peak magnitude at  $\omega = 0$  goes down as  $t'$  decreases due to stronger correlation. (a)  $t' = 0.4$ . (b)  $t' = 0.2$ . (c)  $t' = 0$ .

the results. The band dispersion including next-nearest-neighbor hopping is  $\varepsilon_k = -2t[\cos(k_x a_0) + \cos(k_y a_0)] - 4t' \cos(k_x a_0) \cos(k_y a_0)$ , and  $\Psi$  and  $\chi$  are two self-energy parts. These are given by [16]

$$\Psi(k) = - \sum_{pq} (\varepsilon'_p + \varepsilon'_q + J_{k-p}) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k), \quad (4)$$

$\chi = \chi_0 + \lambda \chi_1$  with  $\chi_0 = - \sum_p \mathbf{g}(p) (\varepsilon'_p + \frac{1}{2} J_{k-p})$ , and

$$\chi_1(k) = - \sum_{pq} \left[ \varepsilon'_p + \varepsilon'_q + \frac{1}{2} (J_{k-p} + J_{k-q}) \right] \times (\varepsilon'_{p+q-k} + J_{k-q}) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k), \quad (5)$$

where  $\sum_k \equiv \frac{k_B T}{N_s} \sum_{k_x, k_y, \omega_n}$ ,  $N_s$  is the number of sites, and  $J_k = 2J(\cos k_x a_0 + \cos k_y a_0)$  is the nearest-neighbor exchange.

Denoting the particle and hole density per site by  $n$  and  $\delta = 1 - n$ , respectively, the two chemical potentials  $\mu$  and  $u_0$  are determined through the number sum rules

$$\sum_k \mathbf{g}(k) e^{i\omega_n 0^+} = \frac{n}{2} = \sum_k \mathcal{G}(k) e^{i\omega_n 0^+}. \quad (6)$$

After analytically continuing  $i\omega_n \rightarrow \omega + i0^+$  we determine the interacting electron spectral function  $\rho_{\mathcal{G}}(\vec{k}, \omega) = -\frac{1}{\pi} \text{Im} \mathcal{G}(\vec{k}, \omega)$ . The set of equations (1)–(6) was solved iteratively on  $L \times L$  lattices with  $L = 61, 131, 181$  and a

frequency grid with  $N_\omega = 2^{14}$  points.  $L > 61$  is usually for  $t' \geq 0$  at low temperatures where the spectral function peak is higher and sharper than the negative  $t'$  cases; therefore it requires better  $k$  resolution.

## B. Parameters in the programs

In this calculation, we set  $t = 1$  as the energy scale and  $t'$  is varied between  $-0.4$  and  $0.4$ . We fix the superexchange to  $J = 0.17$  unless otherwise specified because  $J$  usually is estimated to be in the region from 0 to 0.4, and has a small effect on the  $k$ -dependent behavior and barely influences the averaged physical quantities like resistivity, since the calculation includes a summation in  $k$  space. This argument will be further justified in the last part of Sec. III. Besides, we also explore a large region of doping  $\delta$  from 0.11 to 0.3, where the second-order ECFL theory is reliable [16], and present the  $\delta$ -dependent behavior at different  $t'$ . If not specified,  $\omega$  is in units of  $t$ . According to Ref. [2], we assume  $t = 0.45$  eV when using the absolute temperature scale.

## C. The sign of $t'$

The significance of the sign of  $t'$  should be kept in mind, and the case  $t' > 0$  is believed to correspond to electron-doped cuprate superconductors whereas  $t' < 0$  is the hole-doped cuprates. The hole-doped case appears highly

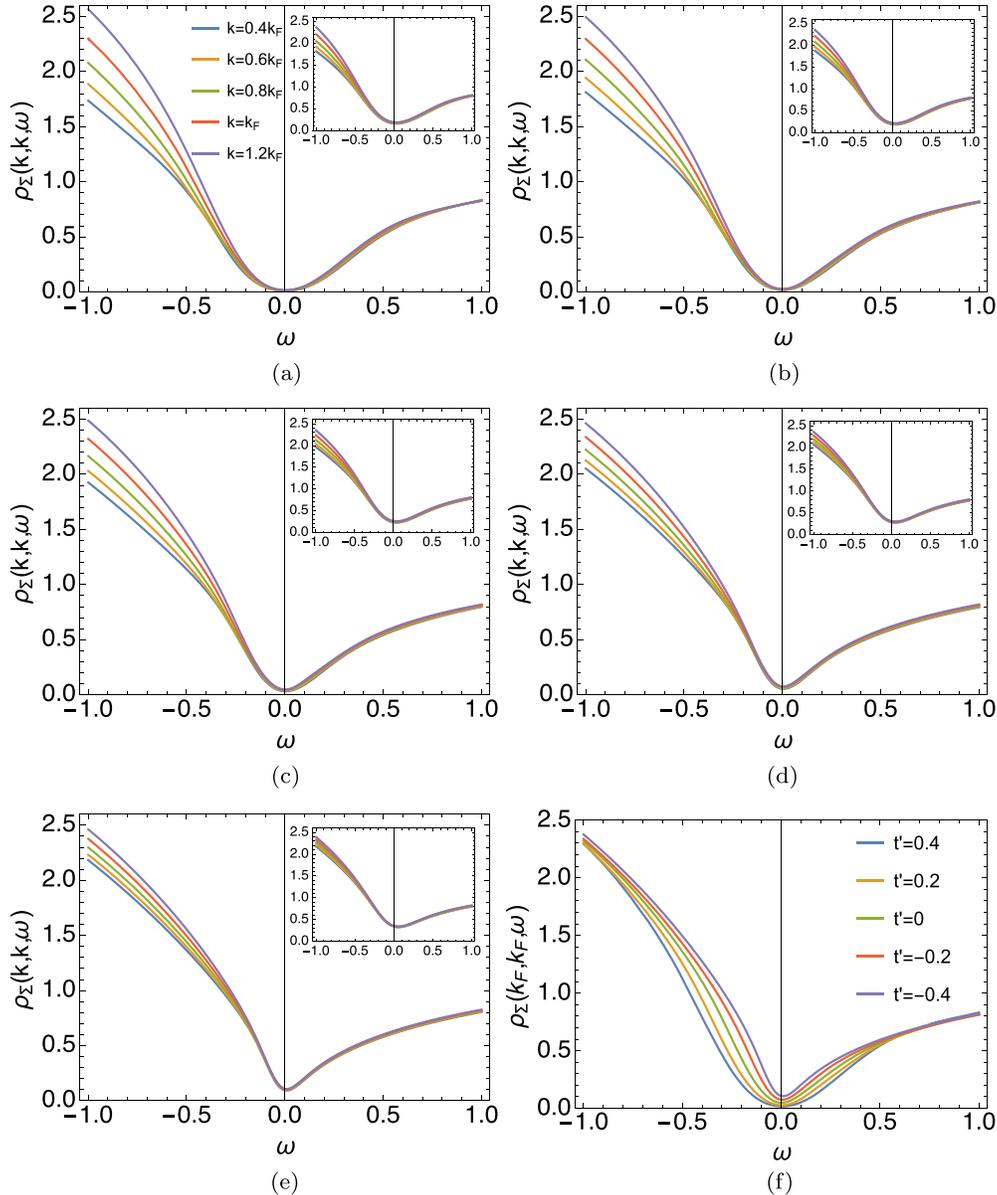


FIG. 3. (a)–(e) The negative imaginary part of self-energy  $\rho_{\Sigma}$  at different  $k$  in nodal ( $\Gamma \rightarrow X$ ) direction with several  $t'$ . Here  $\delta = 0.15$ ,  $T = 105$  K and 400 K (inset). In all cases,  $\rho_{\Sigma}$  has a weak  $k$  dependence and differs mostly at high energy on the unoccupied side. Increasing temperature raises the bottom of the self-energy while leaving its high-energy part almost unchanged. (f)  $\rho_{\Sigma}$  at fixed  $k = k_F$  in nodal direction varying  $t'$ . Increasing  $t'$  lowers the bottom of  $\rho_{\Sigma}$  and makes its low-energy part more rounded (Fermi-liquid-like). (a)  $t' = 0.4$ . (b)  $t' = 0.2$ . (c)  $t' = 0$ . (d)  $t' = -0.2$ . (e)  $t' = -0.4$ . (f) Varying  $t'$ .

non-Fermi-liquid-like as compared to the electron-doped case in experiments, and our earlier calculations as well as the present ones give a microscopic understanding of this important basic fact. We emphasize that despite this, the  $t' > 0$  case is also strongly correlated, when we view the  $T$  dependence of the spectral features, where the effective Fermi scale is much reduced from the bare (band structure) value.

### III. RESULTS

#### A. Spectral properties

##### 1. Spectral function and self-energy

In earlier studies [23], the ECFL spectral function obtained phenomenologically [11,23,28] has been compared with

experimental data measured with angle-resolved photoemission spectroscopy (ARPES) at optimal doping, leading to very good fits. Later we calculated the spectral function from the raw second-order ECFL equations in the symmetrized model [29] but it is only valid for doping  $\delta \gtrsim 0.25$ . Here we present the result at optimal doping  $\delta = 0.15$  from a microscopic calculation of ECFL by numerically solving the improved set of second-order equations [15,16].

We display the energy distribution curves (EDCs) in Fig. 1 and Fig. 2, obtained by fixing  $k$  and scanning  $\omega$  at optimal doping and various  $t'$ . These quantities can be measured in ARPES experiments. Figure 1 shows the EDCs for several constant  $k$  along the nodal ( $\Gamma \rightarrow X$ ) and Fig. 2 for the antinodal direction ( $\Gamma \rightarrow M$  for  $t' > 0$ ). Note that the value of

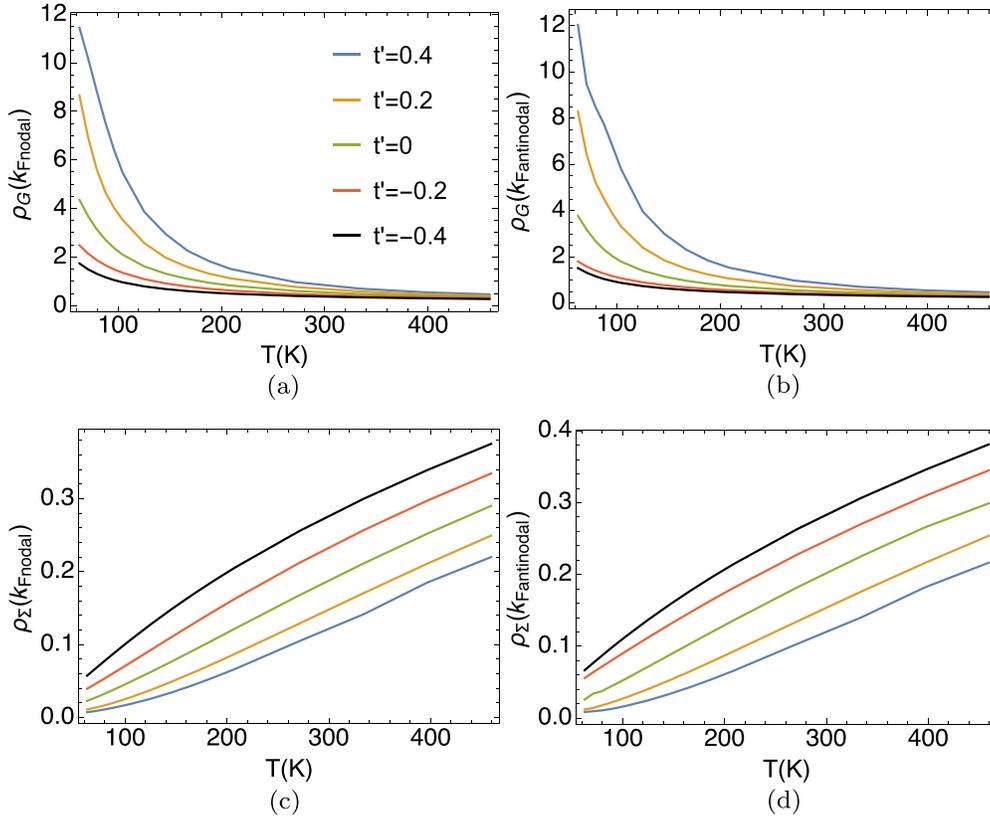


FIG. 4. The spectral functions at  $\omega = 0$ :  $\rho_G$  and  $\rho_\Sigma$  at  $k_F$  (in nodal and antinodal directions) vs  $T$  with varying  $t'$  at  $\delta = 0.15$ ; legend is the same for each figure. (a)  $\rho_G(k_F \text{ nodal})$ . (b)  $\rho_G(k_F \text{ antinodal})$ . (c)  $\rho_\Sigma(k_F \text{ nodal})$ . (d)  $\rho_\Sigma(k_F \text{ antinodal})$ .

$k_F$  depends on  $t'$  and direction in  $k$  space. The fixed value of  $k$  is given in terms of  $k_F$  based on the specific  $t'$  and direction. The antinodal ( $M \rightarrow X$ )  $k_F$  for  $t' \leq -0.2$  is close to zero. The corresponding EDCs are too close to resolve clearly; hence they are not presented.

We observe that at low temperatures the EDC peak gets sharper as  $k$  approaches the Fermi surface. The insets show that a small heating ( $\Delta T \sim 0.06t$ ) strongly suppresses the region around the Fermi surface  $k \sim k_F$  while it leaves the region away from Fermi surface almost unchanged. As a result, a weaker  $k$  dependence of peak height can be viewed in the higher temperature. It also shows that the EDC line shape is asymmetric for  $k < k_F$ , consistent with ARPES experiments. As  $t'$  decreases from positive (electron doped) to negative (hole doped), the correlation becomes stronger, and therefore the spectral peak gets lower. Slight anisotropy is found for  $t' \leq 0.2$  in that the peak at the Fermi surface is a bit higher in the nodal direction than in the antinodal direction, indicating a weak  $k$  dependence of self-energy.

The spectral function of the Dyson self-energy is defined as

$$\rho_\Sigma(\vec{k}, \omega) = -\frac{1}{\pi} \text{Im} \Sigma(\vec{k}, \omega). \quad (7)$$

It is calculated from the spectral function obtained from solving the set of ECFL equations (1)–(6):

$$\rho_\Sigma(\vec{k}, \omega) = \frac{\rho_G(\vec{k}, \omega)}{\pi^2 \rho_G^2(\vec{k}, \omega) + [\text{Re}G(\vec{k}, \omega)]^2}, \quad (8)$$

where  $\text{Re}G$  is calculated through Hilbert transform of  $\rho_G$ . As observed in Figs. 3(a)–3(e), the self-energy shows asymmetry from intermediate frequencies at essentially all values of  $t'$  and  $k$ , which is consistent with previous studies [16,29], unlike the symmetric curves in standard Fermi liquid theory. Further they all appear to depend weakly on  $k$ . This is qualitatively different from the strong  $k$  dependence of the low-energy behaviors of the self-energy in one dimension [21]. This weak  $k$  dependence supports our approximation of resistivity formula ignoring vertex correction in the next section. The inset indicates that the heating makes the most difference in the low-energy region by lifting the bottom. In Fig. 3(f),  $\rho_\Sigma$  at  $k_F$  for different  $t'$  are put together. As  $t'$  increase from negative to positive, its minimum goes down, indicating a lower decay rate, and the bottom region becomes rounded and more Fermi-liquid-like.

We also study the temperature-dependent  $\rho_G(k_F)$  and  $\rho_\Sigma(k_F)$  at  $\omega = 0$  for  $k_F$  in the nodal and antinodal direction in Fig. 4. Also, panels (a) and (b) show that the spectral function peak is very sensitive to temperature changes. A sharp drop happens over a small temperature region ( $<1\%$  bare bandwidth), wiping out the quasiparticle peak for  $T > 400$  K in either direction. Another angle to observe this phenomenon is through the self-energy,  $\rho_\Sigma(k_F) = 1/[\pi^2 \rho_G(k_F)]$ , describing the decay rate of a quasiparticle. The huge increase of  $\rho_\Sigma(k_F)$  upon small warming shows a rapid drop in the lifetime of a quasiparticle. Note that the  $\rho_\Sigma$  curvature dependence on  $t'$  is similar to that of the plane resistivity in Fig. 4 of Ref. [15].

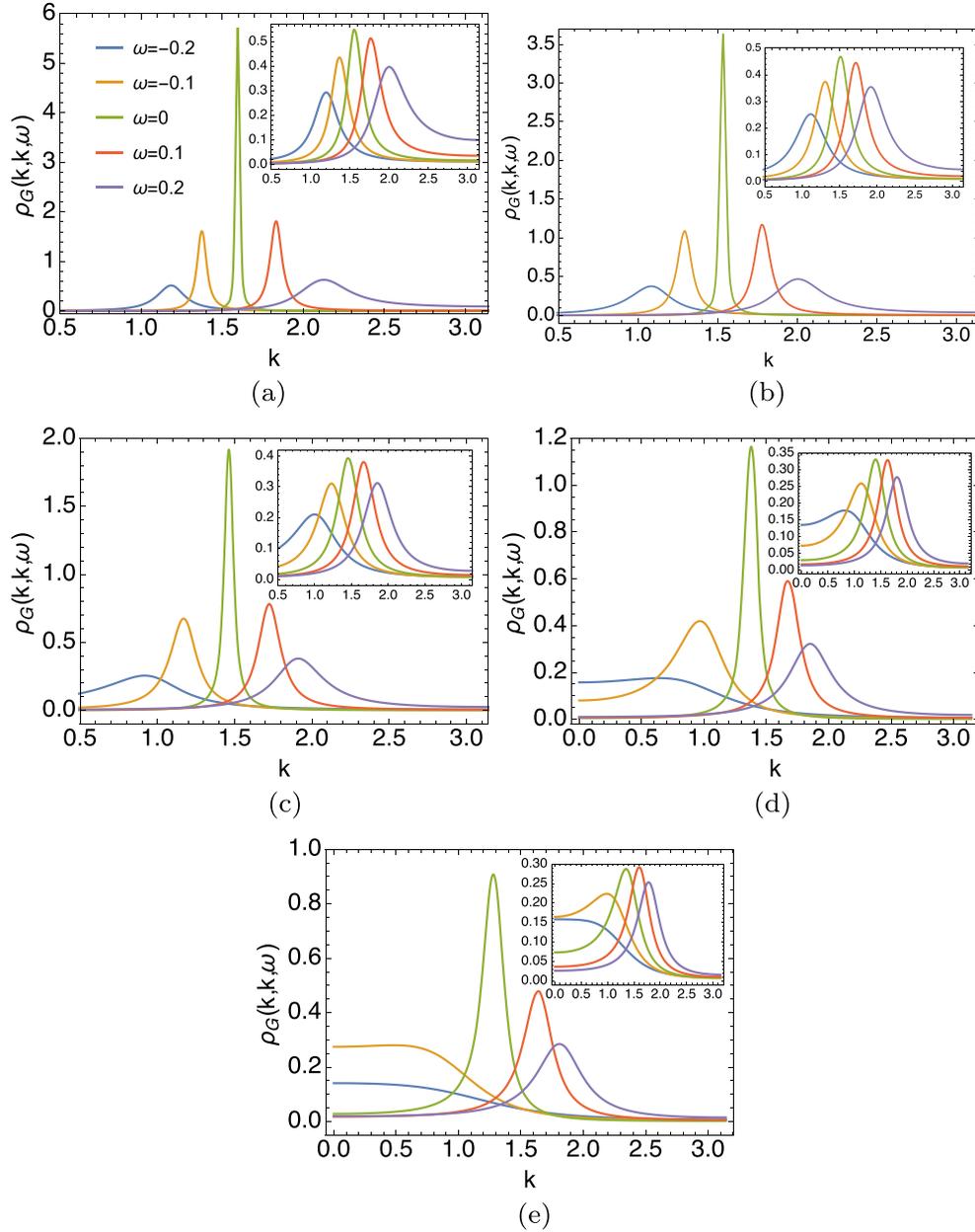


FIG. 5. MDC line shapes at different fixed values of frequency  $\omega$  in each curve. All figures including insets share the legend. Here the parameters are set as  $\delta = 0.15$ ,  $T = 105$  K and 400 K (inset).  $k$  is scanned along the nodal ( $\Gamma \rightarrow X$ ) direction. In all cases, they have a highest peak with a symmetric shape at  $\omega = 0$ . Consistently, the peak height decreases with smaller  $t'$ , or stronger correlation. (a)  $t' = 0.4$ , nodal ( $\Gamma \rightarrow X$ ). (b)  $t' = 0.2$ , nodal ( $\Gamma \rightarrow X$ ). (c)  $t' = 0$ , nodal ( $\Gamma \rightarrow X$ ). (d)  $t' = -0.2$ , nodal ( $\Gamma \rightarrow X$ ). (e)  $t' = -0.4$ , nodal ( $\Gamma \rightarrow X$ ).

The momentum distribution curves (MDCs) are plotted in Fig. 5 and Fig. 6, obtained by fixing  $\omega$  and scanning  $k$  in nodal and antinodal directions, respectively, at optimal doping and various  $t'$ . As expected from the EDC case, the MDC peak is highest at the Fermi surface  $\omega = 0$ , which gets broadened the most upon warming. However, unlike the EDC case, the MDC peaks that are far away from  $k = 0$  or  $\pi$  look more symmetric. This difference is consistent with experimental findings. The spectral function in the early phenomenological versions of ECFL, Refs. [23,28], leads to a somewhat exaggerated asymmetry in MDC curves, and has been the subject of further phenomenological adjustments in Ref. [30], to reconcile with experiments. The present microscopic results

show that the greater symmetry of the MDC spectral lines comes about naturally, without the need for any adjustment of the parameters.

## 2. Fermi surface

The Fermi surface (FS) structure can be observed in the momentum distribution of spectral function peak height. We present the case for  $t' = -0.2$ , which is roughly the parameter describing the LSCO cuprate material [31], and vary the doping  $\delta$  in Fig. 7. The FS is hole-like (open) for low doping [panels (a) and (b)] and becomes electron-like (closed) for high doping in panels (d) and (e). The transition

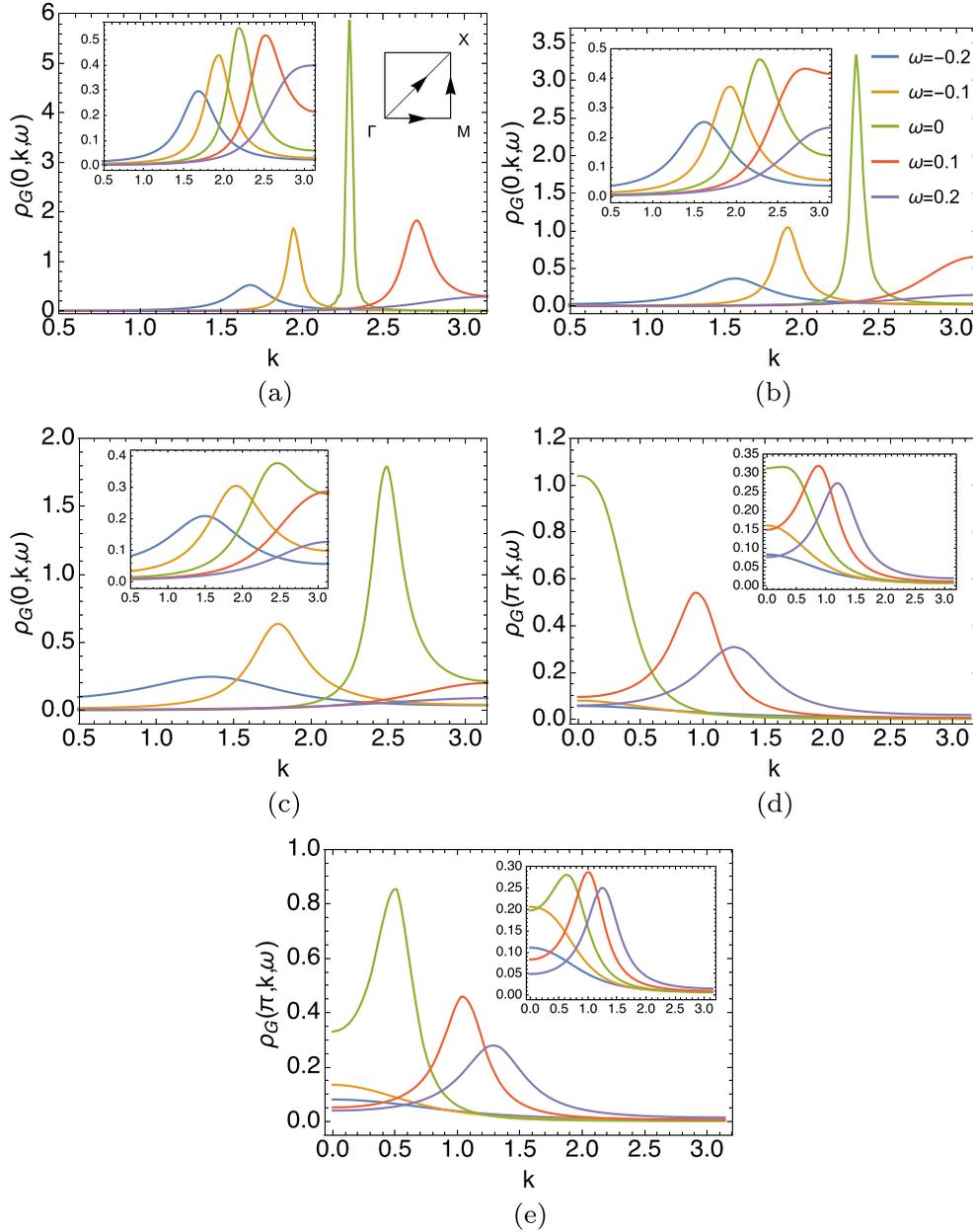


FIG. 6. MDC line shapes at different fixed values of frequency  $\omega$  in each curve. All figures including insets share the same legend. Here the parameters are set as  $\delta = 0.15$ ,  $T = 105$  K and 400 K (inset).  $k$  is scanned along the antinodal ( $\Gamma \rightarrow M$  for  $t' \geq 0$  or  $M \rightarrow X$  for  $t' < 0$ ) directions. (a)  $t' = 0.4$ , antinodal ( $\Gamma \rightarrow M$ ). (b)  $t' = 0.2$ , antinodal ( $\Gamma \rightarrow M$ ). (c)  $t' = 0$ , antinodal ( $\Gamma \rightarrow M$ ). (d)  $t' = -0.2$ , antinodal ( $M \rightarrow X$ ). (e)  $t' = -0.4$ , antinodal ( $M \rightarrow X$ ).

point  $\delta \approx 0.17$  can be explicitly seen in Fig. 8(a) which is close to the noninteracting case with the tight-binding model in Fig. 8(e), consistent with experimental findings [31–33]. At higher (hole) doping which leads to a weaker effective correlation [15], the quasiparticle peak height increases and becomes more Fermi-liquid-like.

The FS is only well defined at zero temperature. Following Ref. [34] we can define a pseudo-FS at finite temperature, by examining a specifically weighted first moment of the energy:

$$\gamma_{k\sigma}(\mu, T) = - \int \rho_G(k, \omega) \times \frac{d\omega \omega}{\cosh(\beta\omega/2)} \Big/ \int \rho_G(k, \omega) \frac{d\omega}{\cosh(\beta\omega/2)}. \quad (9)$$

We define a pseudo-FS as the surface in  $\vec{k}$  space where  $\gamma_{k\sigma}$  changes sign from positive to negative. Shastry has recently shown [34] that at  $T = 0$ , the pseudo-FS becomes the exact Luttinger-Ward FS. It is further suggested that it is useful to study a  $T$ -dependent effective carrier density

$$N_{\text{eff}} = \sum_{k\sigma} \Theta(\gamma_{k\sigma}(\mu, T)), \quad (10)$$

where  $\Theta$  is the Heaviside step function, such that  $N_{\text{eff}} = N$  at zero temperature. At finite temperatures we expect that  $N_{\text{eff}} \neq N$ , and the difference between the two gives insights into the different  $T$  scales at play. This is especially applicable in strongly correlated materials, where it is well known [17–19]

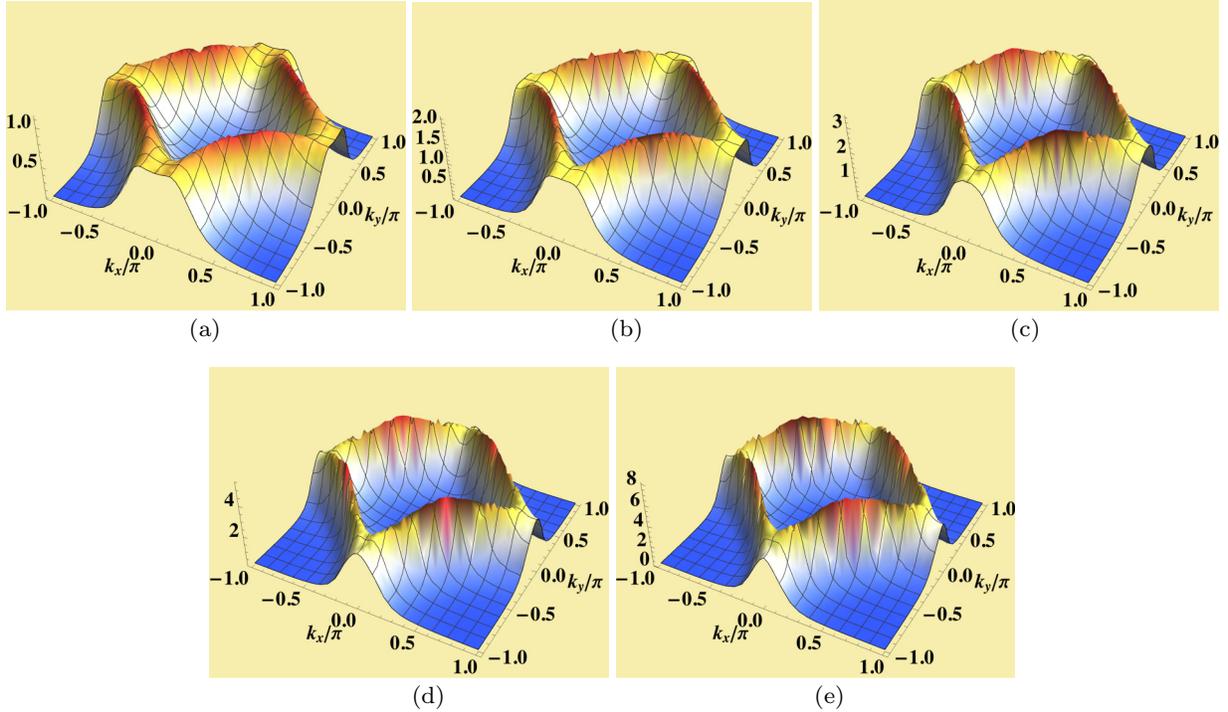


FIG. 7. The 3D plot of the spectral function peak height at several dopings at  $t' = -0.2$ ,  $T = 63$  K. The ridge in the spectral function peak tracks the Fermi surface. As  $\delta$  increases, we find that the Fermi surface changes from open (hole-like) to close (electron-like), with the critical  $\delta \approx 0.17$ . The ridge height increases generally as  $\delta$  goes up, showing decreasing correlation strength. (a)  $\delta = 0.11$ . (b)  $\delta = 0.14$ . (c)  $\delta = 0.17$ . (d)  $\delta = 0.2$ . (e)  $\delta = 0.23$ .

that Gutzwiller correlations result in the Fermi liquid regime, the strange-metal regime, and the bad-metal regime, followed by a high- $T$  regime, with three crossover temperatures. In Fig. 9, we show how  $N_{\text{eff}}/N$  changes with temperature for different  $t'$ . For  $t' \geq 0$ ,  $N_{\text{eff}}$  increases monotonically toward  $N$  as  $T$  goes down. And for  $t' < 0$ ,  $N_{\text{eff}}$  decreases from larger to smaller than  $N$  upon cooling. With further lowering  $T$  one expects that  $N_{\text{eff}}$  equals  $N$ .

At low temperatures ( $T \ll t$ ), we find that the roots of  $\gamma_k$  are close to the location of the ridge of spectral peak height shown in Fig. 11, and hence it can be taken as an approximate or a pseudo-finite-temperature FS. Figure 10 shows that the pseudo-FS is getting close to the true FS at zero temperature as  $T$  goes down for both electron-doped and hole-doped systems.

To understand better the deviations at finite  $T$  seen in Fig. 9, Fig. 10, and Fig. 11, it is helpful to recall a phenomenological spectral function [48] (see Eq. (9) and Eqs. (SI-20) and (SI-21) in Ref. [48]). This function is obtained by expanding the two self-energies in Eq. (2) and Eq. (3) at low energies in a power series. It captures many features of the ECFL calculations in terms of a few parameters, and is given as

$$A(\hat{k}, \omega) = \frac{z_0}{\pi} \frac{\Gamma(\omega)}{\Gamma(\omega)^2 + (\omega - V_L \hat{k})^2} \left( 1 - \frac{\xi}{\sqrt{1 + c_\alpha \xi^2}} \right), \quad (11)$$

where  $\hat{k}$  is the component of  $\vec{k}$  normal to the FS;  $\xi = \frac{1}{\Delta_0}(\omega - r V_L \hat{k})$ ;  $\Gamma(\omega) = \eta + \frac{\pi}{\Omega_\Phi}(\omega^2 + \pi^2 k_B^2 T^2)$ ;  $\Delta_0$  and  $\Omega_\Phi$  are the low- and high-energy scales;  $V_L$  is the Fermi velocity; and  $z_0$ ,  $r$ , and  $c_\alpha$  are numerical constants. The important variable  $r \in [0, 2]$  determines the location of a feature in the dispersion

known as the ‘‘kink.’’ It is analyzed using this model spectral function in Ref. [48]. Here  $r = 1$  is at the border of two regimes  $r < 1$  with kinks in the unoccupied side, and  $r > 1$  with kinks in the occupied side of the distribution. In Fig. 12 we plot the location of the peak in the spectral function Eq. (11) against  $T$ , for three values  $r = 0.5, 1$ , and  $1.5$ . From this we see that these regimes display either a shrinking or an enlargement of the FS with increasing  $T$ . This corresponds to the types of behavior seen in Fig. 10 and Fig. 11.

### 3. Local density of states

The local density of states (LDOS) is calculated by  $\sum_{\vec{k}} (1/N_s) \rho_G(\vec{k}, \omega)$  and plotted in Figs. 13 and 14, varying  $t'$  with fixed  $\delta = 0.15$  and varying  $\delta$  with fixed  $t' = 0, -0.4$ , respectively. This quantity can be measured by scanning tunneling microscopy [35–39].

In Fig. 13, comparing panels (a) and (c), we observe that the LDOS peak gets smoothed and also broadened by the electron-electron interaction. Although the relative position for different  $t'$  remains unchanged after turning on interaction, the strong correlation brings them closer by renormalizing the bare band into the effective one, as shown in the inset of Fig. 22. From panel (a) to (b), raising temperature tends to have a stronger suppression on the peak with lower  $t'$ . This means the system with higher  $t'$  has a higher Fermi liquid temperature scale, and therefore it is more robust to heating, which is consistent with the previous findings of the spectral function.

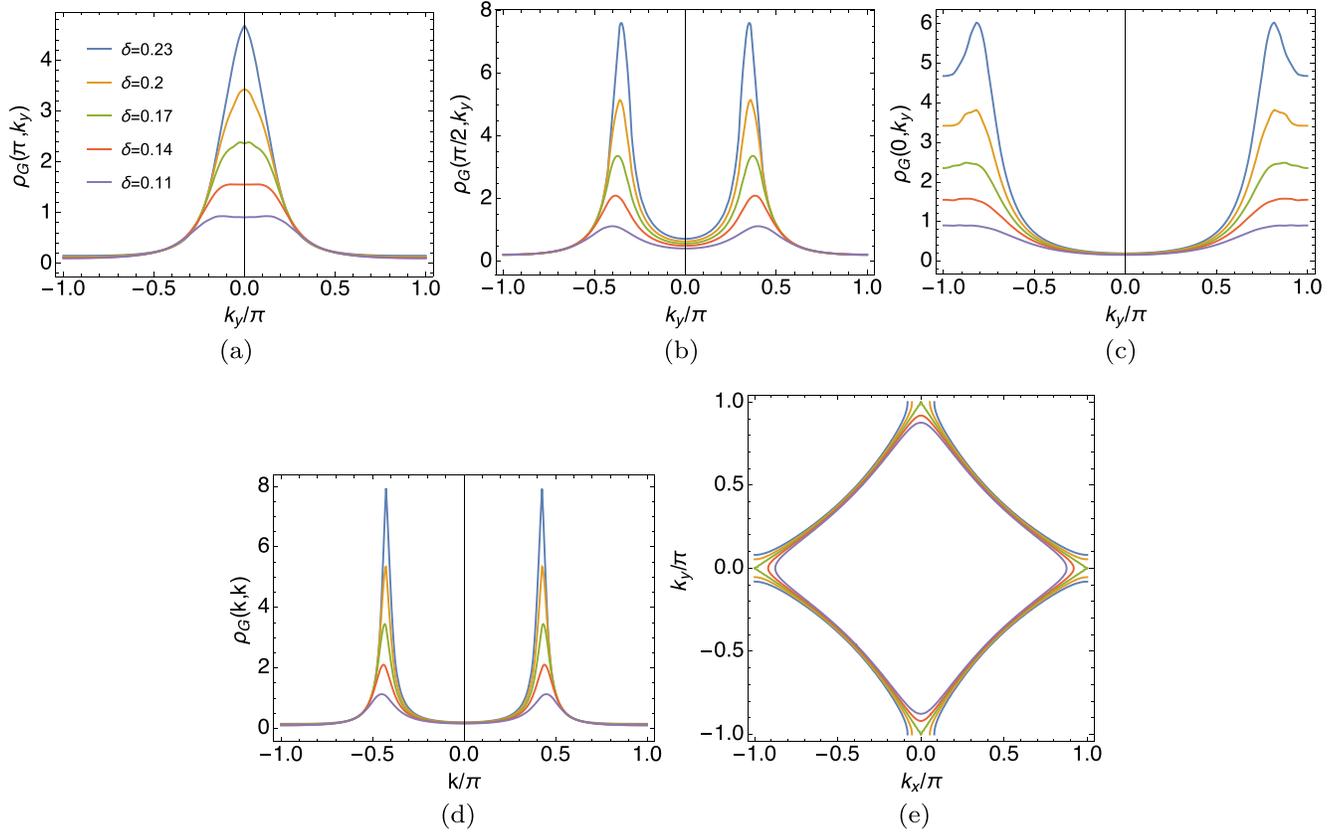


FIG. 8. The spectral function peak height in typical directions of momentum space at several dopings at  $t' = -0.2$  and  $T = 63$  K. All panels share the same legend. Panel (a) shows evidence of Lifshitz transition (Fermi surface changed from opened to closed) at  $\delta \approx 0.17$ , similar to the tight-binding model case shown in panel (e). Panels (b), (c), (d) provide other angles to observe this transition, in complementarity with the 3D plots in Fig. 7. (a)  $\rho_G(\pi, k_y)$ ,  $M \rightarrow X$ . (b)  $\rho_G(\pi/2, k_y)$ . (c)  $\rho_G(0, k_y)$ ,  $\Gamma \rightarrow M$ . (d)  $\rho_G(k, k)$ , nodal ( $\Gamma \rightarrow X$ ). (e) Tight-binding model.

In Fig. 14, from the electron-like panels [(a), (c), (e)] to the strongly hole-like panels [(b), (d), (f)], the LDOS peak shifts from  $\omega > 0$  to  $\omega < 0$ . In contrast to the noninteracting tight-binding model in (e) and (f) where the peak height is independent of doping, (a)–(d) have smaller peaks in general and show that the height decreases at smaller doping with more weight in the lower Hubbard band (insets). This is again a feature of strong correlation. As the system approaches the half-filling limit ( $\delta \rightarrow 0$ ), the correlation enhances and further suppresses the quasiparticle peak, which contributes to the central peak of the LDOS. We also observe that (a) is similar to the density dependence of the location of Kondo or Abrikosov-Suhl resonance in the Anderson impurity problem [16]. It can be understood as a generic characteristic in strongly correlated matter given the relation between density and the effective interaction.

### B. Resistivity

We next present the resistivity under strong electron-electron interaction. The popular bubble approximation is used and the current correlator is written as  $\langle J(t)J(0) \rangle \sim \sum_k v_k^2 \mathcal{G}^2(k)$ . Here the velocity  $\hbar v_k^\alpha = \frac{\partial \epsilon_k}{\partial k_\alpha}$  represents the bare current vertex. In tight-binding theory the sign

oscillation in  $v_k^\alpha$  leads to a reduction in the average over the Brillouin zone and therefore diminishes the magnitude of the vertex corrections. Also the weak  $k$  dependence of self-energy in Fig. 3 reduces the importance of vertex corrections.

In our picture of a quasi-two-dimensional metal, there are 2D well-separated sheets, by a distance  $c_0$  in the  $c$

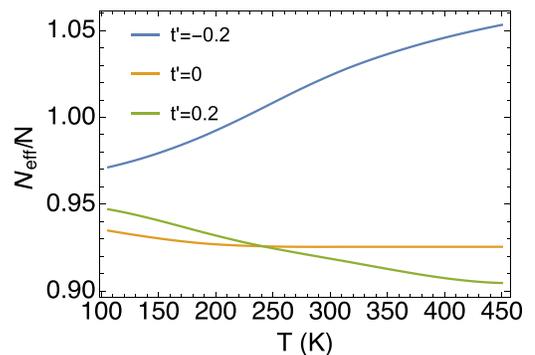


FIG. 9.  $N_{\text{eff}}/N$  vs  $T$  at  $\delta = 0.15$  and various  $t'$ . For electron-doped ( $t' \geq 0$ ) case,  $N_{\text{eff}}$  increases as one lowers the temperature, while in the hole-doped ( $t' = -0.2$ ) case,  $N_{\text{eff}}$  decreases upon cooling down. At lower temperature, one expects that  $N_{\text{eff}}$  equals  $N$ .

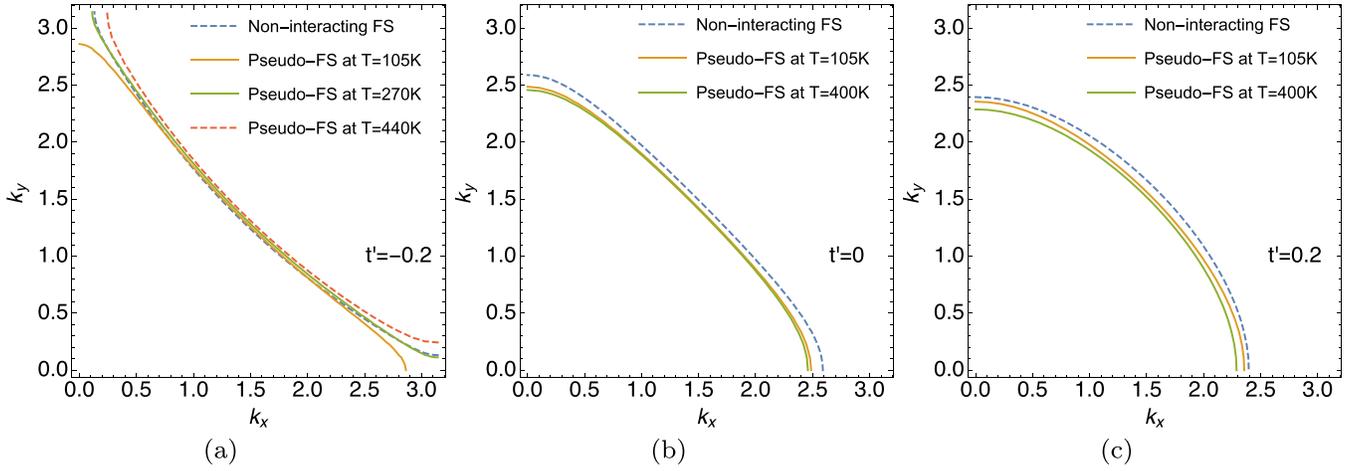


FIG. 10. Comparison between the noninteracting FS and pseudo-FS at low and high temperature. Here we fix  $\delta = 0.15$  and vary  $t'$ . Generally, as we cool down the system, the pseudo-FS approaches the noninteracting system or FS from the right ( $t' = -0.2$ ) or left ( $t' = 0, 0.2$ ) side. The exception is that at  $T = 105$  K and  $t' = -0.2$  the pseudo-FS turns out to be closed (electron-like). This delicate effect is a consequence of the redistribution of weight in the spectral function, and its thermal sensitivity is presumably related to the nearby Lifshitz transition point for the choice of  $t' = -0.2$ . We cannot access very low  $T$  for our system sizes, but it is expected that the pseudo-FS flips back to being hole-like at a low  $T$ .

direction. Thus each sheet can be effectively characterized by the 2D  $t$ - $J$  model. Its dc resistivity  $\rho_{xx}$  can be written as follows:

$$\rho_{xx} = \rho_0 \bar{\rho}_{xx} = \frac{\rho_0}{\bar{\sigma}_{xx}}, \quad (12)$$

$$\bar{\sigma}_{xx} = (2\pi)^2 \int_{-\infty}^{\infty} d\omega \left( -\frac{\partial f}{\partial \omega} \right) \left\langle \rho_G^2(\vec{k}, \omega) \frac{(\hbar v_k^x)^2}{a_0^2} \right\rangle_k, \quad (13)$$

where  $\bar{\rho}_{xx}$  and  $\bar{\sigma}_{xx}$  represent dimensionless resistivity and conductivity, respectively;  $\rho_0 \equiv c_0 h / e^2$  ( $\sim 1.718$  m $\Omega$  cm)

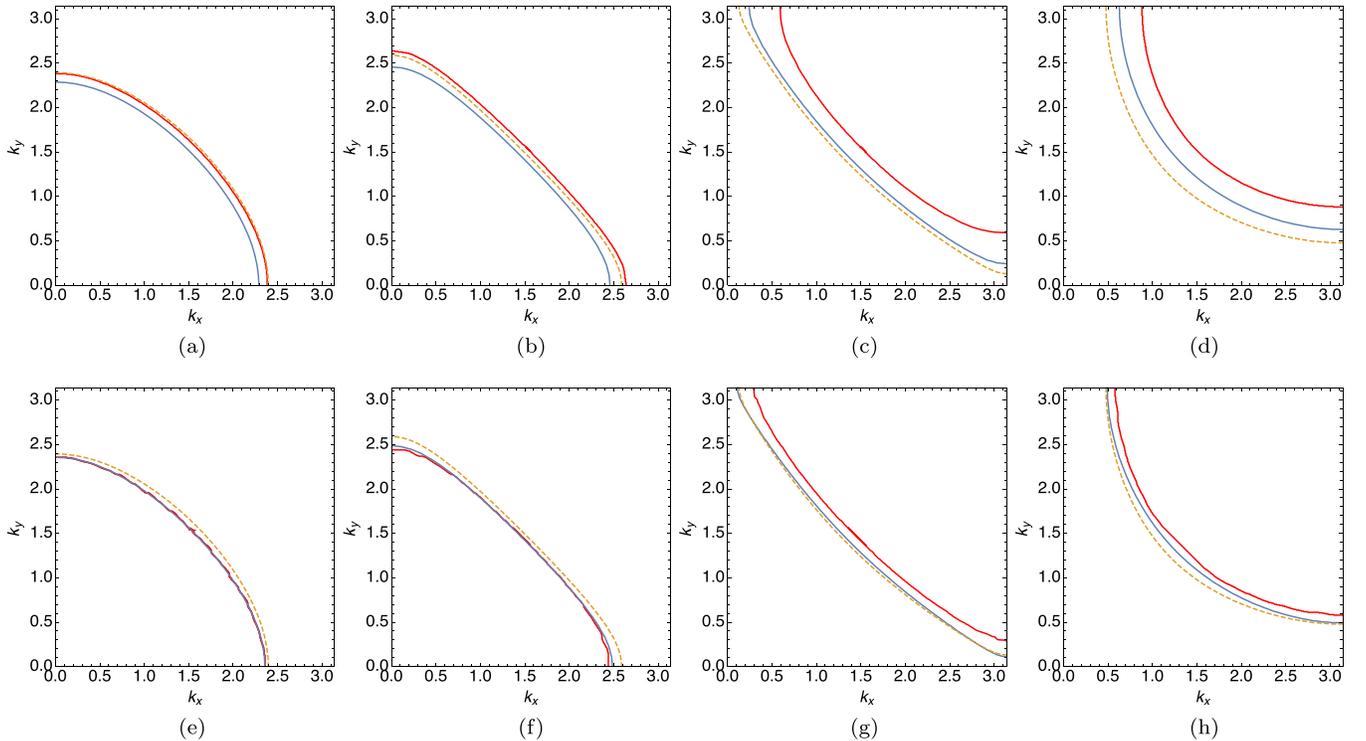


FIG. 11. Comparison between the pseudo-FS from  $\gamma_k$  (blue), the spectral peak (red), and the noninteracting FS (dashed) at various  $t'$  and fixed  $\delta = 0.15$ . Note that the spectral peak (location) and the pseudo-FS are not exactly the same, but deviate from the noninteracting FS in the same direction. As  $T$  decreases, the difference between them gets smaller. (a)  $t' = 0.2$ ,  $T = 400$  K. (b)  $t' = 0$ ,  $T = 400$  K. (c)  $t' = -0.2$ ,  $T = 440$  K. (d)  $t' = -0.4$ ,  $T = 420$  K. (e)  $t' = 0.2$ ,  $T = 105$  K. (f)  $t' = 0$ ,  $T = 105$  K. (g)  $t' = -0.2$ ,  $T = 270$  K. (h)  $t' = -0.4$ ,  $T = 105$  K.

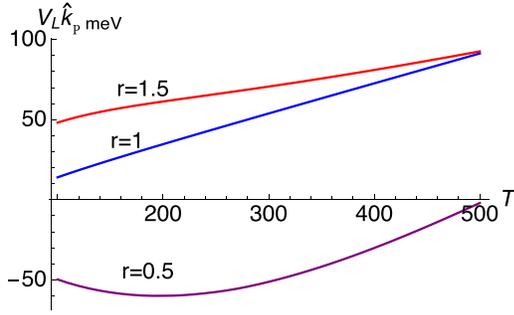


FIG. 12. The location of the peak of the spectral function  $A(\hat{k}, \omega)$  in Eq. (11) in units of  $\hat{k}_p V_L$  versus  $T$ , at three values of  $r$ . The model spectral function, Eq. (11), is from Ref. [48]. It is obtained by a low-energy expansion of the two ECFL self-energies  $\Psi$  and  $\Phi$  (equivalently  $\chi$ ) in Eq. (2) and Eq. (3). As  $T \rightarrow 0$  all the curves move towards  $\hat{k} = 0$  as one expects, but the approach from finite  $T$  displays significant differences depending on the value of  $r$ . The values of the parameters used here are  $\eta = 0.01$ ,  $\Delta_0 = 50$ ,  $\Omega_\Phi = 5000$  (in meV), and  $c_\alpha = 10$ . An estimated [48]  $V_L \sim 2 \text{ eV \AA}$  gives the shift in wave vector  $\Delta \hat{k} \sim .05 \text{ \AA}$ , at 500 K for  $r = 1.5$ .

serves as the scale of resistivity;  $\langle A \rangle_k \equiv \frac{1}{N_s} \sum_{\vec{k}} A(\vec{k})$ ;  $f$  is the Fermi distribution function. We present our results in absolute units in Fig. 15 by putting the measured values of the lattice constant into the formula and converting the energy unit using  $t = 0.45 \text{ eV} \approx 5220 \text{ K}$ . The scale of ECFL resistivity is consistent with the experimental findings in cuprates [24].

TABLE I. The Fermi liquid temperature  $T_{FL}$  obtained from fitting the data with Eq. (14). Increasing either  $t'$  (horizontally) or doping  $\delta$  (vertically) increases  $T_{FL}$ , signaling weaker correlations.

Fermi liquid temperature $T_{FL}$ (K)					
$\downarrow \delta \rightarrow t'$	-0.2	-0.1	0	0.1	0.2
0.12	10.0	18.4	33.1	68.2	117.6
0.15	15.8	31.1	66.3	135.4	218.0
0.18	24.4	53.7	117.4	245.2	420.9
0.21	37.3	78.8	189.5	360.3	618.4
0.24	56.8	145.2	274.4	569.5	820.5

In our previous study [15], a significant finding was that the curvature of resistivity changes when  $t'$  varies. Here we focus more on the  $\delta$ -dependent behavior of resistivity as shown in Fig. 15. For a given  $t'$ , decreasing the hole doping changes the curves from concave to linear then to convex and varying  $t'$  shifts the crossover doping region. This phenomenon signals a change of the effective Fermi temperature scale. In higher hole doping (lower electron density), there is less influence of the Gutzwiller projection. Hence the system has less correlation effectively and displays more Fermi-liquid-like behavior, namely,  $T^2$  dependence, and hence positive curvature. In the case with low hole doping, i.e., closer to the Mott-insulating limit, the correlation is relatively stronger and suppresses the Fermi liquid state into a much lower temperature region, which is usually masked

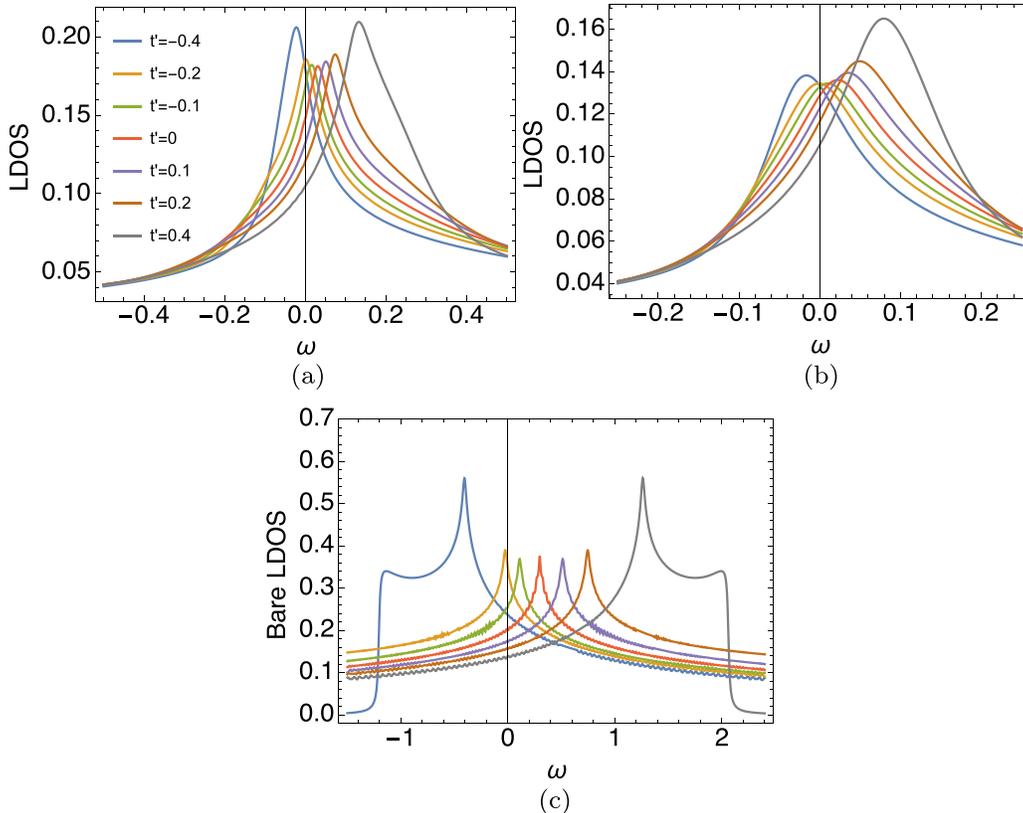


FIG. 13. Local density of states with varying  $t'$  while fixing  $\delta = 0.15$ , at  $T = 105 \text{ K}$  and  $400 \text{ K}$  from ECFL and at  $T = 0$  from the bare case. All figures share the same legend. (a)  $T = 105 \text{ K}$ . (b)  $T = 400 \text{ K}$ . (c) Tight-binding model for reference,  $T = 0$ .

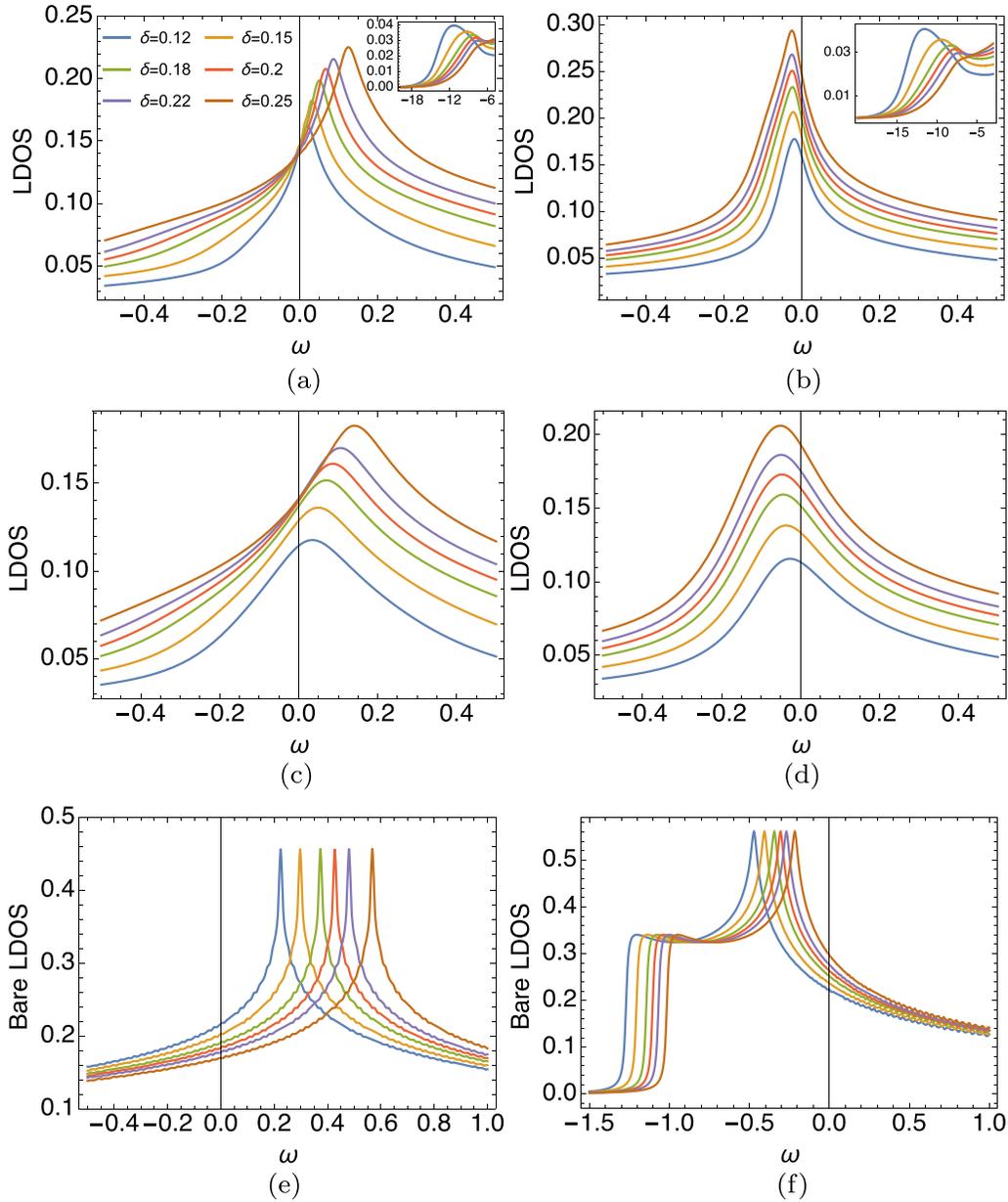


FIG. 14. Local density of states with varying  $\delta$  while fixing  $t' = 0$  and  $-0.4$ , at  $T = 105$  K and  $400$  K from ECFL and at  $T = 0$  from the bare case. All figures share the same legend. (a)  $t' = 0$ ,  $T = 105$  K. (b)  $t' = -0.4$ ,  $T = 105$  K. (c)  $t' = 0$ ,  $T = 400$  K. (d)  $t' = -0.4$ ,  $T = 400$  K. (e)  $t' = 0$ , tight-binding model for reference,  $T = 0$ . (f)  $t' = -0.4$ , tight-binding model for reference,  $T = 0$ .

by superconductivity. In the displayed temperature range of Fig. 15, the system shows strange-metal or even bad-metal behaviors [17] instead, and hence negative curvature. The curvature can be explicitly calculated as the second derivative of  $\rho_{xx}$  with respect to  $T$  shown in Fig. 16, which displays features qualitatively similar to the experiments [24,40–43].

To explore the crossover from the Fermi liquid ( $\rho_{xx} \propto T^2$ ) at low  $T$  to the strange metal ( $\rho_{xx} \propto T$ ) at higher  $T$ , we define a simple fitting model:

$$\rho_{\text{approx}} = \text{const.} \times \frac{T^2}{T_{FL} + T}. \quad (14)$$

This fit gives Fermi liquid behavior for  $T \lesssim T_{FL}$  and then crosses over to strange-metal linear behavior at  $T \gtrsim T_{FL}$ . Thus,  $T_{FL}$  serves as a crossover scale describing the boundary of the Fermi liquid region as well as estimating the strength of correlation. We find our data fit into this model well up to intermediate temperature with fitted coefficient and  $T_{FL}$ .

Table I shows the value of  $T_{FL}$  in various sets of  $\delta$  and  $t'$ . In all cases, the  $T_{FL}$  is considerably smaller than the Fermi temperature in the noninteracting case at the order of the bandwidth, as a result of strong correlation. In experiment, a small enough  $T_{FL}$  prevents the observation of the Fermi liquid because at low enough temperature the superconducting state shows up instead [24]. Relatively,  $T_{FL}$  is further suppressed

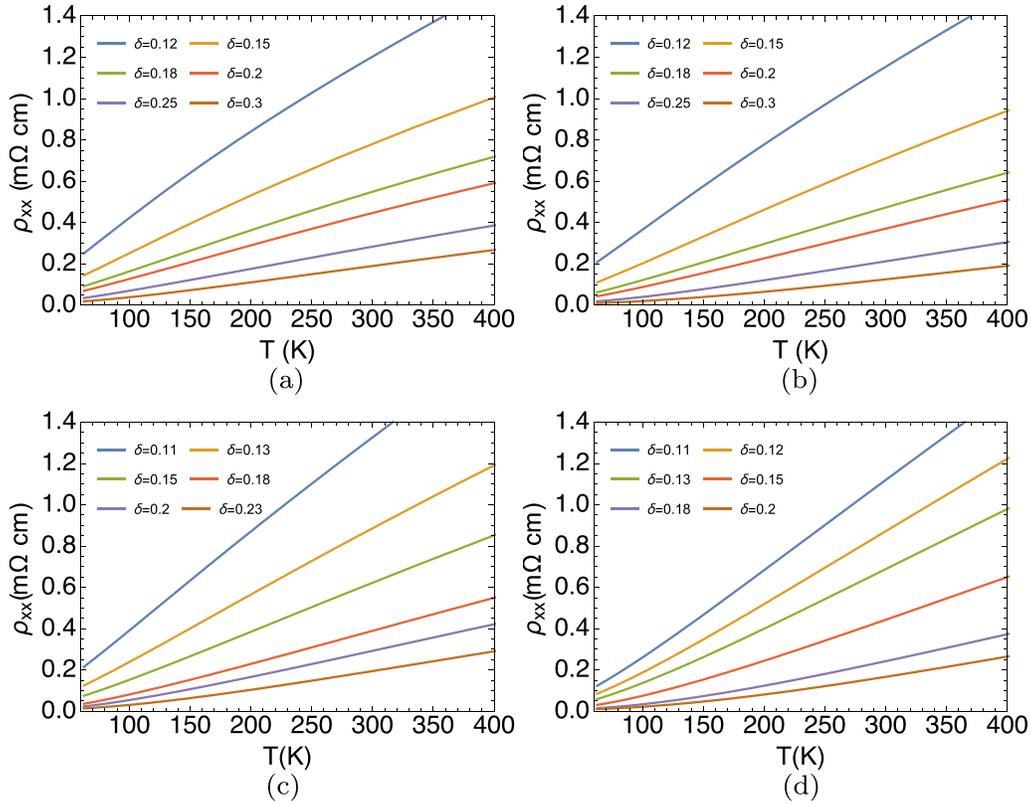


FIG. 15. Resistivity versus  $T$  for varying hole doping  $\delta$  and  $t' = -0.2, -0.1, 0, 0.2$  (some data in (a), (b), and (d) can be found in Ref. [15]). The curvature tends to change from negative (convex) to positive (concave) with increasing doping. (a)  $t' = -0.2$ . (b)  $t' = -0.1$ . (c)  $t' = 0$ . (d)  $t' = 0.2$ .

by smaller second-neighbor hopping  $t'$  or smaller doping  $\delta$ , either of which strengthens the effective correlation. Negative  $t'$  increases the resistivity and shrinks the temperature region for the Fermi liquid. In this sense, decreasing  $t'$  turns up the effective correlation by depressing the hopping process. On the other hand, decreasing doping leaves less space for electron movement, which also effectively increases the correlation and suppresses  $T_{FL}$ .  $\delta$  and  $t'$  both control the effective correlation strength and hence  $T_{FL}$ , as shown in Table I. Their similar role can also be understood in the fact that they both change the geometry of the Fermi surface which determines the conductivity at  $T \ll W$ , where  $W = 8t$  is the bare bandwidth. In general, either increasing  $\delta$  with fixed  $t'$  or increasing  $t'$  with fixed  $\delta$  changes the Fermi surface from hole-like to electron-like.

### C. Hall number

Within the bubble scheme, we also calculate the Hall conductivity [19,44–46] as  $\sigma_{xy} = (-2\pi^2/\rho_0)(\frac{\Phi}{\Phi_0})(\bar{\sigma}_{xy})$ . The dimensionless conductivity can be written as

$$\bar{\sigma}_{xy} = \frac{4\pi^2}{3} \int_{-\infty}^{\infty} d\omega (-\partial f / \partial \omega) \langle \rho_G^3(k, \omega) \eta(k) \rangle_k, \quad (15)$$

where  $\eta(k) = \frac{\hbar^2}{a_0^4} \{ (v_k^x)^2 \frac{\partial^2 \varepsilon_k}{\partial k_y^2} - (v_k^x v_k^y) \frac{\partial^2 \varepsilon_k}{\partial k_x \partial k_y} \}$ ;  $\Phi = Ba_0^2$  is the flux [47], and  $\Phi_0 = hc/(2|e|)$  is the flux quantum. In these

terms, we can compute the Hall number as

$$n_H = -\frac{1}{4\pi^2} \frac{\bar{\sigma}_{xx}^2}{\bar{\sigma}_{xy}}. \quad (16)$$

Note that in this definition, the sign of the Hall number is opposite to that in Ref. [15]. In this definition,  $n_H$  shares the same sign with the Hall coefficient  $R_H$ , consistent with the experimental convention [24,40–43,49–53]. We present the ECFL Hall number  $n_H$  in Fig. 17 together with the non-interacting one  $n_{H0}$  for comparison. In all cases of different  $t'$ ,  $n_H$  is around 60% of  $n_{H0}$  and decreasing  $t'$  suppresses the scale of  $n_H$ . This indicates the reduction of effective charge carrier due to strong correlation. Therefore, the Hall number increases when the effective correlation turns down either by increasing  $t'$  or increasing  $\delta$ , as shown in Fig. 17. In panel (d),  $n_H$  remains smooth when crossing the Lifshitz transition  $\delta \approx 0.17$ , where the Fermi surface changes from opened to closed as presented in Sec. III A, while  $n_{H0}$  shows a crossover to a steeper region.

### D. Spin susceptibility and the NMR relaxation rate

The imaginary part of spin susceptibility can also be calculated in the Bubble approximation:

$$\chi''(k, \omega) = \int_{-\infty}^{\infty} dy \langle \rho_G(p, y) \rho_G(p+k, y+\omega) \rangle_p \times [f(y) - f(y+\omega)], \quad (17)$$

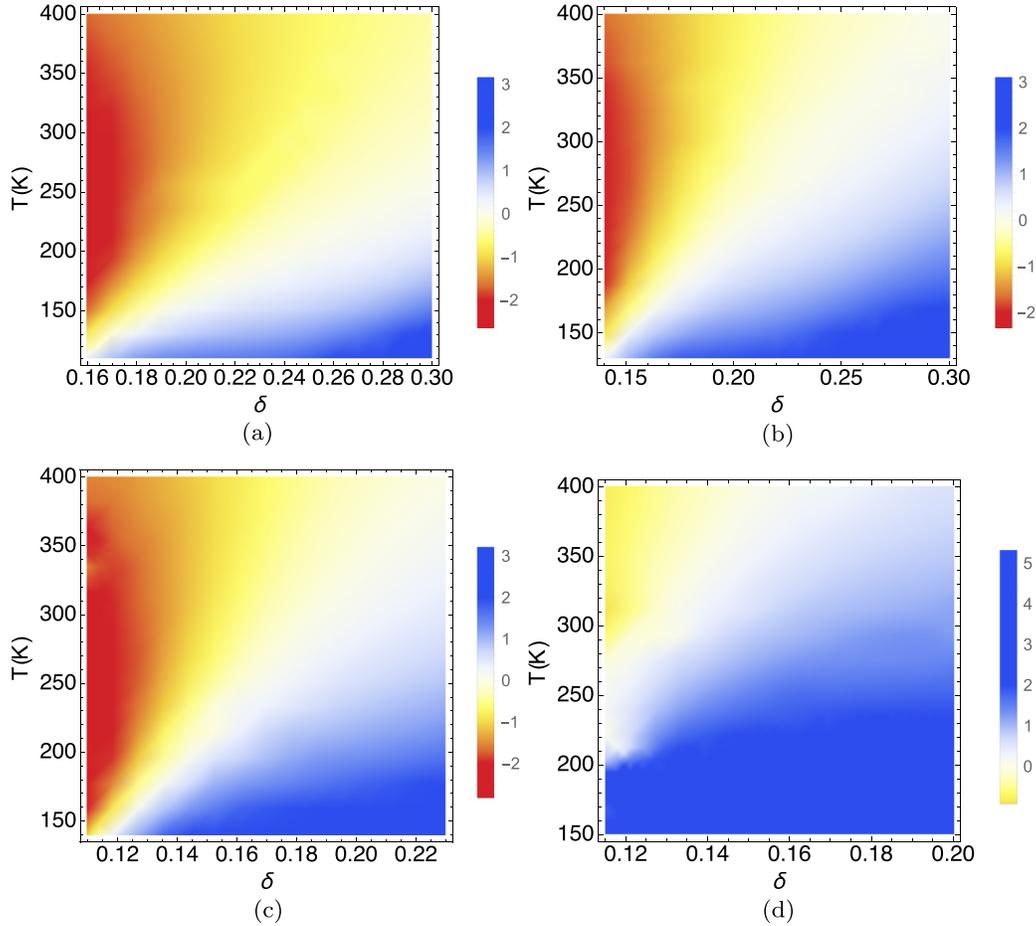


FIG. 16. Curvature of resistivity versus  $T$  for a range of doping  $\delta$  and  $t' = -0.2, -0.1, 0$ , and  $0.2$ . For most values of  $t'$ , there is a blue area towards the right bottom representing positive (concave) curvature akin to a Fermi liquid. Towards the left top we find a red area with negative (convex) curvature resembling a strange (or bad) metal [17]. This trend is consistent with experimental results [24]. (a)  $t' = -0.2$ . (b)  $t' = -0.1$ . (c)  $t' = 0$ . (d)  $t' = 0.2$ .

while the real part  $\chi'$  can be obtained from calculating the Hilbert transform of  $\chi''$ .  $\chi''$  is shown in Fig. 18 for hole-doped ( $t' = -0.2$ ) and electron-doped ( $t' = 0.2$ ) cases at various fixed  $k$ . In both cases, we see the quasielastic peaks in the occupied region for small  $k$  which disappears gradually as  $k$  increases.

Figure 19 presents the  $k$ -dependent  $\chi'$  at zero frequency, in comparison with the noninteracting  $\chi'_0$  in the inset. We observe that  $\chi'$  is much smaller than  $\chi'_0$  due to the broadening in the spectral function as a result of strong interaction. Despite the scale difference, the  $k$ -dependent  $\chi$  seems closer to  $\chi_0$  in the electron-doped case ( $t' = 0.2$ ) than the hole-doped case ( $t' = -0.2$ ), consistent with the previous discussion that the system is more Fermi-liquid-like for positive  $t'$ . The Knight shift  $\chi'(k = 0, \omega = 0)$  of the system is almost independent of temperature and therefore not shown specifically in figure.

The relaxation rates for cuprates are given by [25,26,55]

$$\frac{1}{T_1} = \frac{\gamma^2 k_B T}{\mu_B^2} \sum_q A_q^2 \frac{\chi''(q, \omega_0)}{\omega_0}, \quad (18)$$

where  $A_q$  is a form factor that is determined by the local geometry of the nucleus [25,26,55], and  $\omega_0$  is nuclear

frequency which is assumed to be very small. Our scheme of calculation is not yet refined enough to capture the detailed difference between the copper and oxygen relaxation rates in cuprates. Hence, we will content ourselves by presenting the case with  $A_q = 1$ , which should correspond to the inelastic neutron scattering (INS) derived relaxation rate in Ref. [25] from Walstedt *et al.* We plot  $1/T_1$  vs  $T$  at  $\delta = 0.15$  and various  $t'$  in Fig. 20. For  $t' = -0.2$ ,  $1/T_1$  increases sublinearly with temperature. It shows roughly the same trend as the copper rates shown in Ref. [25], but is somewhat steeper than the derived INS rate therein.

### E. $J$ variation

Above we have discussed the ECFL results at  $J = 0.17$ . We next address the question of variation with  $J$ . Figure 21 shows the EDCs and MDCs at different  $J$  fixing  $t' = 0$ . Turning on  $J$  raises the peak in EDC [(a)  $\rightarrow$  (c)  $\rightarrow$  (e)] and MDC [(b)  $\rightarrow$  (d)  $\rightarrow$  (f)] slightly. Also, increasing  $J$  separates the other EDC lines farther away from  $k = k_F$  while bringing the other MDC lines closer to  $\omega = 0$ .

We find that  $J$  has an effect on the effective bandwidth. This can be seen in the EDC and MDC dispersion relation in

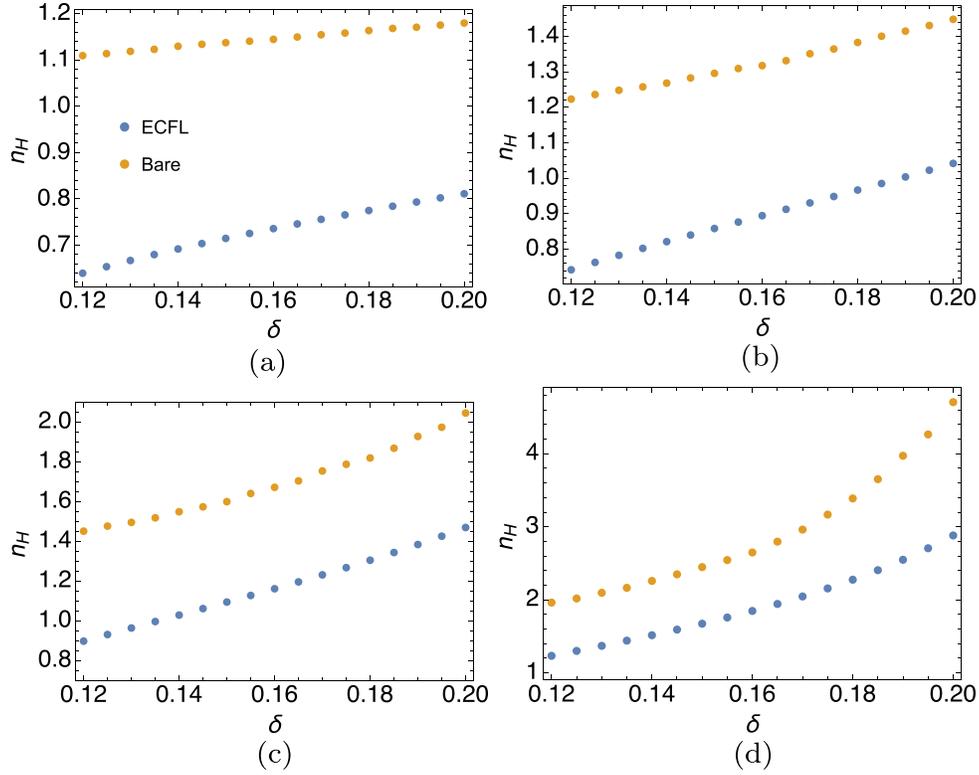


FIG. 17. Hall number vs doping at different  $t'$  and  $T = 105$  K, where  $t'$  controls the scale of  $n_H$ . (a)  $t' = -0.4$ . (b)  $t' = -0.3$ . (c)  $t' = -0.25$ . (d)  $t' = -0.2$ .

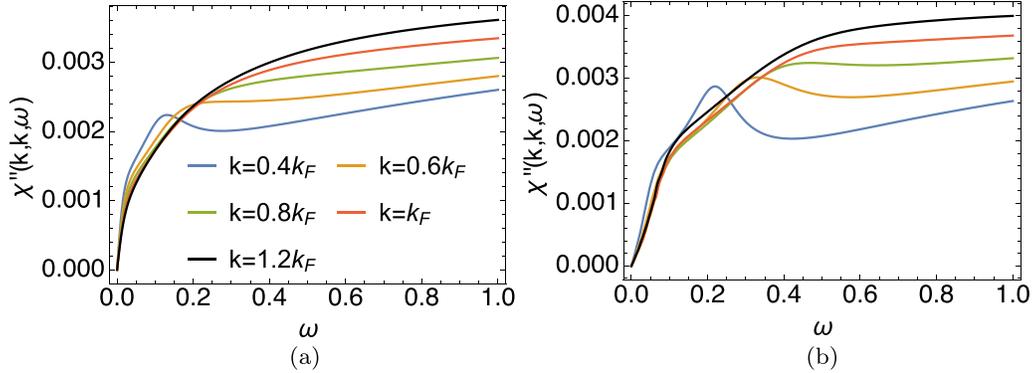


FIG. 18.  $\chi''$  at different  $k$  for  $\delta = 0.15$ ,  $T = 63$  K, and  $t' = \pm 0.2$ . (a)  $t' = -0.2$ . (b)  $t' = 0.2$ .

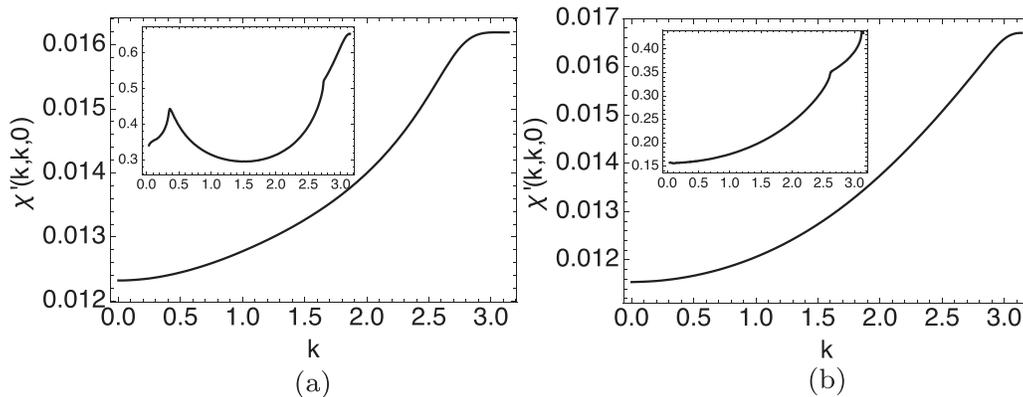


FIG. 19.  $\chi'$  at  $\omega = 0$  for  $\delta = 0.15$ ,  $T = 63$  K, and  $t' = \pm 0.2$ . Inset shows the corresponding noninteracting  $\chi'_0$ .  $\chi'$  is largely suppressed from the bare case due to strong interaction. (a)  $t' = -0.2$ . (b)  $t' = 0.2$ .

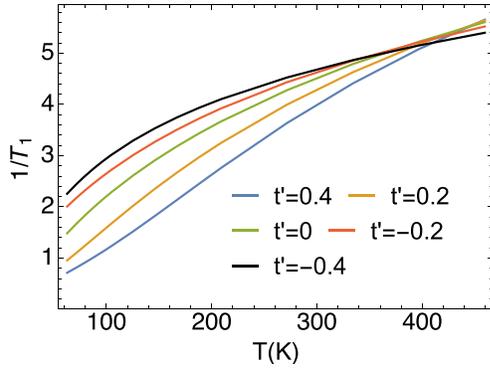


FIG. 20. Relaxation rate from Eq. (18) (arb. units) at  $\delta = 0.15$  and different  $t'$ . The curve becomes more sublinear as  $t'$  decreases from positive to negative. The sublinear curve at  $t' = -0.2$  looks similar to the copper relaxation rate in Ref. [25].

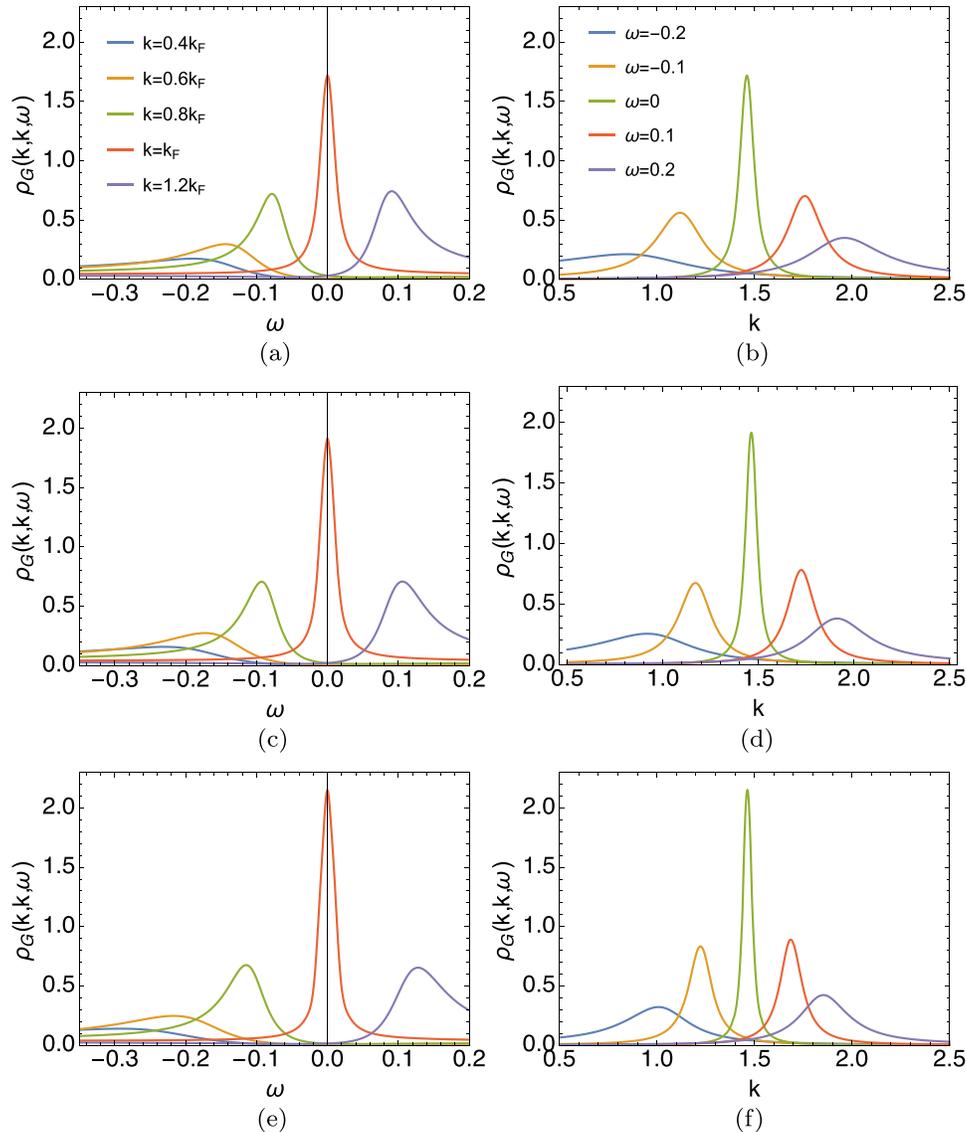


FIG. 21. EDC and MDC line shapes at different values of superexchange  $J$ . All EDC figures [(a), (c), (e)] or MDC figures [(b), (d), (f)] share the same respective legend. Here the parameters are set as  $\delta = 0.15$ ,  $t' = 0$ ,  $T = 105$  K, and  $J = 0, 0.17, 0.4$ , in the nodal ( $\Gamma \rightarrow X$ ) direction. Increasing  $J$ , the peak at the chemical potential becomes somewhat higher, but it remains qualitatively similar at all  $J$ . Besides, increasing  $J$  separates the EDC lines farther away from  $k = k_F$  and brings the MDC lines closer to  $\omega = 0$ . (a)  $J = 0$ , EDC. (b)  $J = 0$ , MDC. (c)  $J = 0.17$ , EDC. (d)  $J = 0.17$ , MDC. (e)  $J = 0.4$ , EDC. (f)  $J = 0.4$ , MDC.

Fig. 22. As  $J$  increases, the EDC and MDC bands separate out more widely, though they are still very narrow (due to strong correlations) compared to the bare bandwidth. The MDC dispersion shows a high-energy feature, namely the kink (or waterfall). Due to the finite lattice size and to the second-order approximation made in the present work, the low-energy kink discussed in Ref. [48] cannot be resolved clearly. Another angle to view the effect of  $J$  is through the 3D plot of the nodal direction spectral function  $\rho_G(k, k, \omega)$  in Fig. 23. It appears that turning on  $J$  rotates the spectral function counterclockwise with respect to the  $z$  axis with  $k = k_F$  and  $\omega = 0$  if viewed from above. In other words, increasing  $J$  extended the renormalized bandwidth with no effect on the Fermi surface location since all curves cross at the same  $k_F$ . That said, small variation of  $J$  does not change the system behavior qualitatively, and only slightly in

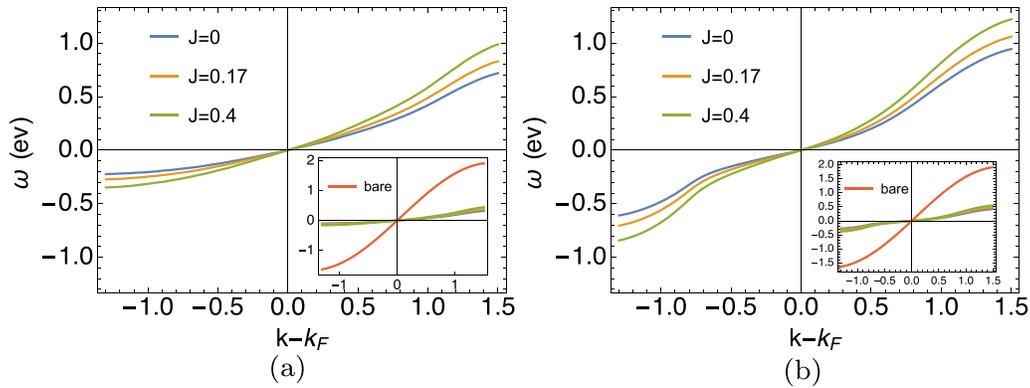


FIG. 22. (a) EDC and (b) MDC dispersion relation at different values of superexchange  $J$ . In both cases, increasing  $J$  expands the renormalized bandwidth, consistent with Fig. 21 of EDC and MDC lines. Both insets show that the renormalized band is strongly suppressed by correlation compared with the bare one. The energy and  $k$  resolution in the present study is not fine enough to deduce the detailed properties of the low-energy kinks (for  $\omega \sim 0.07$  eV) discussed phenomenologically within ECFL in Ref. [48].

quantitative detail. Therefore it is reasonable to set  $J = 0.17$  from experiment as a representative number and to explore the  $k$ ,  $\omega$ ,  $t'$ , and  $\delta$  dependence of the system.

From the discussion above, we expect the  $k$ -average physical quantity like resistivity with significant contribution from the area around the Fermi surface to be insensitive to  $J$  variation. Figure 24 shows the resistivity at different  $J$  for fixed  $t'$ . As expected, varying  $J$  from 0 to 0.4 does not make a qualitative difference in the resistivity of the normal state,

although it has a relatively stronger effect on the case with larger  $|t'|$ .

#### IV. CONCLUSION

We apply the recently developed second-order ECFL scheme [15,16] to studying the 2D  $t$ - $J$  model with second-nearest-neighbor hopping  $t'$ . We have presented the spectral function, self-energy, LDOS, resistivity, Hall number, and dy-

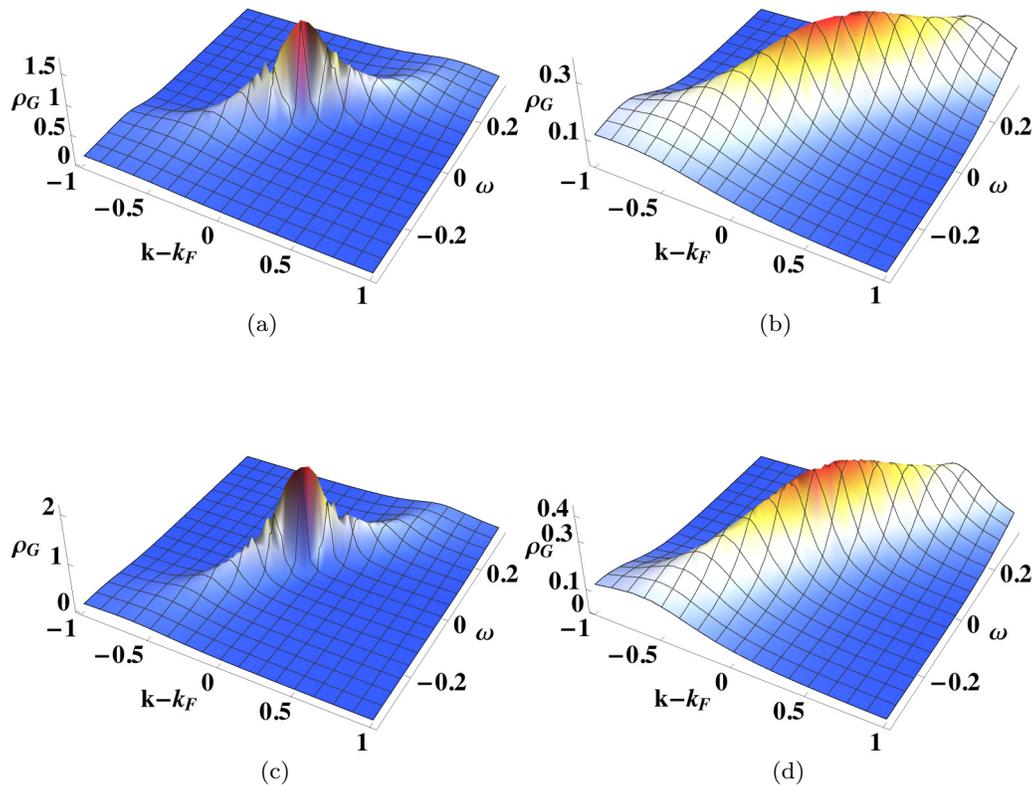


FIG. 23. 3D plot of the nodal direction spectral function  $\rho_G(k, k, \omega)$ . Consistent with Fig. 21, turning on  $J$  increases the peak height and rotates  $\rho_G$  counterclockwise with respect to the  $z$  axis with  $k = k_F$  and  $\omega = 0$  if viewed from above. This is another facet of the steeper dispersion with  $J$  noted in Fig. 22. (a)  $J = 0$ ,  $T = 105$  K. (b)  $J = 0$ ,  $T = 400$  K. (c)  $J = 0.4$ ,  $T = 105$  K. (d)  $J = 0.4$ ,  $T = 400$  K.

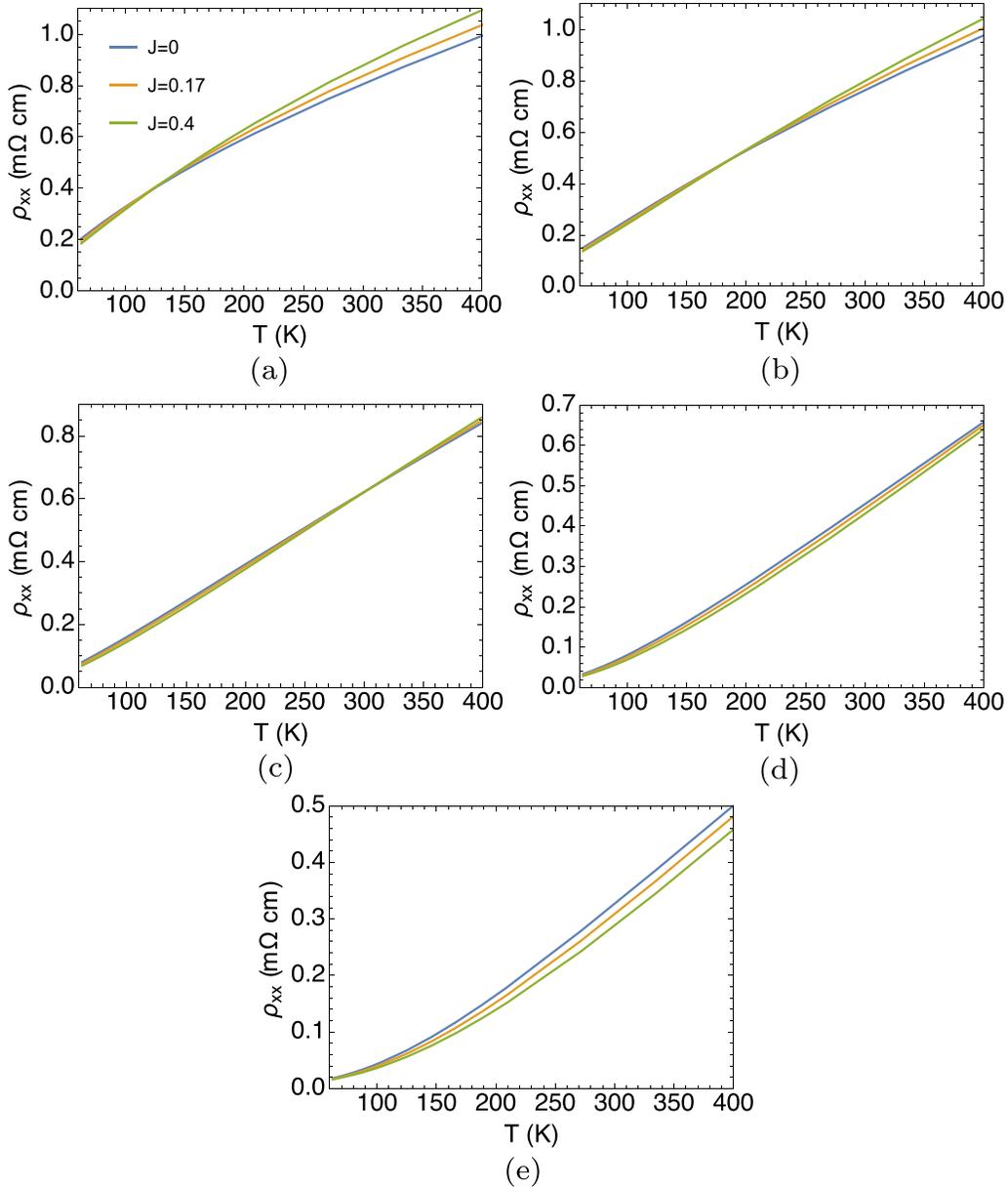


FIG. 24. Resistivity at  $\delta = 0.15$  versus  $T$  for various  $J$  and  $t'$  (same legend for all panels). In all  $t'$ , we observe that  $J$  variation of the resistivity is small. As  $|t'|$  becomes large  $J$  has a somewhat larger influence on the resistivity. (a)  $t' = -0.4$ . (b)  $t' = -0.2$ . (c)  $t' = 0$ . (d)  $t' = 0.2$ . (e)  $t' = 0.4$ .

namical susceptibility at low and intermediate temperatures, with  $t'$  varying from  $-0.4$  to  $0.4$  and within a large density region around optimal doping.

The spectral properties are shown to be consistent with ARPES experiments [56–60] on correlated material. The asymmetric EDCs and more symmetric MDCs are observed as expected from the previous study on the phenomenological model of simplified ECFL theory [23]. The weak  $k$  dependence of self-energy indicates the relative unimportance of vertex corrections at the densities considered, and gives credence to the use of the bubble approximation for transport.

The curvature change on the resistivity  $\rho$ - $T$  curve arises from varying  $t'$  and  $\delta$ , signaling different strength of effective

correlation. Both  $t'$  and  $\delta$  affect the effective electron-electron correlation because  $t'$  controls the second-neighbor hopping process and  $\delta$  leaves more or less space for electron movement. As a feature in 2D, the combination of them determines the geometry of the Fermi surface and therefore the low-energy behaviors.

#### ACKNOWLEDGMENTS

We thank Edward Perepelitsky and Michael Arciniaga for helpful comments on the manuscript. We thank Shawdong Dong for helpful suggestions with the computations. The work at UCSC was supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences, un-

der Award No. DE-FG02-06ER46319. Computations reported here used the XSEDE environment [54] (TG-DMR170044)

supported by National Science Foundation Grant No. ACI-1053575.

- [1] P. W. Anderson, *Science* **235**, 1196 (1987).
- [2] M. Ogata and H. Fukuyama, *Rep. Prog. Phys.* **71**, 036501 (2008).
- [3] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, *Rev. Mod. Phys.* **78**, 865 (2006).
- [4] K. Haule and G. Kotliar, *Europhys. Lett.* **77**, 27007 (2007).
- [5] K. Haule and G. Kotliar, *Phys. Rev. B* **76**, 104509 (2007).
- [6] K. Bouadim, N. Paris, F. Hebert, G. G. Batrouni, and R. T. Scalettar, *Phys. Rev. B* **76**, 085112 (2007).
- [7] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, *Rev. Mod. Phys.* **77**, 1027 (2005).
- [8] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011).
- [9] A. Go and A. J. Millis, *Phys. Rev. Lett.* **114**, 016402 (2015).
- [10] X. Wang, H. T. Dang, and A. J. Millis, *Phys. Rev. B* **84**, 014530 (2011).
- [11] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).
- [12] B. S. Shastry, *Ann. Phys.* **343**, 164 (2014); **373**, 717(E) (2016).
- [13] F. J. Dyson, *Phys. Rev.* **102**, 1230 (1956); S. V. Maleev, *Zh. Eksp. Teor. Fiz.* **33**, 1010 (1958) [*Sov. Phys. JETP* **6**, 776 (1958)].
- [14] E. Perepelitsky and B. S. Shastry, *Ann. Phys.* **357**, 1 (2015).
- [15] B. S. Shastry and P. Mai, *New J. Phys.* **20**, 013027 (2017).
- [16] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016); R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *ibid.* **88**, 235132 (2013); B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *ibid.* **88**, 205108 (2013).
- [17] W. Ding, R. Žitko, P. Mai, E. Perepelitsky, and B. S. Shastry, *Phys. Rev. B* **96**, 054114 (2017); W. Ding, R. Žitko, and B. S. Shastry, *ibid.* **96**, 115153 (2017).
- [18] X. Y. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [19] W. Xu, K. Haule, and G. Kotliar, *Phys. Rev. Lett.* **111**, 036401 (2013).
- [20] A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [21] P. Mai, S. R. White, and B. S. Shastry, *Phys. Rev. B* **98**, 035108 (2018).
- [22] P. Mai and B. S. Shastry, *Phys. Rev. B* **98**, 115101 (2018).
- [23] G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).
- [24] Y. Ando, S. Komiya, K. Segawa, S. Ono, and Y. Kurita, *Phys. Rev. Lett.* **93**, 267001 (2004).
- [25] R. E. Walstedt, T. E. Mason, G. Aeppli, S. M. Hayden, and H. A. Mook, *Phys. Rev. B* **84**, 024530 (2011).
- [26] R. E. Walstedt, *The NMR Probe of High- $T_c$  Materials and Correlated Electron Systems*, Springer Tracts in Modern Physics (Springer, New York, 2017).
- [27] Observe that in these equations, an arbitrary shift of the band  $\varepsilon_k \rightarrow \varepsilon_k + c$  can be absorbed into  $u_0$ . Thus the *shift invariance* is manifest to second order in  $\lambda$ .
- [28] B. S. Shastry, *Phys. Rev. B* **84**, 165112 (2011); **86**, 079911(E) (2012).
- [29] D. Hansen and B. S. Shastry, *Phys. Rev. B* **87**, 245101 (2013).
- [30] K. Matsuyama and G.-H. Gweon, *Phys. Rev. Lett.* **111**, 246401 (2013).
- [31] J. Chang, M. Shi, S. Pailh s, M. M nsson, T. Claesson, O. Tjernberg, A. Bendounan, Y. Sassa, L. Patthey, N. Momono, M. Oda, M. Ido, S. Guerrero, C. Mudry, and J. Mesot, *Phys. Rev. B* **78**, 205103 (2008); N. Doiron-Leyraud, O. Cyr-Choini re, S. Badoux, A. Ataei, C. Collignon, A. Gourgout, S. Dufour-Beaus jour, F. F. Tafti, F. Lalibert , M.-E. Boulanger, M. Matusiak, D. Graf, M. Kim, J.-S. Zhou, N. Momono, T. Kurosawa, H. Takagi, and L. Taillefer, *Nat. Commun.* **8**, 2044 (2017).
- [32] T. Yoshida, X. J. Zhou, K. Tanaka, W. L. Yang, Z. Hussain, Z.-X. Shen, A. Fujimori, S. Sahrakorpi, M. Lindroos, R. S. Markiewicz, A. Bansil, Seiki Komiya, Yoichi Ando, H. Eisaki, T. Kakeshita, and S. Uchida, *Phys. Rev. B* **74**, 224510 (2006).
- [33] T. Yoshida, X. J. Zhou, M. Nakamura, S. A. Kellar, P. V. Bogdanov, E. D. Lu, A. Lanzara, Z. Hussain, A. Ino, T. Mizokawa, A. Fujimori, H. Eisaki, C. Kim, Z.-X. Shen, T. Kakeshita, and S. Uchida, *Phys. Rev. B* **63**, 220501(R) (2001).
- [34] B. S. Shastry, [arXiv:1808.00405](https://arxiv.org/abs/1808.00405).
- [35] F. Ming, S. Johnston, D. Mulugeta, T. S. Smith, P. Vilmercati, G. Lee, T. A. Maier, P. C. Snijders, and H. H. Weitering, *Phys. Rev. Lett.* **119**, 266802 (2017).
- [36] Y. J. Yan, M. Q. Ren, H. C. Xu, B. P. Xie, R. Tao, H. Y. Choi, N. Lee, Y. J. Choi, T. Zhang, and D. L. Feng, *Phys. Rev. X* **5**, 041018 (2015).
- [37] P. Choubey, A. Kreisel, T. Berlijn, B. M. Andersen, and P. J. Hirschfeld, *Phys. Rev. B* **96**, 174523 (2017).
- [38] A. Kreisel, P. Choubey, T. Berlijn, W. Ku, B. M. Andersen, and P. J. Hirschfeld, *Phys. Rev. Lett.* **114**, 217002 (2015).
- [39] K. Fujita, A. R. Schmidt, E.-A. Kim, M. J. Lawler, D. H. Lee, J. C. Davis, H. Eisaki, and S.-i. Uchida, *J. Phys. Soc. Jpn.* **81**, 011005 (2012).
- [40] S. Martin, A. T. Fiory, R. M. Fleming, L. F. Schneemeyer, and J. V. Waszczak, *Phys. Rev. Lett.* **60**, 2194 (1988).
- [41] H. Takagi, T. Ido, S. Ishibashi, M. Uota, S. Uchida, and Y. Tokura, *Phys. Rev. B* **40**, 2254 (1989).
- [42] Y. Onose, Y. Taguchi, K. Ishizaka, and Y. Tokura, *Phys. Rev. B* **69**, 024504 (2004).
- [43] Y. Li, W. Tabis, G. Yu, N. Bari i , and M. Greven, *Phys. Rev. Lett.* **117**, 197001 (2016).
- [44] P. Voruganti, A. Golubentsev, and S. John, *Phys. Rev. B* **45**, 13945 (1992); H. Fukuyama, H. Ebisawa, and Y. Wada, *Prog. Theor. Phys.* **42**, 494 (1969); H. Kohno and K. Yamada, *ibid.* **80**, 623 (1988).
- [45] For this we additionally assume that the magnetic field vertex also assumes its bare value. This assumption requires further validation in 2 dimensions within the  $t$ - $J$  model; hence the results for the Hall conductivity are less reliable than the longitudinal conductivity.
- [46] L.-F. Arsenault and A. M. S. Tremblay, *Phys. Rev. B* **88**, 205109 (2013).
- [47] The numerics assume a bct unit cell  $(a, a, c)$  with  $a = 3.79 \text{ \AA}$  and  $c = 13.29 \text{ \AA}$ . In the expression for  $\rho_0$ ,  $c_0$  corresponds to the interlayer separation  $c_0 = c/2$ .

- [48] K. Matsuyama, E. Perepelitsky, and B. S. Shastry, *Phys. Rev. B* **95**, 165435 (2017).
- [49] H. Y. Hwang, B. Batlogg, H. Takagi, H. L. Kao, J. Kwo, R. J. Cava, J. J. Krajewski, and W. F. Peck Jr., *Phys. Rev. Lett.* **72**, 2636 (1994).
- [50] Y. Ando, Y. Kurita, S. Komiyama, S. Ono, and K. Segawa, *Phys. Rev. Lett.* **92**, 197001 (2004).
- [51] F. F. Balakirev, J. B. Betts, A. Migliori, S. Ono, Y. Ando, and G. S. Boebinger, *Nature (London)* **424**, 912 (2003).
- [52] F. F. Balakirev, J. B. Betts, A. Migliori, I. Tsukada, Y. Ando, and G. S. Boebinger, *Phys. Rev. Lett.* **102**, 017004 (2009).
- [53] J. Takeda, T. Nishikawa, and M. Sato, *Phys. C (Amsterdam, Neth.)* **231**, 293 (1994) (see especially Fig. 4 therein).
- [54] J. Town, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, N. W-Diehr, *Comput. Sci. Eng.* **16**, 62 (2014).
- [55] B. S. Shastry, *Phys. Rev. Lett.* **63**, 1288 (1989).
- [56] A. Damascelli, Z. Hussain, and Z.-X. Shen, *Rev. Mod. Phys.* **75**, 473 (2003).
- [57] W. S. Lee, I. M. Vishik, D. H. Lu, and Z.-X. Shen, *J. Phys.: Condens. Matter* **21**, 164217 (2009).
- [58] J. D. Koralek, J. F. Douglas, N. C. Plumb, Z. Sun, A. V. Federov, M. M. Murnane, H. C. Kapteyn, S. T. Cundiff, Y. Aiura, K. Oka, H. Eisaki, and D. S. Dessau, *Phys. Rev. Lett.* **96**, 017005 (2006).
- [59] T. Yoshida, X. J. Zhou, D. H. Lu, S. Komiyama, Y. Ando, H. Eisaki, T. Kakeshita, S. Uchida, Z. Hussain, Z.-X. Shen, and A. Fujimori, *J. Phys.: Condens. Matter* **19**, 125209 (2007).
- [60] N. P. Armitage, D. H. Lu, C. Kim, A. Damascelli, K. M. Shen, F. Ronning, D. L. Feng, P. Bogdanov, X. J. Zhou, W. L. Yang, Z. Hussain, P. K. Mang, N. Kaneko, M. Greven, Y. Onose, Y. Taguchi, Y. Tokura, and Z.-X. Shen, *Phys. Rev. B* **68**, 064517 (2003).

# Theory of anisotropic elasto-resistivity of two-dimensional extremely strongly correlated metals

Michael Arciniaga<sup>1</sup>, Peizhi Mai<sup>2,1</sup> and B. Sriram Shastry<sup>1</sup>

<sup>1</sup>Physics Department, University of California, Santa Cruz, California 95064, USA

<sup>2</sup>Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6494, USA



(Received 13 September 2019; revised manuscript received 17 May 2020; accepted 18 May 2020; published 19 June 2020)

There is considerable recent interest in the phenomenon of anisotropic electroresistivity of correlated metals. While some interesting work has been done on the iron-based superconducting systems, not much is known for the cuprate materials. Here we study the anisotropy of elasto-resistivity for cuprates in the normal state. We present theoretical results for the effect of strain on resistivity and additionally on the optical weight and local density of states. We use the recently developed extremely strongly correlated Fermi liquid theory in two dimensions, which accounts quantitatively for the unstrained resistivities for three families of single-layer cuprates. The strained hoppings of a tight-binding model are roughly modeled analogously to strained transition metals. The strained resistivity for a two-dimensional  $t$ - $t'$ - $J$  model are then obtained, using the equations developed in recent work. Our quantitative predictions for these quantities have the prospect of experimental tests in the near future, for strongly correlated materials such as the hole-doped and electron-doped high- $T_c$  materials.

DOI: [10.1103/PhysRevB.101.245149](https://doi.org/10.1103/PhysRevB.101.245149)

## I. INTRODUCTION & MOTIVATION

Understanding the temperature and doping dependent electrical conductivity of very strongly correlated metals in two dimensions (2D) is a very important problem in condensed matter physics. Recent interest in *elasto-resistivity*, i.e., the strain dependence of resistivity has been triggered by the discovery of strong nematicity in iron based superconductors [1–3]. The nematic susceptibility is defined as

$$\chi_{\text{nem}} = \lim_{\epsilon_{xx} \rightarrow 0} \frac{\rho'_{xx} - \rho_{xx}}{\rho_{xx} \epsilon_{xx}}, \quad (1)$$

where  $\rho'_{xx}(\rho_{xx})$  is the  $x$ -axis resistivity in the presence (absence) of a small strain  $\epsilon_{xx}$ . The large magnitude of this dimensionless susceptibility ( $|\chi_{\text{nem}}| \gtrsim 200$ ), and the peaklike features in its temperature dependence, suggest enhanced nematic fluctuations in the pnictides.

The situation for cuprates is less studied thus motivating the present work. The recently developed extremely correlated Fermi liquid theory (ECFL) [4] accounts quantitatively for the (unstrained) normal state resistivities of three families of single layer cuprates [5–7]. This theory treats correlation effects within the well-defined  $t$ - $t'$ - $J$  model. The model lacks any explicit mechanism to drive large nematic fluctuations, but it is possible that these fluctuations are emergent. It is thus natural to ask if the theory can provide a benchmark scale for elasto-resistivity effects in cuprates, as well as to examine if nematic fluctuation are encouraged. Towards this goal we present results for the anisotropic elasto-resistivity in various geometries for cuprate materials in the normal state within the extremely correlated Fermi liquid theory (ECFL) [4] as applied to the  $t$ - $t'$ - $J$  model for spin- $\frac{1}{2}$  electrons on a square

lattice given by the Hamiltonian

$$H = - \sum_{ij\sigma} t_{ij} \tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma} - \mu \hat{N} + \frac{1}{2} \sum_{ij} J_{ij} \left( \tilde{S}_i \cdot \tilde{S}_j - \frac{1}{4} n_i n_j \right). \quad (2)$$

Here  $t_{ij} = t(t')$  for nearest (next-nearest) neighbor sites  $ij$  and is zero otherwise on the square lattice [8],  $\hat{N}$  is the number operator,  $\tilde{C}_{i\sigma} = P_G C_{i\sigma} P_G$ , and  $P_G$  is the Gutzwiller projection operator which projects out the doubly occupied states. Also the superexchange  $J_{ij} = J$  when acting on nearest neighbor sites and is zero otherwise. The other symbols have their usual meaning.

While the ECFL theory accounts for the variation of resistivity with a change of hopping parameters, we need another piece of information to calculate elasto-resistivity. That is a solution to the independent problem of describing the effects of strain on the hopping parameters of the underlying tight-binding model. In cuprates the  $t$ - $t'$ - $J$  model arises as an effective low energy model from downfolding from a three band (or in general multiband) description obtained within band structure calculations[9–11]. This procedure is not unique since the extent of correlations included in the band structure can differ among different calculations. We take the practical view that the hopping parameters can be chosen to depend parametrically on the distance between atoms, in parallel to the treatment of volume effects in transition metals by V. Heine[12]. Thus in our approach, a small strain can be parametrized through a single variable  $\alpha$  relating the hopping to the separation  $R$  via the relation

$$t(R) \sim \frac{A}{R^\alpha}. \quad (3)$$

From tight-binding theory  $\alpha = l_1 + l_2 + 1$ , where  $l_1, l_2$  are the angular momenta of the overlapping orbitals [12]. Within

this scheme we expect that compression enhances overlap and hence the magnitude of hopping, and conversely stretching reduces overlap. Excluding very strong multiband effects we may take  $\alpha \in \{2, 5\}$  for cuprates. The single parameter needed for our purpose is  $\alpha$ , since  $A$  is reabsorbed in the unstrained hopping. We further suggest that one may more realistically estimate this single parameter  $\alpha$  by measuring other  $\alpha$  dependent variation of physical variables with strain, as described below.

This modeling neglects the possible three-dimensional effects, where the  $c$ -axis propagation could in certain situations influence the two-dimensional bands indirectly. Also cuprates with many layers per unit cell may have more complex dependence on strain as compared to single layer systems. Despite the above caveats in place, it is still worthwhile to study the model Eq. (2) together with the relation Eq. (3) for understanding the elastoresistivity of single layer cuprates.

The problem of (unstrained) normal state resistivity has been explored in various experiments [13–15] on different materials over the last few decades. Experiments reveal interesting and challenging transport regimes, termed the *strange metal* and the *bad metal* regime [16], whose existence is inexplicable within the standard Fermi liquid theory of metals. These results have attracted several numerical studies using the techniques of dynamical mean field theory [17–19], determinant quantum Monte-Carlo method [20,21], and dynamical cluster approximation [22,23] etc. These studies indicate that the unusual regimes are indicative of very strong correlations of the Mott-Hubbard variety.

Despite the numerical progress, few analytical techniques are available to extract the low temperature transport behavior and thus better understand the various regimes. This is due to the inherent difficulties of treating strong correlations, i.e., physics beyond the scope of perturbation theory. Recently, the extremely correlated Fermi liquid theory (ECFL) [4,24,25] has been developed by Shastry and coworkers. This theory consists of a basic reformulation of strong correlation physics, and its many applications have been reported for the  $t$ - $t'$ - $J$  model in dimensions  $d = 1, 2, \infty$ . This is a minimal and fundamental model to describe extreme correlations. The ECFL theory leads to encouraging results which are in close accord with experiments such as spectral line shape in angle-resolved photoemission spectroscopy (ARPES) [6,7,26–31], Raman susceptibility [32,33], and particularly, resistivity [6,7,25,34]. A recent work [5] shows that the ECFL theory gives a quantitatively consistent account of the  $T$  and density dependence of the resistivity for single layer hole-doped and electron-doped correlated materials. Here we explore the strain dependence of the resistivity within the same scheme.

In the ECFL theory, the resistivity arises from (Umklapp-type) inelastic scattering between strongly correlated electrons. Here the hopping amplitudes of electrons play a dual role. The first one, that of propagating the fragile quasiparticles, is standard in all electronic systems. They provide a simple model for the band structure. Additionally, for very strong correlations the ECFL theory shows that the hopping parameters are also involved in the scattering of quasiparticles off each other [35]. A surprisingly low characteristic temperature scale [7,34] emerges from the strong correlations, above which the resistivity crosses over from Fermi liquid type,

i.e.,  $\rho \sim T^2$  behavior, to an almost linear type, i.e.,  $\rho \sim T$  behavior [6,7,15].

From the above we argue that strain effects could provide a test of the underlying mechanism for resistivity within the ECFL theory to include strain dependence. Experiments probing these strain effects are likely in the near future, thus enabling an important test of the theory. For the purpose of independently estimating the strain-hopping parameter  $\alpha$  in Eq. (3), we have identified two experimentally accessible variables. Firstly we study the integrated weight of the anisotropic electrical optical conductivity, i.e., the f-sum rule weight, accessible in optical experiments [36,37]. Secondly we study the local density of states (LDOS), measurable through scanning tunneling microscopy (STM) [38–42]. The f-sum rule weight in tight-binding systems is related to the expectation of the kinetic energy, or hopping, and can be obtained from the Green's function. The LDOS can also be calculated from the local Green's function easily.

The plan of the paper is as follows: In Sec. II (A) we introduce the  $t$ - $t'$ - $J$  model and summarize the second order ECFL equations and the corresponding Green's functions and self-energies. (B) We describe how to convert the lattice constants and hopping parameters for a system under strain. (C) We outline the parameters for the program. In Sec. III, we present the detailed calculation for and results of (A) the resistivity, (B) the kinetic energy, and (C) the LDOS and their associated susceptibilities with respect to strain. We provide a brief summary and discussion of our results and future work in Sec. IV.

## II. METHODS & PARAMETERS

### A. The model

It has been argued that the  $t$ - $t'$ - $J$  model is key to describing the physics of high- $T_c$  superconducting materials [43]. This model is composed of two terms:  $H_{tJ} = H_t + H_J$  where  $H_t$  is derived by taking the infinite- $U$  limit of the Hubbard model plus an additional term  $H_J$  which introduces antiferromagnetic coupling. The general Hamiltonian Eq. (2) can be rewritten in terms of the Hubbard  $X$  operators [4] as

$$\begin{aligned} H_t &= - \sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i\sigma} X_i^{\sigma\sigma}, \\ H_J &= \frac{1}{2} \sum_{ij\sigma} J_{ij} X_i^{\sigma\sigma} \\ &\quad + \frac{1}{4} \sum_{ij\sigma_1\sigma_2} J_{ij} \{ X_i^{\sigma_1\sigma_2} X_j^{\sigma_2\sigma_1} - X_i^{\sigma_1\sigma_1} X_j^{\sigma_2\sigma_2} \}. \end{aligned} \quad (4)$$

Here  $t_{ij}$  and  $J_{ij}$  are already defined below Eq. (2). We present results for both vanishing and nonvanishing  $J_{ij}$ . The operator  $X_i^{ab} = |a\rangle \langle b|$  takes the electron at site  $i$  from the state  $|b\rangle$  to the state  $|a\rangle$  where  $|a\rangle$  and  $|b\rangle$  are one of the three allowed states: two occupied states  $|\uparrow\rangle, |\downarrow\rangle$ , or the unoccupied state  $|0\rangle$ —the appropriate  $X$  operator referring to the doubly occupied state  $|\uparrow\downarrow\rangle$  is excluded in both the Hamiltonian and state space. The  $X$  operator relates to the alternative representation used in Eq. (2) as follows:  $X_i^{\sigma 0} \rightarrow \tilde{C}_{i\sigma}^\dagger, X_i^{0\sigma} \rightarrow \tilde{C}_{i\sigma}$ , and  $\sum_{\sigma} X_i^{\sigma\sigma} \rightarrow n_i$ .

### B. The ECFL equations

In this section, we briefly introduce the ECFL equations for the  $t$ - $t'$ - $J$  model. More details can be found in [4,6,24,25]. In the ECFL theory, the one-electron Green's function  $\mathcal{G}$  is found using the Schwinger method [44] and in momentum space is factored as a product of an auxiliary Green's function  $\mathbf{g}$  and a "caparison" function  $\tilde{\mu}$ :

$$\mathcal{G}(k) = \mathbf{g}(k) \times \tilde{\mu}(k), \quad (5)$$

where  $k \equiv (\vec{k}, i\omega_k)$ , and  $\omega_k = (2k+1)\pi k_B T$  is the fermionic Matsubara frequency and subscript  $k$  is an integer. The auxiliary  $\mathbf{g}(k)$  plays the role of a Fermi-liquid type Green's function whose asymptotic behavior is  $1/\omega$  as  $\omega \rightarrow \infty$ , and  $\tilde{\mu}$  is an adaptive spectral weight that mediates between two conflicting requirements [24]: (1) the high frequency behavior of the noncanonical fermions and (2) the Luttinger-Ward volume theorem at low frequencies.

The Schwinger equation of motion for the physical Green's function can be symbolically written as [24]

$$(\mathbf{g}_0^{-1} - \hat{X} - Y_1) \cdot \mathcal{G} = \delta (1 - \gamma), \quad (6)$$

where  $\hat{X}$  represents a functional derivative and  $Y_1$  describes a Hartree-type energy, i.e.,  $\mathcal{G}$  convoluted with hopping and exchange interactions. The left hand side of Eq. (6) is analogous to that of the Schwinger-Dyson equation for Hubbard model [45]:  $(\mathbf{g}_0^{-1} - U\delta/\delta\mathcal{V} - U\mathcal{G}) \cdot \mathcal{G} = \delta \mathbf{1}$ . Observe on the right side of Eq. (6), the essential difference is the  $\gamma$  term which is proportional to a local  $\mathcal{G}$  and originates from the non-canonical algebra of creation and annihilation operators. The noncanonical nature of operators and the lack of an obvious small parameter for expansion present the main difficulties towards solving this equation.

To tackle these difficulties, the ECFL theory inserts into Eq. (6) the  $\lambda$  parameter

$$(\mathbf{g}_0^{-1} - \lambda\hat{X} - \lambda Y_1) \cdot \mathcal{G} = \delta (1 - \lambda\gamma), \quad (7)$$

where  $\lambda \in [0, 1]$  interpolates from a noninteracting to fully interacting system. This parameter plays a parallel role to that of inverse spin parameter  $1/2S$  in quantum magnets, where  $S$  is the magnitude of the spin. Then we expand Eq. (7) systematically with respect to  $\lambda$  up to a finite order and at the end set  $\lambda = 1$  to recover the full  $t$ - $t'$ - $J$  physics. The introduction of  $\lambda$  bound to  $[0, 1]$  in ECFL makes it possible that a low-order expansion could be enough to describe low-energy excitations in a large region of doping. This argument has been justified in one [46] and infinite [25] dimensions by benchmarking against exact numerical techniques and in two [6,7] dimensions by comparing well with experiments.

In the following, we use the minimal version of second order (in  $\lambda$ ) ECFL equations [6]:

$$\tilde{\mu}(k) = 1 - \lambda \frac{n}{2} + \lambda \psi(k) \quad (8)$$

$$\mathbf{g}^{-1}(k) = i\omega_k + \boldsymbol{\mu} - \epsilon_{\vec{k}} + \lambda \frac{n}{2} \epsilon_{\vec{k}} - \lambda \phi(k) \quad (9)$$

where  $\boldsymbol{\mu}$  is the chemical potential (denoted in boldface) and  $\epsilon_{\vec{k}}$  is the bare band energy found by taking the Fourier transformation of the hopping parameter. The physical Green's function features two self-energy terms: the usual Dyson-like self-energy denoted  $\phi(k)$  in the denominator and a second

self-energy in the numerator  $\psi(k)$ . The self-energy  $\phi(k)$  can conveniently be decomposed as follows:  $\phi(k) = \chi(k) + \epsilon'_{\vec{k}} \psi(k)$  where  $\chi(k)$  denotes a self-energy part,  $\epsilon'_{\vec{k}} = \epsilon_{\vec{k}} - u_0/2$  and  $\psi(k)$  the second self-energy. Here  $u_0$  acts as a Lagrange multiplier, enforcing the shift invariance [4,6,24] of the  $t$ - $t'$ - $J$  model at every order of  $\lambda$ . The two self-energies functions  $\psi$  and  $\chi$  expanded formally in  $\lambda$  to second order approximation  $\mathcal{O}(\lambda^2)$  are  $\psi = \psi_{[0]} + \lambda\psi_{[1]} + \dots$  and  $\chi = \chi_{[0]} + \lambda\chi_{[1]} + \dots$ . The expression for these self-energies in the expansion are

$$\psi_{[0]}(k) = 0, \quad \chi_{[0]}(k) = - \sum_p \left( \epsilon'_{\vec{p}} + \frac{1}{2} J_{\vec{k}-\vec{p}} \right) \mathbf{g}(p) \quad (10)$$

and

$$\psi_{[1]}(k) = - \sum_{pq} (\epsilon'_{\vec{p}} + \epsilon'_{\vec{q}} + J_{\vec{k}-\vec{p}}) \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k) \quad (11)$$

$$\begin{aligned} \chi_{[1]}(k) = & - \sum_{pq} (\epsilon'_{\vec{p}} + \epsilon'_{\vec{q}} + J_{\vec{k}-\vec{q}}) (\epsilon'_{\vec{p}+\vec{q}-\vec{k}} + J_{\vec{k}-\vec{p}}) \\ & \times \mathbf{g}(p) \mathbf{g}(q) \mathbf{g}(p+q-k), \end{aligned} \quad (12)$$

where  $\sum_k \equiv \frac{k_B T}{N_s} \sum_{\vec{k}, \omega_k}$  and  $J_{\vec{q}}$  is the Fourier transform of  $J_{ij}$  [47]. By setting  $\lambda$  to 1, the resulting expressions for the ECFL equations expanded to  $\mathcal{O}(\lambda^2)$  are

$$\tilde{\mu}(k) = 1 - \frac{n}{2} + \psi(k) \quad (13)$$

$$\begin{aligned} \mathbf{g}^{-1}(k) = & i\omega_k + \boldsymbol{\mu} - \epsilon_{\vec{k}} + \frac{n}{2} \epsilon_{\vec{k}} - \chi_{[0]}(k) \\ & - \chi_{[1]}(k) - \epsilon'_{\vec{p}} \psi_{[1]}(k). \end{aligned} \quad (14)$$

We can verify that an arbitrary shift of  $\epsilon_{\vec{k}} \rightarrow \epsilon_{\vec{k}} + \mathbf{c}_0$  leaves the above expression invariant by shifting  $\boldsymbol{\mu} \rightarrow \boldsymbol{\mu} + \mathbf{c}_0$  and  $u_0 \rightarrow u_0 + 2\mathbf{c}_0$ . In this sense, we may take  $u_0$  as a second chemical potential. We can determine the two chemical potentials  $\boldsymbol{\mu}$  and  $u_0$  by satisfying the following number sum rules

$$\sum_k \mathbf{g}(k) e^{i\omega_k 0^+} = \frac{n}{2} = \sum_k \mathcal{G}(k) e^{i\omega_k 0^+}, \quad (15)$$

where  $n$  is the particle density. We find the spectral function  $\rho_{\mathcal{G}}(k) = -1/\pi \Im m \mathcal{G}(k)$  by analytically continuing (i.e.,  $i\omega_k \rightarrow \omega + i\eta$ ) and by solving Eq. (5) and Eqs. (10)–(15) iteratively. We remind the reader that the spectral function  $\rho_{\mathcal{G}}(\vec{k}, \omega)$  is referred to in most experimental literature by the symbol  $A(\vec{k}, \omega)$ . We can recover the interacting Green's function from  $\rho_{\mathcal{G}}$  using

$$\mathcal{G}(\vec{k}, i\omega_k) = \int_{-\infty}^{\infty} \frac{\rho_{\mathcal{G}}(\vec{k}, \nu)}{i\omega_k - \nu} d\nu. \quad (16)$$

### C. Strain effects on hopping and exchange

#### 1. Converting lattice constant changes to hopping changes

The  $t$ - $t'$ - $J$  model in two dimensions describes the hopping of electrons between copper atoms in the 2d plane. In this model, the hopping parameters with strain and without strain are denoted as

$$\{t_x, t_y, t_d\} \rightarrow \{t, t, t'\}. \quad (17)$$

Thus under strain  $t_x$  and  $t_y$  refer to nearest neighbor hops along  $x$  and  $y$  axes, and  $t_d$  is the second neighbor hopping along the diagonal of the square lattice. We start with the tetragonal symmetry case  $t_x = t_y = t$  where there are just two parameters  $t, t'$ .

At the level of a single bond between two coppers, any generic hopping  $t(R)$  for a bond with length  $R$  can be represented by [12]

$$t(R) \sim \frac{A}{R^\alpha}, \quad (18)$$

where  $A$  is a constant. In the simplest cases, the exponent  $\alpha$  is given by the angular momentum  $l_1, l_2$  of the relevant atomic shells of the two atoms by the formula

$$\alpha = l_1 + l_2 + 1. \quad (19)$$

Thus for two copper atoms  $l_1 = l_2 = 2$  and hence we might expect

$$\alpha \sim 5, \quad (20)$$

whereas for copper oxygen bonds  $l_1 = 2, l_2 = 1$ , therefore

$$\alpha \sim 3. \quad (21)$$

For the effective single band description of the cuprate materials, it is not entirely clear what value of  $\alpha$  is most appropriate. Comparisons with experiments might be the best way to decide on this question, when the results become available. Until then we can bypass this issue by presenting the theoretical results in terms of  $\frac{\delta t}{t}$  rather than the strain itself. Towards this end Eq. (18) is a very useful result. We rewrite it as

$$\frac{\delta t(R)}{t(R)} = -\alpha \frac{\delta R}{R}, \quad (22)$$

thus enabling us to convert a change of the lattice constant to that of the corresponding hopping, using only the value of  $t$  and  $\alpha$ . Throughout this paper we will refer to  $\delta t/t$  as ‘‘strain’’ or with emphasis as ‘‘hopping strain’’ in order to distinguish it from ‘‘conventional strain’’  $\delta R/R$ . Strain will always refer to variations along the  $x$  axis unless otherwise noted.

## 2. Geometrical aspects of the strain variation

Our calculation studies a few variations of parameters. We start on a lattice with tetragonal symmetry at  $t \sim 5220$  K (0.45 eV), and we vary  $t'$  to capture both electron-doped ( $t' > 0$ ) and hole-doped ( $t' < 0$ ) cuprates. The magnitude of  $t$  is only a crude estimate, it is refined for different single layer cuprate systems in [5].

On the distorted lattice with orthorhombic symmetry and lattice constants  $a$  and  $b$ , the three distances of interest (two sets of nearest neighbors and one set of second neighbors) are

$$a, b, \rho = \sqrt{a^2 + b^2}. \quad (23)$$

For the tetragonal case we refer to the undistorted lattice parameter as  $a_0$ , thus  $a = b = a_0, \rho = \sqrt{2}a_0$ . We next study the effect of stretching ( $\delta a > 0$ ) or compressing ( $\delta a < 0$ ) the  $x$ -axis lattice constant, leaving the  $y$  axis unchanged. The changes in the lattice constants then read as

$$a \rightarrow a_0 + \delta a; \quad b \rightarrow a_0; \quad \rho \rightarrow \sqrt{2}a_0 + \frac{\delta a}{\sqrt{2}}. \quad (24)$$

We denote the strain in the  $x$  direction as

$$\epsilon_{xx} = \frac{\delta a}{a_0}. \quad (25)$$

In terms of the strain, we can rewrite the distances to neighbors as

$$a = a_0(1 + \epsilon_{xx}), \quad b = a_0, \quad \rho = \sqrt{2}a_0 \left(1 + \frac{\epsilon_{xx}}{2}\right), \quad (26)$$

so that  $\epsilon_{xx} > 0$  is regarded as stretching and  $\epsilon_{xx} < 0$  as compression. The single particle (tight-binding) energies for the distorted lattice are given by

$$\epsilon_{\vec{k}} = -2t_x \cos(k_x a) - 2t_y \cos(k_y b) - 4t_d \cos(k_x a) \cos(k_y b).$$

In terms of the band parameters of the unstrained system  $t$  and  $t'$ , we can write the anisotropic band parameters as

$$t_x = (1 - \alpha \epsilon_{xx}) t, \quad t_y = t, \quad t_d = \left(1 - \alpha \frac{\epsilon_{xx}}{2}\right) t', \quad (27)$$

where the factor of  $\frac{1}{2}$  for  $t_d$  comes about due to a shorter stretching of  $\rho$  as in Eq. (26). Their strain variations are denoted by

$$\frac{\delta t_x}{t_x} \equiv \frac{\delta t}{t} = -\alpha \epsilon_{xx}, \quad \frac{\delta t_y}{t_y} = 0, \quad \frac{\delta t_d}{t_d} = -\frac{1}{2} \alpha \epsilon_{xx}. \quad (28)$$

These formulas relate the change in hopping to the physical strain, and thus involve the parameter  $\alpha$  which is somewhat uncertain. For that reason, we actually vary  $\frac{\delta t}{t}$  in this study. We also go beyond the linear response regime, i.e., we use larger values of  $\frac{\delta t}{t}$  than those attainable in the laboratory. In such a case we set  $\frac{\delta t_d}{t_d} = \frac{\delta t}{2t}$ . To summarize the sign convention used in this work,

$$\begin{aligned} \text{compress: } & \frac{\delta t}{t} > 0, \quad \epsilon_{xx} < 0 \\ \text{stretch: } & \frac{\delta t}{t} < 0, \quad \epsilon_{xx} > 0. \end{aligned} \quad (29)$$

## 3. Converting hopping changes into exchange changes

In this model, the superexchange interaction maps to hopping as follows:  $J = t^2/U$  where  $U$  is the on site energy of the Hubbard model. As we vary the hopping parameter, we find  $\delta J = 2(\delta t/t)J$  since  $U$  does not vary with strain. In this model the first neighbor exchange parameters with and without strain, similar to Eq. (17), are denoted as

$$\{J_x, J_y\} \rightarrow \{J, J\}, \quad (30)$$

where  $J_x$  and  $J_y$  refer to the first neighbor exchange interactions along the  $x$  and  $y$  axes. In terms of hopping changes we can rewrite the exchange parameters as

$$J_x = \left(1 + 2 \frac{\delta t_{xx}}{t_{xx}}\right) J, \quad J_y = J. \quad (31)$$

## D. Parameters in the program

The model considered applies to several classes of materials, such as the cuprates, the sodium cobaltates, and presumably also to the iron arsenide superconductors. We shall restrict our discussion to the cuprates where the parameters are fairly well agreed upon in the community [5,43,48].

In this calculation, we set  $t = 1$  as our energy scale and we allow  $t'/t$  to vary between  $-0.4$  and  $0.4$ , to cover the full range of cuprate materials. The hopping strain  $\delta t/t$  is varied from  $-0.15$  to  $0.15$ . The exchange parameter  $J$  is set to zero except where otherwise noted. We convert the energy to physical units by setting  $t = 0.45$  eV, and hence the bandwidth is  $W = 8t = 3.6$  eV. If one wants to make a different choice for  $t$ , this can be done by rescaling the energies and  $T$ 's by the same scaling factor.

We focus on the optimal doping case  $\delta = 0.15$  for cuprate materials [49]. Here  $\delta$  refers to the hole doping and relates to the particle density as follows  $\delta \equiv (1 - n)$ . The temperature range is set to  $T \in [37, 450]$  K. Lower temperatures than this lie outside the range of convergence for the current scheme. For the interacting system we solve the ECFL equations (10)–(15) iteratively on a real frequency grid of size  $N_\omega = 2^{14}$  within the range  $[-2.5W, 2.5W]$ , where  $W$  is the bare bandwidth, and a lattice  $L \times L$  with  $L = 61, 79, 135$ . The scale of the frequency grid is tuned to capture the low- $T$  physics. A frequency grid of size  $N_\omega = 2^{16}$  only slightly improves our results at much larger computational costs. We primarily use an  $L > 61$  for  $t' > 0$  at low temperatures (i.e.,  $T < 100$  K) in order to get sufficient resolution to converge electrical resistivity calculation. The need for a high resolution lattice at low temperatures is a product of the spectral function which features higher, sharper peaks for  $t' > 0$ , to which the resistivity calculation is sensitive [6], i.e., a larger grid is required to settle the unphysical oscillations in the resistivity calculation. For the noninteracting system we compute LDOS using a system of size  $N_\omega = 2^{12}$  and  $L = 271$ .

### III. RESULTS

Here we present the effects of strain along the  $x$  axis on electrical resistivity, kinetic energy, and LDOS and their associated susceptibilities in response to a compressive ( $\delta t/t > 0$ ) and tensile ( $\delta t/t < 0$ ) hopping strain.

#### A. Resistivity for an $x$ -axis strain

We now study the response of electrical resistivity  $\rho_\alpha$  characterized by electron-electron scattering [6] in the presence of a strain. We use the bubble approximation, factoring the current correlator as  $\langle J(t)J(0) \rangle \sim \sum_k v_k^2 \mathcal{G}^2(k)$  with suitable vertices  $v_k$  and dressed Green's function  $\mathcal{G}$ , to compute the conductivity  $\sigma_\alpha$ . Our picture of a quasi-2D metal consists of well separated Cu-O planes and hence each plane can be characterized using the 2D  $t$ - $J$  model. The weak  $k$  dependence of the self-energy as seen in Fig. 3 of Ref. [7] diminishes the significance of vertex corrections. In fact the self-energy is completely  $k$  independent in the  $d = \infty$  limit, and studies in this limit [34] have successfully implemented the bubble approximation while completely ignoring vertex corrections. We shall calculate and quote the following objects denoting the irreducible representations of the  $D_{4h}$  point group by the standard names [2,50–52]

(i)  $\rho'_{xx}(T)$  the strained version of resistivity along the  $x$  axis.

(ii)  $\rho'_{yy}(T)$  the strained version of resistivity along the  $y$  axis.

(iii)  $\rho_{xx}$  without a prime refers to the tetragonal result, which is the same as  $\rho_{yy}$ .

(iv)  $XX$  component variations:

$$-(\rho'_{xx} - \rho_{xx})/(\rho_{xx}\delta t/t) \text{ vs } T$$

(v)  $YY$  component variations:

$$-(\rho'_{yy} - \rho_{yy})/(\rho_{xx}\delta t/t) \text{ vs } T$$

(vi)  $A_{1g}$  symmetry variations:

$$-\frac{\rho'_{xx} + \rho'_{yy} - 2\rho_{xx}}{2\rho_{xx}\delta t/t} \text{ vs } T$$

(vii)  $B_{1g}$  symmetry variations:

$$-\frac{\rho'_{xx} - \rho'_{yy}}{\rho_{xx}\delta t/t} \text{ vs } T$$

Of special interest are the  $\rho'_{xx} + \rho'_{yy}$  response which corresponds to the  $A_{1g}$  irreducible representation (irrep) and the  $\rho'_{xx} - \rho'_{yy}$  response, corresponding to the  $B_{1g}$  irrep.

#### 1. Computation of the anisotropic resistivity

To find the anisotropic resistivity, we compute the dimensionless conductivity [6] for the anisotropic case

$$\sigma_{xx} = \langle \Upsilon_{\vec{k}} (\hbar v_{\vec{k}}^x)^2 / (ab) \rangle_k, \quad (32)$$

$$\sigma_{yy} = \langle \Upsilon_{\vec{k}} (\hbar v_{\vec{k}}^y)^2 / (ab) \rangle_k, \quad (33)$$

where  $\langle A \rangle_k = \frac{1}{N_s} \sum_{\vec{k}} A$ ,  $N_s = L \times L$  and

$$\Upsilon_{\vec{k}} = (2\pi)^2 \int_{-\infty}^{\infty} d\omega (-\partial f / \partial \omega) \rho_G^2(\vec{k}, \omega), \quad (34)$$

where  $f(\omega) \equiv 1/(1 + \exp(\beta\omega))$  is the Fermi function,  $\rho_G(k)$  is the spectral function from from ECFL theory up to  $\mathcal{O}(\lambda^2)$ , and  $v_{\vec{k}}^x, v_{\vec{k}}^y$  are the bare vertices, which are defined as

$$v_{\vec{k}}^x = \frac{1}{\hbar} \frac{\partial \epsilon_{\vec{k}}}{\partial k_x} = \frac{a}{\hbar} \frac{\partial \epsilon_{\vec{k}}}{\partial k_1}, \quad (35)$$

$$v_{\vec{k}}^y = \frac{1}{\hbar} \frac{\partial \epsilon_{\vec{k}}}{\partial k_y} = \frac{b}{\hbar} \frac{\partial \epsilon_{\vec{k}}}{\partial k_2}, \quad (36)$$

where  $k_1 = k_x a$  and  $k_2 = k_y b$  denote the components of the dimensionless momenta. Inserting the dimensionless momenta into Eq. (33), we obtain

$$\sigma_{xx} = \left\langle \Upsilon_{\vec{k}} \left( \frac{d\epsilon_{\vec{k}}}{dk_1} \right)^2 (a/b) \right\rangle_k, \quad (37)$$

$$\sigma_{yy} = \left\langle \Upsilon_{\vec{k}} \left( \frac{d\epsilon_{\vec{k}}}{dk_2} \right)^2 (b/a) \right\rangle_k \quad (38)$$

for the dimensionless conductivity. The corresponding dimensionless resistivities are  $\rho_{xx} = 1/\sigma_{xx}$  and  $\rho_{yy} = 1/\sigma_{yy}$ .

The electrical resistivity can be converted to physical units as follows:  $\rho_{\text{physical},\alpha} = \rho_\alpha \times \rho_0$  where  $\rho_0 = c_0 \hbar / e^2 (\sim 1.171 \text{ m}\Omega \text{ cm})$  sets the scale for the resistivity, and  $\alpha = xx$  describes the longitudinal (i.e., current  $\parallel \epsilon_{xx}$ ) resistivity and  $yy$  describes the transverse (i.e., current  $\perp \epsilon_{xx}$ ) resistivity. Here  $c_0 \sim 6.645 \text{ \AA}$  is the typical separation

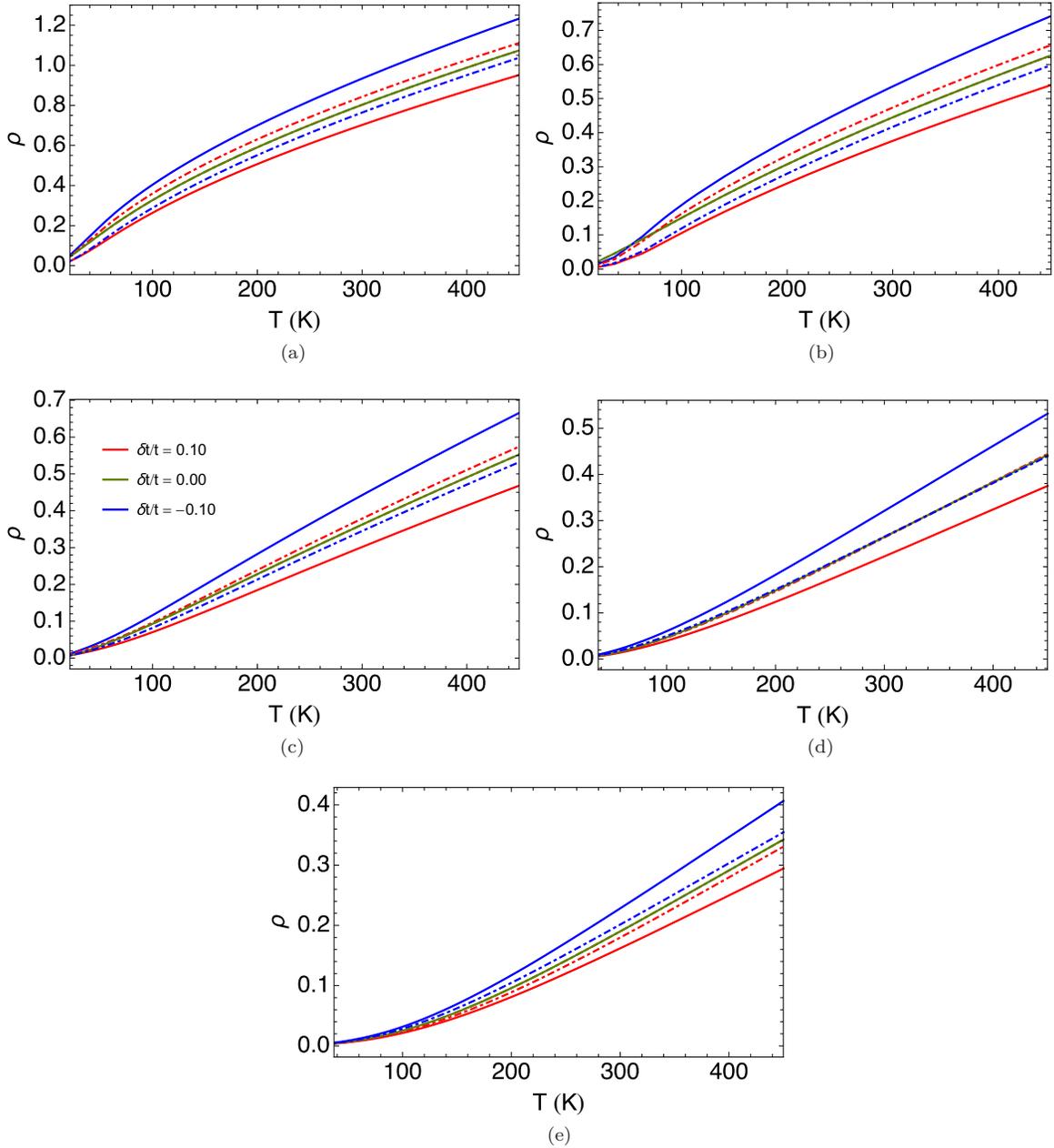


FIG. 1. The longitudinal and transverse resistivities,  $\rho_{xx}$  (solid) and  $\rho_{yy}$  (dashed), respectively, versus temperature at filling  $n = 0.85$ , for various  $t'$ , and at representative strains. While the green curves are for the unstrained case, the red curves correspond to a compressive strain  $\delta t/t = 0.10$  (i.e.  $\epsilon_{xx} \sim -.02$ ), and the blue curves correspond to a tensile strain  $\delta t/t = -0.10$  (i.e.  $\epsilon_{xx} \sim .02$ ), if we take  $\alpha \sim 5$ . (a)  $t'/t = -0.4$ ,  $n = 0.85$ ; (b)  $t'/t = -0.2$ ,  $n = 0.85$ ; (c)  $t'/t = 0.0$ ,  $n = 0.85$ ; (d)  $t'/t = 0.2$ ,  $n = 0.85$ ; (e)  $t'/t = 0.4$ ,  $n = 0.85$ . All figures share the legend. The resistivity in physical units can be found by  $\rho_{\text{physical}} = \rho \times \rho_0$ , where  $\rho_0 = c_0 h/e^2 \sim 1.17 \text{ m}\Omega \text{ cm}$ .

between parallel Cu-O planes [5,53]. In order to estimate the magnitude of the inelastic scattering, we can relate the dimensionless resistivity to  $\langle k_F \rangle \ell$  as follows  $\langle k_F \rangle \ell = 1/\rho_\alpha$  as argued in Refs. [15,54] for quasi-2D materials, where  $\langle k_F \rangle$  is an (angle averaged) effective Fermi momentum and  $\ell$  is the mean-free-path. Hence we expect  $\rho_\alpha/\rho_0 < 1$  in a good metal.

## 2. The raw resistivities

We first present the effects of hopping strain  $\delta t/t$  on resistivity. In Fig. 1, we study the anisotropy of the raw dimensionless resistivity over a broad range of temperatures

at the optimal density  $n = 0.85$ . Figure 1 displays the longitudinal resistivity  $\rho_{xx}$  (solid) and the transverse resistivity  $\rho_{yy}$  (dashed) for a compressive strain (red) and tensile strain (blue) in comparison to the unstrained tetragonal system (green). Here we used a representative magnitude of compressive strain  $\delta t/t = 0.10$  (i.e.,  $\epsilon_{xx} \sim -.02$ ). We observe that longitudinal resistivity under a compressive strain ( $\delta t/t > 0$ ) is reduced, and conversely, under a tensile strain ( $\delta t/t < 0$ ) it is enhanced across the displayed temperature range for all  $t'$ . The response for transverse resistivity is less than the longitudinal one in magnitude. An interesting new feature lies in the  $t'$  dependence; we note that magnitude and sign of the change

in transverse resistivity is controlled by  $t'$ , e.g., for  $t' = 0.2t$  the resistivity is almost unchanged for all strains.

These behaviors can be understood qualitatively in the following ways. First, let us look at the simplest case with  $t' = 0$  as in Fig. 1(c). When the system is compressed in the  $x$  axis, the hopping  $t_x$  rises according to Eq. (27) and so does the conductivity along the same direction, and vice versa. Hence, the longitudinal resistivity gets suppressed (enhanced) under compressive (tensile) strains. One can also consider isolating the strain-induced effects in Eqs. (37) and (38) from the band structure, contained in  $v_k^\alpha$ , and from the spectral function  $\rho_G$ , which accounts for the influence of the Gutzwiller correlations on resistivity. (Changes in the resistivity due to variation of the explicit lattice constants are small.) When we exert a compressive strain, this produces additive changes to the longitudinal resistivity due to in equal parts (1) changes in vertex and (2)  $T$ -dependent changes in spectral function, both arising from the enhancement of  $t_x$ . Whereas for the transverse resistivity the hopping parameter  $t_y$  is unchanged and hence changes to resistivity from the band structure become less important and as a result the transverse resistivity is dominated by strain-induced effects on the spectral function. For this reason, the transverse response to compressive strain is generally smaller in magnitude than the longitudinal response and likewise for a tensile strain both shown in panel (c). We also find that the transverse strain response has a different sign than the longitudinal one when there is no second neighbor hopping.

Now let us turn on  $t'$ . According to Eq. (27), the strain has a longitudinal-like effect, only smaller, on the magnitude of the second neighbor hopping. Turning on a positive  $t'$  strengthens longitudinal response and “counters” the transverse response from  $t_y$  hopping. Therefore we see that the longitudinal curves depart further from the unstrained one in panels (d) and (e), and it also explains why the transverse change almost vanishes for  $t' = 0.2$  in panel (d) and switches to the same sign as the longitudinal one for  $t' = 0.4$  in panel (e). Likewise turning on a negative  $t'$  weakens the longitudinal response and enhances the transverse response, so that the longitudinal response gets smaller in panels (a) and (b) while the transverse shifts more explicitly to the same side as  $t' = 0$ . Further analysis of these effects can be found in the Supplemental Material (SM) [55].

Next we discuss how strain affects the effective interaction and the characteristic temperature scale. We mainly use the longitudinal resistivity in this discussion because the longitudinal response is more explicit. In our recent work [6,7], a significant finding was the  $t'$  dependence of the curvature of the  $\rho$ - $T$  lines. We observe that this  $t'$ -dependent curvature persists under strain, i.e., the curvature changes from positive (concave up like  $+T^2$ ) to negative (convex up like  $-T^2$ ) as  $t'/t$  is varied upward.

Recall that strain is effectively a small change in the hopping parameter, so we ought to expect strain to change the  $t'$  dependence of the curvature only quantitatively but not qualitatively. Phenomenologically, varying  $t'$  signals a change in the effective Fermi temperature scale  $T_{FL}$  where for  $T < T_{FL}$  the system is in the Fermi liquid regime  $\rho \propto T^2$  and hence has a positive curvature. Moreover, as we decrease  $t'$  from positive to negative, the Fermi liquid temperature regime is compressed into a smaller temperature regime down

to temperatures where resistivity is usually hidden by the superconducting state. We want to focus on the crossover between Fermi liquid and strange metal which is covered by the following empirical relation

$$\rho \sim C \frac{T^2}{T_{FL} + T}. \quad (39)$$

Here  $C$  is a constant that defines the slope of linear regime and  $T_{FL}$  marks the crossover from the Fermi-liquid regime. For example when  $t' = -0.2t$  as found in typical hole-doped cuprates [56], we observe that a compressive strain extends the Fermi-liquid regime for the longitudinal resistivity, and flipping the strain reduces the Fermi-liquid regime. Qualitatively speaking, a compressive strain enhances the longitudinal hopping so that the effective interaction reduces relatively to the hopping. Likewise, a tensile strain increases the effective interaction in the unit of longitudinal hopping and suppresses the Fermi liquid temperature scale. Besides, we observe that a compressive strain suppresses the linear constant  $C$  while a tensile strain enhances it, as shown more obviously in Fig. 2. That can be verified in the experiment by measuring the slope of  $\rho$ - $T$  for a *strange metal* under strain.

### 3. Susceptibilities for anisotropic resistivities

It has been argued [1] that cuprates are candidates for an electron nematic phase, in which nematic order might coexist with high temperature superconductivity, that is, the electronic system breaks a discrete rotational symmetry while leaving the translational symmetry intact. Here the normalized resistivity response plays the role of the order parameter in the phase transition. Since it is possible to experimentally identify continuous phase transitions through observation of a diverging thermodynamic susceptibility across a phase boundary this makes the temperature profile of elastoresistance, i.e., normalized resistivity response with respect to an arbitrary strain, an interest observable to explore. For that reason, we shall examine linear response function for the longitudinal and transverse components of the elastoresistivity tensor constructed in terms of the hopping strain as:

$$\chi_{XX} \equiv -\left(\frac{\rho'_{xx} - \rho_{xx}}{\rho_{xx}}\right) / \left(\frac{\delta t}{t}\right), \quad (40)$$

$$\chi_{YY} \equiv -\left(\frac{\rho'_{yy} - \rho_{yy}}{\rho_{xx}}\right) / \left(\frac{\delta t}{t}\right), \quad (41)$$

respectively. The susceptibility as defined is positive if compression along the  $x$  axis leads to a reduction of the resistivity in the specified direction. We note the connection of these susceptibilities with the nematic susceptibility Eq. (1) on using Eq. (28) as

$$\chi_{nem} = \alpha \lim_{\epsilon_{xx} \rightarrow 0} \chi_{XX}. \quad (42)$$

We compute the susceptibility for small values of strain  $\delta t/t \gtrsim .05$ . However, even these values of strain pick up some nonlinear components of the response function. These are also of interest, and we comment on these below.

The linear response function for strain-resistivity curves is plotted as a function of temperature in Fig. 2 for the longitudinal and transverse components at optimal density

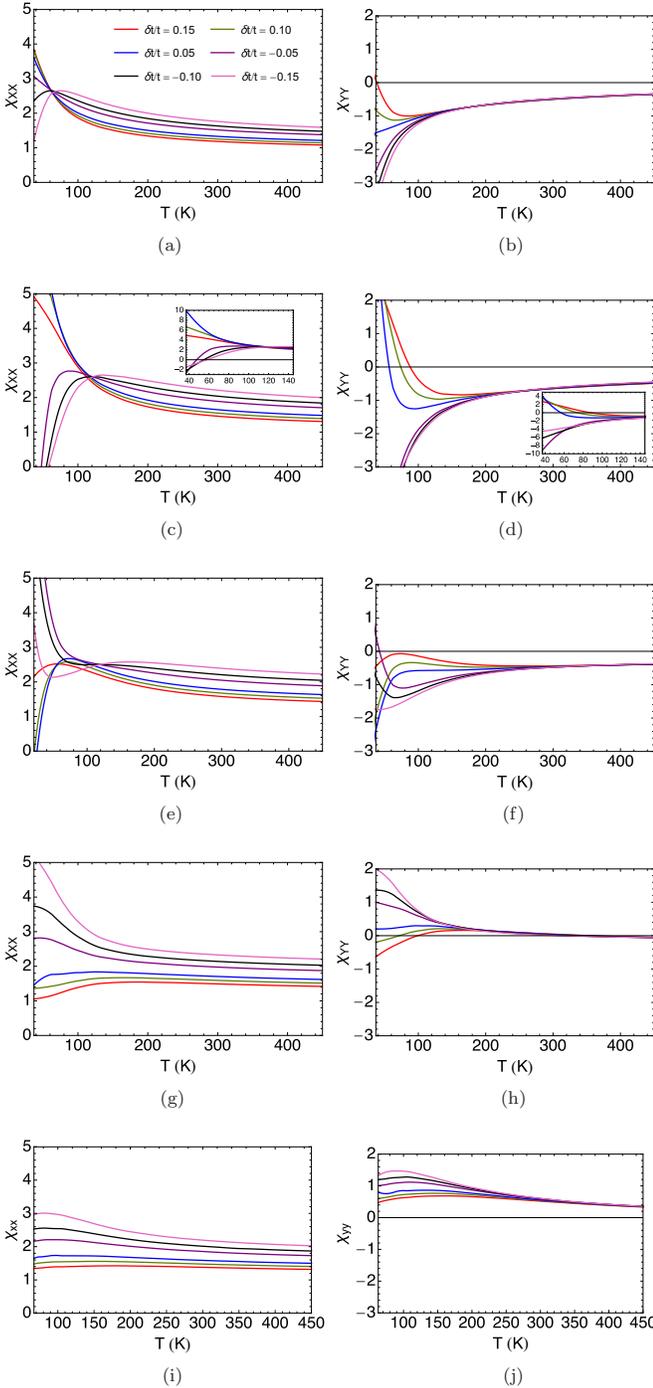


FIG. 2. The longitudinal [Eq. (40)] and transverse [Eq. (41)] strain-resistivity susceptibilities versus temperature at filling  $n = 0.85$ , for various  $t'$  and  $\delta t/t$ . (a), (c), (e), (g), (i)  $t'/t = -0.4, -0.2, 0.0, 0.2, 0.4$  longitudinal, respectively; (b), (d), (f), (h), (j)  $t'/t = -0.4, -0.2, 0.0, 0.2, 0.4$  transverse, respectively. All figures share a legend. For various  $\delta t/t$ , the susceptibilities  $\chi_{xx}$  for  $T \gtrsim 100$  approach each other in two sets, one for  $\delta t/t > 0$  and another slightly displaced set for  $\delta t/t < 0$ . They splay apart at low  $T$  thus displaying strong nonlinearity in the Fermi liquid regime. The susceptibilities  $\chi_{yy}$  approach a single set for  $T \gtrsim 100$  and splay apart for low  $T$  thus also displaying strong nonlinearity in the Fermi liquid regime.

$n = 0.85$  for various  $t'$  and  $\delta t/t$ . Note that since the resistivity vanishes as  $T \rightarrow 0$ , there is an enhancement of the normalized susceptibility at low  $T$ .

In Figs. 2(a), 2(c), 2(e), 2(g), and 2(i), we see that the linear response function for the longitudinal resistivity  $\chi_{xx}$  is mostly positive and shows nonlinear (in  $\delta t/t$ ) behavior at a fixed  $T$  (as can be identified by the separation of the strain curves) with respect to strain across the entire temperature range. This nonlinearity will be measured directly in Fig. 5 for  $t' = -0.2$ . The response function for  $T \gtrsim 100$  K is highly ordered in that varying the strain from positive (compressive) to negative (tensile) increases the strength of the response function for all  $t'$ . Conversely as we cool the system, we observe that strain dependence of the response function becomes increasingly nonlinear, i.e., showing a wider separation between strain curves, the forms of which are strongly  $t'$  dependent. Now if we vary  $t'$  to survey the range of cuprate materials, we find at low  $T$  for holelike ( $t' < 0$ ) materials a significant enhancement in and an inversion of the strain dependence that is absent in electronlike ( $t' > 0$ ) materials, though for both material types the strength of the response function remains approximately invariant at high  $T$ .

We next discuss the transverse linear response function  $\chi_{yy}$  shown in Figs. 2(b), 2(d), 2(f), 2(h), and 2(j). This response is potentially interesting since the affects of strain on the band structure are found to play a less significant role, hence the correlation effects dominate. We find that the features of transverse response function are different from that of the longitudinal response function mainly in two ways: (1) the  $\chi_{yy}$  collapses at high  $T$ , showing strong linearity with respect to the strain and (2) it changes sign from negative to positive as we vary  $t'/t$  across 0.2 from below, consistent with Fig. 1. Measurements confirming this linear behavior and sign change would be potentially interesting results.

#### 4. Resistivity with nonzero $J$

In this section we examine the role of exchange parameter  $J$  (nearest neighbor exchange energy) on resistivity and the susceptibilities, setting  $J = 0.17t$  which is the typical value for LSCO cuprate materials [48]. We take  $J = t^2/U$  where  $U$  is the on site energy of the Hubbard model and  $U$  does not vary with strain and hence  $\delta J = 2(\delta t/t)J$  [47]. Now, if we turn on the exchange parameter  $J$ , we find that at low temperatures the resistivity is reduced by the exchange energy and at high temperatures the resistivity is slightly enhanced as seen in Figs. 3(a) and 3(b). In Figs. 3(c) and 3(d) we see the longitudinal and transverse susceptibility with exchange interaction is further enhanced at low temperatures whereas at higher temperatures the response is unchanged. The  $J$  effects are magnified in the low- $T$  response since  $\rho \rightarrow 0$  as  $T \rightarrow 0$ . We can say the effects of  $J$  on the response are negligible at high  $T$ .

#### 5. Susceptibilities for $A_{1g}$ and $B_{1g}$ irreps

Experimentally, it is possible to identify the irrep to which the order parameter belongs by applying a strain with a

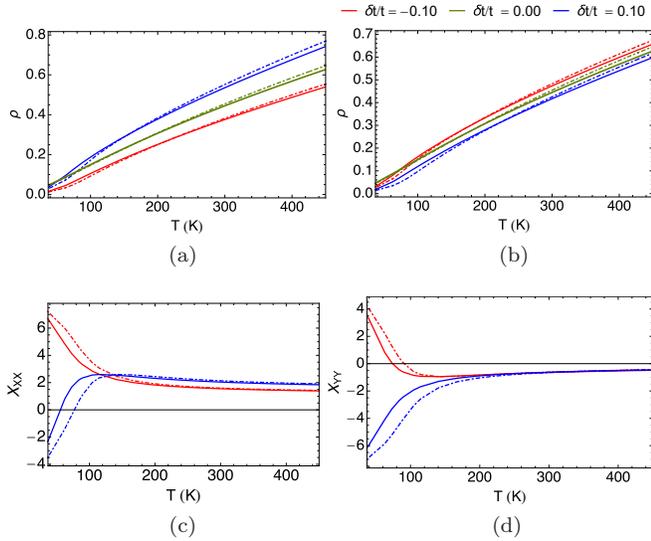


FIG. 3. (a), (b) The strain-induced resistivity at optimal filling  $n = 0.85$ ,  $t'/t = -0.2$  for three representative strain types with exchange parameter  $J = 0.0$  (solid) and  $J = 0.17$  (dashed). (c), (d) The strain-resistivity susceptibility for the same parameter set as above. All figures share a legend. We note that a nonvanishing  $J$  enhances somewhat the magnitude of the susceptibilities in the low temperature Fermi liquid regime. (a),(c) Longitudinal; (b),(d) Transverse.

particular irrep of strain and searching for a divergence in the temperature profile. In the case of uniaxial strain along the  $x$  axis the strain can be decomposed into the  $A_{1g}$  and  $B_{1g}$  irreps. In this section we examine the strain-resistivity linear response function for the  $A_{1g}$  and  $B_{1g}$  irreps defined in terms of the hopping strain as

$$\chi_{A_{1g}} \equiv -\left(\frac{\rho'_{xx} + \rho'_{yy} - 2\rho_{xx}}{2\rho_{xx}}\right) \bigg/ \left(\frac{\delta t}{t}\right) = \frac{\chi_{XX} + \chi_{YY}}{2},$$

$$\chi_{B_{1g}} \equiv -\left(\frac{\rho'_{xx} - \rho'_{yy}}{\rho_{xx}}\right) \bigg/ \left(\frac{\delta t}{t}\right) = \chi_{XX} - \chi_{YY}, \quad (43)$$

respectively.

In Fig. 4 we present the normalized strain-resistivity response functions at optimal density  $n = 0.85$  for various  $t'$  and  $\delta t/t$ . In this picture the  $A_{1g}$  and  $B_{1g}$  irreps play the roles of a center of mass coordinate and a relative coordinate, respectively. Together the two susceptibilities characterize the shift of in-plane resistivity as a result of an arbitrary in-plane strain. Recall that since the resistivity vanishes as  $T \rightarrow 0$ , the  $A_{1g}$  and  $B_{1g}$  susceptibilities are also enhanced at low  $T$ .

Examining the  $A_{1g}$  susceptibilities in Figs. 4(a), 4(c) 4(e), and 4(g), one important feature stands out, namely, that for  $T \gtrsim 100$  K the response function is positive for all  $t'$  and strains  $\delta t/t$ . This indicates that increasing a tensile (compressive) strain for  $T \gtrsim 100$  K enhances (suppresses) the average of the anisotropic resistivities.

We also see that at  $T \sim 100$  K with hole doping, i.e.,  $t' \leq 0$ , the normalized susceptibilities become independent of the strain, and hence the response is in the linear regime (signaled by the convergence of all strain curves). The nonlinear response at lower  $T$  is interesting and potentially observable in experiments with varying strain. On the other hand for

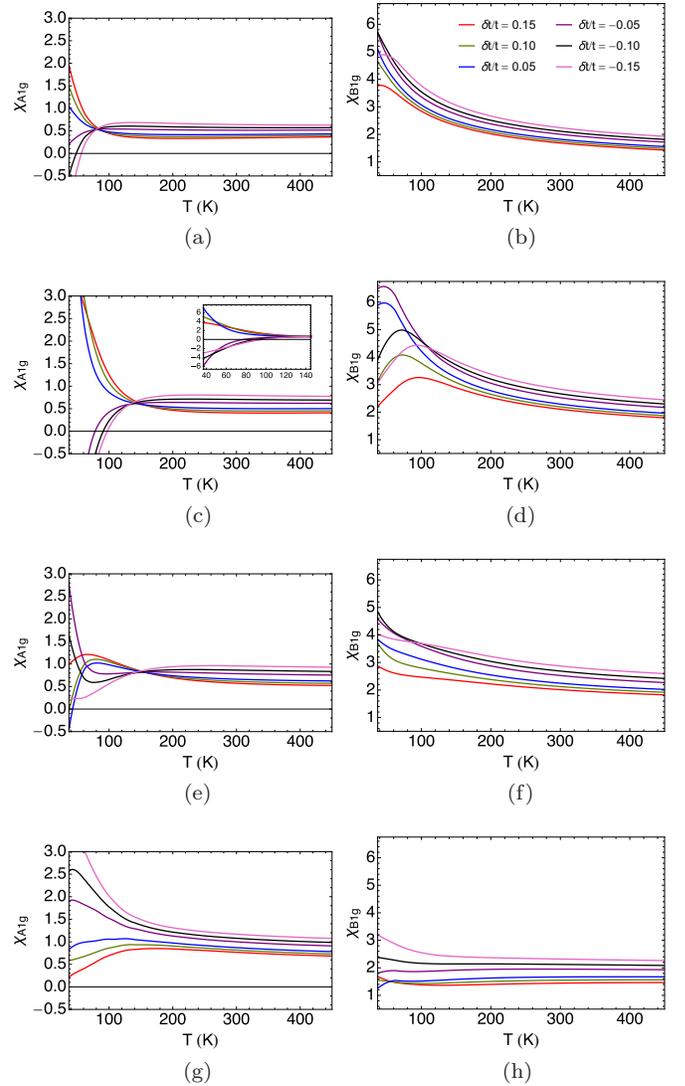


FIG. 4. The normalized strain-resistivity susceptibilities from Eq. (43) versus  $T$  for the  $A_{1g}$  and  $B_{1g}$  irreducible representations at filling  $n = 0.85$  at various  $t'$  and  $\delta t/t$ . (a), (c), (e), (g)  $t'/t = -0.4, -0.2, 0.0, 0.2$   $\chi_{A_{1g}}$ , respectively; (b), (d), (f), (h)  $t'/t = -0.4, -0.2, 0.0, 0.2$   $\chi_{B_{1g}}$ , respectively; All the figures share a legend. For various  $\delta t/t$ , and for  $T \gtrsim 100$  all the susceptibilities approach each other in two sets, one for  $\delta t/t > 0$  and another slightly displaced set for  $\delta t/t < 0$ . They splay apart at low  $T$  thus displaying strong nonlinearity in the Fermi liquid regime.

electron doping, i.e.,  $t' > 0$ , we see nonlinear behavior even at high  $T$ . Its origin is the extended Fermi-liquid regime which has a higher crossover temperature scale. Summarizing, we find that the early departure from Fermi liquid behavior into a strange metallic behavior in the hole doping favors an apparent linear response above 100 K due to a change in scale. Conversely we expect to see nonlinearity extending to much higher  $T$ 's in electron-doped systems.

From Fig. 4, we observe that the  $B_{1g}$  susceptibilities for  $T < 100$  K are strongly dependent on the value of  $t'$  of the system. We find in holelike materials ( $t' \lesssim 0.0$ ) there is a strong enhancement (the details of which depend on the  $\delta t/t$ ) in the susceptibility at low  $T$ . In contrast, this feature is absent

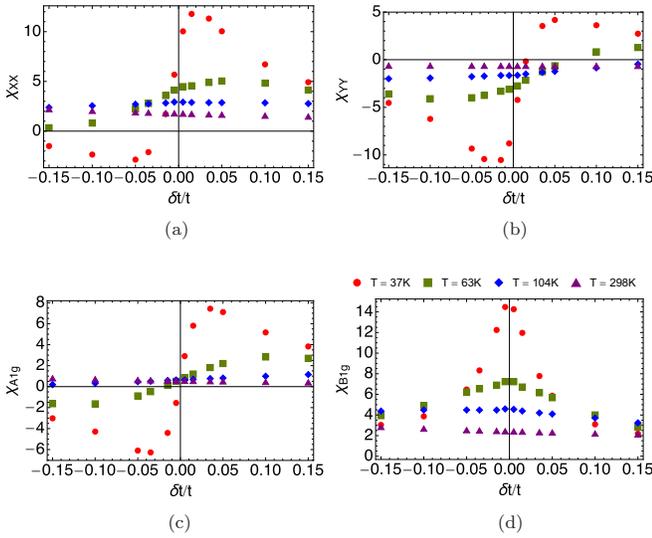


FIG. 5. The strain-resistivity susceptibilities for various symmetries as a function of strain  $\delta t/t$  at filling  $n = 0.85$  and  $t'/t = -0.2$ . (a)  $XX$ , longitudinal, Eq. (40); (b)  $YY$ , transverse, Eq. (41); (c)  $A_{1g}$  irrep, Eq. (43); (d)  $B_{1g}$  irrep, Eq. (43). The susceptibilities are relatively strain independent above 100 K but show strong nonlinearity at low  $T$ . It is noteworthy that for the lowest  $T$  shown, the susceptibilities  $\chi_{xx}$ ,  $\chi_{yy}$ ,  $\chi_{A_{1g}}$  change sign at or close to  $\delta t/t = 0$ . At higher  $T$  this change of sign is lost. The behavior of the nematic susceptibility  $\chi_{nem} = \lim_{\epsilon_{xx} \rightarrow 0} (\alpha \chi_{xx})$  at low  $T$  has thus the potential for a change of sign, depending on how we choose a sufficiently small  $|\epsilon_{xx}|$  or  $|\delta t/t|$  for the purpose of taking the limit  $\lim_{\epsilon_{xx} \rightarrow 0}$ .

in electronlike materials ( $t' > 0.0$ ) where there is weaker correlation, higher  $T_{FL}$ , and hence stronger quasiparticles.

Focusing on the strain dependence, we see that at high- $T$  the susceptibilities are relatively insensitive to  $t'$  and generally increases as we vary from a compressive to a tensile strain. There is also asymmetry in rate of change of susceptibilities between a compressive and tensile strain as  $|\delta t/t|$  is varied, i.e., the response function changes more rapidly for tensile than compressive strains. Therefore the degree of anisotropy is higher for tensile strain than compressive strains of equal magnitude.

Also, the  $B_{1g}$  curves under compressive strain ( $\delta t/t > 0$ ) are closer to each other than those under tensile strain for electron-doped systems, yet this spacing difference is less obvious in the hole-doped case. It means that a tensile response tends to show stronger nonlinearity, especially in electron-doped systems.

### 6. Susceptibilities versus strain

In Fig. 5, we display the strain-resistivity response functions versus hopping strain for various symmetries at  $t' = -0.2t$  and  $n = 0.85$  (which is roughly the parameter set for LSCO cuprate material [56] at optimal density) at four representative temperatures. Here we approximate the variance in the linear response function as follows

$$\chi(T) = c_0(T) + c_1(T)(\delta t/t) + c_2(T)(\delta t/t)^2 + \dots \quad (44)$$

In panels (a) and (b) we have longitudinal and transverse linear response functions, respectively, showing nonlinear

behavior at low temperature which becomes more linear (as indicated by horizontal line) as the system warms. This nonlinear behavior at low  $T$  can be understood as a result of the increasing importance of correlations as the system is cooled. Although the longitudinal and transverse response functions differ considerably in magnitude, the curves are approximately symmetric under inversion of the axes. In panels (a), (b), and (c) there is a wavelike oscillation which indicates the presence of higher order terms, e.g., the  $T = 37$  K curve in panel (a) appears to have  $(\delta t/t)^3$  term competing with a linear term. Another interesting result we find that as the system cools the  $B_{1g}$  response function appears diverge at  $\delta t/t = 0$  as  $T \rightarrow 0$  suggests that any deviation from the point group symmetry of the square lattice produces a finite resistivity response.

### B. Kinetic energy for an $x$ -axis strain

In this section we explore the kinetic energy anisotropy induced by strain along the  $x$  axis using ECFL theory. Since the anisotropic kinetic energy can be related to measurements of the optical conductivity using the f-sum rule on the  $t$ - $t'$ - $J$  model, this makes it another interesting observable to explore.

The total kinetic energy for a system under strain is computed as

$$K_{tot} = \left\langle \int_{-\infty}^{\infty} \rho_G(\vec{k}, \omega) \epsilon_{\vec{k}} d\omega \right\rangle_k. \quad (45)$$

This may be decomposed as follows:

$$K_{tot} = K_{xx} + K_{yy} + K_{xy}, \quad (46)$$

where the cross kinetic energy  $K_{xy}$  comes from the second neighbor interactions and is related to the dynamic Hall conductivity. Additional information on the total kinetic energy can be found in the SM [55]. The longitudinal, transverse, and cross kinetic energies are given by

$$K_{xx} = \left\langle \int_{-\infty}^{\infty} d\omega \rho_G(\vec{k}, \omega) \epsilon_{k_x} \right\rangle_k \quad (47)$$

$$K_{yy} = \left\langle \int_{-\infty}^{\infty} d\omega \rho_G(\vec{k}, \omega) \epsilon_{k_y} \right\rangle_k \quad (48)$$

$$K_{xy} = \left\langle \int_{-\infty}^{\infty} d\omega \rho_G(\vec{k}, \omega) \epsilon_{k_{xy}} \right\rangle_k \quad (49)$$

where

$$\epsilon_{k_x} = -2t_x \cos(k_x a) \quad (50)$$

$$\epsilon_{k_y} = -2t_y \cos(k_y b) \quad (51)$$

$$\epsilon_{k_{xy}} = -4t_d \cos(k_x a) \cos(k_y b). \quad (52)$$

In the  $t$ - $t'$ - $J$  model the anisotropic kinetic energies  $K_{\alpha}$ , where  $\alpha = xx, yy,$  and  $xy$ , are related to the optical conductivity  $\sigma_{\alpha}$  by the following sum rule

$$\Re \int_0^{\infty} \sigma_{\alpha}(\omega) d\omega = -K_{\alpha} e^2, \quad (53)$$

where  $e$  is the electrical charge.  $K_{\alpha} e^2$  sets the scale of the optical conductivity, i.e.,

$$-\frac{1}{K_{\alpha} e^2} \Re \int_0^{\infty} \sigma_{\alpha}(\omega) d\omega = 1. \quad (54)$$

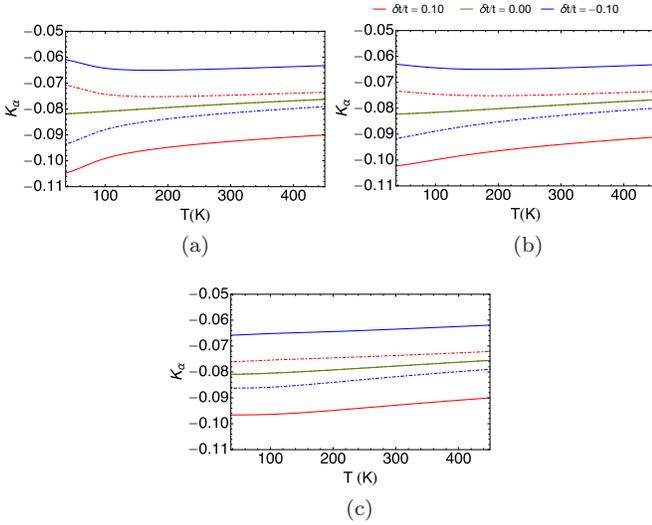


FIG. 6. Anisotropic kinetic energies  $K_\alpha$  versus  $T$  for  $K_{xx}$  (solid) and  $K_{yy}$  (dashed) at filling  $n = 0.85$ , for various  $t'$  and at three representative strains: compressive strain, no strain, and tensile strain. Note that  $K_{xx} = K_{yy}$  in the absence of strain. (a)  $t'/t = -0.2$ , (b)  $t'/t = 0.0$ , (c)  $t'/t = 0.2$ . All figures share a legend.

The optical conductivity in the DC limit  $\sigma_\alpha(0)$  relates to the DC resistivity as follows:  $\rho_\alpha(0) = 1/\sigma_\alpha(0)$ . For the anisotropic kinetic energy, we calculate and quote the following objects:

- (i)  $K'_{xx}$  is the strained version of longitudinal kinetic energy.
- (ii)  $K'_{yy}$  is the strained version of transverse kinetic energy.
- (iii) We call  $K_{xx}$  without a prime the tetragonal result. It is the same as  $K_{yy}$ .
- (iv) We present  $A_{1g}$  :

$$-\frac{K'_{xx} + K'_{yy} - 2K_{xx}}{2K_{xx}(\delta t/t)} \text{ vs } T.$$

- (v) We present  $B_{1g}$  :  $-(K'_{xx} - K'_{yy})/(K_{xx}\delta t/t)$  vs  $T$ .

### 1. Raw kinetic energies

From Eq. (47) we calculate the anisotropic kinetic energies  $K_\alpha$  as a function of temperature at optimal density for a representative range of cuprate materials  $t'$  and hopping strains  $\delta t/t$  as shown in Fig. 6. The main observation is that a compressive (tensile) strain suppresses (enhances) the longitudinal kinetic energy and vice versa for the transverse kinetic energy response with a smaller magnitude of variation. The variation in the longitudinal kinetic energy can be understood as a combination of changes in the band structure parameter  $t_x$  and correlations. On the other hand, the transverse kinetic energy is dominated by changes to the correlation function since the parameter  $t_y$  is unmodified by  $x$ -axis strain. There is little  $T$  dependence with the exception of a slight broadening of the range of the response at low  $T$  as the  $T_{FL}$  is reduced. The  $t'$  dependence is also weak because  $K_{xx}$  and  $K_{yy}$  do not explicitly depend on  $t'$  but through the spectral function.

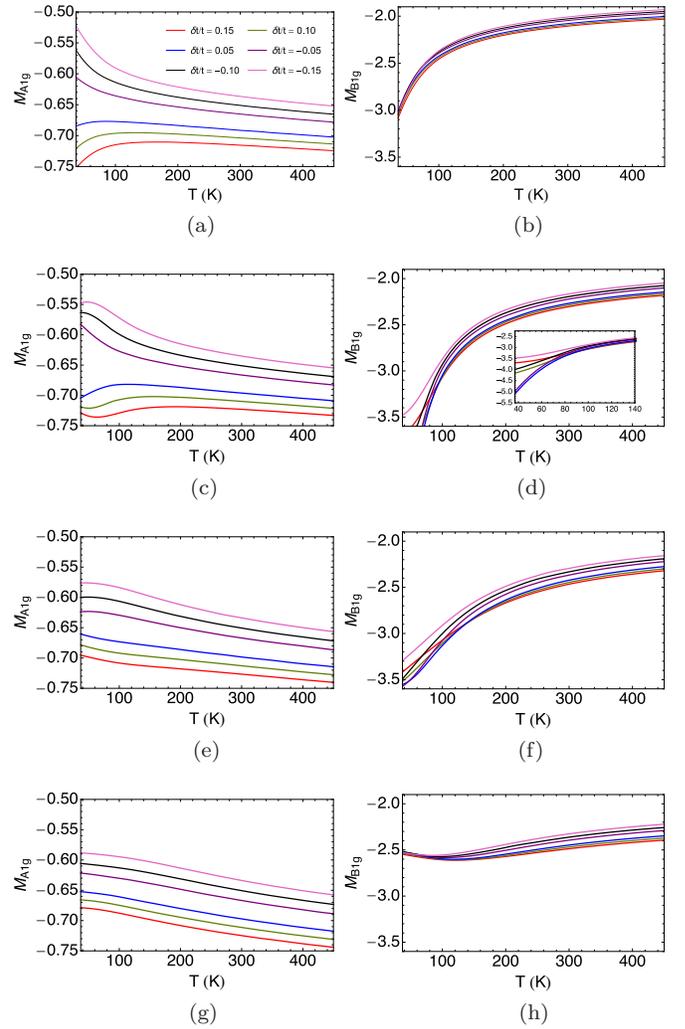


FIG. 7. The normalized strain-kinetic-energy susceptibilities vs  $T$  for the  $A_{1g}$  and  $B_{1g}$  irrep as defined in Eqs. (55) and (56) at filling  $n = 0.85$ , for various  $t'$  and  $\delta t/t$ . (a), (c), (e), (g)  $t'/t = -0.4, -0.2, 0.0, 0.2$   $M_{A_{1g}}$ , respectively; (b), (d), (f), (h)  $t'/t = -0.4, -0.2, 0.0, 0.2$   $M_{B_{1g}}$ , respectively; All figures share a legend.

### 2. Strain-kinetic-energy susceptibilities

In analogy with elastoresistance, we compute the so-called normalized strain-kinetic-energy response function, which measures the change in kinetic energy with respect to a strain. We shall focus on the normalized strain-kinetic-energy response functions for the  $A_{1g}$  and  $B_{1g}$  irrep since measurements of these symmetries are sensitive to a break in the fourfold rotation symmetry of a square lattice. Explicitly the response functions are defined in terms of hopping strain as

$$M_{A_{1g}} \equiv -\left(\frac{K'_{xx} + K'_{yy} - 2K_{xx}}{2K_{xx}}\right) \Big/ \left(\frac{\delta t}{t}\right), \quad (55)$$

$$M_{B_{1g}} \equiv -\left(\frac{K'_{xx} - K'_{yy}}{K_{xx}}\right) \Big/ \left(\frac{\delta t}{t}\right), \quad (56)$$

where the sign is imposed so that susceptibility defined in terms of hopping strain matches its counterpart defined in terms of conventional strain. Figure 7 displays the normalized strain-kinetic-energy susceptibilities as a function of

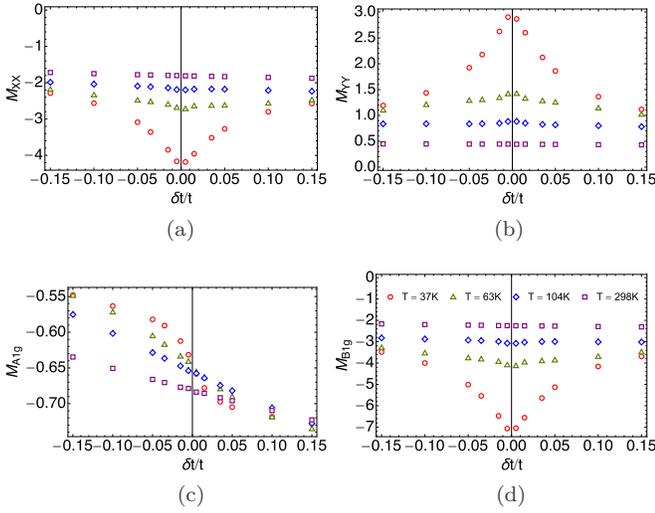


FIG. 8. The strain-kinetic-energy susceptibilities versus  $\delta t/t$  at filling  $n = 0.85$  and  $t'/t = -0.2$  at four representative temperatures. All figures share a legend. (a)  $XX$ , longitudinal, Eq. (57); (b)  $YY$ , transverse, Eq. (58); (c)  $A_{1g}$  irrep, Eq. (55); (d)  $B_{1g}$  irrep, Eq. (56).

temperature for the  $A_{1g}$  and  $B_{1g}$  irrep at optimal density for various  $t'$  and  $\delta t/t$ . The  $A_{1g}$  irrep susceptibility signals a change in the sum of anisotropic kinetic energies  $K_{xx} + K_{yy}$  with respect to the hopping change. The  $A_{1g}$  susceptibility shows that tuning the strain from tensile to compressive increases rather uniformly the magnitude of the anisotropic kinetic energy, i.e., strain enhances the overall optical weight from Eq. (53). Analysis of the longitudinal and transverse components are in the SM [55].

The  $B_{1g}$  susceptibility is characterized as the difference in the kinetic energies  $K_{xx} - K_{yy}$  with respect to the hopping change. Thus a nonzero value for the  $B_{1g}$  irrep signals an anisotropy between the two directions. We observe that the response function for the  $B_{1g}$  irrep is strongly  $t'$  dependent. For  $t' = -0.4$ , the response functions is nearly linear at all temperatures. We point out a curious feature for the  $t' = -0.2$  curve where at high  $T$  the system is linear whereas at low  $T$  the system is nonlinear, but it is nearly symmetric with respect to a compressive or tensile strain of similar magnitude. At high  $T$  for all  $t'$  the system is monotonic with respect to strain. For  $t' \geq 0$  there is little variation in the response function across the temperature range and it appears to become increasingly nonlinear as the system is warmed due to the reduction in the scale of variation.

### 3. Strain-kinetic-energy susceptibility versus strain

We now present strain-kinetic-energy susceptibility as a function of strain at optimal density ( $n = 0.85$ ) and  $t' = -0.2t$  for  $XX$ ,  $YY$ ,  $A_{1g}$ ,  $B_{1g}$  symmetries at various  $T$  (see Fig. 8), where we define the longitudinal and transverse response functions as

$$M_{XX} \equiv -\left(\frac{K'_{xx} - K_{xx}}{K_{xx}}\right) / \left(\frac{\delta t}{t}\right), \quad (57)$$

$$M_{YY} \equiv -\left(\frac{K'_{yy} - K_{yy}}{K_{xx}}\right) / \left(\frac{\delta t}{t}\right), \quad (58)$$

respectively. Like the resistivity case,  $M_{A_{1g}} = 0.5 \times (M_{XX} + M_{YY})$  and  $M_{B_{1g}} = M_{XX} - M_{YY}$ .

We find that at low temperatures, decreasing the magnitude of the strain increases the strength of the longitudinal response function in panel (a) and the response function is symmetric with respect to both strain types. The transverse response function in panel (b) shows a similar symmetry between tensile and compressive strains with a flipped sign. Therefore we find that a compressive strain for the  $A_{1g}$  response function [panel (c)] depletes the in-plane optical weight and vice versa for a tensile strain. The  $B_{1g}$  response function is similar to the longitudinal and transverse only more intensive and it signals an enhanced (suppressed) anisotropy between in-plane kinetic energies for a compressive (tensile) strain. In all cases the response function is approximately linear at room temperature (297 K) and becomes increasingly nonlinear as the system cools. In comparing panels (b)–(d) we see strong similarity between their respective responses. This is expected since strain merely shifts kinetic energy versus temperatures curves up and down. Also, it appears to diverge for small strains as  $T \rightarrow 0$ .

### C. The local density of states for an $x$ -axis strain

The local density of states (LDOS) is also very interesting since it can be measured using STM probes. We present results on how the LDOS changes with strain and the related susceptibilities. We argue that if experiments are done on resistivity variation as well as LDOS variation with strain, we can bypass the need for measuring strain accurately and of estimating the parameter  $\alpha$  in Eq. (19). The LDOS is calculated as  $\rho_{G\text{loc}}(\omega) = \langle \rho_G(\vec{k}, \omega) \rangle_k$  where averaging over the Brillouin zone is implied, and  $G \rightarrow g$  is the free Green's function (i.e., band structure) which gives the bare LDOS and the ECFL Green's function  $G \rightarrow \mathcal{G}$  gives the LDOS for the  $t$ - $t'$ - $J$  model.

In this section we calculate the normalized change in the local density of states and quote the following:

- (i)  $\rho'_{g\text{loc}}(\omega) = \langle \rho_g(\vec{k}, \omega) \rangle_k$  is the bare LDOS for a strain along the  $x$  axis.
- (ii)  $\rho'_{\mathcal{G}\text{loc}}(\omega) = \langle \rho_{\mathcal{G}}(\vec{k}, \omega) \rangle_k$  is the interacting LDOS for an  $x$  axis strain.
- (iii)  $\rho_{g\text{loc}}$  without a prime refers to the tetragonal result and similarly for  $\rho_{\mathcal{G}\text{loc}}$ .
- (iv) We present  $(\rho'_{g\text{loc}} - \rho_{g\text{loc}}) / (\rho_{g\text{loc}} \delta t/t)$  vs  $\omega$ .
- (v) We present  $(\rho'_{\mathcal{G}\text{loc}} - \rho_{\mathcal{G}\text{loc}}) / (\rho_{\mathcal{G}\text{loc}} \delta t/t)$  vs  $\omega$ .

#### 1. $T$ variation

In Fig. 9, we display the LDOS at optimal density ( $n = 0.85$ ) and  $t' = -0.2$  for various temperatures at three characteristic strains: a compressive strain (thick dashed), unstrained (solid), and tensile strain (thin dashed). We compare the LDOS for a noninteracting system [panel (a)] to a system with electron-electron interaction [panels (b)–(d)]. We find over large temperature scales that curves for the bare LDOS shifts to left along the  $\omega$  spectrum upon warming, leaving the line shape intact. In contrast with the bare LDOS, we see that warming the LDOS for the interacting system in panel (c) completely smooths and broadens the LDOS peaks for

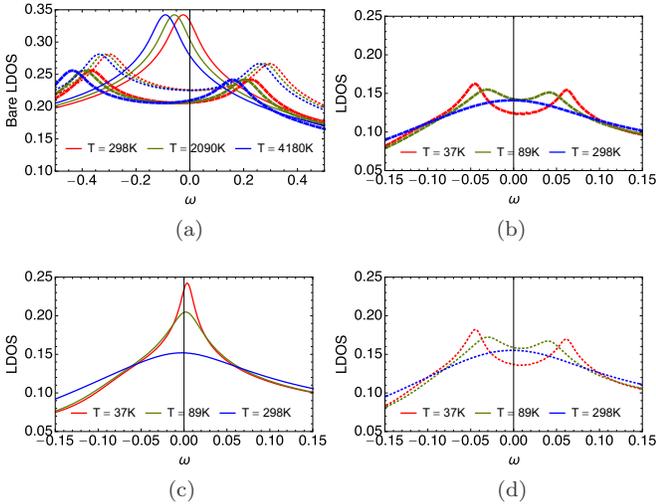


FIG. 9. The local density of states for (a) the noninteracting (band-structure) and (b)–(d) interacting system ( $t$ - $J$  model) at optimal filling ( $n = 0.85$ ),  $t' = -0.2t$ , for various temperatures and at three characteristic strains:  $\delta t/t = 0.15, 0.00, -0.15$  (thick dashed, solid, thin dashed), respectively.

all strains and slightly shifting them left. This is consistent with previous findings that interactions significantly lower the Fermi liquid temperature  $T_{FL}$  [6]. We note that strain inverts the LDOS peak at low  $T$ , leaving behind a pair of cusps at a reduced height. This is an artifact of the anisotropy of hopping parameters since it also shows up in the bare case.

### 2. $J$ variation

In Fig. 10, we turn on the exchange parameter  $J$  and examine the LDOS. We also find it useful to examine the self-energy of the system. We define the Dyson self-energy  $\Sigma$  as

$$\mathcal{G}(k) = \frac{1}{\omega + \mu - \epsilon_{\vec{k}} - \Sigma(k)}. \quad (59)$$

Here we use the shorthand  $\Sigma = \Sigma' + i\Sigma''$  to denote the real and imaginary parts of a complex function. In terms of the spectral function, self-energy imaginary part is

$$\Sigma''(k) = \frac{-\pi \rho_{\mathcal{G}}(k)}{[\mathcal{G}'(k)]^2 + [\pi \rho_{\mathcal{G}}(k)]^2}, \quad (60)$$

where  $\Re \mathcal{G} = \mathcal{G}'$  is found by taking the Hilbert transform of  $\Im m \mathcal{G} = \mathcal{G}''$  and we can find  $\Sigma'$  in the same manner. In Figs. 10(c)–10(f) we display the Dyson self-energy averaged over the Brillouin zone  $\Sigma_{loc}(\omega) = \langle \Sigma(\vec{k}, \omega) \rangle_k$ .

Turning on the exchange parameter in Fig. 10(a) has a small but visible effect on LDOS at low  $\omega$  when compared to Fig. 6(b) of the SM [55]. For panel (c) we see that varying strain from compressive ( $\delta t/t > 0$ ) to tensile ( $\delta t/t < 0$ ) shifts the average quasiparticle states to higher energies and panel (e) shows that increasing the intensity of the strain produces quasiparticles with higher and sharper peaks. In panels (b), (d), and (f) we see that varying  $J$  from ferromagnetic (negative) to antiferromagnetic (positive) splits a single LDOS

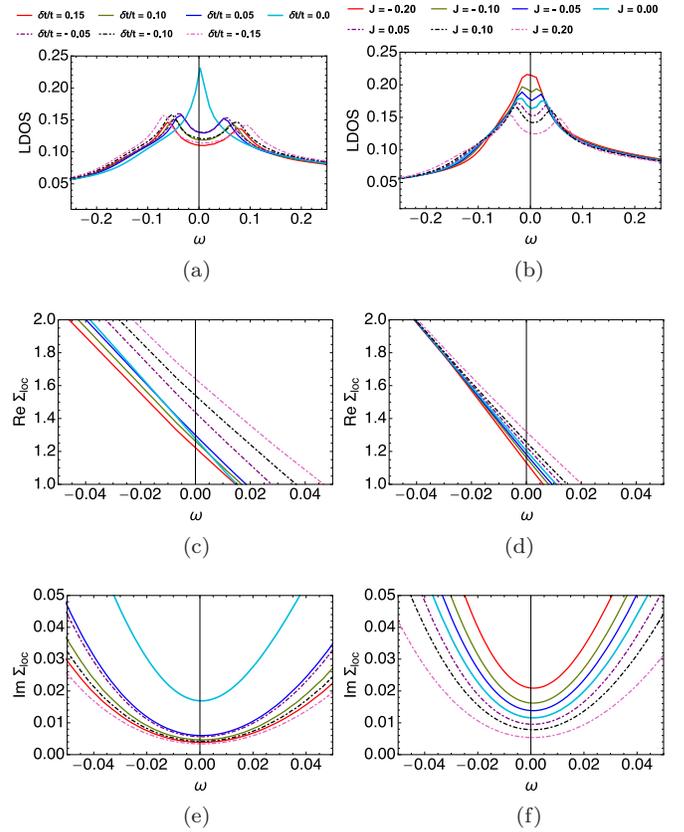


FIG. 10. (a), (b) The LDOS; (b), (c) real part of local Dyson self-energy; (e), (f) the imaginary part of local Dyson self-energy for parameter set  $n = 0.85$ ,  $T = 37\text{K}$ ,  $t' = -0.2t$  with varying  $\delta t/t$  (LHS) and varying  $J$  (RHS). Figures (a), (c), (e)  $J = 0.17t$  and (b), (d), (f)  $\delta t/t = 0.05$  share a legend, respectively.

peak into two, shifts the average quasiparticle states to higher energies, and narrows the quasiparticle peaks.

### 3. $t'$ variation

In Fig. 11, we examine the LDOS from a different vantage point by looking at the  $t'$  dependence for a system at optimal density ( $n = 0.85$ ), for a compressive strain of  $\delta t/t = 0.15$ , at various  $t'/t$ . In panel (c), we show the bare LDOS at room temperature as a reference for the interacting system. In panels (a) and (b), we display the interacting system at  $T = 37\text{K}$  and  $T = 298\text{K}$ , respectively. Upon inspection it appears the primary role that  $t'$  plays is to shift the energy band along the spectrum. As previously noted, warming the interacting system to room temperature smooths and broadens the characteristic LDOS peaks for all strain types and at all  $t'$  while leaving their position in the spectrum fixed. Even though the relative position of different  $t'$  curves remain unchanged as the interactions are turned on, we note that strong correlations renormalize the bare band into a smaller energy region. Comparing panels (a) and (b) fixed at  $t' = -0.4, -0.2$ , we observe that LDOS peak height is more strongly suppressed at a lower  $t'$ . This is consistent with previous studies [7] on the unstrained interacting system, and it indicates that a smaller  $t'$  has a lower Fermi-liquid temperature scale  $J$  and hence it is less

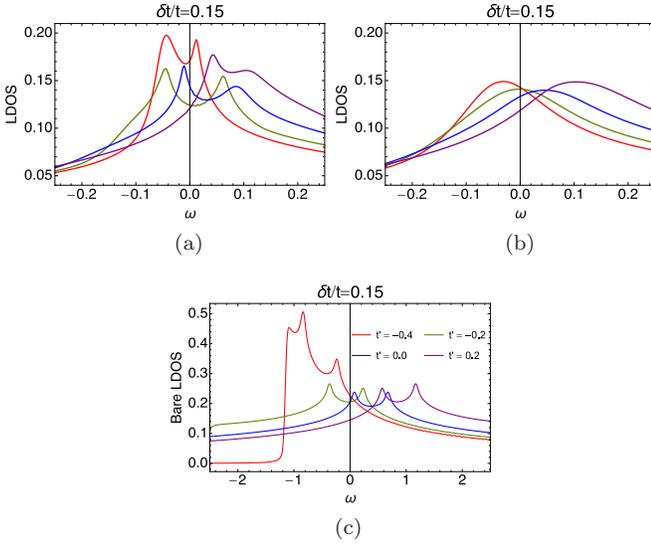


FIG. 11. The local density of states versus frequency at optimal filling ( $n = 0.85$ ), for a compressive strain ( $\delta t/t = 0.15$ ) at various  $t'$ . (a), (b) The interacting system ( $t$ - $t'$ - $J$  model) at  $T = 37$  K and  $T = 298$  K, respectively. (c) The noninteracting (band-structure) system at  $T = 298$  K. All figures share the same legend.

robust to heating. For further analysis of the strain dependence see the SM [55].

#### 4. Susceptibilities

Next, we examine the normalized response function of LDOS of the noninteracting and interacting system, respectively, defined as

$$N_g \equiv \left( \frac{\rho'_{g\text{loc}} - \rho_{g\text{loc}}}{\rho_{g\text{loc}}} \right) / \left( \frac{\delta t}{t} \right), \quad (61)$$

$$N_G \equiv \left( \frac{\rho'_{G\text{loc}} - \rho_{G\text{loc}}}{\rho_{G\text{loc}}} \right) / \left( \frac{\delta t}{t} \right). \quad (62)$$

In Fig. 12, we plot the LDOS susceptibility for a noninteracting and interacting system at room temperature at optimal density for various  $t'$ . We observe that the response function is linear at all frequencies except near the LDOS peak and, although not shown in the figure, at the band edges. Regardless of the presence of interaction, we note that the susceptibility is enhanced by tensile strain near the LDOS peak and reduced by a compressive strain.

#### 5. Susceptibility versus strain

Changing up the perspective, we explore the LDOS susceptibility now as a function of strain, at four representative frequencies as seen in Fig. 13. We can approximate the variance in the linear response function in Eqs. (61) and (62) as

$$N(T) = c_0(T) + c_1(T)(\delta t/t) + c_2(T)(\delta t/t)^2 + \dots, \quad (63)$$

where  $c_0$  is the linear term,  $c_1$  is the second order term, and  $c_2$  is the third order term of the response. We see that for the bare LDOS, Fig. 13(a), at  $\omega = 0.45$  the system is nearly linear with  $c_0 \approx -0.5$  and  $c_1 \approx 3$ . The other presented frequencies

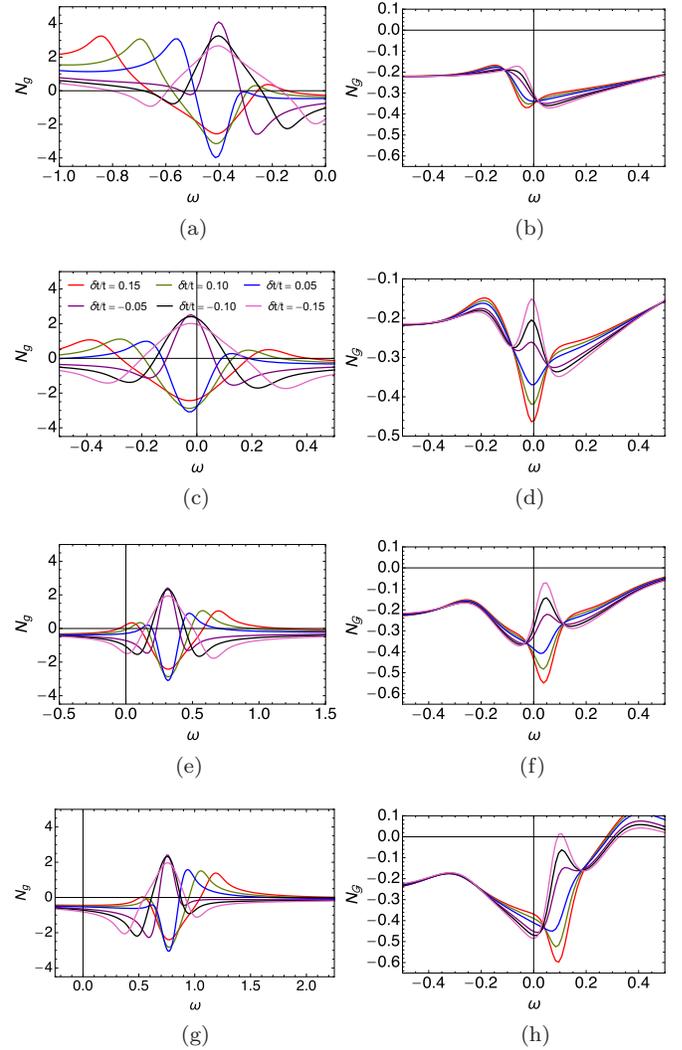


FIG. 12. The LDOS susceptibility versus frequency at optimal filling  $n = 0.85$ , at room temperature ( $T = 297$  K), for various  $t'$  and  $\delta t/t$ . The noninteracting (band-structure) system (LHS) and interacting system ( $t$ - $t'$ - $J$  model) (RHS) from Eq. (61) and Eq. (62), respectively. (a), (c), (e), (g)  $t'/t = -0.4, -0.2, 0.0, 0.2$   $N_g$ , respectively; (b), (d), (f), (h)  $t'/t = -0.4, -0.2, 0.0, 0.2$   $N_G$ , respectively; All the figures share a legend.

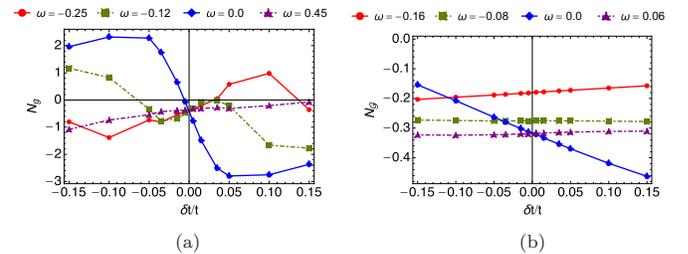


FIG. 13. The LDOS susceptibility versus strain at optimal filling  $n = 0.85$ , at room temperature ( $T = 298$  K), for  $t' = -0.2t$ , at a few representative frequencies  $\omega$  in units of  $t$ . (a) The noninteracting system (band-structure) in Eq. (61). (b) The interacting system ( $t$ - $t'$ - $J$  model) in Eq. (62).

appear to be nonlinear with significant second and third order terms. The LDOS susceptibility for the interacting system [panel (b)] appears to be nearly linear everywhere except at the location of the LDOS peak ( $\omega = 0$ ) which has a strong quadratic response, suggesting that at temperatures relevant to experiments nonlinear behavior is only observable at energies near the Fermi surface. Note that the second order scheme used here is good for low energies but somewhat less reliable at high energies,  $|\omega| \gtrsim k_B T_{FL}$ .

#### IV. SUMMARY AND COMMENTS

##### A. Summary

In this work, we have applied the ECFL theory to study the effect of small strain on the resistivity, kinetic energy, LDOS, and their associated susceptibilities in the  $t$ - $t'$ - $J$  model Eq. (2) with various  $t'$  at  $n = 0.85$ . These results are expected to be relevant to cuprate superconductors, especially single layered materials, where the calculated unstrained resistivities are in good accord with the experimental data [5].

Based on comparisons carried out earlier, the second order scheme of ECFL used here is expected to be reasonable in the density range  $0.85 \gtrsim n \gtrsim 0.80$  spanning an experimentally accessible range in cuprates. With improvements in the theoretical scheme, we expect that while resistivities themselves might not change too much, the related susceptibilities [involving division by the small resistivity as in Eq. (1)] could be more sensitive.

Our results exhibit in considerable detail the theoretically expected strain dependence of resistivity and LDOS as well as optical weight. The derived susceptibilities depend sensitively on the magnitude and sign of  $t'$ . Our results in Figs. 2 and 3 illustrate the quantitative change of the strain dependence due to varying the magnitude and sign of  $t'$ . We should stress that the absolute scale of  $t$  is important in determining the  $T$  dependence. For illustration we have used  $t = 0.45$  eV in the present paper while the more fine-tuned estimates in Ref. [5] suggest a material dependent and somewhat larger value of  $t \sim 1$  eV in most cases.

Our results can be converted to actual strains as in Eq. (42), with  $\alpha$  in the range  $\alpha \in \{2, 5\}$ . If data is available one may ideally eliminate  $\alpha$  by measuring the strain dependence of the LDOS or the optical conductivity sum rule.

##### B. Comments on experiments

The results found in Fig. 2 yield a magnitude of the nematic susceptibility  $\chi_{nem} \sim (1-5)\alpha$  for cuprates. Using the expected range of  $\alpha \in \{2, 5\}$ , we find  $\chi_{nem} \sim 2-25$ . On the other hand, iron based pnictide superconductors appear to have a considerably larger value for  $\chi_{nem}$ , e.g., in Fig. 3 of [1] the range  $|\chi_{nem}| \lesssim 650$  is reported, thus an order of magnitude greater than our theoretical estimate for cuprates. While fluctuations may drive the magnitude of nematicity further upwards, especially at some densities and temperatures, it appears that the baseline magnitude of this object is itself much larger than expected in cuprates. For example in the four featureless curves of Fig. 3 of [1] we see that  $|\chi_{nem}| \sim 200$ .

This magnitude indicates that the downfolding of the many bands of the pnictides to an effective single (or few) band model must yield hopping parameters that are much more sensitive to strain than in cuprates. The different type of quantum overlap of relevant atomic orbitals from those in cuprates are presumably the origin of this difference. We also note that the sharp peaks in  $|\chi_{nem}|$  on varying  $T$ , as reported in Refs. [1,3], are missing in our results. Instead we have a monotonic increase of  $|\chi_{nem}|$  and related susceptibilities as we cool the system, as seen in Fig. 2 and Fig. 4. This increase is largely due to the decrease of the (unstrained) resistivity with lowering  $T$  in the Fermi liquid regime.

The sign of  $\chi_{nem}$  presents a more subtle problem. In iron pnictides it is known to be sensitive to effective mass anisotropy. In fact it changes sign with doping in certain hole-doped iron pnictides [57]. Our single band model lacks such an anisotropy and is therefore not appropriate to describe the elastoresistivity of iron pnictide materials.

Recently, we came across the measurement of the elastoresistivity nematic susceptibility in [58] on the two layer cuprate Bi2212. In this experiment, the magnitude of the nematic susceptibility is found to be in the range  $|\chi_{nem}| \in \{2.5, 5\}$ . This range is consistent with our theoretical estimate. It is also smaller than the nematic susceptibility in iron pnictides by about two orders of magnitude.

The sign of the nematic susceptibility  $\chi_{nem}$  [Eq. (1)] reported in [58] implies that the resistivity *increases* in the direction of compression. This result has the opposite sign to our theoretical result as seen in Fig. 2. There we see that the theoretical resistivity *decreases* in the direction of compression, although it does *increase* in the transverse direction. It is possible that the two layer nature of Bi2212 might be responsible for this opposite sign. Also as noted in Fig. 5, the behavior of the nematic susceptibility  $\chi_{nem} = \lim_{\epsilon_{xx} \rightarrow 0} (\alpha \chi_{xx})$  at sufficiently low  $T$  has the potential for a change of sign, depending on how we choose a sufficiently small  $|\epsilon_{xx}|$  or  $|\delta t/t|$  for the purpose of taking the limit  $\lim_{\epsilon_{xx} \rightarrow 0}$ . On the experimental side, a more detailed  $T$  variation and examining the various susceptibilities listed in Fig. 5 should yield a more complete picture.

The results found here should also motivate further studies of the strain variation of the three-dimensional electronic bands of cuprates, towards computing strain variation of the resulting two-dimensional bands found from projecting to a  $t$ - $t'$ - $J$  model. These would test the simple assumptions made here between strain and hopping parameters of a reduced two-dimensional model as presented in Eqs. (22), (27), (28), and (29). It is also possible that under certain situations, the sign of  $\alpha$  can even be changed, as a naive interpretation of the experiments of Ref. [58] suggests.

We believe that it is important to study a more extensive set of samples including single layer cuprates at various compositions in the future. It would also be useful to study the variations of resistivity along different axes, parallel *and* transverse to the strain axis and extend the studies to various  $T$ 's. This type of measurements would enable the construction of the symmetry adapted susceptibilities as in Fig. 5, which provide a greater insight into the results. It would also be of considerable interest to measure the variations of the LDOS and optical weight with strain, as emphasized above.

## ACKNOWLEDGMENTS

We thank Prof. I. R. Fisher and Prof. S. A. Kivelson for stimulating our interest in this problem and for helpful discussions. We thank Prof. T. Shibauchi for a helpful explanation of features of [58]. The work at UCSC was supported by the US Department of Energy (DOE), Of-

fice of Science, Basic Energy Sciences (BES), under Award No. DE-FG02-06ER46319. The computation was done on the Comet system, a resource operated by the San Diego Supercomputer Center at UC San Diego, available through the XSEDE program [59] (Allocation No. TG-DMR170044), which is supported by National Science Foundation Grant No. ACI-1053575.

- [1] J. H. Chu, H. Kuo, J. G. Analytis, and I. R. Fisher, *Science* **337**, 710 (2012).  
 [2] M. C. Shapiro, P. Hlobil, A. T. Hristov, A. V. Maharaj, and I. R. Fisher, *Phys. Rev. B* **92**, 235147 (2015).  
 [3] H. Kuo, J. H. Chu, J. C. Palmstrom, S. A. Kivelson, and I. R. Fisher, *Science* **352**, 958 (2016).  
 [4] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).  
 [5] B. S. Shastry and P. Mai, *Phys. Rev. B* **101**, 115121 (2020).  
 [6] B. S. Shastry and P. Mai, *New J. Phys.* **20**, 013027 (2018).  
 [7] P. Mai and B. S. Shastry, *Phys. Rev. B* **98**, 205106 (2018).  
 [8] In the  $t$ - $t'$ - $J$  model the first and last term of Eq. (2), referred to as the hopping  $H_t$  and exchange  $H_J$  term, respectively, explicitly expanded as sum over neighbors is

$$H_t + \mu\hat{N} = -t \sum_{\langle ij \rangle \sigma} \tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma} - t' \sum_{\langle\langle ij \rangle\rangle \sigma} \tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma},$$

$$H_J = \frac{J}{2} \sum_{\langle ij \rangle} \left( \tilde{S}_i \cdot \tilde{S}_j - \frac{1}{4} n_i n_j \right),$$

where  $\langle ij \rangle$  is the sum over nearest neighbors, and  $\langle\langle ij \rangle\rangle$  is the sum over second nearest neighbors. In terms of the Hubbard operators—see Eq. (4)—this becomes

$$H_t + \mu \sum_{i\sigma} X_i^{\sigma\sigma} = -t \sum_{\langle ij \rangle \sigma} X_i^{\sigma 0} X_j^{0\sigma} - t' \sum_{\langle\langle ij \rangle\rangle \sigma} X_i^{\sigma 0} X_j^{0\sigma},$$

$$H_J = \frac{J}{2} \sum_{\langle ij \rangle \sigma} X_i^{\sigma\sigma} + \frac{J}{4} \sum_{\langle\langle ij \rangle\rangle \sigma_1 \sigma_2} \{ X_i^{\sigma_1 \sigma_2} X_j^{\sigma_2 \sigma_1} - X_i^{\sigma_1 \sigma_1} X_j^{\sigma_2 \sigma_2} \}.$$

- [9] F. C. Zhang and T. M. Rice, *Phys. Rev. B* **41**, 7243 (1990).  
 [10] B. S. Shastry, *Phys. Rev. Lett.* **63**, 1288 (1989).  
 [11] R. S. Markiewicz, S. Sahrakorpi, M. Lindroos, Hsin Lin, and A. Bansil, *Phys. Rev. B* **72**, 054519 (2005).  
 [12] V. Heine, *Phys. Rev.* **153**, 673 (1967).  
 [13] Y. Onose, Y. Taguchi, K. Ishizaka, and Y. Tokura, *Phys. Rev. B* **69**, 024504 (2004).  
 [14] R. L. Greene, P. R. Mandal, N. R. Poniatowski, and T. Sarkar, *Ann. Rev. Condens. Matter Phys.* **11**, 213 (2020).  
 [15] Y. Ando, S. Komiya, K. Segawa, S. Ono, and Y. Kurita, *Phys. Rev. Lett.* **93**, 267001 (2004).  
 [16] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).  
 [17] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, *Rev. Mod. Phys.* **78**, 865 (2006).  
 [18] W. Xu, K. Haule, and G. Kotliar, *Phys. Rev. Lett.* **111**, 036401 (2013).  
 [19] X. Y. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).

- [20] K. Bouadim, N. Paris, F. Hebert, G. G. Batrouni, and R. T. Scalettar, *Phys. Rev. B* **76**, 085112 (2007).  
 [21] E. W. Huang, R. Sheppard, B. Moritz, and T. P. Devereaux, *Science* **366**, 987 (2019).  
 [22] T. A. Maier, M. Jarrell, T. Prushke, and M. H. Hettler, *Rev. Mod. Phys.* **77**, 1027 (2005).  
 [23] Y. Zhang, Y. F. Zhang, S. X. Yang, K.-M. Tam, N. S. Vidhyadhiraja, and M. Jarrell, *Phys. Rev. B* **95**, 144208 (2017).  
 [24] B. S. Shastry, *Phys. Rev. B* **87**, 125124 (2013).  
 [25] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016); R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges and B. S. Shastry, *ibid.* **88**, 235132 (2013); B. S. Shastry, E. Perepelitsky and A. C. Hewson, *ibid.* **88**, 205108 (2013).  
 [26] A. Damascelli, Z. Hussain, and Z.-X. Shen, *Rev. Mod. Phys.* **75**, 473 (2003).  
 [27] W. S. Lee, I. M. Vishik, D. H. Lu, and Z.-X. Shen, *J. Phys.: Condens. Matter* **21**, 164217 (2009).  
 [28] J. D. Koralek, J. F. Douglas, N. C. Plumb, Z. Sun, A. V. Federov, M. M. Murnane, H. C. Kapteyn, S. T. Cundiff, Y. Aiura, K. Oka, H. Eisaki, and D. S. Dessau, *Phys. Rev. Lett.* **96**, 017005 (2006).  
 [29] T. Yoshida, X. J. Zhou, D. H. Lu, S. Komiya, Y. Ando, H. Eisaki, T. Kakeshita, S. Uchida, Z. Hussain, Z.-X. Shen, and A. Fujimori, *J. Phys.: Condens. Matter* **19**, 125209 (2007).  
 [30] N. P. Armitage, D. H. Lu, C. Kim, A. Damascelli, K. M. Shen, F. Ronning, D. L. Feng, P. Bogdanov, X. J. Zhou, W. L. Yang, Z. Hussain, P. K. Mang, N. Kaneko, M. Greven, Y. Onose, Y. Taguchi, Y. Tokura, and Z.-X. Shen, *Phys. Rev. B* **68**, 064517 (2003).  
 [31] G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).  
 [32] P. Mai and B. S. Shastry, *Phys. Rev. B* **98**, 115101 (2018).  
 [33] A. Koitzsch, G. Blumberg, A. Gozar, B. S. Dennis, P. Fournier, and R. L. Greene, *Phys. Rev. B* **67**, 184522 (2003).  
 [34] W. Ding, R. Žitko, P. Mai, E. Perepelitsky and B. S. Shastry, *Phys. Rev. B* **96**, 054114 (2017); W. Ding, Rok Žitko, and B. S. Shastry, *ibid.* **96**, 115153 (2017).  
 [35] This can be seen, e.g., Eq. (11), where the “interaction” term in the self-energy is determined by the  $\varepsilon_{\vec{k}}$ ’s, the Fourier transform of the hopping matrix elements.  
 [36] D. N. Basov and T. Timusk, *Rev. Mod. Phys.* **77**, 721 (2005).  
 [37] D. N. Basov, R. D. Averitt, D. van der Marel, M. Dressel, and K. Haule, *Rev. Mod. Phys.* **83**, 471 (2011).  
 [38] F. Ming, S. Johnston, D. Mulugeta, T. S. Smith, P. Vilmercati, G. Lee, T. A. Maier, P. C. Snijders, and H. H. Weitering, *Phys. Rev. Lett.* **119**, 266802 (2017).

- [39] Y. J. Yan, M. Q. Ren, H. C. Xu, B. P. Xie, R. Tao, H. Y. Choi, N. Lee, Y. J. Choi, T. Zhang, and D. L. Feng, *Phys. Rev. X* **5**, 041018 (2015).
- [40] P. Choubey, A. Kreisel, T. Berlijn, B. M. Andersen, and P. J. Hirschfeld, *Phys. Rev. B* **96**, 174523 (2017).
- [41] A. Kreisel, P. Choubey, T. Berlijn, W. Ku, B. M. Andersen, and P. J. Hirschfeld, *Phys. Rev. Lett.* **114**, 217002 (2015).
- [42] K. Fujita, A. R. Schmidt, E.-A. Kim, M. J. Lawler, D. H. Lee, J. C. Davis, H. Eisaki, and S.-I. Uchida, *J. Phys. Soc. Jpn.* **81**, 011005 (2012).
- [43] P. W. Anderson, *Science* **235**, 1196 (1987).
- [44] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problem* (Benjamin, NY, 1962).
- [45] T. Arai and M. H. Cohen, *Phys. Rev. B* **21**, 3300 (1980); **21**, 3309 (1980); T. Arai, *ibid.* **21**, 3320 (1980).
- [46] P. Mai, S. R. White, and B. S. Shastry, *Phys. Rev. B* **98**, 035108 (2018).
- [47] Here  $J_{\vec{q}}$  is the Fourier transform of  $J_{ij}$  over nearest neighbors. For a homogeneous and translationally invariant system, the Fourier transform gives  $J_{\vec{q}} = 2J_x \cos(q_x a) + 2J_y \cos(q_y b)$  on a rectangular lattice where  $J_x$  and  $J_y$  are the exchange parameter along the principle axes.
- [48] M. Ogata and H. Fukuyama, *Rep. Prog. Phys.* **71**, 036501 (2008).
- [49] We refer to  $n = 0.85$  as the optimal density since in a typical single-layer cuprate superconductor such as  $\text{La}_{2-x}\text{SrCuO}_4$  (LSCO)—to which the 2D  $t$ - $J$  model is most relevant to the system—obtains its maximum critical temperature  $T_c$  at hole doping  $\delta = 0.15$ . Similarly,  $\text{Ba}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$  (Bi2212), a double-layer cuprate superconductor, obtains its max  $T_c$  at doping  $\delta = 0.15$ .
- [50] M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).
- [51] M. Hamermesh, *Group Theory And Its Application To Physical Problems* (Addison-Wesley Pub. Co., Reading, Mass, 1962).
- [52] L. D. Landau, and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Pergamon Press, New York, 1991), Vol. 3, p. 366.
- [53] The unstrained system assumes a body-centered orthorhombic unit cell  $(a, b, c)$  where  $a, b$  are the lattice constants of the base and  $c$  is that of the height. In the expression for resistivity  $\rho_0$ ,  $c_0$  corresponds to the interlayer separation  $c_0 = c/2$  between copper-oxide planes.
- [54] Y. Ando, in *High Tc Superconductors and Related Transition Metal Oxides*, edited by A. Bussmann-Holder and H. Keller (Springer, Berlin, 2007), pp. 17–28.
- [55] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.101.245149> for further discussion and supplemental figures on the effects of an  $x$ -axis strain on the resistivity, kinetic energy, and local density of states.
- [56] In high  $T_c$  systems estimates give  $t' \lesssim -0.27t$  Bi2212 and LSCO  $t' \sim 0.16t$ . We take  $t' = -0.2t$  as a compromise between the two extremes.
- [57] E. C. Blomberg, M. A. Tanatar, R. M. Fernandes, I. I. Mazin, Bing Shen, Hai-Hu Wen, M. D. Johannes, J. Schmalian, and R. Prozorov, *Nat. Commun.* **4**, 1914 (2013).
- [58] K. Ishida, S. Hosoi, Y. Teramoto, T. Usui, Y. Mizukami, K. Itaka, Y. Matsuda, T. Watanabe, and T. Shibauchi, *J. Phys. Soc. Jpn.* **89**, 064707 (2020).
- [59] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkins-Diehr, *Comput. Sci. Engineering* **16**, 62 (2014).

## Aspects of the normal state resistivity of cuprate superconductors

B. Sriram Shastry<sup>1,\*</sup> and Peizhi Mai<sup>2,1</sup>

<sup>1</sup>*Physics Department, University of California, Santa Cruz, California 95064, USA*

<sup>2</sup>*CNMS, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6494, USA*

 (Received 21 November 2019; revised manuscript received 14 January 2020; accepted 11 February 2020; published 13 March 2020)

Planar normal state resistivity data taken from three families of cuprate superconductors are compared with theoretical calculations from the recent extremely correlated Fermi liquid theory (ECFL) [B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011)]. The two hole-doped cuprate materials LSCO and BSLCO and the electron-doped material LCCO have yielded rich data sets at several densities  $\delta$  and temperatures  $T$ , thereby enabling a systematic comparison with theory. The recent ECFL resistivity calculations for the highly correlated  $t$ - $t'$ - $J$  model by us give the resistivity for a wide set of model parameters [B. S. Shastry and P. Mai, *New J. Phys.* **20**, 013027 (2018); P. Mai and B. S. Shastry, *Phys. Rev. B* **98**, 205106 (2018)]. After using x-ray diffraction and angle-resolved photoemission data to fix parameters appearing in the theoretical resistivity, only one parameter, the magnitude of the hopping  $t$ , remains undetermined. For each data set, the slope of the experimental resistivity at a *single* temperature-density point is sufficient to determine  $t$ , and hence the resistivity on absolute scale at all remaining densities and temperatures. This procedure is shown to give a fair account of the entire data.

DOI: [10.1103/PhysRevB.101.115121](https://doi.org/10.1103/PhysRevB.101.115121)

### I. INTRODUCTION

Understanding the normal state resistivity of high- $T_c$  cuprate superconductors and other strongly correlated materials is a challenging problem. The resistivity reveals the nature of the lowest energy charge excitations and therefore constitutes a relatively simple and yet fundamental probe of matter. In cuprates the different chemical compositions, conditions of preparation, and temperatures and a wide range of electronic densities lead to a complex variety of data sets. These are almost impossible to understand within the standard Fermi liquid theory of metals. Major puzzles are the almost  $T$ -linear planar resistivity of the hole-doped cuprates, the  $T^2$  resistivity of the closely related electron-doped cuprates, and the intermediate behavior at various densities. Indeed one of the larger questions about the cuprates is whether the differing  $T$  dependence of the electron-doped and hole-doped cases can possibly arise from a common physical model. Equally puzzling is the drastic reduction of the observed  $T$  scale of the resistivity variation ( $\sim 100$ – $400$  K) from a bare bandwidth ( $\sim eV$ 's) by a few orders of magnitude for both electron-doped and hole-doped cuprates. This situation has generated an upsurge of often radically new theoretical work on correlated systems in the last three decades, amounting to something like a revolution in condensed matter physics. In this new class of theories the planar resistivity stands at the center [1–14]; its unusual temperature dependence is most often emphasized.

In this work we bring theory face to face with experimental data on resistivity. We focus on the extremely correlated Fermi liquid theory (ECFL) proposed by Shastry [1,15,16], where a detailed and meaningful comparison has become

possible, as explained below. Starting from a microscopic Hamiltonian, the ECFL theory yields the resistivity on an absolute scale with a very few parameters determining the underlying model. The resistivity is calculated starting from the  $t$ - $t'$ - $J$  model [17–19] containing four parameters, of which three parameters can be fixed using ARPES and x-ray crystal structure data; thus only *one* parameter remains undetermined. The theory works in 2 dimensions without introducing any redundant degrees of freedom, and therefore the results can be meaningfully tested against data on a variety of cuprates, including both hole-doped and electron-doped cases.

### II. SUMMARY OF THE ECFL THEORY

A summary of the basic ideas and context of the ECFL theory is provided here; readers familiar with these ideas may skip to the later sections giving the results. The ECFL formalism is applicable in any dimension to doped Mott-Hubbard systems described by the  $t$ - $t'$ - $J$  model [17–19]

$$H = -t \sum_{\langle i,j \rangle} (\tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma} + \text{H.c.}) - t' \sum_{\langle\langle i,j \rangle\rangle} (\tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma} + \text{H.c.}) + J \sum_{\langle i,j \rangle} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j \right), \quad (1)$$

where  $\langle i,j \rangle$  ( $\langle\langle i,j \rangle\rangle$ ) denotes a sum over nearest (next-nearest) neighbors  $i, j$ , the Gutzwiller projector is given by  $P_G = \prod_i (1 - n_{i\uparrow} n_{i\downarrow})$ , the operator  $\tilde{C}_{i\sigma} = P_G C_{i\sigma} P_G$  is the Gutzwiller-projected version of the standard (canonical) fermion operator, and  $\vec{S}_i$  ( $n_i$ ) the spin (density) operator at site  $i$ . This model is in essence obtained from the Hubbard model by a canonical transformation implementing the large- $U$  limit [17]. The transformation preserves the physics of the strong-coupling Hubbard

\*sriram@physics.ucsc.edu

model at the lowest energies. The large energy scale  $U$  of the Hubbard model is traded for noncanonical anticommutation relations between Gutzwiller-projected electrons in the  $t$ - $t'$ - $J$  model. Standard (Feynman) diagrammatic many-body techniques do not apply to the  $t$ - $J$  model due to the effect of the Gutzwiller projection on the anticommutation relations. For the relevant operators  $\tilde{C}, \tilde{C}^\dagger$  of the model Eq. (1), the canonical fermionic anticommutator  $\{C_{i\sigma_i}, C_{j\sigma_j}^\dagger\} = \delta_{ij}\delta_{\sigma_i\sigma_j}$  is replaced by a noncanonical anticommuting Lie algebra

$$\{\tilde{C}_{i\sigma_i}, \tilde{C}_{j\sigma_j}^\dagger\} = \delta_{ij}(\delta_{\sigma_i\sigma_j} - \sigma_i\sigma_j\tilde{C}_{i\bar{\sigma}_i}^\dagger\tilde{C}_{j\bar{\sigma}_j}), \quad (2)$$

where  $\bar{\sigma}_i = -\sigma_i$ . An immediate resulting problem is that Wick's theorem simplifying products of operators into pairwise contractions is now invalid. Hence a formally exact and systematic Feynman-Dyson series expansion of the Green's functions in a suitable parameter is unavailable. On the other hand, in the Hubbard model with canonical fermions, the Feynman-Dyson series exists but is not controllable since  $U$  the parameter of expansion is very large for strong correlations. In trading the Hubbard model for the  $t$ - $J$  model in the Gutzwiller-projected subspace, we gain the tactical advantage of avoiding accounting for the large energy scale. However this advantage is lost unless we succeed in finding a corresponding formally exact expansion to replace the Feynman-Dyson series. The ECFL formalism solves this problem by replacing the Feynman-Dyson series with an alternate  $\lambda$  series. This series is formally exact and is an expansion of the Green's functions in a parameter  $\lambda$ . This parameter lives in a finite domain  $\lambda \in [0, 1]$ , interpolating between the free Fermi gas at  $\lambda = 0$  and the fully Gutzwiller-projected limiting case  $\lambda = 1$ . One way is to introduce  $\lambda$  as the coefficient of the noncanonical term in the anticommutator Eq. (2). For analogy it is useful to compare Eq. (2) with the contrast between the commutators of canonical bosons and the usual rotation group  $[SU(2)]$  Lie algebra of spin- $S$  particles. One finds [20] that  $\lambda$  plays a parallel role to the inverse spin, in the theory of quantum spin systems, i.e.,  $\lambda \leftrightarrow \frac{1}{2S}$ , where  $S = \frac{1}{2}, 1, \dots$ . For computing the Green's functions, we note the *exact* functional differential equation of the canonical Hubbard model and the  $t$ - $J$  model written in shorthand space-time-spin matrix notation [1,2] as

$$\left(g_0^{-1} - U \frac{\delta}{\delta \mathcal{V}} - UG\right) \cdot G = \delta \mathbb{1}, \quad (3)$$

$$(g_0^{-1} - \lambda \hat{X} - \lambda \hat{Y}_1) \cdot G = \delta(1 - \lambda\gamma), \quad (4)$$

where  $g_0^{-1}$  is the noninteracting Green's function,  $\gamma$  is a local version of  $G$ , and the remaining terms [of a similar character to the 2nd and 3rd terms in Eq. (3)] are detailed in [1,2]. Here Eq. (3) is the functional differential equation for the Hubbard model. By inverting the operator multiplying  $G$  and expanding in  $U$ , one generates the complete Feynman series in powers of  $U$  for the Hubbard model. In Eq. (4)  $\lambda$  is set at unity to obtain the exact equation for the  $t$ - $J$  model. Its iteration of the above type is not straightforward due to the extra time-dependent term on the right-hand side. These are the equations of motion in the presence of a space-time-spin dependent potential  $\mathcal{V}$ , which is set at zero at the end as prescribed in the Schwinger-Tomonaga method of field theory. The fermionic antiperiodic boundary conditions on

$G$  in the imaginary-time variable complete the mathematical statement of the problem. The ECFL formalism converts the noncanonical equation (4) into a pair of equations of the type Eq. (3) by introducing a decomposition of the Green's function  $G = g \cdot \tilde{\mu}$  into auxiliary Green's function  $g$  and a caparison function  $\tilde{\mu}$ . These pieces satisfy the exact equations

$$(g_0^{-1} - \lambda \overline{X} \cdot g \cdot g^{-1} - \lambda \hat{Y}_1) \cdot g = \delta \mathbb{1}, \quad (5)$$

$$\tilde{\mu} = \delta(1 - \lambda\gamma) + \lambda \overline{X} \cdot g \cdot \tilde{\mu}, \quad (6)$$

where the contraction symbol indicates that the functional derivative contained in  $X$  acts on the term at the other end of the symbol, while other terms satisfy matrix product rules. Notice that Eq. (5) looks similar to Eq. (3) with a unit matrix on the right-hand side, and is thus essentially like a canonical Green's function expression. The second equation, Eq. (6), must be solved simultaneously with Eq. (5), since  $\hat{Y}_1$  depends on both  $g$  and  $\tilde{\mu}$ . This task is done by expanding all variables systematically in powers of  $\lambda$  and writing down a set of successive equations to each order. The solution thus found is continuously connected to the free Fermi gas, and satisfies the Luttinger-Ward volume theorem at  $T = 0$ . The latter is an essential part of claiming that the resulting theory is a variety of Fermi liquid, being notoriously difficult to satisfy in uncontrolled approximations such as the truncations of Green's function equations. On setting the time-dependent potential to zero we get the frequency-dependent Green's function as

$$\begin{aligned} G(k, i\omega_j) &= g(k, i\omega_j) \times \tilde{\mu}(k, i\omega_j) \\ &= \frac{1 - \lambda \frac{n}{2} + \lambda \Psi(\vec{k}, i\omega_j)}{g_0^{(-1)}(\vec{k}, i\omega_j) - \lambda \Phi(\vec{k}, i\omega_j)}, \end{aligned} \quad (7)$$

where the two self-energies  $\Psi, \Phi$  determine  $G$ . The ECFL formalism has a systematic expansion of these equations in powers of  $\lambda$ , starting with the free Fermi gas as the lowest term and finally setting  $\lambda = 1$ . An expansion in  $\lambda$  thus provides a controlled framework for explicit calculations [1,15]. The current version of the theory [1–3,15] is valid to  $O(\lambda^2)$  and has been benchmarked against other standard techniques for strong coupling in limiting cases of infinite dimensionality [i.e., dynamical mean field theory (DMFT)] and the single-impurity limit [15]. Higher-order terms in  $\lambda$  are expected to impact the results outside the regime considered here, namely  $0.13 \lesssim \delta \lesssim 0.2$ . It has been recently applied to several objects of experimental interest such as angle-resolved photoemission (ARPES), Raman scattering, optical conductivity, the Hall constant, and recently the resistivity [2,3,16].

One of the main effects of strong correlations is to reduce significantly the quasiparticle weight  $Z$  from its Fermi gas value of unity. It is worth commenting that the exact DMFT studies of the Hubbard model in  $d = \infty$  using a mapping to a self-consistent Anderson impurity model yield a very small  $Z$  for  $U > 2.918D$  ( $2D$  is the bandwidth) as one approaches the insulating limit  $n \rightarrow 1$ . This is seen, e.g., in Fig. 1(a) of [21], where  $Z$  is plotted versus  $\delta = 1 - n$  for various  $U$ . One sees that  $Z$  decreases upon with increasing  $U$ , taking a nonzero value in the  $U = \infty$  limit. In this limit its density dependence is close to the empirical formula  $Z \sim \delta^{1.39}$ . In the case of the 2-d  $t$ - $t'$ - $J$  model the ECFL results [2] have

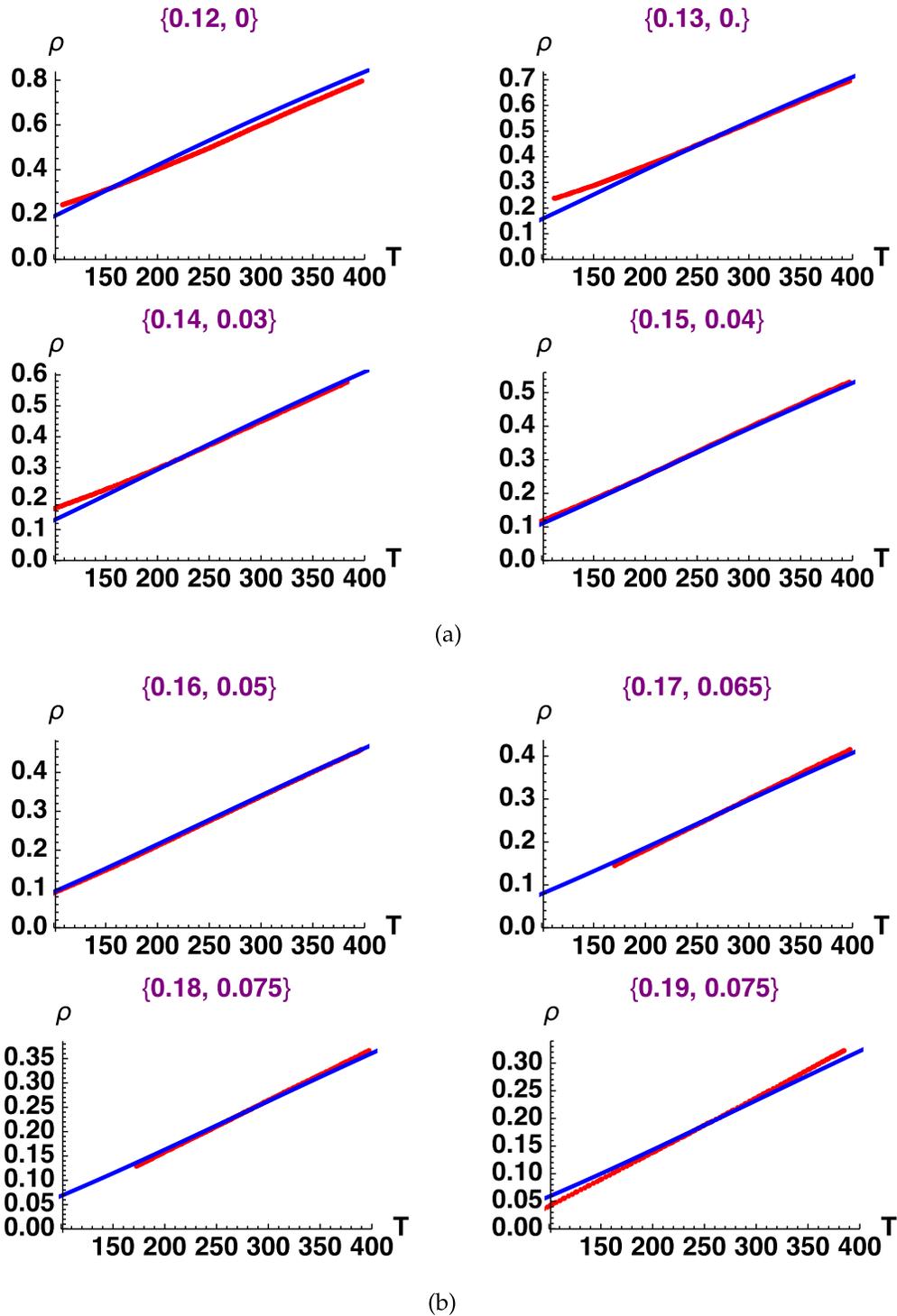


FIG. 1. (a) LSCO: Slightly underdoped to optimally doped. (b) LSCO: Near-optimal doping. The resistivity  $\rho$  is in units of  $\text{m}\Omega\text{cm}$  and  $T$  is in kelvins. Dotted (red) line is data extracted from Fig. 2(b) of Ando and co-workers [22], and solid (blue) curve is the theoretical curve with  $t'/t = -0.2$ . Panels (a) and (b) focus on densities in the slightly underdoped and near-optimal doping ranges. The displayed pair of numbers  $\{\delta, \rho_{\text{imp}}\}$  indicates the hole density and estimated impurity resistivity. The parameter  $t = 0.9$  eV was fixed using  $\Gamma(T^\Phi)$ , the slope of the resistivity [see Eq. (9)] at  $\delta = 0.15$ ,  $T^\Phi = 250$  K in panel (a). The resistivities at every other density in other panels and in the three panels of Fig. 2 are then predicted by the theory.

a similar character. The reduction of  $Z$  from unity occurs as we approach the insulating limit  $n \rightarrow 1$ . Additionally, it is very sensitive to the sign and magnitude of  $t'/t$  [23]. The dependence of  $Z$  on  $n$  and  $t'/t$  is most clearly seen in

Fig. 1 of [2]. Qualitatively we find that  $Z$  decreases when  $t'/t$  is negative and growing in magnitude, whereas a positive  $t'/t$  enhances its value. Within the theory, reduction of the magnitude of  $Z$ , i.e., the loss of weight of the quasiparticles, is

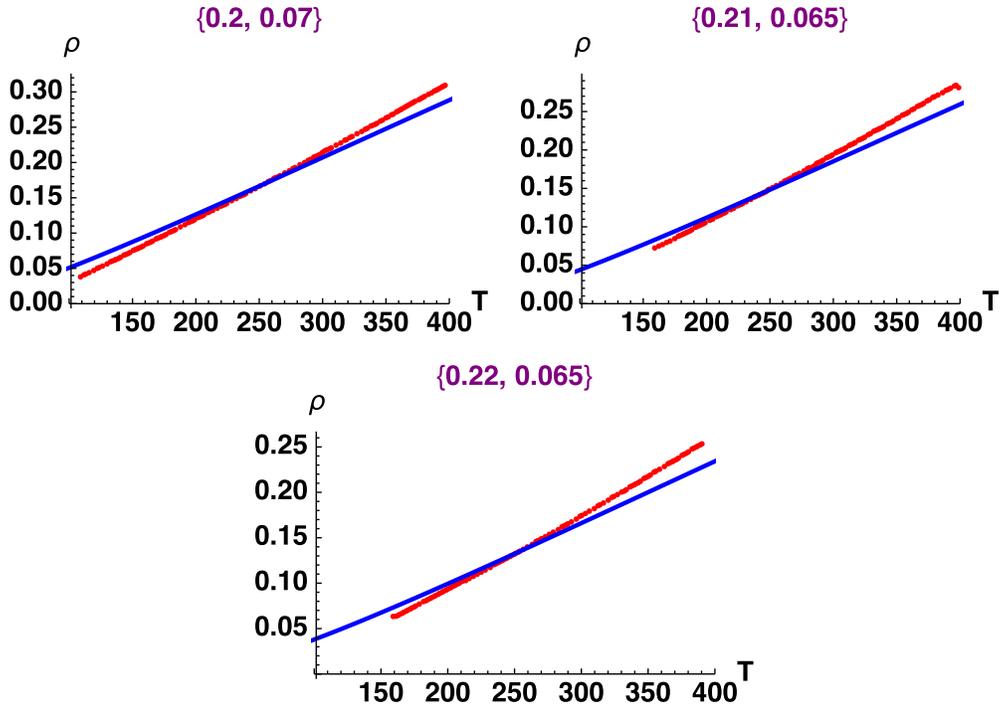


FIG. 2. The resistivity  $\rho$  is in units of  $\text{m}\Omega \text{ cm}$  and  $T$  is in kelvins. Dotted (red) line is data extracted from Fig. 2(b) of Ando and co-workers [22] for the slightly overdoped cases of LSCO, and solid (blue) curve is the theoretical curve with  $t'/t = -0.2$  and  $t = 0.9 \text{ eV}$ .

compensated exactly by the growth of the background pieces of the spectral function, as seen in Figs. 1–2 of [3]. We note that experiments on cuprates strongly indicate the growth of background weight, and indeed the ECFL theoretical results closely match experiments in regard to the shapes of spectral functions [16].

The resistivity calculations in Refs. [2,3] were performed for a typical set of model parameters chosen for illustrative purposes. In these works we noted that the resulting resistivities are broadly comparable to experiments in their magnitude and on the scale of temperature variation. In the present paper we push this observation to a more explicit and quantitative level, by comparing the ECFL results of [2,3] with experiments on a few representative high- $T_c$  materials with both hole and electron doping. Although broken symmetries of various types are possible within the methodology, we focus here on the properties of the paramagnetic normal state.

### III. PARAMETERS OF THE MODEL

The ECFL theory results used here [2,3] are valid for a quasi-two-dimensional correlated metal, with separation  $c_0$  between layers. The resistivity in the calculations [2,3] arises from intrinsic inelastic  $e$ - $e$  scattering with the umklapp processes, inherent in the tight-binding model, relaxing the momentum efficiently. The (smaller)  $a$  and  $b$  axis lattice constants cancel out in the formula for resistivity. The theory gives the planar resistivity in the form

$$\rho = R_{vK} \times c_0 \times \bar{\rho} \left( \frac{t'}{t}, \frac{J}{t}, \frac{k_B T}{t}, \delta \right), \quad (8)$$

where  $R_{vK} = \frac{h}{e^2} = 25813 \Omega$  is the von Klitzing resistance. The (dimensionless) theoretical resistivity  $\bar{\rho}$  is a function of

the four displayed dimensionless variables. Detailed formulas leading to this expression can be found in Eqs. (45) and (46) of [2] and Eqs. (12) and (13) of [3]. More precisely  $\delta$  is the concentration of holes measured from half filling, i.e.,  $\delta = 1 - n$  and  $n = \frac{N_e}{N_s}$ , where  $N_e$  ( $N_s$ ) is the number of electrons (copper sites). At  $\delta = 0$  ( $n = 1$ ) the model describes a Mott-Hubbard insulator. We discuss below the exchange parameter  $J/t$ , which plays a secondary role at the densities considered here. While three parameters  $c_0$ ,  $\delta$ ,  $T$  are obtained from experiments directly, ARPES constrains the parameter  $t'/t$  from the shape of the Fermi surface in most cases. Given these, the remaining single parameter  $t$  fixes the resistivity on an absolute scale. In addition the usually small and  $T$ -independent (extrinsic) impurity resistance, usually arising from scatterers located off the 2-d planes, must be estimated separately.

In addition to  $c_0$ , the basic parameters of the model are the nearest-neighbor hopping  $t$ , the second-neighbor hopping  $t'$ , and a superexchange energy  $J$  within a tight-binding description of the copper  $d$ -like bands. The parameter  $t'$  plays an important role in distinguishing between hole-doped superconductors ( $t' < 0$ ) with a positive Hall constant and the electron-doped superconductors ( $t' > 0$ ) with a negative Hall constant. The shape of the Fermi surface is sensitive to the ratio of the bare hopping parameters  $t'/t$ , if one assumes that interactions do not change its shape very much; this is largely borne out in ECFL theory. For this reason ARPES can most often provide us with a good estimate for this parameter  $t'/t$ , although  $t$  itself is not fixed by knowing the shape of the Fermi surface. We fix  $J$  at a typical value  $0.17t$ . At the densities we study here we find that the magnitude of  $J$  has a very limited influence on the calculated resistivity, as seen, e.g., in Fig. 24 of [3]. For the single-layer cuprate systems, one has two Cu-O

layer per unit cell and therefore the separation  $c_0$  equals half the  $c$ -axis lattice constant  $c_L$  [24,25]. The applicability of the theoretical calculations to systems with a higher number of layers per unit cell, such as Bi-2212 or  $\text{YBa}_2\text{Cu}_3\text{O}_{6-\delta}$ , is less direct. It requires making further assumptions relating  $c_0$  to the lattice constants. In order to avoid this we confine ourselves to single-layer systems.

The theoretical results tested here are found by ignoring a possible superconducting or magnetic state. We have produced a grid of theoretical calculations for  $t'/t = -0.4, -0.3, \dots, 0.3$  at several densities in the range  $0.12 \leq \delta \leq 0.22$  surrounding the interesting regime of optimal doping  $\delta \sim 0.15$ . Since the theory is smooth in most theoretical parameters we can interpolate in it, when necessary. Calculations are carried out in a wide range of  $T$  with a lower end  $T \sim 100$  K with a system size of  $62 \times 62$ . Lower  $T$  calculations require bigger system sizes which are computationally expensive and alternate methods are possible for estimating the resistivity. For example at lower  $T \lesssim 50$  K the resistivity can be extrapolated to a quadratic in  $T$  quite accurately using  $\rho = \alpha T^2 / (1 + T/T_0)$  with suitable constants  $\alpha, T_0$ . This form is consistent with the  $T \rightarrow 0$  Fermi liquid character of the theory below the (already low)  $T_0$

#### IV. THE CHOICE OF SYSTEMS

The lattice structure of the cuprates allows for a systematic change in carrier concentration by chemical substitution of elements situated away from the copper oxide planes, without severely impacting the impurity resistance. The role of block layers or charge reservoirs in hosting the donors away from the copper oxide planes plays an important role in achieving this property of the cuprates [24,25]. This feature also provides a useful handle in our analysis; we can access data on families of cuprates that contain a reasonably large range of electron densities. Since the basic parameters of the theoretical model can be assumed unchanged with doping [26], such a family provides a systematic proving ground for theory. Thus the experimental data used for testing the theory are narrowed down to the available systematic sets of resistivity data on single-layer cuprates with varying densities.

In Table I we list the single-layer cuprate compounds where data sets with several densities are available. The hole-doped LSCO and BSLCO materials are well studied by many authors, and the data set from Ref. [22] used here reports a very extensive set of densities for each family. This provides us with 11 densities for LSCO in the range  $0.12 \leq \delta \leq 0.22$  and 7 densities for BSLCO in the range  $0.12 \leq \delta \leq 0.18$  which are essentially within the range treatable by theory. We include recent thin-film data on the electron-doped LCCO from Ref. [27]. Here 4 densities are available in the theoretical range and the very regular  $\rho \sim +T^2$  type behavior of the data allows for easily eliminating the impurity contribution. For a more balanced representation of the electron-doped materials, we included data on NCCO from Ref. [28]. The NCCO family contains only two densities in the theoretically accessible range, of which one is impacted by 2-d localization effects. It is therefore not as constraining as the other families. The choice of the above four families of single-layer cuprates with

TABLE I. The single-layer cuprates analyzed in this work. For the first three materials the values of  $t'/t$  are obtained from ARPES experiments where the Fermi surface shape is fitted to a tight-binding model. For LCCO the ARPES data on the Fermi surface do not exist. The quoted  $t'/t$  is chosen to be the same as NCCO. The resistivity data for LCCO are from thin films while the other data are from single crystals. Here  $c_L$  is the  $c$ -axis lattice constant. In all the above cases the unit cell contains two copper oxide layers, and hence their separation  $c_0$  entering Eq. (8) is half the lattice constant  $\frac{1}{2}c_L$ . The last column lists the values of  $t$  determined in this work. The single adjustable parameter, the hopping  $t$ , is found using the slope of the experimental resistivity at 200 K at a single density  $\delta = 0.15$  as in Eq. (9). Band structure estimates of the  $t'/t$  ratio [29,30] are quite close to the ones used here, but the estimates of  $t$  differ somewhat. It must be kept in mind that the quoted parameter  $t$  is the bare one, i.e., prior to many-body renormalization.

Single-Layer High- $T_c$ Compounds			
Material	$c_L$ (Å)	$t'/t$	$t$ (eV)
$\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (LSCO)	13.25 [22,31]	-0.2 [32,33]	0.9
$\text{Bi}_2\text{Sr}_{2-x}\text{La}_x\text{CuO}_6$ (BSLCO)	24.3 [22]	-0.25 [33]	1.35
$\text{Nd}_{2-x}\text{Ce}_x\text{CuO}_4$ (NCCO)	12.01 [24,34]	+0.2 [35]	0.9
$\text{La}_{2-x}\text{Ce}_x\text{CuO}_4$ (LCCO)	12.45 [36]	+0.2	0.76

$\gtrsim 20$  sample densities seemed sufficiently representative for our task.

In addition to the above set of materials there are a few others belonging to the single-layer class with data provided for several densities. Among these we have excluded from our analysis the mercury compound Hg1201 ( $\text{HgBa}_2\text{CuO}_{4+\delta}$ ) [37] and the thallium compound Tl2201 ( $\text{Tl}_2\text{Ba}_2\text{CuO}_{6+\delta}$ ) [38,39]. In the literature for these compounds, the value of  $T_c$  for different samples is quoted and one needs to extract the electron density from other measurements, e.g., the Hall constant. This was hard for the authors to achieve, with a required accuracy in density  $\Delta\delta \sim 0.1$  necessary for the present analysis.

In Table I we quote the  $c$ -axis lattice constant  $c_L$  taken from x-ray diffraction data. The ratio  $t'/t$  is taken from angle-resolved photoemission (ARPES) experiments on the shape of the Fermi surface, fitted to a tight-binding band. In some cases the experimental fits include a small further neighbor hopping as well; we neglect it here since the corrections only fine-tune the shapes of the Fermi surface while preserving their basic topology. Theoretical estimates from band structure [29,30] are roughly consistent with the above experimentally guided choices of  $t'/t$ .

#### V. PROTOCOL FOR FIXING $t$ AND ESTIMATING THE IMPURITY RESISTIVITY

We determine the magnitude of  $t$  for each material by collating a data set consisting of experimental  $\rho_{\text{exp}}(T, \delta)$  points at various densities  $\delta = \delta_1, \delta_2, \dots$ . From this set we extract the slope of the resistivity

$$\Gamma(T^\Phi) = \left( \frac{d\rho_{\text{exp}}(T, \delta = 0.15)}{dT} \right)_{T=T^\Phi}. \quad (9)$$

Equating  $\Gamma(T^\Phi)$  to the corresponding theoretical slope at  $T^\Phi$  determines the single parameter  $t$ . The density is chosen as  $\delta = 0.15$  since it is in a regime where the calculation is quite reliable.  $T^\Phi$  is chosen as the midpoint of the temperature range of the data set, so that  $T^\Phi = 250$  K for LSCO and  $T^\Phi = 200$  K for BSLCO and LCCO in the following analysis.

We next need to estimate the  $T$ -independent impurity contribution to the resistivity at each density  $\rho_{\text{imp}}(\delta)$  for LSCO and BSLCO [40]. For LCCO the impurity contribution  $\rho_{\text{imp}}$  has been eliminated by the authors of [27]; thus this task is already done. For the others we shift down the experimental resistivity  $\rho_{\text{exp}}(T^\Phi, \delta)$  to match the theoretical resistivity; the magnitude of the shift gives us the estimated  $\rho_{\text{imp}}(\delta)$  at each density. We are thus using the relation  $\rho_{\text{exp}}(T^\Phi, \delta) - \rho_{\text{imp}}(\delta) = \rho_{\text{th}}(T^\Phi, \delta)$ , where  $\rho_{\text{th}}$  is from Eq. (8). The impurity contribution is displayed in all figures and is a small fraction of the total resistivity in all cases.

In summary fixing the magnitude of  $t$  for a data set requires a comparison with experiments at a *single density* ( $\delta = 0.15$ ) and a *single temperature* ( $T = T^\Phi$ ). The impurity contribution is estimated at each density at the same temperature  $T = T^\Phi$ . Checking these against data constitutes the essence of the test carried out here. The final two columns in Table I report the fitted value of the single undetermined parameter  $t$ . The bare bandwidth is estimated as  $W \sim 8t$ . Slightly different choices of the density and  $T^\Phi$  lead to comparable results for  $t$ .

Before looking at the results, we make a few comments about the analysis. (a) The requirement that the fitted values of  $t$  and  $t'/t$  remain unchanged for different densities  $\delta$  gives added significance to the fits. It is clearly an important and nontrivial requirement from any theory as well. In this sense matching the experimental resistivity at a single density of any particular compound is less significant than doing so at a sequence of different densities. (b) The impurity shifts reported in each curve are seen to be on a typically expected scale  $\sim 50\text{--}150 \mu\Omega \text{ cm}$ . The data on LCCO [27] are available with the impurity contribution already removed by the authors. (c) At low electron densities the effects of (2-d) electron localization are visible in some data sets. In these cases the impurity contribution leads to an upturn at low  $T$ . This upturn has been discussed extensively in the literature [22,28] and also manipulated with magnetic fields [41]. Since the ECFL theory excludes any strong-disorder effects, we do not expect to capture these in the fits.

For LCCO the digital data were provided by the authors of [27]. For the other data sets studied here the published resistivity data were digitized using the commercial software program DigitizeIt [42]. We found that the program works quite well provided the experimental curves do not overlap or cross. This feature limited our data extraction to some extent, as the reader might notice from the low-temperature truncation in the experimental data in the figures presented below.

We next describe the comparison for different systems.

## VI. LSCO

In Figs. 1(a) and 1(b) and Fig. 2 the extensive data set from Fig. 2(b) of [22] is compared with theoretical predic-

tions. The parameters in Table I are used here. The band parameter  $t = 0.9$  eV is found from the slope of the resistivity  $\delta = 0.15$ ,  $T^\Phi = 250$  K. All other densities are then predicted by theory on an absolute scale. While some deviations at low density  $\delta = 0.12$  and also at high density  $\delta \gtrsim 0.2$  are visible, the overall agreement seems fair. For the same parameters Fig. 6(a) shows the theoretical resistivity over an enlarged temperature window. Here subtle changes of curvature are visible at high and low  $T$ .

## VII. BSLCO and Bi-2201

In Fig. 3 the data for the BSLCO family of compounds  $\text{Bi}_2\text{Sr}_{2-x}\text{La}_x\text{CuO}_6$  from Ando [22] is compared with theory. The band parameter  $t = 1.35$  eV is found from the slope of the resistivity at  $\delta = 0.15$ ,  $T^\Phi = 200$  K. All other densities are then predicted by theory. For these parameters Fig. 6(b) shows the theoretical resistivity over an enlarged temperature window. The larger value of  $t$  in BSLCO relative to that in LSCO can be understood from comparing Figs. 6(a) and 6(b). The almost doubled value of  $c_0$  increases by a similar factor the resistance of BSLCO over that of LSCO, provided one is at the same scaled temperature  $T/t$ . A larger  $t$  spreads this increase over a larger  $T$  window.

## VIII. NCCO AND LCCO

The NCCO family of materials with composition  $\text{Nd}_{2-x}\text{Ce}_x\text{CuO}_4$  and the closely related LCCO family  $\text{La}_{2-x}\text{Ce}_x\text{CuO}_4$  are of considerable interest as counterpoints to the other two families studied above. Both have the opposite sign of the Hall constant from the hole-doped cases and display a pronounced  $T^2$ -type resistivity.

In a single band model description, such as the  $t$ - $t'$ - $J$  model used here, these materials can also be treated as having a filling less than half. The filling of these materials in the original electron picture is greater than half. Starting with a Hubbard model one can perform a particle-hole transformation of both spin species to map the model to less than half filling. For  $U$  large enough the  $t$ - $J$  model is once again introduced in the place of the Hubbard model. This process generates some  $U$ -dependent constant terms that are absorbed into the chemical potential. It also flips the sign of all hopping matrix elements. While the nearest-neighbor hopping  $t$  can be flipped back to the standard (positive) sign using a simple unitary transformation (exploiting the square lattice geometry), the second-neighbor hopping  $t'$  is now positive and the Fermi surface is electron-like.

On the materials side, the available data on NCCO [28] [see Fig. 9(b) therein] is relatively sparse in the metallic range containing only two samples. One of these is afflicted with strong-disorder effects at low  $T$ . In Fig. 4 we compare the data from Onose and co-workers [28] with theory. While the density  $\delta = 0.15$  is perfectly matched with theory, the lower density  $\delta = 0.125$  curve shows a distinct upturn at low  $T$ , as discussed in [28]. A systematic treatment of strong-disorder effects in the ECFL theory is currently missing.

The data on LCCO [27] give us four densities within the range covered by theory. In the absence of ARPES data we choose  $t'/t = 0.2$ , i.e., the same value as in NCCO. We have

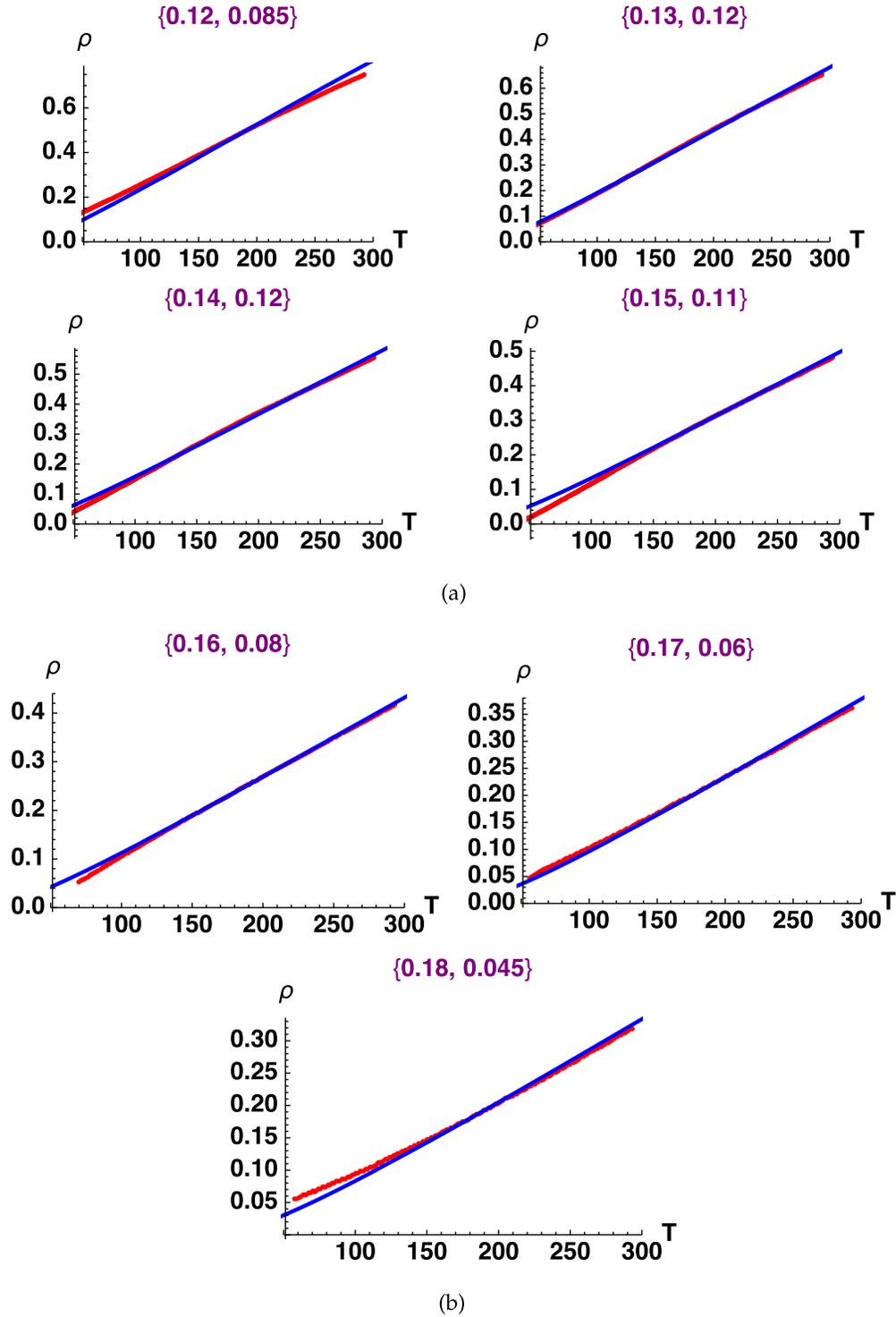


FIG. 3. (a) BSLCO: Slightly underdoped. (b) BSLCO: Near-optimal doping. The resistivity  $\rho$  is in units of  $\text{m}\Omega \text{ cm}$  and  $T$  is in kelvins. Dotted (red) line is data extracted from Fig. 1(a) of Ando and co-workers [22], and solid (blue) curve is the theoretical curve with  $t'/t = -0.25$ . Panels (a) and (b) focus on densities in the slightly underdoped and near-optimal doping ranges. The displayed pair of numbers  $\{\delta, \rho_{\text{imp}}\}$  indicates the hole density and estimated impurity resistivity. The parameter  $t = 1.35 \text{ eV}$  was fixed using  $\Gamma(T^\Phi)$ , the slope of the resistivity [see Eq. (9)] at  $\delta = 0.15$ ,  $T^\Phi = 200 \text{ K}$  in panel (a). The resistivity at every density in the other panels is then predicted by theory.

verified that nearby values to  $t'/t$  lead to a similar quality of fits after adjusting the parameter  $t$ , and hence this choice not final. The authors conveniently present the resistivity in Fig. 2(b) of [27] requiring no further impurity corrections. In

Fig. 5 we compare theory and experiment, and in Fig. 6(c) we present the theoretical resistivity on an extended  $T$  scale at several densities. The discrepancy in LCCO between theory and experiment at  $\delta = 0.17$  at  $T = 200$  is  $\sim 0.01$ , and is quite

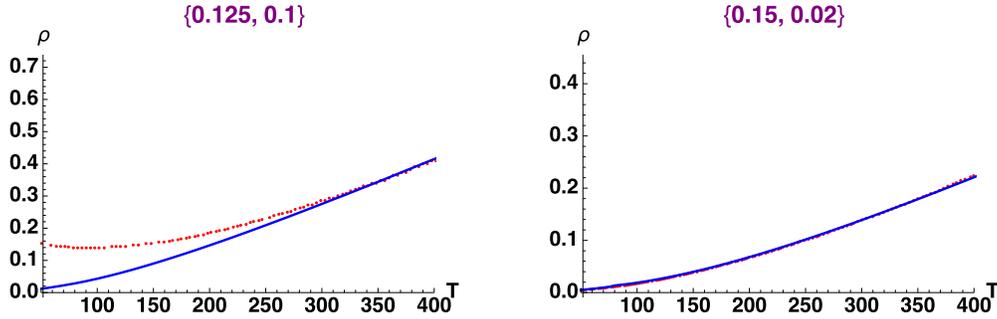


FIG. 4. The resistivity  $\rho$  is in units of  $m\Omega$  cm and  $T$  is in kelvins. Dotted (red) line is data extracted from Fig. 9(b) of Onose and co-workers [28], and solid (blue) curve is the theoretical curve with  $t'/t = +0.2$ . The parameter  $t = 0.9$  eV was fixed using the slope of the  $\delta = 0.15$  data at 200 K. The data set contains only these two densities within the range accessible to theory. The upturn in the lower density curve and the larger magnitude of impurity resistivity are due to strong disorder effects, as already noted in [27]. The sign of  $t'/t$  is reversed between this figure and Fig. 1 for LSCO, while other parameters  $c_0, t$  are essentially unchanged. Both the experimental and theoretical resistance display a resistance with a positive upward curvature (i.e.,  $\rho \sim +T^2$ ).

visible. However we should keep in mind that at corresponding densities the absolute scale of the resistivity for LCCO is considerably smaller than that for LSCO and BSLCO. This can be seen in Figs. 2 and 3. As a consequence a similar scale of absolute error leads to a much larger relative error.

### IX. DISCUSSION

We have presented a comparison of theoretical resistivity with extensive data on three families of cuprate superconductors. It is also feasible to fit data on noncuprate strongly correlated systems such as  $Sr_2RuO_4$  from [43], where data over a large range  $T \leq 1000$  are available. However data are

available at only one composition in this case, and the value of  $t'/t$  is hard to find from experiments. Since a single density within a family does not test the theory stringently, we omit the comparison here.

Overall we have shown that the ECFL theory gives a reasonable account of data in the three families discussed above. A small number of parameters taken from experimental data fix the model completely. It is encouraging that the resulting resistivity affords a reasonable fit to a collection of resistivity data at various densities, both in terms of the  $T$  dependence and its magnitude. It is also encouraging that upon using different model parameters, the same calculation

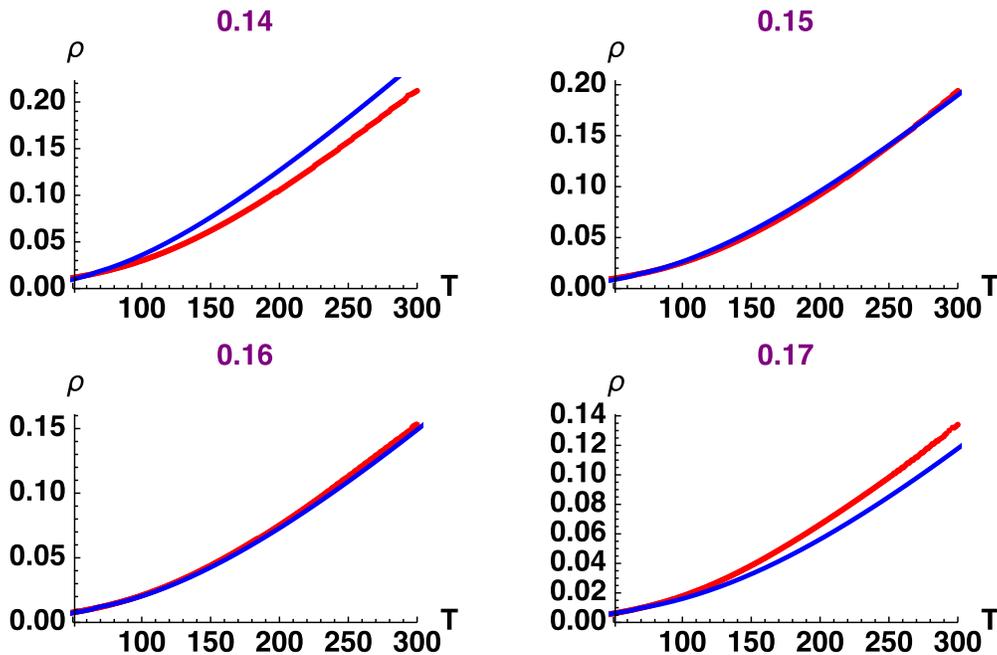


FIG. 5. The resistivity  $\rho$  is in units of  $m\Omega$  cm and  $T$  is in kelvins. Data are from Fig. 2(b) of Sarkar and co-workers [27] as the dotted red line. The impurity contribution in this data set has been removed by the authors in [27]. The theoretical curve is in solid blue, with  $t'/t = +0.2$ . The hole density is marked at the top in each plot. The parameter  $t = 0.76$  eV was fixed using  $\Gamma(T^\Phi)$  [see Eq. (9)], the slope of the resistivity at  $\delta = 0.15$ ,  $T^\Phi = 200$  K. The sign of  $t'/t > 0$  is common to NCCO and reversed from that in LSCO and BSLCO. Both experiments and theory find a resistance with a positive curvature (i.e.,  $\rho \sim +T^2$ ), as in NCCO. This is in striking contrast to LSCO and BSLCO as seen in Figs. 1, 2, and 3.

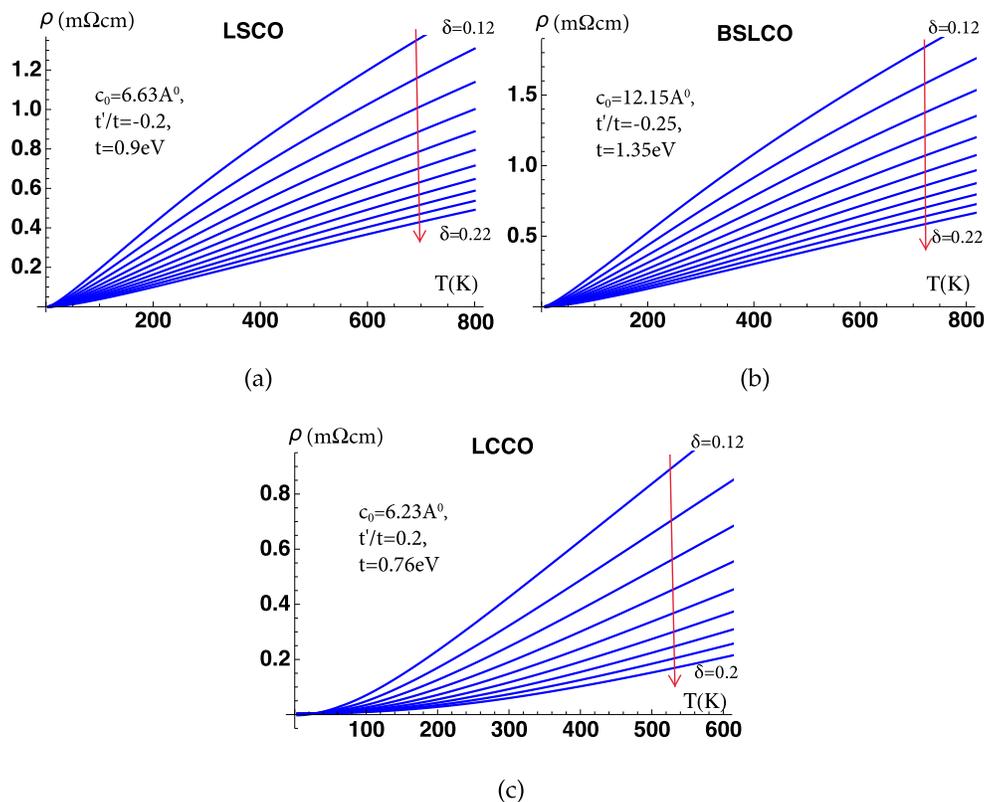


FIG. 6. (a) LSCO:  $\delta = 0.12 \rightarrow 0.22$  (increasing  $\downarrow$ ). (b) BSLCO:  $\delta = 0.12 \rightarrow 0.22$  (increasing  $\downarrow$ ). (c) LCCO:  $\delta = 0.12 \rightarrow 0.2$  (increasing  $\downarrow$ ). Theoretical resistivity curves for LSCO [panel (a)], BSLCO [panel (b)], and LCCO [panel (c)] over an extended temperature range. The hole densities increase downward at intervals  $\Delta\delta = 0.01$ . In going from LSCO with BSLCO the separation between the layers, i.e.,  $c_0$ , is almost doubled while  $t'/t$  changes only slightly. The resistivity at a comparable  $(\delta, T)$  here, and also in the data, changes by a smaller factor than  $c_0$ . In order to reconcile with this feature of the data, the deduced hopping parameter  $t$  is greater by  $\sim 50\%$  for BSLCO relative to LSCO. The distinct almost pure  $T^2$  behavior of the resistivity of LCCO relative to the other two systems is striking. Additionally it is noteworthy that the magnitude of the intrinsic resistivity of the electron-doped LCCO is considerably smaller than that of the hole-doped LSCO. Since these have roughly the same  $c_0, t, |t'/t|$  values, the difference is attributable to the different sign of  $t'/t$ .

fits the resistivity of both hole-doped and electron-doped materials.

In Fig. 6 we display the theoretical resistivities on a larger  $T$  scale and for more densities, using parameters of the three families separately. We found that the data are fitted almost equally well by making nearby choices of the pair  $t'/t$  and  $t$ . The differences between different choices do exist and show up but only at higher  $T$ , especially in the location of subtle kinks of the sort seen in Fig. 6.

## X. CONCLUSIONS

From the above exercise it appears that the extremely correlated Fermi liquid theory has the necessary ingredients to explain the variety of data seen in the above materials. Other materials, some of them with a higher number of layers, do display further subtle features which are missing in the theory. However these features are also missing in the displayed data from the above materials. We have thus made a fair beginning with the above “standard” cuprate materials, but further challenges from more complex behavior are to be expected.

A few comments on the results and their implications are appropriate. Let us first discuss the hole-doped materials. Here the quasilinear resistivity seen near  $\delta \sim 0.15$  is remarkable, as noted by many authors. We should also pay attention to the underlying suppression of scale. By this we refer to the fact that the temperature scale of resistivity variation is as low as  $\sim 100$ - $300$  K, starting from a bare bandwidth of almost  $10$  eV. The three orders of magnitude reduction in scale is nontrivial, reminiscent of the emergence of the low-energy Kondo scale in magnetic impurity systems. Starting from wide energy bands with a width of  $\sim 10$  eV, the ECFL theory systematically generates low energy and temperature scales, a few orders of magnitude smaller than the bare ones [1,15,16]. The low energy scales depend sensitively on the density and a few other parameters, especially the sign and magnitude of  $t'/t$ .

A major part of this scale suppression is due to the small quasiparticle weight  $Z \lesssim 0.1$  at relevant densities that arise in the theory [1–3]. More physically we can attribute this suppression to the profound role of Gutzwiller projection on the electron propagators near the Mott-Hubbard half-filled limit. It is captured to a good extent by the ECFL theory,

and is visible in the detailed structure of the electron spectral functions [1–3].

For the electron-doped materials, it is interesting that the theoretical resistivity matches experiments essentially as well as for the hole-doped materials. The two classes of materials have the opposite sign of the parameter  $t'/t$ , which is disconnected from the extent of correlations. This finding has a bearing on the frequently debated topic of the Fermi liquid nature of electron-doped cuprates. The ECFL theory says that both hole-doped and electron-doped systems are (extremely correlated) Fermi liquids at the lowest temperature. Additionally the theory quantifies the range of  $T$  where a Fermi liquid type behavior  $\rho \sim T^2$  holds good. Going further it also identifies regimes succeeding the Fermi liquid [1–3,15,44,45] upon warming.

In order to better understand the origin of the difference between hole and electron doping within the theory, the following observation may be helpful. It is known that the sign and magnitude of the parameter  $t'/t$  directly influences the magnitude of the already small quasiparticle weight  $Z$  (see Fig. 1 of [2]) [23]. A positive  $t'/t$  leads to a small  $Z$ , while a negative  $t'/t$  leads to an even smaller but nonvanishing  $Z$ . For the electron-doped case this distinction ultimately results in an enhanced thermal range displaying a positive curvature

of the  $\rho$ - $T$  plots. The effect on resistivity of the sign of  $t'/t$  can be seen explicitly by comparing the theoretical resistivity curves for the hole-doped cases Figs. 6(a) and 6(b) with the electron-doped case in Fig. 6(c).

As a cross-check on the theory, it would be interesting to compare other physical variables with data for the systems considered here, using the deduced parameters. Finally we should note that future technical developments in the implementation of the ECFL theory are likely to refine some of the theoretical results presented here.

#### ACKNOWLEDGMENTS

We are grateful to Professor A. J. Leggett for stimulating discussions and for valuable comments on the manuscript. We thank Prof. R. L. Greene and Dr. T. Sarkar for providing the digitized resistivity data of [27]. We also thank Prof. Y. Ando, Prof. M. Greven, and Prof. A. Ramirez for useful comments. The work at UCSC was supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award No. DE-FG02-06ER46319. The computation was done on the Comet in XSEDE [46] (TG-DMR170044) supported by National Science Foundation Grant No. ACI-1053575.

- 
- [1] B. S. Shastry, *Phys. Rev. Lett.* **107**, 056403 (2011).  
 [2] B. S. Shastry and P. Mai, *New J. Phys.* **20**, 013027 (2018).  
 [3] P. Mai and B. S. Shastry, *Phys. Rev. B* **98**, 205106 (2018).  
 [4] P. W. Anderson, *Phys. Rev. B* **78**, 174505 (2008); P. W. Anderson and Z. Zou, *Phys. Rev. Lett.* **60**, 132 (1988).  
 [5] C. M. Varma, P. B. Littlewood, S. Schmitt-Rink, E. Abrahams, and A. E. Ruckenstein, *Phys. Rev. Lett.* **63**, 1996 (1989).  
 [6] T. Moriya, Y. Takahashi, and K. Ueda, *J. Phys. Soc. Jpn.* **59**, 2905 (1990).  
 [7] P. Monthoux, A. V. Balatsky, and D. Pines, *Phys. Rev. Lett.* **67**, 3448 (1991).  
 [8] A. S. Alexandrov, V. V. Kabanov, and N. F. Mott, *Phys. Rev. Lett.* **77**, 4796 (1996).  
 [9] R. Zeyher and M. L. Kulić, *Phys. Rev. B* **53**, 2850 (1996); **54**, 8985 (1996); A. Greco and R. Zeyher, *Europhys. Lett.* **35**, 115 (1996).  
 [10] N. M. Plakida and V. S. Oduenko, *Phys. Rev. B* **59**, 11949 (1999).  
 [11] G. Kastinakis, *Phys. C (Amsterdam)* **340**, 119 (2000); *Phys. Rev. B* **71**, 014520 (2005).  
 [12] P. Phillips, *Philos. Trans. R. Soc. A* **369**, 1574 (2011).  
 [13] A. A. Patel and S. Sachdev, *Phys. Rev. Lett.* **123**, 066601 (2019).  
 [14] I. J. Robinson, P. D. Johnson, T. M. Rice, and A. M. Tsvelik, *Rep. Prog. Phys.* **82**, 126501 (2019).  
 [15] B. S. Shastry and E. Perepelitsky, *Phys. Rev. B* **94**, 045138 (2016); P. Mai, S. R. White, and B. S. Shastry, *ibid.* **98**, 035108 (2018).  
 [16] P. Mai and B. S. Shastry, *Phys. Rev. B* **98**, 115101 (2018); K. Matsuyama, E. Perepelitsky, and B. S. Shastry, *ibid.* **95**, 165435 (2017); G.-H. Gweon, B. S. Shastry, and G. D. Gu, *Phys. Rev. Lett.* **107**, 056404 (2011).  
 [17] The  $J$  part of the  $t$ - $J$  model for the insulator was introduced by P. W. Anderson, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1963), Vol. 14, p. 99; Using Kohn's general canonical transformation, A. B. Harris and R. Lange [*Phys. Rev.* **157**, 295 (1967)] derived the correct form of the kinetic energy in the lower Hubbard band [see Eq. (4.49b)]. The bare-bones version of the  $t$ - $J$  model used here, and in much of the literature, consists of this reduced kinetic energy and the superexchange term. We ignore the additional three-site correlated hopping terms arising from a canonical transform discussed by K. A. Chao, J. Spalek, and A. M. Oles [*J. Phys. C* **10**, L271 (1977)].  
 [18] P. W. Anderson, *Science* **235**, 1196 (1987); The glare of attention on the  $t$ - $J$  model appears to have begun with this paper.  
 [19] M. Ogata and H. Fukuyama, *Rep. Prog. Phys.* **71**, 036501 (2008) give a useful review of the  $t$ - $J$  model and some results using early techniques.  
 [20] B. S. Shastry, *Ann. Phys.* **343**, 164 (2014).  
 [21] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, *Phys. Rev. B* **88**, 235132 (2013); B. S. Shastry, E. Perepelitsky, and A. C. Hewson, *ibid.* **88**, 205108 (2013).  
 [22] Y. Ando, S. Komiya, K. Segawa, S. Ono, and Y. Kurita, *Phys. Rev. Lett.* **93**, 267001 (2004).  
 [23] The quasiparticle weight  $Z$  is given at various fillings and  $t'/t$  in Fig. 1 of [2]. The rapid plunge of the spectral function peaks in the Brillouin zone with warming are shown in Fig. 3 of [2]. Note that  $T$  is scaled by  $t$ , and the value of  $t$  in this figure is fixed at  $t = 0.45$  eV in [2]. Therefore  $T$  (and  $t$ ) must be rescaled for different materials as per the values discussed in Table I.  
 [24] Y. Tokura and T. Arima, *Jpn. J. Appl. Phys.* **29**, 2388 (1990).  
 [25] A. J. Leggett, *Quantum Liquids* (Oxford University Press, Oxford, 2006); see especially Sec. 7.2.

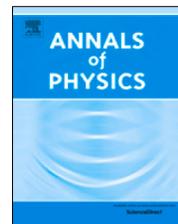
- [26] The lattice constant in many cuprates actually changes by a few percent with doping, as discussed by N. R. Khasanova and E. V. Antipov, *Phys. C (Amsterdam)* **246**, 241 (1995) and in [31,34]. While it is easy enough to accommodate this change in our calculation, it makes a relatively small difference and we neglect it here. Our comment on fixed parameters refers more strongly to the value of  $t'/t$  which can be taken as unchanged in a given family of compounds with varying densities.
- [27] T. Sarkar, R. L. Greene, and S. Das Sarma, *Phys. Rev. B* **98**, 224503 (2018).
- [28] Y. Onose, Y. Taguchi, K. Ishizaka, and Y. Tokura, *Phys. Rev. B* **69**, 024504 (2004) (see Fig. 10).
- [29] R. Markiewicz, *Phys. Rev. B* **72**, 054519 (2005).
- [30] E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, and O. K. Andersen, *Phys. Rev. Lett.* **87**, 047003 (2001).
- [31] S. Kanbe, K. Kishio, K. Kitazawa, K. Fueki, H. Takagi, and S. Tanaka, *Chem. Lett.* **547** (1987).
- [32] T. Yoshida, T. Yoshida, X. J. Zhou, D. H. Lu, S. Komiya, Y. Ando, H. Eisaki, T. Kakeshita, S. Uchida, Z. Hussain, Z.-X. Shen, and A. Fujimori, *J. Phys.: Condens. Matter* **19**, 125209 (2007).
- [33] M. Hashimoto, T. Yoshida, H. Yagi, M. Takizawa, A. Fujimori, M. Kubota, K. Ono, K. Tanaka, D. H. Lu, Z.-X. Shen, S. Ono, and Y. Ando, *Phys. Rev. B* **77**, 094516 (2008).
- [34] P. K. Mang, S. Larochelle, A. Mehta, O. P. Vajk, A. S. Erickson, L. Lu, W. J. L. Buyers, A. F. Marshall, K. Prokes, and M. Greven, *Phys. Rev. B* **70**, 094507 (2004).
- [35] D. M. King, Z.-X. Shen, D. S. Dessau, B. O. Wells, W. E. Spicer, A. J. Arko, D. S. Marshall, J. DiCarlo, A. G. Loeser, C. H. Park, E. R. Ratner, J. L. Peng, Z. Y. Li, and R. L. Greene, *Phys. Rev. Lett.* **70**, 3159 (1993).
- [36] A. Sawa, M. Kawasaki, H. Takagi, and Y. Tokura, *Phys. Rev. B* **66**, 014531 (2002).
- [37] D. Pelc, M. J. Veit, C. J. Dorow, Y. Ge, N. Barišić, and M. Greven, [arXiv:1902.00529](https://arxiv.org/abs/1902.00529).
- [38] T. Manako, Y. Kubo, and Y. Shimakawa, *Phys. Rev. B* **46**, 11019 (1992).
- [39] J. Kokalj, N. E. Hussey, and R. H. McKenzie, *Phys. Rev. B* **86**, 045132 (2012).
- [40] Extrapolating the normal state resistivity to  $T \rightarrow 0$  is possible for some densities where the  $T$  dependence is purely linear (or purely quadratic) over most of the range. In such a case extrapolation gives the same result as our procedure. For general cases with a more complex  $T$  dependence, extrapolation requires an indeterminate fitting function. Our procedure then corresponds to choosing the fitting function as in the panels of Fig. 6 (all vanishing as  $T \rightarrow 0$ ) plus an impurity contribution.
- [41] G. S. Boebinger, Y. Ando, A. Passner, T. Kimura, M. Okuya, J. Shimoyama, K. Kishio, K. Tamasaku, N. Ichikawa, and S. Uchida, *Phys. Rev. Lett.* **77**, 5417 (1996).
- [42] I. Bormann, DigitizeIt (version 2.0), <http://www.digitizeit.de>.
- [43] H. Berger, L. Forro, and D. Pavuna, *Eur. Phys. Lett.* **41**, 531 (1998).
- [44] W. Ding, R. Žitko, P. Mai, E. Perepelitsky, and B. S. Shastry, *Phys. Rev. B* **96**, 054114 (2017); W. Ding, R. Žitko, and B. S. Shastry, *ibid.* **96**, 115153 (2017).
- [45] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, *Phys. Rev. Lett.* **110**, 086401 (2013).
- [46] J. Town *et al.*, *Comput. Sci. Eng.* **16**, 62 (2014).



ELSEVIER

Contents lists available at ScienceDirect

## Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)

## Extremely correlated superconductors

B. Sriram Shastry

Physics Department, University of California, Santa Cruz, CA, 95064, United States of America



## ARTICLE INFO

## Article history:

Received 2 September 2021

Accepted 8 September 2021

Available online 17 September 2021

## Keywords:

The  $t$ - $J$  model

Cuprates d-wave superconductivity

Correlated Gor'kov equation

## ABSTRACT

Superconductivity in the  $t$ - $J$  model is studied by extending the recently introduced extremely correlated fermi liquid theory. Exact equations for the Greens functions are obtained by generalizing Gor'kov's equations to include extremely strong local repulsion between electrons of opposite spin. These equations are expanded in a parameter  $\lambda$  representing the fraction of double occupancy, and the lowest order equations are further simplified near  $T_c$ , resulting in an approximate integral equation for the superconducting gap. The condition for  $T_c$  is studied using a model spectral function embodying a reduced quasiparticle weight  $Z$  near half-filling, yielding an approximate analytical formula for  $T_c$ . This formula is evaluated using parameters representative of single layer High- $T_c$  systems. In a narrow range of electron densities that is necessarily separated from the Mott-Hubbard insulator at half filling, we find a typical  $T_c \sim 10^2$  K.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The single band  $t$ - $J$  model Eq. (1), [1,2], and the closely related strong coupling Hubbard model have attracted much attention in recent years. In large part the interest is due to the potential relevance of these models in describing the phenomenon of High  $T_c$  superconductivity, discovered in cuprate materials in 1987 [3] and later, in other materials. These models lead to a single sheet of the fermi surface, and are specified by fixing the band hopping  $t$  and the exchange energy  $J$  for the  $t$ - $J$  model, or equivalently  $4t^2/U$  for the strong coupling ( $U \gg t$ ) Hubbard model, where the interaction is given by  $V_{Hub.} = U \sum_i n_{i\downarrow} n_{i\uparrow}$ . The exotic possibility of superconductivity arising from such inherently repulsive systems, is surprising from a theoretical perspective, and also challenging. Significant theoretical work using a variety of tools on the strong coupling Hubbard model and

E-mail address: [sriram@physics.ucsc.edu](mailto:sriram@physics.ucsc.edu).

<https://doi.org/10.1016/j.aop.2021.168614>

0003-4916/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the extremely strong coupling  $t$ - $J$  models [4–16] has given useful insights into the role of strong correlations in cuprate superconductivity. However given the non-triviality of the theoretical task of analytically solving these models, progress in that direction has been slow.

In this work we extend the extremely correlated fermi liquid theory (ECFL) [17,18] recently formulated to overcome the analytical difficulties of the strong coupling models, to include superconducting type broken symmetry. Upon cooling the normal metallic state, a superconducting instability is expected to arise, and our main goal in this study is to determine the conditions for the occurrence of this state, and to provide its detailed description.

In order to motivate these calculations of the superconducting state, it is useful to summarize the main features of ECFL theory as applied to the normal (non-superconducting) state so far. We provide a broad overview next, further details can be found in Refs. [17,18].

The methodology developed in this theory starts with exact functional differential equations for the various Greens function, obtained using the Tomonaga–Schwinger approach of external potentials. These equations incorporate the modification of the anti-commutation relations between the fermion operators due to Gutzwiller projection (see Eq. (5)). While providing a formally exact starting point for us, these equations are not yet amenable to systematic approximations. The core difficulty is that an additional set of terms arise from this modified non-canonical anticommutator structure Eq. (5). These non-canonical terms multiply the most singular term in the equation, namely the Dirac delta function (originating in the time derivative of the time ordering  $\Theta$  functions in the Greens functions). For an explicit example, note the  $\gamma$  term multiplying the delta function in Eq. (34).

In order to make progress, we therefore need to go beyond the established framework of Tomonaga–Schwinger. The first development in ECFL is that the above inconvenient feature of a non-canonical coefficient of the delta function, is eliminated by factoring the Greens function into two parts, the auxiliary Greens function  $\mathbf{g}$  and the caparison function  $\tilde{\mu}$  (see Eq. (46) and the discussion in the text following it). The auxiliary Greens function  $\mathbf{g}$  now satisfies a canonical equation (as in Eq. (51) by ignoring the term involving  $\mathbf{f}$ ), while the caparison function  $\tilde{\mu}$  accounts for the non-canonical nature of the original equation (as in Eq. (52)). This factorization process and the resulting equations are exact.

As the next development, we introduce a parameter  $\lambda$  in the range  $0 \leq \lambda \leq 1$  into these exact equations. Setting  $\lambda=0$  gives the uncorrelated system, while  $\lambda=1$  gives the exact equations of the strongly correlated system. The  $\lambda$  parameter has a formal similarity to the expansion parameter  $\frac{1}{2S}$  used in the Dyson–Maleev (or Holstein–Primakoff) formulations [19,20] of the spin-wave theory of magnets. The magnetic models involve spin operators satisfying the SU(2) (angular momentum) Lie algebra. They can be approached using different strategies. On the one hand we may think of spins as canonical bosons with a constraint on their occupation number  $n_i^b$  at any site  $i$ , namely  $n_i^b = 0, 1, \dots, 2S$ . This constraint can be implemented using a repulsive interaction between bosons  $Un_i^b(n_i^b - 1) \dots (n_i^b - 2S)$ , and finally letting  $U \rightarrow \infty$ . This bosonic Hubbard model is difficult to solve, since the large energy scale  $U$  makes the use of perturbation theory impractical. On the other hand we can employ the Dyson–Maleev (or Holstein–Primakoff) non-linear mappings to bosons, and expand the relevant Heisenberg equations of motion in a series in  $\frac{1}{2S}$ . This gives an efficient way of solving the models to considerable precision at fairly low orders in  $\frac{1}{2S}$ . This latter method is parallel to the  $\lambda$  expansion employed here, since the modified anticommutators Eq. (5) also yield a (non-canonical) Lie algebra. This analogy is discussed further in Ref. [18] (Sec. 6). In a different setting, the parameter  $\lambda$  can also be related to the fraction of doubly occupied states [21] (see Appendix A).

The parameter  $\lambda$  serves two important and related objectives. Firstly it provides a continuous path between the uncorrelated and the fully correlated system equations. Since  $0 \leq \lambda \leq 1$ , dialing it up from 0 does not involve invoking a large energy scale, unlike for example, dialing up  $U$  in the Hubbard model. This (isothermal) continuity enables the ECFL method to retain the ideal (i.e. non-interacting) fermi surface volume at low  $T$ . This ideal volume is expected for weakly interacting fermi systems from the Luttinger–Ward perturbative arguments [22], and importantly, survives the transition to extremely strongly correlated regions, as argued recently using non-perturbative arguments [23]. Lastly, the ideal volume is also seen in photoemission studies of overdoped and

optimally doped cuprate superconductors in the normal state [24], which provide a useful starting point for our study.

The second aspect of  $\lambda$  is that it can be used to organize a systematic power series expansion, analogous in spirit to the skeleton graph expansion of Dyson [25] in perturbative theories. This  $\lambda$  expansion can be carried out order by order, leading to a set of successive equations that are amenable to numerical study. A question might arise, whether a low order calculation in this expansion can capture the strongly correlated limit. For answering this, it is useful to examine the results for the  $d = \infty$  Hubbard model at  $U = \infty$ , where numerically exact results are available from the dynamical mean field theory [26]. The  $\lambda$  expansion to  $\mathcal{O}(\lambda^2)$  is compared with the exact numerical result from the dynamical mean field theory [27], in Fig. (6) of Ref. [28]. This shows that the calculated quasiparticle weight  $Z$  vanishes upon approaching a density of 1-particle per site, i.e. half filling. This vanishing is a hallmark of the strong correlation limit, where the Mott–Hubbard insulating state is realized. In the above  $d = \infty$  study, and also in the case of the 2-dimensional  $t$ - $J$  model [29–31], the  $\lambda$  expansion describes an extremely correlated Fermi liquid state, characterized by a small quasiparticle weight that vanishes near the Mott–Hubbard insulator, accompanied by a rich set of low energy scales located above the (strongly suppressed) effective fermi temperature. The  $\mathcal{O}(\lambda^2)$  equations for the normal state have been applied to calculations of the asymmetric photoemission lines [30–32], and most recently the calculation of the almost T-linear resistivity in single layer cuprates [29].

In this paper we extend the above formalism to the case where superconducting order emerges at low temperatures. This requires a non-trivial generalization to the superconducting state of the various steps of the ECFL theory highlighted above. In a similar fashion to the normal state, we first obtain exact equations for the normal and anomalous Greens functions for the  $t$ - $J$  model. These equations generalize Gor’kov’s equations for BCS type weak coupling superconductivity [33] by including the effect of extremely strong local repulsion between electrons. These equations are studied further using a specific decomposition of the Greens functions into two pieces (see Eq. (46)). This step is followed by a systematic expansion in a parameter  $\lambda$ . This leads to a set of equations Eqs. (51), (52), (54), iterating these in  $\lambda$  to all orders constitutes the exact answer. In the present work, we perform a leading order calculation.

In order to obtain explicit results, Eqs. (51), (52), (54) are further simplified near  $T_c$  where the order parameter is small, leading to simplified versions of these in Eqs. (55)–(57). These are treated to  $\mathcal{O}(\lambda^2)$ , and the lowest order condition for  $T_c$  is formulated in Eq. (68). In summary Eq. (68) is the leading order term near  $T_c$ , within the  $\lambda$  expansion, and constitutes an important formal result of the present work. In principle it should be possible to find further systematic equations to higher order, and also to extend the results for  $T \ll T_c$  following the procedure laid out here. In this work we are content to study this first set in detail. The transition temperature is given from Eq. (68), which is expressed in terms of the electronic Greens function, renormalized by strong correlations. In this renormalization the short ranged Hubbard–Gutzwiller terms are dominant, and the pairing energy causing the instability, is provided by the much smaller exchange energy  $J$ . This equation exhibits both a tendency towards an insulating state due to a diminished quasiparticle weight, and a tendency towards superconductivity due to the exchange term  $J$ . Their competing tendencies play out in Eq. (68) and the closely related Eq. (70). These equations determine whether superconductivity is found at all, and further identifies the model parameters that promote it. When the superconducting state is found, they also provide an estimate of the range of densities and temperatures which favor it.

The conditions Eqs. (68) and (70) are evaluated using a simple phenomenological electronic spectral function, modeling strong correlations near half filling in terms of a density dependent quasiparticle weight  $Z$  and a wide background. This model has the advantage of leading to an explicit analytical formula for  $T_c$ , in terms of the various parameters of the  $t$ - $J$  model, thus allowing for a thorough understanding of the role of different parameters on the result. Evaluating this expression we find that the model supports a d-wave superconducting phase consistent with data [34,35], located away from half filling. The  $T_c$  is found to be typically  $\sim 10^2$  K, i.e. an order of magnitude smaller than that of the model Eq. (2) where the sole difference from the  $t$ - $J$  model is that short ranged Hubbard–Gutzwiller type correlations are ignored, in a range of densities

determined by the band parameters. The temperature–density phase diagram has the form of a tapered tower Fig. 1. A smooth dome structure reported in cuprates, is replaced here by a somewhat narrow density range and an exaggerated height near the peak. The location of the peak can be varied by choosing the hopping parameters, but always remains well-separated from the insulating limit.

The paper is organized as follows. In Section 2 we define the  $t$ - $J$  Hamiltonian, express it in terms of the correlated fermionic operators, and outline the method of external potentials employed to generate the exact dynamical equations for the electron Greens function  $\mathcal{G}$  and the Gor'kov anomalous Greens function  $\mathcal{F}$ . In Section 3 the equation is expanded in  $\lambda$  and further simplified near  $T_c$ . In Section 4 the condition for  $T_c$  is evaluated using a model spectral function. This section contains expressions that involve only the electronic spectral function, and might be directly accessible to readers who are more interested in the concrete results. In Section 5 we conclude with a discussion of the results.

## 2. Theoretical preliminaries

The  $t$ - $J$  Hamiltonian [1,2] is

$$\begin{aligned} H_{tJ} &= H_t + H_J \quad (1) \\ H_t &= - \sum_{ij\sigma} t_{ij} \tilde{c}_{i\sigma}^\dagger \tilde{c}_{j\sigma} - \mu \sum_i n_i \\ H_J &= \frac{1}{2} \sum_{ij} J_{ij} (\vec{S}_i \cdot \vec{S}_j - \frac{n_i n_j}{4}) \end{aligned}$$

where  $t_{ij}$  are the band hopping matrix elements detailed below,  $J_{ij}$  the nearest neighbor exchange and  $\mu$  the chemical potential, with the density operator  $n_i = \sum_\sigma \tilde{c}_{i\sigma}^\dagger \tilde{c}_{i\sigma}$ , and spin density operator  $S_i^\alpha = \frac{1}{2} \sum_{\sigma\sigma'} \tilde{c}_{i\sigma}^\dagger \tau_{\sigma\sigma'}^\alpha \tilde{c}_{i\sigma'}$ ,  $\tau^\alpha$  is a Pauli matrix and the correlated fermi destruction operator  $\tilde{c}_i$  is found from the plain (i.e. canonical or unprojected) operators  $c_i$ , by sandwiching it between two Gutzwiller projection operators  $\tilde{c}_{i\sigma} = P_G c_{i\sigma} P_G$ , where  $P_G \equiv \prod_j (1 - n_{j\uparrow} n_{j\downarrow})$  [36]. It acts by eliminating all states with double occupancy in the state space. The creation operators follow by taking their hermitean conjugate. The physical meaning of this sandwiching process is that the fermi operators act within the subspace where projector  $P_G$  enforces single occupancy at each site. The  $t$ - $J$  model may be obtained by taking the large  $U$  limit of the Hubbard model [1]. It has also been argued [2] to be the low energy effective Hamiltonian for an underlying three-band model, describing the copper oxygen lattice of the cuprate superconductors, where it is found by eliminating high energy states of the model.

In the following work we will also find it useful to study the model

$$H_{\text{unc-tJ}} = - \sum_{ij} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} - \mu \sum_i n_i + \frac{1}{2} \sum_{ij} J_{ij} (\vec{S}_i \cdot \vec{S}_j - \frac{n_i n_j}{4}). \quad (2)$$

We may view it as an *uncorrelated*  $t$ - $J$  model in contrast to the correlated version Eq. (1), here the ultra strong short ranged Hubbard–Gutzwiller correlations with  $U \gg \max\{|t_{ij}|\}$  are turned off, while the relatively weak exchange term  $J \ll \max\{|t_{ij}|\}$  is retained. All operators that appear in Eq. (2), including the density and spin, are defined by the same expression as Eq. (1) but with the unprojected fermion operators  $c_{i\sigma}, c_{i\sigma}^\dagger$ 's. In this model the exchange term, which is usually viewed as the mechanism for antiferromagnetism, doubles up to play the role of a superconducting pairing potential. This fruitful observation of Anderson, Baskaran and Zou [5,6] follows from viewing the interaction in the crossed or Cooper channel. It is paralleled in our discussion later (see paragraph below Eq. (30)), where the exchange term, after a rearrangement amounting to a crossed channel, leads to a mean Cooper pair expectation in Eq. (31). Its superconducting solution, found by standard BCS–Gor'kov meanfield theory, is presented below (see Eqs. (74), (75)), and serves as a useful reference point in the study of the strongly correlated  $t$ - $J$  model.

It is convenient for our calculations to use the operators invented by Hubbard Refs. [37,38] to represent this projection process. Ref. [39] (Sec.8) discusses the origin of difficulties of the early work employing the Hubbard operators, in reproducing the Luttinger–Ward Fermi surface volume at low temperatures. In contrast the present ECFL formalism achieves this goal successfully, using continuity with the Fermi gas and the  $\lambda$  expansion described in [17,39] and below. We denote

$$\tilde{c}_{i\sigma}^\dagger \leftrightarrow X_i^{\sigma 0}, \quad \tilde{c}_{i\sigma} \leftrightarrow X_i^{0\sigma}, \quad \tilde{c}_{i\sigma}^\dagger \tilde{c}_{i\sigma'} \leftrightarrow X_i^{\sigma\sigma'}. \quad (3)$$

These operators satisfy the following fundamental anti-commutation relations and their adjoints:

$$\{X_i^{0\sigma_i}, X_j^{0\sigma_j}\} = 0 \quad (4)$$

$$\{X_i^{0\sigma_i}, X_j^{\sigma_j 0}\} = \delta_{ij} \left( \delta_{\sigma_i\sigma_j} - \sigma_i\sigma_j X_i^{\bar{\sigma}_i\bar{\sigma}_j} \right), \quad \bar{\sigma} = -\sigma. \quad (5)$$

In physical terms, for a given site index  $i$  and with  $\{a, b\} \in \{0, \uparrow, \downarrow\}$  limited to the three allowed initial and final states of the projected Hilbert space, the symbol  $X_i^{ab}$  represents an operator representing all allowed matrix elements. To yield the correct fermion antisymmetry, the creation operator  $X_i^{\sigma_i 0}$  anti-commutes with creation or destruction operators at *different* sites with any spin. In terms of these operators we can rewrite

$$H_t = - \sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i\sigma} X_i^{\sigma\sigma} \quad (6)$$

$$H_j = - \frac{1}{4} \sum_{ij\sigma_i\sigma_j} J_{ij} \sigma_i \sigma_j X_i^{\sigma_i\sigma_j} X_j^{\bar{\sigma}_i\bar{\sigma}_j}. \quad (7)$$

In the following we employ a convenient repeated internal spin summation convention. We shall follow the convention that in an equation defining any object, often (but not always) indexed by external spin indices, all the *internal and repeated* spin indices are to be summed over. As an example, we could drop the explicit summation over spins in Eqs. (6) and (7), but not in Eq. (5) where  $\sigma_i, \sigma_j$  are external spin indices that appear on the left hand side. We also use a repeated internal site index below.

In order to calculate the Greens functions for this model, we add an imaginary time  $\tau$  dependent external potential (or source term)  $\mathcal{A}$  to the definition of thermal averages. The expectation of an arbitrary observable  $Q(\tau_1, \dots)$ , composed e.g. of a product of several (imaginary) time ordered Heisenberg picture operators, is written in the notation

$$\langle\langle Q(\tau_1, \dots) \rangle\rangle = \text{Tr } P_\beta T_\tau \{e^{-\mathcal{A}} Q(\tau_1, \dots)\}. \quad (8)$$

Here  $T_\tau$  is the time-ordering operator, an external potential term  $\mathcal{A} = \int_0^\beta d\tau \mathcal{A}(\tau)$ , and  $P_\beta = e^{-\beta H} / \text{Tr} (e^{-\beta H} T_\tau e^{-\mathcal{A}})$  is the Boltzmann weight factor including  $\mathcal{A}$ . Here  $\mathcal{A}(\tau)$  is a sum of two terms,  $\mathcal{A}_\rho(\tau)$  involving a density-spin dependent external potential  $\mathcal{V}$ , and  $\mathcal{A}_c(\tau)$  involving  $\mathcal{J}$  ( $\mathcal{J}^*$ ) Cooper pair generating (destroying) external potentials. These are given by

$$\begin{aligned} \mathcal{A}_\rho(\tau) &= \sum_i \mathcal{V}_i^{\sigma_i\sigma_j}(\tau) X_i^{\sigma_i\sigma_j}(\tau) \\ \mathcal{A}_c(\tau) &= \frac{1}{2} \sum_{ij} \left( \mathcal{J}_{j\sigma_j i\sigma_i}^*(\tau) X_i^{0\sigma_i}(\tau) X_j^{0\sigma_j}(\tau) + \mathcal{J}_{i\sigma_i j\sigma_j}(\tau) X_i^{\sigma_i 0}(\tau) X_j^{\sigma_j 0}(\tau) \right), \end{aligned} \quad (9)$$

where the repeated internal spin convention implies summing over  $\sigma_i, \sigma_j$ , and where we require the antisymmetry  $\mathcal{J}_{i\sigma_i; j\sigma_j} = -\mathcal{J}_{j\sigma_j; i\sigma_i}$  and likewise for  $\mathcal{J}^*$ . The external potentials  $\mathcal{J}, \mathcal{J}^*$  in Eq. (9) couple to operators that add and remove Cooper pairs of correlated electrons, and are essential to describe the superconducting phase. At the end of the calculations, the external potentials are switched off, so that the average in Eq. (8) reduces to the standard thermal average. Tomonaga [40] in 1946 and Schwinger [41] in 1948 (TS) pioneered the use of such external potentials [25,42]. We next illustrate this technique for the present problem.

## 2.1. Using external potentials

The advantage of introducing these external potential (or “sources”) is that we can take the (functional) derivatives of Greens function with respect to the added external potentials in order to generate higher order Greens functions. If we abbreviate the external term as  $\mathcal{A} = \sum_i \mathcal{U}_i(\tau) V_j(\tau)$ , where  $\mathcal{U}_i(\tau)$  is one of the above  $c$ -number potential, and  $V_j(\tau)$  is the corresponding operator in the imaginary-time Heisenberg picture, and  $Q_i(\tau)$  an arbitrary observable, straightforward differentiation leads to the TS identity

$$\text{Tr} P_{\beta} T_{\tau} \{ e^{-\mathcal{A}} Q_i(\tau') V_j(\tau) \} = \langle \langle Q_i(\tau') \rangle \rangle \langle \langle V_j(\tau) \rangle \rangle - \frac{\delta}{\delta \mathcal{U}_i(\tau)} \langle \langle Q_i(\tau') \rangle \rangle \quad (10)$$

This important identity can be found by taking the functional derivative of Eq. (8) with respect to  $\mathcal{U}_i(\tau)$  (see e.g. Ref. [21] Eq. (18)), and is now illustrated with various choices of the external potential.

From Eq. (10) we note the frequently used result

$$\langle \langle \sigma_i \sigma_j X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau) Q(\tau') \rangle \rangle = (\gamma_{\sigma_i \sigma_j}(i\tau) - \mathcal{D}_{\sigma_i \sigma_j}(i\tau)) \langle \langle Q(\tau') \rangle \rangle \quad (11)$$

where

$$\begin{aligned} \gamma_{\sigma_i \sigma_j}(i, \tau) &= \sigma_i \sigma_j \langle \langle X_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau) \rangle \rangle \\ \mathcal{D}_{\sigma_i \sigma_j}(i, \tau) &= \sigma_i \sigma_j \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_i \bar{\sigma}_j}(\tau)}, \end{aligned} \quad (12)$$

The singlet Cooper pair operator is

$$\left( X_i^{0\uparrow} X_j^{0\downarrow} - X_i^{0\downarrow} X_j^{0\uparrow} \right) = \sigma X_i^{0\sigma} X_j^{0\bar{\sigma}}, \quad (13)$$

where summation over  $\sigma$  is implied on the right hand side, and its Hermitean conjugate

$$-\left( X_i^{\uparrow 0} X_j^{\downarrow 0} - X_i^{\downarrow 0} X_j^{\uparrow 0} \right) = \bar{\sigma} X_i^{\sigma 0} X_j^{\bar{\sigma} 0}. \quad (14)$$

We define the (singlet) Cooper pair correlation functions at time  $\tau$  as

$$C_{ij}(\tau) = \langle \langle \sigma X_i^{0\sigma}(\tau) X_j^{0\bar{\sigma}}(\tau) \rangle \rangle \quad (15)$$

$$C_{ij}^*(\tau) = \langle \langle \bar{\sigma} X_i^{\sigma 0}(\tau) X_j^{\bar{\sigma} 0}(\tau) \rangle \rangle, \quad (16)$$

where  $\sigma$  is summed over. We note that  $C_{ij}^*$  equals the complex conjugate of  $C_{ij}$  only after the external potentials are finally turned off, but not so in the intermediate steps.

The basic equation Eq. (10) for the Cooper pair operators with an arbitrary operator  $Q$  is

$$\frac{\delta}{\delta \mathcal{J}_{i\sigma;j\sigma_j}^*(\tau)} \langle \langle Q \rangle \rangle = \langle \langle X_j^{0\sigma_j}(\tau) X_i^{0\sigma_i}(\tau) \rangle \rangle \langle \langle Q \rangle \rangle - \langle \langle X_j^{0\sigma_j}(\tau) X_i^{0\sigma_i}(\tau) Q \rangle \rangle \quad (17)$$

$$\frac{\delta}{\delta \mathcal{J}_{i\sigma;j\sigma_j}(\tau)} \langle \langle Q \rangle \rangle = \langle \langle X_i^{0\sigma_i}(\tau) X_j^{0\sigma_j}(\tau) \rangle \rangle \langle \langle Q \rangle \rangle - \langle \langle X_i^{0\sigma_i}(\tau) X_j^{0\sigma_j}(\tau) Q \rangle \rangle \quad (18)$$

From these relations the Cooper-pair correlations can be found by summing over the spins

$$\langle \langle \sigma X_i^{0\sigma}(\tau) X_j^{0\bar{\sigma}}(\tau) Q \rangle \rangle = [C_{ij}(\tau) - \kappa_{ij}(\tau)] \langle \langle Q \rangle \rangle \quad (19)$$

$$\langle \langle \bar{\sigma} X_i^{\sigma 0}(\tau) X_j^{\bar{\sigma} 0}(\tau) Q \rangle \rangle = [C_{ij}^*(\tau) - \kappa_{ij}^*(\tau)] \langle \langle Q \rangle \rangle \quad (20)$$

where

$$\kappa_{ij}(\tau) = \bar{\sigma} \frac{\delta}{\delta \mathcal{J}_{i\sigma;j\bar{\sigma}}^*(\tau)} \quad (21)$$

$$\kappa_{ij}^* = \bar{\sigma} \frac{\delta}{\delta \mathcal{J}_{i\sigma;j\bar{\sigma}}(\tau)}, \quad (22)$$

where  $\sigma$  is summed over.

## 2.2. Greens functions and their dynamical equations

We are interested in the electron Greens function (see e.g. Ref. [21] Eq. (17)) expressed compactly by

$$\mathcal{G}_{i\sigma j\sigma_f}(\tau, \tau') = -\langle\langle X_i^{0\sigma_i}(\tau) X_j^{\sigma_f 0}(\tau') \rangle\rangle, \quad (23)$$

where the Dyson time ordering  $T_\tau$  and the external potential factor  $e^{-\mathcal{A}}$  are included in the definition of the brackets Eq. (8). To describe the superconductor, following Gor'kov [33] we define the anomalous Greens function :

$$\mathcal{F}_{i\sigma j\sigma_f}(\tau, \tau') = \bar{\sigma}_i \langle\langle X_i^{\bar{\sigma}_i 0}(\tau) X_j^{\sigma_f 0}(\tau') \rangle\rangle \quad (24)$$

where  $\bar{\sigma} \equiv -\sigma$ , and as in Eq. (23), the Dyson time ordering  $T_\tau$  and the external potential factor  $e^{-\mathcal{A}}$  are included in the definition of the brackets Eq. (8)

We note that the Cooper pair correlation functions Eq. (16), which plays a crucial role in defining the order parameter of the superconductor, can be expressed in terms of the anomalous Greens function using

$$C_{ij}^*(\tau) = \mathcal{F}_{i\sigma j\sigma}(\tau, \tau), \quad (25)$$

where  $\sigma$  is to be summed over, as per the convention used. We will also need the equal time correlation of creation operators  $C_{ij}(\tau)$  Eq. (15). It is straightforward to show that when the external potentials  $\mathcal{A}$  are switched off, this object is independent of  $\tau$  and can be obtained by complex conjugation of  $C_{ij}^*$ . It is possible to add another anomalous Greens function with two destruction operators as in Eq. (24), corresponding to Nambu's generalization of Gor'kov's work. In the present context it adds little to the calculation and is avoided by taking the complex conjugate of  $C_{ij}^*$  to evaluate  $C_{ij}$ .

### 2.2.1. Greens function $\mathcal{G}$

The equations for the Greens functions follow quite easily from the Heisenberg equations, followed by the use of the identity Eq. (10), and has been discussed extensively by us earlier. There is one new feature, concerning an alternate treatment of the  $H_j$  (exchange) term, necessary for describing superconductivity described below. In this section we make use of the internal repeated site index summation convention quite extensively.

Taking the  $\tau$  derivative of  $\mathcal{G}$  we obtain

$$\begin{aligned} \partial_\tau \langle\langle X_i^{0\sigma_i}(\tau) X_j^{\sigma_f 0}(\tau') \rangle\rangle &= \delta(\tau - \tau') \delta_{ij} (\delta_{\sigma_i \sigma_f} - \gamma_{\sigma_i \sigma_f}(i\tau)) \\ &+ \langle\langle [H_t + H_j + \mathcal{A}(\tau), X_i^{0\sigma_i}(\tau)] X_j^{\sigma_f 0}(\tau') \rangle\rangle \end{aligned} \quad (26)$$

We work on the terms on the right hand side. At time  $\tau$  we note

$$[H_t + \mathcal{A}_\rho, X_i^{0\sigma_i}] = \mu X_i^{0\sigma_i} - \mathcal{V}_i^{\sigma_i \sigma_j} X_i^{0\sigma_j} + t_{ij} (\delta_{\sigma_i \sigma_j} - \sigma_i \sigma_j X_i^{\bar{\sigma}_i \bar{\sigma}_j}) X_j^{0\sigma_j}, \quad (27)$$

where the repeated internal indices  $\sigma_j$  and  $j$  are summed over. From this basic commutator, using Eq. (10), Eq. (11) and the definitions Eq. (12) we obtain

$$\begin{aligned} \langle\langle [H_t + \mathcal{A}_\rho(\tau), X_i^{0\sigma_i}(\tau)] X_j^{\sigma_f 0}(\tau') \rangle\rangle &= (\mu \delta_{\sigma_i \sigma} - \mathcal{V}_i^{\sigma_i \sigma}) \langle\langle X_i^{0\sigma}(\tau) X_j^{\sigma_f 0}(\tau') \rangle\rangle \\ &+ t_{ij} \langle\langle X_j^{0\sigma_i}(\tau) X_j^{\sigma_f 0}(\tau') \rangle\rangle - t_{ij} (\gamma_{\sigma_i \sigma}(i, \tau) - \mathcal{D}_{\sigma_i \sigma}(i, \tau)) \langle\langle X_j^{0\sigma}(\tau) X_j^{\sigma_f 0}(\tau') \rangle\rangle, \end{aligned} \quad (28)$$

where the repeated spin index  $\sigma$ , and the site index  $j$  are summed over, while  $\sigma_i, \sigma_f$  and site indices  $i, f$  are held fixed.

For the exchange term

$$[H_j, X_i^{0\sigma_i}] = \frac{1}{2} J_{ij} \sigma_i \sigma X_i^{0\sigma} X_j^{\bar{\sigma}_i \bar{\sigma}} \quad (29)$$

$$= -\frac{1}{2} J_{ij} \sigma_i X_j^{\bar{\sigma}_i 0} (X_i^{0\uparrow} X_j^{0\downarrow} - X_i^{0\downarrow} X_j^{0\uparrow}), \quad (30)$$

where the repeated internal indices  $\sigma$  and  $j$  are summed over. In order to obtain Eq. (30) from Eq. (29), we used  $X_j^{\bar{\sigma}_i^0} = X_j^{\bar{\sigma}_i^0} X_j^{0\bar{\sigma}}$  and anticommuted the equal time operators  $X_i^{0\sigma} X_j^{\bar{\sigma}_i^0}$  into  $-X_j^{\bar{\sigma}_i^0} X_i^{0\sigma}$ , followed by an explicit sum over  $\sigma$ . This subtle step is essential for obtaining the superconducting phase, as discussed (para following Eq. (2)) in the Introduction, since the role of exchange in promoting Cooper pairs manifests itself here. Using Eq. (19) we find

$$\langle\langle [H_j, X_i^{0\sigma_i}(\tau)] X_f^{\sigma_f^0}(\tau') \rangle\rangle = -\frac{1}{2} J_{ij} \sigma_i (C_{ij}(\tau^+) - \mathcal{K}_{ij}(\tau^+)) \langle\langle X_j^{\bar{\sigma}_i^0}(\tau) X_f^{\sigma_f^0}(\tau') \rangle\rangle, \quad (31)$$

where the repeated internal index  $j$  is summed over, and with  $\eta$  is a positive infinitesimal we indicate here and elsewhere  $\tau^+ \equiv \tau + \eta$  and  $\tau^- \equiv \tau - \eta$ .

In treating this term we could have proceeded differently by sticking to Eq. (29), using Eq. (10) with a different external potential term as in Eq. (11) to write

$$\begin{aligned} \langle\langle [H_j, X_i^{0\sigma_i}(\tau)] X_f^{\sigma_f^0}(\tau') \rangle\rangle &= \frac{1}{2} J_{ij} \sigma_i \sigma \langle\langle X_i^{0\sigma}(\tau) X_j^{\bar{\sigma}_i^0}(\tau) X_f^{\sigma_f^0}(\tau') \rangle\rangle \\ &= -\frac{1}{2} J_{ij} (\gamma_{\bar{\sigma}_i^0}(j, \tau) - \mathcal{D}_{\bar{\sigma}_i^0}(j, \tau)) \langle\langle X_i^{0\sigma}(\tau) X_f^{\sigma_f^0}(\tau') \rangle\rangle, \end{aligned} \quad (32)$$

where the repeated spin index  $\sigma$ , and the site index  $j$  are summed over, while  $\sigma_i, \sigma_f$  and site indices  $i, f$  are held fixed. These two expressions Eq. (31) and Eq. (32) are alternate ways of writing the higher order Greens functions [43]. In order to describe a broken symmetry solution with superconductivity, we are required to use Eq. (31), since using the other alternative disconnects the normal and anomalous Greens functions altogether, thereby precluding a superconducting solution.

The term  $\langle\langle [\mathcal{A}_C(\tau), X_i^{0\sigma_i}(\tau)] X_f^{\sigma_f^0}(\tau') \rangle\rangle$  generates a term that is linear in  $\mathcal{J}$  which is treated similarly and the final result quoted in Eq. (34).

We summarize these equations compactly by defining

$$\begin{aligned} G_{0i\sigma_j}^{-1} &= \delta_{ij} \delta_{\sigma_i \sigma_j} (\boldsymbol{\mu} - \partial_\tau) + t_{ij} \delta_{\sigma_i \sigma_j} - \delta_{ij} \mathcal{V}_i^{\sigma_i \sigma_j} \\ Y_{i\sigma_j} &= t_{ij} \gamma_{\sigma_i \sigma_j}(i, \tau) \\ X_{i\sigma_j} &= -t_{ij} \mathcal{D}_{\sigma_i \sigma_j}(i, \tau), \end{aligned} \quad (33)$$

and write the exact equation

$$\begin{aligned} (G_{0i\sigma_j}^{-1} - Y_{i\sigma_j} - X_{i\sigma_j}) \mathcal{G}_{j\sigma_j \sigma_f}(\tau, \tau') &= \delta(\tau - \tau') \delta_{if} (\delta_{\sigma_i \sigma_f} - \gamma_{\sigma_i \sigma_f}(i, \tau)) \\ &+ \frac{1}{2} J_{ij} (C_{ij}(\tau) - \mathcal{K}_{ij}(\tau)) \mathcal{F}_{j\sigma_j \sigma_f}(\tau, \tau') \\ &+ \mathcal{J}_{\sigma_j; i\sigma_k} (\delta_{\sigma_i \sigma_k} - \gamma_{\sigma_i \sigma_k}(i, \tau) + \mathcal{D}_{\sigma_i \sigma_k}(i, \tau)) \sigma_j \mathcal{F}_{j\bar{\sigma}_i \sigma_f}(\tau, \tau'), \end{aligned} \quad (34)$$

where the spins  $\sigma_j, \sigma_k$  and the site index  $j$  are summed over, while  $\sigma_i, \sigma_f$  and site indices  $i, f$  are held fixed. The final term drops off when we switch off the external potential  $\mathcal{J}$ . Viewing the spin and site indices as joint matrix indices, these equations and their counterparts Eq. (40), are transformed into matrix equations below.

### 2.2.2. Greens function $\mathcal{F}$

The Gor'kov Greens function  $\mathcal{F}$  in Eq. (24) satisfies an exact equation that can be found as follows. First we note

$$\partial_\tau \langle\langle X_i^{\bar{\sigma}_i^0}(\tau) X_f^{\sigma_f^0}(\tau') \rangle\rangle = \langle\langle [H_t + H_j + \mathcal{A}(\tau), X_i^{\bar{\sigma}_i^0}(\tau)] X_f^{\sigma_f^0}(\tau') \rangle\rangle \quad (35)$$

A part of the right hand side satisfies

$$\begin{aligned} \langle\langle [H_t + \mathcal{A}_\rho(\tau), X_i^{\bar{\sigma}_i^0}(\tau)] X_f^{\sigma_f^0}(\tau') \rangle\rangle &= -(\boldsymbol{\mu} \delta_{\sigma_i \sigma} - \mathcal{V}_i^{\bar{\sigma}_i^0}) \langle\langle X_i^{\bar{\sigma}_i^0}(\tau) X_f^{\sigma_f^0}(\tau') \rangle\rangle \\ &- t_{ij} \langle\langle X_j^{\bar{\sigma}_i^0}(\tau) X_f^{\sigma_f^0}(\tau') \rangle\rangle + t_{ij} (\gamma_{\bar{\sigma}_i^0}(i\tau) - \mathcal{D}_{\bar{\sigma}_i^0}(i\tau)) \langle\langle X_i^{\sigma_i^0}(\tau); X_f^{\sigma_f^0}(\tau') \rangle\rangle, \end{aligned} \quad (36)$$

where the repeated spin index  $\sigma$ , and the site index  $j$  are summed over, while  $\sigma_i$ ,  $\sigma_f$  and site indices  $i, f$  are held fixed. The exchange term is treated similarly to Eq. (29)

$$[H_J, X_i^{\bar{\sigma}_i 0}] = \frac{1}{2} J_{ij} \left( X_i^{\uparrow 0} X_j^{\downarrow 0} - X_i^{\downarrow 0} X_j^{\uparrow 0} \right) \sigma_i X_j^{0 \sigma_i} \quad (37)$$

so that using Eq. (20) we get

$$\langle\langle [H_J, X_i^{\bar{\sigma}_i 0}] X_f^{\sigma_f 0}(\tau') \rangle\rangle = -\frac{1}{2} J_{ij} \sigma_i \left( C_{ij}^*(\tau^-) - \mathcal{K}_{ij}^*(\tau^-) \right) \langle\langle X_j^{0 \sigma_i}(\tau) X_f^{\sigma_f 0}(\tau') \rangle\rangle, \quad (38)$$

where the repeated internal index  $j$  is summed over

We gather and summarize these equations in terms of the variables that are “time-reversed” partners of Eq. (34) and hence denoted with hats:

$$\begin{aligned} \widehat{G}_{0i\sigma_j\sigma_j}^{-1} &= \delta_{ij} \delta_{\sigma_i\sigma_j} (\boldsymbol{\mu} + \partial_\tau) + t_{ij} \delta_{\sigma_i\sigma_j} - \delta_{ij} \mathcal{V}_i^{\bar{\sigma}_i\bar{\sigma}_j} \\ \widehat{Y}_{i\sigma_j\sigma_j} &= t_{ij} \gamma_{\bar{\sigma}_j\bar{\sigma}_i}(i, \tau) \\ \widehat{X}_{i\sigma_j\sigma_j} &= -t_{ij} \mathcal{D}_{\bar{\sigma}_j\bar{\sigma}_i}(i, \tau) \end{aligned} \quad (39)$$

So that

$$\begin{aligned} \left( \widehat{G}_{0i\sigma_j\sigma_j}^{-1} - \widehat{Y}_{i\sigma_j\sigma_j} - \widehat{X}_{i\sigma_j\sigma_j} \right) \mathcal{F}_{j\sigma_j\sigma_f}(\tau, \tau') &= -\frac{1}{2} J_{ij} (C_{ij}^* - \mathcal{K}_{ij}^*) \mathcal{G}_{j\sigma_j\sigma_f}(\tau, \tau') \\ &+ \sigma_i \sum_m \mathcal{J}_{i\bar{\sigma}_n m \sigma_m}^* (\delta_{\sigma_i, \sigma_n} - \gamma_{\bar{\sigma}_n \bar{\sigma}_i}(i, \tau) + \mathcal{D}_{\bar{\sigma}_n \bar{\sigma}_i}(i, \tau)) \mathcal{G}_{m\sigma_m\sigma_f}(\tau, \tau') \end{aligned} \quad (40)$$

where the repeated spin indices  $\sigma_j$ ,  $\sigma_n$ ,  $\sigma_m$  and site index  $j$  are summed over, while  $\sigma_i$ ,  $\sigma_f$  and  $i, f$  are held fixed. The final term arising from  $\langle\langle [\mathcal{A}_C, X_i^{\bar{\sigma}_i 0}] X_f^{\sigma_f 0}(\tau') \rangle\rangle$  drops off when we switch off the external potential  $\mathcal{J}^*$ .

### 2.2.3. Summary of equation in symbolic notation

Eq. (34) and Eq. (40) are exact in the strong correlation limit. Noting that all terms containing  $\gamma$  and  $\mathcal{D}$  in Eq. (34) and Eq. (40) arise from Gutzwiller projection, we obtain the corresponding equations for the uncorrelated  $t$ - $J$  model in Eq. (2) by dropping these terms. Recall also that the external potentials  $\mathcal{J}$ ,  $\mathcal{J}^*$  represent the imposed symmetry-breaking terms that force superconductivity, and are meant to be dropped at the end. In this uncorrelated case, let us understand the role of the terms with the Cooper pair derivatives  $\mathcal{K}$ ,  $\mathcal{K}^*$ . If we ignore these terms and also set  $\mathcal{J}$ ,  $\mathcal{J}^* \rightarrow 0$  right away, Eq. (34) and Eq. (40) reduce to the Gor'kov mean-field equations for the uncorrelated model [33], with the equation Eq. (25) providing a self consistent determination of  $C_{ij}^*$  in terms of  $\mathcal{F}$ . Thus by neglecting the terms with  $\mathcal{K}$ ,  $\mathcal{K}^*$ , the role of the exchange  $J$  is confined to providing the lowest order electron–electron attraction in the Cooper channel. This amounts to neglecting the  $\mathcal{O}(J^2)$  dressings of the electron self energies and irreducible interaction i.e. the pairing kernel in Eq. (64). When retained, the normal state studies (see Ref. [31] Figs. (22,23,24-(a))) show that the self energy terms arising from  $J$  change the spectral functions of the model only slightly. Regarding the irreducible interaction in the superconducting channel, the  $\mathcal{O}(J)$  term is already attractive. Since we are in the regime of  $J \ll \max\{|t_{ij}|\}$  the retained term is expected to dominate the neglected higher order term. In summary, strong Hubbard–Gutzwiller type short ranged interactions renormalize the Greens function to  $\mathcal{G}$  from  $G_0$ , and the self energy terms due to  $J$  are minor [17,31]. The role of  $J$  is significant only insofar as it provides a mechanism for superconducting pairing, and potentially magnetic instabilities close to half filling. Keeping these considerations in mind, we drop the terms involving  $\mathcal{K}$ ,  $\mathcal{K}^*$ ,  $\mathcal{J}$ ,  $\mathcal{J}^*$  in Eq. (34) and Eq. (40). This suffices for our initial goal, of generalizing a Gor'kov type [33] mean-field treatment of Eq. (2) to the strongly correlated problem Eq. (1).

Multiplying the  $\gamma$  and  $\mathcal{D}$  terms, or equivalently the  $X$  and  $Y$  terms with  $\lambda$  and expanding the resulting equations systematically in this parameter constitutes the  $\lambda$ -expansion that we discuss below.

With these remarks in mind we make the following changes to Eq. (34) and Eq. (40):

- (i) We drop the terms proportional to  $\mathcal{J}$ ,  $\mathcal{J}^*$  and the corresponding derivative terms  $\mathcal{K}$ ,  $\mathcal{K}^*$ .

(ii) Defining the gap functions:

$$\Delta_{ij} = \frac{1}{2} J_{ij} C_{ij} \quad \text{and} \quad \Delta_{ij}^* = \frac{1}{2} J_{ij} C_{ij}^* \quad (41)$$

(iii) We scale the each occurrence of  $\gamma, X, Y, \widehat{X}, \widehat{Y}$  by  $\lambda$ .

With these changes we write the modified Eq. (34) and Eq. (40):

$$\begin{aligned} & (\mathbf{G}_{0i\sigma_j\sigma_j}^{-1} - \lambda Y_{i\sigma_j\sigma_j} - \lambda X_{i\sigma_j\sigma_j}) \mathcal{G}_{j\sigma_j f \sigma_f} \\ & = \delta(\tau - \tau') \delta_{if} (\delta_{\sigma_i\sigma_f} - \lambda \gamma_{\sigma_i\sigma_f}(i, \tau)) + \Delta_{ij} \mathcal{F}_{j\sigma_i f \sigma_f} \end{aligned} \quad (42)$$

$$\left( \widehat{\mathbf{G}}_{0i\sigma_j\sigma_j}^{-1} - \lambda \widehat{Y}_{i\sigma_j\sigma_j} - \lambda \widehat{X}_{i\sigma_j\sigma_j} \right) \mathcal{F}_{j\sigma_j f \sigma_f} = -\Delta_{ij}^* \mathcal{G}_{j\sigma_i f \sigma_f}, \quad (43)$$

where  $\sigma_j$  is summed over in both Eq. (42) and Eq. (43). Note that the self consistency condition Eq. (16) and Eq. (25) fix the correlation functions  $C$ 's in terms of  $\mathcal{F}$ . As  $\lambda \rightarrow 0$  we get back the meanfield equations of Gor'kov for the uncorrelated-J model. The  $\lambda$  parameter governs the density of doubly occupied states, and hence a series expansion in this parameter builds in Gutzwiller type correlations systematically. We expand the Greens functions to required order in  $\lambda$  and finally set  $\lambda = 1$ .

We write Eq. (42) and Eq. (43) symbolically as

$$(\mathbf{g}_0^{-1} - \lambda Y - \lambda X) \cdot \mathcal{G} = (\mathbf{1} - \lambda \gamma) + \Delta \cdot \mathcal{F} \quad (44)$$

$$(\widehat{\mathbf{g}}_0^{-1} - \lambda \widehat{Y} - \lambda \widehat{X}) \cdot \mathcal{F} = -\Delta^* \cdot \mathcal{G} \quad (45)$$

where the symbols  $\mathcal{G}, \mathcal{F}$  etc are regarded as matrices in the space, spin and time variables,  $\mathbf{1}$  is the Dirac delta function in time and a Kronecker delta in space and spin, with the dot indicating matrix multiplication or time convolution. In the case of  $X, \widehat{X}$  it also indicates taking the necessary functional derivatives.

### 3. Expansion of the equations in $\lambda$

We decompose both Greens functions in Eq. (44) and Eq. (45) as

$$\mathcal{G} = \mathbf{g} \cdot \widetilde{\mu}, \quad \mathcal{F} = \mathbf{f} \cdot \widetilde{\mu} \quad (46)$$

where  $\widetilde{\mu}$  is a function of spin, space and time that is common to both Greens function. As an example of the notation, the equation  $\mathcal{G} = \mathbf{g} \cdot \widetilde{\mu}$  stands for  $\mathcal{G}_{i\sigma_j\sigma_j}(\tau_i, \tau_j) = \sum_{k\sigma_k} \int_0^\beta d\tau_k \mathbf{g}_{i\sigma_j k\sigma_k}(\tau_i, \tau_k) \widetilde{\mu}_{k\sigma_k j\sigma_j}(\tau_k, \tau_j)$ . Here  $\widetilde{\mu}$  is called the caparison (i.e. a further dressing) function, in a similar treatment of the normal state Greens function. The terms  $\mathbf{g}$  and  $\mathbf{f}$  are called the auxiliary Greens function. The basic idea is that this type of factorization can reduce Eq. (44), to a canonical type equation for  $\mathbf{g}$ , where the terms  $\mathbf{1} - \lambda \gamma$  are replaced by  $\mathbf{1}$ . We remark that this is a technically important step since the term  $\mathbf{1} - \lambda \gamma$  modifies the coefficient of the delta function in time, and encodes the distinction between canonical and non-canonical fermions.

To simplify further, we note that  $X$  contains a functional derivative with respect to  $\nu$ , acting on objects to its right. When acting on a pair of objects, e.g.  $X \cdot \mathcal{G} = X \cdot \mathbf{g} \cdot \widetilde{\mu}$ , we generate two terms. One term is  $(X \cdot \mathbf{g}) \cdot \widetilde{\mu}$ , where the bracket, temporarily provided here, indicates that the operation of  $X$  is confined to it. The second term has the derivative acting on  $\widetilde{\mu}$  only, but the matrix product sequence is unchanged from the first term. We write the two terms together as

$$X \cdot \mathbf{g} \cdot \widetilde{\mu} = \overline{X \cdot \mathbf{g}} \cdot \widetilde{\mu} + X \cdot \overline{\mathbf{g} \cdot \widetilde{\mu}}, \quad (47)$$

so that the 'contraction' symbol refers to the differentiation by  $X$ , and the ' $\cdot$ ' symbol refers to the matrix structure. We may view this as the Leibnitz product rule.

Let us now operate with  $X$  on the identity  $\mathbf{g}\mathbf{g}^{-1} = \mathbf{1}$ , where  $\mathbf{g}^{-1}$  is the matrix inverse of  $\mathbf{g}$ . Using the Leibnitz product rule, we find

$$\overline{X}\cdot\mathbf{g} = -\left(\overline{X}\cdot\mathbf{g}\mathbf{g}^{-1}\right)\cdot\mathbf{g} \quad (48)$$

and hence we can rewrite Eq. (47) in the useful form

$$X\cdot\mathbf{g}\cdot\tilde{\mu} = -\left(\overline{X}\cdot\mathbf{g}\mathbf{g}^{-1}\right)\cdot\mathbf{g} + \overline{X}\cdot\mathbf{g}\cdot\tilde{\mu}. \quad (49)$$

With this preparation we rewrite Eq. (44) the equation for  $\mathcal{G}$  as

$$\left(\mathbf{g}_0^{-1} - \lambda Y + \lambda\left(\overline{X}\cdot\mathbf{g}\mathbf{g}^{-1}\right)\right)\cdot\mathbf{g}\cdot\tilde{\mu} = (\mathbf{1} - \lambda\gamma) + \Delta\mathbf{f}\cdot\tilde{\mu} + \lambda\overline{X}\cdot\mathbf{g}\cdot\tilde{\mu} \quad (50)$$

We now choose  $\mathbf{g}, \mathbf{f}$  such that

$$\left(\mathbf{g}_0^{-1} - \lambda Y + \lambda\left(\overline{X}\cdot\mathbf{g}\mathbf{g}^{-1}\right)\right)\cdot\mathbf{g} = \mathbf{1} + \Delta\mathbf{f}. \quad (51)$$

Substituting Eq. (51) into Eq. (50), we find that  $\tilde{\mu}$  satisfies the equation

$$\tilde{\mu} = (\mathbf{1} - \lambda\gamma) + \lambda\overline{X}\cdot\mathbf{g}\cdot\tilde{\mu}. \quad (52)$$

Note that Eq. (51) has the structure of a canonical equation since we replaced the  $\mathbf{1} - \lambda\gamma$  term by  $\mathbf{1}$  in Eq. (50). Thus the non-canonical Eq. (44) for  $\mathcal{G}, \mathcal{F}$  is replaced by a pair of canonical equations for  $\mathbf{g}, \tilde{\mu}$ . In Eq. (51) we note that the action of  $X$  is confined to the bracket  $\lambda\left(\overline{X}\cdot\mathbf{g}\mathbf{g}^{-1}\right)$ , unlike the term  $\lambda X\cdot\mathcal{G}$  in the initial Eq. (44). We may thus view the term in bracket in Eq. (51) as a proper self energy for  $\mathbf{g}$ .

For treating the equation for  $\mathcal{F}$  Eq. (45) we use the same scheme Eq. (46) and find

$$\widehat{X}\cdot\mathcal{F} = \widehat{X}\mathbf{f}\cdot\tilde{\mu} = -\left(\widehat{X}\cdot\mathbf{f}\mathbf{f}^{-1}\right)\cdot\mathbf{f}\cdot\tilde{\mu} + \widehat{X}\mathbf{f}\cdot\tilde{\mu} \quad (53)$$

With this we rewrite Eq. (45) after canceling an overall right multiplying factor  $\tilde{\mu}$

$$\left(\widehat{\mathbf{g}}_0^{-1} - \lambda\widehat{Y} + \lambda\widehat{X}\cdot\mathbf{f}\mathbf{f}^{-1}\right)\cdot\mathbf{f} = -\Delta^*\cdot\mathbf{g} + \lambda\widehat{X}\cdot\mathbf{f}\cdot\tilde{\mu}\cdot\tilde{\mu}^{-1} \quad (54)$$

Summarizing we need to solve for  $\mathbf{f}, \mathbf{g}, \tilde{\mu}, \Delta^*$  from Eqs. (51), (52), (54) by iteration in powers of  $\lambda$ .

### 3.1. Simplified equations near $T_c$

For the present work, we note that Eq. (54) simplifies considerably, if we work close to  $T_c$ . In this regime  $\mathbf{f}$  may be assumed to be very small, enabling us to throw away all terms of  $\mathcal{O}(f^2)$  and also to discard terms of  $\mathcal{O}(\lambda f)$ . This truncation scheme is sufficient to determine  $T_c$  for low orders in  $\lambda$ .

When  $T \sim T_c$ , throwing away terms of  $\mathcal{O}(f^2)$  and  $\mathcal{O}(\lambda f)$ , we obtain the simplified version of Eq. (54)

$$\mathbf{f} = -\widehat{\mathbf{g}}_0\cdot\Delta^*\cdot\mathbf{g} + o(\lambda\mathbf{f}), \quad (55)$$

so that Eq. (51) can be written as

$$\mathbf{g}^{-1} = \mathbf{g}_0^{-1} - \lambda Y + \lambda\left(\overline{X}\cdot\mathbf{g}\mathbf{g}^{-1}\right) + \Delta\cdot\widehat{\mathbf{g}}_0\cdot\Delta^* \quad (56)$$

In this limit the above two are the  $\mathcal{O}(\lambda^2)$  equations required to be solved, together with Eq. (52) and the self consistency condition Eq. (41), Eq. (25). The latter can be combined with Eq. (24) as

$$\Delta_{ij}^* = \frac{1}{2}J_{ij}C_{ij}^* = -\frac{1}{2}J_{ij}\sum_{\sigma}\mathcal{F}_{i\sigma,j\sigma}(\tau^+, \tau) \quad (57)$$

and further reduced using Eq. (46). On turning off the external potentials we regain time translation invariance. We next perform a fourier transform to fermionic Matsubara frequencies  $\omega_n = \frac{\pi}{\beta}(2n+1)$  using the definition  $\mathcal{F}(\tau) = \frac{1}{\beta} \sum_n e^{-i\omega_n \tau} \mathcal{F}(i\omega_n)$ , and write Eq. (46) in the frequency domain as

$$\mathcal{F}_{p\sigma}(i\omega_n) = \mathbf{f}_{p\sigma}(i\omega_n) \tilde{\mu}_{p\sigma}(i\omega_n). \quad (58)$$

Thus taking spatial fourier transforms with the definition

$$J(q) = 2J (\cos q_x + \cos q_y), \quad (59)$$

so that the self consistency condition Eq. (57) finally reduces to

$$\Delta^*(k) = -\frac{1}{2\beta} \sum_{p\sigma\omega_n} J(k-p) \mathbf{f}_{p\sigma}(i\omega_n) \tilde{\mu}_{p\sigma}(i\omega_n) \quad (60)$$

We may write Eq. (55) as

$$\mathbf{f}_{p\sigma}(i\omega_n) = -\widehat{\mathbf{g}}_{0\sigma}(p, i\omega_n) \Delta^*(p) \mathbf{g}_\sigma(p, i\omega_n) \quad (61)$$

where the time reversed free Greens function

$$\widehat{\mathbf{g}}_0(p, i\omega_n) = \frac{1}{-i\omega_n + \mu_0 - \varepsilon_{-p}} = \frac{1}{-i\omega_n - \xi_p} \quad (62)$$

with  $\xi = \varepsilon_p - \mu_0$  and by using  $\varepsilon_p = \varepsilon_{-p}$ , and  $\mu_0$  is taken as the non-interacting system chemical potential, discarding the corrections of  $\mu$  due to  $\lambda$ . Therefore Eq. (60) becomes

$$\Delta^*(k) = \frac{1}{2\beta} \sum_{p\sigma\omega_n} J(k-p) \widehat{\mathbf{g}}_{0\sigma}(p, i\omega_n) \Delta^*(p) \mathbf{g}_\sigma(p, i\omega_n) \tilde{\mu}_{p\sigma}(i\omega_n) \quad (63)$$

Here  $\mathbf{g}$  is taken from Eq. (56), i.e. the  $\mathcal{O}(\lambda^2)$  Greens function with a small correction (for  $T \sim T_c$ ) from the gap  $\Delta$ . Performing the spin summation and recombining  $\mathbf{g} \cdot \tilde{\mu} = \mathcal{G}$ , we get the equation in terms of the physical electron Greens function

$$\Delta^*(k) = \frac{1}{\beta} \sum_{p\omega_n} J(k-p) \Delta^*(p) \widehat{\mathbf{g}}_0(p, i\omega_n) \mathcal{G}(p, i\omega_n). \quad (64)$$

This is an important result of our formalism, it represents the leading order Gor'kov equation for the  $t$ - $J$  model. It is analogous to a refinement of Gor'kov's equation [33], usually called the Eliashberg equation [44], valid for strong electron-phonon coupling superconductivity. Our  $\lambda$  expansion plays the role of the Migdal theorem [45] in that problem. The analogy with Migdal [45] and Eliashberg's [44] work is only superficial, since the strongly correlated problem does not share the physics of the separation of the electronic and phonon time scales, underlying those results.

In Eq. (64) the physical electron Greens function  $\mathcal{G}$  is taken from the  $\mathcal{O}(\lambda^2)$  theory if we neglect the corrections from the gap, which vanishes above  $T_c$  anyway. We express the physical Greens function in terms of its spectral function  $A(p, \nu)$

$$\mathcal{G}(p, i\omega_n) = \int d\nu \frac{A(p, \nu)}{i\omega_n - \nu} \quad (65)$$

The frequency integral in Eq. (57) can be performed as

$$\frac{1}{\beta} \sum_{\omega_n} \widehat{\mathbf{g}}_0(p, i\omega_n) \mathcal{G}(p, i\omega_n) = \int d\nu A(p, \nu) \frac{1 - f(\nu) - f(\xi_p)}{\nu + \xi_p}. \quad (66)$$

where  $f$  is the fermi distribution  $f(\nu) = 1/(1 + \exp \beta\nu)$ . Hence

$$\Delta^*(k) = \sum_p J(k-p) \Delta^*(p) \int d\nu A(p, \nu) \frac{1 - f(\nu) - f(\xi_p)}{\nu + \xi_p}. \quad (67)$$

In summary this eigenvalue type equation for  $\Delta^*(k)$ , together with the spectral function  $A(p, \nu)$  determined from the  $\mathcal{O}(\lambda^2)$  Greens function in Eq. (56), gives the self-consistent gap near  $T_c$ . At

sufficiently high temperatures, i.e. in the normal state  $T > T_c$   $\Delta^*$  vanishes, so that  $A$  is independent of  $\Delta^*$ . In this case Eq. (67) reduces to a linear integral equation for  $\Delta^*$ . We may then determine  $T_c$  from the condition that the largest eigenvalue crosses 1. For this purpose we only need the normal state electron spectral function of the strongly correlated metal.

#### 4. Estimate of $T_c$

##### 4.1. Equation for determining $T_c$

The condition for obtaining a d-wave superconducting state is given by setting  $T = T_c^+$  in Eq. (67) writing  $\Delta^*(k) = \Delta_0(\cos k_x - \cos k_y)$ , using the normal state spectral function for  $A$  and canceling an overall factor  $\Delta_0(\cos k_x - \cos k_y)$ . Following these steps we get

$$1 = J \sum_p \{ \cos(p_x) - \cos(p_y) \}^2 \int dv \frac{1 - f(v) - f(\epsilon_p - \mu_0)}{v + \epsilon_p - \mu_0} A(p, v) \Big|_{T_c}. \quad (68)$$

Instead of working with Eq. (68), it is convenient to make a useful simplification for the average over angles. Since Eq. (68) is largest when  $\vec{p}$  is on the fermi surface, we factorize the two terms and write

$$1 = J \Psi(\mu_0) \Gamma \quad (69)$$

$$\Gamma = \sum_p \int dv \frac{1 - f(v) - f(\epsilon_p - \mu_0)}{v + \epsilon_p - \mu_0} A(p, v) \Big|_{T_c} \quad (70)$$

where  $\Gamma$  is a particle–particle type susceptibility. Here  $\Psi(\mu_0)$  is more correctly the weighted average of  $\{ \cos(p_x) - \cos(p_y) \}^2$  with a weight function that is the integrand in Eq. (70). We simplify it to the fermi surface averaged momentum space d-wavefunction

$$\Psi(\mu_0) = \frac{1}{n(\mu_0)} \sum_p \{ \cos(p_x) - \cos(p_y) \}^2 \delta(\epsilon_p - \mu_0) \quad (71)$$

where  $n(\epsilon)$  is the band density of states (DOS) per spin and per site, at energy  $\epsilon$ ,

$$n(\epsilon) = \frac{1}{N_s} \sum_p \delta(\epsilon_p - \epsilon). \quad (72)$$

Using this simplification and performing the angular averaging over the energy surface  $\epsilon_{\vec{p}} = \epsilon$  we write the (particle–particle) susceptibility  $\Gamma$  (Eq. (70)) as

$$\Gamma = \int d\epsilon \int dv n(\epsilon) A(\epsilon, v) \frac{1 - f(v) - f(\epsilon - \mu_0)}{v + \epsilon - \mu_0} \Big|_{T_c}. \quad (73)$$

where  $A(\epsilon, v)$  is the angle-averaged version of the spectral function  $A(p, v)$ . We estimate this expression below for the extremely correlated fermi liquid, by using a simple model for the spectral function  $A$ .

In Eq. (73) if we replace the spectral function  $A$  by the (fermi gas) non-interacting result  $A_0(\epsilon, \mu_0) = \delta(v - \epsilon + \mu_0)$ , we obtain the Gor'kov–BCS mean-field theory, where the susceptibility  $\Gamma$  reduces to  $\int d\epsilon n(\epsilon) \frac{\tanh \frac{1}{2} \beta_c (\epsilon - \mu_0)}{2(\epsilon - \mu_0)}$ . This is evaluated by expanding around the fermi energy, and utilizing the low T formula  $\int_0^{W_0} \frac{d\epsilon}{\epsilon} \tanh \frac{1}{2} \beta_c \epsilon \sim \log \left[ \frac{\zeta_0 W_0}{k_B T_c} \right]$ , where  $W_0$  is the half-bandwidth and  $\zeta_0 = 1.13387 \dots$ . Equating  $\Gamma$  to  $1/J\Psi(\mu_0)$  gives the d-wave superconducting transition temperature for the uncorrelated  $t$ - $J$  model

$$k_B T_c^{(un)} \sim 1.134 W_0 e^{-\frac{1}{g}}, \quad (74)$$

with the superconducting coupling constant

$$g = J\Psi(\mu_0)n(\mu_0). \quad (75)$$

#### 4.2. Model spectral function

We next use a simple model spectral function to estimate these integrals. It has the great advantage that we can carry out most integrations analytically and get approximate but closed form analytical expressions for  $T_c$ , which provide useful insights. The model spectral function contains the following essential features of strong correlations namely:

- A quasiparticle part with fermi liquid type parameters, where the quasiparticle weight  $Z$  goes to 0 at half filling  $n = 1$ , and
- A wide background.

The model spectral function used is in the spirit of Landau's fermi liquid theory [46–48] with suitable modifications due to strong correlation effects [17]. We take the spectral function as

$$A(\epsilon, \nu) = Z\delta(\nu - \frac{m}{m^*}\epsilon) + (1 - Z)\frac{1}{2W_0}\Theta(W_0 - |\nu|). \quad (76)$$

Here  $\Theta(x) = \frac{1}{2}(1 + \frac{x}{|x|})$ ,  $W_0$  the half-bandwidth  $\frac{m}{m^*}$  is the renormalized effective mass of the fermions, and  $Z$  is the fermi liquid renormalization factor. The first term is the quasiparticle part with weight  $Z$ , and second part represents the background modeled as an inverted square-well. Integration over  $\nu$  gives unity at each energy  $\epsilon$ .  $Z$  is chosen to reflect the fact that we are dealing with a doped Mott–Hubbard insulator so it must vanish at  $n = 1$ . For providing a simple estimate we use Gutzwiller's result [36,49]

$$Z = 1 - n. \quad (77)$$

The effective mass is related to  $Z$  and the  $k$ -dependent Dyson self energy  $\Sigma$  through the standard fermi liquid theory [46–48] formula

$$\frac{m}{m^*} = Z \times \left( 1 + \left. \frac{\partial \Sigma(\vec{k}, \mu)}{\partial \epsilon_k} \right|_{k_F} \right). \quad (78)$$

The Landau fermi liquid renormalization factor  $\frac{m}{m^*}$  can be inferred from heat capacity experiments provided the bare density of states is assumed known.

Using Eq. (76) in Eq. (73) and decomposing the susceptibility  $\Gamma$  into a quasiparticle and background part, the equation determining  $T_c$  is:

$$(\Gamma_{QP} + \Gamma_B) \Big|_{T \rightarrow T_c} = \frac{1}{J\Psi(\mu_0)} \quad (79)$$

$$\Gamma_{QP} = Z \int d\epsilon n(\epsilon) \frac{1 - f(\epsilon - \mu_0) - f(\frac{m}{m^*}(\epsilon - \mu_0))}{(\epsilon - \mu_0)(1 + \frac{m}{m^*})} \quad (80)$$

$$\Gamma_B = \frac{(1 - Z)}{2W_0} \int d\epsilon n(\epsilon) \int_{-W_0}^{W_0} d\nu \frac{1 - f(\epsilon - \mu_0) - f(\nu)}{(\epsilon - \mu_0) + \nu}. \quad (81)$$

Using the same approximations that lead to Eq. (74) the  $\Gamma_{QP}$  can be evaluated as

$$\Gamma_{QP} = \frac{Zn(\mu_0)}{1 + \frac{m}{m^*}} \int_0^{W_0} \frac{d\epsilon}{\epsilon} \left( \tanh \frac{\epsilon}{2k_B T} - \tanh \frac{\epsilon/m^*}{2k_B T} \right), \quad (82)$$

and hence at low enough  $T$  the estimate

$$\Gamma_{QP} \sim n(\mu_0) \frac{2Z}{1 + \frac{m}{m^*}} \log \left[ \frac{\zeta_0 W_0 \sqrt{\frac{m}{m^*}}}{k_B T} \right]. \quad (83)$$

Unlike the quasiparticle part with this  $\log T$  behavior at low  $T$ , the background part is nonsingular as  $T \rightarrow 0$ , since a double integral over the region of small  $\epsilon - \mu_0$  and  $\nu$  is involved. It can be estimated

by setting  $T = 0$ ,  $\epsilon - \mu_0 \sim \epsilon$  and replacing  $n(\epsilon) \sim n(\mu_0)$ . With

$$\Gamma_B \equiv n(\mu_0)\gamma_B, \quad (84)$$

$$\gamma_B = \frac{(1-Z)}{2W_0} \int_{-W_0}^{W_0} \int_{-W_0}^{W_0} \frac{1}{2} \frac{\text{sign}(\epsilon) + \text{sign}(\nu)}{\epsilon + \nu} d\epsilon d\nu. \quad (85)$$

Integrating this expression we obtain

$$\gamma_B = (1-Z) \log 4. \quad (86)$$

Combining Eqs. (79), (83), (86) we find

$$k_B T_c \sim 1.134 W_0 \times \sqrt{\frac{m}{m^*}} \times e^{-\frac{1}{g_{\text{eff}}}} \quad (87)$$

where the effective superconducting coupling:

$$g_{\text{eff}} = \frac{2Z}{\left(1 + \frac{m}{m^*}\right)} \{J_{\text{eff}} \Psi(\mu_0) n(\mu_0)\} \quad (88)$$

and an effective exchange

$$J_{\text{eff}} = \frac{J}{1 - \gamma_B J \Psi(\mu_0) n(\mu_0)}, \quad (89)$$

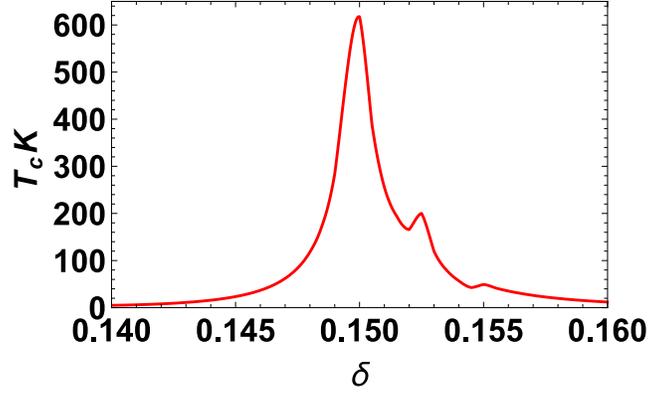
where the denominator represents an enhancement due to the background spectral weight. In comparing Eq. (87) with the uncorrelated result Eq. (74) several changes are visible. The bandwidth prefactor is reduced by correlations due to the factor of  $\sqrt{\frac{m}{m^*}} \ll 1$ . This factor vanishes as  $n \rightarrow 1$  thereby diminishing superconducting  $T_c$  in the close proximity of the insulator. A similar but even more drastic effect arises from multiplying factor  $\frac{2Z}{\left(1 + \frac{m}{m^*}\right)}$  in the coupling  $g_{\text{eff}}$  Eq. (88). This term reflects the quasiparticle weight in the pairing process, and since  $Z$  vanishes near the insulating state, it leads to an essential singularity in  $T_c$  as a function of hole density. Being situated in the exponential, it kills superconductivity much more effectively than the bandwidth prefactor. Away from the close proximity of the insulator other terms in  $g_{\text{eff}}$  become prominent, allowing for the possibility of superconductivity. Amongst them is the replacement of the exchange energy by  $J_{\text{eff}}$ . In a density range where  $\Psi(\mu_0) n(\mu_0)$  is appreciable, this enhances  $J_{\text{eff}}$  over  $J$  due to the feedback nature of Eq. (89), and has an important impact on determining the phase region with superconductivity.

#### 4.3. Numerical estimates of $T_c$

We turn to the task of estimating the order of magnitude of the  $T_c$  in this model. When we take typical values for cuprate systems:  $W_0 \sim 10^4$  K (i.e.  $\sim 1$  eV) and  $J \sim 10^3$  K (i.e.  $\sim 0.1$  eV), the transition temperature of the uncorrelated model  $T_c^{(um)}$  Eq. (74) is a few thousand K, at most densities. It remains robustly non-zero at half filling, since in this formula correlation effects are yet to be built in and the Mott–Hubbard insulator is missing. For the correlated system, we estimate  $T_c$  from Eq. (87) using similar values of model parameters. The terms arising from correlations in Eq. (87) are guaranteed to suppress superconductivity near the insulating state, since  $Z \rightarrow 0$  and the quasiparticle is lost. A more refined question is whether an intermediate density regime ( $\delta > 0$ ) can support superconductivity. And if so, whether the temperature scales are robust enough to be observable. Within the context and confines of the simplified model spectral function considered, we answer both questions positively here.

##### 4.3.1. Choice of model parameters

In order to estimate the order of magnitude of the  $T_c$ , its dependence on  $J$  and band parameters, we choose parameters similar to those used in contemporary studies for the single layer High  $T_c$  compound  $La_{2-x}Sr_xCuO_4$ . The hopping Hamiltonian  $-\sum_{ij} t_{ij} \tilde{C}_{i\sigma}^\dagger \tilde{C}_{j\sigma}$ , gives rise to band energy dispersion  $\varepsilon(\vec{k}) = -2t(\cos k_x + \cos k_y) - 4t' \cos k_x \cos k_y - 2t''(\cos 2k_x + \cos 2k_y)$  on a square lattice.



**Fig. 1.** The superconducting transition temperature for the correlated model  $T_c$  (Eq. (87)) ( $t'/t = -0.159$ ,  $t''/t = 0.01$ ,  $\frac{m}{m^*} = Z$ ). The scale of the maximum transition temperature is smaller by an order of magnitude from the uncorrelated model. As the insulator is approached  $\delta \rightarrow 0$ , and  $T_c$  decreases drastically. This is easy to understand since the quasiparticle weight  $Z$  shrinks on approaching the insulating state, killing the coupling  $g_{eff}$  Eq. (88). When  $\delta$  goes beyond the peak (optimum) value, the effective superconducting coupling  $g_{eff}$  again falls off as seen in Fig. 2 and in Fig. 5 due to the other factors in Eq. (88). When  $g_{eff}$  drops below  $\sim 0.12$ , the resulting  $T_c$  is negligible.

Thus the hopping amplitudes  $t_{ij}$  are equal to  $t$  when  $i, j$  are nearest neighbors,  $t'$  when  $i, j$  are second-nearest neighbors, and  $t''$  when  $i, j$  are third-nearest neighbors. For this system we will use the values [14,50]

$$t = 0.45 \text{ eV}, \quad t'/t = -0.16 \pm 0.02, \quad t''/t = .01. \quad (90)$$

This parameter set is roughly consistent with the experimentally determined fermi surface of  $La_{2-x}Sr_xCuO_4$  [50], we comment below on considerations leading to a more precise choice. The tight binding band extends from  $-W_0 \leq \epsilon \leq W_0$ , where  $W_0 = 4t$ , neglecting a small shift due to  $t'$ . The exchange energy is chosen to be

$$J/t = 0.3, \quad \text{or } J/k_B \sim 1550 \text{ K}, \quad (91)$$

as determined from two magnon Raman experiments [51] on the parent insulating  $La_2CuO_4$ . Note that the  $t$ - $J$  model is obtainable from the Hubbard model by performing a large  $U/t$  super-exchange expansion, giving  $J = \frac{4t^2}{U}$ . Thus our choice of  $J$  corresponds to a strong coupling type magnitude of  $U/t \sim 13.3$  in the Hubbard model, placing it in a perturbatively inaccessible regime of that model.

We now discuss the enhancement of effective mass  $\frac{m}{m^*}$  [52]. In the proximity of the Mott-Hubbard insulating state  $n \rightarrow 1$ , an enhancement in  $\frac{m}{m^*}$  is expected on general grounds, reflecting a diminished thermal excitation energy scale due to band narrowing. For illustrating the role of this parameter we use two complementary estimates

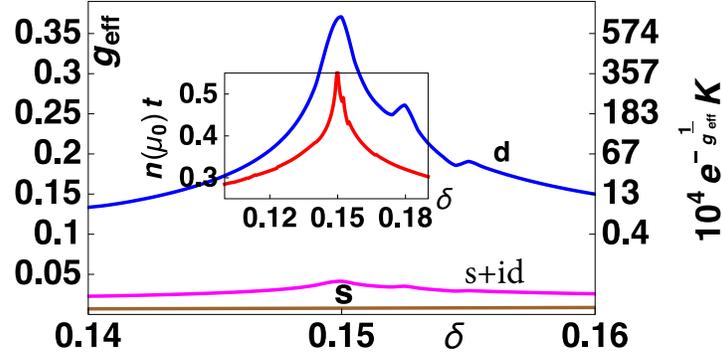
$$\frac{m}{m^*} \sim 3.4 (1 - n), \quad (a) \quad (92)$$

$$\frac{m}{m^*} = Z, \quad (b) \quad (93)$$

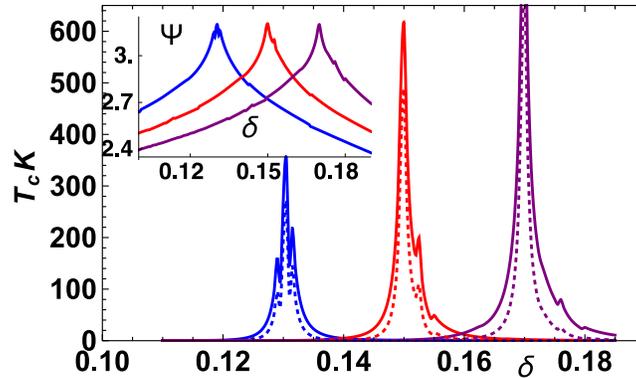
where estimate (a) gives a two-fold enhancement of  $\frac{m}{m^*}$  at  $\delta = .15$ , while estimate (b), obtained by neglecting the  $k$  dependence of the self energy in Eq. (78), gives a seven-fold enhancement. The formulas used are simple enough so that the effect of other estimates for  $\frac{m}{m^*}$  should be easy for the reader to gauge.

#### 4.3.2. Results

In Fig. 1 the superconducting transition temperature for d-wave symmetry is shown as a function of the hole density  $\delta = 1 - n$  where the band parameters are indicated in the caption. It shows that  $T_c$  is maximum at  $\delta \sim 0.15$  and falls off rapidly as one moves away from that density in either



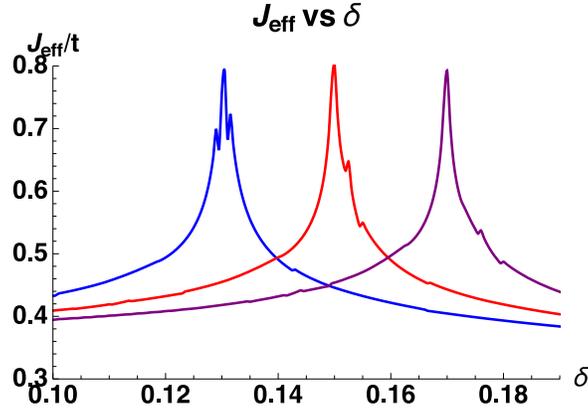
**Fig. 2.** The figure and inset use  $t'/t = -0.159$ ,  $t''/t = 0.01$ , and  $\frac{m}{m^*} = Z$ . The effective superconducting coupling  $g_{eff}$  (Eq. (88)) for three Cooper pair symmetries: (i) (blue) d-wavefunction  $\langle \{\cos(k_x) - \cos(k_y)\}^2 \rangle_{FS}$ , (ii) (brown) extended s-wavefunction  $\langle \{\cos(k_x) + \cos(k_y)\}^2 \rangle_{FS}$ , and (iii) (magenta)  $s + id$  wavefunction  $\langle \{\cos^2(k_x) + \cos^2(k_y)\} \rangle_{FS}$ . For the d-wavefunction, the drastic decrease of  $T_c$  on both sides of the peak values in Fig. 1 can be understood by referring to the second y-axis, giving the temperature scale  $T_c^{approx} = 10^4 \times e^{-\frac{1}{g_{eff}}}$  K. This scale provides an order of magnitude of  $T_c$  at a given  $g_{eff}$  by assuming a prefactor  $10^4$  K. It illustrates the rapid reduction of  $T_c$  when  $g_{eff} \lesssim 0.12$ . The other two symmetries lead to much smaller couplings and are therefore ineffective. **Inset:** The band DOS at the fermi energy shows an enhancement around the hole density  $\delta = 0.15$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** The superconducting transition temperature for the correlated model  $T_c$  (Eq. (87)) for three parameter sets – (i) (red)  $t'/t = -0.159$ ,  $t''/t = 0.01$  with  $\delta_{peak} = 0.15$ , (ii) (blue)  $t'/t = -0.137$ ,  $t''/t = .01$  with  $\delta_{peak} = 0.13$  and (iii) (purple)  $t'/t = -0.181$ ,  $t''/t = 0.01$  with  $\delta_{peak} = 0.17$ . The solid lines use  $\frac{m}{m^*} = Z$  and the dashed lines  $\frac{m}{m^*} = 3.4\delta$ . **Inset:**  $\Psi(\mu_0)$  the fermi surface averaged d-wavefunction  $\langle \{\cos(k_x) - \cos(k_y)\}^2 \rangle_{FS}$  is shown for the three sets of band parameters. The peak values occur at the densities where  $T_c$  is highest. Their peak magnitude  $\sim 3.2$  indicates a strong constructive interference effect from  $\vec{k} \sim \{\pm\pi, 0\}$ ,  $\{0, \pm\pi\}$ , where  $|\cos(k_x) - \cos(k_y)| \sim 2$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

direction. The scale of  $T_c$  is a few hundred K, which is an order of magnitude lower than that of the uncorrelated system. The small kink-like features to the right of the peak reflect structure in the DOS shown as inset in Fig. 2. In Fig. 2 the effective superconducting coupling  $g_{eff}$  is shown for three different symmetries of the Cooper pairs: d-wave, extended s-wave, and  $s + id$ -wave. It is clear that within this theory, only d-wave symmetry leads to robust superconductivity, the other two symmetries lead to effects too small to be observable. From Fig. 3 we see that the peak density is shifted by varying the band hopping parameters. As the peak density moves towards small  $\delta$ , its height falls rapidly. This is understandable as the effect of the quasiparticle weight  $Z$  in the formulas Eqs. (87), (88). We also note that the use of different expressions for the effective mass in Eqs. (92), (93) change the width of the allowed regions somewhat, but are quite comparable.

The inset in Fig. 3 displays the d-wavefunction averages corresponding to the same sets of parameters. It is interesting to note that the height of the peaks,  $\Psi_{max} \sim 3.2$ , is close to their upper



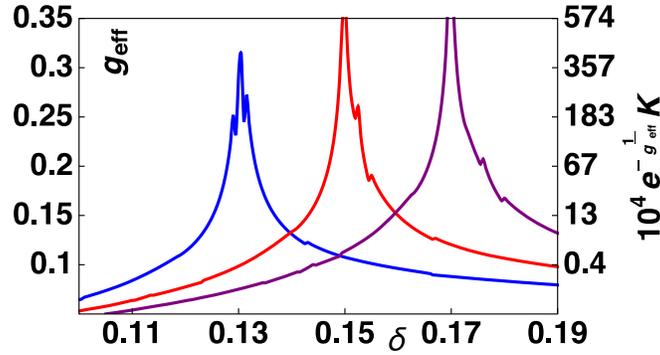
**Fig. 4.** The effective exchange  $J_{eff}$  from Eq. (89) for the three parameter sets – (i) (red)  $t'/t = -0.159$ ,  $t''/t = 0.01$  with  $\delta_{peak} = 0.15$ , (ii) (blue)  $t'/t = -0.137$ ,  $t''/t = 0.01$  with  $\delta_{peak} = 0.13$  and (iii) (purple)  $t'/t = -0.181$ ,  $t''/t = 0.01$  with  $\delta_{peak} = 0.17$ . Since we assumed  $J/t \sim 0.3$  (Eq. (89)),  $J_{eff}/t$  is considerably enhanced in the range of densities exhibiting high  $T_c$ . This enhancement in turn boosts up  $g_{eff}$ , via Eq. (88), and hence plays an important role in giving an observable magnitude of  $T_c$  in Figs. 1 and 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

bound 4, from a type of constructive interference that requires comment. Note first that the DOS can be expressed as a line integral in the octant of the Brillouin zone  $n(\mu_0) = \frac{2}{\pi^2} \int_0^{k_{max}} \frac{dk_x}{|v^y(k)|}$ , where the velocity  $v^y(\vec{k}) = 2 \sin(k_y)(t + 2t' \cos k_x + 4t'' \cos k_y)$  is evaluated with  $k_y \rightarrow k_y(k_x, \mu_0)$  on the fermi surface. Thus the region of small  $|v^y|$  dominates the integral. If  $v^y$  vanishes on the fermi surface, we get a (logarithmic van Hove) peak in the DOS. Now the average of  $\Psi(\mu_0)$  is largest, when  $k$  is close to  $\{\pm\pi, 0\}$  and  $\{0, \pm\pi\}$ . Therefore if the fermi surface passes through  $\{\pm\pi, 0\}$  and  $\{0, \pm\pi\}$  for an “ideal density”, then we simultaneously maximize the average of  $\Psi$ , and obtain a large DOS. The condition for this is found by equating the band energy at  $\{\pm\pi, 0\}$  to the chemical potential  $\mu_0 = 4t' - 4t''$ , thereby fixing the corresponding density  $\delta$ . It follows that a given  $\delta$  can be found from several different sets of the parameters  $t', t''$ . The inset of Fig. 3 shows the average  $\Psi(\mu_0)$  displays peaks, the middle one (red) coincides in location with the peak in the DOS in the inset of Fig. 2.

In Fig. 4 we illustrate the role of the feedback enhancement of the exchange  $J$  due to the background spectral function discussed in Eq. (89). For each set of parameters, there is a density region where both the DOS at the fermi energy and the averaged d-wavefunction are enhanced, and the confluence directly enhances  $J_{eff}$ . In turn this is reflected in the superconducting coupling  $g_{eff}$ . In Fig. 5 we see how the confluence of enhancements in the DOS and in the d-wavefunction  $\Psi(\mu_0)$ , further boosts the superconducting coupling  $g_{eff}$  and offsets to some extent the suppression due to a small magnitude of  $Z$ , as seen in Eq. (88). As a result of this competition  $T_c$  turns out to be in the observable range. The additional y-axis in Fig. 5 translates the superconducting coupling  $g_{eff}$  to an order of magnitude type transition temperature  $T_c^{appx} = e^{-1/g_{eff}} \times 10^4$  K. This scale helps us to understand why  $T_c$  falls off so rapidly when  $\delta$  increases beyond the peak value where the coupling  $g_{eff}$  falls below  $\sim 0.12$ , thereby rapidly suppressing  $T_c$ .

## 5. Conclusions

This work presents a new methodology for treating extremely correlated superconductors. The exact equations of the normal and anomalous Greens functions in the superconductor are derived. These are further expanded in powers of a control parameter  $\lambda$  related to the density of double occupancy, and the second order equations are given in Eqs. (51), (52), (54), together with the self consistency conditions Eqs. (25), (41). A further simplification is possible for  $T \sim T_c$  where the anomalous terms are small. This leads to a tractable condition for  $T_c$  given in Eq. (69), expressed in terms of the electron spectral function. Further analysis uses a model spectral function Eq. (76),



**Fig. 5.** The effective superconducting coupling  $g_{\text{eff}}$  (Eq. (88)) for the three curves in Fig. 3, with parameter sets – (i) (red)  $t'/t = -0.159$ ,  $t''/t = 0.01$  with  $\delta_{\text{peak}} = 0.15$ , (ii) (blue)  $t'/t = -0.137$ ,  $t''/t = .01$  with  $\delta_{\text{peak}} = 0.13$  and (iii) (purple)  $t'/t = -0.181$ ,  $t''/t = 0.01$  with  $\delta_{\text{peak}} = 0.17$ . The drastic decrease of  $T_c$  on both sides of the peak values in Fig. 3 can be understood by referring to the second y-axis, giving the approximate temperature scale  $T_c^{\text{appx}} = 10^4 \times e^{-\frac{1}{g_{\text{eff}}}}$  K. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

which is simple enough to yield an explicit expression for  $T_c$  in Eq. (87). More elaborate calculations should be feasible upon the availability of reliable spectral functions, when one may directly solve Eq. (68).

Our calculation delineates the regime of parameters where superconductivity is possible in the  $t$ - $J$  model within the ECFL theory. This regime turns out to be quite constrained. The calculation highlights the requirement of a substantial magnitude of the d-wavefunction average and the DOS at the fermi energy. It shows that  $T_c$  is maximal at a density where  $n(\mu_0)$ , the bare DOS is peaked, and is co-located with the peak of the fermi surface average of the d-wavefunction Eq. (71) (inset Fig. 3). The latter aspect is understandable, since the passing of the Fermi level energy dispersion through the zone boundary points  $\{\pm\pi, 0\}$ ,  $\{0, \pm\pi\}$ , promotes a peak in the DOS, and also leads to the maximization of  $\langle (\cos p_x - \cos p_y)^2 \rangle_{\text{FS}}$ . The prediction of a correlation between the peak in  $T_c$  with a peak in the d-wavefunction average is testable, since the latter is amenable to measurement using angle resolved photoemission.

In the approximation used here, the maximum  $T_c$  is nominally unbounded in a narrow density range here due to the logarithmic singularity of the DOS. It is expected to be cutoff to a finite value of  $\mathcal{O}(10^2 \text{ K})$  due to a more exact integration over energies, when using a reliable spectral function, in the place of the model used here. Such an integration would also supersede the Gor'kov-type approximation of expanding around the fermi surface ( $\int d\epsilon n(\epsilon) \sim n(\mu_0) \int d\epsilon$ ) employed here, thereby flattening out the sharp peak into a smoother shape. Finally this mean-field description of the superconductor is expected to be corrected by fluctuations of the phase, in a strictly two dimensional case, and by interlayer coupling, in the physically realistic case of a three dimensional system of weakly coupled layers.

In conclusion this work contains the essential outline of a new and controlled formalism to treat superconducting states of models with extremely strong correlations, such as the  $t$ - $J$  model. A transparent calculation within a low order approximation is presented here. It demonstrates that the exchange energy  $J$  can indeed provide the fundamental binding force between electrons forming Cooper pairs. It leads to superconductivity with  $T_c$ 's of  $\mathcal{O}(10^2 \text{ K})$ , in a finite range of densities located away from the insulator, as also experimentally found in cuprate superconductors.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

The work at UCSC was supported by the US Department of Energy (DOE), Office of Science, USA, Basic Energy Sciences, USA (BES), under Award No. DE-FG02-06ER46319.

## References

- [1] P.W. Anderson, *Solid State Physics*, vol. 12, Academic Press Inc., New York, 1963, p. 99, W. Kohn, *Phys. Rev.* 133 (1964) A171; A.B. Harris, R.V. Lange, *Phys. Rev.* 157 (1967) 295; W.F. Brinkman, T.M. Rice, *Phys. Rev. B* 2 (1970) 1324, (esp Eq. (4.1)); K.A. Chao, J. Spalek, A.M. Oles, *J. Phys. C* 10 (1977) L271.
- [2] F.C. Zhang, T.M. Rice, *Phys. Rev. B* 37 (1988) 3759; B.S. Shastry, *Phys. Rev. Lett.* 63 (1989) 1288.
- [3] J.G. Bednorz, K.A. Muller, *Z. Phys. B* 64 (1986) 188.
- [4] J.E. Hirsch, *Phys. Rev. Lett.* 54 (1985) 1317.
- [5] P.W. Anderson, *Science* 235 (1987) 1196.
- [6] G. Baskaran, Z. Zou, P.W. Anderson, *Solid State Commun.* 63 (1987) 973.
- [7] G. Kotliar, *Phys. Rev. B* 37 (1988) 3664.
- [8] C. Gros, *Ann. Physics* 189 (1989) 53.
- [9] H. Yokoyama, H. Shiba, *J. Phys. Soc. Japan* 57 (1988) 2482.
- [10] C.S. Hellberg, E.J. Mele, *Phys. Rev. Lett.* 67 (1991) 2080.
- [11] A. Paramekanti, M. Randeria, N. Trivedi, *Phys. Rev. Lett.* 87 (2001) 217002.
- [12] F.C. Zhang, T.M. Rice, *Phys. Rev. B* 37 (1988) 3759.
- [13] T. Giamarchi, C. Lhuillier, *Phys. Rev. B* 43 (1991) 12943.
- [14] M. Ogata, H. Fukuyama, *Rep. Progr. Phys.* 71 (2008) 036501.
- [15] T.-S. Huang, C.L. Baldwin, M. Hafezi, V. Galitski, 2020, arXiv:2004.10825.
- [16] S. Banerjee, T.V. Ramakrishnan, C. Dasgupta, *Phys. Rev. B* 83 (2011) 024510; *Phys. Rev. B* 84 (2011) 144525.
- [17] B.S. Shastry, *Phys. Rev. Lett.* 107 (2011) 056403, <http://physics.ucsc.edu/~sriram/papers/ECFL-Reprint-Collection.pdf>.
- [18] B.S. Shastry, *Ann. Physics* 343 (2014) 164–199.
- [19] F.J. Dyson, *Phys. Rev.* 102 (1956) 1217; S.V. Maleev, *Sov. Phys.—JETP* 6 (1958) 776.
- [20] T. Holstein, H. Primakoff, *Phys. Rev.* 58 (1940) 1098.
- [21] B.S. Shastry, *Phys. Rev. B* 87 (2013) 125124.
- [22] J.M. Luttinger, J.C. Ward, *Phys. Rev.* 118 (1960) 1417; J.M. Luttinger, *Phys. Rev.* 119 (1960) 1153.
- [23] B.S. Shastry, *Ann. Physics* 405 (2019) 155.
- [24] A. Damascelli, Z. Hussain, Z.-X. Shen, *Rev. Modern Phys.* 75 (2003) 473.
- [25] F.J. Dyson, *Phys. Rev.* 75 (1949) 486.
- [26] A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, *Rev. Modern Phys.* 68 (1996) 13; W. Metzner, D. Vollhardt, *Phys. Rev. Lett.* 62 (1989) 324.
- [27] R. Zitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, B.S. Shastry, *Phys. Rev. B* 88 (235) (2013) 132.
- [28] B.S. Shastry, E. Perepelitsky, *Phys. Rev. B* 94 (2016) 045138.
- [29] B.S. Shastry, P. Mai, *Phys. Rev. B* 101 (2020) 115121.
- [30] B.S. Shastry, P. Mai, *New J. Phys.* 20 (2018) 013027, arXiv:1703.08142.
- [31] P. Mai, B.S. Shastry, *Phys. Rev. B* 98 (2018) 205106, arXiv:1808.09788.
- [32] G.-H. Gweon, B.S. Shastry, G.D. Gu, *Phys. Rev. Lett.* 107 (2011) 056404.
- [33] L.P. Gor'kov, *Sov. Phys. JETP* 7 (1958) 505.
- [34] C.C. Tsuei, J.R. Kirtley, M. Rupp, J.Z. Sun, A. Gupta, M.B. Ketchen, C.A. Wang, Z.F. Ren, J.H. Wang, M. Bhusan, *Science* 271 (1996) 329.
- [35] D.J. Scalapino, *Phys. Rep.* 250 (1995) 329.
- [36] M. Gutzwiller, *Phys. Rev. Lett.* 10 (1963) 159.
- [37] J. Hubbard, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 277 (1964) 237.
- [38] S.G. Ovchinnikov, V.V. Val'kov, *Hubbard Operators in the Theory of Strongly Correlated Electrons*, Imperial College Press, London, 2004.
- [39] E. Perepelitsky, B.S. Shastry, *Ann. Physics* 357 (2015) 1.
- [40] S. Tomonaga, *Progr. Theoret. Phys.* 1 (1946) 27.
- [41] J. Schwinger, *Phys. Rev.* 74 (1948) 1439.
- [42] P.C. Martin, J. Schwinger, *Phys. Rev.* 115 (1959) 1342.
- [43] S. Engelsberg, *Phys. Rev.* 126 (1962) 1251.
- [44] G.M. Eliashberg, *Sov. Phys. JETP* 11 (1960) 696.
- [45] A.B. Migdal, *Sov. Phys. JETP* 34 (1958) 996.
- [46] L.D. Landau, *Sov. Phys. JETP* 30 (1956) 1058; L.D. Landau, *Sov. Phys. JETP* 32 (1956) 59.
- [47] P. Nozières, *Theory of Interacting Fermi Systems*, W. A. Benjamin, New York, 1964.

- [48] A.A. Abrikosov, L. Gor'kov, I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- [49] W.F. Brinkman, T.M. Rice, *Phys. Rev. B* 2 (1970) 4302.
- [50] T. Yoshida, T. Yoshida, X.J. Zhou, D.H. Lu, S. Komiya, Y. Ando, H. Eisaki, T. Kakeshita, S. Uchida, Z. Hussain, Z.-X. Shen, A. Fujimori, *J. Phys.: Condens. Matter* 19 (2007) 125209;  
M. Hashimoto, T. Yoshida, H. Yagi, M. Takizawa, A. Fujimori, M. Kubota, K. Ono, K. Tanaka, D.H. Lu, X., Z. -Shen, S. Ono, Y. Ando, *Phys. Rev. B* 77 (2008) 094516.
- [51] R.R.P. Singh, P.A. Fleury, K. Lyons, *Phys. Rev. Lett.* 62 (1989) 2736.
- [52] For this purpose we could, in principle use the Sommerfeld type low T formula for a fermi liquid. Estimating  $\frac{m}{m^*}$  from data turns out to be somewhat problematic. One must not only choose the bare band hopping parameters, but also neglect features of the data, such as the strong  $T$  dependence of the Sommerfeld coefficient  $\gamma$  in the data for  $La_{2-x}Sr_xCuO_4$ , as reported in, J.W. Loram, J. Luo, J.R. Cooper, W.Y. Liang, J.L. Tallon, *J. Phys. C Solids* 62 (2001) 59.