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 ω, 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

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 $U \to \infty$ 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 $1/S \to 0$ corresponds to the limit of free-fermion limit $\lambda \to 0$. Continuity in λ 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 λ, 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 $d \to \infty$ 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.

 ©2016 American Physical Society

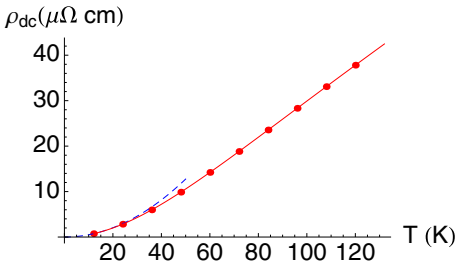

FIG. 1. ρ_{dc} on absolute scale vs T in Kelvin for particle density $n = 0.85$. We have used the estimates $D = 12\,000$ K and $\rho_0 = 258\ \mu\Omega$ cm. The latter is obtained by using [57] $\rho_0 \approx \frac{ha_0 d}{e^2}$, where we estimate $a_0 d \approx 10^{-8}$ cm. The Fermi liquid behavior with quadratic resistivity in the blue dotted line breaks down above $T_{\rm FL} \approx 30$ 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 $0.7 \leqslant n \leqslant 1$. 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 ω, 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 ω, 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

$$\mathcal{G}(k,i\omega_n) = \mathbf{g}(k,i\omega_n) \times \widetilde{\mu}(k,i\omega_n), \tag{1}$$

a product of the auxiliary Green's function $\mathbf{g}$ and the "caparison" function [44] given in terms of a second self-energy $\Psi(k,i\omega_n)$ and the particle density n as $\widetilde{\mu}(k,i\omega_n) = \{1 - n/2 + \Psi(k,i\omega_n)\}$. The auxiliary Green's function $\mathbf{g}(k,i\omega_n)$ given by

$$\mathbf{g}(k,i\omega_n) = \frac{1}{i\omega_n + \boldsymbol{\mu} - \{1 - n/2\}\,\varepsilon_k - \Phi(k,i\omega_n)}, \tag{2}$$

where $\boldsymbol{\mu}$ is the chemical potential and ε_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

$$\Psi(k,i\omega_n) = \Psi(i\omega_n), \tag{3}$$

$$\Phi(k,i\omega_n) = \chi(i\omega_n) + \varepsilon_k \Psi(i\omega_n), \tag{4}$$

where both Ψ and χ 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 $i\omega_n \to \omega + i0^+$, we may express the Green's function in the standard Dyson representation

$$\mathcal{G}(k,\omega + i0^+) = \frac{1}{\omega + i0^+ + \boldsymbol{\mu} - \varepsilon_k - \Sigma(\omega + i0^+)}, \tag{5}$$

where the Dyson self-energy is now manifestly momentum independent, and given by

$$\Sigma(\omega + i0^+) = \boldsymbol{\mu} + \omega + \frac{\chi(\omega + i0^+) - \boldsymbol{\mu} - \omega}{1 - \frac{n}{2} + \Psi(\omega + i0^+)}. \tag{6}$$

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 λ expansion to compute the two self-energies Ψ and χ. This λ expansion is described in detail in [2,46,47], in brief the parameter λ 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 Ψ and χ in a formal power series in λ, with terms that are functionals of $\mathbf{g}$ and the band energies ε_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 λ, 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

$$\mathcal{G}(k,i\omega_n) = \mathbf{g}(k,i\omega_n)\{a_G + \lambda\Psi(k,i\omega_n)\}, \tag{7}$$

$$\mathbf{g}^{-1}(k,i\omega_n) = i\omega_n + \boldsymbol{\mu}' - \lambda\,\chi(k,i\omega_n) - \{a_G + \lambda\Psi(k,i\omega_n)\}\left(\varepsilon_k - \frac{u_0}{2}\right), \tag{8}$$

with

$$a_G = 1 - \lambda\mathcal{G}(j,j^-) = 1 - \lambda\sum_k \mathcal{G}(k,i\omega_n)\,e^{i\omega_n 0^+}, \tag{9}$$

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 $j^- \equiv (\vec{R}_j,\tau_j + i\,0^-)$. The two self-energy functions Ψ and χ are expanded formally in λ as $\Psi = \Psi_{[0]} + \lambda\Psi_{[1]} + \cdots$ and $\chi = \chi_{[0]} + \lambda\chi_{[1]} + \ldots$. A systematic expansion in λ 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

$$\Psi_{[1]}(k) = -\sum_{pq}(\varepsilon_p + \varepsilon_q - u_0)\mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(p+q-k), \tag{10}$$

$$\chi_{[1]}(k) = -\sum_{pq}\left(\varepsilon_{p+q-k} - \frac{u_0}{2}\right)(\varepsilon_p + \varepsilon_q - u_0) \times \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(p+q-k). \tag{11}$$

In view of the explicit factors of λ in Eqs. (7) and (8), this leads to an $O(\lambda^2)$ approximation for $\mathcal{G}$; the recipe further requires that the parameter λ 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 λ, namely, $t_{ij} \to t_{ij} + c\,\delta_{ij}$ with an arbitrary constant c. For instance, we see in Eq. (11) that a shift of the energies $\varepsilon_k \to c + \varepsilon_k$ 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 $\mathcal{G}$:

$$\sum_k \mathbf{g}(k)\,e^{i\omega_n 0^+} = \frac{n}{2} = \sum_k \mathcal{G}(k)\,e^{i\omega_n 0^+}. \tag{12}$$

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

$$a_G^{(I)} \to 1 - \lambda\frac{n}{2}, \tag{13}$$

reducing to the exact answer $a_G^{\text{exact}} = 1 - \frac{n}{2}$ as $\lambda \to 1$.

Alternately, we could expand the $\mathcal{G}$ in powers of λ, a procedure we followed in [6,49]. We expanded $\mathcal{G}$ out to first order in λ 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

$$a_G^{(II)} = 1 - \lambda\frac{n}{2} + \lambda^2\frac{n^2}{4} + O(\lambda^3). \tag{14}$$

Setting $\lambda \to 1$, we thus get two alternate approximate skeleton versions of Eq. (7):

$$\mathcal{G}^{(I)}(k,i\omega_n) = \mathbf{g}^{(I)}(k,i\omega_n)\{1 - n/2 + \Psi(k,i\omega_n)\}, \tag{15}$$

$$\mathcal{G}^{(II)}(k,i\omega_n) = \mathbf{g}^{(II)}(k,i\omega_n) \times \{1 - n/2 + n^2/4 + \Psi(k,i\omega_n)\}, \tag{16}$$

where both expressions involve *the same* approximate Ψ given in Eq. (10), and the auxiliary $\mathbf{g}^{(\cdots)}$ 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 χ and ψ, has the correct functional form. Moreover, the unusual and important feature of particle-hole asymmetry, i.e., the presence of a strong ω^3 term in the Im Σ, 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 $n \to 1$, as one expects in a Mott insulator.

Within the spirit of Eq. (16) one might expect that further approximations involving higher-order terms in λ 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 λ 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 $Z^{(II)}$ 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,

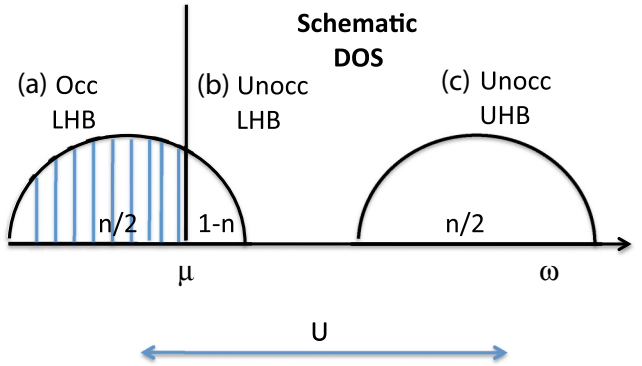

FIG. 2. A schematic depiction of the local spectral density of states $\rho_{GLocal}(\omega)$ [popularly called $A_{Local}(\omega)$] for the large-U Hubbard model, where the correlation split Hubbard bands are clearly separated. It shows three regions: (a) occupied electronic states, (b) unoccupied lower 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., $\rho_G(\omega) \equiv \sum_k \rho_G(\vec{k},\omega)$. The local Green's function itself is given by

$$G(\omega + i0^+) = \int d\nu \frac{\rho_G(\nu)}{\omega - \nu + i0^+}, \tag{17}$$

and so the $\omega \to \infty$ asymptotic behavior is determined by this sum rule as $G(\omega) \to \frac{\int d\nu \rho_G(\nu)}{\omega} = \frac{1}{\omega}$. This can be partitioned into three sum rules as depicted in Fig. 2:

$$\int_{-\infty}^{0} d\omega\, \rho_G(\omega) = n/2, \quad \int_{0}^{\Omega_*} d\omega\, \rho_G(\omega) = 1 - n,$$
$$\int_{\Omega_*}^{\infty} d\omega\, \rho_G(\omega) = n/2, \tag{18}$$

where Ω_* 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 $\rho_{\mathcal{G}}(\omega)$ satisfies

$$\int_{-\infty}^{0} d\omega\, \rho_{\mathcal{G}}(\omega) = n/2, \quad \int_{0}^{\infty} d\omega\, \rho_{\mathcal{G}}(\omega) = 1 - n, \tag{19}$$

where the upper Hubbard band (and Ω_*) 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 $\omega \to \infty$ asymptotic form $\lim_{\omega\to\infty} \mathcal{G}(\omega) \to \frac{1-n/2}{\omega}$, and gives us a relation of importance to this study:

$$\left(\lim_{\omega\to\infty} \mathcal{G}(\omega) \to \frac{1 - n/2}{\omega}\right) \leftrightarrow \left(\int_0^{\infty} d\omega\, \rho_{\mathcal{G}}(\omega) = 1 - n\right). \tag{20}$$

To see its relevance, we note that as $n \to 1$, 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 $\lim_{\omega\to\infty} (\Psi(\omega), \chi(\omega)) \to 0$ and also $\lim_{\omega\to\infty} \mathbf{g}(\omega) \to \frac{1}{\omega}$, we combine these to obtain

$$\lim_{\omega\to\infty} \mathcal{G}^{(II)}(\omega) \to \frac{1 - n/2 + n^2/4}{\omega},$$

whereby the unoccupied region $\int_0^{\infty} d\omega\, \rho_{\mathcal{G}^{(II)}}(\omega) = 1 - n + n^2/4$, in conflict with the condition (20) for a vanishing Z, as $n \to 1$.

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 $n \to 1$; 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

$$\mathbf{g}^{-1}(\epsilon,z) = z + \boldsymbol{\mu}' - (\epsilon - u_0/2)\left(1 - \frac{n}{2} + \Psi_{[1]}(z)\right) - \chi_{[1]}(z), \tag{21}$$

$$\mathbf{g}_{Loc,m}(z) = \int d\epsilon\, D(\epsilon)\mathbf{g}(\epsilon,z)\epsilon^m = \int d\nu \frac{\rho_{\mathbf{g}L,m}(\nu)}{z-\nu}. \tag{22}$$

The chemical potential $\boldsymbol{\mu}'$ absorbs all constants such as $\chi_{[0]}$, leading to

$$\boldsymbol{\mu} = \boldsymbol{\mu}' + \frac{u_0}{2}\left(1+\frac{n}{2}\right) - \int d\omega\, f(\omega)\rho_{\mathbf{g}L,1}(\omega), \tag{23}$$

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

$$\rho_{\mathbf{g}L,m}(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\mathbf{g}_{Loc,m}(\omega + i0^+). \tag{24}$$

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 $\Sigma(z) = z + \boldsymbol{\mu} - \varepsilon - \mathcal{G}^{-1}(\epsilon,z)$. We define its local version $\mathcal{G}_{Loc}(\omega)$ and its density through a band integration

$$\mathcal{G}_{Loc,m}(z) = \int d\epsilon\, D(\epsilon)\epsilon^m\, \mathcal{G}(\epsilon,z),$$
$$\rho_{\mathcal{G}L,m}(\omega) = -\frac{1}{\pi}\mathrm{Im}\mathcal{G}_{Loc,m}(\omega + i0^+). \tag{25}$$

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 Ψ,χ. These can in turn be found from expressions involving the fundamental convolution:

$$\rho^{(\mathcal{I})}_{abc}(u) = \int_{u_1,u_2,u_3} \delta(u+u_3-u_1-u_2)\{f(u_1)f(u_2)\bar{f}(u_3) + \bar{f}(u_1)\bar{f}(u_2)f(u_3)\}\rho_{\mathbf{g}L,a}(u_1)\rho_{\mathbf{g}L,b}(u_2)\rho_{\mathbf{g}L,c}(u_3), \tag{26}$$

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

$$\mathcal{I}_{abc}(z) = \mathcal{P}\int d\nu\, \frac{\rho^{(\mathcal{I})}_{abc}(\nu)}{z-\nu}. \tag{27}$$

From this object, the two self-energies can be found as the combinations

$$\Psi_{[1]}(z) = 2\mathcal{I}_{010}(z) - u_0\,\mathcal{I}_{000}(z),$$
$$\chi_{[1]}(z) = 2\mathcal{I}_{011}(z) - u_0[\mathcal{I}_{010}(z) + \mathcal{I}_{001}(z)] + \frac{u_0^2}{2}\mathcal{I}_{000}(z). \tag{28}$$

In summary, we can compute $\mathbf{g}$ in terms of χ,Ψ from Eq. (21). Having done so, we compute χ,Ψ in terms of the $\mathbf{g}$ from Eq. (28), thus defining the second part of the loop. The two chemical potentials $\boldsymbol{\mu}$ and u_0 are found from Eq. (23) and the two particle-number sum rules:

$$\int d\omega f(\omega)\,\rho_{\mathbf{g}L,0}(\omega) = \frac{n}{2}, \tag{29}$$
$$\int d\omega f(\omega)\,\rho_{\mathcal{G}L,0}(\omega) = \frac{n}{2}, \tag{30}$$

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 $n \to 1$ 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

$$\boldsymbol{\mu} \sim k_B T \ln\{n/[2(1-n)]\}, \tag{31}$$

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 $(\partial S/\partial N)_T = -(\partial\boldsymbol{\mu}/\partial T)_N$, and the initial condition $S(n\to 0) = 0$, we find

$$S \sim -k_B N_s\{n\ln n/2 + (1-n)\ln(1-n)\}, \tag{32}$$

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 $T \to 0$, or high $T \gg D$ with moderated densities $n \geqslant 0.7$, 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 $d\boldsymbol{\mu}/dT$ 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 $T \to 0$ 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 λ 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 λ expansion is underway currently, and we expect to present the details in a forthcoming paper. However, at the 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 $\chi \to 0$ and $\Psi \to 0$) 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 $-3D \leqslant \omega \leqslant 3D$ 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 $|\omega| \lesssim \frac{D}{3}$

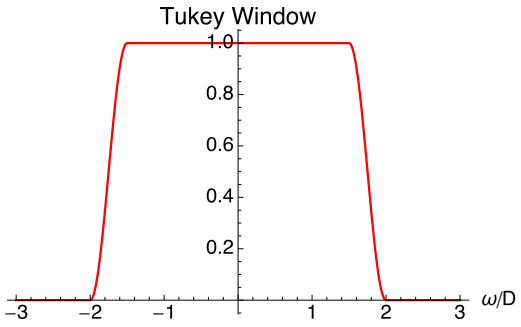

FIG. 3. Multiplication through the Tukey window $\mathbf{W}_T(\omega)$ [Eq. (34)] is used for providing a cutoff in our scheme (33). It is applied only to the auxiliary local Green's function $\rho_{\mathbf{g}L,m}(\omega)$, while the physical spectral functions $\rho_{\mathcal{G}L}(\omega)$ are unconstrained, apart from an overall window $|\omega| \leqslant 5D$ 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:

$$\hat{\rho}_{\mathbf{g}L,m}(\omega) = \frac{1}{\mathcal{V}}\, \rho_{\mathbf{g}L,m}(\omega)\mathbf{W}_T(\omega), \tag{33}$$

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 $|\omega| \leqslant \Omega_c^{(-)}$, 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)

$$\begin{aligned}
\mathbf{W}_T(\omega) &= 1 \quad \text{for} \quad \Omega_c^{(-)} \geqslant |\omega| \\
&= \frac{1}{2}\left(1 + \sin\left\{\pi/2\, \frac{\Omega_c^{(+)} + \Omega_c^{(-)} - 2|\omega|}{\Omega_c^{(+)} - \Omega_c^{(-)}}\right\}\right) \\
&\quad \text{for } \Omega_c^{(+)} \geqslant |\omega| \geqslant \Omega_c^{(-)} \\
&= 0 \quad \text{for } |\omega| > \Omega_c^{(+)}.
\end{aligned} \tag{34}$$

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 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 $\rho_{\mathcal{G}}$ is computed as per the rules without any further assumptions. It typically

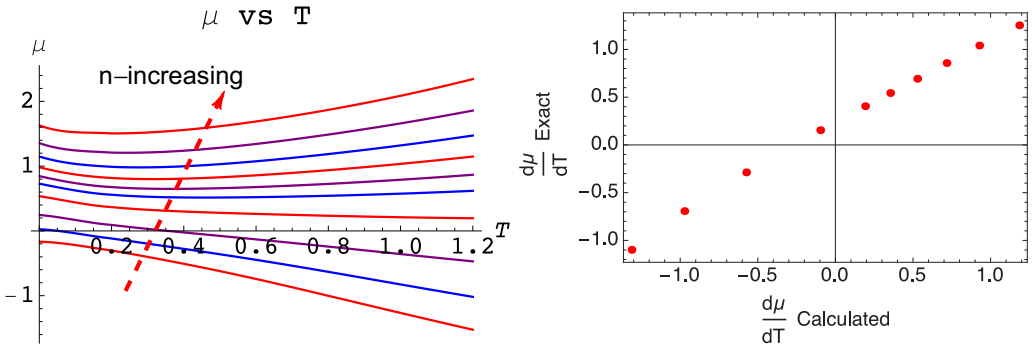

FIG. 4. Left: The chemical potential at particle densities $n = 0.4,0.5,0.6,0.7,0.75,0.775,0.8,0.825,0.85,0.875$ increasing from bottom to top. We confine ourselves to the limited regime $T \leqslant 1.2$ since higher T requires further adjustment of the cutoffs. Note that within this regime, the $\mu(T)$ curve turns around at a density around $n \sim 0.7$. For lower densities, μ decreases monotonically with increasing T, whereas at higher densities we have a shallow minimum followed by a regime of rising μ. 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 Ω_* 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 μ 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 μ are quite reasonable in the hole-rich region $n \geqslant 0.75$ (i.e., $\delta \leqslant 0.25$) 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 $\delta \to 0$. 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 δ. 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 $|\omega| \leqslant ZD$ through a Fermi liquid form with the expected particle-hole asymmetric correction [8]

$$-\mathrm{Im}\Sigma(\omega) = \frac{\omega^2}{\Omega_0}\left(1 - \frac{\omega}{\Delta}\right), \tag{35}$$

whereby introducing two energies: Ω_0, which determines the magnitude of $\mathrm{Im}\Sigma$ and Δ the asymmetry scale. In [6] and also in [5], it was pointed out that Ω_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 Ω_0 is computed by averaging $\mathrm{Im}\Sigma(\omega)$ in the domain $|\omega| \leqslant ZD$. In the bottom right panel of Fig. 7, we show the variation of Ω_0 versus Z^2 and in the inset with δ^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 Ω_0 with Z^2, rather than δ^2 at the lowest δ. 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 $-\mathrm{Im}\Sigma$ versus ω/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 $\mathrm{Im}\Sigma$ into symmetric and

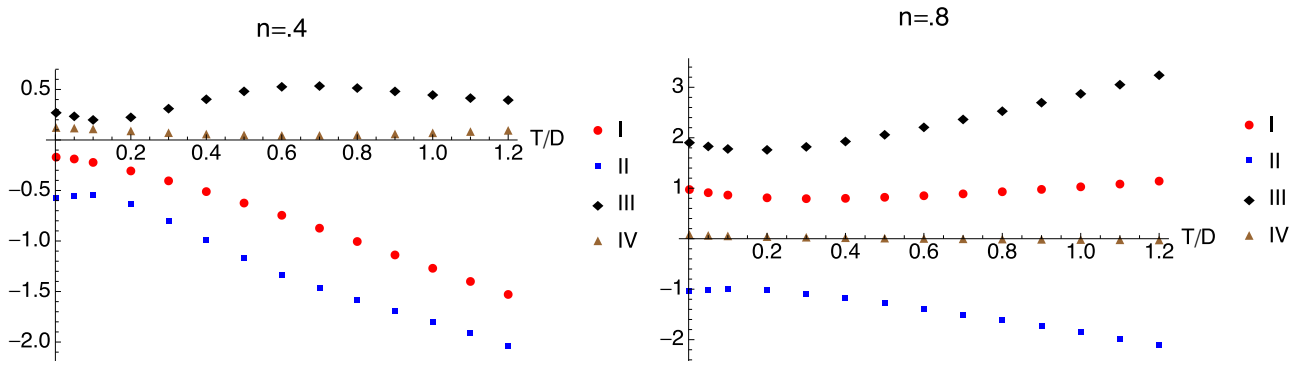

FIG. 5. The T dependence of the chemical potential $\boldsymbol{\mu}$ and its three additive contributions from Eq. (23) at two densities. The physical chemical potential $\boldsymbol{\mu}$ (I, red), the auxiliary part: $\boldsymbol{\mu}'$ (II, blue), the u_0 contribution: $(1/2 + n/4)u_0$ (III, purple), and the small part from the integral $-\int f\rho_{\mathbf{g}L,1}$ (IV, magenta). The observed upturn in $\boldsymbol{\mu}$ 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 Δ in Eq. (35). We find that Δ is proportional to Z again, but with a weak density-dependent correction:

$$\Delta(\delta) = Z(\delta)\{3.38 - 15.6\delta + 27.1\delta^2\}. \tag{36}$$

The region beyond the straight line is captured on average, by extending Eq. (35) to

$$-\,\mathrm{Im}\Sigma(\omega) = \frac{\omega^2}{\Omega_0}\left(1 - \frac{\omega}{\Delta\sqrt{1 + 2\,\omega^2/Z^2}}\right). \tag{37}$$

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 ω where the spectral functions and self-energy vary, shrinks rapidly with increasing n; this is indirectly a reflection of variation of the Z with density in Fig. 6 since the scale of variation of Σ is set by Z. We also note that the spectral asymmetry in $\mathrm{Im}\Sigma$ is very clearly visible here.

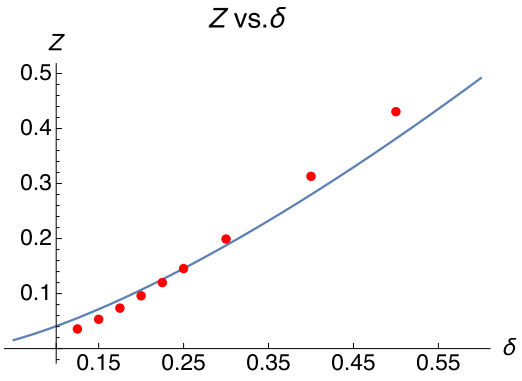

FIG. 6. The computed quasiparticle weight Z (dots) versus the hole density $\delta = 1 - n$, compared with the exact numerical results from DMFT ([6] solid curve), which fits very well to the formula $Z \sim \delta^{1.39}$. We see that the present scheme accounts well the suppression of Z near $\delta \sim 0$, 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.

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 $\mathrm{Im}\Sigma$ at small ω, 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

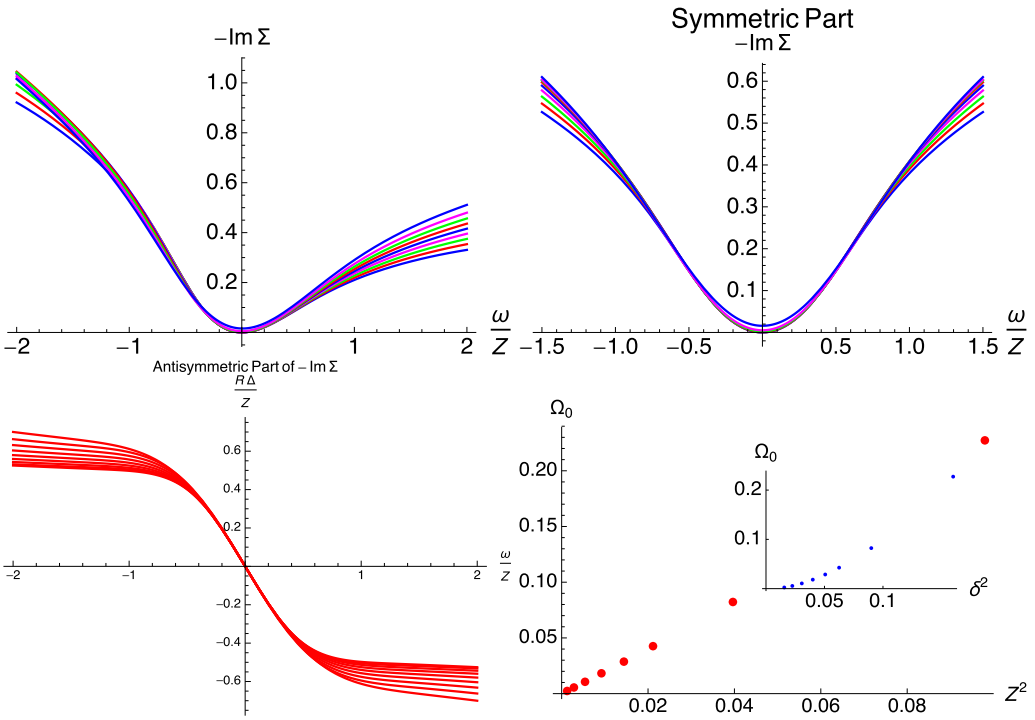

FIG. 7. Top left: $-\mathrm{Im}\Sigma(\omega)$ versus ω/Z at several densities $n = 0.7, 0.725, 0.75, 0.775, 0.8, 0.825, 0.85, 0.875, 0.9$ from bottom to top. We see that the frequency dependence scales well with Z, with better behavior on the occupied side $\omega \leqslant 0$. Top right: The symmetrized 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 Δ/Z at the above densities versus ω/Z, with $n = 0.9$ at the top and $n = 0.7$ at the bottom for $\omega \leqslant 0$. These collapse to a straight line with slope -1 in the range $|\omega| \leqslant Z$, provided we allow for an additional mild density dependence of the ratio Δ/Z, as in Eq. (36). Bottom right: The energy scale Ω_0 [Eq. (35)] determining the magnitude of the $\mathrm{Im}\Sigma$ at $T = 0$ is shown versus Z^2, and in the inset versus the hole density δ^2. Here, Ω_0 is seen to scale better with Z^2 rather than with δ^2.

arising from $d \to \infty$:

$$\sigma_{DC} = \frac{2\pi\hbar e^2}{V} \sum_k \left(v_k^x\right)^2 \int d\omega\, (-\partial f/\partial\omega)\, \rho_{\mathcal{G}}^2(\epsilon_k, \omega), \tag{38}$$

where the band velocity is given as $\hbar v_k^x = \partial\varepsilon_k/\partial k_x$. We wrap the velocity into a useful function

$$\begin{aligned}\Phi(\epsilon) &= \frac{1}{a_0}\frac{1}{N_s}\sum_k \delta(\varepsilon - \varepsilon_k)\left(v_k^x\right)^2/a_0^2 \\ &= \frac{1}{a_0} D(\epsilon)\left\langle \frac{(v_k^x)^2}{a_0^2}\right\rangle_{\varepsilon_k=\epsilon},\end{aligned} \tag{39}$$

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

$$\frac{\Phi(\epsilon)}{\Phi(0)} = \Theta(1 - \epsilon^2/D^2)\ \sqrt[3/2]{1 - \epsilon^2/D^2}, \tag{40}$$

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

$$\sigma_{DC} = \sigma_0 2\pi D \iint d\epsilon\, d\omega\, (-\partial f/\partial\omega)\left(\frac{\Phi(\epsilon)}{\Phi(0)}\right)\rho_{\mathcal{G}}^2(\epsilon,\omega). \tag{41}$$

We write the (inverse) Green's function at real ω as

$$\mathcal{G}_\pm^{-1}(\epsilon,\omega) = A(\omega) - \epsilon \pm i B(\omega), \tag{42}$$

where the retarded case corresponds to $\mathcal{G}_+$, and

$$\begin{aligned}A(\omega,T) &= \omega + \boldsymbol{\mu} - \mathrm{Re}\ \Sigma(\omega,T), \\ B(\omega,T) &= \pi\rho_\Sigma(\omega,T) = -\mathrm{Im}\ \Sigma(\omega,T),\end{aligned} \tag{43}$$

and Σ is the Dyson self-energy. Setting $D = 1$ and using the identities $\rho_{\mathcal{G}} = i/(2\pi)(\mathcal{G}_- - \mathcal{G}_+)$ and $\mathcal{G}_\pm^2 = \partial_\epsilon\mathcal{G}_\pm$, and further integrating by parts over ϵ we obtain

$$\begin{aligned}\sigma &= \sigma_0 \int d\omega\, (-\partial f/\partial\omega)\xi(\omega), \\ \xi(\omega) &= \frac{1}{2\pi}\int d\epsilon \left\{\frac{i}{B}(\mathcal{G}_+ - \mathcal{G}_-)\frac{\Phi(\epsilon)}{\Phi(0)} + (\mathcal{G}_+ + \mathcal{G}_-)\frac{\Phi'(\epsilon)}{\Phi(0)}\right\}.\end{aligned} \tag{44}$$

Using the explicit form of Φ and $\mathcal{G}_\pm$, we reexpress ξ exactly as

$$\xi(\omega) = \frac{1}{\pi}\int_{-1}^{1} d\epsilon\ \sqrt{1-\epsilon^2}\,\frac{1 - 3\epsilon A + 2\epsilon^2}{B^2 + (A-\epsilon)^2}. \tag{45}$$

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

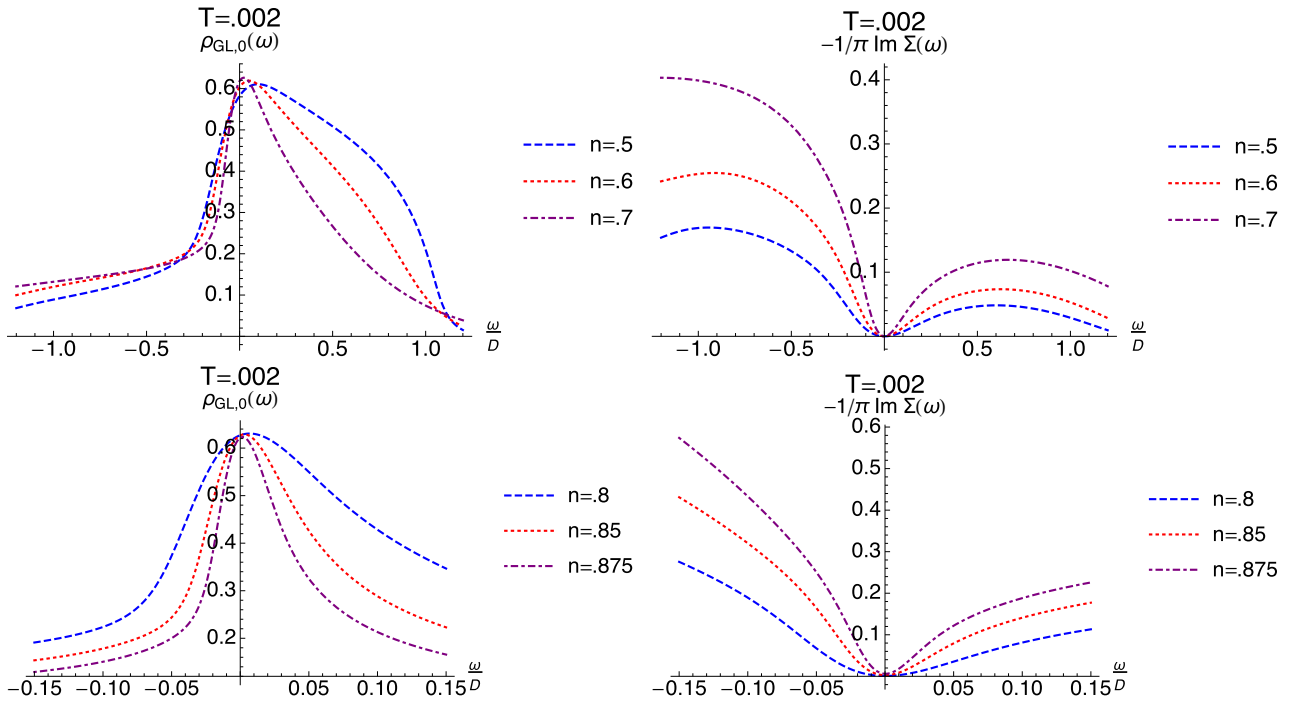

FIG. 8. The two figures on the left display the physical local spectral function $\rho_{\mathcal{G}L,0} = -\frac{1}{\pi}\,\mathrm{Im}\mathcal{G}_{Loc,0}(\omega + i0^+)$ from Eq. (25), and the two figures on the right show the Dyson self-energy $-\frac{1}{\pi}\mathrm{Im}\Sigma(\omega)$, plotted against the frequency ω/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 ω, using the observation that scaling ω with Z collapses Σ''.

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 $B \to 0$ in the remainder. With this, we obtain the asymptotic approximation

$$\lim_{B\ll 1} \xi(\omega) \sim \frac{[1 - A^2(\omega)]^{3/2}}{B(\omega)}\,\Theta[1 - A^2(\omega)]. \qquad (46)$$

In Fig. 10, we use Eq. (46) to plot $\frac{\rho_{dc}}{\rho_0}$ versus $\frac{T}{D}$ for $0.75 \leqslant n \leqslant 0.85$, 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 $0 < T < T_{\mathrm{FL}}$, where $T_{\mathrm{FL}} = (c\ D) \times Z(T = 0)$, and $c \approx 0.05$. Furthermore, $\frac{\rho_{dc}}{\rho_0}$ is a function of $\frac{T}{DZ(T=0)}$ for $T \lesssim 2T_{\mathrm{FL}}$ [Fig. 10(c)]. An important scale emphasized in DMFT studies [42,43] is the Brinkman-Rice scale ($T_{\mathrm{BR}} = D\delta$), 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 δ. As T is increased above T_{FL}, the Fermi liquid regime is followed by a linear regime for $T_{\mathrm{FL}} < T \lesssim 0.01D$. 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

$$\rho_{DC} \sim \rho_0 \frac{-\mathrm{Im}\Sigma(0,T)}{[1 - \{\boldsymbol{\mu} - \mathrm{Re}\Sigma(0,T)\}^2]^{3/2}}. \qquad (47)$$

The constituent objects $-\mathrm{Im}\Sigma(0,T)$ and $A(0,T)$ are plotted along with $Z(T)$ in the relevant temperature range in Fig. 11. We first examine $A(0,T) = \boldsymbol{\mu} - \mathrm{Re}\Sigma(0,T)$, displayed in Fig. 11(c). For $T_{\mathrm{FL}} \lesssim T \lesssim 0.01D$, 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 $[-\mathrm{Im}\Sigma(0,T)]$ in this low-temperature range. Accordingly, in Fig. 11(a), we see that $[-\mathrm{Im}\Sigma(0,T)]$ is quadratic for $T \lesssim T_{\mathrm{FL}}$ (tracked by the blue dashed parabola) and linear for $T_{\mathrm{FL}} \lesssim T \lesssim 0.01D$ (tracked by the magenta dashed line). Finally, in Fig. 11(b), we see that $Z(T)$ is approximately constant for $T \lesssim T_{\mathrm{FL}}$, and grows linearly for $T_{\mathrm{FL}} \lesssim T \lesssim 0.01D$, 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 $T \gtrsim T_{\mathrm{FL}}$. As emphasized in [43], the temperature dependence of $[-\mathrm{Im}\Sigma(0,T)]$ and $Z(T)$ lead to a quasiparticle scattering rate, defined as $[-\mathrm{Im}\Sigma(0,T)] \times Z(T)$, 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

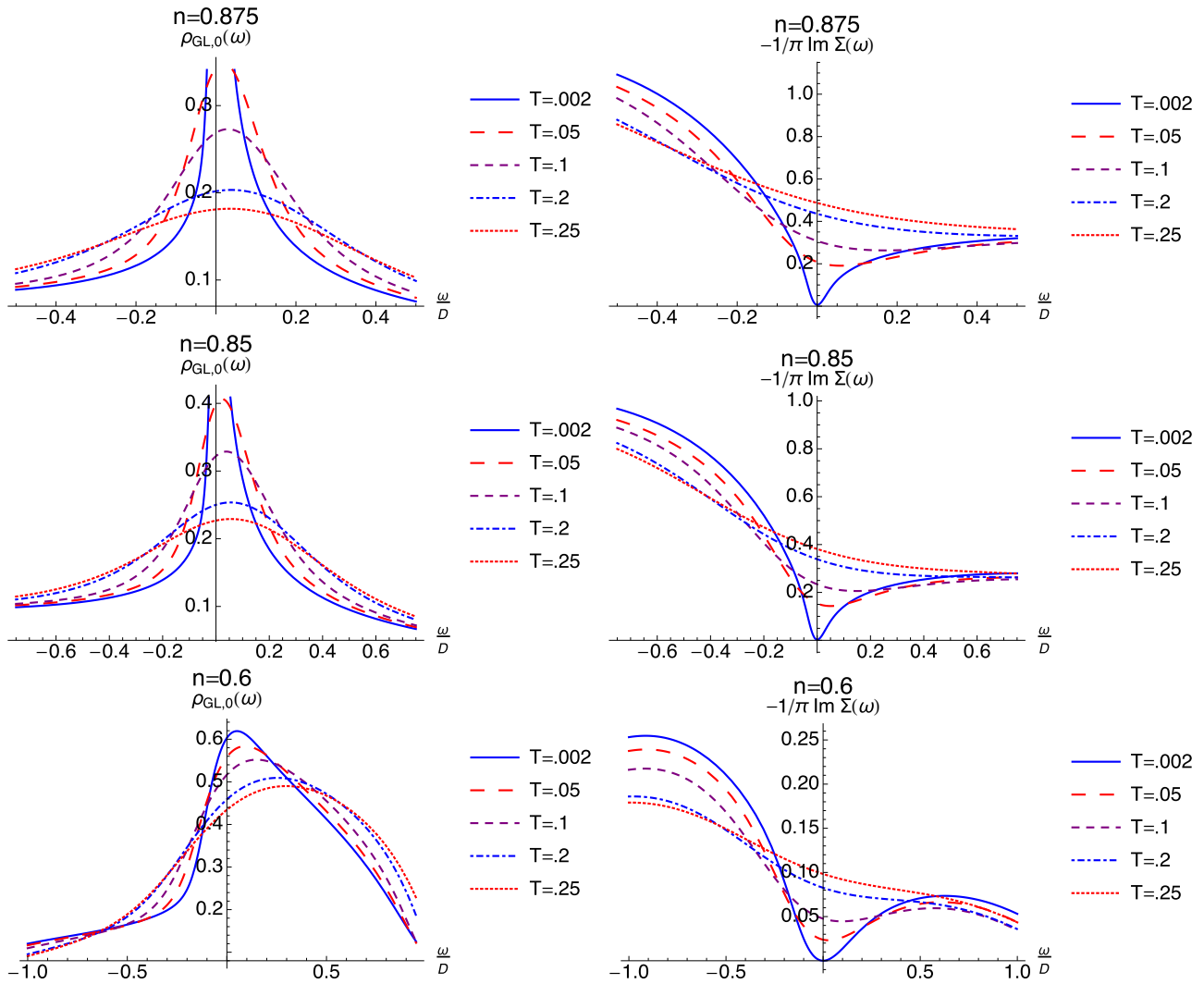

FIG. 9. The temperature variation with the frequency ω/D, of the spectral function $\rho_{\mathcal{G}L,0}$ on the left and the Dyson self-energy $-\frac{1}{\pi}\mathrm{Im}\Sigma$ on the right, at density $n = 0.875$ (top), $n = 0.85$ (middle), and at $n = 0.6$ (bottom). With increasing T we note the rapid broadening and shifting of $\rho_{\mathcal{G}L,0}$. Here, $-\frac{1}{\pi}\mathrm{Im}\Sigma$ 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 $-\frac{1}{\pi}\mathrm{Im}\Sigma$ has a strong asymmetric correction to the quadratic ω 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.

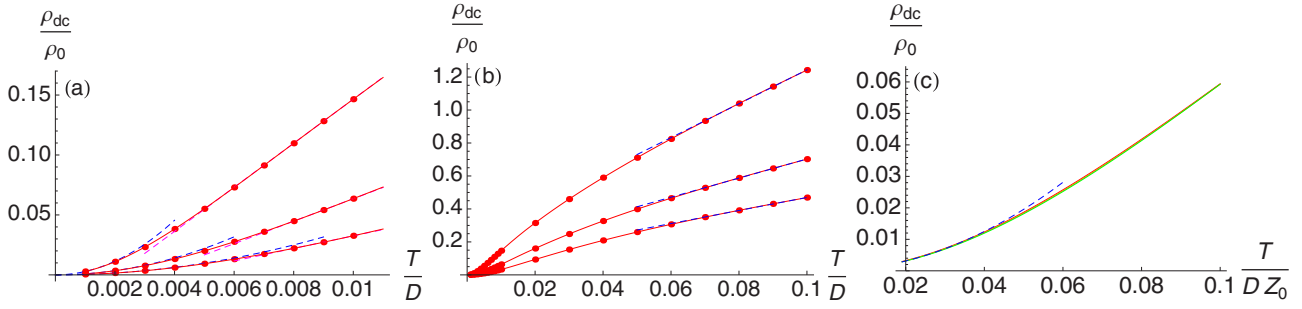

FIG. 10. (a), (b) $\frac{\rho_{dc}}{\rho_0}$ vs $\frac{T}{D}$ for $n = 0.75$, 0.8, 0.85 from bottom to top. In (a), the blue dashed parabola tracks the FL regime $0 < T < T_{\mathrm{FL}}$ where $\frac{\rho_{dc}}{\rho_0} \propto (\frac{T}{D})^2$. The magenta dashed line tracks the first linear regime $T_{\mathrm{FL}} < T \lesssim 0.01D$. 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 $T_{\mathrm{FL}} = (c\ D) \times Z(T = 0)$, with $c \approx 0.05$, and that $\frac{\rho_{dc}}{\rho_0}$ is a function of $\frac{T}{DZ(T=0)}$ for $T \lesssim 2T_{\mathrm{FL}}$.

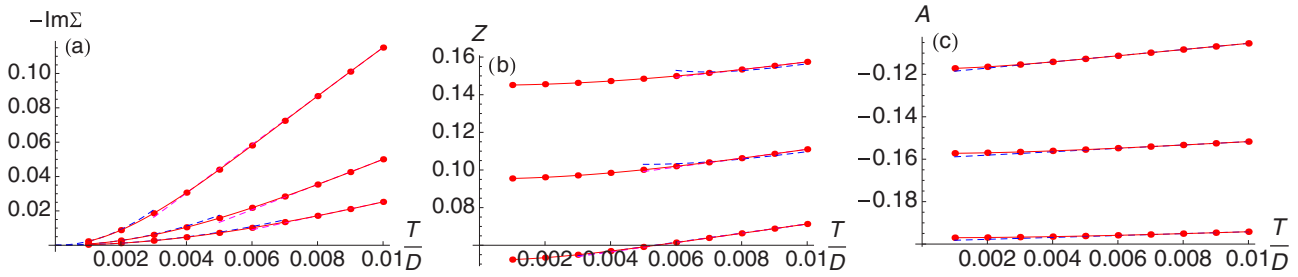

FIG. 11. (a) $[-\mathrm{Im}\Sigma(0,T)]$ vs $\frac{T}{D}$ for $n = 0.75,\ 0.8,\ 0.85$ from bottom to top. $[-\mathrm{Im}\Sigma(0,T)]$ is quadratic for $T \lesssim T_{\mathrm{FL}}$ (tracked by the blue dashed parabola) and linear for $T_{\mathrm{FL}} \lesssim T \lesssim 0.01D$ (tracked by the magenta 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 $T \lesssim T_{\mathrm{FL}}$, and grows linearly for $T_{\mathrm{FL}} \lesssim T \lesssim 0.01D$, 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 $T \gtrsim T_{\mathrm{FL}}$. (c) $A(0,T) = \boldsymbol{\mu}(T) - \mathrm{Re}\ \Sigma(0,T)$ vs $\frac{T}{D}$ for $n = 0.75,\ 0.8,\ 0.85$ from bottom to top. For $T_{\mathrm{FL}} \lesssim T \lesssim 0.01D$, it is linear, as tracked by the dashed blue line.

$T > 0.01D$, $[-\mathrm{Im}\Sigma(0,T)]$ 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 $T \gtrsim T_{\mathrm{FL}}$. 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 $T_F^{\mathrm{eff}} \sim \delta\, T_F$.

Within the terms of its limitations of $d \to \infty$ 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 ~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

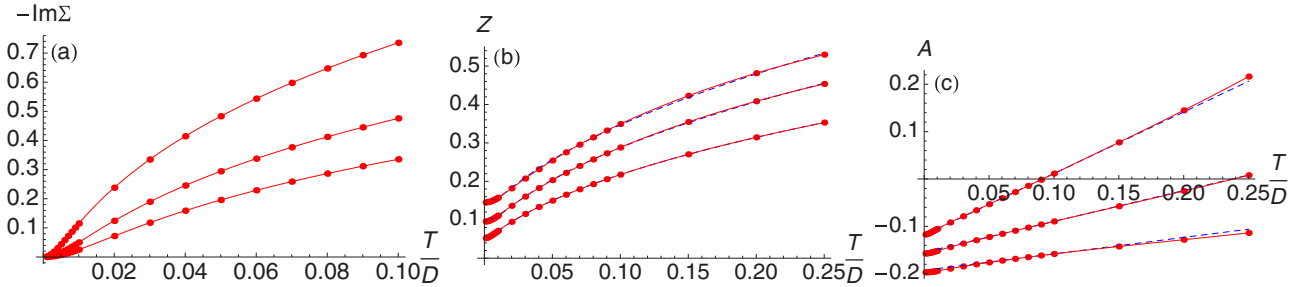

FIG. 12. Same plots as in Fig. 11 over a broader range of temperatures. (a) $[-\mathrm{Im}\Sigma(0,T)]$ continues to grow as T is increased beyond $0.01D$, until it finally begins to saturate at higher temperatures. (b) The blue dashed curve is the fit to the functional form $Z(T) = \sqrt{\frac{1+aT+bT^2}{c+dT}}$, which works well for $T \gtrsim T_{\mathrm{FL}}$. 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. Krishnamurthy, 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 Palsson 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 $\omega \geqslant D$.
[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 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 **g**, namely, by Φ.

[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 $J \to 0$. 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 γ is used exactly and hence we retain only the first term (64a) and abandon the higher-order λ 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

$$\Phi(0) = \frac{D^3}{2\pi(a_0 d\hbar^2)} \frac{\varepsilon_F}{\left(D^2 - \varepsilon_F^2\right)^{3/2}},$$

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).

Annals of Physics 357 (2015) 1–39

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: www.elsevier.com/locate/aop

Diagrammatic λ 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:
λ 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 λ-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 λ 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 (E. Perepelitsky).

http://dx.doi.org/10.1016/j.aop.2015.03.010
0003-4916/© 2015 Elsevier Inc. All rights reserved.

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 $\widetilde{\mu}[i,f]$ (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 $\widetilde{\mu}[i,f]$ individually satisfy the shift invariance theorem [10], and together with the original chemical potential μ, 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 $\widetilde{\mu}[i,f]$ are defined through their respective equations of motion, and the expansion parameter λ is inserted directly into the equation of motion. The practical issue of computing objects to various orders in λ 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 $\widetilde{\mu}[i,f]$ 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 λ expansion. This expansion allows one to calculate the Greens function and related objects to any order in λ 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 λ series has the following advantages:

- It allows one to calculate the nth 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 λ, 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 $\widetilde{\mu}[i,f]$. 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 λ, 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 λ. 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 $\widetilde{\mu}$. These are in turn expressed in terms of the two self-energies Φ and Ψ. 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 λ 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 $\mathbf{g}$, $\widetilde{\mu}$, Φ, and Ψ 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 λ 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 Φ consists of two independent pieces. The first can be obtained by adding a single interaction line to the terminal point of a Ψ diagram, while the second one is completely independent of Ψ. We denote the second piece by the letter χ, 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 λ 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 $\Psi(\vec{k}, i\omega_k) \to \Psi(i\omega_k)$ and $\chi(\vec{k}, i\omega_k) \to \chi(i\omega_k)$. We also derive the Schwinger EOM defining the object χ in finite dimensions.

We derive diagrammatic rules for the three point vertices Λ and $\mathcal{U}$, defined as functional derivatives of $\mathbf{g}^{-1}$ and $\widetilde{\mu}$ (Eq. (11)). Diagrammatically, their relationship to Φ and Ψ 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 λ 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 $\widetilde{\mu}$, 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 λ 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 λ. 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 $\mathbf{g}$, $\widetilde{\mu}$, Φ, Ψ, χ, γ, Λ, 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 λ 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 λ for the objects $\mathbf{g}$ and $\widetilde{\mu}$. 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 λ 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 λ expansion

The Greens function is the fundamental object in this theory and is defined as usual by

$$\mathcal{G}_{\sigma_i,\sigma_f}[i,f] \equiv -\langle\langle X_i^{0\sigma_i} X_f^{\sigma_f 0}\rangle\rangle = -\frac{1}{Z[\mathcal{V}]}\mathrm{tr}\, e^{-\beta H} T_\tau \left(e^{-\mathcal{A}} X_i^{0\sigma_i}(\tau_i) X_f^{\sigma_f 0}(\tau_f)\right), \tag{3}$$

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 $Z[\mathcal{V}] = \mathrm{tr}\, e^{-\beta H} T_\tau e^{-\mathcal{A}}$. 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].

$$\begin{aligned}
&\left\{\left[(\partial_{\tau_i} - \mu)\delta[i,\mathbf{j}] - t[i,\mathbf{j}]\right]\delta_{\sigma_1,\sigma_j} + \mathcal{V}_i^{\sigma_1,\sigma_j}\delta[i,\mathbf{j}]\right\}\ \mathcal{G}_{\sigma_j,\sigma_2}[\mathbf{j},f] \\
&\quad = -\delta[i,f]\delta_{\sigma_1,\sigma_2} + \lambda\delta[i,f]\gamma_{\sigma_1,\sigma_2}[i] - \lambda t[i,\mathbf{j}]\gamma_{\sigma_1,\sigma_a}[i]\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f] \\
&+ \lambda t[i,\mathbf{j}]\sigma_1\sigma_a\frac{\delta}{\delta\mathcal{V}_i^{\bar{\sigma}_1,\bar{\sigma}_a}}\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f] + \frac{\lambda}{2}J[i,\mathbf{j}]\left(\gamma_{\sigma_1,\sigma_a}[\mathbf{j}]\mathcal{G}_{\sigma_a,\sigma_2}[i,f] - \sigma_1\sigma_a\frac{\delta}{\delta\mathcal{V}_{\mathbf{j}}^{\bar{\sigma}_1,\bar{\sigma}_a}}\mathcal{G}_{\sigma_a,\sigma_2}[i,f]\right),
\end{aligned} \tag{4}$$

where the bold repeated indices are summed over. The functional derivative takes place at time (τ_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 ∂_{τ_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

$$\{X_i^{0\sigma_i}, X_j^{\sigma_j 0}\} = \delta_{ij}\left(\delta_{\sigma_i\sigma_j} - \lambda\,\sigma_i\sigma_j X_i^{\bar{\sigma}_i\bar{\sigma}_j}\right), \tag{5}$$

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 λ expansion. From this, we get $\delta_{if}\delta(\tau_i - \tau_f)\left(\delta_{\sigma_i\sigma_f} - \lambda\sigma_i\sigma_f\langle X_i^{\bar{\sigma}_i\bar{\sigma}_f}(\tau_i)\rangle\right)$. This is expressed back in terms of the Greens function by writing $\langle X_i^{\bar{\sigma}_i\bar{\sigma}_f}\rangle \to \mathcal{G}_{\bar{\sigma}_f\bar{\sigma}_i}[i\tau_i, i\tau_i^+] = \sigma_i\sigma_f\gamma_{\sigma_i\sigma_f}[i]$, and thus to the first two terms on the right hand side of Eq. (4).

Another contribution arises from the τ_i dependence in the lower and upper limits of the time integrals in the expression

$$T_\tau e^{-\mathcal{A}_S} X_i^{0\sigma_i}(\tau_i) = e^{-\sum_j \int_{\tau_i}^{\beta} X_j^{\sigma\sigma'}(\tau')\mathcal{V}_j^{\sigma\sigma'}(\tau')d\tau'} X_i^{0\sigma_i}(\tau_i) e^{-\sum_j \int_0^{\tau_i} X_j^{\sigma\sigma'}(\tau')\mathcal{V}_j^{\sigma\sigma'}(\tau')d\tau'}, \tag{6}$$

involving the equal time commutator $\sum_j \mathcal{V}_j^{\sigma\sigma'}(\tau_i)[X_j^{\sigma\sigma'}(\tau_i), X_i^{0\sigma_i}(\tau_i)] = \mathcal{V}_i^{\sigma_i\sigma'}(\tau_i)X_i^{0\sigma'}(\tau_i)$. 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

$$[X_i^{0\sigma}, H] = -\mu X_i^{0\sigma} - t_{ij}X_j^{0\sigma_i} + \lambda\sum_{j\sigma_j} t_{ij}(\sigma_i\sigma_j)X_i^{\bar{\sigma}_i\bar{\sigma}_j}X_j^{0\sigma_j} - \frac{1}{2}\lambda\sum_{j\neq i} J_{ij}(\sigma_i\sigma_j)X_j^{\bar{\sigma}_i\bar{\sigma}_j}X_i^{0\sigma_j}. \tag{7}$$

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 λ, 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

$$\begin{aligned}\langle\langle X_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i)X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f)\rangle\rangle &= \langle\langle X_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i)\rangle\rangle\langle\langle X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f)\rangle\rangle \\ &- \frac{\delta}{\delta\mathcal{V}_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i)}\langle\langle X_j^{0\sigma_j}(\tau_i), X_f^{\sigma_f 0}(\tau_f)\rangle\rangle.\end{aligned} \tag{8}$$

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 $J \to 0$ 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:

$$\mathcal{G}[i,f] = \mathbf{g}[i,\mathbf{j}].\widetilde{\mu}[\mathbf{j},f], \tag{9}$$

where $\mathbf{g}[i,f]$ is the auxiliary Green's function, $\widetilde{\mu}[i,f]$ is the caparison factor, all objects have been represented as 2×2 matrices in spin space, and matrix multiplication has been indicated by a dot. $\mathbf{g}[i,f]$ and $\widetilde{\mu}[i,f]$ are defined by the their respective Schwinger equations of motion.

$$\begin{aligned}&\mathbf{g}^{-1}[i,m] = (\mu - \partial_{\tau_i} - \mathcal{V}_i)\,\delta[i,m] + t[i,m]\,(1 - \lambda\gamma[i]) - \lambda\Phi[i,m].\\ &\widetilde{\mu}[i,m] = (1 - \lambda\gamma[i])\delta[i,m] + \lambda\Psi[i,m]\\ &\Phi[i,m] = -t[i,\mathbf{j}]\,\xi^*.\mathbf{g}[\mathbf{j},\mathbf{n}].\Lambda_*[\mathbf{n},m;i]; \qquad \Psi[i,m] = -t[i,\mathbf{j}]\,\xi^*.\mathbf{g}[\mathbf{j},\mathbf{n}].\mathcal{U}_*[\mathbf{n},m;i].\end{aligned} \tag{10}$$

These exact relations give the required objects $\mathbf{g}$ and $\widetilde{\mu}$ 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

$$\gamma[i] = \widetilde{\mu}^{(k)}[\mathbf{n}, i^+].\mathbf{g}^{(k)}[i, \mathbf{n}]; \quad \Lambda[n, m; i] = -\frac{\delta}{\delta \mathcal{V}_i}\mathbf{g}^{-1}[n, m]; \quad \mathcal{U}[n, m; i] = \frac{\delta}{\delta \mathcal{V}_i}\widetilde{\mu}[n, m], \tag{11}$$

where we have used the notation $M^{(k)}_{\sigma_1,\sigma_2} = \sigma_1\sigma_2 M_{\bar{\sigma}_2,\bar{\sigma}_1}$ 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 $\widetilde{\mu}$. The vertices defined above (Λ and $\mathcal{U}$) have four spin indices, those of the object being differentiated and those of the source. For example, $\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_a\sigma_b}[n, m; i] = \frac{\delta}{\delta \mathcal{V}_i^{\sigma_a\sigma_b}}\widetilde{\mu}_{\sigma_1\sigma_2}[n, m]$. 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: $(\xi^*.\mathbf{g}[j, \mathbf{n}].\mathcal{U}_*[\mathbf{n}, m; i])_{\sigma_1\sigma_2} = \sigma_1\sigma_{\mathbf{a}}\, \mathbf{g}_{\sigma_{\mathbf{a}},\sigma_{\mathbf{b}}}[j, \mathbf{n}]\frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_1\bar{\sigma}_{\mathbf{a}}}}\widetilde{\mu}_{\sigma_{\mathbf{b}},\sigma_2}[\mathbf{n}, m]$.

The λ expansion is obtained by expanding Eqs. (10) and (11) iteratively in the continuity parameter λ. The $\lambda = 0$ limit of these equations is the free Fermi gas. Therefore, a direct expansion in λ will lead to a series in λ 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 λ, 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.

$$\frac{\delta \mathbf{g}[i, m]}{\delta \mathcal{V}_r} = \mathbf{g}[i, \mathbf{x}].\Lambda[\mathbf{x}, \mathbf{y}, r].\mathbf{g}[\mathbf{y}, m]. \tag{12}$$

Within the λ expansion, the LHS is evaluated to a certain order in λ by taking the vertex Λ on the RHS to be of that order in λ.

3. Heuristic discussion of λ expansion diagrams

3.1. Numerical representations of Feynman diagrams

Before deriving the precise rules for the λ 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 λ 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).

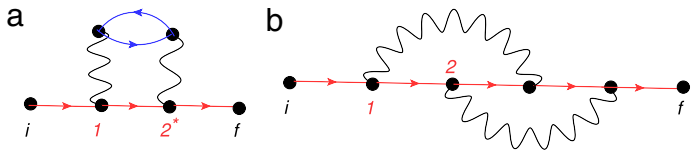

Fig. 1. Second order Feynman diagrams for $\mathcal{G}[i, f]$. The zeroth Fermi loop, which is the chain running from i to f is colored in red. In panel (a), the first Fermi loop is colored in blue. The numerical representation of the diagram in panel (a) is $1\,2^*\,f;\ 2^*:0\,{}_1\,f$, while that of 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.)

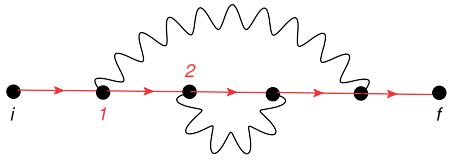

Fig. 2. This Feynman diagram results from reversing the order of the subscripts in the numerical representation of the Feynman diagram in Fig. 1(b). Therefore, the numerical representation of this diagram is $1\,2\,{}_{2\,1}\,f$.

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.

$$\begin{array}{ll} & 1\,2^*\,f \\ 2^*: & 0\,f. \end{array}$$

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

$$\begin{array}{ll} & 1\,2^*\,f \\ 2^*: & 0\,{}_1\,f. \end{array}$$

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.

$$1\,2\,3\,4^*\,f.$$

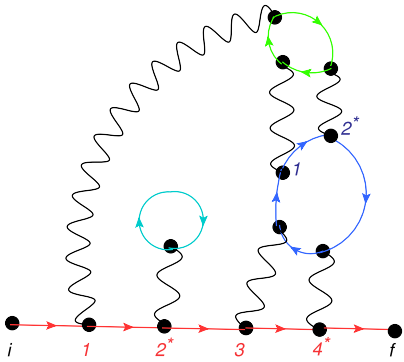

Fig. 3. Sixth order Feynman diagram. The zeroth, first, second, and third Fermi loops are drawn in red, blue, green, and turquoise respectively. Interaction vertices introduced in a particular Fermi loop are numbered in the same color as that loop. An interaction vertex is starred if it spawns a new 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

$$\begin{array}{rl} & 1\,2\,3\,4^*\,f \\ 4^*: & 0\,1\,2\,f. \end{array}$$

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

$$\begin{array}{rl} & 1\,2\,3\,4^*\,f \\ 4^*: & 0\,1\,2^*\,f \\ (4,2)^*: & 0\,f. \end{array}$$

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

$$\begin{array}{rl} & 1\,2^*\,3\,4^*\,f \\ 4^*: & 0\,1\,2^*\,f \\ (4,2)^*: & 0\,f \\ 2^*: & 0\,f. \end{array}$$

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

$$\begin{array}{rl} & 1\,2^*\,3\,4^*\,f \\ 4^*: & 0_{\,3}\,1\,2^*\,f \\ (4,2)^*: & 0_{\,(4,1)\,1}\,f \\ 2^*: & 0\,f\,. \end{array}$$

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 $1 \ldots m\,f$ 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 ith lower row also consists of integers $0 \ldots m_i\,f$, where $m_i \geq 0$, e.g.

$$\begin{array}{rl} & 1\,2^*\,3\,4^*\,f \\ 4^*: & 0\,1\,2\,f \\ 2^*: & 0\,f\,. \end{array}$$

- (4) In the lower rows, assign a star to any of the integers excluding 0, e.g.

$$\begin{array}{rl} & 1\,2^*\,3\,4^*\,f \\ 4^*: & 0\,1\,2^*\,f \\ 2^*: & 0\,f\,. \end{array}$$

- (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.

$$\begin{array}{rl} & 1\,2^*\,3\,4^*\,f \\ 4^*: & 0\,1\,2^*\,f \\ (4,2)^*: & 0\,f \\ 2^*: & 0\,f\,. \end{array}$$

- (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.

$$\begin{array}{rl} & 1\,2^*\,3\,4^*\,f \\ 4^*: & 0_{\,3}\,1\,2^*\,f \\ (4,2)^*: & 0_{\,(4,1)\,1}\,f \\ 2^*: & 0\,f\,. \end{array}$$

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.

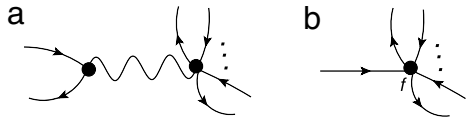

Fig. 4. The 2 elements used for constructing the λ expansion diagrams.

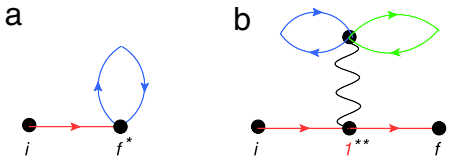

Fig. 5. Two of the simplest non-Feynman diagrams in the λ expansion. A non-Feynman diagram occurs when the terminal point of a Fermi loop spawns another Fermi loop. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Topologies of λ expansion diagrams

The exact rules for drawing diagrams for the λ expansion, as defined in Section 2 will be derived in Section 4. There, it will be shown that the λ 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 λ 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.

$$\begin{array}{rl} & f^* \\ f^*: & 0 f. \end{array}$$

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.

$$\begin{array}{rl} & 1^{**} f \\ 1^{**}: & 0 f^* \\ (1,f)^*: & 0 f. \end{array}$$

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 λ expansion diagrams. However, as shown below, the allowed topologies of λ 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 λ 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

$$1\,{}_1 f.$$

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

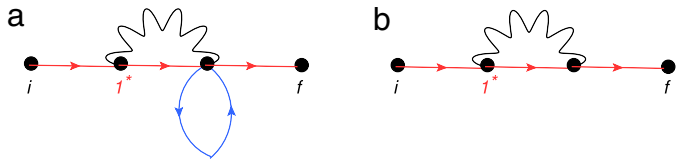

Fig. 6. The diagram in panel (a) is not allowed in the λ expansion. This is because first Fermi loop emanates from a point which is represented by a subscript in the Fock diagram displayed in panel (b).

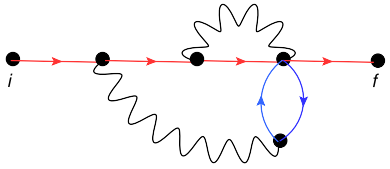

Fig. 7. A more elaborate version of the diagram in Fig. 6(a), which is also not allowed in the λ expansion.

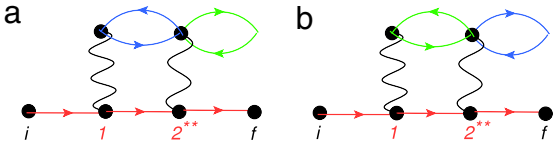

Fig. 8. A demonstration that unlike Feynman diagrams, λ expansion diagrams are not in one-to-one correspondence with their numerical representations. In the diagram in panel (a), the interaction vertex 1 of the zeroth Fermi loop also connects to the first Fermi loop. In the diagram in panel (b), it connects 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 λ expansion diagram is drawn in Fig. 7.

Another feature of the λ 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

$$\begin{array}{rl} & 1\,2^{**}\,f \\ 2^{**}: & 0\,{}_1\,f^{*} \\ (2,f)^{*}: & 0\,f\,. \end{array}$$

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

$$\begin{array}{rl} & 1\,2^{**}\,f \\ 2^{**}: & 0\,f^{*} \\ (2,f)^{*}: & 0\,{}_1\,f\,. \end{array}$$

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.

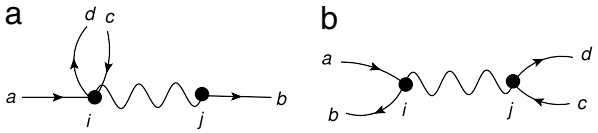

Fig. 9. t-vertices in panel (a) versus J-vertices in panel (b) for the λ expansion diagrams which are also Feynman diagrams. The corresponding lines are marked with corresponding letters.

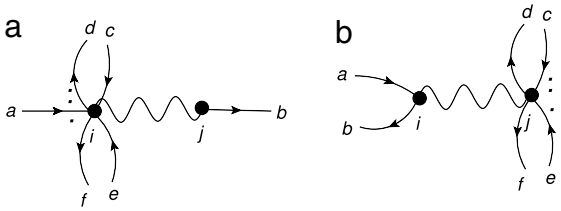

Fig. 10. t-vertices in panel (a) versus J-vertices in panel (b) for the more general λ expansion diagrams. The corresponding lines are marked with corresponding letters.

A final point to mention in this discussion of the λ 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

$$-\sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} = -\sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger (1 - n_{i\bar{\sigma}}) c_{j\sigma} = -\sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger n_{i\bar{\sigma}} c_{j\sigma}. \tag{13}$$

Here, we have used the non-Hermitean mapping described in Ref. [11]

$$X_i^{\sigma 0} \to c_{i\sigma}^\dagger (1 - n_{i\bar{\sigma}}); \qquad X_j^{0\sigma} \to c_{j\sigma}. \tag{14}$$

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

$$\frac{1}{2}\sum_{ij\sigma} J_{ij} n_{i\sigma} + \frac{1}{4}\sum_{ij\sigma_1\sigma_2} \{J_{ij} c_{i\sigma_1}^\dagger c_{i\sigma_2} c_{j\sigma_2}^\dagger c_{j\sigma_1} - n_{i\sigma_1} n_{j\sigma_2}\}. \tag{15}$$

The first term amounts to a shift in the chemical potential μ, 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 λ 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 λ 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 λ 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

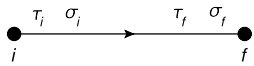

Fig. 11. The zeroth order contribution to the Green's function: $\mathbf{g}^{(0)}[i,f]$.

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 λ 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 λ 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 λ expansion for $\mathcal{G}[i,f]$.

4.1. Integral equation of motion and the first order λ expansion.

As can be seen from Eq. (4), the parameter λ adiabatically connects the free Fermi gas at $\lambda = 0$ with the fully projected model at $\lambda = 1$. Therefore, in the bare λ series for $\mathcal{G}$, to each order in λ, $\mathcal{G}[i,f]$ 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 nth 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 λ. 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

$$-\mathbf{g}^{-1(0)}_{\sigma_1,\sigma_j}[i,\mathbf{j}]\mathcal{G}_{\sigma_j,\sigma_2}[\mathbf{j},f] = -\delta[i,f]\delta_{\sigma_1,\sigma_2} + \lambda\times\delta[i,f]\gamma_{\sigma_1,\sigma_2}[i] - \lambda\times t[i,\mathbf{j}]\gamma_{\sigma_1,\sigma_a}[i]\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f] + \lambda\times t[i,\mathbf{j}]\sigma_1\sigma_a\frac{\delta}{\delta\mathcal{V}_i^{\bar\sigma_1,\bar\sigma_a}}\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f], \tag{16}$$

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 $\mathcal{G}[i,f]$.

$$\mathcal{G}_{\sigma_1,\sigma_2}[i,f] = \mathbf{g}^{(0)}_{\sigma_1,\sigma_2}[i,f] - \lambda\,\mathbf{g}^{(0)}_{\sigma_1,\sigma_b}[i,f]\sigma_b\sigma_2\mathcal{G}_{\bar\sigma_2,\bar\sigma_b}[f,f^+] - \lambda\times\mathbf{g}^{(0)}_{\sigma_1,\sigma_b}[i,\mathbf{k}]\left(-t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a\mathcal{G}_{\bar\sigma_a,\bar\sigma_b}[\mathbf{k},\mathbf{k}^+]\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f] + t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a\frac{\delta}{\delta\mathcal{V}_{\mathbf{k}}^{\bar\sigma_b,\bar\sigma_a}}\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f]\right), \tag{17}$$

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): $-\lambda\sigma_b\sigma_2\mathbf{g}^{(0)}_{\substack{if\\ \sigma_1\sigma_b}}[\tau_i,\tau_f]\mathbf{g}^{(0)}_{\substack{ff\\ \bar\sigma_2\bar\sigma_b}}[\tau_f,\tau_f^+]$;

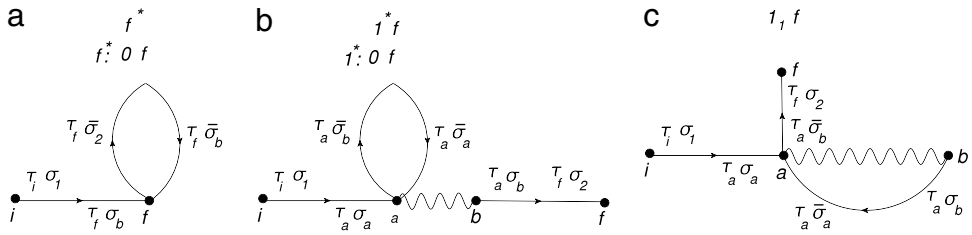

Fig. 12. The first order contribution to the Green's function: $\mathcal{G}^{(1)}[i,f]$. The diagrams in panels (a), (b), and (c) come from the first, second, and third terms on the RHS of Eq. (17), respectively.

(b): $\lambda\sigma_a\sigma_b \mathbf{g}^{(0)}_{\substack{ia\\ \sigma_1\sigma_a}}[\tau_i,\tau_a]\mathbf{g}^{(0)}_{\substack{aa\\ \bar\sigma_b\bar\sigma_a}}[\tau_a,\tau_a^+]t_{ab}\mathbf{g}^{(0)}_{\substack{bf\\ \sigma_b\sigma_2}}[\tau_a,\tau_f]$; and (c): $-\lambda\sigma_a\sigma_b \mathbf{g}^{(0)}_{\substack{ia\\ \sigma_1\sigma_a}}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{\substack{ba\\ \sigma_b\bar\sigma_a}}[\tau_a,\tau_a^+]\mathbf{g}^{(0)}_{\substack{af\\ \bar\sigma_b\sigma_2}}[\tau_a,\tau_f]$. In drawing the diagram in Fig. 12(c), we have used the Schwinger identity

$$\frac{\delta \mathbf{g}^{(0)}_{\sigma_a,\sigma_b}[i,f]}{\delta \mathcal{V}_r^{\sigma_c\sigma_d}} = -\mathbf{g}^{(0)}_{\sigma_a,\sigma_x}[i,\mathbf{x}]\frac{\delta \mathbf{g}^{-1(0)}_{\sigma_x,\sigma_y}[\mathbf{x},\mathbf{y}]}{\delta \mathcal{V}_r^{\sigma_c\sigma_d}}\mathbf{g}^{(0)}_{\sigma_y,\sigma_b}[\mathbf{y},f] = \mathbf{g}^{(0)}_{\sigma_a,\sigma_c}[i,r]\mathbf{g}^{(0)}_{\sigma_d,\sigma_b}[r,f]. \tag{18}$$

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 nth 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 λ. Moreover, by noticing the pattern in the iterative process, we can derive the rules for obtaining the nth 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 λ expansion introduced in Ref. [9] and the equivalent diagrammatic one developed here. We now present the diagrammatic rules for calculating the nth order contribution to $\mathcal{G}$.

- (1) Write a row of consecutive integers followed by the letter f, i.e. $1\ldots m\,f$, 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 ν 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.

$$
\begin{array}{rl}
 & 1^{**}\ 2\ 3\ {}_{3\ 2}f^{*} \\
1^{**}: & 0\ 1\ 2^{***}\ f^{*} \\
(1,2)^{***}: & 0\ 1\ f^{**} \\
(1,2,f)^{**}: & 0\ {}_{(1,2,1)}\ f^{*} \\
(1,2,f,f)^{*}: & 0\ 1\ {}_{(1,2,f,f,1)}\ f \\
(1,f)^{*}: & 0\ 1\ {}_{(1,f,1)}\ {}_{(1,1)}\ f \\
f^{*}: & 0\ 1\ 2\ {}_{(f,1)}\ {}_{(f,2)}\ f.
\end{array}
$$

- (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) σ_a and σ_d contribute to the sign while σ_b and σ_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 σ_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 $\sigma_a \to \sigma_2$. 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

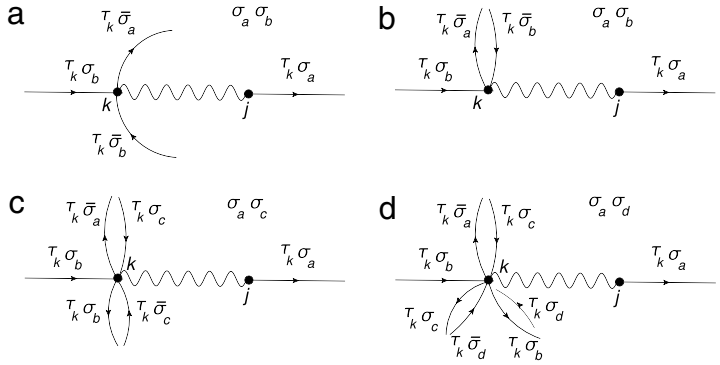

Fig. 13. Interaction vertices appearing in the diagrams. Panels (a), (b), (c), and (d) correspond to 0, 1, 2, and 3 stars on the number representing the interaction vertex, respectively. Note that the lines are broken into pairs based on spin. A pair of two incoming or two outgoing lines share opposite 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.

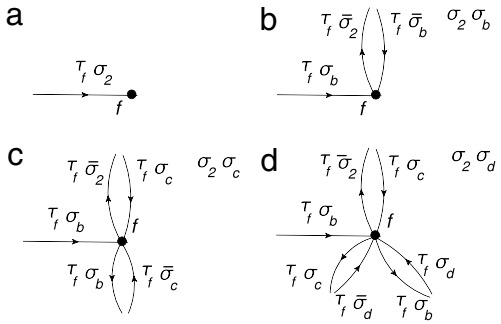

Fig. 14. Terminal point in the diagram corresponding to the f in the top row. Panels (a), (b), (c), and (d) correspond to 0, 1, 2, and 3 stars on the f in the top row, respectively. Same comments regarding spin apply as in Fig. 13. Note that in the case of one or more stars on the f in the top row, the 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 σ_c (incoming) correspond to the row with 2 stars on its f, the lines σ_b (outgoing) and σ_d correspond to the row with 1 star on its f, and the lines σ_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 σ_n (incoming) are part of the row $(1, 2)^{***}$ (3rd row) in the numerical representation, the lines σ_l (outgoing) and σ_o are part of the row $(1, 2, f)^{**}$ (4th row) in the numerical representation, and the lines σ_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 $(-1)^{n}(-1)^{s}(-1)^{s_f-1} \times$ (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

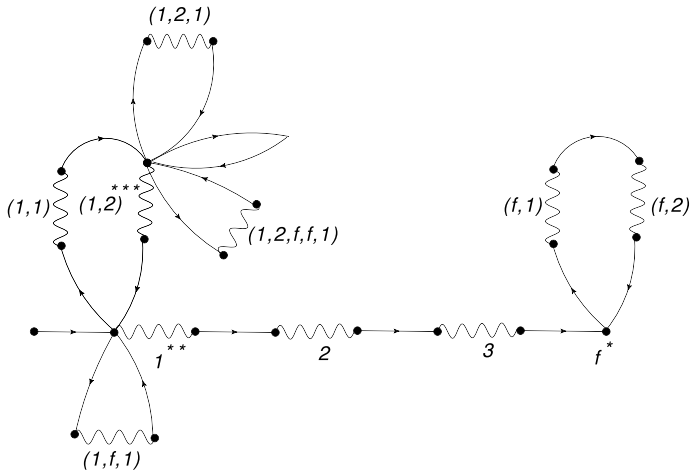

Fig. 15. Intermediate step in the process of drawing the diagram corresponding to the numerical representation in step 5 of the rules. All of the interaction vertices are drawn in. To complete the diagram, we must split some of the Green's function lines through the unused points in the interaction vertices in a manner indicated by the numerical representation.

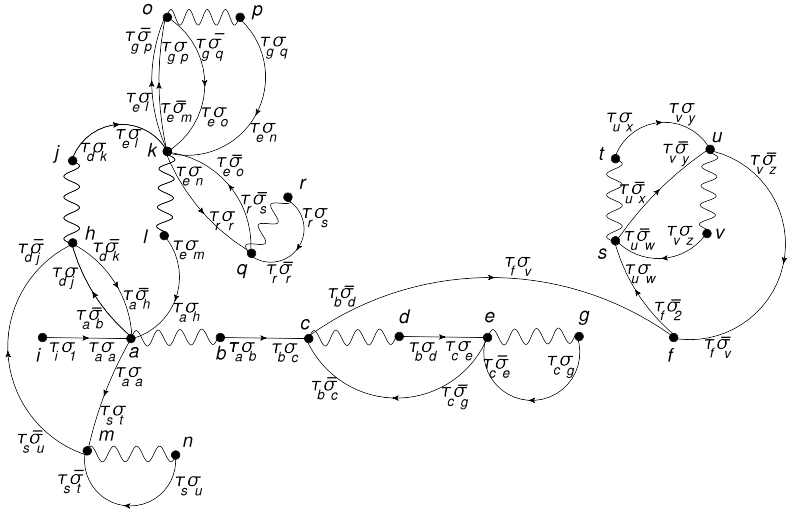

Fig. 16. Diagram corresponding to the numerical representation in step 5 of the rules.

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

$$
\begin{aligned}
&-\sigma_b\sigma_h\sigma_c\sigma_d\sigma_e\sigma_g\sigma_v\sigma_2\sigma_w\sigma_x\sigma_y\sigma_z\sigma_t\sigma_u\sigma_j\sigma_k\sigma_m\sigma_o\sigma_p\sigma_q\sigma_r\sigma_s\\
&\underset{\sigma_1\sigma_a}{\mathbf{g}^{(0)}_{ia}}[\tau_i,\tau_a]t_{ab}\underset{\sigma_b\sigma_c}{\mathbf{g}^{(0)}_{bc}}[\tau_a,\tau_b]t_{cd}\underset{\sigma_d\sigma_e}{\mathbf{g}^{(0)}_{de}}[\tau_b,\tau_c]t_{eg}\underset{\sigma_g\bar{\sigma}_e}{\mathbf{g}^{(0)}_{ge}}[\tau_c,\tau_c^{+}]\underset{\bar{\sigma}_g\bar{\sigma}_c}{\mathbf{g}^{(0)}_{ec}}[\tau_c,\tau_b]\\
&\underset{\bar{\sigma}_d\sigma_v}{\mathbf{g}^{(0)}_{cf}}[\tau_b,\tau_f]\underset{\bar{\sigma}_b\bar{\sigma}_j}{\mathbf{g}^{(0)}_{ah}}[\tau_a,\tau_d]t_{hj}\underset{\sigma_k\sigma_l}{\mathbf{g}^{(0)}_{jk}}[\tau_d,\tau_l]t_{kl}\underset{\sigma_m\sigma_n}{\mathbf{g}^{(0)}_{la}}[\tau_l,\tau_a]\underset{\sigma_e\bar{\sigma}_p}{\mathbf{g}^{(0)}_{ko}}[\tau_e,\tau_g]t_{op}
\end{aligned}
$$

$$\mathbf{g}^{(0)}_{\substack{pk\\ \sigma_q\sigma_n}}[\tau_g,\tau_e]\mathbf{g}^{(0)}_{\substack{ko\\ \bar\sigma_m\sigma_p}}[\tau_e,\tau_g]\mathbf{g}^{(0)}_{\substack{ok\\ \bar\sigma_q\sigma_o}}[\tau_g,\tau_e]\mathbf{g}^{(0)}_{\substack{kq\\ \sigma_n\bar\sigma_r}}[\tau_e,\tau_r]t_{qr}$$
$$\mathbf{g}^{(0)}_{\substack{rq\\ \sigma_s\bar\sigma_r}}[\tau_r,\tau_r^+]\mathbf{g}^{(0)}_{\substack{qk\\ \bar\sigma_s\bar\sigma_o}}[\tau_r,\tau_e]\mathbf{g}^{(0)}_{\substack{am\\ \sigma_a\sigma_t}}[\tau_a,\tau_s]t_{mn}\mathbf{g}^{(0)}_{\substack{nm\\ \sigma_u\bar\sigma_t}}[\tau_s,\tau_s^+]\mathbf{g}^{(0)}_{\substack{mh\\ \bar\sigma_u\bar\sigma_j}}[\tau_s,\tau_d]\mathbf{g}^{(0)}_{\substack{ha\\ \bar\sigma_k\bar\sigma_h}}[\tau_d,\tau_a]$$
$$\mathbf{g}^{(0)}_{\substack{fs\\ \bar\sigma_2\sigma_w}}[\tau_f,\tau_u]t_{st}\mathbf{g}^{(0)}_{\substack{tu\\ \sigma_x\sigma_y}}[\tau_u,\tau_v]t_{uv}\mathbf{g}^{(0)}_{\substack{vs\\ \sigma_z\bar\sigma_w}}[\tau_v,\tau_u]\mathbf{g}^{(0)}_{\substack{su\\ \bar\sigma_x\bar\sigma_y}}[\tau_u,\tau_v]\mathbf{g}^{(0)}_{\substack{uf\\ \bar\sigma_z\bar\sigma_v}}[\tau_v,\tau_f].$$

Upon turning off the sources, the Green's functions become spin diagonal, i.e. $\mathbf{g}^{(0)}_{\sigma_1\sigma_2}[i,f] = \delta_{\sigma_1\sigma_2}\mathbf{g}^{(0)}_{\uparrow\uparrow}[i,f] = \delta_{\sigma_1\sigma_2}\mathbf{g}^{(0)}_{\downarrow\downarrow}[i,f] \equiv \delta_{\sigma_1\sigma_2}\mathbf{g}^{(0)}[i,f]$. 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 σ_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 λ 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 σ_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 σ_2 from the zeroth spin loop now also appears explicitly in the sign. Therefore, flipping all of the spins in the zeroth loop once again does not change the sign of the diagram. In Fig. 16, we find

$$(\sigma_1) = (\sigma_a) = \sigma_t = \bar\sigma_u = \bar\sigma_j = \sigma_b = \sigma_c = \sigma_g = \bar\sigma_e = \bar\sigma_d = \sigma_v = \sigma_z = \bar\sigma_w = \sigma_2; \qquad \sigma_x = \sigma_y;$$
$$\sigma_h = \sigma_k = (\sigma_l) = \bar\sigma_p = \sigma_m; \qquad \sigma_q = \bar\sigma_o = \bar\sigma_s = \sigma_r = (\sigma_n),$$

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

$$-8\times\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_a,\tau_b]t_{cd}\mathbf{g}^{(0)}_{de}[\tau_b,\tau_c]t_{eg}\mathbf{g}^{(0)}_{ge}[\tau_c,\tau_c^+]\mathbf{g}^{(0)}_{ec}[\tau_c,\tau_b]$$
$$\mathbf{g}^{(0)}_{cf}[\tau_b,\tau_f]\mathbf{g}^{(0)}_{ah}[\tau_a,\tau_d]t_{hj}\mathbf{g}^{(0)}_{jk}[\tau_d,\tau_l]t_{kl}\mathbf{g}^{(0)}_{la}[\tau_l,\tau_a]\mathbf{g}^{(0)}_{ko}[\tau_e,\tau_g]t_{op}\mathbf{g}^{(0)}_{pk}[\tau_g,\tau_e]\mathbf{g}^{(0)}_{ko}[\tau_e,\tau_g]$$
$$\mathbf{g}^{(0)}_{ok}[\tau_g,\tau_e]\mathbf{g}^{(0)}_{kq}[\tau_e,\tau_r]t_{qr}\mathbf{g}^{(0)}_{rq}[\tau_r,\tau_r^+]\mathbf{g}^{(0)}_{qk}[\tau_r,\tau_e]\mathbf{g}^{(0)}_{am}[\tau_a,\tau_s]$$
$$t_{mn}\mathbf{g}^{(0)}_{nm}[\tau_s,\tau_s^+]\mathbf{g}^{(0)}_{mh}[\tau_s,\tau_d]\mathbf{g}^{(0)}_{ha}[\tau_d,\tau_a]\mathbf{g}^{(0)}_{fs}[\tau_f,\tau_u]t_{st}\mathbf{g}^{(0)}_{tu}[\tau_u,\tau_v]t_{uv}\mathbf{g}^{(0)}_{vs}[\tau_v,\tau_u]\mathbf{g}^{(0)}_{su}[\tau_u,\tau_v]\mathbf{g}^{(0)}_{uf}[\tau_v,\tau_f],$$

where all sites and times other than i and f, and τ_i and τ_f are summed/integrated over.

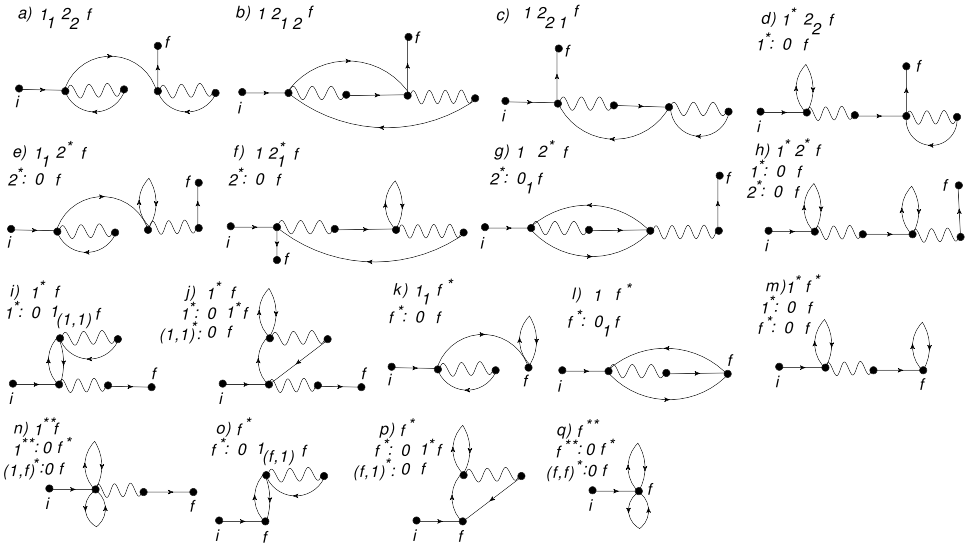

Fig. 17. The second order diagrams contributing to the Green's function: $\mathcal{G}^{(2)}[i,f]$, and their corresponding numerical representations. Note that the diagrams (a) through (j) are the standard second order Feynman diagrams. The other diagrams are not.

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

$$(a)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{ba}[\tau_a,\tau_a^+]\mathbf{g}^{(0)}_{ac}[\tau_a,\tau_b]t_{cd}\mathbf{g}^{(0)}_{dc}[\tau_b,\tau_b^+]\mathbf{g}^{(0)}_{cf}[\tau_b,\tau_f]$$

$$(b)\ -2\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_a,\tau_c]t_{cd}\mathbf{g}^{(0)}_{da}[\tau_c,\tau_a]\mathbf{g}^{(0)}_{ac}[\tau_a,\tau_c]\mathbf{g}^{(0)}_{cf}[\tau_c,\tau_f]$$

$$(c)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_a,\tau_b]t_{cd}\mathbf{g}^{(0)}_{dc}[\tau_b,\tau_b^+]\mathbf{g}^{(0)}_{ca}[\tau_b,\tau_a]\mathbf{g}^{(0)}_{af}[\tau_a,\tau_f]$$

$$(d)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]\mathbf{g}^{(0)}_{aa}[\tau_a,\tau_a^+]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_a,\tau_b]t_{cd}\mathbf{g}^{(0)}_{dc}[\tau_b,\tau_b^+]\mathbf{g}^{(0)}_{cf}[\tau_b,\tau_f]$$

$$(e)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{ba}[\tau_a,\tau_a^+]\mathbf{g}^{(0)}_{ac}[\tau_a,\tau_b]\mathbf{g}^{(0)}_{cc}[\tau_b,\tau_b^+]t_{cd}\mathbf{g}^{(0)}_{df}[\tau_b,\tau_f]$$

$$(f)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_b,\tau_b^+]t_{cd}\mathbf{g}^{(0)}_{da}[\tau_b,\tau_a]\mathbf{g}^{(0)}_{af}[\tau_a,\tau_f]$$

$$(g)\ -2\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_a,\tau_b]\mathbf{g}^{(0)}_{ca}[\tau_b,\tau_a]\mathbf{g}^{(0)}_{ac}[\tau_a,\tau_b]t_{cd}\mathbf{g}^{(0)}_{ia}[\tau_b,\tau_f]$$

$$(h)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]\mathbf{g}^{(0)}_{aa}[\tau_a,\tau_a^+]t_{ab}\mathbf{g}^{(0)}_{bc}[\tau_a,\tau_b]\mathbf{g}^{(0)}_{cc}[\tau_b,\tau_b^+]t_{cd}\mathbf{g}^{(0)}_{df}[\tau_b,\tau_f]$$

$$(i)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bf}[\tau_a,\tau_f]\mathbf{g}^{(0)}_{ac}[\tau_a,\tau_b]t_{cd}\mathbf{g}^{(0)}_{dc}[\tau_b,\tau_b^+]\mathbf{g}^{(0)}_{ca}[\tau_b,\tau_a]$$

$$(j)\ \mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bf}[\tau_a,\tau_f]\mathbf{g}^{(0)}_{ac}[\tau_a,\tau_b]\mathbf{g}^{(0)}_{cc}[\tau_b,\tau_b^+]t_{cd}\mathbf{g}^{(0)}_{da}[\tau_b,\tau_a]$$

$$(k)\ -\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{ba}[\tau_a,\tau_a^+]\mathbf{g}^{(0)}_{af}[\tau_a,\tau_f]\mathbf{g}^{(0)}_{ff}[\tau_f,\tau_f^+]$$

$$(l)\ 2\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]t_{ab}\mathbf{g}^{(0)}_{bf}[\tau_a,\tau_f]\mathbf{g}^{(0)}_{fa}[\tau_f,\tau_a]\mathbf{g}^{(0)}_{af}[\tau_a,\tau_f]$$

$$(m)\ -\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]\mathbf{g}^{(0)}_{aa}[\tau_a,\tau_a^+]t_{ab}\mathbf{g}^{(0)}_{bf}[\tau_a,\tau_f]\mathbf{g}^{(0)}_{ff}[\tau_f,\tau_f^+]$$

$$(n)\ -\mathbf{g}^{(0)}_{ia}[\tau_i,\tau_a]\mathbf{g}^{(0)}_{aa}[\tau_a,\tau_a^+]\mathbf{g}^{(0)}_{aa}[\tau_a,\tau_a^+]t_{ab}\mathbf{g}^{(0)}_{bf}[\tau_a,\tau_f]$$

$$(o)\ -\mathbf{g}^{(0)}_{if}[\tau_i,\tau_f]\mathbf{g}^{(0)}_{fa}[\tau_f,\tau_a]t_{ab}\mathbf{g}^{(0)}_{ba}[\tau_a,\tau_a^+]\mathbf{g}^{(0)}_{af}[\tau_a,\tau_f]$$

$(p)\ -\mathbf{g}^{(0)}_{if}[\tau_i,\tau_f]\mathbf{g}^{(0)}_{fa}[\tau_f,\tau_a]\mathbf{g}^{(0)}_{aa}[\tau_a,\tau_a^+]t_{ab}\mathbf{g}^{(0)}_{bf}[\tau_a,\tau_f]$

$(q)\ \mathbf{g}^{(0)}_{if}[\tau_i,\tau_f]\mathbf{g}^{(0)}_{ff}[\tau_f,\tau_f^+]\mathbf{g}^{(0)}_{ff}[\tau_f,\tau_f^+].$

5. Diagrammatic λ 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 $\mathbf{g}[i,f]$.

$$\mathbf{g}_{\sigma_1,\sigma_2}[i,f]=\mathbf{g}^{(0)}_{\sigma_1,\sigma_2}[i,f]$$
$$-\lambda\,\mathbf{g}^{(0)}_{\sigma_1,\sigma_b}[i,\mathbf{k}]\left(-t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a\mathcal{G}_{\bar\sigma_a,\bar\sigma_b}[\mathbf{k},\mathbf{k}^+]\mathbf{g}_{\sigma_a,\sigma_2}[\mathbf{j},f]+t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a\frac{\delta}{\delta\mathcal{V}^{\bar\sigma_b,\bar\sigma_a}_{\mathbf{k}}}\mathbf{g}_{\sigma_a,\sigma_2}[\mathbf{j},f]\right),\quad(19)$$

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 Σ_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 Σ_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 Σ_a by Σ^*_a.

Now consider those diagrams which do have a star on the f in the top row. The second self-energy, Σ_b, is obtained from these diagrams by removing the external line coming from the point i. Once again, if a diagram for Σ_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 Σ_b by Σ^*_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 $\widetilde{\mu}[i,f]=\delta[i,f]+\Sigma^*_b[i,f]$. 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 $\widetilde{\mu}$ (Eq. (10)) in Section 5.2.

5.2. $\mathbf{g}^{-1}$ and $\widetilde{\mu}$

We shall now prove, starting with the equations of motion in Eq. (10), the observations already made in Section 5.1, that

$$\widetilde{\mu}[i,f]=\delta[i,f]+\Sigma^*_b[i,f];\qquad \mathbf{g}^{-1}[i,f]=\mathbf{g}^{-1(0)}[i,f]-\Sigma^*_a[i,f].\quad(20)$$

This is equivalent to showing that

$$\Sigma^*_b[i,f]=-\lambda\gamma[i]\delta[i,f]+\lambda\Psi[i,f];\qquad \Sigma^*_a[i,f]=\lambda\gamma[i]t[i,f]+\lambda\Phi[i,f].\quad(21)$$

We rewrite the EOM for $\widetilde{\mu}[i,f]$ (Eq. (10)) in expanded form.

$$\widetilde{\mu}_{\sigma_1\sigma_2}[i,f]=(\delta_{\sigma_1\sigma_2}-\lambda\sigma_1\sigma_2\mathcal{G}_{\bar\sigma_2\bar\sigma_1}[i,i^+])\delta[i,f]+\lambda\Psi_{\sigma_1\sigma_2}[i,f]$$
$$\Psi_{\sigma_1\sigma_2}[i,f]=-t[i,\mathbf{j}]\,\sigma_1\sigma_a\mathbf{g}_{\sigma_a,\sigma_b}[\mathbf{j},\mathbf{n}]\frac{\delta}{\delta\mathcal{V}^{\bar\sigma_1\bar\sigma_a}_i}\widetilde{\mu}_{\sigma_b\sigma_2}[\mathbf{n},f].\quad(22)$$

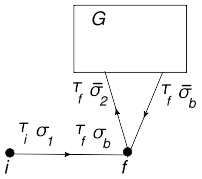

Fig. 18. Schematic representation for a Green's function diagram with only the number $f^{*\cdots*}$ in the top row. Upon removing the incoming external line, it contributes to $\Sigma^*_{b1}[i,f]$.

We now proceed to prove the first of Eq. (20) using induction in λ. The lowest order contribution to $\Sigma^*_b[i,f]$ comes from diagram (a) in Fig. 12. Removing the incoming external line, we obtain $\Sigma^{*(1)}_{\substack{b\\ \sigma_1\sigma_2}}[i,f] = -\lambda\sigma_1\sigma_2\delta[i,f]\mathbf{g}^{(0)}_{\substack{ff\\ \bar{\sigma}_2\bar{\sigma}_1}}[\tau_f,\tau_f^+]$. Using Eq. (22) to obtain the first order contribution to $\widetilde{\mu}[i,f]$, we get $\widetilde{\mu}^{(1)}_{\sigma_1\sigma_2}[i,f] = -\lambda\sigma_1\sigma_2\mathbf{g}^{(0)}_{\bar{\sigma}_2,\bar{\sigma}_1}[i,i^+]\delta[i,f]$. Clearly, these two are equal, and we have that $\Sigma^{*(1)}_{\substack{b\\ \sigma_1\sigma_2}}[i,f] = \widetilde{\mu}^{(1)}_{\sigma_1\sigma_2}[i,f]$.

Now consider the mth order contribution $\Sigma^{*(m)}_{\substack{b\\ \sigma_1\sigma_2}}[i,f]$. This will be obtained from the corresponding mth 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 $f^{*\cdots*}$ 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 $\mathcal{G}^{(m)}_{\sigma_1\sigma_2}[i,f] = -\lambda\,\mathbf{g}^{(0)}_{\sigma_1\sigma_b}[i,f]\sigma_b\sigma_2\mathcal{G}^{(m-1)}_{\bar{\sigma}_2\bar{\sigma}_b}[f,f^+]$ (see Fig. 18). The resulting contribution to Σ^*_b, which we shall denote by Σ^*_{b1}, is

$$\Sigma^{*(m)}_{\substack{b1\\ \sigma_1\sigma_2}}[i,f] = -\lambda\,\sigma_1\sigma_2\mathcal{G}^{(m-1)}_{\bar{\sigma}_2\bar{\sigma}_1}[f,f^+]\delta[i,f]. \tag{23}$$

Alternatively, there are numbers other than $f^{*\cdots*}$ in the top row of the numerical representation of the corresponding mth order $\mathcal{G}[i,f]$ diagram. Then, the top row reads $1\ldots f^{*\cdots*}$ (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 $\mathbf{g}[\mathbf{j},\mathbf{n}].\Sigma^*_b[\mathbf{n},f]$ (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 Σ_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

$$-\lambda\mathbf{g}^{(0)}_{\sigma_1\sigma_b}[i,\mathbf{k}]t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a\mathbf{g}_{\sigma_a\sigma_c}[\mathbf{j},\mathbf{n}]\frac{\delta}{\delta\mathcal{V}^{\bar{\sigma}_b\bar{\sigma}_a}_{\mathbf{k}}}\Sigma^*_{\substack{b\\ \sigma_c\sigma_2}}[\mathbf{n},f]. \tag{24}$$

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 $\Sigma^*_{b2}[i,f]$.

$$\Sigma^{*(m)}_{\substack{b2\\ \sigma_1\sigma_2}}[i,f] = -\lambda t[i,\mathbf{j}]\sigma_1\sigma_a\mathbf{g}^{(m_1)}_{\sigma_a\sigma_b}[\mathbf{j},\mathbf{n}]\frac{\delta}{\delta\mathcal{V}^{\bar{\sigma}_1\bar{\sigma}_a}_i}\widetilde{\mu}^{(m_2)}_{\sigma_b\sigma_2}[\mathbf{n},f], \tag{25}$$

where $m = m_1 + m_2 + 1$. Comparing Eq. (25) with Eq. (22), we see that

$$\Sigma^*_{\substack{b2\\ \sigma_1\sigma_2}}[i,f] = \lambda\Psi_{\sigma_1\sigma_2}[i,f]. \tag{26}$$

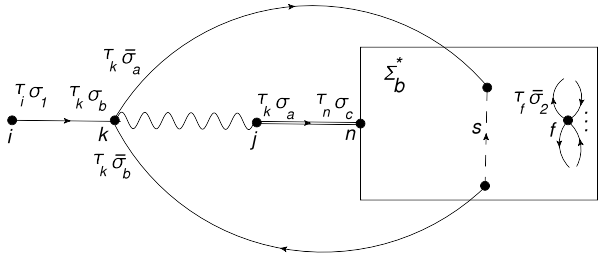

Fig. 19. Schematic representation for a Green's function diagram whose top row is $1\ldots f^{*\cdots *}$. Upon removing the incoming external line, it contributes to $\Sigma^*_{b2}[i,f]$.

Combining Eqs. (23) and (26), we find that

$$\Sigma^*_{b\ \sigma_1\sigma_2}[i,f] = \Sigma^*_{b1\ \sigma_1\sigma_2}[i,f] + \Sigma^*_{b2\ \sigma_1\sigma_2}[i,f] = -\lambda\,\sigma_1\sigma_2 \mathcal{G}_{\bar\sigma_2\bar\sigma_1}[f,f^+]\delta[i,f] + \lambda\Psi_{\sigma_1\sigma_2}[i,f]. \tag{27}$$

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.

$$\begin{aligned}\mathbf{g}^{-1}_{\sigma_1\sigma_2}[i,f] &= \mathbf{g}^{-1(0)}_{\sigma_1\sigma_2}[i,f] - \lambda\, t[i,f]\sigma_1\sigma_2\mathcal{G}_{\bar\sigma_2\bar\sigma_1}[i,i^+] - \lambda\Phi_{\sigma_1\sigma_2}[i,f]\\ \Phi_{\sigma_1\sigma_2}[i,f] &= t[i,\mathbf{j}]\,\sigma_1\sigma_a\mathbf{g}_{\sigma_a,\sigma_b}[\mathbf{j},\mathbf{n}]\frac{\delta}{\delta\mathcal{V}_i^{\bar\sigma_1\bar\sigma_a}}\mathbf{g}^{-1}_{\sigma_b\sigma_2}[\mathbf{n},f].\end{aligned} \tag{28}$$

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 $1\ldots c^{*\cdots *}f$. 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 Σ^*_a by Σ^*_{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 Σ^*_a by Σ^*_{a2}. Comparing Fig. 18 with Fig. 20 and Fig. 19 with Fig. 21, and removing the external lines, we find that

$$\Sigma^*_{a1\ \sigma_1\sigma_2}[i,f] = -\Sigma^*_{b1\ \sigma_1\sigma_2}[i,\mathbf{j}]t[\mathbf{j},f]; \qquad \Sigma^*_{a2\ \sigma_1\sigma_2}[i,f] = -\Sigma^*_{b2\ \sigma_1\sigma_2}[i,\mathbf{j}]t[\mathbf{j},f]. \tag{29}$$

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 $f^{*\cdots *}$ in Σ^*_b) and $(-1)^s$ (applicable to $c^{*\cdots *}$ in Σ^*_a). Using Eqs. (23) and (26), we find that

$$-\Sigma^*_{a1\ \sigma_1\sigma_2}[i,f] = -\lambda\,\sigma_1\sigma_2\mathcal{G}_{\bar\sigma_2\bar\sigma_1}[i,i^+]t[i,f]; \qquad -\Sigma^*_{a2\ \sigma_1\sigma_2}[i,f] = \lambda\Psi_{\sigma_1\sigma_2}[i,\mathbf{j}]t[\mathbf{j},f]. \tag{30}$$

Motivated by this observation, we define a new object $\chi_{\sigma_1\sigma_2}[i,f]$ defined by the formula

$$\Phi_{\sigma_1\sigma_2}[i,f] = -\Psi_{\sigma_1\sigma_2}[i,\mathbf{j}]t[\mathbf{j},f] + \chi_{\sigma_1\sigma_2}[i,f]. \tag{31}$$

Plugging this formula into Eq. (28), we obtain

$$\mathbf{g}^{-1}_{\sigma_1\sigma_2}[i,f] = \mathbf{g}^{-1(0)}_{\sigma_1\sigma_2}[i,f] - \lambda\, t[i,f]\sigma_1\sigma_2\mathcal{G}_{\bar\sigma_2\bar\sigma_1}[i,i^+] + \lambda\Psi_{\sigma_1\sigma_2}[i,\mathbf{j}]t[\mathbf{j},f] - \lambda\chi_{\sigma_1\sigma_2}[i,f] \tag{32}$$

Plugging Eq. (32) into the equation for Φ (Eq. (28)), we obtain

$$\begin{aligned}\Phi_{\sigma_1\sigma_2}[i,f] &= -t[i,\mathbf{j}]\sigma_1\bar\sigma_2\mathbf{g}_{\bar\sigma_2\bar\sigma_1}[\mathbf{j},i]\delta[i,f] - \Psi_{\sigma_1\sigma_2}[i,\mathbf{j}]t[\mathbf{j},f]\\ &\quad - \lambda t[i,\mathbf{j}]\sigma_1\sigma_a\mathbf{g}_{\sigma_a\sigma_b}[\mathbf{j},\mathbf{n}]\frac{\delta}{\delta\mathcal{V}_i^{\bar\sigma_1\bar\sigma_a}}\chi_{\sigma_b\sigma_2}[\mathbf{n},f],\end{aligned} \tag{33}$$

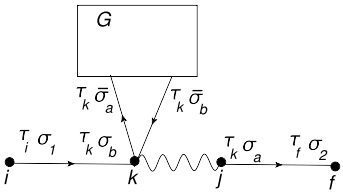

Fig. 20. Schematic representation for Green's function diagram whose top row is $1^{*\cdots *}f$. Upon removing the incoming and outgoing external lines, it contributes to $\Sigma^*_{a1}[i,f]$. One can obtain $\Sigma^*_{b1}[i,f]$ displayed in Fig. 18 by also removing the interaction line exiting the point k.

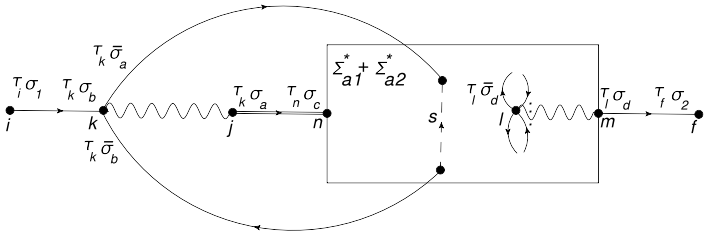

Fig. 21. Schematic representation for Green's function diagram whose top row is $1\ldots c^{*\cdots *}f$. Upon removing the outgoing and incoming external lines, it contributes to $\Sigma^*_{a2}[i,f]$. One can obtain $\Sigma^*_{b2}[i,f]$ displayed in Fig. 19 by also removing the interaction line exiting 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 $\chi_{\sigma_1\sigma_2}[i,f]$.

$$\chi_{\sigma_1\sigma_2}[i,f] = -t[i,\mathbf{j}]\sigma_1\bar\sigma_2\mathbf{g}_{\bar\sigma_2\bar\sigma_1}[\mathbf{j},i]\delta[i,f] - \lambda t[i,\mathbf{j}]\sigma_1\sigma_a\mathbf{g}_{\sigma_a\sigma_b}[\mathbf{j},\mathbf{n}]\frac{\delta}{\delta\mathcal{V}_i^{\bar\sigma_1\bar\sigma_a}}\chi_{\sigma_b\sigma_2}[\mathbf{n},f]. \tag{34}$$

Comparing Eqs. (32) and (30), we see that $-\underset{\sigma_1\sigma_2}{\Sigma^*_{a1}}[i,f]$ and $-\underset{\sigma_1\sigma_2}{\Sigma^*_{a2}}[i,f]$ 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 $1\ldots c_{\ldots}f$ (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 $1\ldots c_{\ldots 1}f$ (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 $1\ldots c_{\ldots d}f$, where $d\neq 1$ (e.g. panels (a), (b), and (d) of Fig. 17). Our goal is to show that $\lambda\chi_{\sigma_1\sigma_2}[i,f] = \underset{\sigma_1\sigma_2}{\Sigma^*_{a3}}[i,f] + \underset{\sigma_1\sigma_2}{\Sigma^*_{a4}}[i,f]$.

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

$$-\lambda\sigma_a\sigma_b\underset{\sigma_1\sigma_a}{\mathbf{g}^{(0)}_{\mathbf{ia}}}[\tau_i,\tau_{\mathbf{a}}]t_{\mathbf{ab}}\underset{\sigma_b\bar\sigma_a}{\mathbf{g}_{\mathbf{ba}}}[\tau_{\mathbf{a}},\tau^+_{\mathbf{a}}]\underset{\bar\sigma_b\sigma_2}{\mathbf{g}^{(0)}_{\mathbf{a}f}}[\tau_{\mathbf{a}},\tau_f]. \tag{35}$$

After removing the two external lines, we find that

$$\underset{\sigma_1\sigma_2}{\Sigma^*_{a3}}[i,f] = -\lambda\sigma_1\bar\sigma_2 t_{i\mathbf{b}}\underset{\bar\sigma_2\bar\sigma_1}{\mathbf{g}_{\mathbf{bi}}}[\tau_i,\tau_i^+]\delta[i,f]. \tag{36}$$

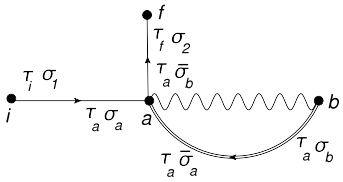

Fig. 22. Schematic representation for Green's function diagram whose top row is $1 \ldots c_{\ldots 1} f$. Upon removing the incoming and outgoing external lines, it contributes to $\Sigma^*_{a3}[i, f]$. This is the analog of the Fock diagram in the standard Feynman skeleton expansion.

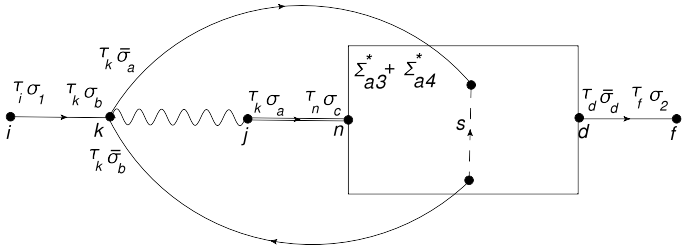

Fig. 23. Schematic representation for Green's function diagram whose top row is $1 \ldots c_{\ldots d} f$, where $d \neq 1$. Upon removing the incoming and outgoing external lines, it contributes to $\Sigma^*_{a4}[i, f]$.

Thus, $\Sigma^*_{\substack{a3 \\ \sigma_1\sigma_2}}[i, f]$ 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 $\Sigma^*_{\substack{a4 \\ \sigma_1\sigma_2}}[i, f]$ equals the second term on the RHS of Eq. (34). The $\mathbf{g}[i, f]$ diagrams contributing to $\Sigma^*_{\substack{a4 \\ \sigma_1\sigma_2}}[i, f]$ 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 $\mathbf{g}_{\substack{\mathbf{jn} \\ \sigma_a\sigma_c}}[\tau_{\mathbf{k}}, \tau_{\mathbf{n}}]$, then the resulting Σ_a (of the overall diagram) would be reducible, while if s was the bare line $\mathbf{g}^{(0)}_{\substack{\mathbf{df} \\ \bar{\sigma}_d\sigma_2}}[\tau_{\mathbf{d}}, \tau_f]$, the diagram would contribute to Σ^*_{a3} (see Fig. 22). The box can be either a Σ^*_{a3} insertion or a Σ^*_{a4} insertion, but cannot be a Σ^*_{a1} insertion or a Σ^*_{a2} insertion, since in this case the diagram would contribute to Σ^*_{a2} (see Fig. 21). The analytical contribution of Fig. 23 is

$$-\lambda \mathbf{g}^{(0)}_{\sigma_1\sigma_b}[i, \mathbf{k}] t[\mathbf{k}, \mathbf{j}] \sigma_b \sigma_a \mathbf{g}_{\sigma_a\sigma_c}[\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}^{\bar{\sigma}_b\bar{\sigma}_a}_{\mathbf{k}}} \left(\Sigma^*_{\substack{a3 \\ \sigma_c\bar{\sigma}_d}}[\mathbf{n}, \mathbf{d}] + \Sigma^*_{\substack{a4 \\ \sigma_c\bar{\sigma}_d}}[\mathbf{n}, \mathbf{d}] \right) \mathbf{g}^{(0)}_{\bar{\sigma}_d\sigma_2}[\mathbf{d}, f]. \tag{37}$$

Dropping the external lines, and using the inductive hypothesis, we obtain

$$\Sigma^*_{\substack{a4 \\ \sigma_1\sigma_2}}[i, f] = -\lambda^2 t[i, \mathbf{j}] \sigma_1 \sigma_a \mathbf{g}_{\sigma_a\sigma_b}[\mathbf{j}, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}^{\bar{\sigma}_1\bar{\sigma}_a}_i} \chi_{\sigma_b\sigma_2}[\mathbf{n}, f]. \tag{38}$$

Combining this with Eq. (36) and comparing with Eq. (34), we find that

$$\Sigma^*_{\substack{a3 \\ \sigma_1\sigma_2}}[i, f] + \Sigma^*_{\substack{a4 \\ \sigma_1\sigma_2}}[i, f] = \lambda \chi_{\sigma_1\sigma_2}[i, f]. \tag{39}$$

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

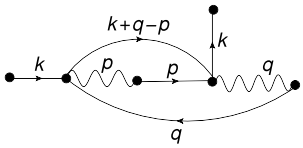

Fig. 24. Momentum space representation of diagram for $\mathbf{g}(k)$ from Fig. 17(b). Upon removing the incoming and outgoing external lines, it contributes to $\chi(k)$.

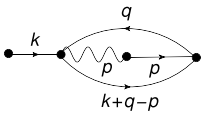

Fig. 25. Momentum space representation of diagram for $\mathcal{G}(k)$ from Fig. 17(l). Upon removing the incoming external line, it contributes to $\Psi(k)$.

by the generic symbol $Q[i,f]$, as

$$Q[i,f] = \frac{1}{N_s\beta}\sum_k e^{ik(i-f)}Q(k), \tag{40}$$

where N_s is the number of sites on the lattice, β 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

$$-2\,\mathbf{g}^{(0)}(k)\epsilon_p\mathbf{g}^{(0)}(p)\epsilon_q\mathbf{g}^{(0)}(q)\mathbf{g}^{(0)}(k+q-p)\mathbf{g}^{(0)}(k) \tag{41}$$

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 $\chi(k)$:

$$-2\,\epsilon_p\mathbf{g}^{(0)}(p)\epsilon_q\mathbf{g}^{(0)}(q)\mathbf{g}^{(0)}(k+q-p). \tag{42}$$

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

$$-2\,\mathbf{g}^{(0)}(k)\epsilon_p\mathbf{g}^{(0)}(p)\mathbf{g}^{(0)}(q)\mathbf{g}^{(0)}(k+q-p) \tag{43}$$

Upon removing the incoming external line, we obtain the following contribution to $\Sigma_b^*(k)$, or equivalently to $\Psi(k)$:

$$-2\,\epsilon_p\mathbf{g}^{(0)}(p)\mathbf{g}^{(0)}(q)\mathbf{g}^{(0)}(k+q-p). \tag{44}$$

5.4. The vertices Λ 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.

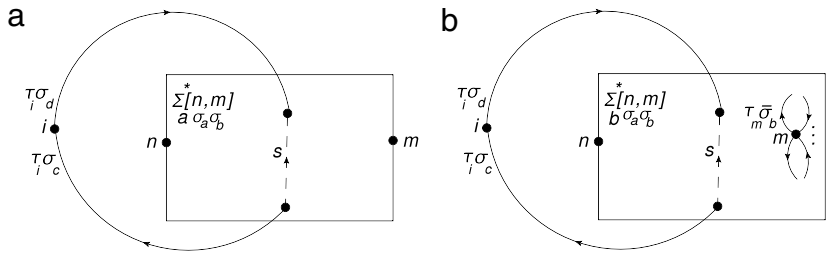

Fig. 26. Schematic diagram for the vertices. $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ is displayed in panel (a) while $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ is displayed in panel (b).

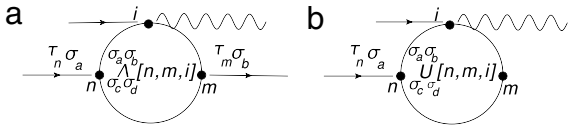

Fig. 27. Schematic diagram for the vertices. $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ is displayed in panel (a) while $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ is displayed in panel (b).

$$\mathbf{g}^{-1}[i, m] = (\mu - \partial_{\tau_i} - \mathcal{V}_i)\,\delta[i, m] + t[i, m]\,(1 - \lambda\gamma[i]) - \lambda\Phi[i, m].$$
$$\widetilde{\mu}[i, m] = (1 - \lambda\gamma[i])\delta[i, m] + \lambda\Psi[i, m]$$
$$\Phi[i, m] = -t[i, \mathbf{j}]\,\xi^*.\mathbf{g}[\mathbf{j}, \mathbf{n}].\Lambda_*[\mathbf{n}, m; i]; \quad \Psi[i, m] = -t[i, \mathbf{j}]\,\xi^*.\mathbf{g}[\mathbf{j}, \mathbf{n}].\mathcal{U}_*[\mathbf{n}, m; i]. \tag{45}$$
$$\gamma[i] = \widetilde{\mu}^{(k)}[\mathbf{n}, i^+].\mathbf{g}^{(k)}[i, \mathbf{n}]; \quad \Lambda[n, m; i] = -\frac{\delta}{\delta\mathcal{V}_i}\mathbf{g}^{-1}[n, m]; \quad \mathcal{U}[n, m; i] = \frac{\delta}{\delta\mathcal{V}_i}\widetilde{\mu}[n, m]. \tag{46}$$

We now examine the vertices $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i] \equiv -\frac{\delta}{\delta\mathcal{V}_i^{\sigma_c\sigma_d}}\mathbf{g}^{-1}_{\sigma_a\sigma_b}[n, m]$ and $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i] \equiv \frac{\delta}{\delta\mathcal{V}_i^{\sigma_c\sigma_d}}\widetilde{\mu}_{\sigma_a\sigma_b}[n, m]$ in more detail. The zeroth order vertices, also called the bare vertices, are given by

$$\Lambda^{(0)\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i] = \delta[n, m]\delta[n, i]\delta_{\sigma_a\sigma_c}\delta_{\sigma_b\sigma_d}; \qquad \mathcal{U}^{(0)\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i] = 0. \tag{47}$$

The higher order terms contributing to $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ arise from splitting a line in $\Sigma^*_{\substack{a\\ \sigma_a\sigma_b}}[n, m]$ through the point i. The higher order terms contributing to $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ arise from splitting a line in $\Sigma^*_{\substack{b\\ \sigma_a\sigma_b}}[n, m]$ through the point i. These terms can be represented schematically as in Fig. 26.

From Fig. 26(a), we see that in $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$, 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 $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$, 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 Φ and Ψ are expressed in terms of the vertices Λ 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

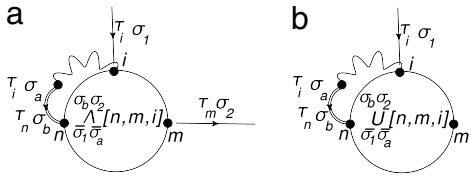

Fig. 28. Schematic diagram for the self-energies in terms of the vertices. In panel (a) $\Phi_{\sigma_1\sigma_2}[i, m]$ is expressed in terms of $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$ and in panel (b) $\Psi_{\sigma_1\sigma_2}[i, m]$ is expressed in terms of $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[n, m; i]$.

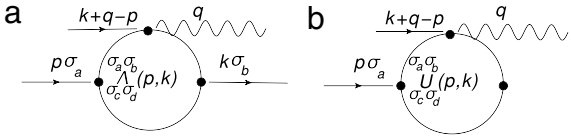

Fig. 29. Schematic diagram for the vertices in momentum space. $\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}(p, k)$ is displayed in panel (a) while $\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}(p, k)$ is displayed in panel (b).

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:

$$Q[n, m, i] = \sum_{kp} e^{ipn} e^{-ikm} e^{i(k-p)i} Q(p, k) = \sum_{kp} e^{ip(n-i)} e^{ik(i-m)} Q(p, k). \tag{48}$$

Furthermore, there are only four non-zero spin configurations contributing to the vertex. These are $Q^{(1)} \equiv Q^{\sigma\sigma}_{\sigma\sigma}$, $Q^{(2)} \equiv Q^{\sigma\bar{\sigma}}_{\bar{\sigma}\sigma}$, $Q^{(3)} \equiv Q^{\sigma\bar{\sigma}}_{\sigma\bar{\sigma}}$, and $Q^{(4)} \equiv Q^{\sigma\bar{\sigma}}_{\bar{\sigma}\sigma}$. These four spin configurations are related by the equation

$$Q^{(1)} - Q^{(2)} = Q^{(3)} + Q^{(4)}. \tag{49}$$

We shall now state the rules for computing the $Q^{(i)}$ and derive Eq. (49). Recall that to obtain a diagram for Λ ($\mathcal{U}$), we must split a line in the self-energy Σ^*_a (Σ^*_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 Σ^*_a (Σ^*_b) are partitioned into anywhere between 0 and F_s spin loops, where the zeroth loop contains the lines with the labels σ_1 and σ_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 σ_1).

Now, in the case that the line split in going from $\Sigma^* \to Q$ 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 Σ^*. In the case that the line split in going from $\Sigma^* \to Q$ is from the zeroth loop, the line split could either carry spin σ_1 in the original Σ^* 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 Σ^*. 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 Σ^*. 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, $Q^{(4)} \to 0$, 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:

$$\Phi(k) = \sum_p \epsilon_p \mathbf{g}(p) \Lambda^{(a)}(p, k); \qquad \Psi(k) = \sum_p \epsilon_p \mathbf{g}(p) \mathcal{U}^{(a)}(p, k). \tag{50}$$

These relations are represented diagrammatically in Fig. 30.

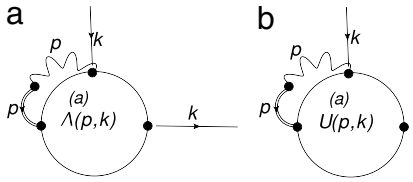

Fig. 30. Schematic diagram for the self-energies in terms of the vertices. In panel (a) $\Phi(k)$ is expressed in terms of $\Lambda^{(a)}(p, k)$ and in panel (b) $\Psi(k)$ is expressed in terms of $\mathcal{U}^{(a)}(p, k)$.

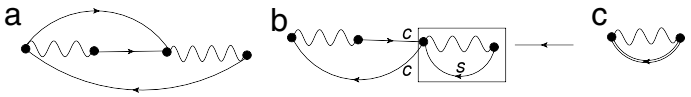

Fig. 31. Examples of skeleton and non-skeleton diagrams for the irreducible self-energy Σ_a^*. The diagram in panel (a) is a skeleton diagram. The diagram in panel (b) is not, since we can isolate a self-energy insertion by cutting the two lines labeled by c. The non-skeleton diagram in panel (b) can 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 Λ.

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 Σ_a from the rest of the diagram by cutting two Green's function lines. For example, consider the Σ_a^* diagrams in Fig. 31 (the same considerations will apply to Σ_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 Σ_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 Σ_a self-energy insertion enclosed in the box. In contrast, the Σ_a^* diagram in panel (a) of Fig. 31 is a skeleton diagram, since it is impossible to isolate a Σ_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 Σ_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 $\mathbf{g}^{(0)}[i, f] \to \mathbf{g}[i, f]$.

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 Σ_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 Λ 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 Λ.

Taking this reasoning a step further, consider the diagram for Σ_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 Λ will be a non-skeleton diagram, while if we split the line

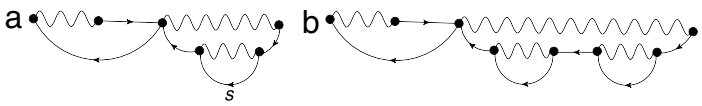

Fig. 32. In this case both Σ_a^* diagrams displayed in panels (a) and (b) are non-skeleton diagrams. However, the diagram in panel (a) contains only irreducible self-energy insertions, while the one in panel (b) contains a reducible self-energy insertion. One can obtain a skeleton diagram for the vertex Λ 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 Λ from the diagram in panel (b) regardless of which line we split.

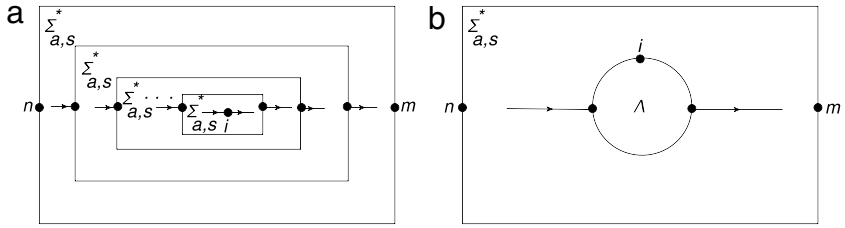

Fig. 33. Panel (a) demonstrates the general procedure for obtaining a skeleton diagram for the vertex Λ from a Σ_a^* diagram consisting of a sequence of Σ_a^* skeleton diagrams. The original Σ_a^* diagram is itself a skeleton diagram only if there is only one skeleton diagram in the sequence, i.e. the Σ_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 Λ. Therefore, to obtain a skeleton diagram for Λ, one must insert the full vertex Λ into a green's function line of a skeleton diagram for Σ_a^*. This is displayed in panel (b). In the case that the original Σ_a^* diagram in panel (a) is itself a skeleton diagram (i.e. there is only one skeleton diagram in the sequence), the Λ vertex in panel (b) is a bare vertex.

labeled by s, the resulting diagram for Λ will be a skeleton diagram. Meanwhile, for the Σ_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 Λ 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 Σ_a^*. Then, insert into at most line of that diagram, a skeleton diagram for Σ_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.

$$\mathbf{g}^{-1} \equiv \mathbf{g}^{-1}[\mathbf{g}]; \qquad \widetilde{\mu} \equiv \widetilde{\mu}[\mathbf{g}], \tag{51}$$

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).

$$\Lambda \equiv \Lambda[\mathbf{g}]; \qquad \mathcal{U} \equiv \mathcal{U}[\mathbf{g}]. \tag{52}$$

The third is the skeleton expansion for the vertices in terms of the auxiliary Green's function and the full vertex Λ. This is the skeleton expansion represented in Fig. 33(b).

$$\Lambda \equiv \Lambda[\mathbf{g}, \Lambda]; \qquad \mathcal{U} \equiv \mathcal{U}[\mathbf{g}, \Lambda]. \tag{53}$$

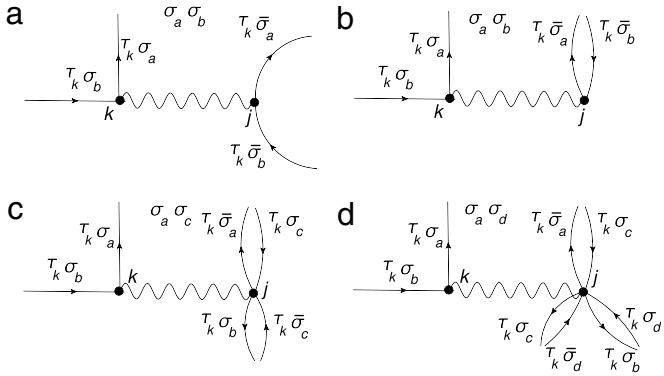

Fig. 34. The J-vertices in the diagrams of the λ expansion. They are more reminiscent of the Feynman diagram vertices than the t-vertices displayed in Fig. 13. The two types of vertices can be obtained from each other by interchanging lines between the two points of the vertex.

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 mth order in λ. Then, plugging the mth 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 kth 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.

$$\begin{aligned}
\mathcal{G}_{\sigma_1,\sigma_2}[i,f] &= \mathbf{g}^{(0)}_{\sigma_1,\sigma_2}[i,f] - \lambda\, \mathbf{g}^{(0)}_{\sigma_1,\sigma_b}[i,f]\sigma_b\sigma_2 \mathcal{G}_{\bar{\sigma}_2,\bar{\sigma}_b}[f,f^+] \\
&- \lambda\, \mathbf{g}^{(0)}_{\sigma_1,\sigma_b}[i,\mathbf{k}] \left(-t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a \mathcal{G}_{\bar{\sigma}_a,\bar{\sigma}_b}[\mathbf{k},\mathbf{k}^+]\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f] + t[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a \frac{\delta}{\delta \mathcal{V}^{\bar{\sigma}_b,\bar{\sigma}_a}_{\mathbf{k}}} \mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{j},f] \right), \\
&- \lambda\, \mathbf{g}^{(0)}_{\sigma_1,\sigma_b}[i,\mathbf{k}] \left(\frac{1}{2} J[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a \mathcal{G}_{\bar{\sigma}_a,\bar{\sigma}_b}[\mathbf{j},\mathbf{j}^+]\mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{k},f] - \frac{1}{2} J[\mathbf{k},\mathbf{j}]\sigma_b\sigma_a \frac{\delta}{\delta \mathcal{V}^{\bar{\sigma}_b,\bar{\sigma}_a}_{\mathbf{j}}} \mathcal{G}_{\sigma_a,\sigma_2}[\mathbf{k},f] \right). \qquad (54)
\end{aligned}$$

The λ-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 σ_a from the point j to the point k, and moving all lines but the one labeled by σ_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.

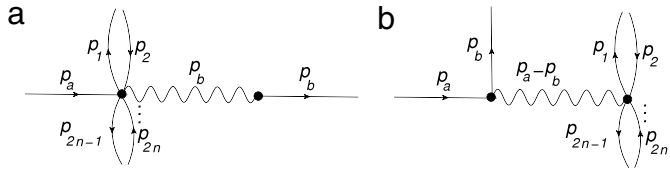

Fig. 35. The t and J vertices are displayed in momentum space in panels (a) and (b) respectively. For each interaction vertex of a diagram, we can choose to use either one as long as we associate with it the factor $\epsilon_{p_b} + \frac{1}{2}J_{p_a-p_b}$.

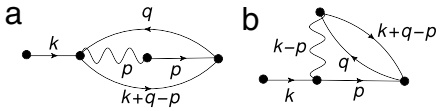

Fig. 36. The $\mathcal{G}(k)$ diagrams drawn in panel (a) and panel (b) correspond to the same diagram. The one in panel (a) is drawn using a t-vertex, while the one in panel (b) is drawn using a J-vertex. In both cases, we associate the factor $\epsilon_p + \frac{1}{2}J_{k-p}$ with the interaction vertex.

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

$$p_a + \sum_{m=1}^{n} p_{2m} = p_b + \sum_{m=1}^{n} p_{2m-1}. \tag{55}$$

In Fig. 35(a), the interaction line contributes a factor of ϵ_{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].

$$\begin{aligned}
&\mathbf{g}^{-1}[i,m] = (\mu - \partial_{\tau_i} - \mathcal{V}_i)\,\delta[i,m] + t[i,m]\,(1-\lambda\gamma[i]) + \frac{\lambda}{2}J[i,\mathbf{j}]\gamma[\mathbf{j}]\delta[i,m] - \lambda\Phi[i,m],\\
&\widetilde{\mu}[i,m] = (1-\lambda\gamma[i])\delta[i,m] + \lambda\Psi[i,m],\\
&\Phi[i,m] = \mathbf{L}[i,\mathbf{n}].\mathbf{g}^{-1}[\mathbf{n},m]; \qquad \Psi[i,m] = -\mathbf{L}[i,\mathbf{n}].\widetilde{\mu}[\mathbf{n},m],
\end{aligned} \tag{56}$$

where the operator $\mathbf{L}$ is given by:

$$\mathbf{L}_{\sigma_1\sigma_2}[i,m] = t[i,\mathbf{j}]\sigma_1\sigma_a\mathbf{g}_{\sigma_a\sigma_2}[\mathbf{j},m]\frac{\delta}{\delta\mathcal{V}_i^{\bar{\sigma}_1\bar{\sigma}_a}} - \frac{1}{2}J[i,\mathbf{j}]\sigma_1\sigma_a\mathbf{g}_{\sigma_a\sigma_2}[i,m]\frac{\delta}{\delta\mathcal{V}_{\mathbf{j}}^{\bar{\sigma}_1\bar{\sigma}_a}}. \tag{57}$$

Using the same decomposition as in Eq. (31), i.e.

$$\Phi[i,m] = -\Psi[i,\mathbf{j}]t[\mathbf{j},m] + \chi[i,m], \tag{58}$$

we find that

$$\chi[i,m] = \mathbf{L}[i,\mathbf{n}].\mathbf{g}^{-1(0)}[\mathbf{n},m] + \frac{\lambda}{2}J[m,\mathbf{k}]\mathbf{L}[i,m]\gamma[\mathbf{k}] - \lambda\mathbf{L}[i,\mathbf{n}]\chi[\mathbf{n},m], \tag{59}$$

where

$$(\mathbf{L}[i,\mathbf{n}].\mathbf{g}^{-1(0)}[\mathbf{n},m])_{\sigma_1\sigma_2} = t[i,\mathbf{j}]\sigma_1\sigma_2\mathbf{g}_{\bar{\sigma}_2\bar{\sigma}_1}[\mathbf{j},i]\delta[i,m] - \frac{1}{2}J[i,m]\sigma_1\sigma_2\mathbf{g}_{\bar{\sigma}_2\bar{\sigma}_1}[i,m]. \tag{60}$$

Finally, we note that the equations associated with Fig. 30 now become

$$\Phi(k) = \sum_p \left(\epsilon_p + \frac{1}{2}J_{k-p}\right)\mathbf{g}(p)\Lambda^{(a)}(p,k); \qquad \Psi(k) = \sum_p \left(\epsilon_p + \frac{1}{2}J_{k-p}\right)\mathbf{g}(p)\mathcal{U}^{(a)}(p,k). \tag{61}$$

7. Finite order calculations

7.1. Zeroth through third order calculation

In this section, we compute the skeleton expansion for the objects γ, Ψ, and χ through second order in λ in momentum space. As can be seen from Eq. (63), this yields the skeleton expansion for $\mathbf{g}^{-1}$ and $\widetilde{\mu}$ through third order in λ. 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 μ. However, the same is not true of the constituent factors $\mathbf{g}$ and $\widetilde{\mu}$, 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 $\widetilde{\mu}$ (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 $\widetilde{\mu}$ 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 μ) is a consequence of the shift identity. Furthermore, the two chemical potentials μ and u_0 can now be used to satisfy the two sum rules

$$\sum_k \mathcal{G}(k) = \frac{n}{2}; \qquad \sum_k \mathbf{g}(k) = \frac{n}{2}. \tag{62}$$

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 $\widetilde{\mu}$ through third order in λ. Taking the Fourier transform of Eqs. (9) and (56), and using Eq. (58), we obtain

$$\begin{aligned}
&\mathbf{g}^{-1}(k) = i\omega_k + \mu' - (\epsilon_k - \frac{u_0}{2})\widetilde{\mu}(k) - \lambda\chi(k),\\
&\mu' = \mu - \frac{u_0}{2} + \frac{\lambda}{2}\gamma J_0\\
&\widetilde{\mu}(k) = (1 - \lambda\gamma) + \lambda\Psi(k),\\
&\mathcal{G}(k) = \mathbf{g}(k)\widetilde{\mu}(k),
\end{aligned} \tag{63}$$

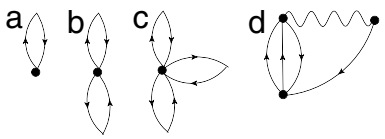

Fig. 37. Second order skeleton expansion for γ. Only the diagram in panel (a) is a standard Feynman diagram. $\gamma^{(0)}$ is given by the diagram in panel (a), $\gamma^{(1)}$ is given by the diagram in panel (b), and $\gamma^{(2)}$ is given by diagrams in panels (c) and (d). We conserve momentum at 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 γ, Ψ, and χ through second order in λ (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 μ and u_0 are then determined by the sum rules in Eq. (62).

In Fig. 37, we have drawn the skeleton diagrams for γ (which is just a constant when the sources are off) through second order in λ. Therefore, γ is the sum of the following terms

$$
\begin{aligned}
&(a)\ \frac{n}{2}\\
&(b)\ -\lambda\left(\frac{n}{2}\right)^2\\
&(c)\ \lambda^2\left(\frac{n}{2}\right)^3\\
&(d)\ -2\lambda^2\sum_{plq}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(p+l-q)\left(\epsilon_q-\frac{u_0}{2}+\frac{1}{2}J_{p-q}\right).
\end{aligned}
\tag{64}
$$

In Fig. 38, we have done the same for $\Psi(k)$. Therefore, $\Psi(k)$ is the sum of the following terms

$$
\begin{aligned}
&(a)\ -2\lambda\sum_{pq}\mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)\\
&(b)\ -4\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+q-l)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)\left(\epsilon_l-\frac{u_0}{2}+\frac{1}{2}J_{p-l}\right)\\
&(c)\ -\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(q+l-p)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)\left(\epsilon_l-\frac{u_0}{2}+\frac{1}{2}J_{p-l}\right)\\
&(d)\ -\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(p+l-q)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)\left(\epsilon_q-\frac{u_0}{2}+\frac{1}{2}J_{p-q}\right)\\
&(e)\ -\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)\left(\epsilon_q-\frac{u_0}{2}+\frac{1}{2}J_{l-q}\right)\\
&(f)\ -\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(l+p-k)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)\left(\epsilon_l-\frac{u_0}{2}+\frac{1}{2}J_{p-k}\right)\\
&(g)\ \lambda^2\frac{n}{2}\sum_{pq}\mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p)\left(\epsilon_p-\frac{u_0}{2}+\frac{1}{2}J_{k-p}\right)
\end{aligned}
\tag{65}
$$

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 ϵ 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,

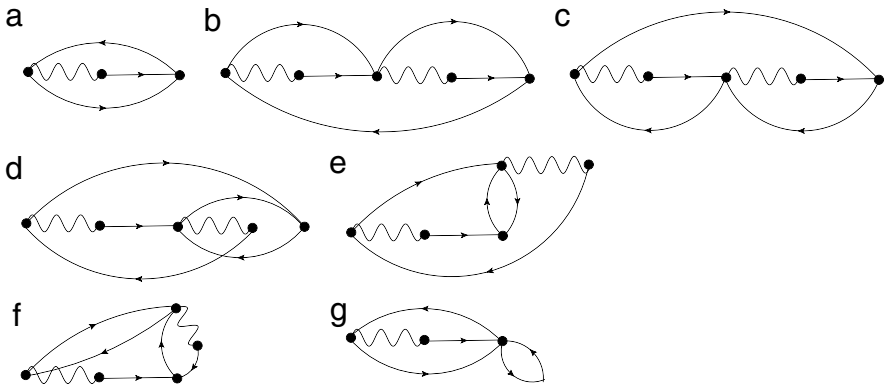

Fig. 38. Second order skeleton expansion for $\Psi(k)$. All diagrams but the one in panel (g) are standard Feynman diagrams (with one interaction line set to unity). $\Psi^{(1)}$ is given by the diagram in panel (a), and $\Psi^{(2)}$ is given by the diagrams in panels (b) through (g). We conserve momentum at each interaction vertex as indicated in Figs. 35 and 36.

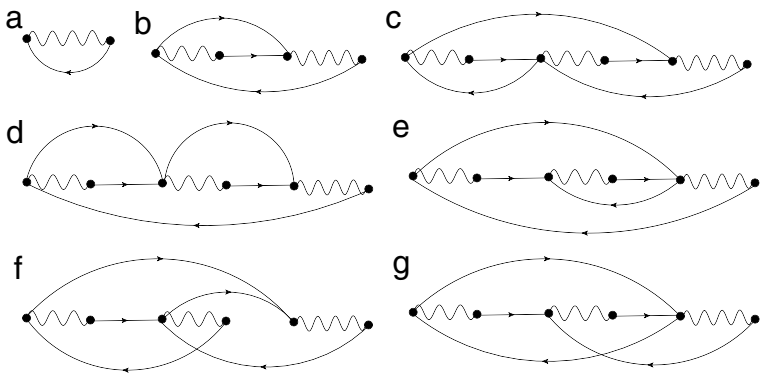

Fig. 39. Second order skeleton expansion for $\chi_A(k)$. These diagrams are independent of those for $\Psi(k)$. All diagrams are standard Feynman diagrams. The diagram in panel (a) contributes to $\chi^{(0)}$, the diagram in panel (b) contributes to $\chi^{(1)}$, and the diagrams in panels (c) through (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).

$$(a)\ -\sum_p \mathbf{g}(p)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)$$

$$(b)\ -2\lambda\sum_{pq}\mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q}\right)$$

$$(c)\ -\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(l+q-p)$$
$$\times\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l}\right)\left(\epsilon_{l+q-p} - \frac{u_0}{2} + \frac{1}{2}J_{p-q}\right)$$

$$(d)\ -4\lambda^2\sum_{pql}\mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+q-l)$$

$$\times \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l}\right)\left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q}\right)$$

$$(e)\ -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p)$$

$$\times \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l}\right)\left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q}\right)$$

$$(f)\ -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(p+l-q)$$

$$\times \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q}\right)\left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{k+q-p-l}\right)$$

$$(g)\ -\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p)$$

$$\times \left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l}\right)\left(\epsilon_{k+l-p} - \frac{u_0}{2} + \frac{1}{2}J_{p-k}\right) \tag{66}$$

$$(a)\ -\lambda \sum_{pq} \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)J_{p-k}$$

$$(b)\ -2\lambda^2 \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+q-l)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l})J_{l-k}$$

$$(c)\ -\frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(q+l-p)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-l})J_{l-k}$$

$$(d)\ -\frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(p+l-q)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{p-q}\right)J_{q-p}$$

$$(e)\ -\frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(k+l-p)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_q - \frac{u_0}{2} + \frac{1}{2}J_{l-q}\right)J_{p-k}$$

$$(f)\ -\frac{\lambda^2}{2} \sum_{pql} \mathbf{g}(p)\mathbf{g}(l)\mathbf{g}(q)\mathbf{g}(k+q-p)\mathbf{g}(l+p-k)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)\left(\epsilon_l - \frac{u_0}{2} + \frac{1}{2}J_{p-k}\right)J_{p-k}$$

$$(g)\ \lambda^2 \frac{n}{4} \sum_{pq} \mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(k+q-p)\left(\epsilon_p - \frac{u_0}{2} + \frac{1}{2}J_{k-p}\right)J_{p-k} \tag{67}$$

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 $\lim_{i\omega_k\to\infty}\ \mathcal{G}(k) = \frac{1-\frac{n}{2}}{i\omega_k}$. From Eq. (63), we see that the high frequency limit of the Green's function can also be expressed as $\lim_{i\omega_k\to\infty}\ \mathcal{G}(k) = \frac{1-\lambda\gamma}{i\omega_k}$. 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 $\widetilde{\mu}(k)$ to mth order in λ, we must calculate γ, $\Psi(k)$, 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 μ and u_0 in λ [25].

$$\mu = \mu^{(0)} + \mu^{(1)} + \cdots \qquad u_0 = u_0^{(0)} + u_0^{(1)} + \cdots, \tag{68}$$

where $\mu^{(0)}$ is zeroth order in λ, $\mu^{(1)}$ is first order in λ, etc. Denoting $\mathbf{g}$, $\widetilde{\mu}$, γ, Ψ, and χ by the generic symbol Q, and plugging the expansions from Eq. (68) into the bare expansion for $Q^{(m)} = Q^{(m)}(\mu, u_0)$,

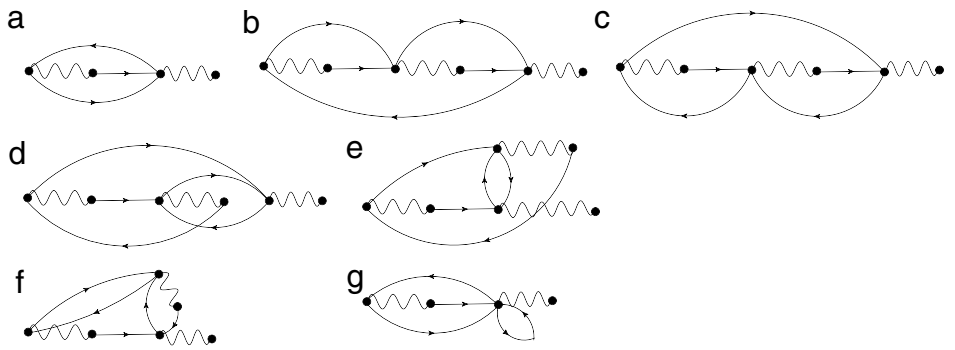

Fig. 40. Second order skeleton expansion for $\chi_B(k)$, which vanishes when $J = 0$. These diagrams can be obtained from those for $\Psi(k)$ in Fig. 38 by adding an interaction line to the terminal point of those diagrams. However, this interaction line contributes only a factor of J, and not a factor of ϵ. 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 $\mathbf{g}^{(m)} = \mathbf{g}^{(m)}(n)$,

$$\gamma^{(m)} = \delta_{m,0}\frac{n}{2}; \qquad \sum_k \mathbf{g}^{(m)}(k) = \delta_{m,0}\frac{n}{2}. \tag{69}$$

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 $\lim_{i\omega_k\to\infty} \mathcal{G}(k) = \frac{1-\lambda\gamma}{i\omega_k} = \frac{1-\lambda\frac{n}{2}}{i\omega_k}$.

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 $\mathbf{g}$, μ, and u_0. The other objects are then obtained from these. In this case, if the skeleton expansions for γ, $\Psi(k)$, 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 $\lim_{i\omega_k\to\infty} \mathcal{G}(k) = \frac{1-\sum_{l=0}^{m-1}\gamma^{(l)}}{i\omega_k}$. Therefore, the error in the high frequency limit is equal to $\gamma^{(m)}$, and we have that

$$\lim_{i\omega_k\to\infty} \mathcal{G}(k) = \frac{1-\frac{n}{2}+\gamma^{(m)}}{i\omega_k}. \tag{70}$$

This error vanishes as $m \to \infty$.

7.3. Analysis of the λ expansion: Feynman type diagrams and non-Feynman diagrams.

The λ 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. Ψ forms a subset of Φ (except for a missing interaction line which is not important for the present discussion), and hence all considerations which apply to Φ will apply equally well to Ψ. 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 λ 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 γ must sum to $\frac{n}{2}$, we see that to "leading order", the only difference between the λ 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 $\gamma \to \frac{n}{2}$, 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 $t \to 0$ and $J \to 0$ 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 β. 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

$$\mathcal{G}_\sigma = \frac{\langle F^{\sigma 0}\rangle + \Delta_\sigma}{(G^0_\sigma)^{-1} - \Sigma_\sigma}. \tag{71}$$

We can make the identifications

$$\langle F^{\sigma 0}\rangle \to 1-\gamma; \qquad \Delta_\sigma \to \Psi(k); \qquad (G^0_\sigma)^{-1} \to \mathbf{g}^{-1(0)}; \qquad \Sigma_\sigma \to -\epsilon_k\gamma + \Phi(k). \tag{72}$$

As is the case in the λ 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 λ. In the case of the high-temperature series, it is βt and βJ. To see this more explicitly, consider the simplest diagram in both series, which is the zeroth order diagram for γ. 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 $\mathbf{g}_{t\to 0, J\to 0}(i\omega_k) = \frac{1}{i\omega_k+\mu}$. The big dot indicates the atomic limit value of γ, i.e. $\gamma_{t\to 0, J\to 0} = \frac{\rho}{2}$, where $\rho = \frac{2e^{\beta\mu}}{1+2e^{\beta\mu}}$ is the atomic limit density. The wiggly line indicates a hoping ϵ_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 γ from the high-temperature series is expanded as an infinite series in λ. Here, each loop corresponds to $\sum_{i\omega_k} \mathbf{g}_{t\to 0, J\to 0}(i\omega_k) = \frac{\frac{\rho}{2}}{1-\frac{\rho}{2}}$, 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 γ from the λ series is expanded as an infinite series in the hopping ϵ_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 λ series, one would have to break up all atomic limit objects into an infinite series in terms of λ, 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 λ

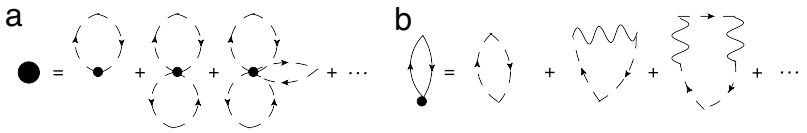

Fig. 41. In panel (a), the zeroth order γ diagram from the high-temperature series (the big dot) is expanded as an infinite series in λ. The dashed lines indicate auxiliary Green's functions in the atomic limit $\mathbf{g}_{t\to 0,J\to 0}(i\omega_k)$. In panel (b), the zeroth order γ diagram from the λ 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 ϵ_k.

expansion at zeroth order, by expanding around the free Fermi gas and by maintaining continuity in λ. 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 λ 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 nth 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 Σ_a and Σ_b. Using the Schwinger equations of motion defining the ECFL objects, $\mathbf{g}$, $\widetilde{\mu}$, γ, Φ, and Ψ, we derived diagrammatic rules for calculating these objects and found that they could be related simply to Σ_a^* and Σ_b^*, the irreducible parts of Σ_a and Σ_b. We also discovered diagrammatically that Ψ diagrams are simply a subset of the Φ diagrams, with an interaction line missing. Denoting the remainder of the Φ diagrams by the symbol χ, 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 $\chi(k) \to \chi(i\omega_k)$ and $\Psi(k) \to \Psi(i\omega_k)$ in Ref. [23], and here we generalized it to finite dimensions. We also derived the Schwinger EOM for the object χ. We derived diagrammatic rules for the three point vertices Λ and $\mathcal{U}$, defined as the functional derivatives of $\mathbf{g}^{-1}$ and $\widetilde{\mu}$ 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 $\widetilde{\mu}$. 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 λ. We also discussed the "deviation" of the λ 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 λ 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 re-summations 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.

Annals of Physics 343 (2014) 164–199

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: 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

HIGHLIGHTS

- 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.

ARTICLE INFO

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

ABSTRACT

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 rederived. 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 Δ_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, 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 Γ_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 $U \to \infty$ 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 $U \to \infty$ 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 λ 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 λ 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

$$X_i^{0\sigma} \to C_{i\sigma}, \qquad X_i^{\sigma 0} \to C_{i\sigma}^\dagger (1 - N_{i\bar{\sigma}}), \qquad X_i^{\sigma\sigma'} \to C_{i\sigma}^\dagger C_{i\sigma'}. \tag{1}$$

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 λ 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 Ψ 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

$$[\psi]'_{final} = Q'_M \ldots Q'_2 . Q'_1 . [\psi]'_{initial}, \tag{2}$$

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

$$[\psi] = \begin{bmatrix} \psi^{ph} \\ \psi^{un} \end{bmatrix}; \qquad \hat{P}_G = \begin{bmatrix} \mathbb{1}^{ph} & 0 \\ 0 & 0 \end{bmatrix}, \tag{3}$$

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:

$$Q_j = \begin{bmatrix} Q_j^{pp} & Q_j^{pu} \\ Q_j^{up} & Q_j^{uu} \end{bmatrix}. \tag{4}$$

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.

$$[\psi]_{final} = Q_M \ldots Q_2 . Q_1 . \hat{P}_G . [\psi]_{initial}, \tag{5}$$

and $[\psi]_{final} = \hat{P}_G . [\psi]_{final}$. 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

$$[\psi]_{final} = \begin{bmatrix} Q_M^{pp} \ldots Q_2^{pp} . Q_1^{pp} . \psi_{initial}^{ph} \\ 0 \end{bmatrix}. \tag{6}$$

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

$$Q_j = \begin{bmatrix} Q_j^{pp} & Q_j^{pu} \\ 0 & Q_j^{uu} \end{bmatrix}. \tag{7}$$

In more formal terms the sufficiency condition with least constraints that leads to Eq. (6) (via the block triangularity Eq. (7)) is

$$(\mathbb{1} - \hat{P}_G) . Q_j . \hat{P}_G = 0. \tag{8}$$

This condition is also expressible as $[Q_j, \hat{P}_G] . \hat{P}_G = 0$; 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 λ. 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 λ 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 λ 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 λ in the Fermion theory, as a parallel of the expansion parameter $\frac{1}{2s}$ 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

$$H_{tJ} = H_t + H_J,$$
$$H_t = -\sum_{ij} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \boldsymbol{\mu} \sum_i X_i^{\sigma\sigma},$$

$$H_J = \frac{1}{2}\sum_{ij} J_{ij}\left(\vec{S}_i.\vec{S}_j - \frac{1}{4}X_i^{\sigma\sigma}X_j^{\sigma'\sigma'}\right). \tag{9}$$

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

$$\hat{\mathcal{A}}_S = \sum_i \int_0^\beta \hat{\mathcal{A}}_S(i,\tau)d\tau,$$

$$\hat{\mathcal{A}}_S(i,\tau) = \left[X_i^{\sigma 0}(\tau)J_{i\sigma}(\tau) + J_{i\sigma}^*(\tau)X_i^{0\sigma}(\tau)\right] + \mathcal{V}_i^{\sigma'\sigma}(\tau)X_i^{\sigma'\sigma}(\tau), \tag{10}$$

and all time dependences are as in $Q(\tau) = e^{\tau H_{tJ}}Qe^{-\tau H_{tJ}}$. The generating functional of Green's functions of the t–J model is

$$Z[J, J^*, \mathcal{V}] \equiv \mathrm{Tr}_{tJ}\, e^{-\beta H_{tJ}}T_\tau\left(e^{-\hat{\mathcal{A}}_S}\right). \tag{11}$$

It reduces to the standard partition function on turning off the indicated source terms. The Green's functions for positive times $0 \le \tau_j \le \beta$, are defined through the Martin–Schwinger prescription [22,23]:

$$\mathcal{G}_{\sigma\sigma'}(i\tau_i, f\tau_f) = -\frac{\left\langle T_\tau\left(e^{-\hat{\mathcal{A}}_S}X_i^{0\sigma}(\tau_i)X_f^{\sigma' 0}(\tau_f)\right)\right\rangle}{\langle T_\tau e^{-\hat{\mathcal{A}}_S}\rangle}. \tag{12}$$

The functional Z conveniently yields the Green's functions upon taking functional derivatives with respect to the sources, e.g.

$$\mathcal{G}_{\sigma\sigma'}(i\tau_i, f\tau_f) = \left(\frac{1}{Z}\frac{\delta^2 Z}{\delta J_{i\sigma}^*(\tau_i)\delta J_{f\sigma'}(\tau_f)}\right), \tag{13}$$

where the sources are turned off at then end. We note that n_σ, the number of particles per site, is determined from the number sum rule:

$$n_\sigma = \mathcal{G}_{\sigma\sigma}(i\tau^-, i\tau), \tag{14}$$

and $\boldsymbol{\mu}$ 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 λ. 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:

$$\left(\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) - Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j)\right)$$
$$\times\, \mathcal{G}_{\sigma_j\sigma_f}(j\tau_j, f\tau_f) = \delta_{if}\delta(\tau_i - \tau_f)\left(\delta_{\sigma_i\sigma_f} - \gamma_{\sigma_i\sigma_f}(i\tau_i)\right), \tag{15}$$

where $\mathbf{g}_0$ is the noninteracting Green's function (Eq. (136)), $\hat{X}$ is a functional derivative operator (Eq. (130)), γ 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 γ (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 λ 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 $\left(\mathbf{g}_0^{-1} - U\delta/\delta\mathcal{V} - UG\right).G = \delta\,\mathbb{1}$. By scaling the interaction $U \to \lambda\, U$, with a parameter λ ($0 \le \lambda \le 1$), the interacting theory is connected continuously to the Fermi gas by tuning λ from 1 to 0. The standard perturbative expansion can be organized by counting the various powers of λ, setting $\lambda = 1$ at the end before evaluating the expressions [26]. Below in Section 5 we provide a more microscopic argument for introducing the λ 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 λ, we rewrite (15) schematically as:

$$\left(\mathbf{g}_0^{-1} - \lambda\hat{X} - \lambda Y_1\right).\mathcal{G} = \delta\,(\mathbb{1} - \lambda\gamma). \tag{16}$$

The strategy of the perturbative λ expansion method is to build up the solution of this equation at $\lambda = 1$ through a suitable expansion in λ, 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 λ 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: $t_{ij} \to t_{ij} + \text{const} \times \delta_{ij}$, 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 $\boldsymbol{\mu}$ (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 λ expansion plays an important "watchdog" role on the λ expansion, in addition to other standard constraints such as the Ward identities.

To summarize some key points of the λ expansion, we first decompose the Greens function into the space time convolution of an auxiliary Greens function and a caparison function as:

$$\mathcal{G} = \mathbf{g}.\mu. \tag{17}$$

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 $\left(\mathbf{g}_0^{-1} - \lambda\hat{X} - \lambda Y_1\right).\mathbf{g} = \delta\,\mathbb{1}$. The λ 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 λ 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 λ expansion satisfying shift invariance for practical calculations.
$\mathcal{G}$	$\partial_\tau \mathcal{G}$	$\mathcal{G}(k) = \mathbf{g}(k)\mu(k)$		$0 \le \lambda \le 1$	$\sum \mathcal{G} = \sum \mathbf{g} = \frac{n}{2}$	

($Y_1 = t\gamma$) in powers of λ. Dyson's skeleton graph idea is implemented by keeping the auxiliary $\mathbf{g}$ intact (i.e. unexpanded in λ), while all other variables are expanded in powers of λ *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 λ. 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 γ 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

$$\mu(\vec{k}, i\omega_n) = 1 - \frac{n}{2} + \Psi(\vec{k}, i\omega_n) \tag{18}$$

$$\mathbf{g}^{-1}(\vec{k}, i\omega_n) = \mathbf{g}_0^{(-1)}(\vec{k}, i\omega) - \Phi(\vec{k}, i\omega_n), \tag{19}$$

where the constant $\frac{n}{2}$ in Eq. (18) is fixed by the condition that Ψ vanishes at infinite frequency. The auxiliary Greens function satisfies a second sumrule analogous to Eq. (14), written in the Fourier domain:

$$(k_B T) \sum_{k,n} e^{i\omega_n 0^+} \mathbf{g}_{\sigma\sigma}(k, i\omega_n) = n_\sigma . \tag{20}$$

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]

$$\mathcal{G}(\vec{k}, i\omega) = \frac{1 - \frac{n}{2} + \Psi(\vec{k}, i\omega)}{\mathbf{g}_0^{(-1)}(\vec{k}, i\omega) - \Phi(\vec{k}, i\omega)}. \tag{21}$$

As argued in [4,6,8,9], simple Fermi liquid type self energies Ψ and Φ 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 $\delta \to 0$.

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 hold density $\delta \to 0$. In reality at very small δ several other broken symmetry states would compete and presumably win over the liquid state, so that this Fermi liquid state would be metastable. It characteristics are of interest and hence we proceed to describe these.

We study Eq. (21) by analytically continuing $i\omega \to \omega + i0^+$ and write

$$\mathbf{g}_0^{(-1)}(\vec{k}, i\omega) = \omega + \boldsymbol{\mu} - \left(1 - \frac{n}{2}\right)\varepsilon_k. \tag{22}$$

Let us define $\hat{k}$ as the *normal deviation* from the Fermi surface i.e. $\hat{k} = (\vec{k} - \vec{k}_F).\vec{k}_F/|\vec{k}_F|$, and the frequently occurring Fermi liquid function

$$\mathcal{R} = \pi\{\omega^2 + (\pi k_B T)^2\}. \tag{23}$$

We carry out a low frequency expansion as follows:

$$1 - \frac{n}{2} + \Psi(\vec{k}, \omega) = \alpha_0 + c_\Psi(\omega + v_\Psi\, \hat{k}\, v_f) + i\mathcal{R}/\gamma_\Psi + O(\omega^3), \tag{24}$$

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

$$\omega + \boldsymbol{\mu} - \left(1 - \frac{n}{2}\right)\varepsilon_k - \Phi(k, \omega) = (1 + c_\Phi)\left(\omega - v_\Phi\, \hat{k}\, v_f + i\mathcal{R}/\Omega_\Phi + O(\omega^3)\right), \tag{25}$$

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 Ψ is dimensionless and Φ is an energy. The dimensionless velocity renormalization constants v_Φ and v_Ψ 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

$$\mathcal{G}(\vec{k}, \omega) \sim \frac{z_0}{\alpha_0}\left(\frac{\alpha_0 + c_\Psi(\omega + v_\Psi\, \hat{k}\, v_f) + i\mathcal{R}/\gamma_\Psi}{\omega - v_\Phi\, \hat{k}\, v_f + i\mathcal{R}/\Omega_\Phi}\right) \tag{26}$$

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:

$$A(\vec{k}, \omega) = \frac{z_0}{\pi}\frac{\Gamma_0}{(\omega - v_\Phi\, \hat{k}\, v_f)^2 + \Gamma_0^2} \times \mu(k, \omega), \tag{27}$$

where the (Fermi liquid) width function

$$\Gamma_0(\hat{k}, \omega) = \eta + \frac{\pi(\omega^2 + (\pi k_B T)^2)}{\Omega_\Phi}, \tag{28}$$

with an extra phenomenological parameter η required to describe elastic scattering [14] in impure systems. The caparison function is

$$\mu(\hat{k}, \omega) = 1 - \frac{\omega}{\Delta_0} + \frac{v_0\, \hat{k}\, v_f}{\Delta_0}, \tag{29}$$

where we introduced an important (emergent) low energy scale combining the other parameters:

$$\Delta_0 = \alpha_0 \frac{\gamma_\Psi}{\Omega_\Phi - c_\Psi \gamma_\Psi} \tag{30}$$

and the dimensionless momentum dependence coefficient

$$\nu_0 = (\nu_\Psi \gamma_\Psi c_\Psi + \nu_\Phi \Omega_\Phi)/(\Omega_\Phi - c_\Psi \gamma_\Psi). \tag{31}$$

A cutoff $\theta\left(\mu(\hat{k},\omega)\right)$ 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, ν_0, ν_Φ, Ω_Φ, Δ_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 η 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 ν_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 $\delta \to 0$, 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:

$$G_{GBR}(\vec{k},\omega) \sim \frac{z}{\omega - z\,\hat{k}\,v_f}, \tag{32}$$

where z vanishes linearly with δ 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 ω.

As $n \to 1$ in Eq. (24) we expect that the constant $\Psi_0 \to -\frac{n}{2}$, 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

$$\begin{aligned} &z_0 \to \overline{z}_0 \times \delta; \qquad \Delta_0 \to \overline{\Delta}_0 \times \delta; \qquad \Omega_\Phi \to \overline{\Omega}_\Phi \times \delta; \\ &\nu_0 \to \overline{\nu}_0 \times \delta; \qquad \nu_\Phi \to \overline{\nu}_\Phi \times \delta; \end{aligned} \tag{33}$$

by using an overline for denoting a non vanishing limit of the stated variable [9,28]. The scaling of the velocity constants ν 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:

$$A(\hat{k}, s\,\omega | s\,T, s\,\delta) = A(\hat{k}, \omega | T, \delta), \tag{34}$$

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 ν's. The scaling holds for $\eta = 0$, and generalizes to a non zero values if we scale $\eta \to s\,\eta$. 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

$$A(\hat{k}, \omega | T, \delta) \sim A\left(\hat{k},\ \omega\frac{\delta_0}{\delta}\middle| T\frac{\delta_0}{\delta}, \delta_0\right), \tag{35}$$

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 ω 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 Γ_0 as a constant. We may then bring Eq. (27) to an interesting form studied in Ref. [5] by defining variables

$$\epsilon = \frac{\omega - v_\Phi\, \hat{k}\, v_f}{\Gamma_0}$$
$$\sinh u_k = \frac{\Delta_0 + (v_0 - v_\Phi)\, \hat{k}\, v_f}{\Gamma_0}, \tag{36}$$

so that the spectral function reduces to the standard form occurring in the ECFL theory:

$$A(u_k, \epsilon) = A_0 \frac{\sinh u_k - \epsilon}{1+\epsilon^2} \times \theta(\sinh u_k - \epsilon) \tag{37}$$

with $A_0 = \frac{z_0}{\Delta_0}$. This expression is valid for small enough ϵ [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

$$A_{Fano}(q_f, \epsilon) \propto \frac{(q_f + \epsilon)^2}{(1+\epsilon^2)}. \tag{38}$$

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

$$A(u_k, \epsilon) \propto \left(A_{Fano}(-e^{u_k}, \epsilon) - A_{Fano}(-e^{u_k}, \infty)\right). \tag{39}$$

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 $u_k \to \infty$, 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 ω, giving the momentum distribution curve (MDC) dispersion E_k. Let us introduce the convenient variables

$$r = \frac{v_0}{v_\Phi}, \tag{40}$$

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_Φ 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:

$$Q(\hat{k}) = \Delta_0 + (v_0 - v_\Phi)\, \hat{k}\, v_f. \tag{41}$$

In terms of these, the two dispersions are obtained as

$$E(k) \;= \frac{1}{2-r}\left(v_\Phi\, \hat{k}\, v_f + \Delta_0 - \sqrt{r(2-r)\, \Gamma_0^2 + Q^2}\right), \tag{42}$$

$$E^*(k) = \left(v_0\, \hat{k}\, v_f + \Delta_0 - \sqrt{\Gamma_0^2 + Q^2}\right). \tag{43}$$

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 ∞, 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 Γ_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 Δ_0, as well as the width Γ_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

$$(\hat{k}\, v_f)_{kink} = \frac{\Delta_0}{v_\Phi(1-r)}, \tag{44}$$

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 $E^*(k) \sim (v_0 + (v_0 - v_\Phi)\, sign(\hat{k}))\, \hat{k}\, v_f$ and $E(k) \sim \frac{1}{2-r}(v_\Phi + (v_0 - v_\Phi)\, sign(\hat{k}))\, \hat{k}\, v_f$. 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 Γ_0 decreases.

The value of the EDC energy at the kink is found by substituting $Q = 0$ and gives

$$E^*(k_{kink}) = -\frac{r}{r-1}\Delta_0 - \Gamma_0. \tag{45}$$

The MDC spectrum shows a kink for $2 \geq r \geq 1$ at the same momentum (44), with energy

$$E(k_{kink}) = -\frac{1}{r-1}\Delta_0 - \Gamma_0\sqrt{\frac{r}{2-r}}, \tag{46}$$

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 Γ_0, determining the kink width, is given in Eq. (28).

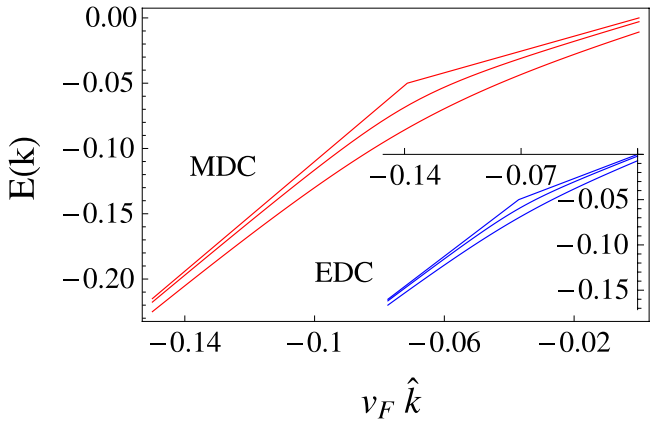

Fig. 1. A kink feature in the MDC dispersion relation $E(k)$ from Eq. (42) and in the inset from the EDC dispersion $E^*(k)$ (Eq. (43)) with parameters $\Delta_0 = .025$ eV, $\nu_0 = 1.05$, $\nu_\Phi = 0.7$ and three values of $\Gamma_0 = 0., .01, .02$ in eV from top to bottom. The kink is more pronounced in the MDC curve as discussed in text.

We observe in Fig. 1 that the kink becomes sharp when Γ_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 Γ_0, in case of laser ARPES, it is argued [14] that η 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 Γ_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 $\Delta_0 \to \delta\overline{\Delta}_0$ 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 ω 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 ω 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

$$\frac{1}{A(k,\omega)} \quad vs \quad (\omega - E_k^*)^2, \tag{47}$$

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

$$\widetilde{\omega}_k = \omega - E_k^*, \tag{48}$$

so that the spectral peak occurs at $\widetilde{\omega}_k = 0$. The inverse spectral function can be computed as a function of $\widetilde{\omega}_k$ and reads:

$$\frac{A(k, E_k^*)}{A(k, E_k^* + \widetilde{\omega}_k)} = 1 + \frac{e^{u_k}}{2\Gamma_0} \times \frac{\widetilde{\omega}_k^2}{\Gamma_0 \cosh(u_k) - \widetilde{\omega}_k}, \tag{49}$$

where the peak value of the spectral function at $\widetilde{\omega}_k = 0$ is:

$$A(k, E_k^*) = \frac{A_0}{2} e^{u_k}. \tag{50}$$

We next construct the object $\mathcal{Q}(\widetilde{\omega}_k)$ from Eq. (49) by subtracting unity and cross multiplying:

$$\mathcal{Q}(\widetilde{\omega}_k) = \frac{\widetilde{\omega}_k^2}{A(k, E_k^*)/A(k, E_k^* + \widetilde{\omega}_k) - 1}. \tag{51}$$

This variable is designed to be a $\widetilde{\omega}_k$ independent constant in a simple Fermi liquid with a Lorentzian line shape (i.e. Eq. (27) without the caparison function μ). Here $\mathcal{Q}$ has dimensions of the square of energy, and when plotted against $\widetilde{\omega}_k$ in the small range surrounding zero i.e. $|\widetilde{\omega}_k| \leq \Gamma_0$ it exhibits a linearly decreasing behavior with $\widetilde{\omega}_k$ within the ECFL spectral function (27)

$$\mathcal{Q}(\widetilde{\omega}_k) = \Gamma_0^2 (1 + e^{-2u_k}) - \left(2\Gamma_0 e^{-u_k}\right) \widetilde{\omega}_k. \tag{52}$$

Note that this function is flat for the usual Fermi liquid state without asymmetric corrections, since in this case $u_k \to +\infty$. If found in data, this linear in $\widetilde{\omega}$ behavior is the distinctive aspect of the asymmetric line shapes. We can then read off various physical quantities once the curve of $\mathcal{Q}(\widetilde{\omega}_k)$ versus $\widetilde{\omega}_k$ is obtained. For this purpose we need the intercept $\mathcal{Q}(0)$ and the slope near the origin $(d\mathcal{Q}(\widetilde{\omega}_k)/d\widetilde{\omega}_k)_0$. Clearly the $\mathcal{Q}(\widetilde{\omega}_k)$ function will deviate from a straight line sufficiently far from $\widetilde{\omega}_k = 0$, 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:

$$\hat{D} = \sum_i \hat{D}_i, \qquad \hat{D}_i \equiv C_{i\uparrow}^\dagger C_{i\uparrow} C_{i\downarrow}^\dagger C_{i\downarrow}, \tag{53}$$

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:

$$\hat{P}_G = \prod_i \left(1 - \hat{D}_i\right). \tag{54}$$

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]

$$\widetilde{C}_{i\sigma} = C_{i\sigma}(1 - N_{i\bar{\sigma}})$$
$$\widetilde{C}^{\dagger}_{i\sigma} = C^{\dagger}_{i\sigma}(1 - N_{i\bar{\sigma}}), \tag{55}$$

where $N_{i\sigma} = C^{\dagger}_{i\sigma} C_{i\sigma}$, and $N_i = \sum_{\sigma} N_{i\sigma}$, with the property that these conserve the number of doubly occupied sites locally:

$$[\widetilde{C}_{i\sigma}, \hat{D}_i] = 0, \qquad [\widetilde{C}^{\dagger}_{i\sigma}, \hat{D}_i] = 0 \tag{56}$$

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^{ab}_i operators as $X^{0\sigma}_i \leftrightarrow \widetilde{C}_{i\sigma}$ and $X^{\sigma 0}_i \leftrightarrow \widetilde{C}^{\dagger}_{i\sigma}$, 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

$$X^{\sigma 0}_i X^{0\sigma}_j \leftrightarrow C^{\dagger}_{i\sigma} C_{j\sigma} \left(1 - N_{i\bar{\sigma}} - N_{j\bar{\sigma}}\right). \tag{57}$$

While the choice

$$X^{\sigma 0}_i X^{0\sigma}_j \leftrightarrow \widetilde{C}^{\dagger}_{i\sigma} \widetilde{C}_{j\sigma} \tag{58}$$

is also a faithful representation, it contains an extra term $C^{\dagger}_{i\sigma} C_{j\sigma} N_{i\bar{\sigma}} N_{j\bar{\sigma}}$, over and above (57), which is redundant since (57) already commutes with (54).

Using (57) we write a canonical expression for the Hamiltonian

$$H_{tJ} \to \hat{H}_{eff} = \hat{H}_t + \hat{H}_J, \tag{59}$$

with

$$\hat{H}_t = \hat{T}_{eff} - \boldsymbol{\mu} \sum_i N_{i\sigma},$$
$$\hat{T}_{eff} = -\sum_{ij\sigma} t_{ij} C^{\dagger}_{i\sigma} C_{j\sigma} \left(1 - N_{i\bar{\sigma}} - N_{j\bar{\sigma}}\right), \tag{60}$$

we call this as the *symmetrized kinetic energy* in view of its obvious symmetry under the exchange $i \leftrightarrow j$, and write $\hat{H}_J \to \frac{1}{2}\sum_{ij} J_{ij} \left(\vec{S}_i.\vec{S}_j - \frac{1}{4} N_i N_j\right)$, 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

$$[\hat{H}_{eff}, \hat{D}] = 0 = [\hat{H}_{eff}, \hat{P}_G], \tag{61}$$

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}:

$$\hat{P}_G(\tau) = \hat{P}_G(0). \tag{62}$$

The partition functional as in (11), now defined with arbitrary τ_0:

$$Z = \mathrm{Tr}\, e^{-\beta \hat{H}_{eff}} T_{\tau} \left(e^{-\hat{\mathcal{A}}_S} \hat{P}_G(\tau_0)\right), \tag{63}$$

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 $\hat{\mathcal{A}}_S$, we use the replacement rules:

$$X^{0\sigma}_i \to \widetilde{C}_{i\sigma}, \qquad X^{\sigma 0}_i \to \widetilde{C}^{\dagger}_{i\sigma}, \qquad X^{\sigma\sigma'}_i \to C^{\dagger}_{i\sigma} C_{i\sigma'}, \tag{64}$$

to convert arbitrary expressions involving X_i^{ab} into those with the $\widetilde{C}$, $\widetilde{C}^\dagger$. 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

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\frac{\left\langle T_\tau \left(e^{-\hat{\mathcal{A}}_S} \widetilde{C}_{i\sigma_i}(\tau_i) \widetilde{C}^\dagger_{f\sigma_f}(\tau_f)\, \hat{P}_G(\tau_0) \right) \right\rangle}{\left\langle T_\tau \left(e^{-\hat{\mathcal{A}}_S} \hat{P}_G(\tau_0) \right) \right\rangle}, \tag{65}$$

evaluated [23] at $\hat{\mathcal{A}}_S \to 0$. 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 $\nabla.\vec{E}(r,t) = 0$ 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 $[H, \vec{\nabla}.\vec{E}] = 0$, 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 τ_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

$$\text{const} \times \sum \langle i|Q_1(\tau_1)\ldots Q_m(\tau_m)\hat{P}_G|i\rangle,$$

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 $\widetilde{C}_{i\sigma}$ 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:

$$[C_{i\sigma}, \hat{P}_G]\hat{P}_G = 0, \tag{66}$$

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 $\hat{f}(\hat{f} \equiv \hat{f}\{C_{i\sigma}\}, \{\widetilde{C}^\dagger_{j\sigma'}\})$ of the operators:

$$[\hat{f}, \hat{P}_G]\hat{P}_G = 0. \tag{67}$$

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

$$Z = \text{Tr}\, e^{-\beta\hat{H}_{eff}}\, T_\tau \left(e^{-\hat{\mathcal{A}}_S} \hat{P}_G(0^-) \right), \tag{68}$$

and introduce the important abbreviation for averages:

$$\langle\langle A(\tau_1)B(\tau_2)\ldots\rangle\rangle \equiv \frac{1}{Z}\text{Tr}\, e^{-\beta\hat{H}_{eff}}\, T_\tau \left(e^{-\hat{\mathcal{A}}_S} A(\tau_1)B(\tau_2)\ldots \hat{P}_G(0^-) \right), \tag{69}$$

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

$$\widetilde{C}_{i\sigma}(\tau) \to C_{i\sigma}(\tau), \tag{70}$$

leaving the result unchanged. Notice that this rule can also be applied to H_{eff}, and the source terms $\hat{\mathcal{A}}_S$ 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

$$X_i^{0\sigma} \to C_{i\sigma}, \qquad X_i^{\sigma 0} \to \widetilde{C}_{i\sigma}^{\dagger} = C_{i\sigma}^{\dagger}(1 - N_{i\bar{\sigma}}), \qquad X_i^{\sigma\sigma'} \to C_{i\sigma}^{\dagger} C_{i\sigma'}. \tag{71}$$

We thus rewrite the sources (10) as:

$$\hat{\mathcal{A}}_S(i,\tau) = \left(\widetilde{C}_{i\sigma}^{\dagger}(\tau) J_{i\sigma}(\tau) + J_{i\sigma}^{*}(\tau) C_{i\sigma}(\tau)\right) + \mathcal{V}_i^{\sigma'\sigma}(\tau) C_{i\sigma'}^{\dagger}(\tau) C_{i\sigma}(\tau), \tag{72}$$

and the Green's function with imaginary time $0 \le \tau_i, \tau_j \le \beta$ is therefore written as:

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle\langle C_{i\sigma_i}(\tau_i) \widetilde{C}_{f\sigma_f}^{\dagger}(\tau_f)\rangle\rangle, \tag{73}$$

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:

$$\hat{T}_{eff} = -\sum_{ij\sigma} t_{ij} \widetilde{C}_{i\sigma}^{\dagger} C_{j\sigma}. \tag{74}$$

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 τ_j are positive and satisfy $0 \le \tau_j \le \beta$. The Green's function (12) satisfies the Kubo–Martin–Schwinger (KMS) anti-periodic boundary conditions [39]

$$\mathcal{G}(a\,\tau_i = 0, b\,\tau_f) = -\mathcal{G}(a\,\tau_i = \beta, b\,\tau_f), \tag{75}$$

$$\mathcal{G}(a\,\tau_i, b\,\tau_f = 0) = -\mathcal{G}(a\,\tau_i, b\,\tau_f = \beta), \tag{76}$$

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:	$X^{\sigma 0}$ $X^{0\sigma}$ $X_i^{\sigma\sigma'}$	$\widetilde{C}_\sigma^\dagger = C_\sigma^\dagger\,(1-N_{\bar\sigma})$ $\widetilde{C}_\sigma = C_\sigma\,(1-N_{\bar\sigma})$ $C_{i\sigma}^\dagger C_{i\sigma'}$	$\widetilde{C}_\sigma^\dagger = C_\sigma^\dagger\,(1-N_{\bar\sigma})$ C_σ $C_{i\sigma}^\dagger C_{i\sigma'}$
Partition Functional: Z	$\mathrm{Tr}_{tJ} e^{-\beta H_{tJ}} T_\tau \left(e^{-\hat{\mathcal{A}}_S}\right)$	$\mathrm{Tr}\, e^{-\beta \hat{H}_{eff}} T_\tau \left(e^{-\hat{\mathcal{A}}_S} \hat{P}_G(\tau_0)\right)$; Arbitrary time τ_0 $(0 \le \tau_0 \le \beta)$.	$\mathrm{Tr}\, e^{-\beta \hat{H}_{eff}} T_\tau \left(e^{-\hat{\mathcal{A}}_S} \hat{P}_G(0^-)\right)$
Green's function: $-\mathcal{G}(1,1')$	$\langle T_\tau (e^{-\hat{\mathcal{A}}_S} X_1^{0\sigma} X_{1'}^{\sigma' 0})\rangle$	$\left\langle T_\tau \left(e^{-\hat{\mathcal{A}}_S} \widetilde{C}_\sigma(1) \widetilde{C}_{\sigma'}^\dagger(2) \hat{P}_G(\tau_0)\right)\right\rangle$ Arbitrary time τ_0 $(0 \le \tau_0 \le \beta)$.	$\left\langle T_\tau \left(e^{-\hat{\mathcal{A}}_S} C_\sigma(1) \widetilde{C}_{\sigma'}^\dagger(2) \hat{P}_G(0^-)\right)\right\rangle$
Remarks:	$H = H^\dagger$ in the defining representation.	Symmetrized theory $H = H^\dagger$	(i) Symmetrized theory: $\hat{H}_{eff} = H_{eff}^\dagger$ (ii) Minimal theory: $\hat{H}_{eff} \neq \hat{H}_{eff}^\dagger$

where the fixed time τ_f (τ_i) in the first (second) equations is assumed to satisfy $0 \le \tau \le \beta$. These conditions are usually proven by using the cyclic invariance of the trace [24], and translates easily to the canonical representation (65), with $\widetilde{C}$ and $\widetilde{C}^\dagger$ replacing the X operators (64).

In using the non-Hermitian representation (71) as in (73), we cannot use cyclicity of trace since the operator $\widetilde{C}$ 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. $\mathcal{A} \to 0$, the more general case follows by a similar argument. From the definitions of the Green's functions, Eq. (76) is true since $\mathrm{Tr}\left(e^{-\beta H_{eff}} C_{a\sigma}(\tau_i)[\widetilde{C}_{b\sigma'}^\dagger(0), \hat{P}_G]\right)$ vanishes identically from Eq. (56).

In order to prove that Eq. (75) remains true, we need to show that the expression

$$\mathrm{Tr}\left(e^{-\beta H_{eff}} \widetilde{C}_{b\sigma'}^\dagger(\tau_f)[C_{a\sigma}(0), \hat{P}_G]\right) \tag{77}$$

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 $\hat{Q}$): $\mathrm{Tr}\left((\mathbb{1} - \hat{P}_G)\hat{Q}\hat{P}_G\right) = 0$, to write the required expression (77) in the form

$$\mathrm{Tr}\left((\hat{P}_G - \mathbb{1}) e^{-\beta H_{eff}} \widetilde{C}_{b\sigma'}^\dagger(\tau_f) \hat{P}_G C_{a\sigma}(0)\right). \tag{78}$$

Using $(\hat{P}_G)^2 = \hat{P}_G$, we rewrite this as:

$$(\hat{P}_G - \mathbb{1}) e^{-\beta H_{eff}} \widetilde{C}_{b\sigma'}^\dagger(\tau_f) \hat{P}_G = [\hat{P}_G, e^{-\beta H_{eff}} \widetilde{C}_{b\sigma'}^\dagger(\tau_f)] \hat{P}_G.$$

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) $\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle\langle C_{i\sigma_i}(\tau_i)\widetilde{C}^\dagger_{f\sigma_f}(\tau_f)\rangle\rangle$, with the averages from Eq. (69). Expanding the $\widetilde{C}^\dagger$ operator this becomes

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle\langle C_{i\sigma_i}(\tau_i)C^\dagger_{f\sigma_f}(\tau_f)\rangle\rangle + \langle\langle C_{i\sigma_i}(\tau_i)C^\dagger_{f\sigma_f}(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle. \tag{79}$$

We next define the auxiliary Green's function as:

$$\mathbf{g}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) = -\langle\langle C_{i\sigma_i}(\tau_i)C^\dagger_{j\sigma}(\tau_j)\rangle\rangle, \tag{80}$$

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

$$\langle\langle C_{i\sigma_i}(\tau_i)C^\dagger_{f\sigma_f}(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle = -\mathbf{g}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f)\langle N_{f\bar{\sigma}_f}(\tau_f)\rangle + \langle\langle C_{i\sigma_i}(\tau_i)C^\dagger_{f\sigma_f}(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle_c. \tag{81}$$

The connected part is written in terms of a second self energy Ψ defined as

$$\Psi_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = \mathbf{g}^{-1}_{\sigma_i\sigma_{\mathbf{k}}}(i\tau_i, \mathbf{k}\tau_{\mathbf{k}}) \times \langle\langle C_{\mathbf{k}\sigma_{\mathbf{k}}}(\tau_{\mathbf{k}})C^\dagger_{f\sigma_f}(\tau_f)N_{f\bar{\sigma}_f}(\tau_f)\rangle\rangle_c, \tag{82}$$

and assembling these we rewrite (79) as the product relation [4]

$$\begin{aligned}\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) &= \mathbf{g}_{\sigma_i\sigma_{\mathbf{k}}}(i\tau_i, \mathbf{k}\tau_{\mathbf{k}})\mu_{\sigma_{\mathbf{k}}\sigma_f}(\mathbf{k}\tau_{\mathbf{k}}, f\tau_f),\\ \mu_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) &= \delta(if)\left(1 - \langle N_{\bar{\sigma}_i}(\tau_i)\rangle\right) + \Psi_{\sigma_i\sigma_f}(i\tau_i, f\tau_f).\end{aligned} \tag{83}$$

There is a slight ambiguity in defining the two objects $\mathbf{g}$ and μ, 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 Ψ 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 λ expansion, where the two self energies are argued to be generically Fermi liquid-like and similar to each other. An energy scale (Δ) emerges from a ratio of their imaginary parts, and controls the significant asymmetry seen in the spectral functions.

5. The λ-Fermions

A natural question is whether Eq. (16), explicitly containing the parameter λ, can arise in a microscopic theory where λ enters in a fundamental way, as opposed to the "engineering approach" in Section 2.3. A set of λ-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:

$$\begin{aligned}X^{\sigma 0}_i(\lambda) &\to C^\dagger_{i\sigma}(1 - \lambda C^\dagger_{i\bar{\sigma}}C_{i\bar{\sigma}})\\ X^{0\sigma}_i(\lambda) &\to C_{i\sigma}\\ X^{\sigma\sigma'}_i(\lambda) &\to C^\dagger_{i\sigma}C_{i\sigma'}.\end{aligned} \tag{84}$$

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. $\left(X^{\sigma 0}_i(\lambda)\right)^\dagger \neq X^{0\sigma}_i(\lambda)$, unless $\lambda = 0$.

These operators satisfy a λ 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:

$$\{X^{0\sigma_i}_i(\lambda), X^{\sigma_j 0}_j(\lambda)\} = \delta_{ij}\{\delta_{\sigma_i\sigma_j} - \lambda\,\sigma_i\sigma_j X^{\bar{\sigma}_i\bar{\sigma}_j}_i(\lambda)\}. \tag{85}$$

We next evaluate the fundamental Bosonic bracket

$$[X_i^{0\sigma_i}(\lambda), X_j^{\sigma_j\sigma_k}(\lambda)] = \delta_{ij}\delta_{\sigma_i\sigma_j}X_i^{0\sigma_k}(\lambda) \tag{86}$$

$$[X_i^{\sigma_i 0}(\lambda), X_j^{\sigma_j\sigma_k}(\lambda)] = -\delta_{ij}\delta_{\sigma_i\sigma_k}X_i^{\sigma_j 0}(\lambda). \tag{87}$$

Here (87) requires a brief calculation [40] invoking the Pauli principle vanishing of $C_\sigma^\dagger C_\sigma^\dagger \to 0$. On the other hand (86) is elementary, due to the absence of λ 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 λ reproduce the "half bracket", or product relations expected for *projection operators*. We find that

$$X_i^{\sigma 0}(\lambda)X_i^{0\sigma'}(\lambda) \neq X_i^{\sigma\sigma'}(\lambda), \tag{88}$$

$$X_i^{0\sigma'}(\lambda)X_i^{\sigma 0}(\lambda) \neq X_i^{00}\delta_{\sigma\sigma'}. \tag{89}$$

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 $\boldsymbol{\mu}$. 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 λ 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 λ-versions, i.e. replacing $X_i^{ab} \to X_i^{ab}(\lambda)$, and thereby obtain equations that yield (16). This procedure then provides a (continuous) interpolation between the free Fermi and extremely correlated regimes by varying λ from 0 to 1. Let us first demonstrate this by a brief calculation.

5.1. The λ-Fermion theory equations of motion

Using the λ Fermions, we define the Green's function as

$$\mathcal{G}^{(\lambda)}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = -\langle T_\tau X_i^{0\sigma_i}(\tau_i, \lambda)X_f^{\sigma_f 0}(\tau_f, \lambda)\rangle_{(\lambda)} \tag{90}$$

where with arbitrary $\hat{A}$

$$\langle\hat{A}\rangle_\lambda \equiv -\frac{\mathrm{Tr}\, e^{-\beta H_{eff}(\lambda)}T_\tau\left(e^{-\hat{\mathcal{A}}_S(\lambda)}\hat{A}\right)}{Z(\lambda)},$$
$$Z(\lambda) = \mathrm{Tr}\, e^{-\beta H_{eff}(\lambda)}T_\tau\left(e^{-\hat{\mathcal{A}}_S(\lambda)}\right). \tag{91}$$

In this expression $H_{eff}(\lambda)$ is given by Eq. (92) and $\hat{\mathcal{A}}_S(\lambda)$ is obtained from (10), with the replacement $X_i^{ab} \to X_i^{ab}(\lambda)$:

$$H_{eff}(\lambda) = -\sum_{ij} t_{ij}X_i^{\sigma 0}(\lambda)X_j^{0\sigma}(\lambda) - \boldsymbol{\mu}\sum_i N_{i\sigma} + \lambda\frac{1}{2}\sum_{ij}J_{ij}\left(\vec{S}_i.\vec{S}_j - \frac{1}{4}N_{i\sigma}N_{j\sigma'}\right)$$
$$+ u_0\,\lambda\sum_i N_{i\uparrow}N_{i\downarrow}, \tag{92}$$

where u_0 is now the "second chemical potential". The scaling of the J term with λ 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 λ 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

$$\begin{aligned}\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j)\mathcal{G}^{(\lambda)}_{\sigma_j\sigma_f}(j\tau_j, f\tau_f) \;=\; & \delta(\tau_i-\tau_f)\delta_{ij}(1-\lambda\,\gamma_{\sigma_i\sigma_f}(i\tau_i)) \\ & -\lambda\sum_{j\sigma_j} t_{ij}(\sigma_i\sigma_j)\,\left\langle T_\tau\left(X_i^{\bar\sigma_i\bar\sigma_j}(\tau_i)X_j^{0\sigma_j}(\tau_i)\,X_f^{\sigma_f 0}(\tau_f)\right)\right\rangle_{(\lambda)} \\ & +\frac{1}{2}\sum_{j\sigma_j} J_{ij}(\sigma_i\sigma_j)\left\langle T_\tau\left(X_j^{\bar\sigma_i\bar\sigma_j}(\tau_i)X_i^{0\sigma_j}(\tau_i)X_f^{\sigma_f 0}(\tau_f)\right)\right\rangle_{(\lambda)} \\ & -\frac{1}{2}\lambda u_0\sum_{\sigma_j}(\sigma_i\sigma_j)\left\langle T_\tau\left(X_i^{\bar\sigma_i\bar\sigma_j}(\tau_i)X_i^{0\sigma_j}(\tau_i)X_f^{\sigma_f 0}(\tau_f)\right)\right\rangle_{(\lambda)}, \end{aligned} \tag{93}$$

where the λ 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 λ through the EOM since it conserves double occupancy. We can choose to additionally scale it with λ as $J_{ij} \to \lambda J_{ij}$, 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 $J_{ij} \to J_{ij} - u_0\delta_{ij}$, with this the resulting equation is

$$\begin{aligned}&\left(\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j) - \lambda\,\hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) - \lambda\,Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j)\right)\times\mathcal{G}^{(\lambda)}_{\sigma_j\sigma_f}(j\tau_j, f\tau_f) \\ &\quad = \delta_{if}\delta(\tau_i-\tau_f)\left(\delta_{\sigma_i\sigma_f} - \lambda\,\gamma_{\sigma_i\sigma_f}(i\tau_i)\right).\end{aligned} \tag{94}$$

The constitutive relation determining the chemical potential is taken as

$$\begin{aligned} n_{i\sigma} \;&=\; \langle X_i^{\sigma 0}(\tau,\lambda)X_i^{0\sigma}(\tau^-,\lambda)\rangle_{(\lambda)}, \\ &=\; \mathcal{G}^{(\lambda)}_{\sigma\sigma}(i,\;\tau^-,\tau), \end{aligned} \tag{95}$$

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 λ it is guided by the requirement of continuity in λ. In the same spirit, we express the function γ in Eq. (16) as

$$\gamma_{\sigma\sigma'}(i\tau) = \sigma\sigma'\mathcal{G}_{\bar\sigma'\bar\sigma}(i\tau^-, i\tau), \tag{96}$$

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 λ and u_0, where the skeleton expansion is avoided. Let us consider the atomic limit of the λ-Fermions theory. We consider the Hamiltonian $H_0 = -\boldsymbol{\mu}\sum_\sigma N_\sigma + \lambda u_0 N_\uparrow N_\downarrow$ with $u_0 \geq 0$. The Green's function in Eq. (90) can be calculated easily using the EOM technique as:

$$\mathcal{G}(i\omega_n) = \frac{1-n_{\bar\sigma}}{i\omega_n + \boldsymbol{\mu}} + \frac{(1-\lambda)n_{\bar\sigma}}{i\omega_n + \boldsymbol{\mu} - \lambda u_0}. \tag{97}$$

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 $\lambda u_0 - \boldsymbol{\mu}$ 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 λ 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 $i\omega_n \to \infty$, the local Green's function falls off as $\mathcal{G}(i\omega_n) \to a_G/i\omega_n$. 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 λ expansion, from Eq. (16) we have $a_G = 1 - \lambda\gamma$, where γ is further expanded in λ. 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 λ 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 λ 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 $\vec{S}_i.\vec{S}_i = s(s+1)$ and b_i, $b_i^\dagger$ are canonical Bose operators. The Boson vacuum state $b_i|vac\rangle = 0$ is mapped as $|vac\rangle \leftrightarrow |\downarrow, \downarrow, \downarrow \ldots \downarrow\rangle$, so that the action of $b_i^\dagger$ creates spin reversals. Their number is cut off such that $n_i \le (2s)$, 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 λ 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	$\vert \downarrow\downarrow \ldots \downarrow\rangle$	$\vert 00 \ldots 0\rangle$	$\vert Vac\rangle$	$\vert 00 \ldots 0\rangle$
Small parameter & its range	$\frac{1}{2s}$	$\frac{1}{2s} \in [0, 1]$	λ	$\lambda \in [0, 1]$
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 Ψ		$\Psi(i,j) = \mathbf{g}^{-1}(i,\mathbf{a})\langle\langle b_{\mathbf{a}} b_j^\dagger n_j\rangle\rangle_c$		$\Psi(i,j) = \mathbf{g}^{-1}(i,\mathbf{a})\langle\langle C_{\mathbf{a}\sigma} C_{j\sigma}^\dagger N_{j\bar{\sigma}}\rangle\rangle_c$

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 $\langle\langle Q \rangle\rangle = \mathrm{Tr}(e^{-\beta H}$

$T_\tau Q P_D)/\mathrm{Tr} e^{-\beta H} P_D$, we write from (98)

$$\frac{1}{2s}\mathcal{G}(i,j) = -\left\langle\!\!\left\langle b_i b_j^\dagger \left(1 - \frac{1}{2s} n_j\right)\right\rangle\!\!\right\rangle. \tag{99}$$

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

$$\frac{1}{2s}\mathcal{G}(i,j) = \mathbf{g}(i,j)\left(1 - \frac{1}{2s}\langle n_j\rangle\right) + \frac{1}{2s}\mathbf{g}(i,\mathbf{k})\,\Psi(\mathbf{k},j), \tag{100}$$

$$\Psi(i,j) = \mathbf{g}^{-1}(i,\mathbf{a})\langle\!\langle b_{\mathbf{a}} b_j^\dagger n_j \rangle\!\rangle_c. \tag{101}$$

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 μ, in the form $\frac{1}{2s}\mathcal{G}(i,j) = \mathbf{g}(i,\mathbf{k})\mu(\mathbf{k},j)$, where $\mu(i,j) = \delta_{ij}(1 - \frac{1}{2s}\langle n_j\rangle) + \frac{1}{2s}\Psi(i,j)$. The auxiliary Green's function is defined in terms of its own self energy Φ 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 λ 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 $\lambda u_o \sum_i N_{i\uparrow} N_{i\downarrow}$, 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}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 $C_{i\sigma} \to c_{i\sigma}$, $\widetilde{C}_{i\sigma} \to \widetilde{c}_{i\sigma} \equiv c_{i\sigma}(1 - c^*_{i\bar{\sigma}} c_{i\bar{\sigma}})$, $C^\dagger_{i\sigma} \to c^*_{i\sigma}$, $\widetilde{C}^\dagger_{i\sigma} \to \widetilde{c}^*_{i\sigma} \equiv c^*_{i\sigma}(1 - c^*_{i\bar{\sigma}} c_{i\bar{\sigma}})$. 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 $M \to \infty$, 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]:

$$Z^{(M)}[J^*, J, \mathcal{V}] = \int_c P_G(\tau_1, \tau_0)\, e^{-\mathcal{A}^{(M)}_{Tot}},$$
$$\mathcal{A}_{Tot} = \mathcal{A}^{(M)}_0 + \mathcal{A}^{(M)}_S + \mathcal{A}^{(M)}_t + \mathcal{A}^{(M)}_J. \tag{102}$$

We detail the various contributions next; the free Fermi term is given by

$$\frac{1}{\Delta\tau}\mathcal{A}^{(M)}_0 = \sum_{j=0}^{M-1} \left[c^*_{i\sigma}(\tau_{j+1})\delta_{\tau_j} c_{i\sigma}(\tau_j) - \boldsymbol{\mu}\; n_{j\sigma}(\tau_j)\right], \tag{103}$$

with the finite difference operator δ_{τ_j} defined through

$$\delta_{\tau_j} F(\tau_j) \equiv \frac{1}{\Delta\tau}\left\{F(\tau_{j+1}) - F(\tau_j)\right\}. \tag{104}$$

As $M \to \infty$, we note that δ_{τ_j} reduces to the derivative operator ∂_τ, and we obtain the integral $\mathcal{A}_0 = \int_0^\beta d\tau\; c^*_{i\sigma}(\tau)(\partial_\tau - \boldsymbol{\mu}) c_{i\sigma}(\tau)$, and in that limit $Z^{(M)} \to Z[J^*, J, \mathcal{V}]$. The source term $\mathcal{A}^{(M)}_S(\tau_{j+1}, \tau_j)$ obtained from (72) is given by

$$\mathcal{A}^{(M)}_S = \sum_i \left[\widetilde{c}^*_{i\sigma}(\tau_{j+1}) J_{i\sigma}(\tau_{j+1}) + J^*_{i\sigma}(\tau_{j+1}) c_{i\sigma}(\tau_j)\right] + \left[\mathcal{V}_i^{\sigma'\sigma}(\tau_{j+1}) c^*_{i\sigma'}(\tau_{j+1}) c_{i\sigma}(\tau_j)\right]. \tag{105}$$

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

$$\mathcal{A}^{(M)}_t = \Delta\tau \sum_j T_{eff}(\tau_j) \to \int_0^\beta d\tau\; T_{eff}(\tau), \tag{106}$$

with T_{eff} from Eq. (60) or Eq. (74):

$$T_{eff}^{Sym}(\tau_j) = -\sum_{lm\sigma} t_{lm}\, c^*_{l\sigma}(\tau_{j+1}) c_{m\sigma}(\tau_j) \times \left(1 - n_{l\bar{\sigma}}(\tau_j) - n_{m\bar{\sigma}}(\tau_j)\right),$$

$$T_{eff}^{Min}(\tau_j) = -\sum_{lm\sigma} t_{lm}\, c^*_{l\sigma}(\tau_{j+1}) c_{m\sigma}(\tau_j) \times \left(1 - n_{l\bar{\sigma}}(\tau_j)\right), \tag{107}$$

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

$$\mathcal{A}_J^{(M)} = \Delta\tau \sum_j H_J(\tau_j) \to \int_0^\beta d\tau\, H_J(\tau),$$

$$H_J(\tau_j) \equiv -\frac{1}{4}\sum_{lm} J_{lm}\, \sigma_1\sigma_2 \times c^*_{l\sigma_1}(\tau_{j+1}) c^*_{m\bar{\sigma}_1}(\tau_{j+1}) c_{m\bar{\sigma}_2}(\tau_j) c_{l\sigma_2}(\tau_j). \tag{108}$$

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 $\mathcal{G}^{(M)}_{\sigma\sigma'}(a\tau_i, b\tau_f) \to \mathcal{G}_{\sigma\sigma'}(a\tau_i, b\tau_f)$ and $Z^{(M)} \to Z$ 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 $\mathcal{H}_{eff}(\tau_{j+1}, \tau_j)$ to $\mathcal{H}_{eff}(\tau_j)$ or even more simply to $\mathcal{H}_{eff}$. 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

$$P_G(\tau_1, \tau_0) \equiv \prod_{i=1}^{N_s} \left(1 - c^*_{i\uparrow}(\tau_1) c_{i\uparrow}(\tau_0) c^*_{i\downarrow}(\tau_1) c_{i\downarrow}(\tau_0)\right), \tag{109}$$

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:

$$\langle\langle Q \rangle\rangle = \frac{\|Q\|}{Z}, \quad \text{with } \|Q\| = \int_c P_G(\tau_1, \tau_0) e^{-\mathcal{A}}\, Q, \tag{110}$$

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 $\delta/\delta J(\tau_j) \to \frac{1}{(\Delta\tau)} d/dJ(\tau_j)$ [50]:

$$\mathcal{G}_{\sigma_i\sigma_f}(i\tau_i, f\tau_f) = \frac{1}{Z} \|\widetilde{c}^*_{f\sigma_f}(\tau_f)\, c_{i\sigma}(\tau_i)\|. \tag{111}$$

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:

$$\int_c P_G(\tau_1, \tau_0) \frac{\delta}{\delta c^*_{i\sigma_i}(\tau_{i+1})} \left[\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})\, e^{-\mathcal{A}_{Tot}}\right] = 0. \tag{112}$$

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 $\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})$ 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

$$\left\| \frac{\delta U}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_S}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_0}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_t}{\delta V} \right\| + \left\| U \frac{\delta \mathcal{A}_J}{\delta V} \right\| = 0. \tag{113}$$

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 $U = \widetilde{c}^*_{f\sigma_f}(\tau_f)$ and $V = c^*_{i\sigma_i}(\tau_i)$. We compute the various pieces of (113) next.

Denoting

$$\widehat{\gamma}_{\sigma_i\sigma_f}(i) \equiv \sigma_i\sigma_f c^*_{i\bar{\sigma}_i}(\tau_{i+1}) c_{\,i\bar{\sigma}_f}(\tau_i), \tag{114}$$

and using the convention that repeated spin indices are summed over, we obtain the first result:

$$\frac{\delta}{\delta c^*_{i\sigma_i}(\tau_{i+1})} \widetilde{c}^*_{f\sigma_f}(\tau_{f+1}) = \delta_{\tau_i\tau_f}\delta_{if} \left\{ \delta_{\sigma_i\sigma_f} - \widehat{\gamma}_{\sigma_i\sigma_f}(i) \right\}. \tag{115}$$

We obtain

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_0}{\delta c^*_{i\sigma_i}(\tau_{i+1})} = \delta_{\tau_i} c_{i\sigma_i}(\tau_i) - \boldsymbol{\mu}\, c_{i\sigma_i}(\tau_i), \tag{116}$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_S}{\delta c^*_{i\sigma_i}(\tau_{i+1})} = \mathcal{V}_i^{\sigma_i\sigma_j}(\tau_{i+1})\, c_{i\sigma_j}(\tau_i) + \left\{ \delta_{\sigma_i\sigma_j} - \widehat{\gamma}_{\sigma_i\sigma_j}(i) \right\} J_{i\sigma_j}(\tau_{i+1}), \tag{117}$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_t^{Sym}}{\delta c^*_{i\sigma_i}(\tau_{i+1})} = -t_{ij} c_{j\sigma_i}(\tau_i) + t_{ij} \left[\hat{\gamma}_{\sigma_i\sigma_j}(i\tau_i)\, c_{j\sigma_j}(\tau_i) + c^*_{j\bar{\sigma}_i} c_{j\bar{\sigma}_i} c_{j\sigma_i} + c^*_{j\bar{\sigma}_i} c_{i\bar{\sigma}_i} c_{i\sigma_i} \right],$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_t^{Min}}{\delta c^*_{i\sigma_i}(\tau_{i+1})} = -t_{ij} c_{j\sigma_i}(\tau_i) + t_{ij} \hat{\gamma}_{\sigma_i\sigma_j}(i\tau_i)\, c_{j\sigma_j}(\tau_i)$$

$$\frac{1}{\Delta\tau} \frac{\delta \mathcal{A}_J}{\delta c^*_{i\sigma_i}(\tau_{i+1})} = -\frac{1}{2} J_{ij}\sigma_i\sigma_j\, c^*_{j\bar{\sigma}_i}(\tau_{i+1}) c_{j\bar{\sigma}_j}(\tau_i) c_{i\sigma_j}(\tau_i), \tag{118}$$

We combine the two terms as:

$$\frac{1}{\Delta\tau} \frac{\delta(\mathcal{A}_t + \mathcal{A}_J)}{\delta c^*_{i\sigma_i}(\tau_{i+1})} = -\sum_j t_{ij} c_{j\sigma_i}(\tau_i) + A_{i\sigma_i}(\tau_{i+1}, \tau_i), \tag{119}$$

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)

$$A^{Min}_{i\sigma_i} = t_{ij} \widehat{\gamma}_{\sigma_i\sigma_j}(i\tau_i)\, c_{j\sigma_j}(\tau_i) - \frac{1}{2} J_{ij} \widehat{\gamma}_{\sigma_i\sigma_j}(j\tau_i) c_{i\sigma_j}(\tau_i), \tag{120}$$

in agreement with Eq. (22) of Ref. [6], and the symmetrized case is obtained in a similar way. Combining these (with $J \to 0$) we get the EOM in discrete time space:

$$\left[\left\{\boldsymbol{\mu} - \delta_{\tau_i} - \mathcal{V}_i^{\sigma_i\sigma_j}(\tau_{i+1})\right\}\delta_{i,j} + t_{ij}\right] \|\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})c_{j\sigma_i}(\tau_i)\| - \|\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})A_{i\sigma_i}(\tau_{i+1},\tau_i)\|$$
$$= \delta_{if}\|(\delta_{\sigma_i\sigma_f} - \widehat{\gamma}_{\sigma_i\sigma_f}(i))\| \frac{\delta_{\tau_i,\tau_f}}{\Delta\tau}. \tag{121}$$

We next take the continuum limit in time τ_i; with $\Delta\tau \to 0$, and using $\frac{\delta_{\tau_i,\tau_f}}{\Delta\tau} \to \delta(\tau_i - \tau_f)$, 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:

$$\|\widetilde{c}^*_{f\sigma_f}(\tau_f)A_{i\sigma_i}(\tau_i)\| = \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j)\|\widetilde{c}^*_{f\sigma_f}(\tau_f)c_{j\sigma_j}(\tau_j)\| \tag{122}$$

where $\hat{X}$ is a functional derivative operator defined more completely below in (130). With this preparation we can rewrite Eq. (121) as

$$\left(\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j)\right) \|\widetilde{c}^*_{f\sigma_f}(\tau_f)c_{j\sigma_j}(\tau_j)\| = \delta_{if}\|(\delta_{\sigma_i\sigma_f} - \widehat{\gamma}_{\sigma_i\sigma_f}(i))\|\delta(\tau_i - \tau_f). \tag{123}$$

Here and elsewhere since τ_j repeats in product, it is assumed to be integrated between $0 \le \tau_j \le \beta$, 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

$$Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j) = \frac{1}{Z}(\hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j)Z), \tag{124}$$

to rewrite (123) in the same form as Eq. (138)

$$\left(\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) - Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j)\right)$$
$$\times \mathcal{G}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) = \delta_{if}\delta(\tau_i - \tau_f)\left[\delta_{\sigma_i\sigma_f} - \gamma_{\sigma_i\sigma_f}(i)\right], \tag{125}$$

where

$$\gamma_{\sigma_i\sigma_f}(i) \equiv \langle\langle\widehat{\gamma}_{\sigma_i\sigma_f}(i)\rangle\rangle. \tag{126}$$

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:

$$\left(\mathbf{g}^{-1}_{0,\tilde{i}\tilde{j}} - \hat{X}_{\tilde{i},\tilde{j}} - Y_{1\tilde{i},\tilde{j}}\right) \mathcal{G}_{\tilde{j},\tilde{f}} = (\delta_{\tilde{i}\tilde{f}} - \gamma_{\tilde{i}\tilde{f}}). \tag{127}$$

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 $D_{\sigma_i\sigma_j}(i) \equiv \sigma_i\sigma_j\delta/\delta\mathcal{V}_i^{\bar{\sigma}_i\bar{\sigma}_j}$ (from Eq. (39) of Ref. [6])

$$D_{\sigma_i\sigma_j}(i)Z[\mathcal{V}] = -\|\hat{\gamma}_{\sigma_i\sigma_j}(i)\|, \tag{128}$$

so that:

$$\|\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})A^{Min}_{i\sigma_i}\| = -t_{ij}D_{\sigma_i\sigma_j}(i)\|\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})c_{j\sigma_j}(\tau_i)\| + \frac{1}{2}J_{ij}D_{\sigma_i\sigma_j}(j)\|\widetilde{c}^*_{f\sigma_f}(\tau_{f+1})c_{i\sigma_j}(\tau_i)\|, \tag{129}$$

and comparing with Eq. (122) we conclude

$$\hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) = \delta(\tau_i - \tau_j)(-t_{ij}D_{\sigma_i\sigma_j}(i) + \delta_{ij}\sum_k \frac{1}{2}J_{ik}D_{\sigma_i\sigma_j}(k\tau_i)), \tag{130}$$

where the derivative $D_{\sigma_i\sigma_j}(k\tau_i)$ is at spatial site k and time τ_i. The corresponding Y_1 (with a similar convention as above) in Eq. (124) is

$$Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j) = -\delta(\tau_i - \tau_j)\left(-t_{ij}\gamma_{\sigma_i\sigma_j}(i) + \delta_{ij}\sum_k \frac{1}{2}J_{ik}\gamma_{\sigma_i\sigma_j}(k\tau_i)\right). \tag{131}$$

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 λ 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]):

$$[H_{tJ}, X_i^{0\sigma_i}] = \sum_j t_{ij}X_j^{0\sigma_i} + \boldsymbol{\mu}X_i^{0\sigma_i} - \sum_{j\sigma_j} t_{ij}(\sigma_i\sigma_j)X_i^{\bar{\sigma}_i\bar{\sigma}_j}X_j^{0\sigma_j} + \frac{1}{2}\sum_{j\neq i} J_{ij}\,(\sigma_i\sigma_j)X_j^{\bar{\sigma}_i\bar{\sigma}_j}X_i^{0\sigma_j}. \tag{132}$$

Temporarily ignoring the Fermionic sources:

$$[\hat{\mathcal{A}}_S(i\tau_i), X_i^{0\sigma_i}] = -\mathcal{V}_i^{\sigma_i\sigma_j}X_i^{0\sigma_j}, \tag{133}$$

and combining with the Heisenberg equation of motion, we see that the Green's function satisfies the EOM

$$\partial_{\tau_i}\mathcal{G}_{\sigma_i\sigma_f}(i,f) = -\delta(\tau_i-\tau_f)\delta_{ij}(1-\gamma_{\sigma_i\sigma_f}(i\tau_i)) - \left\langle T_\tau\left(e^{-\hat{\mathcal{A}}_S}[H_{tJ}+\hat{\mathcal{A}}_S(i,\tau_i), X_i^{0\sigma_i}(\tau_i)]\, X_f^{\sigma_f 0}(\tau_f)\right)\right\rangle \tag{134}$$

where the local Green's function

$$\gamma_{\sigma_a\sigma_b}(i\tau_i) = \sigma_a\sigma_b\mathcal{G}_{\bar{\sigma}_b\bar{\sigma}_a}(i\tau_i^-, i\tau_i). \tag{135}$$

Substituting and using the Fermi gas (i.e. free) Green's function:

$$\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j) = \left\{\delta_{\sigma_i\sigma_j}\left[\delta_{ij}(\boldsymbol{\mu}-\partial_{\tau_i})+t_{ij}\right]-\delta_{ij}\mathcal{V}_i^{\sigma_i\sigma_j}(\tau_i)\right\}\delta(\tau_i-\tau_j), \tag{136}$$

we obtain (using the repeated index summation and integration convention of Ref. [6])

$$\begin{aligned}\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j)\mathcal{G}_{\sigma_j\sigma_f}(j\tau_j, f\tau_f) &= \delta(\tau_i-\tau_f)\delta_{ij}(1-\gamma_{\sigma_i\sigma_f}(i\tau_i)) \\ &\quad - \sum_{j\sigma_j} t_{ij}(\sigma_i\sigma_j)\left\langle T_\tau\left(X_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i)X_j^{0\sigma_j}(\tau_i)X_f^{\sigma_f 0}(\tau_f)\right)\right\rangle \\ &\quad + \frac{1}{2}\sum_{j\sigma_j} J_{ij}(\sigma_i\sigma_j)\left\langle T_\tau\left(X_j^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i)X_i^{0\sigma_j}(\tau_i)X_f^{\sigma_f 0}(\tau_f)\right)\right\rangle.\end{aligned} \tag{137}$$

We next express the higher order Green's function in terms of the derivatives of the lower one to obtain the Schwinger EOM:

$$\begin{aligned}&\left(\mathbf{g}^{-1}_{0,\sigma_i,\sigma_j}(i\tau_i, j\tau_j) - \hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) - Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j)\right)\mathcal{G}_{\sigma_j\sigma_f}(j\tau_j, f\tau_f) \\ &= \delta_{if}\delta(\tau_i-\tau_f)\left(\delta_{\sigma_i\sigma_f}-\gamma_{\sigma_i\sigma_f}(i\tau_i)\right),\end{aligned} \tag{138}$$

where we used the functional derivative operator

$$D_{\sigma_i\sigma_j}(i\tau_i) = \sigma_i\sigma_j\frac{\delta}{\delta\mathcal{V}_i^{\bar{\sigma}_i\bar{\sigma}_j}(\tau_i)} \tag{139}$$

and the composite derivative operator

$$\hat{X}_{\sigma_i\sigma_j}(i\tau_i, j\tau_j) = \delta(\tau_i-\tau_j)\left(-t_{ij}D_{\sigma_i\sigma_j}(i\tau_i)+\delta_{ij}\sum_k\frac{1}{2}J_{ik}D_{\sigma_i\sigma_j}(k\tau_i)\right), \tag{140}$$

where the derivative $D_{\sigma_i\sigma_j}(k\tau_i)$ is at spatial site k and time τ_i. The corresponding Y_1 (with a similar convention as above) in Eq. (124) is

$$Y_{1\sigma_i\sigma_j}(i\tau_i, j\tau_j) = -\delta(\tau_i-\tau_j)\left(-t_{ij}\gamma_{\sigma_i\sigma_j}(i\tau_i)+\delta_{ij}\sum_k\frac{1}{2}J_{ik}\gamma_{\sigma_i\sigma_j}(k\tau_i)\right). \tag{141}$$

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_{eff}. We also note that Eq. (138) can be generalized to include the λ parameter by scaling $\hat{X}_{\sigma_i\sigma_j}, Y_{i\sigma_i\sigma_j}, \gamma_{\sigma_i\sigma_j} \to \lambda\hat{X}_{\sigma_i\sigma_j}, \lambda Y_{i\sigma_i\sigma_j}, \lambda\gamma_{\sigma_i\sigma_j}$.

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 $\left\{c_i, c_j^*\right\} = 0 = \left\{c_i, c_j\right\} = \left\{c_i^*, c_j^*\right\} = \left\{c_i, C_j^\dagger\right\}$. Suppressing j and spin index the Fermi coherent states are given as usual by:

$$
\begin{aligned}
|c\rangle &= e^{-c\,C^\dagger}|vac\rangle = \left(1 - c\,C^\dagger\right)\,|vac\rangle \\
\langle c| &= \langle vac|e^{-C\,c^*} = \langle vac|\left(1 - \,C\,c^*\right) \\
\langle c|c'\rangle &= 1 + c^*c' = e^{c^*c'},
\end{aligned}
\tag{142}
$$

where $|vac\rangle$ is the vacuum state. We use the abbreviation to denote coherent state integrals:

$$
\int_c = \int dc^*\,dc. \tag{143}
$$

The basic integrals are

$$
\begin{aligned}
\int_c \left(1, c^*, c, cc^*\right) &= (0, 0, 0, 1) \\
\int_c e^{-c^*c} &= 1.
\end{aligned}
\tag{144}
$$

The completeness relation reads:

$$
\int_c e^{-c^*c}\,|c\rangle\langle c| = |vac\rangle\langle vac| + C^\dagger|vac\rangle\langle vac|C \equiv \mathbb{1}, \tag{145}
$$

and the trace over Fermionic variables is given by:

$$
\mathrm{Tr}A = \int_c e^{-c^*c}\,\langle -c|A|c\rangle. \tag{146}
$$

Appendix C. Path integral representation

We now introduce path integrals to represent Eq. (68) leading to (102). Towards this end let us write

$$
\begin{aligned}
\beta\hat{H}_{eff} + \hat{\mathcal{A}}_S &= \int_0^\beta \hat{\mathscr{H}}(\tau)d\tau \\
\hat{\mathscr{H}}(\tau) &\equiv \hat{H}_{eff} + \sum_i \hat{\mathcal{A}}_S(i,\tau).
\end{aligned}
\tag{147}
$$

The integral is represented by a finite sum over M intervals, and the limit $M \to \infty$ taken at the end, thus

$$
\int_0^\beta \hat{\mathscr{H}}(\tau)d\tau \to \lim_{M\to\infty} \Delta\tau \sum_{j=1,M} \hat{\mathscr{H}}(\tau_j), \tag{148}
$$

where we defined

$$
\begin{aligned}
\tau_j &= \Delta\tau \times j = \frac{j\,\beta}{M}, \\
\Delta\tau &= \frac{\beta}{M}, \quad j = 1, M.
\end{aligned}
\tag{149}
$$

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

$$\begin{aligned} Z^{(M)} &= \int_{c(0)} e^{-c^*_{i\sigma}(0)c_{i\sigma}(0)} \langle -c(0)|e^{-\Delta\tau\hat{\mathcal{H}}(\tau_M)}\, e^{-\Delta\tau\hat{\mathcal{H}}(\tau_{M-1})} \dots\, e^{-\Delta\tau\hat{\mathcal{H}}(\tau_2)} e^{-\Delta\tau\hat{\mathcal{H}}(\tau_1)}\, \hat{P}_G|c(0)\rangle \\ &= \int_c e^{-\sum_{j=1}^{M} c^*_{i\sigma}(j)c_{i\sigma}(j)}\, \langle c(M)|e^{-\Delta\tau\hat{\mathcal{H}}(\tau_M)}|c(M-1)\rangle \dots\, |c(2)\rangle\langle c(2)| \\ &\quad \times e^{-\Delta\tau\hat{\mathcal{H}}(\tau_2)}|c(1)\rangle\langle c(1)|e^{-\Delta\tau\hat{\mathcal{H}}(\tau_1)}\, \hat{P}_G|c(0)\rangle. \end{aligned} \tag{150}$$

Anti periodic boundary conditions are used: $c(\tau_M) = -c(\tau_0)$ and we set at each time slice τ_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 σ are implicitly summed over. Recall that $\hat{P}_G$ is brought to the extreme right of the product. We calculate as usual:

$$\begin{aligned} &\langle c(\tau_{j+1})|c(\tau_j)\rangle = e^{c^*(\tau_{j+1})c(\tau_j)} \\ &\langle c(\tau_{j+1})|e^{-\Delta\tau\hat{\mathcal{H}}(\tau_{j+1})}|c(\tau_j)\rangle \equiv e^{c^*(\tau_{j+1})c(\tau_j) - \Delta\tau\mathcal{H}(\tau_{j+1},\tau_j)} + O(\Delta\tau^2) \\ &\mathcal{H}(\tau_{j+1},\tau_j) \equiv \frac{\langle c(\tau_{j+1})|\hat{\mathcal{H}}(\tau_{\tau_{j+1}})|c(\tau_j)\rangle}{\langle c(\tau_{j+1})|c(\tau_j)\rangle}. \end{aligned} \tag{151}$$

The last term needs careful attention, we note

$$\begin{aligned} \langle c(\tau_1)|e^{-\Delta\tau\hat{\mathcal{H}}(\tau_1)}\, \hat{P}_G|c(\tau_0)\rangle &= \langle c(\tau_1)|(1-\Delta\tau\hat{\mathcal{H}}(\tau_1))\, \hat{P}_G|c(\tau_0)\rangle + O(\Delta\tau^2) \\ &= \langle c(\tau_1)|c(\tau_0)\rangle(1-\Delta\tau\mathcal{H}(\tau_1,\tau_0))\, P_G(\tau_1,\tau_0) + O(\Delta\tau^2) \\ &= \langle c(\tau_1)|c(\tau_0)\rangle e^{-\Delta\tau\mathcal{H}(\tau_1,\tau_0)}\, P_G(\tau_1,\tau_0) + O(\Delta\tau^2), \end{aligned} \tag{152}$$

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 τ

$$\langle \tau|C^\dagger_{i\sigma}|\tau_j\rangle = c^*_{i\sigma}(\tau)e^{c^*_{i\sigma}(\tau)c_{i\sigma}(\tau_j)} = -\frac{\delta}{\delta c_{i\sigma}(\tau_j)}\langle \tau|\tau_j\rangle. \tag{153}$$

In view of this relation we note the following mappings:

$$\begin{aligned} \langle \psi|C^\dagger_{i\sigma}|\tau_j\rangle &\to -\frac{\delta}{\delta c_{i\sigma}(\tau_j)}\langle \psi|\tau_j\rangle, \\ \langle \psi|C_{i\sigma}|\tau_j\rangle &\to c_{i\sigma}(\tau_j)\langle \psi|\tau_j\rangle, \\ \langle \tau_j|C^\dagger_{i\sigma}|\psi\rangle &\to c^*_{i\sigma}(\tau_j)\langle \tau_j|\psi\rangle, \\ \langle \tau_j|C_{i\sigma}|\psi\rangle &\to \frac{\delta}{\delta c^*_{i\sigma}(\tau_j)}\langle \tau_j|\psi\rangle, \end{aligned} \tag{154}$$

Let us show how the commutation works here:

$$(CC^\dagger + C^\dagger C)|c\rangle = \left(-C\frac{\delta}{\delta c} + C^\dagger c\right)|c\rangle = \left(\frac{\delta}{\delta c}C - cC^\dagger\right)|c\rangle = \left(\frac{\delta}{\delta c}c + c\frac{\delta}{\delta c}\right)|c\rangle = |c\rangle. \tag{155}$$

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 spices, 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:

$$\begin{aligned} &c_{i\sigma}(\tau_j)c^*_{i\sigma}(\tau_j) \to 1\\ &c^*_{i\sigma}(\tau_j)c_{i\sigma}(\tau_j) \to -1\\ &c_{i\sigma}(\tau_j)c_{i\sigma}(\tau_j) \to 0\\ &c^*_{i\sigma}(\tau_j)c^*_{i\sigma}(\tau_j) \to 0. \end{aligned} \tag{156}$$

- (II) We denote the number operator as

$$n_{i\sigma}(\tau_j) \equiv c^*_{i\sigma}(\tau_{j+1})\, c_{i\sigma}(\tau_j), \tag{157}$$

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

$$\begin{aligned} &n_{i\sigma}(\tau_{j+1})n_{i\sigma}(\tau_j) = n_{i\sigma}(\tau_j)\\ &c^*_{i\sigma}(\tau_{j+1})n_{i\sigma}(\tau_j) = 0\\ &n_{i\sigma}(\tau_j)c^*_{i\sigma}(\tau_j) = c^*_{i\sigma}(\tau_{j+1})\\ &n_{i\sigma}(\tau_j)c_{i\sigma}(\tau_j) = 0\\ &c_{i\sigma}(\tau_j)n_{i\sigma}(\tau_{j-1}) = c_{i\sigma}(\tau_{j-1}). \end{aligned} \tag{158}$$

(III) We next obtain the important *Gutzwiller exclusion identity*. Calling the ith term in the product (109) as $P^{(i)}_G(\tau_1, \tau_0)$; we see that

$$\begin{aligned} &n_{i\uparrow}(\tau_1)n_{i\downarrow}(\tau_1)P^{(i)}_G(\tau_1, \tau_0) = n_{i\uparrow}(\tau_1)n_{i\downarrow}(\tau_1)\\ &-c^*_{i\uparrow}(\tau_2)c_{i\uparrow}(\tau_0)c^*_{i\downarrow}(\tau_2)c_{i\downarrow}(\tau_0) \sim 0. \end{aligned} \tag{159}$$

The last line follows upon expanding $\tau_2 (\equiv\ \tau_1 + \Delta\tau)$ about τ_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 $\tau_j \geq \tau_1$:

$$c_{i\uparrow}(\tau_j)c_{i\downarrow}(\tau_j)\ldots P^{(i)}_G(\tau_1, \tau_0) = 0, \tag{160}$$

where the dots indicate contributions from intermediate times. These contributions, after Grassmann integration over the terms at intermediate times, must necessarily end up with $\ldots c_{i\uparrow}(\tau_1)c_{i\downarrow}(\tau_1)P^{(i)}_G(\tau_1, \tau_0)$. 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:

$$\xi_{i\sigma_i}(\tau_i) \equiv \frac{1}{Z}\|X^{0\sigma_i}_i(\tau_i)\| = \frac{1}{Z}\|c_{i\sigma_i}(\tau_i)\|, \tag{161}$$

and further abbreviate $\xi_{i\sigma_i}(\tau_i) \leftrightarrow \xi_{\tilde{i}}$. A creation variable average $\xi^*_{\tilde{i}} \equiv \frac{1}{Z}\|X_i^{\sigma_i 0}(\tau_i)\|$ is also useful. The variable $\xi_{\tilde{i}}$ satisfies the functional differential equation that we study next:

$$(\mathbf{g}^{-1}_{0,\tilde{i},\tilde{j}} - \hat{X}_{\tilde{i}\tilde{j}})(Z\xi_{\tilde{j}}) = Z(\delta_{\tilde{i},\tilde{k}} - \gamma_{\tilde{i},\tilde{k}})J_{\tilde{k}}, \tag{162}$$

or using Eq. (124)

$$(\mathbf{g}^{-1}_{0,\tilde{i},\tilde{j}} - \hat{X}_{\tilde{i}\tilde{j}} - Y_{1\,\tilde{i}\tilde{j}})\,\xi_{\tilde{j}} = (\delta_{\tilde{i},\tilde{k}} - \gamma_{\tilde{i},\tilde{k}})J_{\tilde{k}}. \tag{163}$$

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):

$$\mathcal{G}_{\tilde{i}\tilde{f}} - \xi^*_{\tilde{f}}\,\xi_{\tilde{i}} = \frac{\delta\xi_{\tilde{i}}}{\delta J_{\tilde{f}}}, \tag{164}$$

and on taking the limit $J \to 0, J^* \to 0$, all the single Fermi expectations ξ, ξ^* 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 γ 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

$$J_{\tilde{i}} = (\mu^{-1})_{\tilde{i}\tilde{j}}\,\mathcal{J}_{\tilde{j}}, \tag{165}$$

so that

$$\frac{\delta}{\delta J_{\tilde{f}}} = \frac{\delta}{\delta \mathcal{J}_{\tilde{k}}}\mu_{\tilde{k}\tilde{f}}, \tag{166}$$

where (μ^{-1}) is the *matrix inverse* of μ. These equations have inverses that are easily obtained. The matrix μ 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

$$\begin{aligned}\mathcal{G}_{\tilde{i}\tilde{f}} &= \mathbf{g}_{\tilde{i}\tilde{k}}\,\mu_{\tilde{k}\tilde{f}} + \xi^*_{\tilde{f}}\,\xi_{\tilde{i}},\\ \mathbf{g}_{\tilde{i}\tilde{k}} &= \frac{\delta\xi_{\tilde{i}}}{\delta\mathcal{J}_{\tilde{k}}}.\end{aligned} \tag{167}$$

The goal is to choose μ such that the so defined $\mathbf{g}$ satisfies a canonical equation, i.e. the analog of Eq. (127), but without the γ term on the right. For this purpose we can differentiate Eq. (163) with the scaled source field $\mathcal{J}_{\tilde{k}}$, taking care to observe the non commutation of $\mathcal{J}_{\tilde{k}}$ with the derivative term $\hat{X}$. This process yields the equations:

$$(\mathbf{g}^{-1}_{0,\tilde{i},\tilde{j}} - \hat{X}_{\tilde{i}\tilde{j}} - Y_{1\,\tilde{i}\tilde{j}})\,\mathbf{g}_{\tilde{j}\tilde{k}} = (\delta_{\tilde{i},\tilde{j}} - \gamma_{\tilde{i},\tilde{j}})(\mu^{-1})_{\tilde{j}\tilde{k}} - \left[\hat{X}_{\tilde{i}\tilde{j}}(\mu^{-1})_{\tilde{l}\tilde{k}}\right]\mathbf{g}_{\tilde{j}\tilde{m}}\mu_{\tilde{m}\tilde{l}}, \tag{168}$$

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 μ is confined to the right hand side, and hence we set both sides to the identity matrix:

$$(\mathbf{g}^{-1}_{0,\tilde{i},\tilde{j}} - \hat{X}_{\tilde{i}\tilde{j}} - Y_{1\,\tilde{i}\tilde{j}})\,\mathbf{g}_{\tilde{j}\tilde{k}} = \delta_{\tilde{i}\tilde{k}}, \tag{169}$$

and

$$(\delta_{\tilde{i},\tilde{j}} - \gamma_{\tilde{i},\tilde{j}})(\mu^{-1})_{\tilde{j}\tilde{k}} - \left[\hat{X}_{\tilde{i}\tilde{j}}(\mu^{-1})_{\tilde{l}\tilde{k}}\right]\mathbf{g}_{\tilde{j}\tilde{m}}\mu_{\tilde{m}\tilde{l}} = \delta_{\tilde{i}\tilde{k}}. \tag{170}$$

Multiplying through with μ and using

$$\left[\hat{X}_{\tilde{i}\tilde{j}}(\mu^{-1})_{\tilde{l}\tilde{k}}\right]\mu_{\tilde{m}\tilde{l}} = -\left[\hat{X}_{\tilde{i}\tilde{j}}\mu_{\tilde{m}\tilde{l}}\right](\mu^{-1})_{\tilde{l}\tilde{k}}, \tag{171}$$

we rewrite Eq. (170) as

$$\mu_{\tilde{i}\tilde{f}} = (\delta_{\tilde{i}\tilde{f}} - \gamma_{\tilde{i}\tilde{f}}) + \mathbf{g}_{\tilde{j}\tilde{k}}\left[\hat{X}_{\tilde{i}\tilde{j}}\mu_{\tilde{k}\tilde{f}}\right]. \tag{172}$$

We next show that Eqs. (169) and (170) can be rewritten in terms of the two self energies Φ and Ψ used in Refs. [4,6]. We need the relation analogous to Eq. (171) to simplify Eq. (169):

$$\left[\hat{X}_{\tilde{i}\tilde{j}}\mathbf{g}_{\tilde{j}\tilde{k}}\right] = -\mathbf{g}_{\tilde{j}\tilde{k}}\left[\hat{X}_{\tilde{i}\tilde{j}}\mathbf{g}^{-1}_{\tilde{k}\tilde{l}}\right]\mathbf{g}_{\tilde{l}\tilde{k}}. \tag{173}$$

Therefore we write the two equations as

$$\begin{aligned}(\mathbf{g}^{-1}_{0,\tilde{i},\tilde{j}} - \Phi_{\tilde{i}\tilde{j}} - Y_{1\,\tilde{i}\tilde{j}})\,\mathbf{g}_{\tilde{j}\tilde{k}} &= \delta_{\tilde{i}\tilde{k}},\\ \mu_{\tilde{i}\tilde{f}} &= (\delta_{\tilde{i}\tilde{f}} - \gamma_{\tilde{i}\tilde{f}}) + \Psi_{\tilde{i}\tilde{f}},\end{aligned} \tag{174}$$

where the two self energies Φ and Ψ 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

$$\begin{aligned}\Phi_{\tilde{i}\tilde{j}} &= -\mathbf{g}_{\tilde{m}\tilde{k}}\left[\hat{X}_{\tilde{i}\tilde{m}}\mathbf{g}^{-1}_{\tilde{k}\tilde{j}}\right]\\ \Psi_{\tilde{i}\tilde{f}} &= \mathbf{g}_{\tilde{j}\tilde{k}}\left[\hat{X}_{\tilde{i}\tilde{j}}\mu_{\tilde{k}\tilde{f}}\right].\end{aligned} \tag{175}$$

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 $\langle A\rangle = \frac{1}{Z}\mathrm{Tr}_{tJ} e^{-\beta H} A$. Here $H = H_{tJ}$ or $H = H_{eff}$ 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 λ, since typically the interaction strength is an experimentally determined number (e.g. the Coulomb energy in hydrogenic problems). Introducing λ 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 ω. 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. Fedorov1, 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. Willemsen, Phys. Rev. D 17 (1978) 574;
M. Creutz, I.J. Muzinich, T.N. Tudron, Phys. Rev. D19 (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 $[C^\dagger_{\sigma_1}(1-\lambda N_{\bar\sigma_1}), C^\dagger_{\sigma_2} C_{\sigma_3}]$ reduces to $-\delta_{\sigma_1\sigma_3} C^\dagger_{\sigma_2}(1-\lambda N_{\bar\sigma_2})$ as follows. By straightforward expansion of the commutator we get $-\delta_{\sigma_1\sigma_3} C^\dagger_{\sigma_2} + \lambda C^\dagger_{\sigma_2}\left(\delta_{\sigma_1\sigma_3} N_{\bar\sigma_1} + \delta_{\sigma_2\bar\sigma_3}(\delta_{\bar\sigma_1\sigma_2} - \delta_{\bar\sigma_1\sigma_3}) C^\dagger_{\sigma_1} C_{\sigma_3}\right)$. 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 $C^\dagger_{\sigma_2} C^\dagger_{\sigma_2} = 0$, the last term then simplifies to $\lambda\delta_{\sigma_1\sigma_3} C^\dagger_{\sigma_2}(N_{\bar\sigma_2}\delta_{\bar\sigma_1\sigma_2} + N_{\bar\sigma_1})$, 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 $C^\dagger_{\sigma_2} C^\dagger_{\sigma_2} = 0$, the required result follows.
[41] This alternative choice, upon using Eq. (85) leads to

$$n_{i\sigma} = \frac{1}{\lambda}\{1 - \mathcal{G}^{(\lambda)}_{\bar\sigma\bar\sigma}(i\tau^- i\tau) + \mathcal{G}^{(\lambda)}_{\bar\sigma\bar\sigma}(i\tau^+ i\tau)\}.$$

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, Dordrect, 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 $M \to \infty$ at the end, we will simplify the notation from here by writing $\widetilde{c}^*_{f\sigma_f}(\tau_{f+1}) \to \widetilde{c}^*_{f\sigma_f}(\tau_f)$.
[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 $\xi \to (\mathbb{1}-\gamma)\xi'$. Together, these transformations give an exact formulation with a canonical Green's function $G_{i,f} = \frac{\delta\xi'_i}{\delta J_{\bar f}}$ 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 λ parameter, seems to be a worthwhile goal for the future.

Annals of Physics 373 (2016) 717–718

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: 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
$\hat{k} = (\vec{k} - \vec{k}_F).\vec{k}_F/|\vec{k}_F|$
by
$\hat{k} = (\vec{k} - \vec{k}_F).\nabla\varepsilon_k/|\nabla\varepsilon_k|$.

(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.

http://dx.doi.org/10.1016/j.aop.2016.08.015

(3) Page 175 Third line after Eq. (44).
Replace
$sign(\hat{k})$
by
$sign(-Q)$.
(4) Page 175 Fourth line after Eq. (44).
Replace
$sign(\hat{k})$
by
$sign(-Q)$.
(5) Eq. (45) should read

$$E^*(k_{kink}) = -\frac{1}{r-1}\Delta_0 - \Gamma_0.$$

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.

Annals of Physics 338 (2013) 283–301

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: 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 λ.
- 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 ($d \to \infty$) 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 $d \to \infty$, 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 λ, 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 (E. Perepelitsky).

0003-4916/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.aop.2013.09.010

factor μ. 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 λ.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Motivation

The Hubbard model (HM) with the Hamiltonian:

$$H = -\sum_{ij\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow} - \boldsymbol{\mu} \sum_{i} n_i, \tag{1}$$

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 $\mathcal{G}$:

$$\mathcal{G}(k) = \frac{1}{i\omega_k + \boldsymbol{\mu} - \epsilon_k - \Sigma_D(k)}, \tag{2}$$

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 on sets $U \to \infty$ 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 $U \to \infty$ 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]

$$H = -\sum_{ij\sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \boldsymbol{\mu} \sum_{i\sigma} X_i^{\sigma\sigma} + \frac{1}{2} \sum_{ij\sigma} J_{ij} X_i^{\sigma\sigma} + \frac{1}{4} \sum_{ij\sigma_1\sigma_2} J_{ij} \{X_i^{\sigma_1\sigma_2} X_j^{\sigma_2\sigma_1} - X_i^{\sigma_1\sigma_1} X_j^{\sigma_2\sigma_2}\}. \tag{3}$$

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

$$\mathcal{G}_{\sigma_1\sigma_2}(i, f) = -\langle T_\tau X_i^{0\sigma_1}(\tau_i) X_f^{\sigma_2 0}(\tau_f)\rangle, \tag{4}$$

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 $\frac{1-\frac{n}{2}}{i\omega_n}$ 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:

$$\mathcal{G}(k) = \frac{1-\frac{n}{2}}{i\omega_k + \boldsymbol{\mu} - \epsilon_k\left(1-\frac{n}{2}\right) - \Sigma_{DM}(k)}. \tag{5}$$

Just as is the case for Σ_D in the finite-U Hubbard model, Σ_{DM} is finite as $i\omega \to \infty$ 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 $k = (\vec{k}, i\omega_k)$:

$$\mathcal{G}(k) = \mathbf{g}(k) \times \mu(k). \tag{6}$$

These two factors are in turn written in terms of two self-energies, $\Phi(k)$ and $\Psi(k)$.

$$\mathbf{g}^{-1}(k) = i\omega_k + \boldsymbol{\mu} - (1 - n/2)\epsilon_k - \Phi(k), \tag{7}$$

$$\mu(k) = 1 - \frac{n}{2} + \Psi(k). \tag{8}$$

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 $1 - \frac{n}{2}$. Φ and Ψ are then given in terms of the vertices (i.e. functional derivatives w.r.t. the source of the $\mathbf{g}^{-1}$ and μ) 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 λ. Here λ interpolates between the free Fermi gas and the t–J model. The meaning of λ as a partial projection parameter is detailed in [18], and may be summarized in the mapping $X_i^{\sigma 0} \to f_{i\sigma}^\dagger(1-\lambda\, n_{i\bar{\sigma}})$, 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, $J \to 0$, 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 $\frac{U}{d} \to 0$, while in the case of the latter, $\frac{U}{d} \to \infty$. 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 $U \to \infty$ and $d \to \infty$ 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 λ 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.

$$\Psi(k) = \Psi(i\omega_k), \tag{9}$$

$$\Phi(k) = \chi(i\omega_k) + \epsilon_k \Psi(i\omega_k). \tag{10}$$

These in turn show that the Dyson–Mori self energy behaves as

$$\Sigma_{DM}(k) = \Sigma_{DM}(i\omega_k) = \frac{(i\omega_k + \boldsymbol{\mu})\Psi(i\omega_k) + \left(1 - \frac{n}{2}\right)\chi(i\omega_k)}{1 - \frac{n}{2} + \Psi(i\omega_k)}, \tag{11}$$

and is therefore local in the limit of infinite dimensions. We show that to each order in the λ 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

$$\sigma^{\alpha\beta}(\omega) = \frac{2}{i\omega}\sum_{\vec{p}, i\omega_p} \mathcal{G}(\vec{p}, i\omega_p) v_{\vec{p}}^{\alpha} v_{\vec{p}}^{\beta} [\mathcal{G}(\vec{p}, \omega + i\eta + i\omega_p) - \mathcal{G}(\vec{p}, i\eta + i\omega_p)], \tag{12}$$

where $v_{\vec{p}}^{\alpha}$ is the component of the velocity in the α direction (Eq. (39)). We show that this formula can be applied at each order of the λ 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

$$\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) = \sum_{\vec{k}} \frac{|V_{\vec{k}}|^2}{i\omega_n - \widetilde{\epsilon}_{\vec{k}}} \mathbf{g}(i\omega_k), \tag{13}$$

where $\epsilon_{\vec{k}}$ is the dispersion of the lattice in the t–J model, and $V_{\vec{k}}$ and $\widetilde{\epsilon}_{\vec{k}}$ 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 λ 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 λ 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 λ 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 λ, 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 λ. 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 λ.

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 $t_{ij} \to \frac{1}{\sqrt{2d}} t_0$ 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

$$W_{i,\mathbf{n}} V_{\mathbf{n},m} \sim O\left(\frac{1}{\left(\sqrt{d}\right)^{r_{im}}}\right). \tag{14}$$

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_0} 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 λ 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 λ 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 $J \to 0$ (as discussed in Section 1.1), we write the minimal theory ECFL equations of motion in expanded form:

$$\begin{aligned} \mathbf{g}^{-1}[i,m] &= (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i)\delta[i,m] + t[i,m]\,(1 - \lambda\gamma[i]) \\ &\quad + \lambda t[i,\mathbf{j}]\,\xi^* \cdot \mathbf{g}[\mathbf{j},\mathbf{n}] \cdot \Lambda_*[\mathbf{n},m;i], \\ \mu[i,m] &= (1 - \lambda\gamma[i])\delta[i,m] - \lambda t[i,\mathbf{j}]\,\xi^* \cdot \mathbf{g}[\mathbf{j},\mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n},m;i], \end{aligned} \tag{15}$$

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 μ 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

$$\begin{aligned} &\gamma[i] = \mu^{(k)}[\mathbf{n}, i^+] \cdot \mathbf{g}^{(k)}[i,\mathbf{n}]; \qquad \Lambda[n,m;i] = -\frac{\delta}{\delta\mathcal{V}_i}\mathbf{g}^{-1}[n,m]; \\ &\mathcal{U}[n,m;i] = \frac{\delta}{\delta\mathcal{V}_i}\mu[n,m], \end{aligned} \tag{16}$$

where we have used the notation $M^{(k)}_{\sigma_1,\sigma_2} = \sigma_1\sigma_2 M_{\bar{\sigma}_2,\bar{\sigma}_1}$ 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 μ. The vertices defined above (Λ and $\mathcal{U}$) have four spin indices, those of the object being differentiated and those of the source. For example, $\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_a\sigma_b}[n, m; i] = \frac{\delta}{\delta \mathcal{V}_i^{\sigma_a\sigma_b}} \mu_{\sigma_1\sigma_2}[n, m]$. 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: $(\xi^* \cdot \mathbf{g}[j, \mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i])_{\sigma_1\sigma_2} = \sigma_1\sigma_{\mathbf{a}}\mathbf{g}_{\sigma_{\mathbf{a}},\sigma_{\mathbf{b}}}[j, \mathbf{n}] \frac{\delta}{\delta \mathcal{V}_i^{\bar{\sigma}_1\bar{\sigma}_{\mathbf{a}}}} \mu_{\sigma_{\mathbf{b}},\sigma_2}[\mathbf{n}, m]$. Finally, in order to ensure that the shift identities (Ref. [18]) are satisfied, the substitution $t_{ij} \to t_{ij} + \frac{u_0}{2}\delta_{ij}$ 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 λ expansion is obtained by expanding Eqs. (15) and (16) iteratively in the continuity parameter λ. The $\lambda = 0$ limit of these equations is the free Fermi gas. Therefore, a direct expansion in λ will lead to a series in λ 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 $\mathbf{g}_0[i, f] \to \mathbf{g}[i, f]$. However, one can also obtain the skeleton expansion directly by expanding Eqs. (15) and (16) in λ, 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 $\frac{\delta \mathbf{g}}{\delta \mathcal{V}}$. This is done with the help of the following useful formula which stems from the product rule for functional derivatives:

$$\frac{\delta \mathbf{g}[i, m]}{\delta \mathcal{V}_r} = \mathbf{g}[i, \mathbf{x}] \cdot \Lambda[\mathbf{x}, \mathbf{y}, r] \cdot \mathbf{g}[\mathbf{y}, m]. \tag{17}$$

This is an exact formula and will be used extensively in the arguments given below. Within the λ expansion, the LHS is evaluated to a certain order in λ by taking the vertex Λ on the RHS to be of that order in λ.

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:

$$\frac{\delta}{\delta \mathcal{V}_i} L_j = \delta_{ij} L_i', \tag{18}$$

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.

$$\frac{\delta}{\delta \mathcal{V}_{r_1}} \cdots \frac{\delta}{\delta \mathcal{V}_{r_s}} L_i = \delta_{ir_1} \ldots \delta_{ir_s} \frac{\delta}{\delta \mathcal{V}_i(\tau_{r_1})} \cdots \frac{\delta}{\delta \mathcal{V}_i(\tau_{r_s})} L_i. \tag{19}$$

In the presence of the current source κ (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}

$$O_{if} = W_{f,\mathbf{x}} \frac{\delta}{\delta \kappa_i^\alpha} (L_{\mathbf{x}})\, V_{\mathbf{x},f}, \tag{20}$$

where the functions $V_{\mathbf{x},f}$, $W_{f,\mathbf{x}} \sim O\left(\frac{1}{(\sqrt{d})^{r_{\mathbf{x}f}}}\right)$. Then, neglecting terms of higher order than $O\left(\frac{1}{d^{r_{if}}}\right)$ in O_{if}, $\sum_{i-f} O_{if} \to 0$ as $\mathcal{A} \to 0$. Again iterating property (c) and using property (d), the following must hold for any nonnegative integer s:

$$\sum_{i-f} \left(W_{f,\mathbf{x}} \frac{\delta}{\delta \kappa_i^\alpha} \frac{\delta}{\delta \mathcal{V}_{\mathbf{x}}(\tau_{r_1})} \cdots \frac{\delta}{\delta \mathcal{V}_{\mathbf{x}}(\tau_{r_s})} (L_{\mathbf{x}})\, V_{\mathbf{x},f} \right)_{\mathcal{A} \to 0} = 0. \tag{21}$$

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 $\mathbf{g}[i, j] \to \mathbf{g}_{i,j}[\tau_i, \tau_j]$, 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

$$\begin{aligned} \mathbf{g}^{-1}[i, m] &= (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i)\, \delta[i, m] + t[i, m]\, (1 - \lambda \gamma[i]) - \lambda\, \delta_{i,m}\, \chi_i[\tau_i, \tau_m] \\ &\quad + \lambda\, t_{i,m}\, \Psi_i[\tau_i, \tau_m], \\ \mu[i, m] &= \delta[i, m](1 - \lambda \gamma[i]) + \lambda\, \delta_{i,m}\, \Psi_i[\tau_i, \tau_m], \end{aligned} \tag{22}$$

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 Λ, $\mathcal{U}$, and γ. 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.

$$\begin{aligned}
&\Lambda[n, m; i] = \delta_{i,n}\delta_{i,m}\, A_i[\tau_n, \tau_m; \tau_i] + \delta_{i,n} t_{n,m}\, B_i[\tau_n, \tau_m; \tau_i],\\
&\mathcal{U}[n, m; i] = -\delta_{i,n}\delta_{i,m}\, B_i[\tau_n, \tau_m; \tau_i],\\
&\gamma[i] = \left(1 - \lambda\gamma^{(k)}[i]\right)\mathbf{g}^{(k)}[i, i] + \lambda\, \Psi_i^{(k)}[\tau_{\mathbf{j}}, \tau_i]\mathbf{g}_{ii}^{(k)}[\tau_i, \tau_{\mathbf{j}}],
\end{aligned} \tag{23}$$

where we defined two new functions:

$$\begin{aligned}
&A_i[\tau_n, \tau_m; \tau_i] = \delta[\tau_i, \tau_n]\delta[\tau_i, \tau_m]\, \mathbb{1} + \lambda\, \frac{\delta}{\delta \mathcal{V}_i}\chi_i[\tau_n, \tau_m],\\
&B_i[\tau_n, \tau_m; \tau_i] = \lambda\, \delta[\tau_n, \tau_m]\frac{\delta}{\delta \mathcal{V}_i}\gamma_i[\tau_n] - \lambda\, \frac{\delta}{\delta \mathcal{V}_i}\Psi_i[\tau_n, \tau_m].
\end{aligned} \tag{24}$$

Here A_i and B_i are class L functions since they inherit this property from Ψ_i, χ_i, and $\gamma[i]$ by functional differentiation. Substituting Eq. (23) into Eq. (15) and comparing with Eq. (22),

$$\begin{aligned}
&\chi_i[\tau_i, \tau_m] = -t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{n}}] \cdot A_{i,*}[\tau_{\mathbf{n}}, \tau_m; \tau_i],\\
&\Psi_i[\tau_i, \tau_m] = t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{n}}] \cdot B_{i,*}[\tau_{\mathbf{n}}, \tau_m; \tau_i].
\end{aligned} \tag{25}$$

If we can now show that χ_i, Ψ_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 $t_{i,\mathbf{j}}\, \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_m]$ are Class L functions. Taking their functional derivatives we obtain:

$$\begin{aligned}
\frac{\delta}{\delta \mathcal{V}_r} t_{i,\mathbf{j}}\, \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_m] \;=\; & t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},r}[\tau_i, \tau_{\mathbf{k}}]\, A_r[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]\mathbf{g}_{r,i}[\tau_{\mathbf{l}}, \tau_m]\\
& + t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},r}[\tau_i, \tau_{\mathbf{k}}]\, B_r[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]t_{r,\mathbf{l}}\mathbf{g}_{\mathbf{l},i}[\tau_{\mathbf{l}}, \tau_m],
\end{aligned} \tag{26}$$

and

$$\begin{aligned}
\frac{\delta}{\delta \mathcal{V}_r}\mathbf{g}_{i,i}[\tau_i, \tau_m] \;=\; & \mathbf{g}_{i,r}[\tau_i, \tau_{\mathbf{k}}]\, A_r[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]\mathbf{g}_{r,i}[\tau_{\mathbf{l}}, \tau_m]\\
& + \mathbf{g}_{i,r}[\tau_i, \tau_{\mathbf{k}}]\, B_r[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]t_{r,\mathbf{l}}\mathbf{g}_{\mathbf{l},i}[\tau_{\mathbf{l}}, \tau_m].
\end{aligned} \tag{27}$$

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 $r \to i$, 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 $t_{i,\mathbf{j}}\, \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_m]$ 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 λ, since the proof was done with λ present in all of the equations. In the bare expansion, this would imply that they also hold to each order in λ. 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 λ in the skeleton expansion. In proving this, we shall shed more light on the nature of the objects Ψ_i, χ_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 λ, 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 $t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},i}[\tau_n, \tau_m]$ where the time indices are arbitrary. *Inductive hypothesis*: through nth order in λ, Eqs. (22) and (23) hold. Through $n-1$st order in λ, the objects Ψ_i, χ_i, and $\gamma[i]$, and through nth 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):

$$(L_i)^{(n)} = \lambda^n (R_i)^m, \tag{28}$$

where m is arbitrary. We first examine the base case of zeroth order. In this case,

$$A_i^{(0)}[\tau_n, \tau_m; \tau_i] = \delta[\tau_i, \tau_n]\delta[\tau_i, \tau_m]; \qquad B_i^{(0)}[\tau_n, \tau_m; \tau_i] = 0. \tag{29}$$

Clearly the hypothesis is satisfied. Now, we prove the inductive step. Explicitly displaying the order in λ of all objects, the equations for χ, Ψ, and γ (Eqs. (25) and (23)) become

$$\begin{aligned}
\chi_i^{(n)}[\tau_i, \tau_m] &= -t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{n}}] \cdot A_{i,*}^{(n)}[\tau_{\mathbf{n}}, \tau_m; \tau_i], \\
\Psi_i^{(n)}[\tau_i, \tau_m] &= \ t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{n}}] \cdot B_{i,*}^{(n)}[\tau_{\mathbf{n}}, \tau_m; \tau_i], \\
\gamma^{(n)}[i] &= -\lambda\, \gamma^{(k)(n-1)}[i]\mathbf{g}^{(k)}[i, i] + \lambda\, \Psi_i^{(k)(n-1)}[\tau_{\mathbf{j}}, \tau_i]\mathbf{g}_{ii}^{(k)}[\tau_i, \tau_{\mathbf{j}}].
\end{aligned} \tag{30}$$

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

$$\begin{aligned}
A_i^{(n+1)}[\tau_n, \tau_m; \tau_i] &= \lambda\, \left(\sum_{r\le n} \frac{\delta}{\delta \mathcal{V}_i} \chi_i^{(r)}[\tau_n, \tau_m]\right)^{(n)}, \\
B_i^{(n+1)}[\tau_n, \tau_m; \tau_i] &= \lambda\, \delta[\tau_n, \tau_m] \left(\sum_{r\le n} \frac{\delta}{\delta \mathcal{V}_i} \gamma_i^{(r)}[\tau_n]\right)^{(n)} - \lambda\, \left(\sum_{r\le n} \frac{\delta}{\delta \mathcal{V}_i} \Psi_i^{(r)}[\tau_n, \tau_m]\right)^{(n)}.
\end{aligned} \tag{31}$$

To see that $A^{(n+1)}$ and $B^{(n+1)}$ have the required form we note that for all $l \le n$,

$$\begin{aligned}
\left(\frac{\delta}{\delta \mathcal{V}_r} t_{i,\mathbf{j}}\, \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_m]\right)^{(l)} &= t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},r}[\tau_i, \tau_{\mathbf{k}}]\, A_r^{(l)}[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]\mathbf{g}_{r,i}[\tau_{\mathbf{l}}, \tau_m] \\
&\quad + t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},r}[\tau_i, \tau_{\mathbf{k}}]\, B_r^{(l)}[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]t_{r,\mathbf{l}}\mathbf{g}_{\mathbf{l},i}[\tau_{\mathbf{l}}, \tau_m],
\end{aligned} \tag{32}$$

and

$$\begin{aligned}
\left(\frac{\delta}{\delta \mathcal{V}_r} \mathbf{g}_{i,i}[\tau_i, \tau_m]\right)^{(l)} &= \mathbf{g}_{i,r}[\tau_i, \tau_{\mathbf{k}}]\, A_r^{(l)}[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]\mathbf{g}_{r,i}[\tau_{\mathbf{l}}, \tau_m] \\
&\quad + \mathbf{g}_{i,r}[\tau_i, \tau_{\mathbf{k}}]\, B_r^{(l)}[\tau_{\mathbf{k}}, \tau_{\mathbf{l}}; \tau_r]t_{r,\mathbf{l}}\mathbf{g}_{\mathbf{l},i}[\tau_{\mathbf{l}}, \tau_m].
\end{aligned} \tag{33}$$

In the limit of large dimensions, $r \to i$. 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

$$\begin{aligned}
A_i^{(n+1)} &= \sum_{r=0}^{n} \lambda^{r+1} \left(\frac{\delta}{\delta \mathcal{V}_i} (R_i)^m\right)^{(n-r)}, \\
B_i^{(n+1)} &= \sum_{r=0}^{n} \lambda^{r+1} \left(\frac{\delta}{\delta \mathcal{V}_i} (R_i)^m\right)^{(n-r)}.
\end{aligned} \tag{34}$$

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 $t_{i,j} \to t_{i,j} + \frac{u_0}{2}\delta_{i,j}$ 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 $t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},i}[\tau_n, \tau_m]$ with the local function $t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},i}[\tau_n, \tau_m] + \frac{u_0}{2}\mathbf{g}_{i,i}[\tau_n, \tau_m]$. 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, $\gamma[i] \to \frac{n}{2}$. Then, the above results can be summed up in the following formulae (in which we set $\lambda = 1$):

$$\begin{aligned} \mathbf{g}^{-1}(k) &= i\omega_k + \boldsymbol{\mu} - \varepsilon_k\left(1 - \frac{n}{2}\right) - \chi(i\omega_k) - \varepsilon_k \Psi(i\omega_k), \\ \mu(k) &= 1 - \frac{n}{2} + \Psi(i\omega_k), \end{aligned} \tag{35}$$

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

$$\mathcal{G}(k) = \frac{1 - \frac{n}{2} + \Psi(i\omega_k)}{i\omega_k + \boldsymbol{\mu} - \varepsilon_k\left(1 - \frac{n}{2}\right) - \chi(i\omega_k) - \varepsilon_k\Psi(i\omega_k)}. \tag{36}$$

Comparing with the standard form of the Green's function in terms of the Dyson–Mori self energy

$$\mathcal{G}(k) = \frac{1 - \frac{n}{2}}{i\omega_k + \boldsymbol{\mu} - \epsilon_k\left(1 - \frac{n}{2}\right) - \Sigma_{DM}(k)}, \tag{37}$$

we see the momentum independence of the Dyson–Mori self energy $\Sigma_{DM}(k) = \Sigma_{DM}(i\omega_k)$, and

$$\Sigma_{DM}(i\omega_k) = \frac{(i\omega_k + \boldsymbol{\mu})\Psi(i\omega_k) + \left(1 - \frac{n}{2}\right)\chi(i\omega_k)}{1 - \frac{n}{2} + \Psi(i\omega_k)}. \tag{38}$$

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 λ 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 α is defined as follows:

$$J_j^\alpha = i\sum_{k\sigma} v_{k,j}^\alpha X_k^{\sigma 0} X_j^{0\sigma}; \quad v_{k,j}^\alpha = t_{k,j}(\vec{R}_k - \vec{R}_j)_\alpha, \tag{39}$$

so that v is a velocity. Using the notation $J^\alpha[i] = J_i^\alpha(\tau_i)$; $\widetilde{J}^\alpha[i] = J^\alpha[i] - \langle J^\alpha[i]\rangle$, we define the correlation function $\Pi_{\alpha\beta}[i,f]$ and its Fourier transform as

$$\begin{aligned} \Pi_{\alpha\beta}[i,f] &= \langle T_\tau \widetilde{J}^\alpha[i]\widetilde{J}^\beta[f]\rangle; \\ \Pi_{\alpha\beta}(\vec{q}, i\Omega_n) &= \int_0^\beta d(\tau_i - \tau_f)\, e^{i\Omega_n(\tau_i - \tau_f)} \sum_{i-f} e^{-i\vec{q}\cdot(\vec{R}_i - \vec{R}_f)} \Pi_{\alpha\beta}[i,f]. \end{aligned} \tag{40}$$

The optical conductivity can be given in terms of this object as

$$\sigma^{\alpha\beta}(\omega) = \frac{1}{i\omega - \eta}\left[\Pi_{\alpha\beta}(\vec{0}, \omega + i\eta) - \Pi_{\alpha\beta}(\vec{0}, i\eta)\right], \tag{41}$$

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

$$\mathcal{A} \to \mathcal{A} + \sum_{j\alpha} \int_0^\beta d\tau \kappa_j^\alpha(\tau) J_j^\alpha(\tau). \tag{42}$$

In terms of the κ 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

$$\Pi_{\alpha\beta}[i,f] = -i \,\mathrm{Tr}\left(\frac{\delta}{\delta \kappa_i^\alpha} \mathcal{G}[f, \mathbf{j}]\, v^\beta[\mathbf{j}, f^+] \right)_{\mathcal{A}\to 0}, \tag{43}$$

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 κ source derivative), finally obtaining an expression for $\Pi_{\alpha\beta}[i,f]$ in terms of the κ source derivatives of $\mathbf{g}^{-1}$ and μ.

$$\begin{aligned} \Pi_{\alpha\beta}[i,f] \;=\; & i \,\mathrm{Tr}\left(\mathbf{g}[f,\mathbf{x}] \frac{\delta}{\delta \kappa_i^\alpha} \mathbf{g}^{-1}[\mathbf{x},\mathbf{y}]\, \mathbf{g}[\mathbf{y},\mathbf{k}] \mu[\mathbf{k},\mathbf{j}]\, v^\beta[\mathbf{j}, f^+] \right)_{\mathcal{A}\to 0} \\ & - i \,\mathrm{Tr}\left(\mathbf{g}[f,\mathbf{k}] \frac{\delta}{\delta \kappa_i^\alpha} \mu[\mathbf{k},\mathbf{j}]\, v^\beta[\mathbf{j}, f^+] \right)_{\mathcal{A}\to 0}. \end{aligned} \tag{44}$$

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

$$t[i,f] \to t[i,f] - i \sum_\alpha \kappa_f^\alpha \, v^\alpha[i,f]. \tag{45}$$

Thus, the new equations of motion can be read off from Eq. (15) as

$$\begin{aligned} \mathbf{g}^{-1}[i,m] \;=\; & (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i)\, \delta[i,m] + \left(t[i,m] - i \sum_\alpha \kappa_m^\alpha \, v^\alpha[i,m] \right) \\ & \times (1 - \lambda\gamma[i]) + \lambda \left(t[i,\mathbf{j}] - i \sum_\alpha \kappa_{\mathbf{j}}^\alpha \, v^\alpha[i,\mathbf{j}] \right) \xi^* \cdot \mathbf{g}[\mathbf{j},\mathbf{n}] \cdot \Lambda_*[\mathbf{n}, m; i], \\ \mu[i,m] = (1 - \lambda\gamma[i])\delta[i,m] - \lambda & \left(t[i,\mathbf{j}] - i \sum_\alpha \kappa_{\mathbf{j}}^\alpha \, v^\alpha[i,\mathbf{j}] \right) \xi^* \cdot \mathbf{g}[\mathbf{j},\mathbf{n}] \cdot \mathcal{U}_*[\mathbf{n}, m; i]. \end{aligned} \tag{46}$$

Since there is no source derivative with respect to κ 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)):

$$\begin{aligned} \mathbf{g}^{-1}[i,m] \;=\; & (\boldsymbol{\mu} - \partial_{\tau_i} - \mathcal{V}_i)\, \delta[i,m] - \lambda\, \delta_{i,m}\, \chi_i[\tau_i,\tau_m] + \left(t[i,m] - i \sum_\alpha \kappa_m^\alpha \, v^\alpha[i,m] \right) \\ & \times (1 - \lambda\gamma[i]) + \lambda \left(t_{i,m} - i \sum_\alpha \kappa_m^\alpha \, v_{i,m}^\alpha \right) \Psi_i[\tau_i,\tau_m], \\ \mu[i,m] = \delta[i,m](1 - \lambda\gamma[i]) & + \lambda\, \delta_{i,m}\, \Psi_i[\tau_i,\tau_m], \end{aligned} \tag{47}$$

where χ_i, Ψ_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 χ_i, Ψ_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 χ_i, Ψ_i, and $\gamma[i]$. It remains to show that $\mathbf{g}_{x,x}[\tau_n, \tau_m]$ and $(t_{x,\mathbf{j}} - i\sum_\alpha \kappa^\alpha_{\mathbf{j}}(\tau_n) v^\alpha_{x,\mathbf{j}})\, \mathbf{g}_{\mathbf{j},x}[\tau_n, \tau_m]$ (the time indices are arbitrary) satisfy this equation.

Defining the notation $w_{i,f}(\tau_i) \equiv t_{i,f} - i\sum_\alpha \kappa^\alpha_f(\tau_i)\, v^\alpha_{i,f}$, and using (the κ source derivative version of) Eq. (17) as well as Eq. (47), we find that

$$\begin{aligned}\left(\frac{\delta}{\delta\kappa^\alpha_i} w_{x,\mathbf{j}}(\tau_n)\, \mathbf{g}_{\mathbf{j},x}[\tau_n, \tau_m]\right)_{\mathcal{A}\to 0} &= -i\delta[\tau_i, \tau_n] v^\alpha_{x,i}\, \mathbf{g}_{i,x}[\tau_i, \tau_m] + i t_{x,\mathbf{j}} \mathbf{g}_{\mathbf{j},\mathbf{a}}[\tau_n, \tau_{\mathbf{a}}] \\ &\quad \times (1 - \lambda\gamma[\mathbf{a}]\delta[\tau_{\mathbf{a}}, \tau_i] + \lambda\Psi_{\mathbf{a}}[\tau_{\mathbf{a}}, \tau_i]) v^\alpha_{\mathbf{a},i} g_{i,x}[\tau_i, \tau_m] \\ &\quad + \lambda\, t_{x,\mathbf{j}} \mathbf{g}_{\mathbf{j},\mathbf{a}}[\tau_n, \tau_{\mathbf{a}}] \frac{\delta}{\delta\kappa^\alpha_i} (\gamma[\mathbf{a}]\delta[\tau_{\mathbf{a}}, \tau_{\mathbf{b}}] - \Psi_{\mathbf{a}}[\tau_{\mathbf{a}}, \tau_{\mathbf{b}}]) \\ &\quad \times t_{\mathbf{a},\mathbf{b}} g_{\mathbf{b},x}[\tau_{\mathbf{b}}, \tau_m] + \lambda\, t_{x,\mathbf{j}} \mathbf{g}_{\mathbf{j},\mathbf{a}}[\tau_n, \tau_{\mathbf{a}}] \\ &\quad \times \frac{\delta}{\delta\kappa^\alpha_i} (\chi_{\mathbf{a}}[\tau_{\mathbf{a}}, \tau_{\mathbf{b}}])\, g_{\mathbf{a},x}[\tau_{\mathbf{b}}, \tau_m], \end{aligned} \tag{48}$$

where the RHS is also evaluated in the $\mathcal{A} \to 0$ 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 $v^\alpha_{i,i} = 0$ while the latter must vanish due to the sum over $i - f$ and the odd parity of $v^\alpha_{i,f}$. The same reasoning applies to the second term except that in this term the $\mathbf{x} = i$ case vanishes by the odd parity of $v^\alpha_{i,f}$. Hence, we have shown that $(t_{x,\mathbf{j}} - i\sum_\alpha \kappa^\alpha_{\mathbf{j}}(\tau_n) v^\alpha_{x,\mathbf{j}})\, \mathbf{g}_{\mathbf{j},x}[\tau_n, \tau_m]$ satisfies Eq. (21) with $s = 0$. A completely analogous argument shows that this is also the case for $\mathbf{g}_{x,x}[\tau_n, \tau_m]$. 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

$$\sum_{i-f} \Pi_{\alpha\beta}[i, f] = \sum_{i-f} \mathrm{Tr}\left(\mathcal{G}[f, \mathbf{k}] v^\alpha[\mathbf{k}, i] \mathcal{G}[i, \mathbf{j}]\, v^\beta[\mathbf{j}, f^+]\right)_{\mathcal{A}\to 0}. \tag{49}$$

Substituting this equation into Eq. (41), the optical conductivity may be expressed as

$$\sigma^{\alpha\beta}(\omega) = \frac{2}{i\omega} \sum_{\vec{p}, i\omega_p} \mathcal{G}(\vec{p}, i\omega_p) v^\alpha_{\vec{p}} v^\beta_{\vec{p}} [\mathcal{G}(\vec{p}, \omega + i\eta + i\omega_p) - \mathcal{G}(\vec{p}, i\eta + i\omega_p)]. \tag{50}$$

We now want to prove that this result holds to each order in λ. We do this via an inductive argument, in which we assume that through nth order in λ, $\left(\frac{\delta}{\delta\kappa^\alpha_i} L_x\right)^{(n)}_{\mathcal{A}\to 0}$ (where L_i can be Ψ_i, χ_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 $\left(\frac{\delta}{\delta\kappa^\alpha_i} L_x\right)^{(n)}_{\mathcal{A}\to 0}$ to simplify the resulting expressions, thereby proving Eqs. (49) and (50) to each order in λ.

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 nth order in λ, the κ source derivative of the objects Ψ_i, χ_i, and $\gamma[i]$ (denoted below by the generic symbol L_i) can be written as

$$\begin{aligned}\left(\frac{\delta}{\delta\kappa^\alpha_i} L_x\right)^{(n)}_{\mathcal{A}\to 0} &= \lambda^n X_x Y_{x,\mathbf{x}_1} X_{\mathbf{x}_1} Y_{\mathbf{x}_1,\mathbf{x}_2} X_{\mathbf{x}_2} \ldots X_{\mathbf{x}_{m-1}} Y_{\mathbf{x}_{m-1},\mathbf{x}_m} X_{\mathbf{x}_m} \\ &\quad \times v^\alpha_{\mathbf{x}_m,i}\, Y_{i,\mathbf{x}_{m-1}} X_{\mathbf{x}_{m-1}} \ldots X_{\mathbf{x}_1} Y_{\mathbf{x}_1,x} X_x, \end{aligned} \tag{51}$$

where the number m is arbitrary. In the base case of zeroth order, the objects Ψ_i, χ_i, and $\gamma[i]$ are

$$\begin{aligned}
&\Psi_i^{(0)}[\tau_i, \tau_m] = 0; \qquad \gamma^{(0)}[i] = \mathbf{g}^{(k)}[i, i];\\
&\chi_i^{(0)}[\tau_i, \tau_m] = -\left(t_{i,\mathbf{j}} - i\sum_\alpha \kappa_{\mathbf{j}}^\alpha(\tau_i) v_{\mathbf{ij}}^\alpha\right) \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_i]\delta[\tau_i, \tau_m].
\end{aligned} \tag{52}$$

We note that $\left(\frac{\delta}{\delta\kappa_i^\alpha} w_{x,\mathbf{j}}(\tau_n)\, \mathbf{g}_{\mathbf{j},x}[\tau_n, \tau_m]\right)^{(l)}_{\mathscr{A}\to 0}$ is given by Eq. (48) with the appropriate objects on the RHS evaluated to the appropriate order in λ. An analogous formula holds for $\left(\frac{\delta}{\delta\kappa_i^\alpha}\mathbf{g}_{x,x}[\tau_n, \tau_m]\right)^{(l)}_{\mathscr{A}\to 0}$. 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 $t_{i,\mathbf{j}} \to w_{i,\mathbf{j}}(\tau_n)$ (the time index is again arbitrary). Therefore, using the notation $\widetilde{R}_i = [R_i]_{t_{i,\mathbf{j}}\to w_{i,\mathbf{j}}(\tau_n)}$, we may write

$$\left(\frac{\delta}{\delta\kappa_i^\alpha} L_x\right)^{(n+1)}_{\mathscr{A}\to 0} = \sum_{r=0}^{n+1} \lambda^r \left(\frac{\delta}{\delta\kappa_i^\alpha}(\widetilde{R}_x)^m\right)^{(n+1-r)}_{\mathscr{A}\to 0}. \tag{53}$$

Substituting the formulas for $\left(\frac{\delta}{\delta\kappa_i^\alpha} w_{x,\mathbf{j}}(\tau_n)\, \mathbf{g}_{\mathbf{j},x}[\tau_n, \tau_m]\right)^{(l)}_{\mathscr{A}\to 0}$ and $\left(\frac{\delta}{\delta\kappa_i^\alpha}\mathbf{g}_{x,x}[\tau_n, \tau_m]\right)^{(l)}_{\mathscr{A}\to 0}$ (Eq. (48)) for $l \le n+1$ into Eq. (53), and using the inductive hypothesis, shows that $\left(\frac{\delta}{\delta\kappa_i^\alpha}\Psi_x\right)^{(n+1)}_{\mathscr{A}\to 0}$, $\left(\frac{\delta}{\delta\kappa_i^\alpha}\chi_x\right)^{(n+1)}_{\mathscr{A}\to 0}$, and $\left(\frac{\delta}{\delta\kappa_i^\alpha}\gamma[x]\right)^{(n+1)}_{\mathscr{A}\to 0}$ all have the desired form (Eq. (51)). Thus, Eq. (51) holds to all orders in λ.

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 $(t_{\mathbf{x},\mathbf{y}} - i\sum_\alpha \kappa_{\mathbf{y}}^\alpha\, v_{\mathbf{x},\mathbf{y}}^\alpha)$ 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 $\mathbf{x}_1 \ldots \mathbf{x}_{m-1}$ 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 λ.

3.4. $O(\lambda^2)$ theory in the limit of large dimensions

To obtain self-consistent integral equations to any order in λ for the objects $\mathbf{g}^{-1}[i, f]$ and $\mu[i, f]$, we expand Eqs. (22) through (25) iteratively in λ, 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

$$\mathbf{g}_{\mathrm{loc},m}(i\omega_k) \equiv \sum_{\vec{k}} \mathbf{g}(k)\epsilon_{\vec{k}}^m, \tag{54}$$

$$I_{m_1m_2m_3}(i\omega_k) \equiv -\sum_{\omega_p,\omega_q} \mathbf{g}_{\mathrm{loc},m_1}(i\omega_q)\mathbf{g}_{\mathrm{loc},m_2}(i\omega_p)\mathbf{g}_{\mathrm{loc},m_3}(i\omega_q + i\omega_p - i\omega_k), \tag{55}$$

the resulting equations to $O(\lambda^2)$ are

$$a_{\mathscr{g}} \equiv 1 - \lambda\frac{n}{2} + \lambda^2\frac{n^2}{4}, \tag{56}$$

$$\mathbf{g}^{-1}(k) = i\omega_k + \boldsymbol{\mu}' - a_{\mathscr{g}}\left(\varepsilon_k - \frac{u_0}{2}\right) - \lambda\left(\epsilon_{\vec{k}} - \frac{u_0}{2}\right)\Psi(i\omega_k) - \lambda\chi(i\omega_k), \tag{57}$$

$$\mu(i\omega_k) = a_{\mathscr{g}} + \lambda\Psi(i\omega_k), \tag{58}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} - u_0\left(\lambda\frac{n}{2} - \lambda^2\frac{n^2}{8}\right) + \lambda\sum_p \varepsilon_p \mathbf{g}(p) - a_{\mathcal{G}}\frac{u_0}{2}, \tag{59}$$

$$\Psi(i\omega_k) = -\lambda u_0 I_{000}(i\omega_k) + 2\lambda I_{010}(i\omega_k), \tag{60}$$

$$\chi(i\omega_k) = -\frac{u_0}{2}\Psi(i\omega_k) - u_0\lambda I_{001}(i\omega_k) + 2\lambda I_{011}(i\omega_k). \tag{61}$$

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:

$$\sum_k \mathbf{g}(k) = \frac{n}{2}; \qquad \sum_k \mathcal{G}(k) = \frac{n}{2}. \tag{62}$$

The objects $\mathbf{g}_{\mathrm{loc},m}(i\omega_k)$ 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 ϵ (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 $\mathcal{G}(\{\tau_j\})$, $\mathbf{g}(\{\tau_j\})$, $\mu(\{\tau_j\})$ or $\mathcal{G}(\{i\omega_j\})$, $\mathbf{g}(\{i\omega_j\})$, $\mu(\{i\omega_j\})$ 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 μ 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 $U \to \infty$ which has the following Hamiltonian:

$$H = \sum_\sigma \epsilon_d X^{\sigma\sigma} + \sum_{k\sigma} \widetilde{\epsilon}_k n_{k\sigma} + \sum_{k\sigma}(V_k\, X^{\sigma 0}\, c_{k\sigma} + V_k^*\, c_{k\sigma}^\dagger\, X^{0\sigma}), \tag{63}$$

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:

$$\mathcal{G}_{\sigma_i\sigma_f}[\tau_i, \tau_f] = -\langle\!\langle\, X^{0\sigma_i}(\tau_i)\, X^{\sigma_f 0}(\tau_f)\rangle\!\rangle. \tag{64}$$

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:

$$\mathcal{G}[\tau_i, \tau_f] = \mathbf{g}[\tau_i, \tau_{\mathbf{j}}] \cdot \mu[\tau_{\mathbf{j}}, \tau_f]. \tag{65}$$

The equations of motion for $\mathbf{g}$ and μ can be written as

$$\begin{aligned}(\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i))\mathbf{g}[\tau_i, \tau_f] \;=\; & -\delta(\tau_i - \tau_f) - (1 - \lambda\gamma[\tau_i]) \cdot \Delta[\tau_i, \tau_{\mathbf{j}}] \cdot \mathbf{g}[\tau_{\mathbf{j}}, \tau_f] \\ & - \lambda\, \xi^* \Delta[\tau_i, \tau_{\mathbf{j}}] \cdot \mathbf{g}[\tau_{\mathbf{j}}, \tau_{\mathbf{x}}] \cdot \Lambda_*[\tau_{\mathbf{x}}, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}[\tau_{\mathbf{y}}, \tau_f],\end{aligned} \tag{66}$$

$$\mu[\tau_i, \tau_f] = \delta(\tau_i - \tau_f)(\mathbb{1} - \lambda\gamma[\tau_i]) + \lambda\, \xi^* \cdot \Delta[\tau_i, \tau_{\mathbf{j}}] \cdot \mathbf{g}[\tau_{\mathbf{j}}, \tau_{\mathbf{x}}] \cdot \mathcal{U}_*[\tau_{\mathbf{x}}, \tau_f; \tau_i], \tag{67}$$

where the conduction band enters through the ($\mathcal{V}$ independent) function

$$\Delta[\tau_i, \tau_f] = -\mathbb{1} \sum_k |V_k|^2 (\partial_{\tau_i} + \widetilde{\epsilon}_k)^{-1} \delta(\tau_i - \tau_f). \tag{68}$$

We have also made use of the following definitions:

$$\begin{aligned} &\Lambda[\tau_n, \tau_m; \tau_i] = -\frac{\delta}{\delta \mathcal{V}(\tau_i)} \mathbf{g}^{-1}[\tau_n, \tau_m]; \qquad \mathcal{U}[\tau_n, \tau_m; \tau_i] = \frac{\delta}{\delta \mathcal{V}(\tau_i)} \mu[\tau_n, \tau_m]; \\ &\gamma[\tau_i] = \mu^{(k)}[\tau_{\mathbf{n}}, \tau_i^+] \cdot \mathbf{g}^{(k)}[\tau_i, \tau_{\mathbf{n}}]. \end{aligned} \tag{69}$$

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

$$\begin{aligned} (\partial_{\tau_i} - \mu + \mathcal{V}_i(\tau_i))\, \mathbf{g}_{i,i}[\tau_i, \tau_f] \;=\; & -\delta(\tau_i - \tau_f) + (1 - \lambda\gamma[i]) \cdot t_{i,\mathbf{j}}\, \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_f] \\ & + \lambda\, t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{x}}] \cdot A_{i,*}[\tau_{\mathbf{x}}, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_{\mathbf{y}}, \tau_f] \\ & + \lambda\, t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{x}}] \cdot B_{i,*}[\tau_{\mathbf{x}}, \tau_{\mathbf{y}}; \tau_i] \cdot t_{i,\mathbf{y}}\, \mathbf{g}_{\mathbf{y},i}[\tau_{\mathbf{y}}, \tau_f], \end{aligned} \tag{70}$$

$$\mu_{i,i}[\tau_i, \tau_f] = (1 - \lambda\gamma[i])\delta(\tau_i - \tau_f) + \lambda\, t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{x}}] \cdot B_{i,*}[\tau_{\mathbf{x}}, \tau_f; \tau_i]. \tag{71}$$

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 $\mathbf{g}^{-1}_{\mathrm{loc},i}[\tau_i, \tau_f]$ and define it to be the temporal inverse of the local auxiliary Green's function:

$$\mathbf{g}_{i,i}[\tau_i, \tau_{\mathbf{j}}] \cdot \mathbf{g}^{-1}_{\mathrm{loc},i}[\tau_{\mathbf{j}}, \tau_f] = \delta(\tau_i - \tau_f). \tag{72}$$

Note that $\mathbf{g}^{-1}_{\mathrm{loc},i}[\tau_i, \tau_f] \neq \mathbf{g}^{-1}_{i,i}[\tau_i, \tau_f]$. We also define the corresponding vertex:

$$\Lambda_{\mathrm{loc},i}[\tau_n, \tau_m; \tau_i] = -\frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}^{-1}_{\mathrm{loc},i}[\tau_n, \tau_m]. \tag{73}$$

We now make use of the following identity:

$$\Lambda_{\mathrm{loc},i}[\tau_x, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_{\mathbf{y}}, \tau_f] = A_i[\tau_x, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_{\mathbf{y}}, \tau_f] + B_i[\tau_x, \tau_{\mathbf{y}}; \tau_i] \cdot t_{i,\mathbf{y}}\, \mathbf{g}_{\mathbf{y},i}[\tau_{\mathbf{y}}, \tau_f]. \tag{74}$$

This identity is easily proven by considering $\frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{i,i}[\tau_x, \tau_f]$:

$$\frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{i,i}[\tau_x, \tau_f] = \mathbf{g}_{i,i}[\tau_x, \tau_{\mathbf{j}}] \Lambda_{\mathrm{loc},i}[\tau_{\mathbf{j}}, \tau_{\mathbf{y}}; \tau_i] \mathbf{g}_{i,i}[\tau_{\mathbf{y}}, \tau_f]. \tag{75}$$

The LHS can also be expressed as

$$\begin{aligned} \frac{\delta}{\delta \mathcal{V}_i(\tau_i)} \mathbf{g}_{i,i}[\tau_x, \tau_f] \;=\; & \mathbf{g}_{i,i}[\tau_x, \tau_{\mathbf{j}}] \cdot (A_i[\tau_{\mathbf{j}}, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_{\mathbf{y}}, \tau_f] \\ & + B_i[\tau_{\mathbf{j}}, \tau_{\mathbf{y}}; \tau_i] \cdot t_{i,\mathbf{y}}\, \mathbf{g}_{\mathbf{y},i}[\tau_{\mathbf{y}}, \tau_f]). \end{aligned} \tag{76}$$

Left multiplying the above 2 equations by $\mathbf{g}^{-1}_{\mathrm{loc},i}$, we recover the identity Eq. (74). Substituting this identity into Eq. (70), we obtain

$$\begin{aligned} (\partial_{\tau_i} - \mu + \mathcal{V}_i(\tau_i))\mathbf{g}_{i,i}[\tau_i, \tau_f] \;=\; & -\delta(\tau_i - \tau_f) + (1 - \lambda\gamma[i]) \cdot t_{i,\mathbf{j}} \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_f] \\ & + \lambda\, t_{i,\mathbf{j}}\, \xi^* \cdot \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_{\mathbf{x}}] \cdot \Lambda_{\mathrm{loc},i*}[\tau_{\mathbf{x}}, \tau_{\mathbf{y}}; \tau_i] \cdot \mathbf{g}_{i,i}[\tau_{\mathbf{y}}, \tau_f]. \end{aligned} \tag{77}$$

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 μ to $-\epsilon_d$. The following mappings also follow as a consequence of these.

$$\begin{aligned} &\gamma[i] \to \gamma[\tau_i]; \qquad \Lambda_{\mathrm{loc},i}[\tau_n, \tau_m; \tau_i] \to \Lambda[\tau_n, \tau_m; \tau_i]; \\ &B_i[\tau_n, \tau_m; \tau_i] \to -\mathcal{U}[\tau_n, \tau_m; \tau_i]. \end{aligned} \tag{78}$$

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:

$$t_{i,\mathbf{j}}\, \mathbf{g}_{\mathbf{j},i}[\tau_i, \tau_f] = -\Delta[\tau_i, \tau_\mathbf{j}] \cdot \mathbf{g}[\tau_\mathbf{j}, \tau_f]. \tag{79}$$

4.3. Mapping to each order in λ

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 λ as well, and it is therefore interesting to know if the mapping from the previous section holds to each order in λ. 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 λ expansion for the Anderson model from Ref. [22]. There, Eqs. (66) and (67) are written as

$$\begin{aligned} \mathbf{g}^{-1}[\tau_i, \tau_f] \;=\; & -(\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i))\delta(\tau_i - \tau_f) - (1 - \lambda\gamma[\tau_i]) \cdot \Delta[\tau_i, \tau_f] \\ & - \lambda\xi^* \Delta[\tau_i, \tau_\mathbf{j}] \cdot \mathbf{g}[\tau_\mathbf{j}, \tau_\mathbf{x}] \cdot \Lambda_*[\tau_\mathbf{x}, \tau_f; \tau_i], \end{aligned} \tag{80}$$

$$\mu[\tau_i, \tau_f] = \delta(\tau_i - \tau_f)(\mathbb{1} - \lambda\gamma[\tau_i]) + \lambda\xi^* \cdot \Delta[\tau_i, \tau_\mathbf{j}] \cdot \mathbf{g}[\tau_\mathbf{j}, \tau_\mathbf{x}] \cdot \mathcal{U}_*[\tau_\mathbf{x}, \tau_f; \tau_i]. \tag{81}$$

The λ expansion is obtained in the same way as for the t–J model, by iterating the equations in $\mathbf{g}^{-1}$ and μ and keeping track of explicit powers of λ. The details to $O(\lambda^2)$ can be found in Ref. [22]. To relate this to the λ expansion for the infinite-dimensional t–J model, recall from Eq. (28) that to each order in λ, Ψ_i, χ_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 $t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j},i}[\tau_n, \tau_m]$. We can now state our *inductive hypothesis*: through nth order in λ, the λ expansion for the Anderson model has the form

$$\begin{aligned} \mathbf{g}^{-1}[\tau_i, \tau_m] \;=\; & -(\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i))\, \delta[\tau_i, \tau_m] - \lambda\, \chi[\tau_i, \tau_m] \\ & - (1 - \lambda\gamma[\tau_i])\Delta[\tau_i, \tau_m] - \lambda\, \Psi[\tau_i, \tau_\mathbf{j}]\Delta[\tau_\mathbf{j}, \tau_m], \\ \mu[\tau_i, \tau_m] = \;& \delta[\tau_i, \tau_m](1 - \lambda\gamma[\tau_i]) + \lambda\Psi[\tau_i, \tau_m], \\ \Lambda[\tau_n, \tau_m; \tau_i] = \;& A[\tau_n, \tau_m; \tau_i] - B[\tau_n, \tau_\mathbf{j}; \tau_i]\Delta[\tau_\mathbf{j}, \tau_m], \\ \mathcal{U}[\tau_n, \tau_m; \tau_i] = \;& -B[\tau_n, \tau_m; \tau_i], \end{aligned} \tag{82}$$

where through nth order in λ, the objects $A[\tau_n, \tau_m; \tau_i]$ and $B[\tau_n, \tau_m; \tau_i]$, and through $n-1$st order in λ, 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

$$\mathbf{g}_{i,i}[\tau_n, \tau_m] \to \mathbf{g}[\tau_n, \tau_m]; \qquad \boldsymbol{\mu} \to -\epsilon_d; \qquad t_{i,\mathbf{j}}\mathbf{g}_{\mathbf{j}i}[\tau_n, \tau_m] \to -\Delta[\tau_n, \tau_\mathbf{j}] \cdot \mathbf{g}[\tau_\mathbf{j}, \tau_m]. \tag{83}$$

We first examine the base case of zeroth order:

$$A^{(0)}[\tau_n, \tau_m; \tau_i] = \delta[\tau_i, \tau_n]\delta[\tau_i, \tau_m]; \qquad B^{(0)}[\tau_n, \tau_m; \tau_i] = 0. \tag{84}$$

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:

$$\begin{aligned} \chi^{(n)}[\tau_n, \tau_m] &= \xi^* \Delta[\tau_n, \tau_\mathbf{j}] \cdot \mathbf{g}[\tau_\mathbf{j}, \tau_\mathbf{x}] \cdot A^{(n)}_*[\tau_\mathbf{x}, \tau_m; \tau_n], \\ \Psi^{(n)}[\tau_n, \tau_m] &= -\xi^* \Delta[\tau_n, \tau_\mathbf{j}] \cdot \mathbf{g}[\tau_\mathbf{j}, \tau_\mathbf{x}] \cdot B^{(n)}_*[\tau_\mathbf{x}, \tau_m; \tau_n], \\ A^{(n+1)}[\tau_n, \tau_m; \tau_i] &= \lambda \left(\frac{\delta}{\delta \mathcal{V}(\tau_i)} \chi[\tau_n, \tau_m] \right)^{(n)}, \end{aligned}$$

$$B^{(n+1)}[\tau_n, \tau_m; \tau_i] = \lambda\, \delta[\tau_n, \tau_m] \left(\frac{\delta}{\delta \mathcal{V}(\tau_i)} \gamma[\tau_n]\right)^{(n)} - \lambda \left(\frac{\delta}{\delta \mathcal{V}(\tau_i)} \Psi[\tau_n, \tau_m]\right)^{(n)},$$

$$\gamma^{(n)}[\tau_i] = -\lambda\, \gamma^{(k)(n-1)}[\tau_i] \mathbf{g}^{(k)}[\tau_i, \tau_i] + \lambda\, \Psi^{(k)(n-1)}[\tau_{\mathbf{j}}, \tau_i] \mathbf{g}^{(k)}[\tau_i, \tau_{\mathbf{j}}]. \tag{85}$$

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

$$\left(\frac{\delta}{\delta \mathcal{V}(\tau_r)} \mathbf{g}[\tau_i, \tau_m]\right)^{(l)} = \mathbf{g}[\tau_i, \tau_{\mathbf{x}}](A^{(l)}[\tau_{\mathbf{x}}, \tau_{\mathbf{y}}; \tau_r] - B^{(l)}[\tau_{\mathbf{x}}, \tau_{\mathbf{j}}; \tau_r] \Delta[\tau_{\mathbf{j}}, \tau_{\mathbf{y}}]) \mathbf{g}[\tau_{\mathbf{y}}, \tau_m]. \tag{86}$$

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 ($\lambda \to 1$, $\gamma[\tau_i] \to \frac{n_d}{2} \equiv \frac{n}{2}$)

$$\mathbf{g}^{-1}(i\omega_k) = i\omega_k - \epsilon_d - \left(1 - \frac{n}{2}\right) \Delta(i\omega_k) - \chi(i\omega_k) - \Delta(i\omega_k) \Psi(i\omega_k),$$

$$\mu(i\omega_k) = 1 - \frac{n}{2} + \Psi(i\omega_k). \tag{87}$$

Comparing with Eq. (35), it immediately follows that under the mapping Eq. (83), $\mu_{i,i}(i\omega_k) \to \mu(i\omega_k)$. 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 $\mathbf{g}_{i,i}(i\omega_k) \to \mathbf{g}(i\omega_k)$. Therefore, the ECFL solution of the infinite dimensional t–J model maps onto the ECFL solution of the AIM to each order in λ as long as the following self-consistency condition is satisfied:

$$\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) = \sum_{\vec{k}} \frac{|V_{\vec{k}}|^2}{i\omega_n - \widetilde{\epsilon}_{\vec{k}}} \mathbf{g}(i\omega_k). \tag{88}$$

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 $\widetilde{\varepsilon}_k$ 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 caparison 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 caparison factor of the Anderson model $\mu(i\omega_k)$ set equal to $\mu_{i,i}(i\omega_k)$ as long as $\widetilde{\varepsilon}_k$ 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

$$\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) = \frac{-1}{1 - \frac{n}{2} + \Psi(i\omega_k)} [1 - (i\omega_k + \boldsymbol{\mu} - \chi(i\omega_k)) \mathbf{g}(i\omega_k)]. \tag{89}$$

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

$$\Sigma_D(i\omega_k) + \frac{1}{\mathcal{G}(i\omega_k)} - (i\omega_k + \boldsymbol{\mu}) = -\sum_{\vec{k}} \epsilon_{\vec{k}} \mathbf{g}(k) \frac{1}{\mathbf{g}(i\omega_k)}. \tag{90}$$

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 λ 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 $U \to \infty$ and $d \to \infty$, 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 λ, with $\Psi(i\omega_k)$ and $\chi(i\omega_k)$ written as a product of the functions $\mathbf{g}_{\mathrm{loc},m}(i\omega_k)$ (Eq. (54)) with $m = 0$ or $m = 1$. This continues to hold to each order in λ. We have analyzed the optical conductivity and have shown that it is given by Eq. (50) in general and to each order in λ. 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 λ 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 λ 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.

Extremely correlated Fermi liquid theory meets dynamical mean-field theory: Analytical insights into the doping-driven Mott transition

R. Žitko,[1,2] D. Hansen,[3] E. Perepelitsky,[3] J. Mravlje,[1] A. Georges,[4,5,6] and B. S. Shastry[3]

[1]*Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

[2]*Faculty for Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia*

[3]*Physics Department, University of California Santa Cruz, California 95064, USA*

[4]*Centre de Physique Théorique, École Polytechnique, CNRS, 91128 Palaiseau Cedex, France*

[5]*Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*

[6]*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 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.[2–5] 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 λ 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[6,7] 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[8]). 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 δ [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 $\mathrm{Im}\Sigma \propto \omega^2 + (\pi T)^2$ 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, Ψ and χ, 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

 ©2013 American Physical Society

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 Ψ and χ.

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,[8] giving rise to an expansion to order λ^2 in the projection parameter λ. 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 Ψ and χ via an expansion in the projection parameter λ. 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

$$H = \sum_{k\sigma} \varepsilon_k c^\dagger_{k\sigma} c_{k\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \tag{1}$$

where ε_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:

$$\rho_0(\varepsilon) = \frac{2}{\pi D^2}\sqrt{D^2 - \varepsilon^2}, \tag{2}$$

where D is the half bandwidth, and thus any sum over the band energy can be converted to an integral as:

$$\frac{1}{N}\sum_k \mathcal{A}(\varepsilon_k) \to \int_{-D}^{D} d\varepsilon\ \rho_0(\varepsilon)\mathcal{A}(\varepsilon). \tag{3}$$

We note that the Fermi energy ε_F satisfies

$$\sin^{-1}\left(\frac{\varepsilon_F}{D}\right) + \left(\frac{\varepsilon_F}{D}\right)\sqrt{1 - \left(\frac{\varepsilon_F}{D}\right)^2} = -\frac{\pi}{2}(1-n), \tag{4}$$

and vanishes near $n \sim 1$ as $\varepsilon_F = -\frac{\pi}{4}(1-n)\,D$. The hole doping level δ is related to the particle density n ($n = N/N_{\text{sites}}$) as:

$$\delta = 1 - n. \tag{5}$$

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):

$$\rho_Q(\omega) = -\frac{1}{\pi}\mathrm{Im}[Q(\omega + i0^+)]. \tag{6}$$

III. ECFL: GENERAL FRAMEWORK

A. ECFL formalism

The ECFL methodology has been discussed extensively in recent literature;[2,3] here we highlight only the aspects that are of relevance to this work. ECFL deals with Gutzwiller-projected states obtained in the limit of $U \to \infty$, 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[8] by analyzing the infinite-dimensional limit of the Schwinger equations of motion in the ECFL.[11]

The exact mapping of the momentum-independent self-energy of the infinite-dimensional Hubbard model onto that

of a self-consistent Anderson impurity model[12] provides a roadmap for a suitable formulation of the ECFL equations in this limit.

In the simplest version of the ECFL theory,[13] 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 $\widetilde{\mu}(k)$. Thus,

$$\mathcal{G}(k) = \widetilde{\mu}(k) \times \mathbf{g}(k), \tag{7}$$

where $k \equiv (\vec{k}, i\omega_k)$ and ω_k is a fermionic Matsubara frequency. Here $\mathbf{g}(k)$ is a Fermi liquidlike Green's function

$$\mathbf{g}^{-1}(k) = i\omega_k + \mu - \left(1 - \frac{n}{2}\right)\varepsilon_k - \Phi(k), \tag{8}$$

and μ is the chemical potential. The factor $\widetilde{\mu}(k)$ (here distinguished from the chemical potential μ by the tilde), plays the role of an adaptive spectral weight, and is given by

$$\widetilde{\mu}(k) = 1 - \frac{n}{2} + \Psi(k), \tag{9}$$

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 $\widetilde{\mu}$ respectively.[2,3] The term $\widetilde{\mu}(k)$ 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 Φ. 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:[8] Ψ is independent of $\vec{k}$ and Φ is decomposable into two $\vec{k}$-independent functions,

$$\Psi(k) = \Psi(i\omega_k), \tag{10}$$

$$\Phi(k) = \chi(i\omega_k) + \varepsilon_k \Psi(i\omega_k), \tag{11}$$

i.e., the two frequency-dependent (but $\vec{k}$-independent) functions χ and Ψ determine the Green's function.

The single-electron physical (Dyson) self-energy Σ is defined from the single-electron Green's function $\mathcal{G}$ in the usual manner, as (using the analytic continuation $i\omega_k \to \omega + i\eta$, $\eta = 0^+$):

$$\mathcal{G}(k, \omega + i\eta) = \frac{1}{\omega + i\eta + \mu - \varepsilon_k - \Sigma(\omega + i\eta)}. \tag{12}$$

Within the large-dimensional ECFL, the Dyson self-energy Σ can be related to Ψ and χ as follows:

$$\Sigma(\omega + i\eta) - \mu - \omega = \frac{\chi(\omega + i\eta) - \mu - \omega}{1 - \frac{n}{2} + \Psi(\omega + i\eta)}. \tag{13}$$

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 ω as $\omega \to \infty$. This is a consequence[14] of the Gutzwiller projection in the $U \to \infty$ limit. At finite U, this behavior is regularized at high-enough frequency and Σ 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.[3] The theory to $O(\lambda^2)$ is expected to be quantitatively accurate for densities up to $n \lesssim 0.7$.[15] We postpone the description of these equations to Sec. VI, but note an important general insight gained from examining and evaluating such an expansion[2,3,15]; the two self-energies χ and Ψ have simple Fermi liquid functional forms, with a dissipative part that is quadratic in ω, 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 $\mathrm{Im}\chi$ and $\mathrm{Im}\Psi$ is limited to the particle-hole symmetric lowest-order Fermi liquid terms.

Indeed, as mentioned above, the first few terms of a systematic λ expansion of the ECFL equations indicate that the self-energies Ψ and χ are very similar functions and resemble the self-energy of a Fermi liquid at low enough T, ω, with suitable scale constants.

For low ω and low T, up to a low-frequency cutoff scale Ω_c, so that $|\omega| \leqslant \Omega_c \ll D$, we define (with $k_B = 1$)

$$\mathcal{R}(\omega, T) \equiv \pi[\omega^2 + (\pi T)^2], \tag{14}$$

and write a Fermi liquidlike expansion for the complex ECFL self-energies:

$$\Psi(\omega) \sim \Psi_0 + c_\Psi\, \omega + \frac{i}{\gamma_\Psi}\, \mathcal{R}(\omega, T) + \Psi_{\mathrm{rem}}(\omega), \tag{15}$$

$$\chi(\omega) \sim \chi_0 - c_\chi\, \omega - \frac{i}{\Omega_\chi}\, \mathcal{R}(\omega, T) + \chi_{\mathrm{rem}}(\omega), \tag{16}$$

where

$$c_\Psi = 2\frac{\Omega_c}{\gamma_\Psi} \quad \text{and} \quad c_\chi = 2\frac{\Omega_c}{\Omega_\chi}. \tag{17}$$

Ω_χ and γ_Ψ are parameters that determine the curvatures of the two imaginary parts, with Ω_χ having the dimensions of energy, while γ_Ψ has the dimensions of the square of energy. Consequently, c_Ψ has the dimensions of an inverse energy, while c_χ is dimensionless. The terms $\Psi_{\mathrm{rem}}(\omega)$ and $\chi_{\mathrm{rem}}(\omega)$ 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_Ψ and c_χ 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_Ψ and c_χ 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 ω 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* ω. 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, Ω_χ and γ_Ψ turn out to be similar functions of the electron density, in view of their parallel role in the two self-energies within the λ expansion. In the analysis below, we will find that as $n \to 1$, 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 Ω_c, which determines the range of frequencies where the quadratic behavior of Im Ψ and Im χ applies, is itself a function of the density. It shrinks linearly with δ as $n \to 1$, and therefore c_Ψ and c_χ are finite as $n \to 1$. We should note that these are leading-order assumptions in δ.

Thus we find at low T,ω:

$$\mathcal{G}(k,\omega+i\eta) \sim \frac{\alpha_0 + c_\Psi\,\omega + \frac{i}{\gamma_\Psi}\mathcal{R}}{\omega(1+c_\chi) + \mu + \frac{i}{\Omega_\chi}\mathcal{R} - \varepsilon_k\{\alpha_0 + c_\Psi\,\omega + \frac{i}{\gamma_\Psi}\mathcal{R}\}}, \tag{18}$$

where we have introduced

$$\alpha_0 \equiv 1 - \frac{n}{2} + \Psi_0, \tag{19}$$

and χ_0 has been absorbed into the chemical potential μ. The entire momentum dependence is contained in ε_k. At $T = 0$ and $\omega = 0$ we must require $\mathcal{G}^{-1}(k_F,0) = 0$, so we need to set

$$\mu = \alpha_0\,\varepsilon_F. \tag{20}$$

At low $\omega + i\eta$ and a fixed $\vec{k}$, we can write a useful expression

$$\mathcal{G}^{-1} \sim -\varepsilon_k + \frac{\omega(1+c_\chi) + \alpha_0\,\varepsilon_F + \frac{i}{\Omega_\chi}\mathcal{R}}{\alpha_0 + c_\Psi\,\omega + \frac{i}{\gamma_\Psi}\mathcal{R}} \tag{21}$$

and therefore

$$\Sigma(\omega+i\eta) \sim \alpha_0\,\varepsilon_F + \omega - \frac{\omega(1+c_\chi) + \alpha_0\,\varepsilon_F + \frac{i}{\Omega_\chi}\mathcal{R}}{\alpha_0 + c_\Psi\,\omega + \frac{i}{\gamma_\Psi}\mathcal{R}}. \tag{22}$$

Note that we adjusted the self-energy so that 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

$$Z = \frac{\alpha_0}{1 + c_\chi - \varepsilon_F\,c_\Psi}. \tag{23}$$

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 ω and $k \sim k_F$

$$A(k,\omega) \sim \left(\frac{{\alpha_0}^2}{\pi\Omega_\Sigma}\right)\frac{\mathcal{R}\left(1-\frac{\omega}{\Delta}\right)}{\{(1+c_\chi - c_\Psi\,\varepsilon_k)\omega - \alpha_0(\varepsilon_k - \varepsilon_F)\}^2 + \left\{{\alpha_0}^2\mathcal{R}^2/\Omega_\Sigma^2\right\}}, \tag{24}$$

where

$$\Omega_\Sigma \equiv \alpha_0\frac{\Omega_\chi\gamma_\Psi}{\gamma_\Psi - \varepsilon_F\Omega_\chi}, \tag{25}$$

$$\Delta \equiv \frac{{\alpha_0}^2\,\gamma_\Psi\Omega_\chi}{\Omega_\Sigma\{(1+c_\chi)\Omega_\chi - c_\Psi\gamma_\Psi\}}. \tag{26}$$

In terms of the wave-function renormalization factor Z,

$$A(k,\omega) \sim \left(\frac{Z^2}{\pi\Omega_\Sigma}\right)\frac{\mathcal{R}\left(1-\frac{\omega}{\Delta}\right)}{\{\omega - Z(\varepsilon_k - \varepsilon_F)\}^2 + \left\{Z^2\mathcal{R}^2/\Omega_\Sigma^2\right\}}. \tag{27}$$

We thus obtain the following final form for the low-energy expression of the Dysonian self-energy:

$$\begin{aligned}\mathrm{Im}\Sigma(\omega) &\sim -\frac{\mathcal{R}}{\Omega_\Sigma}\frac{1-\frac{\omega}{\Delta}}{\{1+\omega\,c_\Psi/\alpha_0\}^2 + \mathcal{R}^2/(\alpha_0\gamma_\Psi)^2},\\ \mathrm{Re}\Sigma(\omega) &\sim \alpha_0\varepsilon_F + \omega\\ &\quad - \frac{\{\varepsilon_F + \omega(1+c_\chi)/\alpha_0\}(1+\omega c_\Psi\,/\alpha_0) + \mathcal{R}^2/({\alpha_0}^2\Omega_\chi\gamma_\Psi)}{\{1+\omega\,c_\Psi/\alpha_0\}^2 + \mathcal{R}^2/(\alpha_0\gamma_\Psi)^2},\end{aligned} \tag{28}$$

where we recall that $\mathcal{R}(\omega,T)$ is defined in Eq. (14). These expressions, and in particular that of ImΣ, are among the key results of the present paper, and will be used below in order to fit and interpret the DMFT data.

If we take the $\Psi_{\rm rem}$ and $\chi_{\rm rem}$ terms in Eq. (15) and Eq. (16) into account, then Eq. (28) receives higher-order polynomial corrections in ω 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.[2,6,7]

C. Low-doping limit $n \to 1$

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,[16] we find that the self-energy and wave function renormalization factor scale correctly with δ as $\delta \to 0$. This assumption gives $\alpha_0 = \delta$, and a scaling of various energy scales with δ. 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

$$Z = \frac{\alpha_0}{1 + c_\chi}, \tag{29}$$

$$\Omega_\Sigma = \alpha_0\, \Omega_\chi, \tag{30}$$

$$\Delta = \frac{\alpha_0\, \gamma_\Psi}{\{(1 + c_\chi)\Omega_\chi - c_\Psi \gamma_\Psi\}}. \tag{31}$$

Near half filling ($\delta \to 0$), we define

$$Z = \delta\, \overline{Z}, \tag{32}$$

$$\Omega_\Sigma = \delta^2\, \overline{\Omega}_\Sigma, \tag{33}$$

$$\Delta \equiv \delta\, \overline{\Delta}. \tag{34}$$

All objects with an overline, such as $\overline{Q}$, are determined to be finite as $\delta \to 0$. 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 δ. 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 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 ImΣ becomes a scaling function of ω/δ at low doping levels.

IV. DOPED MOTT INSULATOR: SINGLE-SITE DMFT

The dynamical mean-field theory[1] is based on the fact[17] that in the limit of a large number of dimensions d the self-energy becomes a momentum-independent local quantity, $\Sigma(\mathbf{k},\omega) \to \Sigma(\omega)$. This implies that the bulk problem for $d \to \infty$ coincides with the problem of an interacting impurity embedded in an appropriate noninteracting bath.[12] 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.[18–22]

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.[23] The truncation cutoff in the NRG was set in the energy space at $10\omega_N$ (here ω_N is the characteristic energy at the Nth 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 δ 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 Σ, we have used the procedure[25] based on the following exact relation from equations of motion:

$$\Sigma_\sigma(z) = \frac{\langle\langle [d_\sigma, H_{\rm int}]; d_\sigma^\dagger\rangle\rangle_z}{\langle\langle d_\sigma; d_\sigma^\dagger\rangle\rangle_z} = \frac{U\langle\langle n_{\bar\sigma} d_\sigma; d_\sigma^\dagger\rangle\rangle_z}{\langle\langle d_\sigma; d_\sigma^\dagger\rangle\rangle_z}. \tag{35}$$

Here d_σ is the impurity annihilation operator, while $H_{\rm int}$ is the interaction part of the Hamiltonian. The two correlators in this expression were computed using the full-density-matrix NRG algorithm.[24,26] To accelerate the convergence of the DMFT self-consistency loop, the Broyden mixing algorithm has been used.[27] 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.[27]

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

$$F_{\rm UHB}(z) = \frac{w_{\rm UHB}}{z - (\epsilon_d + U)} \underset{U\to\infty}{\longrightarrow} \frac{w_{\rm UHB}}{-U} \tag{36}$$

from the UHB, where $w_{\rm UHB}$ is the total weight of the upper Hubbard band, which in the $U \to \infty$ 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 $U \to \infty$ limit is equal to $-w_{\rm UHB}$. 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 δ

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 δ, give either an insulating or a metallic Fermi liquid state. The key quantity characterizing the metallic state is the quasiparticle residue

$$Z = \left(1 - \frac{\partial \mathrm{Re}\Sigma(\omega)}{\partial \omega}\bigg|_{\omega=0}\right)^{-1}. \tag{37}$$

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[1]—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[28] 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).[1] As $\delta \to 0$ 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 δ 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/δ vs δ focusing on the low-doping region [Figs. 1(c) and 1(d)] reveals, however, that the asymptotic low-δ 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[28–30]: 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[30] $\sim(1 - U_c/U)^{-1/2}$. 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. $\mathrm{Im}\Sigma$ 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[31] $\propto \pm\sqrt{Z}$, their locations in the doped case are no longer symmetric and will be discussed below. In addition, there are two broad humps in $\mathrm{Im}\Sigma$ 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 $\mathrm{Im}\Sigma$ 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, $\mathrm{Im}\Sigma$ 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 $\omega < 0$, $|\mathrm{Im}\Sigma|$ 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-ω FL $\sim \omega^2$ dependence. In contrast, at small doping, the low-ω 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 $\omega > 0$, $|\mathrm{Im}\Sigma|$ 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, $|\mathrm{Im}\Sigma|$ 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 $\rho_{\mathcal{G}}(\omega)$, 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 $|\mathrm{Im}\Sigma|$. The detailed form of the line shape on the more

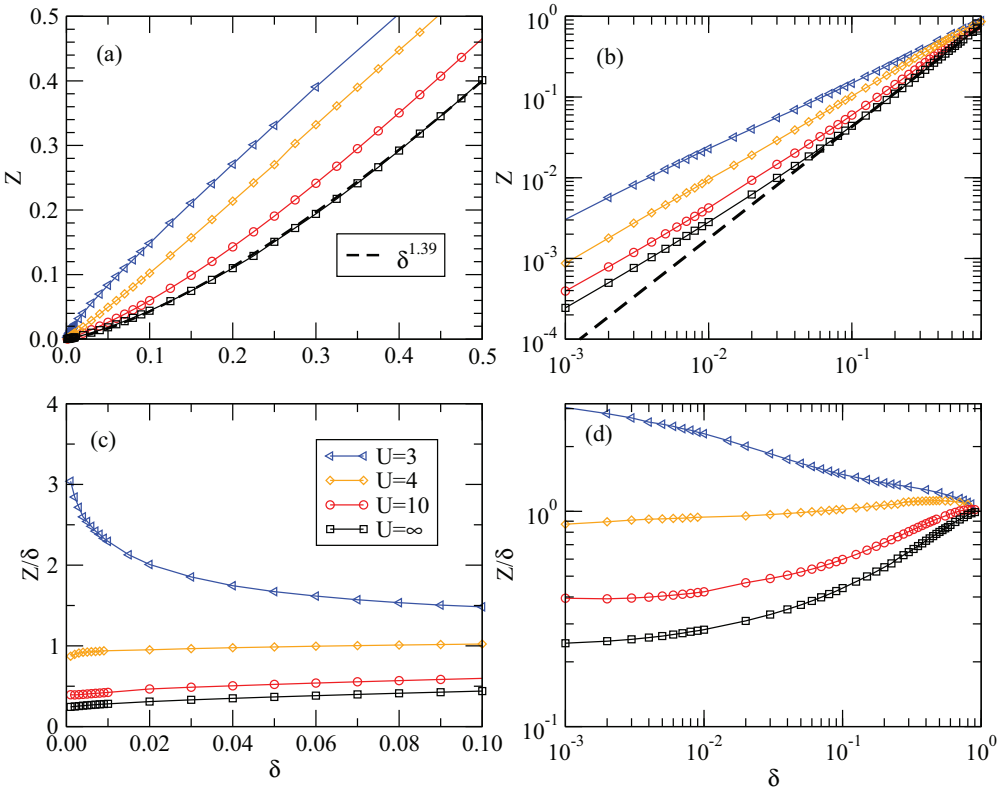

FIG. 1. (Color online) Doping-driven Mott transition within the DMFT: approach to the Mott insulating state for $U > U_c = 2.918D$ with decreasing doping δ. Top panels: doping dependence of Z for a range of δ on a lin-lin (a) and on a log-log (b) plot. The dashed line is a fit to a power-law function $Z = \delta^{\alpha}$ with $\alpha = 1.39$. Bottom panels: Z/δ vs δ(c), and Z/δ vs δ 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 $\mathrm{Re}\Sigma$ 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.[32]

4. ω/Z scaling

Close to the Mott transition, all low-frequency properties are expected to scale with Z, i.e., be described by scaling functions[1,33,34] of ω/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. ω/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.[22,35–37]

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 ω/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 $\mathrm{Im}\Sigma$ is seen to be located at a frequency proportional to doping level ($\omega_{\mathrm{peak}} \simeq -0.7\delta D$). This peak in $\mathrm{Im}\Sigma$ 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 $\mathrm{Re}\Sigma$.

5. Deviation from the low-frequency Fermi liquid behavior

From Figs. 6(b), (c) (for $\mathrm{Re}\Sigma$) and 6(e), (f) (for $\mathrm{Im}\Sigma$), 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: $\mathrm{Re}\Sigma - \mathrm{Re}\Sigma(0) = \omega(1 - 1/Z)$ and $\mathrm{Im}\Sigma \propto -(\omega/Z)^2$.

When visualized on an intermediate frequency scale [Figs. 6(b) and 6(e)] it is seen that deviation from the FL

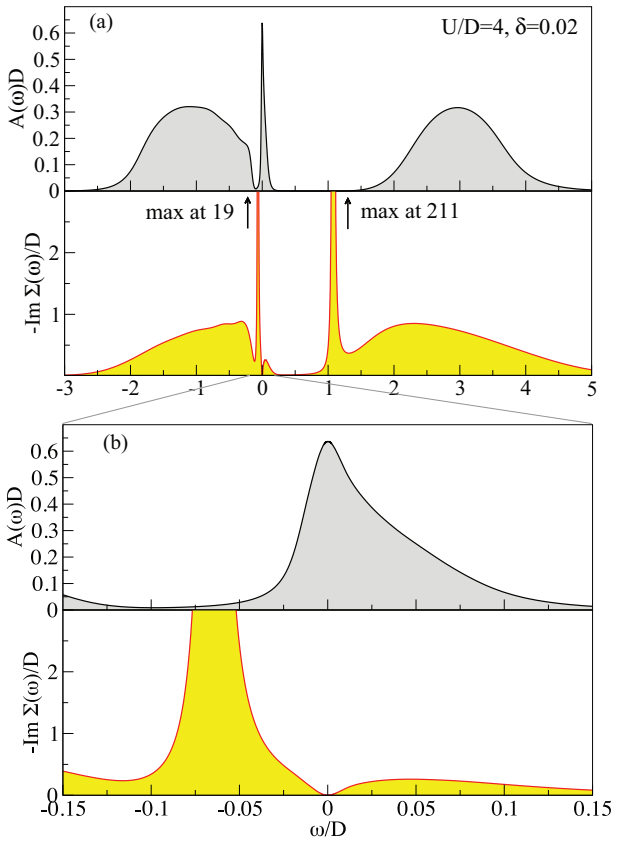

FIG. 2. (Color online) (a) Overview plot showing the full structure of spectra on high-frequency scales (lower Hubbard band, quasiparticle band, upper Hubbard band) at finite $U = 4D$. We show the DMFT local spectral function (top panel) and the imaginary part of the self-energy (bottom panel). (b) Closeup 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.[9,38] $\mathrm{Re}\Sigma$ deviates from linearity and flattens upwards for $\omega > 0$, resulting in the bending of the

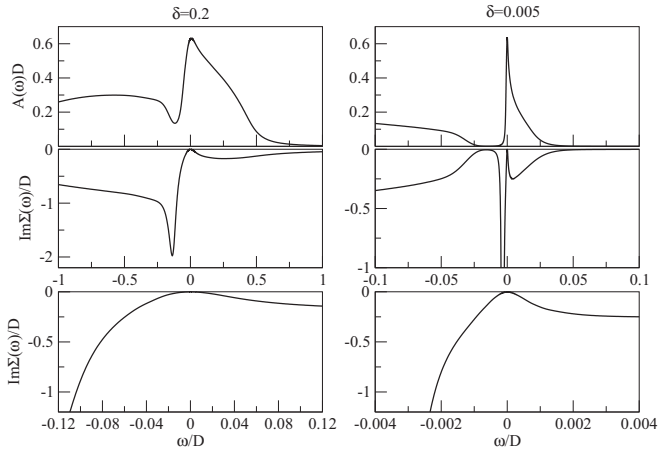

FIG. 3. Local spectral function and the imaginary part of the self-energy for large doping $\delta = 0.2$ (left) and small doping $\delta = 0.005$ (right), for $U = \infty$.

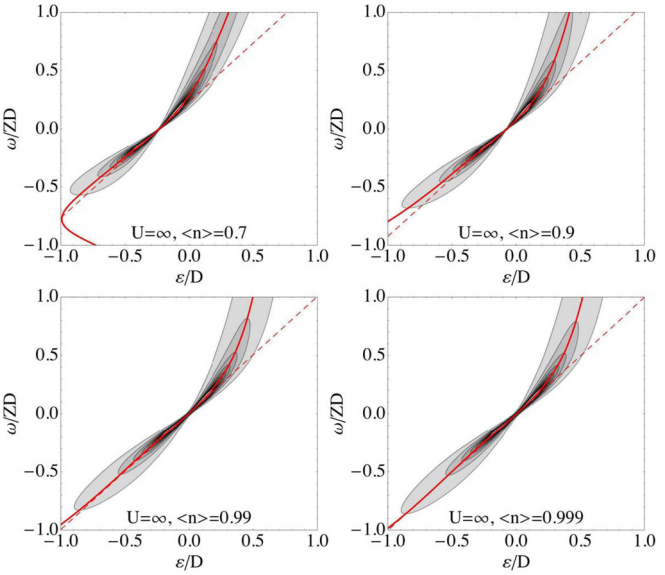

FIG. 4. (Color online) Intensity plots of the momentum (ε_k-) resolved spectral function $A(\varepsilon,\omega)$ for $U = \infty$ at four different doping levels, plotted as a function of ε/D and ω/ZD. The plain line locates the solution of the QP pole equation $\omega + \mu - \mathrm{Re}\Sigma(\omega) - \varepsilon = 0$ (neglecting $\mathrm{Im}\Sigma$). By definition of the QP excitations, this line has slope unity (cf. dashed line) at low-ω when plotted in this manner since $\omega_{\mathrm{QP}} = Z(\varepsilon - \varepsilon_F)$ (i.e., $v_F^\star = Z v_F$) 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 $\mathrm{Im}\Sigma(\omega)$ 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 $\omega_{\mathrm{FL}}^\star \simeq 0.1\,ZD$, which is of order $0.025\delta D$ to $0.05\delta D$ depending on δ. In agreement with previous studies[9] 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),[9] 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 $\pm\omega_{\mathrm{FL}}^\star$. This scale corresponds to a low-energy kink in $\mathrm{Re}\Sigma$. The corresponding low-energy kink in the QP dispersion[39–41] 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

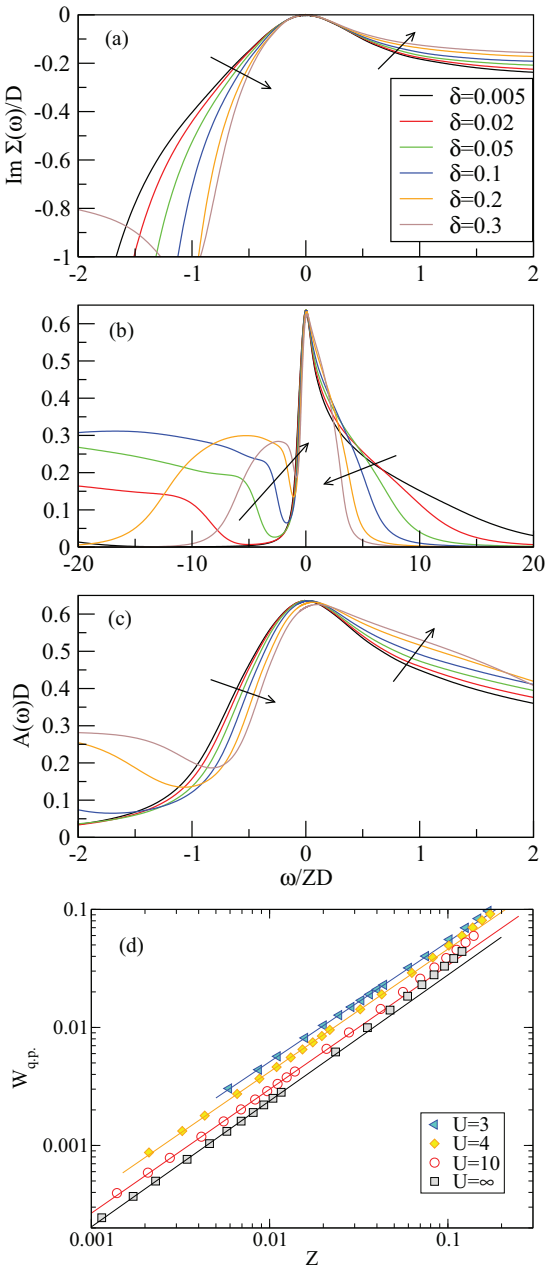

FIG. 5. (Color online) (a) Imaginary part of the self-energy Σ versus the rescaled frequency ω/Z for $U = \infty$. (b), (c) Corresponding local spectral functions. Note that when plotted vs ω/Z, the peak related to the onset of the LHB moves to the left with diminishing δ, as seen in (b). The results for finite U are qualitatively very similar. The arrows indicate the direction of the increasing value of δ. (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 δ, $W_{\text{q.p.}}$ and Z are related by a power-law $W_{\text{q.p.}} = Z^{\gamma}$, with γ close to one, but not exactly 1. More specifically, γ 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.[42] 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 μ 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 μ [Fig. 7(b)]. The charge compressibility κ 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, κ drops to zero as the Mott transition is approached. The asymptotic behavior is a power law δ^{β} 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:

$$x \equiv \omega/\delta \tag{38}$$

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 $\delta \to 0$, e.g., $\overline{P} = P/\delta$, we rewrite Eq. (28) as

$$\text{Im}\Sigma = -\frac{x^2}{\overline{\Omega}_1}\frac{1 - x/\overline{\Delta}_1}{(1 + x/\overline{\Delta}_2)^2 + x^4/\overline{\Omega}_2^2}. \tag{39}$$

This ansatz function is determined by two variables with the meaning of curvature $\overline{\Omega}_{1,2}$ that are simply related to Ω_{Σ} and γ_{Ψ}, respectively, and by two parameters, which adjust the asymmetry $\overline{\Delta}_{1,2}$ that are related to Δ and c_{ψ}.

The numerator of the expression describes a parabolic dependence, with a cubic correction term. This ansatz function has a peak (quasipole) at frequency $x = -\overline{\Delta}_2$ for a finite $\overline{\Omega}_2$, turning into a true pole when $\overline{\Omega}_2 \to \infty$. 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

$$\text{Im}\Sigma = -x^2/\overline{\Omega}_1[1 - (1/\overline{\Delta}_1 + 2/\overline{\Delta}_2)x]. \tag{40}$$

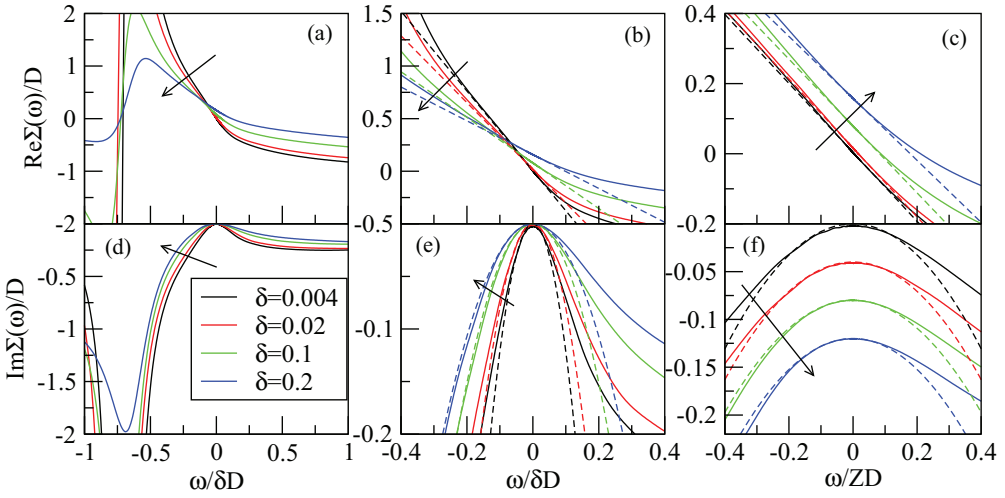

FIG. 6. (Color online) Real and imaginary parts of the DMFT self-energy $\Sigma(\omega)$ for a range of doping for $U = \infty$, as a function of $\omega/\delta D$ (left and center) or as a function of ω/ZD (right). The arrows indicate the direction of the increasing value of δ. The linear (for real 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 Ψ and χ.

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 $\mathrm{Im}\Sigma$ 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 Ψ and χ 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 $\overline{\Delta}_1$ to be discussed below) do not depend substantially on the doping, hence validating our assumptions stated above.

The second observation is that $\overline{\Omega}_1$ is usually found to be very close to $\overline{\Omega}_2$. This supports the conclusions of the λ^2 analysis (to be discussed in the next section), which also finds that γ_Ψ is close to Ω_χ.

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.

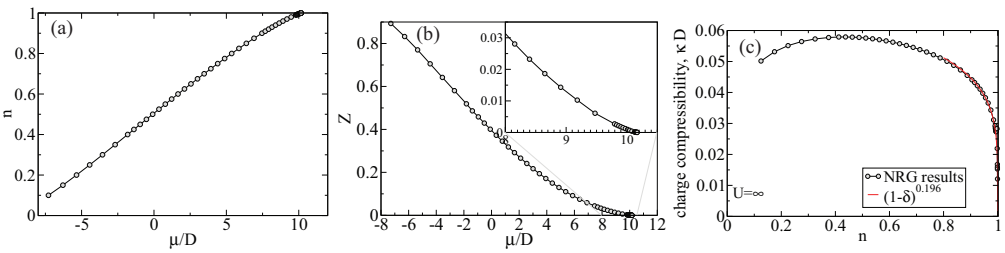

FIG. 7. (Color online) (a) Band filling n vs chemical potential μ for $U = \infty$. (b) Quasiparticle residue Z vs. chemical potential μ for $U = \infty$. (c) Charge compressibility for $U = \infty$.

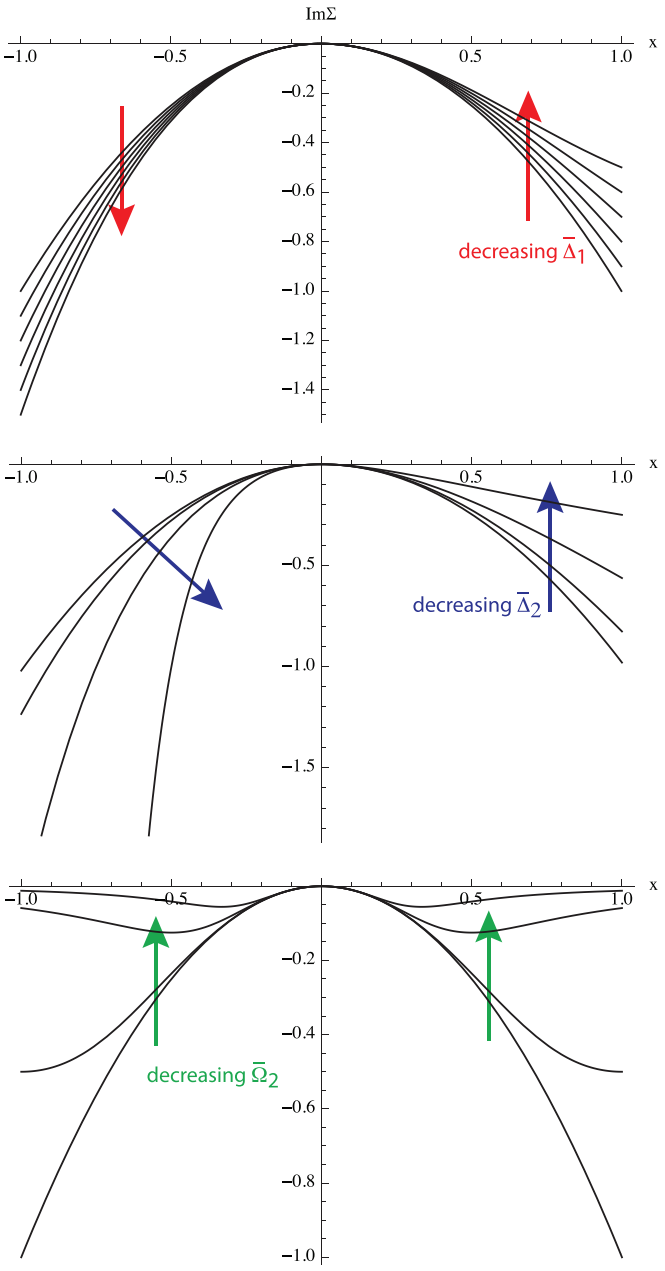

FIG. 8. (Color online) Role of the parameters in the ECFL ansatz function, Eq. (39), used for fitting in Figs. 9 and 10.

C. Summary

To summarize, the ECFL-derived ansatz, Eq. (39), which retains only the lowest-order Fermi liquid terms in Ψ and χ, 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

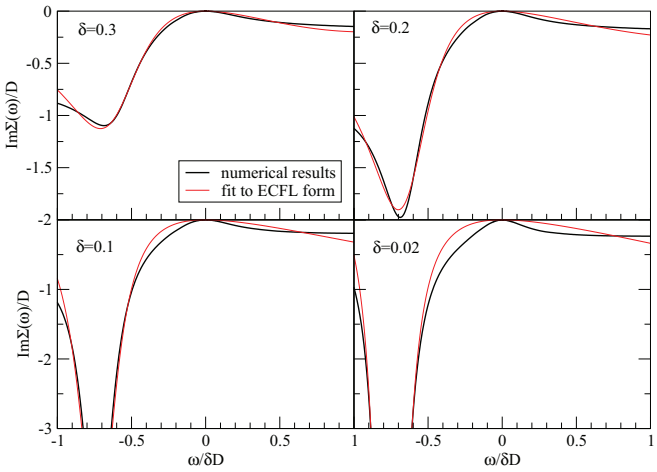

FIG. 9. (Color online) $\mathrm{Im}\Sigma(\omega)$ vs rescaled frequency $\omega/\delta D$ and fits to Eq. (39) for $U = \infty$. The ansatz describes the DMFT self-energy at moderate dopings remarkably well. The nipple structure that becomes pronounced at small doping signals that χ and Ψ develop non-Fermi 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 Ψ, χ.

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

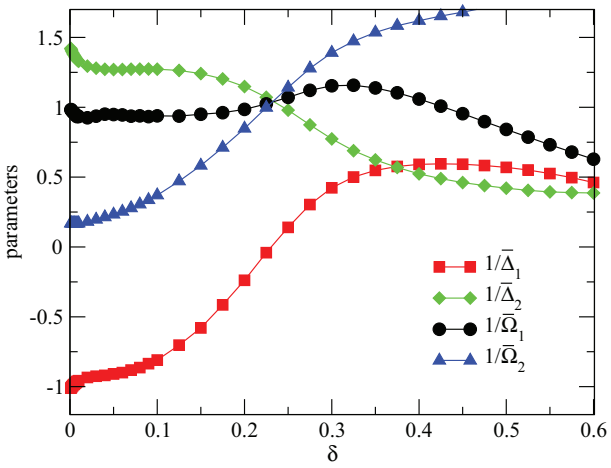

FIG. 10. (Color online) Parameters in the fit function, Eq. (39), for a range of doping δ at $U = \infty$. For a broad fitting energy range, the fitting parameters are smooth as a function of δ, thus the fitting procedure is well defined. At dopings where the fits work best (that is around $\delta = 0.25$), $\bar{\Omega}_1$ is found to be close to $\bar{\Omega}_2$.

one could perform the λ expansion to infinite order in λ, 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 λ expansion against the exact DMFT results. Note that in the $d \to \infty$ 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

$$1 - \frac{n}{2} \to a_{\mathcal{G}} \equiv 1 - \lambda\frac{n}{2} + \lambda^2\frac{n^2}{4} + O(\lambda^3). \tag{41}$$

The second rule is that the explicit self-energy expressions in these equations are multiplied by λ. As an illustration of these rules, we write Eq. (8) and (9) as

$$\mathbf{g}^{-1}(k) = i\omega + \mu - a_{\mathcal{G}}\,\varepsilon_k - \lambda\Phi(k), \tag{42}$$

$$\widetilde{\mu}(k) = a_{\mathcal{G}} + \lambda\,\Psi(k). \tag{43}$$

The factor λ 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 λ:

$$\begin{aligned}
\Psi(i\omega_k) &= -\lambda\sum_{pq}(\varepsilon_p + \varepsilon_q - u_0)\mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(p+q-k),\\
\Phi(k) &= \left(\varepsilon_k - \frac{u_0}{2}\right)\Psi(i\omega_k) + \chi(i\omega_k) - u_0\left(\lambda\frac{n^2}{8} - \frac{n}{2}\right)\\
&\quad - \sum_p \varepsilon_p\mathbf{g}(p)\\
\chi(i\omega_k) &= -\lambda\sum_{pq}\left(\varepsilon_{p+q-k} - \frac{u_0}{2}\right)(\varepsilon_p + \varepsilon_q - u_0)\\
&\quad \times\mathbf{g}(p)\mathbf{g}(q)\mathbf{g}(p+q-k).
\end{aligned} \tag{44}$$

All equations in Eq. (44) are implicitly understood to have $O(\lambda^3)$ corrections, so that the $\mathbf{g}$ and $\widetilde{\mu}$ pieces of $\mathcal{G}$ in Eq. (7) are correct to the stated order. As expected, the functions Ψ, χ 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 μ 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 $\varepsilon_p \to \varepsilon_p + c$. 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 χ and Ψ are independently shift invariant. Combining the expressions, we write

$$\begin{aligned}
\mu' &\equiv \mu + u_0\left(\lambda^2\frac{n^2}{8} - \lambda\frac{n}{2}\right) - \frac{u_0}{2}a_{\mathcal{G}} + \lambda\sum_p \varepsilon_p\mathbf{g}(p),\\
\mathbf{g}^{-1}(k) &= i\omega + \mu' - \left(\varepsilon_k - \frac{u_0}{2}\right)\{a_{\mathcal{G}} + \lambda\Psi(i\omega_k)\} - \lambda\chi(i\omega_k).
\end{aligned} \tag{45}$$

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 $m = 0, 1, \ldots$ using Eq. (45) as

$$\begin{aligned}
\mathbf{g}_{\mathrm{loc},m}(i\omega_k) &\equiv \sum_{\vec{k}}\mathbf{g}(k)\,(\varepsilon_{\vec{k}})^m\\
&= \int_{-D}^{D} d\varepsilon\;\rho_0(\varepsilon)\,\frac{\varepsilon^m}{i\omega + \mu' - (a_{\mathcal{G}} + \Psi)\left(\varepsilon - \frac{u_0}{2}\right) - \chi},
\end{aligned} \tag{46}$$

where χ and Ψ are functions of frequency $i\omega_k$ but not the energy ε. We find that both $\mathbf{g}_{\mathrm{loc},0}$ and $\mathbf{g}_{\mathrm{loc},1}$ 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 $\frac{n}{2} = \sum_{i\omega}\mathbf{g}_{\mathrm{loc},0}(i\omega)$ and $\frac{n}{2} = \sum_{i\omega}\mathcal{G}_{\mathrm{loc},0}(i\omega)$. Where necessary, the usual convergence factor $e^{i\omega 0^+}$ is inserted. The two $\vec{k}$-independent functions Ψ and χ 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:

$$\begin{aligned}
I_{m_1m_2m_3}(i\omega) &= -\frac{1}{\beta^2}\sum_{\nu_1,\nu_2}\mathbf{g}_{\mathrm{loc},m_1}(i\nu_1)\\
&\quad\times\mathbf{g}_{\mathrm{loc},m_2}(i\nu_2)\mathbf{g}_{\mathrm{loc},m_3}(i\nu_1 + i\nu_2 - i\omega).
\end{aligned} \tag{47}$$

After continuation $i\omega \to \omega + i\eta$, 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:

$$\begin{aligned}
\Psi(i\omega) &= -u_0 I_{000}(i\omega) + 2I_{010}(i\omega)\\
\chi(i\omega) &= -\frac{u_0}{2}\Psi(i\omega) - u_0 I_{001} + 2I_{011}(i\omega).
\end{aligned} \tag{48}$$

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 ρ_χ, ρ_ψ, and ρ_Σ (from top to bottom). ρ_ψ and ρ_χ have similar frequency dependence and a Fermi liquid form is obeyed much more accurately than what is found for the Dyson self-energy ρ_Σ. 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.

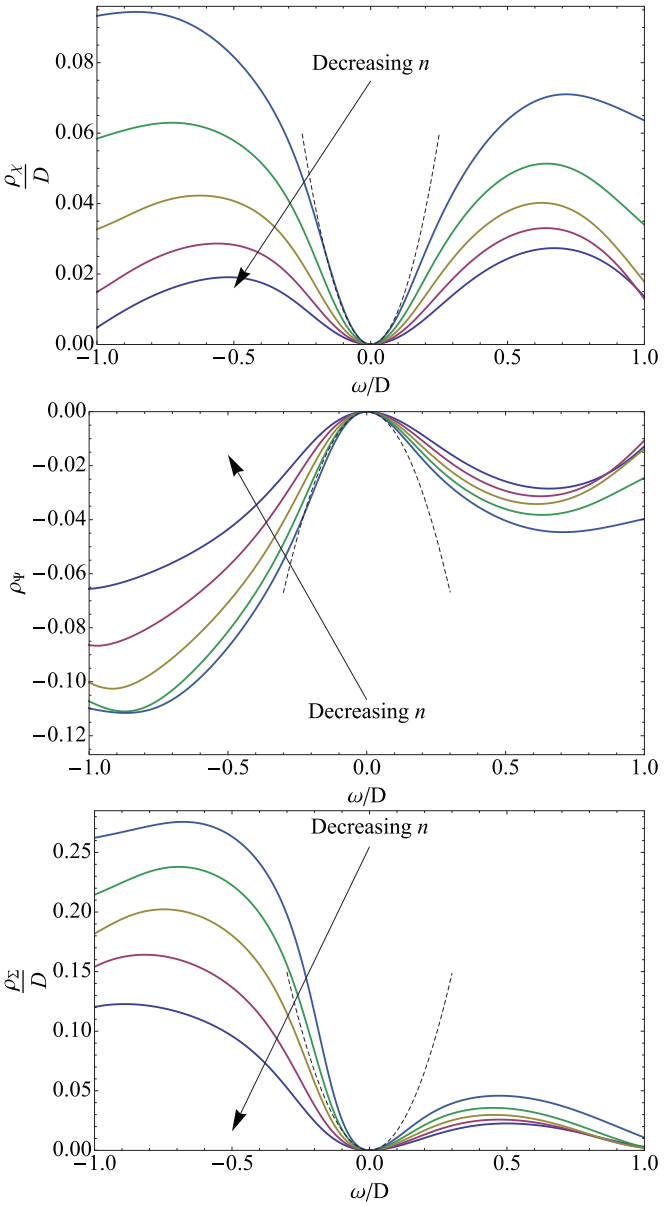

FIG. 11. (Color online) Imaginary parts (spectral densities) of the auxiliary self-energies χ and Ψ, and the Dyson self-energy Σ within the $O(\lambda^2)$ ECFL. The dotted lines are parabolic fits at the highest density. Recall that typical Fermi liquid-type spectral functions exhibit a parabolic and therefore particle-hole symmetric behavior over a large energy range. From these fits one observes that the auxiliary functions Ψ and χ have a Fermi liquid form over a wider energy range than the Dyson self-energy Σ.

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-ω limiting behavior $\mathcal{G} \sim \frac{a_{\mathcal{G}}}{\omega}$, differing from the exact form $\mathcal{G} \sim \frac{1-\frac{n}{2}}{\omega}$ due to the replacement $1-\frac{n}{2} \to a_{\mathcal{G}} \equiv 1-\lambda n/2+\lambda^2 n^2/4$ 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 $n_{\rm eff}$, using the ratio of particle addition and removal states as the relevant metric, so that $\frac{n}{1-n+n^2/4} = \frac{n_{\rm eff}}{1-n_{\rm eff}}$, thus yielding

$$n_{\rm eff} = \frac{n}{1+\frac{n^2}{4}}. \tag{49}$$

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 $n_{\rm eff}$.

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 λ 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 ϵ-resolved spectral functions at two values of the band energy, $\varepsilon_k = -D$ and $\varepsilon_k = \varepsilon_F$. In general, the agreement is encouraging. At $\varepsilon_k = -D$ (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 ε_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 ω/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 $|\omega| \leqslant 0.5DZ$. 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 ω/ZD as in Fig. 14, further emphasizes this similarity. At this level of description, the $U = 4D$ DMFT curve and the

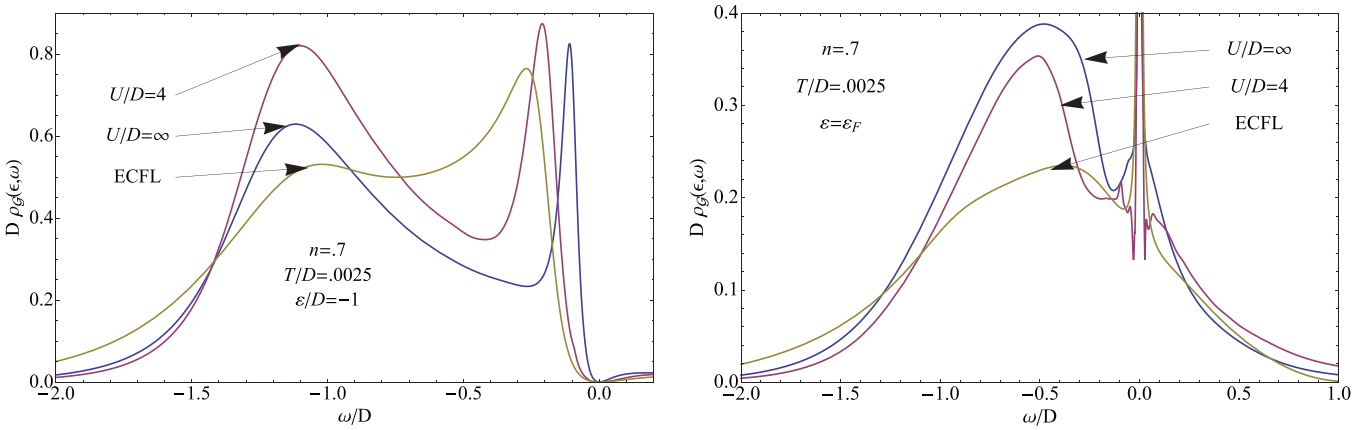

FIG. 12. (Color online) Spectral functions within the DMFT ($U = \infty$ and $U/D = 4$) and ECFL at two typical energies $\varepsilon = -D$ and $\varepsilon = \varepsilon_F$, with $n = 0.7$ and $T = 0.0025D$. The location of the quasiparticle peak near $\omega \sim 0$ and the broad secondary peak for $\omega < 0$ 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

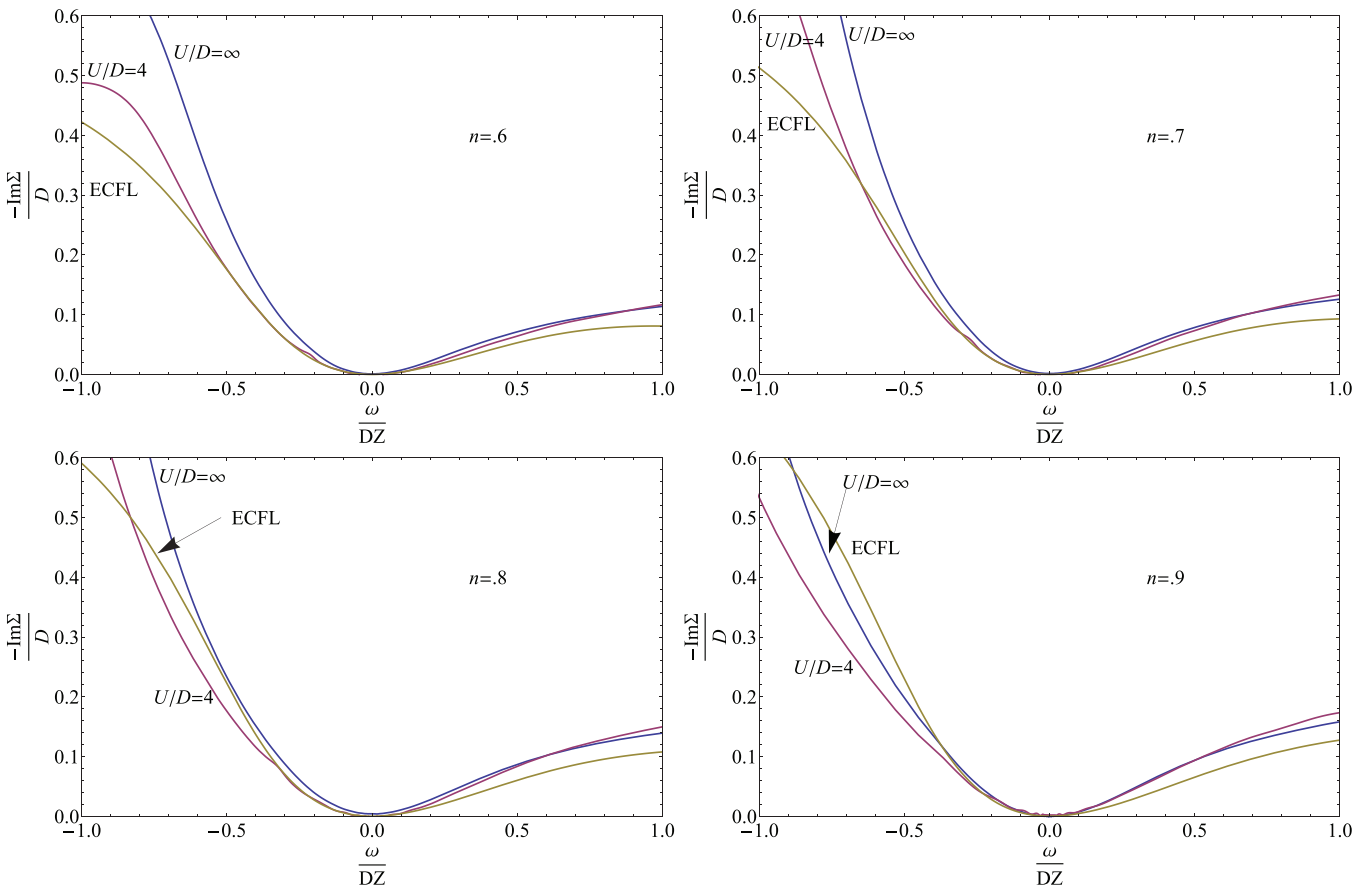

FIG. 13. (Color online) Spectral function (imaginary part) of the Dyson self-energy Σ versus the scaled variable $\omega/(DZ)$ in the ECFL theory at order λ^2, and the DMFT at two values of U. The ECFL predicts a value of Z which is too large at low doping, and significant U dependence creates differences 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.

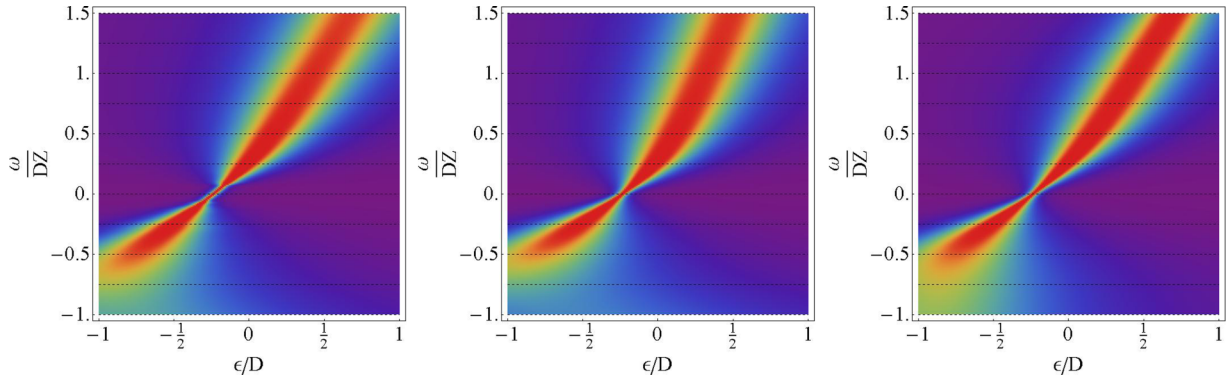

FIG. 14. (Color online) Physical spectral function $A(\epsilon,\omega)$. From left to right: $U = 4$ DMFT, $U = \infty$ DMFT, and ECFL with $n = 0.7$ and $T/D = 0.0025$. Hot colors represent high intensity, while darker blue represents low intensity. Noting from the left panel of Fig. 12 that the QP 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 δ, 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 Ψ and χ. 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 Ψ,χ 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 ω/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.[43] 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 λ.

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[44] and DMFT[9] 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.[9] This dynamical asymmetry has physical implications for the spectral line shapes[9,44] as well as thermopower.[9,45] 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[6,7] 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 caparison 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 λ, and brute-force higher-order calculations in λ 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.[15,46]

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[47] 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 δ, 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

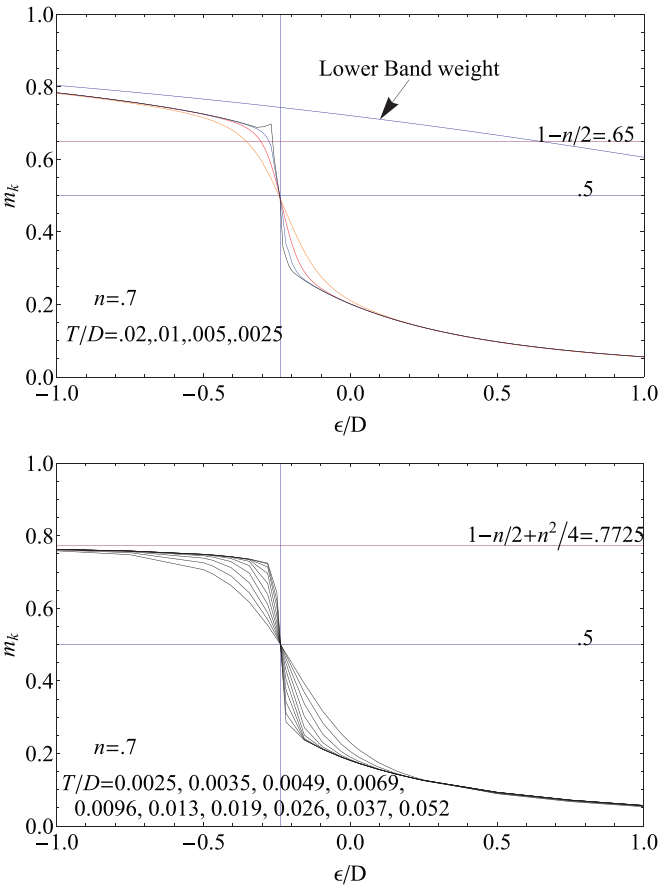

FIG. 15. (Color online) Momentum occupation versus ε within the DMFT at $U/D = 4$ (top) and the ECFL (bottom).

becomes computationally very expensive (several hundreds of DMFT cycles) even with Broyden acceleration.[22] 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

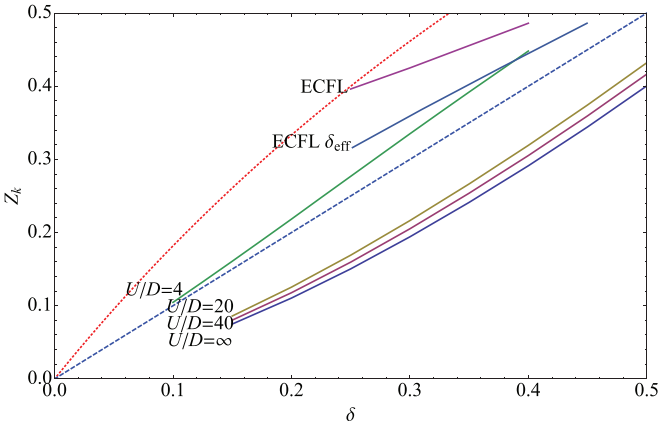

FIG. 16. (Color online) Quasiparticle weight Z as a function of hole doping δ (or $\delta_{\rm eff} = 1 - n_{\rm eff}$, the effective hole doping) from the $O(\lambda^2)$ version of ECFL, and from DMFT for various values of U. More detailed DMFT results are in Fig. 1. The blue dashed line represents $Z = \delta$, 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 Λ, 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

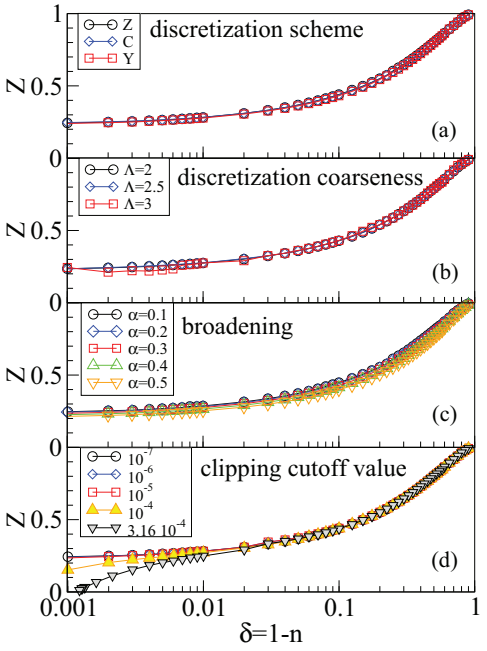

FIG. 17. (Color online) Quasiparticle residue Z for the $U = \infty$ Hubbard model as a function of doping δ. We compare the DMFT results for different choices of the NRG impurity solver parameters.

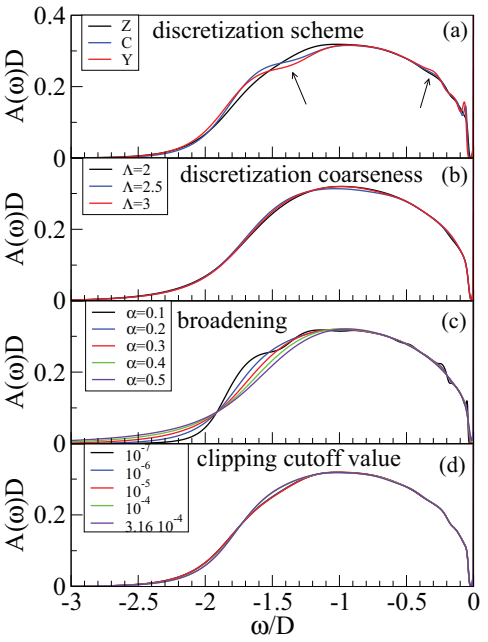

FIG. 18. (Color online) Lower Hubbard band part of the local spectral function $A(\omega)$ for the $U = \infty$ Hubbard model at low doping, $\delta = 0.01$. In (a), the schemes C and Y produce spurious features at the outer band edge and at $\omega = 0.3D$ (indicated by arrows). In addition, the (expected) 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 $\geqslant 10^{-4}$ 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 δ. 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 α, which controls how the raw spectral function in the form of a set of weighted δ 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 α 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 α is decreased, see Fig. 17(c). We find that for $\alpha \leqslant 0.1$, 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 α 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 δ 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 δ, 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.

[1] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[2] B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011).

[3] B. S. Shastry, Phys. Rev. B **87**, 125124 (2013).

[4] B. S. Shastry, Phys. Rev. B **84**, 165112 (2011).

[5] E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, and B. S. Shastry, Phys. Rev. B **87**, 161120 (2013).

[6] G.-H. Gweon, B. S. Shastry, and G. D. Gu, Phys. Rev. Lett. **107**, 056404 (2011).

[7] Kazue Matsuyama and G.-H. Gweon, Phys. Rev. Lett. **111**, 246401 (2013).

[8] E. Perepelitsky and B. S. Shastry, Ann. Phys. (N.Y.) **338**, 283 (2013).

[9] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, Phys. Rev. Lett. **110**, 086401 (2013).

[10] B. S. Shastry, Phys. Rev. Lett. **108**, 029702 (2012).

[11] 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.

[12] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992).

[13] 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 Hermitean,[3] (p. 11).

[14] To avoid the linear growth with ω 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.

[15] D. Hansen and B. S. Shastry, Phys. Rev. B **87**, 245101 (2013).

[16] 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_χ diverge as $\frac{1}{\delta}$. We prefer the current scheme for its simplicity and economy.

[17] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).

[18] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[19] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1003 (1980).

[20] R. Bulla, Phys. Rev. Lett. **83**, 136 (1999).

[21] R. Bulla, T. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008).

[22] R. Žitko and T. Pruschke, Phys. Rev. B **79**, 085106 (2009).

[23] R. Žitko, Phys. Rev. B **84**, 085142 (2011).

[24] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).

[25] R. Bulla, A. C. Hewson, and T. Pruschke, J. Phys.: Condens. Matter **10**, 8365 (1998).

[26] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006).

[27] R. Žitko, Phys. Rev. B **80**, 125125 (2009).

[28] W. F. Brinkman and T. M. Rice, Phys. Rev. B **2**, 4302 (1970).

[29] D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984).

[30] P. Nozières, La méthode de Gutzwiller, Lecture Notes at Collège de France, 1986 (unpublished).

[31] M. J. Rozenberg, G. Kotliar, and X. Y. Zhang, Phys. Rev. B **49**, 10181 (1994).

[32] 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).

[33] G. Moeller, Q. Si, G. Kotliar, M. Rozenberg, and D. S. Fisher, Phys. Rev. Lett. **74**, 2082 (1995).

[34] R. Bulla, T. Pruschke, and A. C. Hewson, Physica B Condens. Matt. **259**, 721 (1999).

[35] M. Karski, C. Raas, and G. S. Uhrig, Phys. Rev. B **72**, 113110 (2005).

[36] M. Karski, C. Raas, and G. S. Uhrig, Phys. Rev. B **77**, 075116 (2008).

[37] Žiga Osolin and R. Žitko, Phys. Rev. B **87**, 245135 (2013).

[38] M. M. Zemljič, P. Prelovšek, and T. Tohyama, Phys. Rev. Lett. **100**, 036402 (2008).

[39] K. Byczuk, M. Kollar, K. Held, Y.-F. Yang, I. A. Nekrasov, T. Pruschke, and D. Vollhardt, Nature Phys. **3**, 168 (2007).

[40] P. Grete, S. Schmitt, C. Raas, F. B. Anders, and G. S. Uhrig, Phys. Rev. B **84**, 205104 (2011).

[41] K. Held, R. Peters, and A. Toschi, Phys. Rev. Lett. **110**, 246402 (2013).

[42] M. Eckstein, M. Kollar, M. Potthoff, and D. Vollhardt, Phys. Rev. B **75**, 125103 (2007).

[43] B. S. Shastry, E. Perepelitsky, and A. C. Hewson, Phys. Rev. B **88**, 205108 (2013).

[44] B. S. Shastry, Phys. Rev. Lett. **109**, 067004 (2012).

[45] 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. Zlatic and A. C. Hewson (Springer, The Netherlands, 2009), pp. 119–131.

[46] R. R. P. Singh and R. L. Glenister, Phys. Rev. B **46**, 14313 (1992).

[47] T. M. Rice and K. Ueda, Phys. Rev. Lett. **55**, 995 (1985).

[48] M. Yoshida, M. A. Whitaker, and L. N. Oliveira, Phys. Rev. B **41**, 9403 (1990).

[49] V. L. Campo and L. N. Oliveira, Phys. Rev. B **72**, 104432 (2005).

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 $U \to \infty$. We develop an expansion in a parameter λ, 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 $n_d \to 1$, 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 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 λ. 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 λ expansion. Presently, the theory to $O(\lambda^2)$ has been evaluated for the t-J model,[2] and higher-order calculations in λ 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 $n \sim 1$.[3] 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[4] 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[5,6] to the powerful numerical renormalization group (NRG) of Wilson,[7] Krishna-murthy *et al.*,[8] 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,[11,12] slave particles,[13] and the Bethe *ansatz*,[14] 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[19] [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.[19,20] 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

 ©2013 American Physical Society

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[21–24] 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[2] 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 λ 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 λ 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 λ as a partial projector can be found from a simple calculation of the atomic limit, with the parameter λ thrown in. Reference 2 (Appendix A1 and especially Fig. 6) explicitly shows that the double occupancy goes from its maximum to zero as λ varies from 0 to 1. Further, it is useful to view the λ parameter in an operator sense, by writing a partially projected (d-orbital) fermion operator $\hat{f}^\dagger_\sigma(\lambda) = (1 - \lambda\ f^\dagger_{\bar\sigma} f_{\bar\sigma}) f^\dagger_\sigma$ and its adjoint (here $\bar\sigma = -\sigma$). The operator $\hat{f}^\dagger_\sigma(\lambda)$ interpolates between the unprojected Fermi operator $f^\dagger_\sigma$ at $\lambda = 0$, and the Gutzwiller projected Hubbard operator $X^{\sigma 0}_i$ at $\lambda = 1$. The Hamiltonian is written in terms of $\hat{f}^\dagger_\sigma(\lambda)$, $\hat{f}_\sigma(\lambda)$, and expanding in λ gives an effective Hamiltonian that generates the auxiliary Green's function $\mathbf{g}$ below. As explained in Ref. 2, the second (caparison) part also has an expansion in λ that follows from the Schwinger equation and the product form Eq. (12). At the end, we set the parameter $\lambda \to 1$ in the formal equations, and only then begin the actual computations. Therefore, the primary use of the parameter λ 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 caparison function μ, 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 μ, and in terms of these, we find exact Schwinger equations of motion for $\mathbf{g}$ and μ in Eq. (14), which are expressed in terms of the two self-energies [Eq. (15)]. The auxiliary Greens's function and the physical impurity Greens's function satisfy the number sum rules in Eqs. (19) or (23), and are then expanded in the parameter λ 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 $U \to \infty$ given by the following Hamiltonian:

$$H = \sum_\sigma \epsilon_d X^{\sigma\sigma} + \sum_{k\sigma} \epsilon_k n_{k\sigma} + \frac{1}{\sqrt{\Omega}} \sum_{k\sigma} (V_k X^{\sigma 0} c_{k\sigma} + V^*_k c^\dagger_{k\sigma} X^{0\sigma}), \quad (1)$$

where Ω 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

$$\mathcal{G}_{\sigma_i\sigma_f}(\tau_i, \tau_f) = -\langle\langle X^{0\sigma_i}(\tau_i)\ X^{\sigma_f 0}(\tau_f)\rangle\rangle, \quad (2)$$

with T_τ the imaginary-time ordering symbol, the definition for an arbitrary time-dependent operator Q: $\langle\langle Q\rangle\rangle = \langle \mathrm{Tr} T_\tau e^{-\mathcal{A}} Q\rangle / \langle \mathrm{Tr} T_\tau e^{-\mathcal{A}}\rangle$, 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 $\mathcal{V}(\tau_i) \to \mathcal{V}_i$. 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 2×2 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)

$$\{(\partial_{\tau_i} + \epsilon_d)\mathbb{1} + \mathcal{V}_i\}\mathcal{G}(\tau_i, \tau_f) = -\delta(\tau_i - \tau_f) \times [\mathbb{1} - \gamma(\tau_i)] - \frac{1}{\sqrt{\Omega}} [\mathbb{1} - \gamma(\tau_i) + \mathcal{D}_i] \cdot \sum_k V_k\, G(k, \tau_i; \tau_f), \quad (3)$$

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 $\gamma(\tau_i) \to \gamma_i$. 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 $(\mathcal{D}_i)_{\sigma_i\sigma_j} = (\sigma_i\sigma_j)\, \delta/\delta \mathcal{V}^{\bar\sigma_i\bar\sigma_j}(\tau_i)$. In the ECFL formalism,[1] 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: $\sigma_i\sigma_j\langle\langle X^{\bar{\sigma}_i\bar{\sigma}_j}\,Q\rangle\rangle = (\gamma_i - \mathcal{D}_i)\langle\langle Q\rangle\rangle$. Proceeding further, we take a time derivative of the mixed Green's function to find

$$\big(\partial_{\tau_i} + \epsilon_k\big)G(k,\tau_i;\tau_f) = -\frac{1}{\sqrt{\Omega}}V_k^*\,\mathcal{G}(\tau_i,\tau_f), \tag{4}$$

so combining with Eq. (3) we find the exact EOM for the localized electron Green's function:

$$\begin{aligned}\big\{\big(\partial_{\tau_i} + \epsilon_d\big)\mathbb{1} + \mathcal{V}_i\big\}\mathcal{G}(\tau_i,\tau_f)\\ = -\delta(\tau_i - \tau_f)(\mathbb{1} - \gamma_i)\\ - (\mathbb{1} - \gamma_i + \mathcal{D}_i)\cdot\Delta(\tau_i - \tau_{\mathbf{j}})\cdot\ \mathcal{G}(\tau_{\mathbf{j}},\tau_f),\end{aligned} \tag{5}$$

with the convention that the time label in bold letters $\tau_{\mathbf{j}}$ is to be integrated over $\in [0,\beta]$. The conduction band enters through the usual ($\mathcal{V}$-independent) function

$$\Delta(\tau_i - \tau_j) = -\frac{\mathbb{1}}{\Omega}\sum_k |V_k|^2\big(\partial_{\tau_i} + \epsilon_k\big)^{-1}\delta(\tau_i - \tau_j), \tag{6}$$

with a Fourier transform

$$\Delta(i\omega_n) = \frac{1}{\Omega}\sum_k \frac{|V_k|^2}{i\omega_n - \epsilon_k} = V_0^2\int\frac{\rho(\epsilon)d\epsilon}{i\omega_n - \epsilon}. \tag{7}$$

We will require below its analytic continuation $i\omega_n \to \omega + i\eta$:

$$\Delta(\omega + i\eta) = \Delta_R(\omega) - i\Gamma(\omega); \tag{8}$$

$$\Gamma(\omega) = \pi V_0^2\ \rho(\omega); \quad \Delta_R(\omega) = \frac{\Gamma_0}{\pi}\ln\frac{|\omega + D|}{|\omega - D|}. \tag{9}$$

Here, $\Gamma_0 = \pi V_0^2\rho_0$. We now use the noninteracting Green's function

$$\mathbf{g}_0^{-1}(\tau_i,\tau_f) = -\big[\partial_{\tau_i} + \epsilon_d + \mathcal{V}(\tau_i)\big]\delta(\tau_i - \tau_f) - \Delta(\tau_i,\tau_f), \tag{10}$$

and rewrite the fundamental equation of motion Eq. (5) as

$$\begin{aligned}\big\{\mathbf{g}_0^{-1}(\tau_i,\tau_{\mathbf{j}}) + (\gamma_i - \mathcal{D}_i)\cdot\Delta(\tau_i - \tau_{\mathbf{j}})\big\}\cdot\mathcal{G}(\tau_{\mathbf{j}},\tau_f)\\ = (\mathbb{1} - \gamma_i)\delta(\tau_i - \tau_f).\end{aligned} \tag{11}$$

Let us note an important *shift invariance* of Eqs. (11) and (10). If we consider a transformation $\Delta(\tau) \to \Delta(\tau) + u_t \times \delta(\tau)$ with an arbitrary u_t, it is possible to show that Eq. (11) is unchanged, except for a shift of ϵ_d by $-u_t$. The added term $u_t(\gamma_i - \mathcal{D}_i)\mathcal{G}(\tau_i,\tau_f)$ 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*

$$\mathcal{G}(\tau_i,\tau_f) = \mathbf{g}(\tau_i,\tau_{\mathbf{j}})\cdot\mu(\tau_{\mathbf{j}},\tau_f), \tag{12}$$

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 $\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(\tau_n,\tau_m;\tau_i) = -\frac{\delta}{\delta\mathcal{V}_i^{\sigma_3\sigma_4}}\mathbf{g}^{-1}_{\sigma_1\sigma_2}(\tau_n,\tau_m)$ and $\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(\tau_n,\tau_m;\tau_i) = \frac{\delta}{\delta\mathcal{V}_i^{\sigma_3\sigma_4}}\mu_{\sigma_1\sigma_2}(\tau_n,\tau_m)$ 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: $\ldots\xi^*_{\sigma_a\sigma_b}\ldots\delta/\delta\mathcal{V}^* = \ldots\sigma_a\sigma_b\ldots\delta/\delta\mathcal{V}^{\bar{\sigma}_a\bar{\sigma}_b}$, or in terms of the vertex functions $\ldots\xi^*_{\sigma_a\sigma_b}\ldots\Lambda_*^{\sigma'\sigma''}\ldots = \ldots\sigma_a\sigma_b\ldots\Lambda^{\sigma'\sigma''}_{\bar{\sigma}_a\bar{\sigma}_b}\ldots$, with the upper indices of Λ 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:

$$\begin{aligned}\Phi(i,j) &= -\mathbf{L}(i,\mathbf{r})\cdot\mathbf{g}^{-1}(\mathbf{r},j) = \xi^*\Delta(i,\mathbf{j})\cdot\mathbf{g}(\mathbf{j},\mathbf{k})\cdot\Lambda_*(\mathbf{k},j;i);\\ \Psi(i,j) &= \mathbf{L}(i,\mathbf{r})\cdot\mu(\mathbf{r},j) = \xi^*\Delta(i,\mathbf{j})\cdot\mathbf{g}(\mathbf{j},\mathbf{k})\cdot\ \mathcal{U}_*(\mathbf{k},j;i).\end{aligned} \tag{13}$$

Summarizing, we may rewrite the exact EOM Eq. (11) symbolically:

$$\big\{\mathbf{g}_0^{-1} + \gamma\Delta - \Phi\big\}\cdot\mathbf{g}\cdot\mu = (\mathbb{1} - \gamma)\delta + \Psi. \tag{14}$$

This equation is split into two parts by requiring $\mathbf{g}$ to be canonical:

$$\mathbf{g}^{-1} = \big\{\mathbf{g}_0^{-1} + \gamma\Delta - \Phi\big\} \quad\text{and}\quad \mu = (\mathbb{1} - \gamma)\delta + \Psi, \tag{15}$$

bringing it into the standard form in the ECFL theory.[1] 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

$$\Phi = \mathcal{L}\cdot\big(-\mathbf{g}_0^{-1} - \gamma\Delta\big) \quad\text{and}\quad \Psi = -\mathcal{L}\cdot\gamma\delta. \tag{16}$$

Therefore, distributing the action of $\mathcal{L}$ over the two terms, we can rewrite

$$\Phi = \chi + \Psi\Delta, \tag{17}$$

with

$$\chi = \mathcal{L}\cdot\big(-\mathbf{g}_0^{-1}\big). \tag{18}$$

Therefore, the self-energy Φ 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

$$\mathbf{g}(\tau,\tau^+) = \frac{n_d}{2}; \quad \mathcal{G}(\tau,\tau^+) = \frac{n_d}{2}, \tag{19}$$

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 $\hat{f}^\dagger_\sigma(\lambda) = (1 - \lambda f^\dagger_{\bar{\sigma}}f_{\bar{\sigma}})f^\dagger_\sigma$ for the correlated electrons, this constraint is understandable as the constraint on the number of "uncorrelated" fermions $\langle f^\dagger_\sigma f_\sigma\rangle$, which must agree with the number of physical (correlated) electrons $\langle\hat{f}^\dagger_\sigma\hat{f}_\sigma\rangle$. Similarly, in the present case, this constraint is needed to fulfill the Friedel sum rule. We also remark that the self-energy Ψ, unlike its counterpart Φ, is dimensionless, and thus interpreted as an adaptive spectral weight.[2]

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 $\gamma \to \frac{n_d}{2}$ we obtain the following expressions in frequency space:

$$\begin{aligned} \mathcal{G}(i\omega_n) &= \mathbf{g}(i\omega_n)\cdot\mu(i\omega_n), \\ \mu(i\omega_n) &= 1 - \frac{n_d}{2} + \Psi(i\omega_n), \\ \mathbf{g}^{-1}(i\omega_n) &= i\omega_n - \epsilon_d - \Delta(i\omega_n)\mu(i\omega_n) - \chi(i\omega_n). \end{aligned} \tag{20}$$

Alternately, this result can be rewritten in terms of the Dyson-Mori self-energy representation as

$$\mathcal{G}(i\omega_n) = \frac{1-\frac{n_d}{2}}{i\omega_n - \epsilon_d - \left(1-\frac{n_d}{2}\right)\Delta(i\omega_n) - \Sigma_{\rm DM}(i\omega_n)} \tag{21}$$

and

$$\begin{aligned} &\Sigma_{\rm DM}(i\omega_n) + \epsilon_d - i\omega_n \\ &\quad = \frac{1-\frac{n_d}{2}}{1-\frac{n_d}{2}+\Psi(i\omega_n)}\left[\chi(i\omega_n) + \epsilon_d - i\omega_n\right]. \end{aligned} \tag{22}$$

The sum rules (19) are

$$\sum_{i\omega_n} \mathcal{G}(i\omega_n)e^{i\omega_n\eta} = \frac{n_d}{2}; \quad \sum_{i\omega_n} \mathbf{g}(i\omega_n)e^{i\omega_n\eta} = \frac{n_d}{2}. \tag{23}$$

These are satisfied at a fixed n_d using two Lagrange multipliers: the localized state energy ϵ_d and the second chemical potential u_0 introduced in Eq. (25). We observe that the usual Dysonian self-energy $\Sigma_{\rm AM}(i\omega_n)$ defined through the usual Dyson equation (valid for finite U) $G^{-1} = i\omega_n - \epsilon_d - \Delta(i\omega_n) - \Sigma_{\rm AM}(i\omega_n)$ in the infinite-U limit can be obtained from

$$\Sigma_{\rm AM}(i\omega_n) = \frac{2}{2-n_d}\Sigma_{\rm DM}(i\omega_n) + \frac{n_d}{2-n_d}(\epsilon_d - i\omega_n). \tag{24}$$

The unlimited growth with ω_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 $i\omega_n \to \omega + i0^+$, the imaginary part of $\Sigma_{\rm AM}$ is well behaved and finite as $\omega \to \infty$. 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 γ and therefore n_d explicitly.

C. Introducing λ 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

$$\left\{\mathbf{g}_0^{-1} + \lambda(\gamma - \mathcal{D})\left(\Delta - \frac{u_0}{2}\delta\right)\right\}\cdot\mathcal{G} = (\mathbb{1} - \lambda\gamma)\delta. \tag{25}$$

As a consequence, in Eq. (14) to Eq. (18) we set $\gamma \to \lambda\gamma$, $\Psi \to \lambda\Psi$, and $\Phi \to \lambda\Phi$, or $\chi \to \lambda\chi$. Second, in Eq. (14) to Eq. (18) we set $\Delta(\tau_i,\tau_f) \to \Delta(\tau_i,\tau_f) - \frac{u_0}{2}\,\delta(\tau_i - \tau_f)$. Note that there is no shift of Eq. (10) implied in Eq. (25).

We write Eq. (15) with λ inserted explicitly and the understanding that $\Delta(\tau_i,\tau_f)$ has been shifted as[25]

$$\begin{aligned} \mathbf{g}^{-1}(\tau_i,\tau_f) &= \mathbf{g}_0^{-1}(\tau_i,\tau_f) + \lambda\gamma(\tau_i)\Delta(\tau_i,\tau_f) - \lambda\,\Phi(\tau_i,\tau_f), \\ \mu(\tau_i,\tau_f) &= \delta(\tau_i-\tau_f)[\mathbb{1} - \lambda\gamma(\tau_i)] + \lambda\,\Psi(\tau_i,\tau_f), \end{aligned} \tag{26}$$

where the two self-energies are given in terms of the vertex functions as

$$\begin{aligned} \Phi(\tau_i,\tau_f) &= \xi^*\Delta(\tau_i,\tau_{\mathbf{j}})\cdot\mathbf{g}(\tau_{\mathbf{j}},\tau_{\mathbf{k}})\cdot\Lambda_*(\tau_{\mathbf{k}},\tau_f;\tau_i), \\ \Psi(\tau_i,\tau_f) &= \xi^*\Delta(\tau_i,\tau_{\mathbf{j}})\cdot\mathbf{g}(\tau_{\mathbf{j}},\tau_{\mathbf{k}})\cdot\mathcal{U}_*(\tau_{\mathbf{k}},\tau_f;\tau_i). \end{aligned} \tag{27}$$

On switching off the sources, these expressions can be spin resolved and expressed as $\Phi = \Delta\mathbf{g}\Lambda^{(a)}$ and $\Psi = \Delta\mathbf{g}\,\mathcal{U}^{(a)}$, with the same time labels as above, and with the usual spin decomposition $\Lambda^{(a)} = \Lambda^{\sigma\sigma}_{\bar\sigma\bar\sigma} - \Lambda^{\sigma\bar\sigma}_{\sigma\bar\sigma}$.

D. λ 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[26]:

$$\begin{aligned} t[i,f] \to -\Delta(\tau_i,\tau_f); \quad \varepsilon_k \to \Delta(i\omega_k), \\ J \to 0, \quad \boldsymbol{\mu} \to -\epsilon_d. \end{aligned} \tag{28}$$

The λ 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

$$a_{\mathcal{G}} = 1 - \lambda\frac{n_d}{2} + \lambda^2\frac{n_d^2}{4}, \tag{29}$$

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

$$\mathcal{G}(i\omega_n) = \mathbf{g}(i\omega_n)\mu(i\omega_n), \quad \mu(i\omega_n) = a_{\mathcal{G}} + \lambda\Psi(i\omega_n), \tag{30}$$

$$\mathbf{g}^{-1}(i\omega_n) = i\omega_n - \epsilon_d' - \left(\Delta(i\omega_n) - \frac{u_0}{2}\right)\mu(i\omega_n) - \lambda\chi(i\omega_n), \tag{31}$$

$$\begin{aligned} \chi(i\omega_n) = -\lambda\sum_{p,q}&[2\Delta(i\omega_p) - u_0] \\ &\times\left[\Delta(i\omega_p + i\omega_q - i\omega_n) - \frac{u_0}{2}\right]\mathcal{R}, \end{aligned} \tag{32}$$

$$\Psi(i\omega_n) = -\lambda\sum_{p,q}[2\Delta(i\omega_p) - u_0]\mathcal{R}. \tag{33}$$

The energy ϵ_d' is given by collecting the static terms in Φ as

$$\epsilon_d' = \epsilon_d + u_0\left(\lambda\frac{n_d}{2} - \lambda^2\frac{n_d^2}{8}\right) + \frac{u_0}{2}a_{\mathcal{G}} - \lambda\sum_{i\omega_p}\Delta(i\omega_p)\mathbf{g}(i\omega_p). \quad (34)$$

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 ϵ_d, enables us to fix the pair of sum rules (19). As explained, we must set $\lambda \to 1$ before using these expressions.

Within the $O(\lambda^2)$ theory, the total spectral weight of the Green's function is $a_{\mathcal{G}}$ 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 $\Sigma_{\rm DM}(\omega)$ retain the feature of being finite as $\omega \to \infty$, it must be slightly redefined (to $\hat{\Sigma}_{\rm DM}$) in the $O(\lambda^2)$ theory:

$$G(\omega) = \frac{a_{\mathcal{G}}}{\omega - \epsilon_d'' - a_{\mathcal{G}}\Delta(\omega) - \hat{\Sigma}_{\rm DM}(\omega)}, \quad (35)$$

where

$$\epsilon_d'' \equiv \epsilon_d' - \frac{u_0}{2}a_{\mathcal{G}}. \quad (36)$$

Using Eqs. (30) and (31), we can relate $\hat{\Sigma}_{\rm DM}(\omega)$ to $\chi(\omega)$ and $\Psi(\omega)$:

$$\hat{\Sigma}_{\rm DM}(\omega) + \epsilon_d' - \omega = \frac{a_{\mathcal{G}}}{a_{\mathcal{G}} + \Psi(\omega)}[\chi(\omega) + \epsilon_d' - \omega]. \quad (37)$$

Since $\Psi(\omega),\chi(\omega) \to 0$ as $\omega \to \infty$, we see explicitly that $\hat{\Sigma}_{\rm DM}(\omega)$ remains finite in this limit. Just as in the case of Im $\Sigma_{\rm DM}(\omega)$, Im $\hat{\Sigma}_{\rm DM}(\omega)$ is related to Im $\Sigma_{\rm AM}(\omega)$ by a multiplicative constant ($1 - \frac{n_d}{2}$ and $a_{\mathcal{G}}$, 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

$$\Sigma_{\rm DM}(\omega) \to \hat{\Sigma}_{\rm DM}(\omega); \quad \epsilon_d \to \epsilon_d''; \quad 1 - \frac{n_d}{2} \to a_{\mathcal{G}}. \quad (38)$$

Keeping these substitutions in mind, we will now only use $\Sigma_{\rm DM}(\omega)$ from the exact theory, with the understanding that the same expressions hold for $\hat{\Sigma}_{\rm DM}(\omega)$ 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[27–29] 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 σ at energy ω:

$$\eta_\sigma(\omega) = \frac{1}{2i}\ln\left[\mathcal{G}_\sigma(\omega + i0^+)\mathcal{G}_\sigma^{-1}(\omega - i0^+)\right], \quad (39)$$

where the logarithm is chosen with a branch cut along the positive real axis, so that $0 \leqslant \eta \leqslant \pi$. 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[30] 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 $U \to \infty$. Using the Dyson-Mori representation (21) to compute the phase shift in Eq. (39), we may rewrite this as

$$n_d = 1 - \frac{2}{\pi}\tan^{-1}\left[\frac{\epsilon_d + \text{Re}\,\Sigma_{\rm DM}(0)}{\Gamma_0\left(1 - \frac{n_d}{2}\right)}\right], \quad (41)$$

with $\epsilon_d + \text{Re}\,\Sigma_{\rm DM}(0) > 0$, in the physical case of $0 \leqslant n_d \leqslant 1$. It is easily seen[32] that this form is equivalent to the standard statement of the Friedel sum rule[15]

$$\rho_{\mathcal{G}}(0) = \frac{1}{\pi\Gamma_0}\sin^2\left(\frac{\pi n_d}{2}\right). \quad (42)$$

Within the approximation of the λ 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 $\rho_{\mathbf{g}}(\omega = 0)$, 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

$$\rho_{\mathbf{g}}(0) = \frac{1}{\pi\Gamma_0\mu(0)}\sin^2\left(\frac{\pi n_d}{2}\right). \quad (44)$$

We check the validity of the Friedel sum rule within the λ 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[30] 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:

$$Q(i\omega_n) = \int_{-\infty}^{\infty} d\nu \frac{\rho_Q(\nu)}{i\omega_n - \nu}, \quad (45)$$

where Q can stand for any object such as $\mathcal{G}$, $\mathbf{g}$, χ, $\Sigma_{\rm DM}$, or Ψ. Therefore, after analytic continuation $i\omega_n \to \omega + i0^+$,

$$\rho_Q(\omega) \equiv -\frac{\text{Im}}{\pi}Q(\omega + i0^+) \quad \text{and} \quad \text{Re}\,Q(\omega) = \mathcal{H}[\rho_Q](\omega), \quad (46)$$

where for any real density $\rho_Q(\omega)$ the Hilbert transform is denoted as $\mathcal{H}[\rho_Q](\omega) = P\int_{-\infty}^{\infty} d\nu \frac{\rho_Q(\nu)}{\omega - \nu}$. From Eq. (33), we

TABLE I. The bare impurity level ϵ_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	$\rho_{G,\mathrm{ECFL}}(0)$	$\epsilon_{d,\mathrm{ECFL}}$	$\epsilon_{d,\mathrm{NRG}}$	z_{ECFL}	z_{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

$$\rho_{\mathcal{G}}(\omega) = \rho_{\mathbf{g}}(\omega)[a_{\mathcal{G}} + \mathrm{Re}\,\Psi(\omega)] + \rho_{\Psi}(\omega)\mathrm{Re}\,\mathbf{g}(\omega). \quad (47)$$

With $f(\omega) = \frac{1}{1+e^{\beta\omega}}$ and $\bar{f}(\omega) = 1 - f(\omega)$, the two sum rules (23) read as

$$\int_{-\infty}^{\infty} f(\omega)\rho_{\mathbf{g}}(\omega)d\omega = \frac{n_d}{2}, \quad \int_{-\infty}^{\infty} f(\omega)\rho_{\mathcal{G}}(\omega)d\omega = \frac{n_d}{2}. \quad (48)$$

We also note $\rho_{\Delta}(\omega) = \frac{\Gamma(\omega)}{\pi}$. It is useful to define a mixed (composite) density

$$\rho_M(x) = \rho_{\mathbf{g}}(x)\left(\Delta_R(x) - \frac{u_0}{2}\right) + \rho_{\Delta}(x)\mathrm{Re}\,\mathbf{g}(x), \quad (49)$$

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

$$\Psi(\omega) = -2\lambda \int_{u,v,w} \frac{\rho_M(u)\rho_{\mathbf{g}}(v)\rho_{\mathbf{g}}(w)}{\omega - u - v + w} \times [f(u)f(v)\bar{f}(w) + \bar{f}(u)\bar{f}(v)f(w)],$$

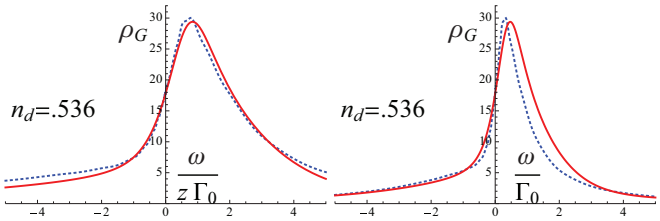

FIG. 2. (Color online) The spectral density for the physical Green's function for the density of $n_d = 0.536$. For the plot on the left, both the ECFL and NRG curves are plotted versus $\frac{\omega}{\Gamma_0 z}$. Since ECFL has a larger z value, the absolute scale of the ω axis differs for the two 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.

$$\chi(\omega) = -2\lambda \int_{u,v,w} \frac{\rho_M(u)\rho_{\mathbf{g}}(v)\rho_M(w)}{\omega - u - v + w} \times [f(u)f(v)\bar{f}(w) + \bar{f}(u)\bar{f}(v)f(w)]. \quad (50)$$

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 Ψ and χ 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 ρ_G, ρ_Σ, ρ_χ, and ρ_Ψ 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 λ 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

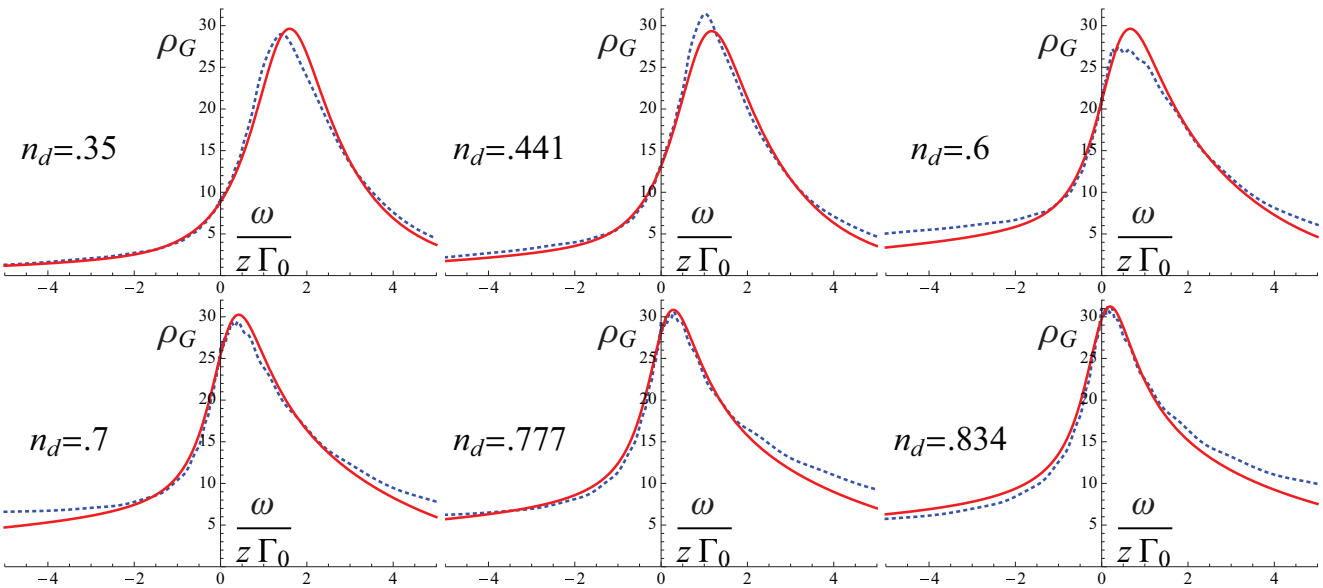

FIG. 1. (Color online) The spectral density for the physical Green's function versus $\frac{\omega}{\Gamma_0 z}$ for densities of $n_d = 0.35, 0.441, 0.6, 0.7, 0.777, 0.834$. The red curve is the ECFL calculation, while the blue curve is the NRG calculation.

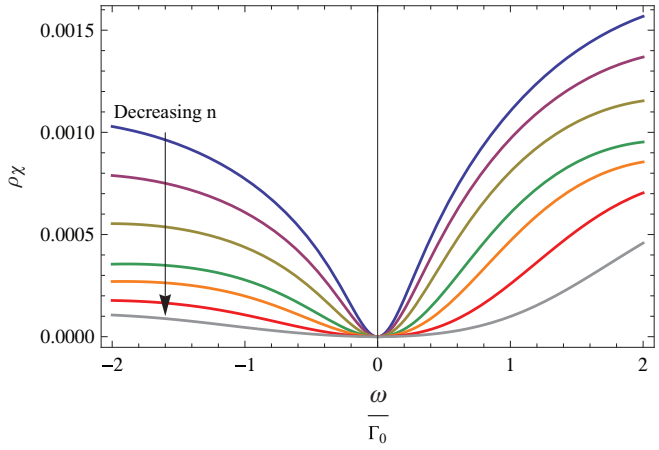

FIG. 3. (Color online) The spectral function for χ for densities of $n_d = 0.834,0.777,0.7,0.6,0.536,0.441,0.35$.

within the λ expansion of the Green's function (33) $\mathcal{G} \sim \frac{a_G}{i\omega}$; 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 ϵ_d and quasiparticle weight z calculated within the ECFL and NRG calculations. The values of ϵ_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 $n_d \to 1$ is expected from the low order in λ 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 λ.

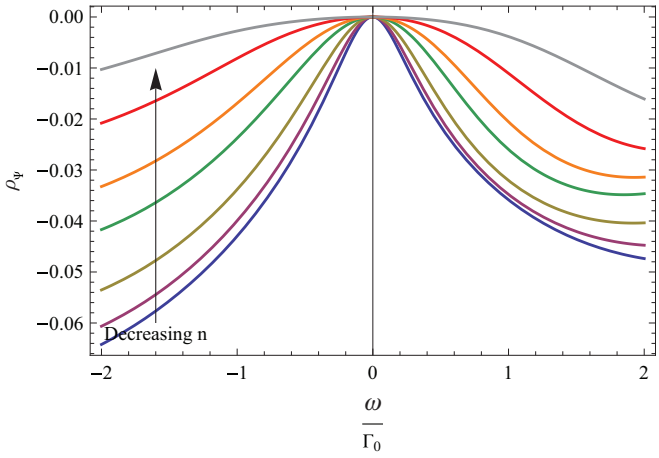

FIG. 4. (Color online) The spectral function for Ψ for densities of $n_d = 0.834,0.777,0.7,0.6,0.536,0.441,0.35$.

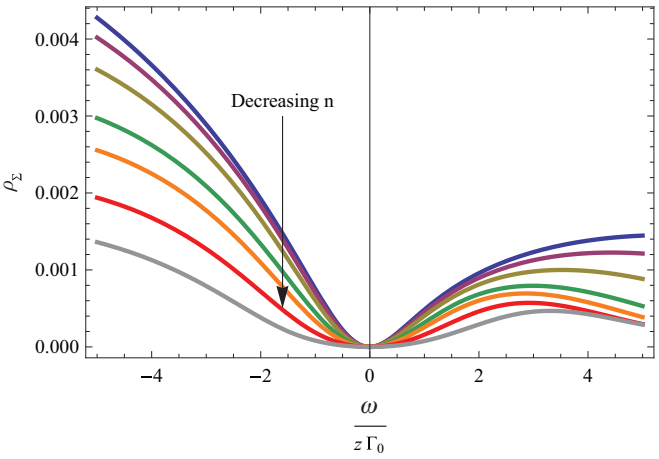

FIG. 5. (Color online) The spectral function for the Dyson-Mori self-energy for densities of $n_d = 0.834,0.777,0.7,0.6,0.536,0.441,0.35$. The curvature of the quadratic minimum becomes larger with increasing density.

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 ω 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 ρ_G is constructed out of the two spectral functions ρ_χ and ρ_Ψ that are shown at various densities in Figs. 3 and 4, exhibiting Fermi liquid type quadratic frequency dependence at low ω.

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[33] and the Hubbard model at large U.[34]

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 $U \to \infty$. In this formalism, the two self-energies of the ECFL theory Ψ and χ 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 Σ, 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 ω, 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 $n_d \to 1$,[31] 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 $n_d \to 1$. It results in the 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 $\Sigma_{\rm DM}(\omega)$ 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 $d \to \infty$, $U \to \infty$ 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 $\epsilon_d + \Sigma_{\rm DM}(0)$ [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 $n_d \to 1$. This can also be understood from the need to have more spectral weight when $\omega \leqslant 0$, 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 ϵ_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 $n_d \to 1$,[31] 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 λ 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 ω 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[33] and the Hubbard model at large U.[34]

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 $[\Lambda]_1$ can be found parallel to Eqs. (B23)–(B28) in Ref. 2 from differentiating

$$[\mathbf{g}^{-1}(i,f)]_1 = \Delta(i,f)\mathbf{g}^{(k)}(i,i) + \delta(i,f)\Delta(i,\mathbf{a})\mathbf{g}^{(k)}(\mathbf{a},f), \tag{A1}$$

as

$$[\Lambda^{(a)}(i,m;j)]_1 = -2\Delta(i,m)\mathbf{g}(i,j)\cdot\mathbf{g}(j,i) - 2\delta(i,m)\Delta(i,\mathbf{k})\mathbf{g}(\mathbf{k},j)\cdot\mathbf{g}(j,i). \tag{A2}$$

Here, the bold labels are integrated over. From this we construct, the time-domain self-energies

$$\Psi(i,f) = -2\lambda\Delta(i,\mathbf{k})\mathbf{g}(\mathbf{k},f)\cdot\mathbf{g}(i,f)\cdot\mathbf{g}(f,i) \tag{A3}$$

and

$$\Phi(i,f) = -\delta(i,f)\Delta(i\mathbf{k})\mathbf{g}(\mathbf{k}i) - 2\lambda\Delta(i\mathbf{j})\mathbf{g}(\mathbf{j}\mathbf{k})\cdot\Delta(\mathbf{k}f)\mathbf{g}(\mathbf{k}i)\cdot\mathbf{g}(i\mathbf{k}) - 2\lambda\Delta(i\mathbf{j})\mathbf{g}(\mathbf{j}f)\cdot\Delta(f\mathbf{k})\mathbf{g}(\mathbf{k}i)\cdot\mathbf{g}(if). \tag{A4}$$

After shifting $\Delta(i,f) \to \Delta(i,f) - \frac{u_0}{2}\delta(i,f)$ 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 μ which are correct to $O(\lambda^2)$. χ can be extracted from Φ 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

$$\mathbf{g}(\tau) = \int_x \rho_{\mathbf{g}}(x)e^{-\tau x}[\theta(-\tau)f(x) - \theta(\tau)\bar{f}(x)], \tag{B1}$$

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

$$X(\tau) = \int_{-\beta}^{\beta} d\bar{\tau}\, \mathbf{g}(\bar{\tau})\left[\Delta(\tau-\bar{\tau}) - \frac{u_0}{2}\delta(\tau-\bar{\tau})\right] = \int_x \rho_M(x)e^{-\tau x}[\theta(-\tau)f(x) - \theta(\tau)\bar{f}(x)], \quad \text{(B3)}$$

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

$$\Psi(\tau) = -2\lambda X(\tau)\mathbf{g}(\tau)\mathbf{g}(-\tau), \qquad \chi(\tau) = -2\lambda X(\tau)X(-\tau)\mathbf{g}(\tau), \quad \text{(B4)}$$

so that taking Fourier transforms is simplest if we first multiply out as in Eq. (B2), leading to Eq. (50).

[1] B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011).

[2] B. S. Shastry, Phys. Rev. B **87**, 125124 (2013); D. Hansen and B. S. Shastry, *ibid.* **87**, 245101 (2013).

[3] 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).

[4] P. W. Anderson, Phys. Rev. **124**, 41 (1961).

[5] P. W. Anderson, Phys. Rev. **164**, 352 (1967); P. W. Anderson and G. Yuval, Phys. Rev. Lett. **23**, 89 (1969).

[6] F. D. M. Haldane, Phys. Rev. Lett. **40**, 416 (1978).

[7] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[8] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1044 (1980).

[9] 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).

[10] 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).

[11] N. Bickers, Rev. Mod. Phys. **59**, 845 (1987).

[12] N. Read and D. M. Newns, J. Phys. C: Solid State Phys. **16**, L1055 (1983); N. Read, *ibid.* **18**, 2651 (1985).

[13] P. Coleman, Phys. Rev. B **29**, 3035 (1984); S. E. Barnes, J. Phys. F: Met. Phys. **6**, 1375 (1976).

[14] N. Andrei, Phys. Rev. Lett. **45**, 379 (1980); P. B. Wiegmann, Sov. Phys.–JETP **31**, 392 (1980).

[15] A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, UK, 1993).

[16] K. Yosida and K. Yamada, Prog. Theor. Phys. Suppl. **46**, 244 (1970).

[17] P. Nozières, J. Low. Temp. Phys. **17**, 31 (1974).

[18] A. C. Hewson, Phys. Rev. Lett. **70**, 4007 (1993).

[19] B. S. Shastry, Phys. Rev. Lett. **109**, 067004 (2012).

[20] G.-H. Gweon, B. S. Shastry, and G. D. Gu, Phys. Rev. Lett. **107**, 056404 (2011).

[21] H. O. Frota and L. N. Oliveira, Phys. Rev. B **33**, 7871 (1986).

[22] O. Sakai, Y. Shimuzu, and T. Kasuya, J. Phys. Soc. Jpn. **58**, 162 (1989).

[23] T. A. Costi and A. C. Hewson, J. Phys.: Condens. Matter **5**, L361 (1993).

[24] 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.

[25] In the notation of Ref. 1, Eq. (58), this corresponds to writing $Y_1(\tau_i,\tau_f) = -\gamma(\tau_i)\Delta(\tau_i,\tau_f)$.

[26] E. Perepelitsky and B. S. Shastry, Ann. Phys. (NY) **338**, 283 (2013).

[27] J. Friedel, Can. J. Phys. **54**, 1190 (1956).

[28] J. S. Langer and V. Ambegaokar, Phys. Rev. **164**, 498 (1967).

[29] D. C. Langreth, Phys. Rev. **150**, 516 (1966).

[30] J. M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960).

[31] J. W. Rasul and A. C. Hewson, J. Phys. C: Solid State Phys. **17**, 3337 (1984).

[32] To recover Eq. (42), we may use Eq. (21) and the Fermi liquid assumption of $\mathrm{Im}\,\Sigma_{\mathrm{DM}}(0) = 0$ so that $\rho_{\mathcal{G}}(0) = \frac{1}{\pi}\frac{\Gamma_0(1-\frac{n_d}{2})^2}{\Gamma_0^2(1-\frac{n_d}{2})^2+[\epsilon_d+\mathrm{Re}\,\Sigma_{\mathrm{DM}}(0)]^2}$, and combine with Eq. (41).

[33] R. Žitko, D. Hansen, E. Perepelitsky, J. Mravlje, A. Georges, and B. S. Shastry, arXiv:1309.5284.

[34] X. Deng, J. Mravlje, R. Zitko, M. Ferrero, G. Kotliar, and A. Georges, Phys. Rev. Lett. **110**, 086401 (2013)

PHYSICAL REVIEW B **87**, 161120(R) (2013)

Electronic spectral properties of the two-dimensional infinite-U Hubbard model

Ehsan Khatami,[1,2] Daniel Hansen,[1] Edward Perepelitsky,[1] Marcos Rigol,[3] and B. Sriram Shastry[1]
[1]*Physics Department, University of California, Santa Cruz, California 95064, USA*
[2]*Department of Physics, Georgetown University, Washington D.C., 20057 USA*
[3]*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 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.[1] 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,[2] the on-site repulsive Coulomb interaction, and hence rendered useless. In contrast, for the t-J model embodying extreme correlations, i.e., $U \to \infty$ 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.[3,4] Experiments using angle-resolved photoemission spectroscopy (ARPES)[5–8] 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,[9,10] and several results of the model pertaining to the detailed line shapes find close agreement with experiment.[5] 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 $0 \leqslant n \lesssim 0.7$. The line shapes of this calculation for $n \geqslant 0.5$ 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[11] 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, $\omega \leqslant 0$, 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 $\mathcal{G}(i,\tau_i;j,\tau_j) = -\langle T_\tau \hat{C}_{i\sigma}(\tau_i)\hat{C}^\dagger_{j\sigma}(\tau_j)\rangle$, where T_τ is the time-ordering operator and $\langle . \rangle$ 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 $\hat{C}_{i\sigma} \equiv X^{0\sigma}_i$, etc. As usual,[12]

 ©2013 American Physical Society

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,[13] and pushed to the eighth order by a highly efficient computer program[11] based on Metzner's linked-cluster formalism.[14] 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 ν is denoted by $\rho_{\mathcal{G}}(k,\nu)$ and determines the Green's function through the relation $\mathcal{G}(k,i\omega_n) = \int_{-\infty}^{+\infty} \frac{\rho_{\mathcal{G}}(k,\nu)}{i\omega_n - \nu} d\nu$. At high frequencies ω_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

$$\mathcal{G}(k,\tau) = \int_{-\infty}^{+\infty} d\nu\, \rho_{\mathcal{G}}(k,\nu) e^{-\nu\tau}[\Theta(-\tau) f(\nu) - \Theta(\tau)\bar{f}(\nu)], \quad (3)$$

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 α_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

$$\alpha_m(k) = \int_{-\infty}^{\infty} \nu^m \chi(\nu)\rho_{\mathcal{G}}(k,\nu)d\nu. \quad (5)$$

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 $\chi(\nu)\rho_{\mathcal{G}}(\nu)$, and naturally leads to three sets of moments at each k, $\varepsilon_m^{\chi}(k) = \alpha_m(k)/\alpha_0(k)$. 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.[15] 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,

$$m_\sigma(k) = \langle \hat{C}^\dagger_{k\sigma}\hat{C}_{k\sigma}\rangle = b_0(k). \quad (8)$$

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 λ, 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,[9] 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 $\rho_{\mathcal{G}}(k,\nu)$ 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 $\varepsilon^{\rm SP}(k)$, 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.[18] The results for the series are obtained from Padé approximations as the bare results show divergent behavior at $T < 1$. The number of terms

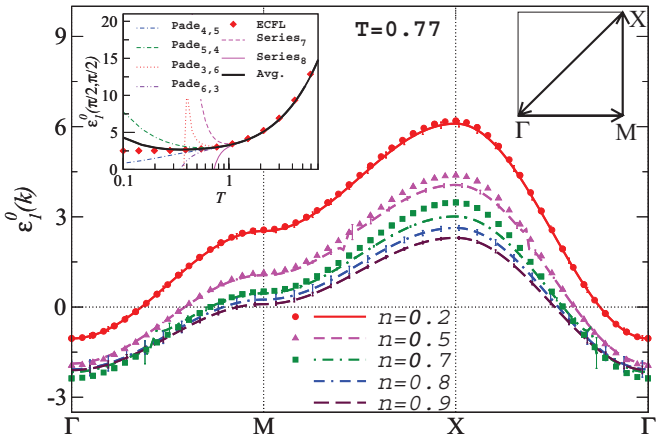

FIG. 1. (Color online) The first symmetric moment $\varepsilon_1^0(k)$ at $T = 0.77$ vs momentum around the irreducible wedge of the Brillouin zone (the path is shown in the right inset). Lines are results from the series and symbols for $n \leqslant 0.7$ are from ECFL calculations. Left inset: $\varepsilon_1^0(k)$ 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$_7$ and Series$_8$), 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$_{\{4,5\}}$ and Padé$_{\{5,4\}}$. In the main panel, the results for the series are either the average between Padé$_{\{4,5\}}$ and Padé$_{\{5,4\}}$ or Padé$_{\{5,5\}}$ and Padé$_{\{5,4\}}$,[16] with the "error bars" defined as the differences between the two.[17]

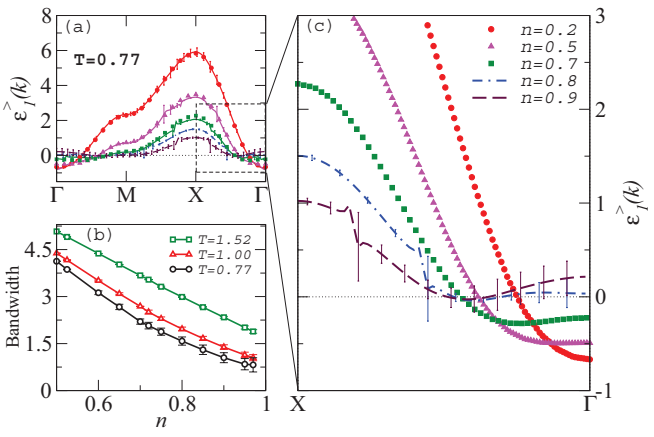

FIG. 2. (Color online) (a) The first *greater* moment $\varepsilon_1^>(k)$ at $T = 0.77$ vs momentum for the same path around the irreducible wedge of the Brillouin zone as in Fig. 1. Lines and symbols are also the same as in Fig. 1. (b) The bandwidth of $\varepsilon_1^>(k)$, defined as the difference 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 $n \leqslant 0.7$, and therefore, we show only the ECFL results for the latter cases.

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 $n \leqslant 0.7$ 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 $n \to 1$ 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 Γ 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 the momentum increases from zero, leading to a negative curvature, or effective mass, at the Γ 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

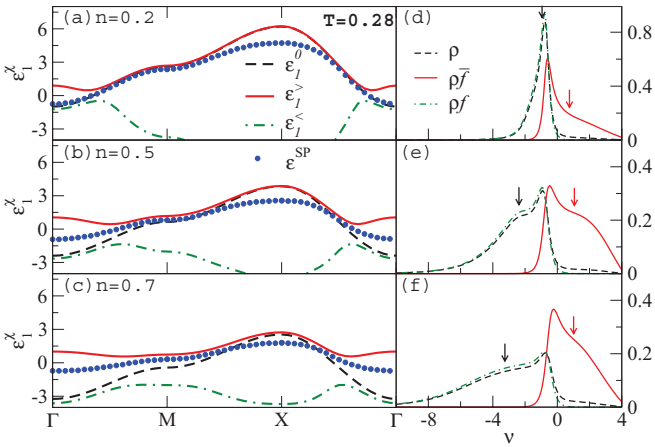

FIG. 3. (Color online) Comparison of the SP location $\varepsilon^{\rm SP}(k)$ (symbols) and the three moments from ECFL at $T = 0.28$ and for (a) $n = 0.2$, (b) $n = 0.5$, and (c) $n = 0.7$. Right panels show the corresponding spectral functions and their products to $\bar{f}(\omega)$ and $f(\omega)$ at Γ for the same densities shown in the left panels. Dark (light) arrows show the values of ε_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 $\varepsilon^{\rm SP}(k)$, 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 $\rho_{\mathcal{G}}(k,\omega)$, $\rho_{\mathcal{G}}(k,\omega)\bar{f}(\omega)$, and $\rho_{\mathcal{G}}(k,\omega)f(\omega)$, 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 $\rho_{\mathcal{G}}$ 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 $\rho_{\mathcal{G}}$ 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 $\varepsilon^{\rm SP}(k)$. 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 $\rho_{\mathcal{G}}$ 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 $\varepsilon^{\rm SP}(k)$ 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.).

[1] 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).

[2] W. Nolting, Z. Phys. **255**, 25 (1972); J. J. Deisz, D. W. Hess, and J. W. Serene, Phys. Rev. B **66**, 014539 (2002).

[3] *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London, 1974), Vol. 3.

[4] 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).

[5] 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.

[6] 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).

[7] 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).

[8] A. Kaminski *et al.*, Phys. Rev. B **69**, 212509 (2004)

[9] 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).

[10] D. Hansen and B. S. Shastry, arXiv:1211.0594.

[11] E. Khatami, E. Perepelitsky, B. S. Shastry, and M. Rigol (unpublished).

[12] A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

[13] E. Perepelitsky and B. S. Shastry (unpublished).

[14] W. Metzner, Phys. Rev. B **43**, 8549 (1991).

[15] B. Farid, Philos. Mag. **84**, 909 (2004).

[16] Note that the zeroth-order terms in the expansion for $\varepsilon_1^0(k)$ is proportional to β^{-1}. Hence, Padé$_{\{l,m\}}$ with $l + m = 9$ will match the

series order by order exactly up to the eighth order. Nevertheless, we find that 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 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 Padé$_{\{5,4\}}$ and Padé$_{\{5,5\}}$ for $n = 0.8$ and Padé$_{\{5,4\}}$ and Padé$_{\{4,5\}}$ for the rest.

[17]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.

[18]We may take the curves of $\varepsilon_1^0(k)$, or more accurately, $\varepsilon_1^>(k)$ as estimates of the SP dispersion $\varepsilon^{\mathrm{SP}}(k)$, 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 $\varepsilon^{\mathrm{SP}}(k)$ and the arrows locating the moments.

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 λ, related to the density, is argued to be quantitatively valid in the overdoped regime for $0 \leqslant n \lesssim 0.75$, 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 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).[1,2] 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,[3–6] 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.[6]

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[2] develops this idea in a systematic fashion, emphasizing the role of expanding in a parameter λ $(0 \leqslant \lambda \leqslant 1)$, related to the particle density, or more closely to $\lambda \sim (1 - \frac{4}{n^2} d)$, where d is the double occupancy $(0 \leqslant d \leqslant \frac{n^2}{4})$. 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 λ expansion. A method for generating a systematic set of equations for the propagator to any orders in λ is given, along with explicit equations to second order in λ 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 mth-order skeleton diagram expansion of the standard Feynman-diagram-based theory, as described in standard texts,[7–9] but generalize to the case of extreme correlations. Summarizing the arguments in Refs. 1 and 2, a low-order theory in λ 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 Φ and Ψ, within the ECFL formalism. The self-energy Ψ 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 λ, namely, $0 \leqslant \lambda \leqslant 1$, 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 $0 \leqslant U \leqslant \infty$. 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 λ seems well worth examining in detail; this is our task here.

 ©2013 American Physical Society

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.[3,4] 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 λ 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 λ 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.[11,12]

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,[13] high-temperature series expansions,[14] variational wave functions,[15–17] and finite temperature Lanczos methods.[18–21] 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.[22] We note that the dynamical mean-field theory (DMFT) for the Hubbard model[23,24] 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[25] 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 $J \to 0$ 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

$$\mathcal{G}(k) = \mathbf{g}(k)\,\mu(k), \quad \text{where } (k) \equiv (\vec{k}, i\omega_k). \tag{1}$$

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 λ, relating to density, and finally we set $\lambda \to 1$. As shown in Eqs. (I-83), (I-84), and (I-85), the second-order equations for the ECFL Green's function are as follows:

$$\mu(k) = 1 - \lambda\frac{n}{2} + \lambda^2\,\frac{n^2}{4} + \lambda^2\Psi(k), \tag{2}$$

$$\Psi(k) = -\sum_{p,q}(\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p} - u_0) \times \mathbf{g}(p)\,\mathbf{g}(q)\,\mathbf{g}(q+k-p), \tag{3}$$

$$\mathbf{g}^{-1}(k) = i\omega_n + \boldsymbol{\mu}' - \overline{\varepsilon}_k - \lambda^2\,\overline{\Phi}(k), \tag{4}$$

$$\overline{\varepsilon}_k = \left(1 - \lambda\, n + \lambda^2\,\frac{3n^2}{8}\right)\varepsilon_k + \lambda\,\sum_q \tfrac{1}{2} J_{k-q}\,\mathbf{g}(q), \tag{5}$$

$$\overline{\Phi}(k) = -\sum_{q,p}\mathbf{g}(q)\,\mathbf{g}(p)\,\mathbf{g}(k+q-p) \times (\varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + J_{k-p} - u_0\,) \times \left\{\varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + \tfrac{1}{2}(J_{k-p} + J_{p-q}) - u_0\right\}, \tag{6}$$

where $\sum_k \equiv \frac{1}{\beta N_s}\sum_{\vec{k},\omega_n}$, with N_s being the number of lattice sites and β is inverse temperature. These expressions for the Green's function satisfy the "shift invariances" described in

Ref. 2, i.e., any uniform shift in ε_k or J_k can be absorbed in $\boldsymbol{\mu}'$ 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 Ψ and $\overline{\Phi}$, respectively. This is written with a slight change of notation $[\Phi(k)]_1 \to \overline{\Phi}(k)$ from (I-85), and we have introduced the effective band energy $\overline{\varepsilon}_k$ in Eq. (5) that gets a static contribution from shrinking of the bare energies ε_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:

$$\frac{n}{2} = \sum_k \mathcal{G}(k)\, e^{i\omega_n 0^+}. \tag{7}$$

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:

$$\frac{n}{2} = \sum_k \mathbf{g}(k)\, e^{i\omega_n 0^+}. \tag{8}$$

In contrast to canonical theories, here we have two independent sum-rule constraints requiring two Lagrange multipliers. The first Lagrange multiplier $\boldsymbol{\mu}'$ is a standard chemical potential in that it sits next to the band energies $\overline{\varepsilon}_k$ 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 Φ and Ψ. Neither of these Lagrange multipliers is the physical thermodynamic chemical potential of the grand canonical ensemble. The physical chemical potential $\mu_{\rm phys}$, denoted by $\boldsymbol{\mu}$, can be obtained as a function of $\boldsymbol{\mu}'$ and u_0, as shown in Eq. (179) of (I):

$$\begin{aligned}\boldsymbol{\mu} = \boldsymbol{\mu}' &+ u_0 \frac{\lambda n}{2}\left(1 - \frac{\lambda n}{4}\right) \\ &- \left[J_0 \frac{\lambda n}{4}\left(1 - \frac{\lambda n}{2}\right) + 2\lambda\left(1 - \frac{\lambda n}{8}\right)\sum_q \varepsilon_q \mathbf{g}(q)\right] \\ &+ O(\lambda^3).\end{aligned} \tag{9}$$

We now discuss the criterion for validity of equations to a second order in λ. 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 λ. 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 λ. 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:

$$\rho_{\mathcal{G}}(k,\omega) = -\frac{1}{\pi}\mathrm{Im}\,[\mathcal{G}(k, i\omega_n \to \omega + i0^+)]. \tag{10}$$

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

$$\mathrm{Re}\,\mathcal{G}(k,\omega) = \mathrm{P.V.}\int_{-\infty}^{\infty} \frac{\rho_{\mathcal{G}}(k,\nu)}{\omega - \nu} d\nu. \tag{11}$$

An analogous definition is given for spectral representation $\rho_{\mathbf{g}}(k,\nu)$, $\rho_{\overline{\Phi}}(k,\nu)$, $\rho_{\Psi}(k,\nu)$ used for $\mathbf{g}$, $\overline{\Phi}$, Ψ, 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

$$\begin{aligned}\rho_{\mathcal{G}}(k,\omega) = \rho_{\mathbf{g}}(k,\omega)&\left(1 - \frac{n}{2} + \frac{n^2}{4} + \mathrm{Re}\,\Psi(k,\omega)\right) \\ &+ \rho_{\Psi}(k,\omega)\,\mathrm{Re}\,\mathbf{g}(k,\omega),\end{aligned} \tag{12}$$

so the two sum rules Eq. (7) and Eq. (8) can be written as

$$\begin{aligned}\frac{n}{2} &= \sum_k \int d\omega \rho_{\mathbf{g}}(k,\omega) f(\omega), \\ \frac{n^2}{4}\left(1 - \frac{n}{2}\right) &= -\sum_k \int d\omega f(\omega)(\rho_{\mathbf{g}}(k,\omega)\,\mathrm{Re}\,\Psi(k,\omega) \\ &\qquad + \mathrm{Re}\,\mathbf{g}(k,\omega)\,\rho_{\Psi}(k,\omega)),\end{aligned} \tag{13}$$

where $f(\omega) = [1 + \exp(\beta\omega)]^{-1}$ and $\overline{f}(\omega) = 1 - f(\omega)$. The auxiliary spectral function is in the usual Dysonian form,

$$\rho_{\mathbf{g}}(k,\omega) = \frac{\rho_{\overline{\Phi}}(k,\omega)}{\{\omega + \boldsymbol{\mu}' - \overline{\varepsilon}_k - \mathrm{Re}\,\overline{\Phi}(k,\omega)\}^2 + (\pi\rho_{\overline{\Phi}})^2}. \tag{14}$$

Using Eqs. (1)–(6), we express the spectral functions for Ψ and $\overline{\Phi}$ as

$$\begin{aligned}\rho_{\overline{\Phi}}(k,\omega) = \frac{1}{N_s^2}\sum_{pq}\int d\nu_1 d\nu_2\, &\rho_{\mathbf{g}}(p,\nu_1)\rho_{\mathbf{g}}(q,\nu_2)\rho_{\mathbf{g}}(p+q-k,\nu_1+\nu_2-\omega) \\ &\times\{f(\nu_1)f(\nu_2)\bar{f}(\nu_1+\nu_2-\omega) + \bar{f}(\nu_1)\bar{f}(\nu_2)f(\nu_1+\nu_2-\omega)\} \\ &\times(\varepsilon_p + \varepsilon_{k+q-p} + \varepsilon_k + \varepsilon_q + J_{k-p} - u_0)\left\{\varepsilon_k + \varepsilon_p + \varepsilon_q + \varepsilon_{k+q-p} + \tfrac{1}{2}(J_{k-p} + J_{k-q}) - u_0\right\},\end{aligned} \tag{15}$$

$$\rho_\Psi(k,\omega) = \frac{1}{N_s^2}\sum_{pq}\int d\nu_1 d\nu_2\ \rho_{\mathbf{g}}(p,\nu_1)\rho_{\mathbf{g}}(q,\nu_2)\rho_{\mathbf{g}}(p+q-k,\nu_1+\nu_2-\omega)$$
$$\times\{f(\nu_1)f(\nu_2)\bar{f}(\nu_1+\nu_2-\omega)+\bar{f}(\nu_1)\bar{f}(\nu_2)f(\nu_1+\nu_2-\omega)\}$$
$$\times(\varepsilon_p+\varepsilon_{k+q-p}+\varepsilon_k+\varepsilon_q+J_{k-p}-u_0). \quad (16)$$

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 $1 \leqslant i,j \leqslant L$ 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 $-\frac{1}{2}\omega_c \leqslant \omega \leqslant \frac{1}{2}\omega_c$, 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 $\rho_{\overline{\Phi}} \rightarrow \rho_{\overline{\Phi}} + \frac{\eta}{\pi}$.

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:

$$m_k \equiv \langle \hat{C}^\dagger_{k\sigma}\hat{C}_{k\sigma}\rangle = \int_{-\infty}^{\infty}\rho_{\mathcal{G}}(\vec{k},\omega)f(\omega)d\omega. \quad (17)$$

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 Re $\mathcal{G}(k,0)$, also given in terms of the spectral function by

$$\text{Re}\ \mathcal{G}(\vec{k},0) = P.V.\int_{-\infty}^{\infty}\frac{\rho_{\mathcal{G}}(\vec{k},\omega)d\omega}{\omega}. \quad (18)$$

At $T = 0$ the FS in $\vec{k}$ space is traced out by 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 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[13] 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[14] 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.[15–17] 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.

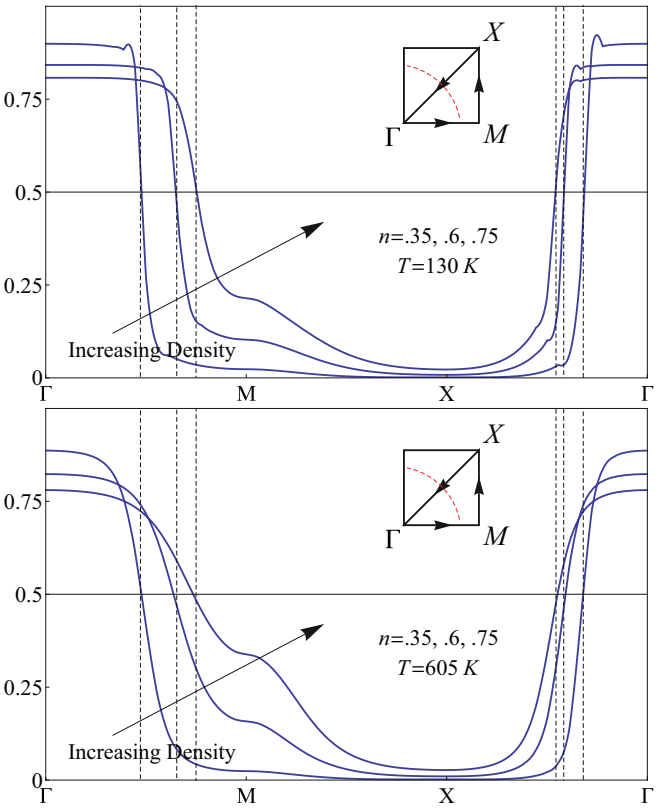

FIG. 1. (Color online) The momentum distribution function m_k is plotted along three principle lines of the BZ. The left and right figures are at 130 and 605 K, respectively. In each case the FS is the same as in the noninteracting problem. The Luttinger-Ward crossing Re $\mathcal{G}^{-1}(\vec{k},0) = 0$ 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 $m_k = \frac{1}{2}$.

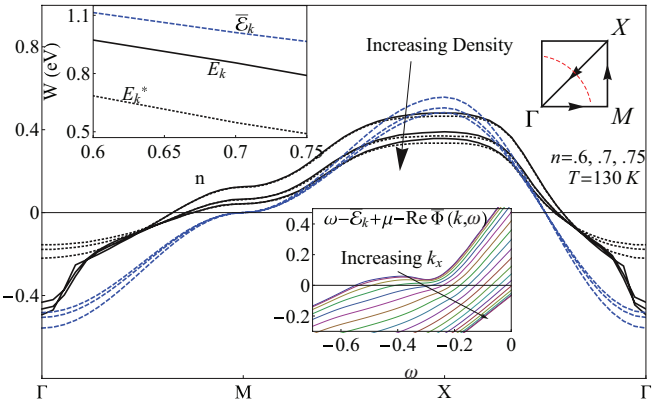

FIG. 2. (Color online) $T = 130$ K. The three dispersions defined in Eq. (19) are plotted along principle directions for three different densities. The upper insets show the bandwidth of the dispersions as a function of the density. The bare bandwidth is 2 eV, but each of these dispersions shrinks compared to that scale. The bandwidth renormalization due to Re $\overline{\Phi}$ in Eq. (19) is k dependent, and so E_k has a different shape than ϵ_k. Note that $E_k \sim E_k^*$ near the FS. However, E_k^* differs from E_k near the Γ 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 ω 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.

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.[27] To understand the effect of the many-body renormalizations, it is fruitful to study three dispersion relations defined in Ref. 4:

$$\begin{aligned} \overline{\varepsilon}_k &= \left(1 - n + \frac{3n^2}{8}\right)\varepsilon_k + \frac{1}{2}\sum_q J_{k-q}\, m_q, \\ E_k &= \overline{\varepsilon}_k - \boldsymbol{\mu}' + \mathrm{Re}\,\overline{\Phi}(k,E_k), \\ E_k^* &= \max[\rho_{\mathcal{G}}(k,\omega) : \omega]. \end{aligned} \tag{19}$$

Here $\overline{\varepsilon}_k$ 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 ω and by scanning k. The so-obtained peak locations yield the fourth dispersion spectrum E_{MDC}. To obtain E_{MDC} in practice, one may invert the MDC peak locations through

$$k^*(\omega) = \max[\rho_{\mathcal{G}}(k,\omega) : k], \quad E_{MDC}(k) = \text{Inverse of } k^*(E). \tag{20}$$

It is worth mentioning that the high-energy kink (or the waterfall) is experimentally defined as the the peeling off of the $E_{MDC}(k)$ from the $E_{EDC}(k) = E_k^*$ spectra.[29]

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 ϵ_k is 2 eV for both cases. Near the FS we see that $E_k \approx E_k^*$, but they differ near the Γ 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 Γ point. Since E_k is defined as the root of Re $\mathbf{g}^{-1}(k,E_k) = 0$, we plot $\omega + \boldsymbol{\mu}' - \overline{\varepsilon}_k - \mathrm{Re}\,\overline{\Phi}(k,\omega)$ at various k as a function of ω in the inset of Fig. 2. A strong ω dependence of Re $\overline{\Phi}(k,\omega)$ 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 Γ point where $k = (0,0)$ the QP becomes incoherent and the bulk of its spectral weight is

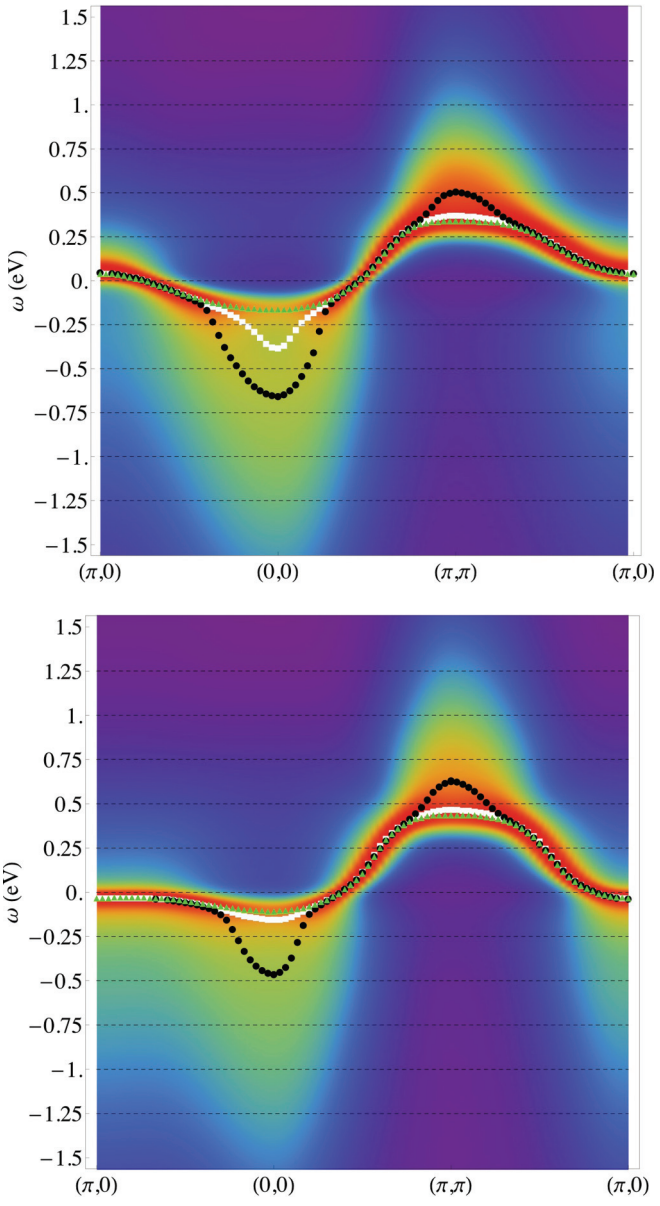

FIG. 3. (Color online) $L = 60$, $(n,T) = (0.75,300)$ K. Density plot of $A(k,\omega)$ of the minimal model (top) and the refined model Ref. 28 (bottom). (Here and below, red denotes high intensity and blue denotes low intensity). E_k, E_k^*, and $E_{MDC}(k)$ spectra are white, green, and black, respectively. 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.

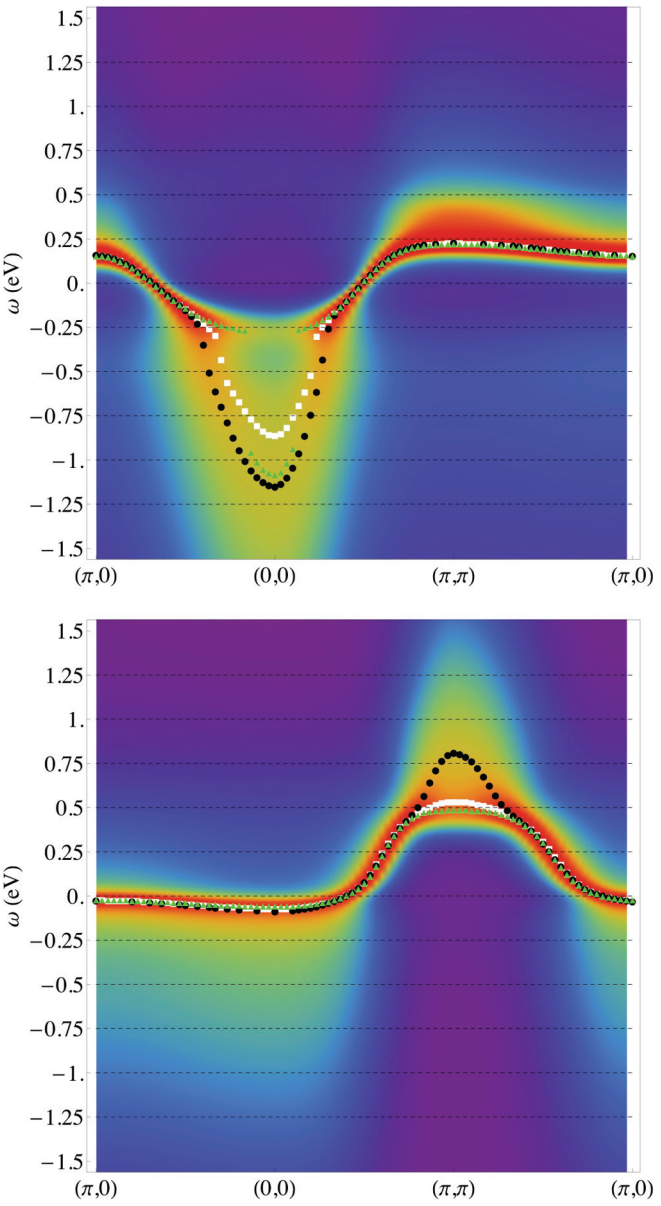

FIG. 4. (Color online) $L = 60$, $(n,T) = (0.75,300)$ K. (Top) $t'/t = 0.4$ is used to model electron-doped high-T_c superconductors. The kink feature is prominent here. (Bottom) Uses $t'/t = -0.4$ to crudely model a holelike FS. In this case the kink near (0,0) is lost, unlike in Fig. 3, correlating with a flat (bare) band dispersion.

spread out to high negative frequencies. In this region E_{MDC} differs considerably from E_k^* and recovers the scale of the bare dispersion ϵ_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 Γ 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 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 Γ 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 Γ. 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,[3] 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)

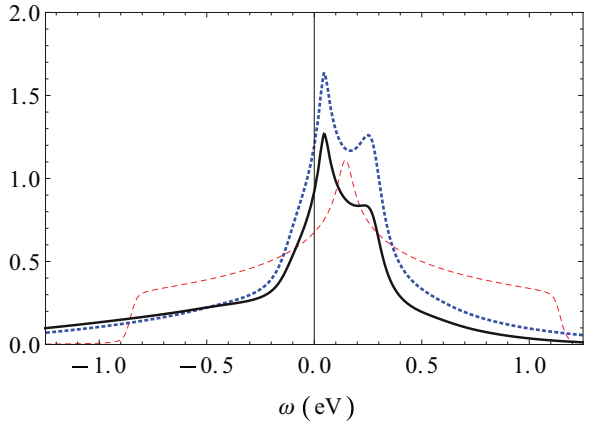

FIG. 5. (Color online) $n = 0.75$. The LDOS of the physical $\mathcal{G}$ (auxiliary $\mathbf{g}$) is in black (dotted blue), and the bare DOS is the dashed red curve. The renormalized band displays narrowing, and a long tail at $\omega < 0$. The LDOS develops a second spectral peak for $\omega > 0$ from a strongly k-dependent feature in the self-energy.

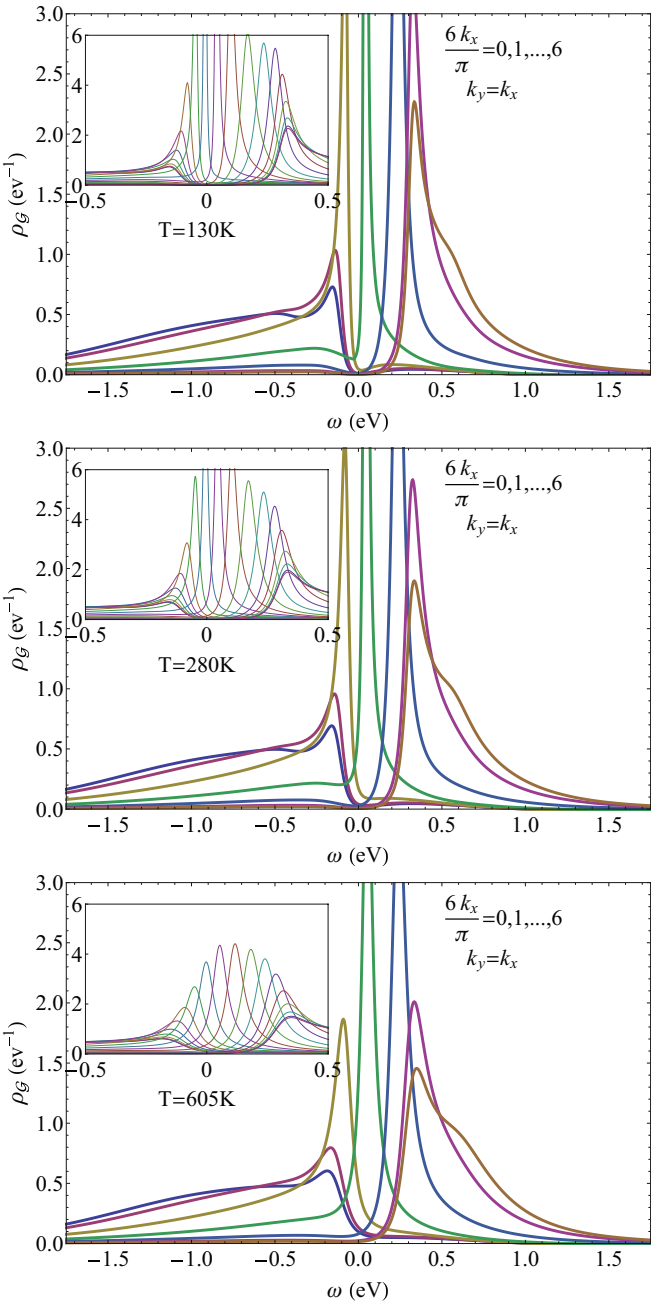

FIG. 6. (Color online) $n = 0.75$. The spectral function $\rho_{\mathcal{G}}$ [$= A(k,\omega)$] at several k points along the $\langle 11\rangle$ direction and T. We used $L_x = 36$; the insets show all positive k_x's and the main figures display a third of the allowed k_x's. The inset in each case zooms 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 Γ 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 **g**:

$$\mathrm{Re}\sigma(\Omega) = \frac{1}{\Omega}\sum_k v_k^2 \int \rho_{\mathbf{g}}(k,\omega)\rho_{\mathbf{g}}(k,\Omega+\omega)\, d\omega \times [f(\omega) - f(\Omega+\omega)]. \quad (21)$$

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 σ. 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

$$\frac{1}{\tau_\sigma} = \frac{4}{\pi}\int_0^{1/\tau_\sigma} \mathrm{Re}\,\sigma(\omega)/\sigma(0) d\omega, \quad (22)$$

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 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 ω, and the phase angle falls off with ω at about 4000 cm^{-1}. At optimum doping, the phase angle is known experimentally to be flat in ω over a much larger range,[32] and differs from the present calculation, whose validity is confined to overdoping. Experimental measurements in the overdoped case of the phase

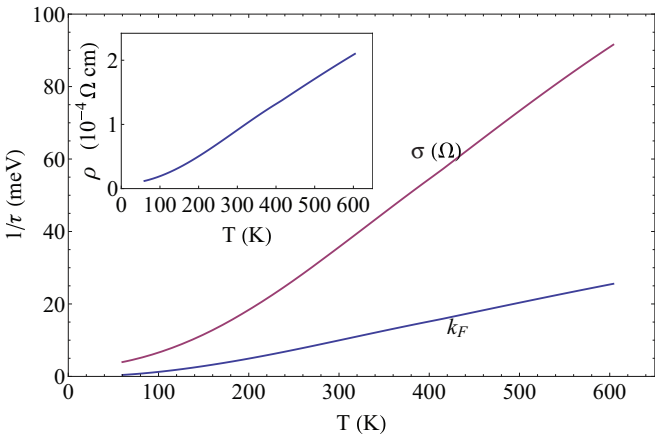

FIG. 7. (Color online) $n = 0.75$. The QP relaxation rate at the FS along $\langle 11\rangle$ obtained from $\rho_\Sigma(k, E_k^*)$, and the rate obtained from the optical conductivity as in Eq. (22). The T^2 behavior of an FL is visible at low temperature, crossing over at a modest temperature ($\sim$150) 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 Re $\sigma(\omega)$ curves along with optical conductivity measurements published by Puchkov *et al.*[12] 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.[25] These results are in qualitative agreement with resistivity measurements in the cuprates and other strongly

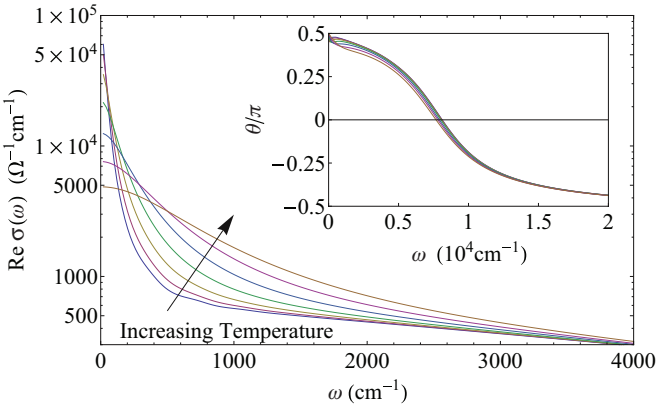

FIG. 8. (Color online) $n = 0.75$, $T = 60,90,130,190,280,410,605$ K. The optical conductivity is calculated on an absolute scale and illustrates how increasing T rapidly fills up the regime $200 \leqslant \omega \leqslant 1000$ cm^{-1}. The rise of conductivity at very low ω is also inferred from the dc resistivity displayed in Fig. 7. The phase of the complex σ falls off rapidly beyond 4000 cm^{-1}.

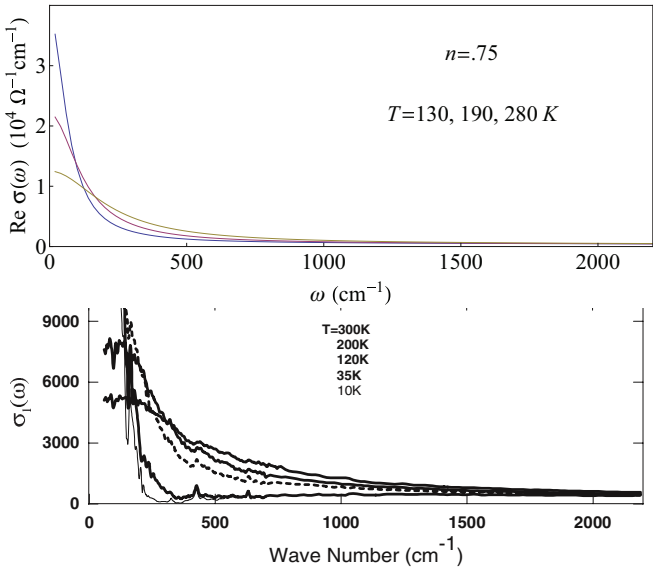

FIG. 9. (Color online) An explicit comparison of optical conductivity with measurements of Puchkov *et al.* from Ref. 12 with the author's kind permission. The data pertains to an overdoped thallium-based cuprate with $T_c = 23$ K, with a density $n \approx 0.75$. We note the similarity of magnitude and variation with ω 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,[27] both $\boldsymbol{\mu}'$ and u_0 rise linearly with T at high temperature. Due to the explicit appearance of u_0 in the expressions for Φ and Ψ, *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 $\rho_{\overline{\Phi}}$ and ρ_Ψ. Both functions exhibit the ω^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

$$\mathcal{G} = \frac{a_\mathcal{G}}{x - \Sigma}, \tag{23}$$

where $a_\mathcal{G}$ is the total spectral weight of the physical $\mathcal{G}$ and $x = \omega + \boldsymbol{\mu}' - \overline{\varepsilon}_k$ such that

$$\Sigma = x + \frac{a_\mathcal{G}}{a_\mathcal{G} + \Psi}(\overline{\Phi} - x). \tag{24}$$

In Fig. 11 we plot the computed imaginary part of the D-M self-energy, ρ_Σ. It exhibits a similar magnitude and k dependence at

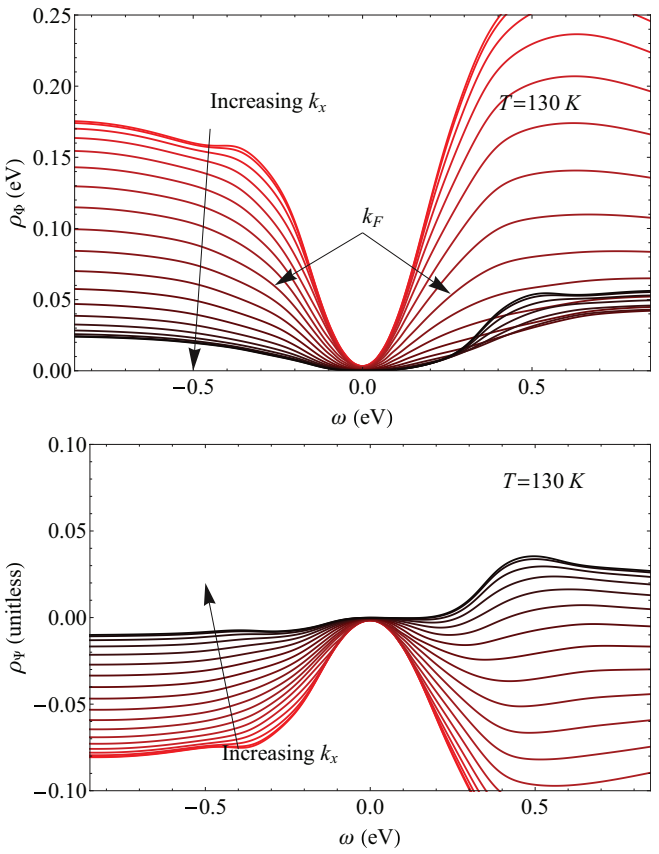

FIG. 10. (Color online) $(n,T) = (0.75,130)$ K. The spectral functions for the two self-energies Φ and Ψ, i.e., $\rho_{\overline{\Phi}}$ (top) and ρ_Ψ (bottom), at several k points along the $\langle 11 \rangle$ direction. Both are roughly quadratic and symmetric at low frequency but have a strongly k-dependent curvature. In the plot of $\rho_{\overline{\Phi}}$, the minimum width η chops off the bottom of the low-frequency minimum.

low frequency to that in $\rho_{\overline{\Phi}}$. 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[4,5] and also in a recent DMFT study of the Hubbard model.[25] 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, ρ_Σ 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 $\varepsilon_{\vec{k}}$, 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 ($|\omega| \leqslant 75$ meV) fit parameters of ρ_Σ 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

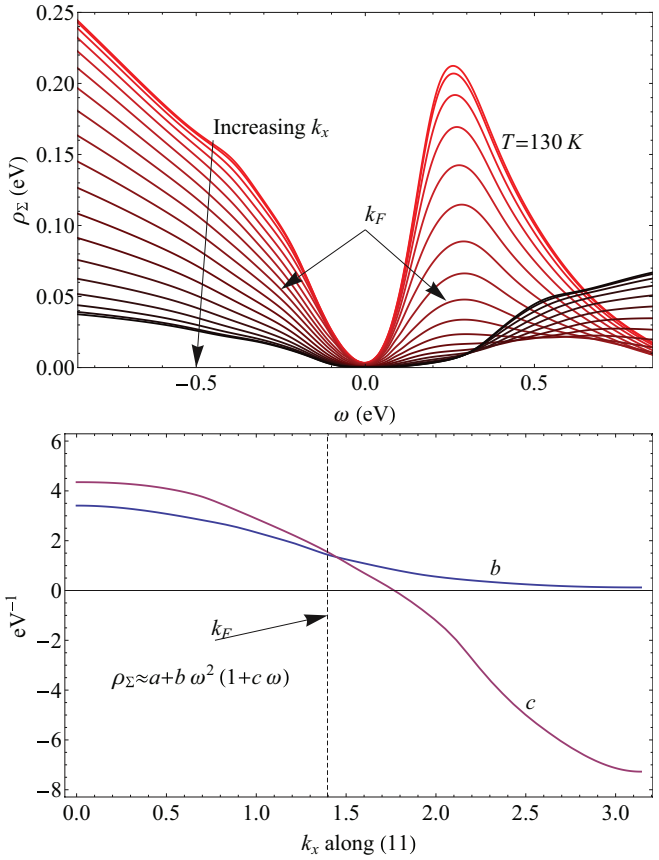

FIG. 11. (Color online) $(n,T) = (0.75,130)$ K. (Top) The spectral function ρ_Σ of the Dyson-Mori self-energy Σ from Eq. (24), at several k points along the $\langle 11 \rangle$ direction. As with $\rho_{\overline{\Phi}}$, ρ_Σ has inherited a strong k dependence. (Bottom) k dependence of the fit parameters from $\rho_\Sigma = a + b\,\omega^2(1 + c\,\omega)$ at low frequencies $|\omega| \leqslant 75$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.

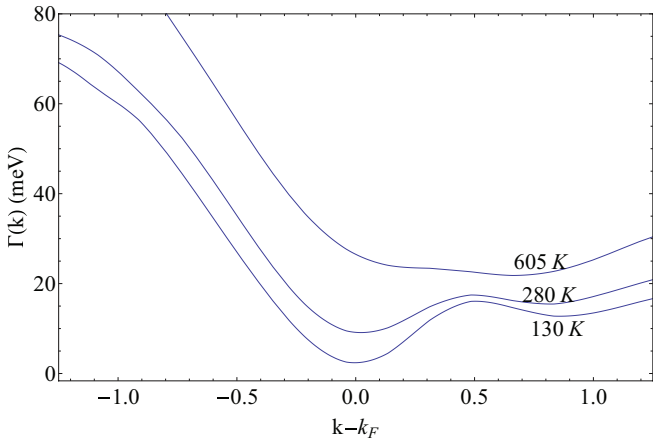

FIG. 12. (Color online) $n = 0.75$. The decay rate of QP near k_F along the nodal line from $\Gamma_k = \rho_\Sigma(k, E_k^*)$. The strong and T-dependent asymmetry makes quasiparticles longer lived at $k > k_F$. With increased T the minimum of Γ moves to $k > k_F$.

summarized by the expression

$$\Gamma(k) \sim b_f \left(1 - \left|\frac{b'_f}{b_f}\right| (k - k_F)\right) V_F^2\,(k - k_F)^2, \qquad (25)$$

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 Γ 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.[27]

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 λ 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 ω. 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 Γ 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.

[1] B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011).
[2] B. S. Shastry, Phys. Rev. B **87**, 125124 (2013).
[3] G.-H. Gweon, B. S. Shastry, and G. D. Gu, Phys. Rev. Lett. **107**, 056404 (2011).
[4] K. Matsuyama and G.-H. Gweon, arXiv:1212.0299.
[5] B. S. Shastry, Phys. Rev. B **84**, 165112 (2011).
[6] B. S. Shastry, Phys. Rev. Lett. **109**, 067004 (2012).
[7] P. Nozières, *Theory of Interacting Fermi Systems* (W. A. Benjamin, Amsterdam, 1964).
[8] A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).
[9] G. D. Mahan, *Many-Particle Physics* (Plenum Press, New York, 1981).
[10] M. Bejas, A. Greco, and A. Foussats, Phys. Rev. B **73**, 245104 (2006); E. Cappelluti and R. Zeyher, *ibid.* **59**, 6475 (1999).
[11] 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.
[12] A. V. Puchkov, D. Basov, and T. Timusk, J. Phys.: Condens. Matter **8**, 10049 (1996).
[13] W. Stephan and P. Horsch, Phys. Rev. Lett. **66**, 2258 (1991).
[14] R. R. P. Singh and R. L. Glenister, Phys. Rev. B **46**, 14313 (1992).
[15] C. Gros, R. Joynt, and T. M. Rice, Phys. Rev. B **36**, 381 (1987).
[16] B. Edegger, V. N. Muthukumar, and C. Gros, Adv. Phys. **56**, 927 (2007).
[17] A. Paramekanti, M. Randeria, and N. Trivedi, Phys. Rev. Lett. **87**, 217002 (2001).
[18] M. M. Zemljic and P. Prelovsek, Phys. Rev. B **72**, 075108 (2005).
[19] M. M. Zemljic, P. Prelovsek, and T. Tohyama, Phys. Rev. B **76**, 012502 (2007); Phys. Rev. Lett. **100**, 036402 (2008).
[20] J. Kokalj and P. Prelovsek, Phys. Rev. B **75**, 045111 (2007).
[21] J. Jaklic and P. Prelovsek, Adv. Phys. **49**, 1 (2000).
[22] 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).
[23] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).
[24] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).
[25] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, Phys. Rev. Lett. **100**, 086401 (2013).
[26] E. Khatami, D. Hansen, E. Perepelitsky, M. Rigol, and B. S. Shastry, Phys. Rev. B **87**, 161120(R) (2013).
[27] 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.
[28] R. S. Markiewicz, S. Sahrakorpi, M. Lindroos, H. Lin, and A. Bansil, Phys. Rev. B **72**, 054519 (2005).
[29] 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).
[30] L. Hozoi, M. S. Laad, and P. Fulde, Phys. Rev. B **78**, 165107 (2008).
[31] 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).
[32] 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[1]

[1] *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, $\boldsymbol{\mu}'$ and u_0, as well as the physical chemical potential, $\boldsymbol{\mu}$, 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 $\boldsymbol{\mu} \approx \boldsymbol{\mu}'$ 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, $\boldsymbol{\mu}$ grows faster than $\boldsymbol{\mu}'$. While $\boldsymbol{\mu}'$ approaches 0, a limit expected for a half-filled FL with particle-hole symmetry, $\boldsymbol{\mu}$ 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 $(\boldsymbol{\mu}', \boldsymbol{\mu}, u_0)$ 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 $\boldsymbol{\mu}'(T) = \boldsymbol{\mu}'(0) - bT^2$ (to focus on the quadratic term we subtract off the $T = 0$ intercepts in Fig. (14)). We can define a temperature scale $T_{\boldsymbol{\mu}'}$ which sets the strength of the $O(T^2)$ term according to

$$T_{\boldsymbol{\mu}'} = \sqrt{\left|\frac{\boldsymbol{\mu}'(0)}{b}\right|}. \tag{S.1}$$

where we have used the auxiliary "chemical potential" $\boldsymbol{\mu}'$ because it is most closely related to the FL aspects of the spectral function. $T_{\boldsymbol{\mu}'}$ is plotted in the inset as a function of the density n. Note the reduction of the scale of $T_{\boldsymbol{\mu}'}$ as we approach half filling. At high T ($> 600K$), $\boldsymbol{\mu}'$ rises linearly with T in typical FL fashion.

We also observe in Fig. (14) that the temperature dependence of $\boldsymbol{\mu}$ comes from $\boldsymbol{\mu}'$ 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, $\Delta\boldsymbol{\mu} \sim 20meV$ on heating from 100K to 300K, and about half this number for the change for $\Delta\boldsymbol{\mu}'$. This large variation should be readily measurable in ARPES, and appears to have been overlooked in most studies so far.

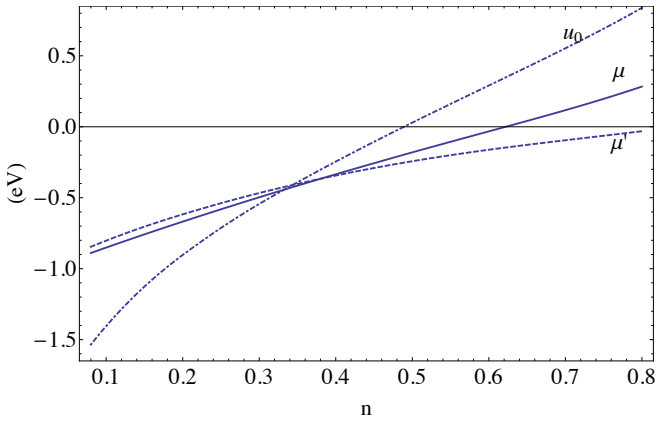

FIG. 13. $T = 280K$. The density dependence of the various chemical potentials, as defined in Eq. (9) of the main paper. The potentials $\boldsymbol{\mu}'$ and u_0, and the net physical chemical potential $\boldsymbol{\mu}$, dashed,dot-dashed, and solid, respectively. The potential $\boldsymbol{\mu}'$ approaches 0 as $n \to 1$, i.e. half filling, while the potential u_0 rises monotonically with density. The physical chemical potential $\boldsymbol{\mu}$ 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 $\boldsymbol{\mu}$ and u_0 corresponds to the scale of the renormalized bandwidth in the high density limit.

Unlike $\boldsymbol{\mu}'$, u_0 does not follow the standard behavior of a FL chemical potential. Whereas the low T temperature dependence of $\boldsymbol{\mu}'$ arises from small changes in occupation at low frequency only, u_0 feels *all frequencies* due to its explicit appearance in $\overline{\Phi}$ and Ψ. 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 Σ is the Dyson self energy. This definition does not immediately work for us as we do not have a conventional Dysonian form

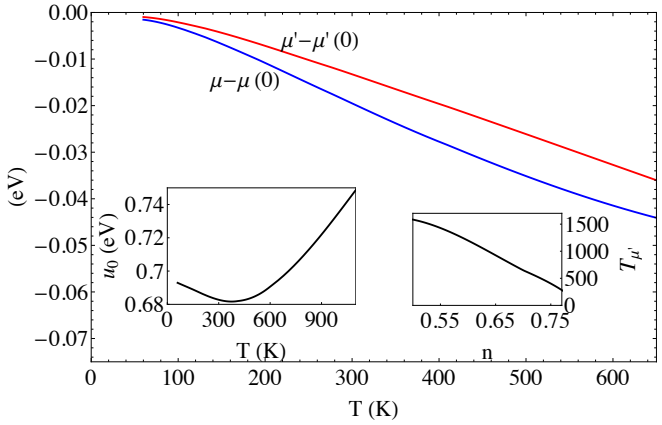

FIG. 14. $\boldsymbol{\mu}'$ and $\boldsymbol{\mu}$ versus T at $n = .75$. An effective Fermi temperature $T_{\mu'}$, found from Eq. (S.1), is shown as a function of the density. Note that the scale of variation of the $\boldsymbol{\mu}'$, $\Delta\boldsymbol{\mu}' \sim 10$ meV in heating from $T = 100$K to $T = 300$K, is quite large and is potentially observable in ARPES. A shrunken 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 $\boldsymbol{\mu}'$.

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

$$\mathcal{G} = \frac{a_{\mathcal{G}} + \Psi}{x - \overline{\Phi}} = \frac{a_{\mathcal{G}}}{x - \Sigma_{DM}} \tag{S.2}$$

where $a_{\mathcal{G}} = 1 - \frac{n}{2}$ exactly and $x = i\omega + \boldsymbol{\mu} - \bar{\epsilon}_k$. However in the present approximation $a_{\mathcal{G}} = 1 - \frac{n}{2} + \frac{n^2}{4}$ owing to the second order in λ approximation. In analogy to the standard FL we now define

$$Z_k = \frac{a_{\mathcal{G}}}{1 - \frac{\partial}{\partial\omega}\Sigma_{DM}}. \tag{S.3}$$

While it may be tempting to drop the factor $a_{\mathcal{G}}$ 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 $a_{\mathcal{G}}$ in the definition of Σ is a convenience employed for two reasons. First, it separates the effects of high and low energy physics formally. The factor $a_{\mathcal{G}}$ 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 $a_{\mathcal{G}}$ results in a self energy Σ_{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 then $1 - n$ 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.

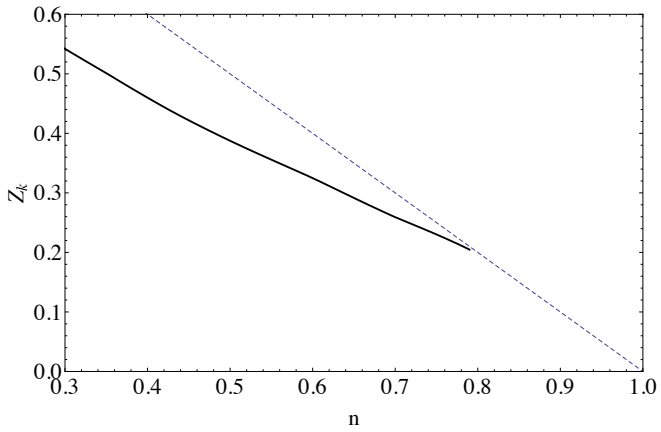

FIG. 15. The QP weight Z_{k_F} is plotted as a function of density. Z_{k_F} decreases with increased density, similar to (1-n) (which is depicted as a dotted line) but somewhat less. The incoherent spectral contribution is therefore already ~ 4 times greater than the QP part at $n \sim .7$, and this 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: $\boldsymbol{\mu}' = \boldsymbol{\mu}_0$, $\rho_\Psi = \rho_{\overline{\Phi}} = 0$, u_0 is set to a plausible value of 2t.

2. Build $\rho_{\mathbf{g}}$ from latest instance of $\boldsymbol{\mu}'$, $\overline{\Phi}$.

3. Calculate $\rho_{\overline{\Phi}}$ 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 $\rho_{\overline{\Phi}}$ have converged to specified tolerance.

6. Calculate ρ_{Ψ} 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 $\rho_{\overline{\Phi}}$. 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 $\rho_{\overline{\Phi}}$ 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 $\rho_{\overline{\Phi}}$ is also well converged. Thus, if $\boldsymbol{\mu}$ has successfully converged to this tolerance we can be sure the $\rho_{\overline{\Phi}}$ 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 Ψ. 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 $\eta \to \Delta\omega$.

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 $\rho_{\overline{\Phi}}$ and ρ_{Ψ} 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 $\rho_{\overline{\Phi}}$ and ρ_{Ψ} 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.

$$\Sigma(k)_{2^{nd}} \sim \sum_{pq} G(p)G(q)G(p+q-k); \quad \Sigma(i,j) \propto G(i,j)^2 G(j,i) \tag{S.4}$$

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", Ψ and Φ have the same frequency convolution structure of $\Sigma(k)_{2^{nd}}$ which appears only through the frequency arguments of $\mathbf{g}$. However, the $\rho_{\overline{\Phi}}$ and ρ_{Ψ} 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 (≈ 70) small FFT's of linear complexity.

To accomplish this we define several $\mathbf{gg}$ correlation functions which are similar to particle-hole bubbles.

$$\chi_0(Q) = \sum_q \mathbf{g}(q)\mathbf{g}(q+Q); \; \chi_1(Q) = \sum_q \varepsilon_q \mathbf{g}(q)\mathbf{g}(q+Q)$$

$$\chi_2(Q) = \sum_q \varepsilon_{q+Q} \mathbf{g}(q)\mathbf{g}(q+Q); \; \chi_3(Q) = \sum_q \varepsilon_{q+Q}^2 \mathbf{g}(q)\mathbf{g}(q+Q)$$

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 $\mathbf{ggg}$ term (except one to be discussed later) found in Ψ and $\overline{\Phi}$ can be written in the form

$$B_{\mathbf{ggg}}(k) = F_1(k) \sum_p F_2(p)\mathbf{g}(p)\chi_n(p-k)F_3(p-k). \tag{S.5}$$

where F_1,F_2, and F_3 are each functions of momentum only and their arguments are carefully matched with the arguments of $B_{\mathbf{ggg}}$, $\mathbf{g}$, and χ_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 $\rho_{\overline{\Phi}}$ into a convolution of one $\mathbf{g}$ and a χ_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_ω frequency bins which cover the support of our spectral functions. In extending the discretized frequency summations to infinity, we follow the standard procedure of padding the N_ω 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.

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 λ, 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 ω of the invariance of the Fermi volume, and at high ω of decaying as $\frac{c_0}{i\omega}$, with a correlation-depleted $c_0 < 1$. The effective Hamiltonian $H_{\rm eff}$ describing the auxiliary fermions is given a natural interpretation with an effective interaction $V_{\rm eff}$ 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 $V_{\rm eff}$ 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 λ 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 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 $U \gtrapprox Zt$, 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 $U \to \infty$.

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[6] 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[7]) 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 λ 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

 ©2013 American Physical Society

by the symbolic equation

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

where $\hat{G}_0^{-1}$ [Eq. (64)] is a noninteracting Green's function and γ [Eq. (35)] is essentially the spatially localized but time-dependent Green's function itself $\sim \mathcal{G}_{\rm local}$. Further, Y_1 is a Hartree-type energy and $X \sim$ (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 γ 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 γ term on the right of Eq. (1); it is a time-dependent Green's function obtainable from $\mathcal{G}$ itself [Eq. (45)]. This γ 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 $X \sim$ (something) $\times \frac{\delta}{\delta \mathcal{V}}$ to write

$$\begin{aligned} X \cdot \mathcal{G} \rightarrow \Sigma \mathcal{G} &\equiv (\text{something}) \times \mathcal{G}\Gamma\mathcal{G}, \\ \text{using } \frac{\delta}{\delta \mathcal{V}}\mathcal{G} &= \mathcal{G}\Gamma\mathcal{G}, \\ \text{following from } \Gamma &\equiv -\frac{\delta}{\delta \mathcal{V}}\mathcal{G}^{-1}, \end{aligned} \tag{2}$$

wherein the vertex Γ is introduced. This gives the Schwinger-Dyson relationship between the self-energy Σ and vertex:

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

This Schwinger-Dyson construction necessarily requires that the vertex Γ 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 γ 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 γ 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 σ, 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.[12] 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 μ 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 μ 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 μ below in Eq. (67) and Eq. (68).

(d) The μ 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

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

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

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

with $0 \leqslant \lambda \leqslant 1$, 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 γ 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 λ lives in a small and finite interval $[0,1]$. The expectation is that low-order perturbation expansion in λ 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 λ can be explicitly related to that of the fraction of double occupancy (and thus also density), so that tuning λ 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 μ 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 λ 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 $H_{\rm eff}$ for canonical electrons $(f_{i\sigma}, f^\dagger_{i\sigma})$, with the property that the (imaginary time) Heisenberg equation of motion for canonical electrons $\dot{f}_{i\sigma} = -[f_{i\sigma}, H_{\rm eff}]$ 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

$$[f_{i\sigma}, H_{\rm eff}] = ([\hat{c}_{i\sigma}, H_{t-J}])_{(\hat{c},\hat{c}^\dagger)\to(f,f^\dagger)} + (\text{expressions involving } f, f^\dagger \text{ that vanish at single occupancy}). \tag{6}$$

We can then add these missing terms with $(f, f^\dagger) \to (\hat{c}, \hat{c}^\dagger)$ 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.[13,14] One should therefore be able to use standard Feynman diagrams (Ref. 13) to compute the properties of this auxiliary theory in powers of λ, if one were so inclined.

We find it straightforward to find such an effective Hamiltonian $H_{\rm eff}$ [Eq. (27)] as described below in Sec. II C. The physical meaning of $H_{\rm eff}$ 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

$$\hat{c}^\dagger_{i\sigma} = X^{\sigma 0}_i \to f^\dagger_{i\sigma}(1 - n_{i\bar{\sigma}}), \quad \hat{c}_{i\sigma} = X^{0\sigma}_i \to f_{i\sigma}(1 - n_{i\bar{\sigma}}) \tag{7}$$

with $\bar{\sigma} = -\sigma$ and $n_{i\sigma} = f^\dagger_{i\sigma} f_{i\sigma}$. 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

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

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 λ here—it will turn out to be the same parameter as in Eq. (5)—and write

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

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

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 $H_{\rm eff}$ in Eq. (27). The introduction of the parameter λ can thus be viewed as replacing Eq. (7) by a "softer" representation of the Gutzwiller projection:

$$\hat{c}^\dagger_{i\sigma} \to f^\dagger_{i\sigma}(1 - \lambda n_{i\bar{\sigma}}), \quad \hat{c}_{i\sigma} \to f_{i\sigma}(1 - \lambda n_{i\bar{\sigma}}). \tag{11}$$

This λ representation discourages but does not completely eliminate double occupancy. However as $\lambda \to 1$, 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 λ in the range $0 \leqslant \lambda \leqslant 1$. Thus Eq. (11) suggests the interpretation of the parameter λ 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 $-\langle\!\langle (1 - n_{i\bar{\sigma}_i}) f_{i\sigma_i}, f^\dagger_{j\sigma_j}(1 - n_{i\bar{\sigma}_j})\rangle\!\rangle$, while $-\langle\!\langle f_{i\sigma_i}, f^\dagger_{j\sigma_j}\rangle\!\rangle$ would represent

the auxiliary Green's function $\mathbf{g}(i,j)$. The caparison factor μ seems hard to interpret in this language, though. The ECFL formalism developed here presents a procedure to splice together $\mathbf{g}$ and μ 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 $H_{\rm eff}$ and the emergence of the second chemical potential u_0.* In $H_{\rm eff}$ [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 $\varepsilon_k \to \varepsilon_k + u_t$, or adding an on-site interaction through $J_{ij} \to J_{ij} + \delta_{ij} u_J$, 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 $H_{\rm eff}$, 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 λ.

Imposing the shift symmetries on $H_{\rm eff}$ 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 $\boldsymbol{\mu}$, this u_0 is a Lagrange multiplier of a term in the Hamiltonian $H_{\rm eff}$. However it multiplies an interaction term that is *quartic* in the canonical fermions, unlike $\boldsymbol{\mu}$ that multiplies the usual (quadratic) number operator. The chemical potential $\boldsymbol{\mu}$ 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 λ 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 $H_{\rm eff}$ 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 μ. 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 λ 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 λ 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 $\hat{c}_{i\sigma} \to X_i^{0\sigma}$, $\hat{c}_{i\sigma}^\dagger \to X_i^{\sigma 0}$, and $\hat{c}_{i\sigma'}^\dagger \hat{c}_{i\sigma} \to X_i^{\sigma'\sigma}$. We study the t-J model given by

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

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:

$$t_{ij} \to t_{ij} - u_t \delta_{ij}, \quad J_{ij} \to J_{ij} + u_J \delta_{ij}, \tag{13}$$

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

$$H \to H + \left(u_t + \tfrac{1}{4} u_J\right) \hat{N}, \tag{14}$$

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 $\boldsymbol{\mu} \to \boldsymbol{\mu} + u_t + \frac{1}{4} u_J$. 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 $\boldsymbol{\mu}$ 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 $\mathcal{A}$:

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

with the usual imaginary-time Heisenberg picture τ dependence of the operators $Q(\tau) = e^{\tau H}Qe^{-\tau H}$, 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

$$\langle\!\langle Q(\tau_1,\tau_2,...)\rangle\!\rangle = \frac{\mathrm{Tr}[e^{-\beta H}T_\tau e^{-\mathcal{A}}\ Q(\tau_1,\tau_2,..)]}{\mathrm{Tr}[e^{-\beta H}T_\tau(e^{-\mathcal{A}})]}, \tag{16}$$

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

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

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

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

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

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

A. The Heisenberg equation of motion

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

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

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 $J_{ij} \to J_{ij} + u_J\delta_{ij}$ 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

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

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

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

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

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

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

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., $\sum_j \mathcal{V}_{ij}^{\sigma_1\sigma_2}f(\ldots,\tau_j,\ldots) \to \sum_{R_j}\int_0^\beta d\tau_j\mathcal{V}_{ij}^{\sigma_1\sigma_2}(\tau_j)\delta(\tau_i - \tau_j)f(\ldots,\tau_j,\ldots)$. 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

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

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

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

We recall from the introduction the discussion regarding suitably generalizing A of Eq. (22), in order to make connection with a Hermitian $H_{\rm eff}$, 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 **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

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

with a Hermitian effective potential $V_{\rm eff}^\dagger = V_{\rm eff}$ (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 terms that vanish under the constraint of single occupancy. The parameter λ 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 $J_{ij} \to J_{ij} - u_0\delta_{ij}$ 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

$$\mathcal{P}^\dagger(i,j) = \sum \sigma f_{i\sigma}^\dagger f_{j\bar{\sigma}}^\dagger = (f_{i\uparrow}^\dagger f_{j\downarrow}^\dagger - f_{i\downarrow}^\dagger f_{j\uparrow}^\dagger),$$
$$\mathcal{P}^\dagger(i,i) = \sum \sigma f_{i\sigma}^\dagger f_{i\bar{\sigma}}^\dagger = 2 f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger, \tag{28}$$

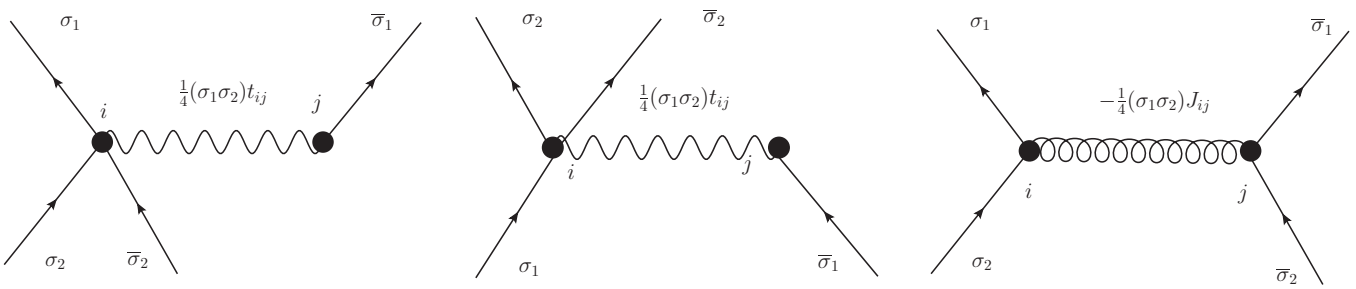

FIG. 1. The pseudopotential $V_{\rm eff}$ in the real-space representation, where the wavy line represents t_{ij} and the coiled line represents J_{ij}. The first two interaction vertices have two undisplayed symmetric partners with the exchange $i \leftrightarrow j$.

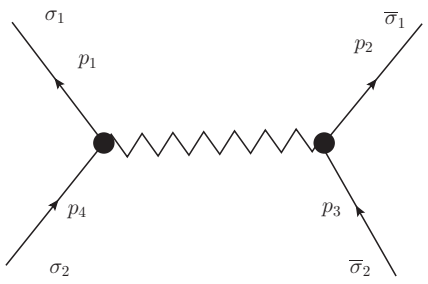

FIG. 2. The pseudopotential $W_{\rm eff}$ in the momentum-space representation. The zigzag line represents $W_{\rm eff}$. Note that the momentum transfer in the argument of J is also expressible as $J_{p_1-p_4}$.

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

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

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

$$\begin{aligned} H_{\rm eff} &= \sum_k \left(\varepsilon_k + \frac{1}{4}J_0 - \boldsymbol{\mu}\right) f^\dagger_{k\sigma} f_{k\sigma} \\ &\quad + \frac{\lambda}{4N_s}\sum_p (\sigma_1\sigma_2) W_{\rm eff}(p_1,p_2;p_3,p_4) \\ &\quad \times f^\dagger_{p_1\sigma_1} f^\dagger_{p_2\bar{\sigma}_1} f_{p_3\bar{\sigma}_2} f_{p_4\sigma_2}, \\ W_{\rm eff}(p_1,p_2;p_3,p_4) &= -\delta_{p_1+p_2,p_3+p_4}\{\varepsilon_{p_1} + \varepsilon_{p_2} + \varepsilon_{p_3} \\ &\quad + \varepsilon_{p_4} + J_{p_2-p_3} - u_0\} \end{aligned} \tag{30}$$

(see Fig. 2), where the momentum-independent term u_0 has been explicitly written out. In this effective Hamiltonian, the band energies ε_{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 $\boldsymbol{\mu}$ 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 $\boldsymbol{\mu} \to \boldsymbol{\mu} + u_t + \frac{1}{4}u_J$ and u_0 as

$$u_0 \to u_0 + 4u_t + u_J. \tag{31}$$

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

Since the standard notation for interaction reads $\sum\langle ab|V|a'b'\rangle f^\dagger_a f^\dagger_b f_{b'} f_{a'}$ for a conventional two-body interaction, our notation corresponds to writing $W_{\rm eff}(p_1,p_2;p_3,p_4) = \langle p_1 p_2|W|p_4 p_3\rangle$. Fermi symmetry implies the invariance $W_{\rm eff}(p_1,p_2;p_3,p_4) = W_{\rm eff}(p_2,p_1;p_4,p_3)$, and Hermiticity implies the invariance $W_{\rm eff}(p_1,p_2;p_3,p_4) = W_{\rm eff}(p_3,p_4;p_1,p_2)$. For this canonical theory, we calculate the commutator:

$$\begin{aligned} \left[f_{i\sigma_i}, H_{\rm eff}\right] &= -\sum_j t_{ij} f_{j\sigma_i} + \left(\frac{1}{4}J_0 - \boldsymbol{\mu}\right) f_{i\sigma_i} + \hat{A}_{i\sigma_i}, \\ \hat{A}_{i\sigma_i} &= \left[f_{i\sigma_i}, V_{\rm eff}\right], \end{aligned} \tag{32}$$

with

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

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^{ab}_i 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:

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

so that A_{i,σ_i} and $\hat{A}_{i,\sigma_i}$ contain terms that are in one to one correspondence. We will use Eq. (34) in 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., $\mathcal{G}_{\sigma_i\sigma_f}[i,f] \to [\mathcal{G}[i,f]]_{\sigma_i\sigma_f}$, so that we may regard $\mathcal{G}$ as a 2×2 matrix. In short, the space-time indices are displayed but the spin indices are hidden in the above matrix structure. We next define γ through

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

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 2×2 dimensional spin space.

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

$$\begin{aligned}\left\langle\!\left\langle \sigma_a\sigma_b X_i^{\bar\sigma_a\bar\sigma_b}(\tau_i)\mathcal{Q}\right\rangle\!\right\rangle &= \left(\gamma_{\sigma_a\sigma_b}[i] - D_{\sigma_a\sigma_b}[i]\right)\langle\!\langle\mathcal{Q}\rangle\!\rangle,\\ \left\langle\!\left\langle \sigma_a\sigma_b X_i^{\bar\sigma_a 0}(\tau_i^+)X_j^{0\bar\sigma_b}(\tau_i)\mathcal{Q}\right\rangle\!\right\rangle &= \left(\gamma_{\sigma_a\sigma_b}[i,j] - D_{\sigma_a\sigma_b}[i,j]\right)\langle\!\langle\mathcal{Q}\rangle\!\rangle,\end{aligned} \tag{36}$$

where we set $\tau_j \to \tau_i^-$ and define

$$\gamma_{\sigma_a\sigma_b}[i,j] = (\sigma_a\sigma_b)\mathcal{G}_{\bar\sigma_b\bar\sigma_a}[j\tau_i^-, i\tau_i] = \left\langle\!\left\langle \sigma_a\sigma_b X_i^{\bar\sigma_a 0}X_j^{0\bar\sigma_b}\right\rangle\!\right\rangle, \tag{37}$$

$$\gamma[i,i] = \gamma[i], \tag{38}$$

and

$$\begin{aligned} D_{\sigma_i\sigma_j}[i] &= \sigma_i\sigma_j\frac{\delta}{\delta\mathcal{V}_i^{\bar\sigma_i\bar\sigma_j}(\tau_i)},\\ D_{\sigma_i\sigma_j}[i,j] &= \sigma_i\sigma_j\frac{\delta}{\delta\mathcal{V}_{i,j}^{\bar\sigma_i\bar\sigma_j}(\tau_i)},\\ \text{and } D[i,i] &= D[i].\end{aligned} \tag{39}$$

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):

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

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

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

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

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

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

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

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 g AND THE CAPARISON FACTOR μ

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

$$\mathcal{G}[a,b] = \mathbf{g}[a,\mathbf{r}]\cdot\mu[\mathbf{r},b], \tag{44}$$

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

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

We define three-point vertex functions

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

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

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

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

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

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 τ_r, τ_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 Λ and $\mathcal{U}$ replaces the single vertex Γ of a canonical many-body system, and we will also find equations determining these below. Clearly in any exact treatment, the

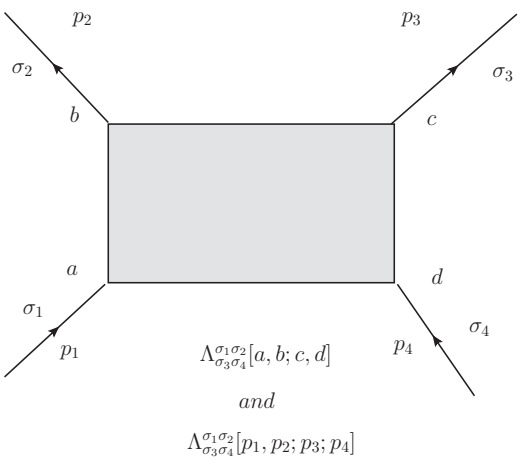

FIG. 3. The four-site vertex can be visualized from its definition for canonical theory: $-\langle\!\langle f_{a\sigma_1} f^\dagger_{b\sigma_2} f^\dagger_{c\sigma_3} f_{d\sigma_4}\rangle\!\rangle + \langle\!\langle f_{a\sigma_1} f^\dagger_{b\sigma_2}\rangle\!\rangle\langle\!\langle f^\dagger_{c\sigma_3} f_{d\sigma_4}\rangle\!\rangle = \frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_{cd}}\mathbf{g}_{\sigma_1\sigma_2}(ab) = \mathbf{g}_{\sigma_1\sigma_1'}[a,\mathbf{a}]\Lambda^{\sigma_1'\sigma_2'}_{\sigma_3\sigma_4}[\mathbf{a},\mathbf{b};c,d]\mathbf{g}_{\sigma_2'\sigma_2}[\mathbf{b},b]$. Therefore we may visualize that apart from the external legs, $\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[a,b;c,d] \sim \langle f_{a\sigma_1} f^\dagger_{b\sigma_2} f^\dagger_{c\sigma_3} f_{d\sigma_4}\rangle$. 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:

$$\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p,q;r) = \Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p,q;r,s\to r), \tag{49}$$

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 $\langle\!\langle X_i^{\sigma_1\sigma_2} X_i^{0\sigma_3}\ldots\rangle\!\rangle = 0$, 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

$$\cdots\xi^*_{\sigma_a\sigma_b}\cdots\frac{\delta}{\delta\mathcal{V}^*_{\mathbf{j}}} = \cdots\sigma_a\sigma_b\cdots\frac{\delta}{\delta\mathcal{V}^{\bar\sigma_a,\bar\sigma_b}_{\mathbf{j}}}, \tag{50}$$

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

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

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

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

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

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

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

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

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

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

Hence we may write Eq. (51) compactly as

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

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 μ as follows:

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

and

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

We write Eq. (43) as

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

We also need to process the object:

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

IV. ASSEMBLING THE EQUATIONS

Let us rewrite the three relevant equations symbolically:

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

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

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

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

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

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

Combining these we rewrite Eq. (60) symbolically as

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

Defining

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

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

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

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

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

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. (65) exactly into a pair of equations that are fundamental to the theory:

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

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

We can usefully invert Eq. (67) and write

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

We see that $\mathbf{g}$ satisfies a canonical equation, with a delta function of weight unity on the right, and μ 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 μ by a common function and its inverse. However, requiring $\mathbf{g}$ to be canonical fixes the function to be unity. The motivation of introducing λ in the above equations is to establish adiabatic, or more properly parametric, continuity with the Fermi gas.[16] 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 λ 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 λ 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 λ 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 λ 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 λ^r, say with $r\geqslant r_0$, 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 λ 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 λ 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 λ.

(5) Note that a λ 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 λ. While it is tempting to freeze this coefficient at the exact value, it would be inconsistent since we take its derivatives to find Ψ, 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 Φ and Ψ 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 Λ as in Ref. 11. We define the three nonvanishing matrix elements as $\Lambda^{(1)} = \Lambda^{\sigma\sigma}_{\sigma\sigma}$, $\Lambda^{(2)} = \Lambda^{\sigma\sigma}_{\bar\sigma\bar\sigma}$, 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 σ (such as in the expression for $\Phi_{\sigma\sigma}$ above):

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

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

$$\begin{aligned}\mathcal{G}[a,b] &= \sum_k e^{ik(a-b)}\mathcal{G}(k),\\ \Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[a,b;c] &= \sum_{p_1,p_2} e^{ip_1(a-c)+ip_2(c-b)}\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p_1,p_2),\\ \Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[a,b;c,d] &= \sum_{p_1+p_4=p_2+p_3} e^{i(p_1a-p_2b-p_3c+p_4d)}\\ &\quad\times\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p_1,p_2;p_3,p_4).\end{aligned} \tag{71}$$

The identity Eq. (49) in momentum space implies

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

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

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

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 $\omega_n = \pi(2n+1)k_BT$, the Green's functions at a fixed λ read

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

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

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

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

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 γ, Y_1, Φ, Ψ can be calculated at various orders in λ as demonstrated below.

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 λ 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.
$\mathbf{g}^{-1}[i,m]$	$\{\mathbf{g}_0^{-1}[i,m] + \lambda\gamma_i\cdot\mathcal{V}_{i,m} - \lambda\xi^*\cdot\mathcal{V}_{i,\mathbf{a}}\cdot\mathbf{g}[\mathbf{a},\mathbf{b}]\cdot\Lambda_*(\mathbf{b},m;i) - \lambda\, Y_1[i,m] - \lambda\Phi[i,m]\}$	Eq. (69)
$\mu[i,m]$	$\delta[i,m]\,(\mathbb{1} - \lambda\gamma[i]) + \lambda\,\Psi[i,m] + \lambda\xi^*\cdot\mathcal{V}_{i,\mathbf{a}}\cdot\mathbf{g}[\mathbf{a},\mathbf{b}]\cdot\mathcal{U}_*(\mathbf{b},m;i)$	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)
$\Phi[i,m]$	$-t[i,\mathbf{j}]\xi^*\cdot\mathbf{g}[\mathbf{j},\mathbf{c}]\cdot\left(\Lambda_*[\mathbf{c},m;i] + \frac{1}{2}\Lambda_*[\mathbf{c},m;\mathbf{j}]\right) - \frac{1}{2}t[i,\mathbf{k}]\,\xi^*\cdot\mathbf{g}[i,\mathbf{c}]\cdot\Lambda_*[\mathbf{c},m;\mathbf{k},i] + \frac{1}{2}J[i,\mathbf{k}]\,\xi^*\cdot\mathbf{g}[i,\mathbf{c}]\cdot\Lambda_*[\mathbf{c},m;\mathbf{k}]$	Eq. (56)
$\Psi[i,m]$	$-t[i,\mathbf{j}]\xi^*\cdot\mathbf{g}[\mathbf{j},\mathbf{c}]\cdot\left(\mathcal{U}_*[\mathbf{c},m;i] + \frac{1}{2}\mathcal{U}_*[\mathbf{c},m;\mathbf{j}]\right) - \frac{1}{2}t[i,\mathbf{k}]\,\xi^*\cdot\mathbf{g}[i,\mathbf{c}]\cdot\mathcal{U}_*[\mathbf{c},m;\mathbf{k},i] + \frac{1}{2}J[i,\mathbf{k}]\,\xi^*\cdot\mathbf{g}[i,\mathbf{c}]\cdot\mathcal{U}_*[\mathbf{c},m;\mathbf{k}]$	Eq. (57)

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.
$\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[i,m;\,j]$	$-\frac{\delta}{\delta\mathcal{V}_j^{\sigma_c\sigma_d}}\mathbf{g}^{-1}_{\sigma_a\sigma_b}[i,m]$	Eq. (46)
$\Lambda^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[i,m;\,j,k]$	$-\frac{\delta}{\delta\mathcal{V}_{j,k}^{\sigma_c\sigma_d}}\mathbf{g}^{-1}_{\sigma_a\sigma_b}[i,m]$	Eq. (48)
$\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[i,m;\,j]$	$\frac{\delta}{\delta\mathcal{V}_j^{\sigma_c\sigma_d}}\mu_{\sigma_a\sigma_b}[i,m]$	Eq. (46)
$\mathcal{U}^{\sigma_a\sigma_b}_{\sigma_c\sigma_d}[i,m;\,j,k]$	$\frac{\delta}{\delta\mathcal{V}_{j,k}^{\sigma_c\sigma_d}}\mu_{\sigma_a\sigma_b}[i,m]$	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

$$\begin{aligned}\Phi_{\sigma_i\sigma_m}[i,m] = &-t[i,\mathbf{j}]\sigma_i\sigma_1\mathbf{g}_{\sigma_1\sigma_2}[\mathbf{j},\mathbf{c}]\\ &\times\left(\Lambda^{\sigma_2\sigma_m}_{\bar\sigma_i\bar\sigma_1}[\mathbf{c},m;i]+\tfrac{1}{2}\Lambda^{\sigma_2\sigma_m}_{\bar\sigma_i\bar\sigma_1}[\mathbf{c},m;\mathbf{j}]\right)\\ &-\tfrac{1}{2}t[i,\mathbf{k}]\;\sigma_i\sigma_1\mathbf{g}_{\sigma_1\sigma_2}[i,\mathbf{c}]\Lambda^{\sigma_2\sigma_m}_{\bar\sigma_i\bar\sigma_1}[\mathbf{c},m;\mathbf{k},i]\\ &+\tfrac{1}{2}J[i,\mathbf{k}]\;\sigma_i\sigma_1\mathbf{g}_{\sigma_1\sigma_2}[i,\mathbf{c}]\Lambda^{\sigma_2\sigma_m}_{\bar\sigma_i\bar\sigma_1}[\mathbf{c},m;\mathbf{k}].\end{aligned}$$

VII. λ 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 λ expansion and the iterative process next. The iterations are analogous to the skeleton diagram expansion in standard many-body theory, where λ plays the role of the interaction constant. Various objects are expanded in terms of λ and $\mathbf{g}$, while $\mathbf{g}$ itself is left intact. Potentially confusing is the treatment of $\mathbf{g}^{-1}$, which *is expanded in* λ *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 useful. Consider $\gamma[i,m]$; we will expand it as

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

keeping $\mathbf{g}$ intact, i.e., unexpanded in λ. 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

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

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:

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

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

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

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 pth-order objects needed are in the second column. Since $\mathbf{g}^{-1}$ and μ 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
$[\mu[i,m]]_{p+1}$	$[\gamma[i],\,\Psi[i,m],\mathcal{U}[a,b;c],\,\mathcal{U}[a,b;c,d]]_p$
$\left[\mathbf{g}^{-1}[i,m]\right]_{p+1}$	$[Y_1[i,m],\Phi[i,m],\Lambda[a,b;c]]_p$

TABLE III. Iteration level p calculations. The auxiliary inverse Green's function $\mathbf{g}^{-1}$ and the adaptive spectral weight μ play the role of seed objects at the pth order. By computing them to pth order in the parameter λ, 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
$[\mu[i,m]]_p$	$[\gamma[i],\gamma[i,m],Y_1[i,m],\mathcal{U}[a,b;c],\,\mathcal{U}[a,b;c,d]]_p$
$[\mathcal{U}[a,b;c],\mathcal{U}[a,b;c,d]]_p$	$[\Psi[i,m]]_p$
$[\mathbf{g}^{-1}[i,m]]_p$	$[\Lambda[a,b;c],\Lambda[a,b;c,d]]_p$
$[\Lambda[a,b;c],\Lambda[a,b;c,d]]_p$	$[\Phi[i,m]]_p$

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

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

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

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

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

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

(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):

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

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

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

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

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

In using this expansion, one must first set $\lambda\to 1$. These expressions satisfy the shift theorem I.1 and shift theorem II, as one can verify by shifting ε_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_{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_{eff} is defined in Fig. 2.

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

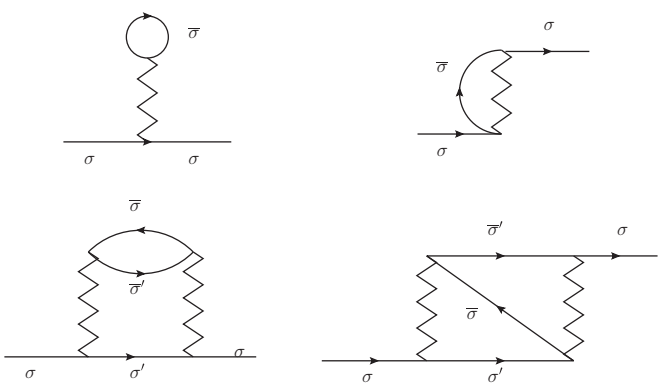

FIG. 4. The self-energy graphs to second order from W_{eff} and the effective Hamiltonian H_{eff}. These determine the Φ self-energy.

i.e., proceeds without requiring the knowledge of μ, and is represented in Feynman diagrammatic terms. The caparison term μ has no obvious interpretation in terms of $H_{\rm eff}$, 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 μ 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,[7,18] 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

$$\mathcal{G}[k] = \mathbf{g}[k]\mu[k]. \tag{89}$$

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

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

while the auxiliary fermion satisfies an identical sum rule,

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

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 Δ_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

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

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 $\mathbf{g}$,

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

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

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

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 λ. The existence of the Ω_Φ 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 Λ 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

$$t[i,j] \to t[i,j](1 + v[j] - v[i]), \tag{95}$$

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

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

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

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

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 σ)

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

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

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

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:

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

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

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

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

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

where we have defined the two Ward functions:

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

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

$$\mathcal{W}_g(p_1,p_2) = 0, \tag{104}$$

$$\mathcal{W}_\mu(p_1,p_2) = 0. \tag{105}$$

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 $\frac{\delta}{\delta\mathcal{V}}\mathbf{g} = \mathbf{g}\Lambda\mathbf{g}$ rather than $\frac{\delta}{\delta\mathcal{V}}\mathbf{g} = \mathbf{g}\mathbf{g}$, 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:

$$\begin{aligned}\mathcal{U}^{(\mathcal{T})}[p_1,p_2] &= -\lambda\sum_q \mathbf{g}[q]\Lambda^{(\mathcal{T})}(q,q+p_2-p_1) \\ &\quad \times \mathbf{g}[q+p_2-p_1] + O(\lambda^2), \\ \mathcal{U}^{(s)}[p_1,p_2] &= -\lambda\sum_q \mathbf{g}[q]\Lambda^{(s)}(q,q+p_2-p_1) \\ &\quad \times \mathbf{g}[q+p_2-p_1] + O(\lambda^2), \\ \mathcal{U}^{(t)}[p_1,p_2] &= \lambda\sum_q \mathbf{g}[q]\Lambda^{(t)}(q,q+p_2-p_1) \\ &\quad \times \mathbf{g}[q+p_2-p_1] + O(\lambda^2),\end{aligned} \tag{106}$$

and similarly for the Λ vertices:

$$\begin{aligned}\Lambda^{(\mathcal{T})}[p_1,p_2] &= (\varepsilon_{p_1} - \varepsilon_{p_2})(1 - \underbrace{\lambda n}) - \lambda\sum_q \mathbf{g}[q]\Lambda^{(\mathcal{T})}(q,q+p_2-p_1)\mathbf{g}[q+p_2-p_1]\mathcal{F}(q,p_1,p_2) + O(\lambda^2), \\ \Lambda^{(s)}[p_1,p_2] &= 1 - \lambda\sum_q \mathbf{g}[q]\Lambda^{(s)}(q,q+p_2-p_1)\mathbf{g}[q+p_2-p_1]\mathcal{F}(q,p_1,p_2) + O(\lambda^2), \\ \Lambda^{(t)}[p_1,p_2] &= 1 + \lambda\sum_q \mathbf{g}[q]\Lambda^{(t)}(q,q+p_2-p_1)\mathbf{g}[q+p_2-p_1]\mathcal{F}(q,p_1,p_2) + O(\lambda^2),\end{aligned} \tag{107}$$

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 $t[i,m] \to t[i,m](1 + v[m] - v[i])$ term in Eq. (B7). It is readily shown by examining the kernel of the integral equations that the solution for $\Lambda^{(\mathcal{T})}(p_1,p_2)$ 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),

$$\begin{aligned}\mathbf{g}[p] &= i\omega_n + \boldsymbol{\mu}' - (1-\lambda n)\varepsilon_k \\ &\quad + \frac{\lambda}{2}\sum_q J_{k-q}\mathbf{g}[q] + O(\lambda^2)\end{aligned} \tag{108}$$

$$\text{and} \quad \mu[p] = 1,$$

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 μ must be chosen to be unity rather than $1 - \lambda\frac{n}{2}$ 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 $S_a^z = \frac{1}{2}\sum_{\sigma_1,\sigma_2}\tau^z_{\sigma_1\sigma_2}X_a^{\sigma_1\sigma_2}$, where τ^z is the usual Pauli matrix. These can be obtained from taking the functional derivatives of the Green's function

$$\Upsilon^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[i,j] = \frac{\delta}{\delta\mathcal{V}_j^{\sigma_3\sigma_4}}\mathcal{G}_{\sigma_1\sigma_2}[i^-,i], \tag{109}$$

and can be conveniently found from taking a limit of the three-site object $\Upsilon^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p,q;r) = \frac{\delta}{\delta\mathcal{V}_r^{\sigma_3\sigma_4}}\mathcal{G}_{\sigma_1\sigma_2}[p,q]$. 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:

$$\begin{aligned}\langle\langle n_a(\tau_a)n_b(\tau_b)\rangle\rangle &= n^2 - 2\Upsilon^{(s)}(a,b), \\ \big\langle\big\langle S_a^z(\tau_a)S_b^z(\tau_b)\big\rangle\big\rangle &= -\tfrac{1}{2}\Upsilon^{(t)}(a,b).\end{aligned} \tag{110}$$

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 $Q \equiv (\vec{Q}, i\Omega_q)$ satisfies the relation:

$$\Upsilon^{(\alpha)}(Q) = \Upsilon^{(\alpha)}(-Q). \tag{111}$$

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 μ as before and we find

$$\begin{aligned}\Upsilon^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(a,b;r) &= \frac{\delta}{\delta \mathcal{V}_r^{\sigma_3\sigma_4}} \left\{\mathbf{g}_{\sigma_1\sigma_a}[a,\mathbf{a}]\mu_{\sigma_a\sigma_2}[\mathbf{a},b]\right\}, \\ &= \left\{\mathbf{g}[a,\mathbf{b}]\Lambda_{\sigma_3\sigma_4}(\mathbf{b},\mathbf{s};r)\mathbf{g}[\mathbf{s},\mathbf{a}]\mu[\mathbf{a},q]\right\}_{\sigma_1\sigma_2} \\ &\quad + \left\{\mathbf{g}[a,\mathbf{b}]\mathcal{U}_{\sigma_3\sigma_4}(\mathbf{b},b;r)\right\}_{\sigma_1\sigma_2}, \end{aligned} \tag{112}$$

where the vertex and Υ carry upper spin indices that are part of the matrix product. Turning off the sources, we find the expressions for singlet and triplet response

$$\begin{aligned}\Upsilon^{(\alpha)}(a,b;r) &= \mathbf{g}[a,\mathbf{b}]\Lambda^{(\alpha)}(\mathbf{b},\mathbf{s})\mathbf{g}[\mathbf{s},\mathbf{a}]\mu[\mathbf{a},b] \\ &\quad + \mathbf{g}[a,\mathbf{b}]\mathcal{U}^{\alpha}(\mathbf{b},b;r), \\ \Upsilon^{(\alpha)}(p_1,p_2) &= \mathbf{g}[p_1]\Lambda^{(\alpha)}(p_1,p_2)\mathbf{g}[p_2]\mu[p_2] \\ &\quad + \mathbf{g}[p_1]\mathcal{U}^{(\alpha)}(p_1,p_2), \end{aligned} \tag{113}$$

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,

$$\begin{aligned} &Q^{(1)} = Q^{\sigma\sigma}_{\sigma\sigma}, \quad Q^{(2)} = Q^{\sigma\bar{\sigma}}_{\bar{\sigma}\sigma}, \quad Q^{(3)} = Q^{\sigma\bar{\sigma}}_{\sigma\bar{\sigma}}, \\ &Q^{(a)} = Q^{(2)} - Q^{(3)}, \quad Q^{(s)} = Q^{(1)} + Q^{(2)}, \\ &Q^{(t)} = Q^{(1)} - Q^{(2)} = Q^{(3)}. \end{aligned} \tag{114}$$

The charge $\alpha = s$ and spin $\alpha = t$ susceptibilities at finite $Q \equiv (\vec{Q}, i\Omega_q)$ are given by setting $p_2 \to p$ and $p_1 \to p+Q$ and summing over p:

$$\begin{aligned}\Upsilon^{(\alpha)}(Q) &\equiv \sum_p \Upsilon^{(\alpha)}(p,p+Q) \\ &= \sum_p [\mathbf{g}[p]\Lambda^{(\alpha)}(p,p+Q)\mathbf{g}[p+Q]\mu[p+Q] \\ &\quad + \mathbf{g}[p]\mathcal{U}^{(\alpha)}(p,p+Q)]. \end{aligned} \tag{115}$$

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

$$\chi^{(\alpha)}_{\Lambda}(Q) \equiv -\sum_q \mathbf{g}[q]\Lambda^{(\alpha)}(q,q+Q)\mathbf{g}[q+Q], \tag{116}$$

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:

$$\mathcal{U}^{(\alpha)}[p_1,p_2] = \lambda\xi_\alpha \chi^{(\alpha)}_{\Lambda}(p_2 - p_1), \tag{117}$$

where ξ_α is 1 for $\alpha =$ singlet and -1 for $\alpha =$ triplet. Therefore we can sum over the p dependence of the second term and rewrite Eq. (115) as

$$(\Upsilon^{(\alpha)}(Q))_{RPA} = -C_\alpha \chi^{(\alpha)}_{\Lambda}(Q), \tag{118}$$

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 $C_{\text{singlet}} = (1-\lambda n) \to 1-n$ and $C_{\text{triplet}} = 1$. The vertices Λ 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 λ by iteration, where explicit results can be obtained. Let us define a few frequently occurring generalized polarizability functions for convenience. We will now reinstate $J_k \to J_k - u_0$:

$$\begin{aligned} \chi_0(Q) &= -\sum_q \mathbf{g}[q]\mathbf{g}[q+Q], \\ \chi_1(Q) &= -\sum_q \mathbf{g}[q]\mathbf{g}[q+Q]\{\varepsilon_q + \varepsilon_{q+Q}\}, \\ \chi_2(Q) &= \frac{1}{2}\sum_{r,p} \mathbf{g}[r]\mathbf{g}[r+Q]\mathbf{g}[p]\mathbf{g}[p+Q]J_{p-r}, \\ F(p+Q,p) &= \sum_r \mathbf{g}[r]\mathbf{g}[r+Q]\Big\{\varepsilon_p + \varepsilon_{p+Q} + \varepsilon_r + \varepsilon_{r+Q} \\ &\quad + \frac{1}{2}(J_Q + J_{p-r})\Big\}. \end{aligned} \tag{119}$$

Here $\chi_0(Q)$ is the standard Lindhard function and is positive in the static limit as $\vec{Q} \to 0$, while the other functions are generalizations thereof.

The answers are

$$\begin{aligned}\Upsilon^{(s)}(Q) &= -(1-\lambda n)\chi_0(Q) - \lambda\big[2\chi_0(Q)\chi_1(Q) \\ &\quad - \big(u_0 - \tfrac{1}{2}J_Q\big)\chi_0^2(Q) + \chi_2(Q)\big], \\ \Upsilon^{(t)}(Q) &= -\chi_0(Q) + \lambda\big[2\chi_0(Q)\chi_1(Q) \\ &\quad + \big(\tfrac{1}{2}J_Q - u_0\big)\chi_0^2(Q) + \chi_2(Q)\big]. \end{aligned} \tag{120}$$

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

$$[\mathcal{U}(Q)]_{\text{phen}} = -\frac{1}{1-n/2}\sum_q \mathcal{G}[q]\mathcal{G}[q+Q], \tag{121}$$

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 μ 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 λ 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 $H_{\rm eff}$, depending parametrically on λ. 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 $H_{\rm eff}$, and is therefore amenable to a perturbative expansion. The form of the $H_{\rm eff}$ 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 $H_{\rm eff}$ 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,

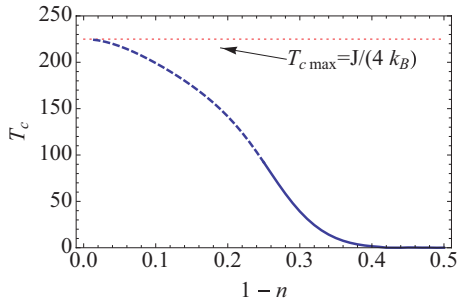

FIG. 5. (Color online) The transition temperature in kelvins, from solving Eq. (122) assuming $t = 3000$ K and $J = 900$ K. The solid line indicates the likely regime of validity of the $O(\lambda)$ theory. Its dotted extension to lower hole density is speculative and is most likely to change with higher 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 $\Delta(k)$:

$$\Delta(k) = \frac{1}{N_s} \sum_p \left\{ \varepsilon_k + \varepsilon_p - u_0 + \frac{1}{2} J_{k-p} \right\} \times \Delta(p) \frac{\tanh \beta E(p)/2}{E(p)}, \tag{122}$$

where $E(p) = \sqrt{\Delta^2(p) + \xi_p^2}$, and $\xi_p = \varepsilon_p(1-n) - \frac{1}{2}\sum J_{q-p} n_q - \boldsymbol{\mu}$. 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 λ from the atomic limit

In order to understand the role of λ we study the atomic limit $t, J \to 0$ where this parameter can be introduced into the physical Green's function in the form

$$\mathcal{G}[\lambda, i\omega_n] = \frac{1 - \lambda\frac{n}{2}}{i\omega_n + \boldsymbol{\mu}}, \tag{A1}$$

and study its dependence on λ in the interval [0,1]. The chemical potential $\boldsymbol{\mu}$ 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

$$\boldsymbol{\mu} = k_B T \ln\left(\frac{n}{2 - (1+\lambda)n}\right). \tag{A2}$$

Thermodynamics tells us that the entropy S can be expressed as

$$S(n) = -N_s \int_0^n dn' \, \frac{\partial \mu(n')}{\partial T}, \tag{A3}$$

and since we know μ from Eq. (A2) we obtain with $y = (1+\lambda)n$

$$\frac{S(n,\lambda)}{k_B N_s} = \frac{1}{1+\lambda}\{\ln 4 - y \ln n - (2-y)\ln(2-y)\}. \tag{A4}$$

We see that its λ 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 λ ranges over its domain $0 \leqslant \lambda \leqslant 1$. 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 λ 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 $\prod_i [1 - (1-g) n_{i\uparrow} n_{j\downarrow}]$ 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

$$\frac{S(n,d)}{k_B N_s} = -d \ln d - (n - 2d)\ln\left(\frac{n}{2} - d\right) - (1 + d - n)\ln(1 + d - n). \tag{A5}$$

An uncorrelated system corresponds to $d = \frac{n^2}{4}$, where the entropy Eq. (A5) is a maximum, while $d = 0$ for the fully projected t-J model. Comparing the two expressions for entropy Eq. (A4) and Eq. (A5), we can express λ in terms of d at any density. We have thus demonstrated that λ is a conjugate variable to the double-occupation density in this limit. Their explicit relationship is illustrated in Fig. 6.

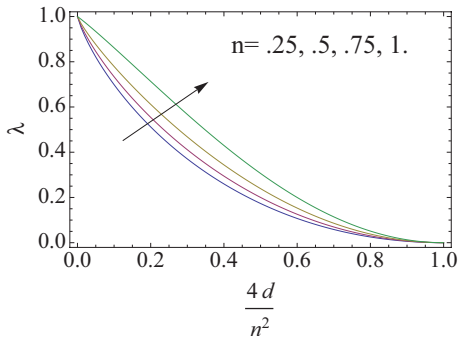

FIG. 6. (Color online) The parameter λ is determined in terms of the double occupancy d at various densities in the atomic limit as described in the text. The arrow indicates increasing density n. Note that the parameter d is scaled into the unit interval.

2. Expansion in λ in the atomic limit

In the atomic limit we set $t \to 0$ and $J \to 0$ so that Eq. (67) and Eq. (68) become

$$\begin{aligned} \mathbf{g}[i,f] &= \mathbf{g}_0[i,f;\boldsymbol{\mu}_a], \\ \mu[i,f] &= \delta[i,f](\mathbb{1} - \lambda\gamma[i]) - \lambda\boldsymbol{\mu}_b \mathbf{g}[i,\mathbf{f}]\cdot\mu[\mathbf{f},f]. \end{aligned} \tag{A6}$$

Here we split the chemical potential into two pieces $\boldsymbol{\mu} = \boldsymbol{\mu}_a + \lambda\boldsymbol{\mu}_b$. Thus in this limit $\mathbf{g}$ is the free Fermi Green's function independent of λ, and $\boldsymbol{\mu}_a$ is the free value $\boldsymbol{\mu}_a \to \boldsymbol{\mu}_0$, 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

$$\begin{aligned} \mathbf{g}[i\omega_n] &= \mathbf{g}_0[i\omega_n;\boldsymbol{\mu}_a] = \frac{1}{i\omega_n + \boldsymbol{\mu}_0}, \\ \mu[i\omega_n] &= \left(1 - \lambda\frac{n}{2}\right) - \lambda\boldsymbol{\mu}_b \mathbf{g}[i\omega_n]\mu[i\omega_n], \\ &= \frac{1 - \lambda\frac{n}{2}}{1 + \lambda\boldsymbol{\mu}_b \mathbf{g}[i\omega_n]}. \end{aligned} \tag{A7}$$

Thus the physical Green's function

$$\mathcal{G}[i\omega_n] = \frac{1 - \lambda\frac{n}{2}}{i\omega_n + \boldsymbol{\mu}_0 + \lambda\boldsymbol{\mu}_b}. \tag{A8}$$

We fix the chemical potentials from the number sum rule as usual and thus

$$\begin{aligned} \frac{n}{2} &= \frac{1}{1 + e^{-\beta\boldsymbol{\mu}_0}}, \\ \frac{n}{2} &= \left(1 - \lambda\frac{n}{2}\right)\frac{1}{1 + e^{-\beta(\boldsymbol{\mu}_0 + \lambda\boldsymbol{\mu}_b)}}. \end{aligned} \tag{A9}$$

We may then solve for $\boldsymbol{\mu}$'s in terms of the density and obtain

$$\begin{aligned} \boldsymbol{\mu}_0 &= k_B T \ln\left(\frac{n}{2-n}\right), \\ \lambda\boldsymbol{\mu}_b &= k_B T \ln\left(\frac{2-n}{2-(1+\lambda)n}\right). \end{aligned} \tag{A10}$$

Thus the chemical potential $\boldsymbol{\mu}_1$ has a power series representation

$$\boldsymbol{\mu}_b = \sum_{m=0}^{\infty} \lambda^m \boldsymbol{\mu}_b^{(m)} = k_B T \sum_{m=0}^{\infty} \frac{\lambda^m}{m+1} \left[\frac{n}{2-n}\right]^{m+1}. \tag{A11}$$

We see explicitly from Eq. (A11) that the λ 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

$$\begin{aligned} \mathbf{g}_0^{-1}[i,m] &= \big\{\big[\big(\boldsymbol{\mu} - \partial_{\tau_i} - \tfrac{1}{4}J_0\big)\mathbb{1} - \mathcal{V}_i\big] \\ &\quad \times \delta[i,m] + t[i,m] - \mathcal{V}_{i,m}\big\}, \\ [\mu[i,f]]_0 &= \mathbb{1}\delta[i,f]. \end{aligned} \tag{B1}$$

Derived objects:

$$\begin{aligned} [\gamma[i]]_0 &= \mathbf{g}^{(k)}[i,i], \\ [\gamma[i,m]]_0 &= \mathbf{g}^{(k)}[m,i], \\ \big[\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]\big]_0 &= 0, \\ \big[\mathcal{U}_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]\big]_0 &= 0. \end{aligned} \tag{B2}$$

$$\begin{aligned} [Y_1[i,m]]_0 &= t[i,m]\big(\mathbf{g}^{(k)}[i,i] + \tfrac{1}{2}\mathbf{g}^{(k)}[m,m]\big) \\ &\quad - \delta[i,m]\tfrac{1}{2}(J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] - t[i,\mathbf{j}]\mathbf{g}^{(k)}[i,\mathbf{j}]), \\ [\Psi[i,m]]_0 &= 0. \end{aligned} \tag{B3}$$

$$\begin{aligned} \big[\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j]\big]_0 &= \delta_{\sigma_a\sigma_c}\delta_{\sigma_b,\sigma_d}\delta[i,j]\delta[j,m], \\ \big[\Lambda_{\sigma_c\sigma_d}^{\sigma_a\sigma_b}[i,m;j,k]\big]_0 &= \delta_{\sigma_a\sigma_c}\delta_{\sigma_b,\sigma_d}\delta[i,j]\delta[m,k]\delta(\tau_j - \tau_k). \end{aligned} \tag{B4}$$

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

$$\begin{aligned} [\Phi[i,m]]_0 &= \delta[i,m]t[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},i] + \tfrac{1}{2}t[i,m]\mathbf{g}^{(k)}[m,m] \\ &\quad + \tfrac{1}{2}\delta[i,m]t[i,\mathbf{j}]\,\mathbf{g}^{(k)}[i,\mathbf{j}] - \tfrac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m]. \end{aligned} \tag{B5}$$

Combining the two we get

$$\begin{aligned} &[Y_1[i,m]]_0 + [\Phi[i,m]]_0 \\ &\quad = \delta[i,m]t[i,\mathbf{j}](\mathbf{g}^{(k)}[i,\mathbf{j}] + \mathbf{g}^{(k)}[\mathbf{j},i]) \\ &\qquad + t[i,m](\mathbf{g}^{(k)}[i,i] + \mathbf{g}^{(k)}[m,m]) - \tfrac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m] \\ &\qquad - \delta[i,m]\tfrac{1}{2}J[i,\mathbf{j}]\mathbf{g}^{(\mathbf{k})}[\mathbf{j},\mathbf{j}]. \end{aligned} \tag{B6}$$

b. Stepping up and final Green's function to $O(\lambda)$

To first order in λ we collect the above results to obtain the Green's function

$$\begin{aligned} [\mathbf{g}^{-1}[i,m]]_1 &= \mathbf{g}^{(k)}[i,i]\cdot\mathcal{V}_{i,m} + \delta[i,m]\mathbf{g}^{(k)}[\mathbf{a},i]\cdot\mathcal{V}_{i,\mathbf{a}}^{(k)} \\ &\quad + \delta[i,m]\tfrac{1}{2}J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] + \tfrac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m] \\ &\quad - t[i,m](\mathbf{g}^{(k)}[i,i] + \mathbf{g}^{(k)}[m,m]) \\ &\quad - \delta[i,m]t[i,\mathbf{j}](\mathbf{g}^{(k)}[i,\mathbf{j}] + \mathbf{g}^{(k)}[\mathbf{j},i]) \end{aligned} \tag{B7}$$

and the caparison factor

$$[\mu[i,m]]_1 = -\mathbf{g}^{(k)}[i^-,i]\delta[i,m]. \tag{B8}$$

The FTs of these on turning off the sources are found using $\mathbf{g}[i^-,i] \to \frac{n}{2}$ as

$$\begin{aligned} [\mathbf{g}^{-1}[k]]_1 &= n\varepsilon_k - \frac{n}{2}u_0 + \frac{1}{2}\sum_q J_{k-q}\mathbf{g}[q] \\ &\quad + \left(\frac{1}{4}J_0 n + 2\sum_q \varepsilon_q \mathbf{g}[q]\right), \end{aligned} \tag{B9}$$

$$[\mu[k]]_1 = -\frac{n}{2}. \tag{B10}$$

The term $-\frac{n}{2}u_0$ in Eq. (B9) arises when we reinstate $J[i,j] \to J[i,j] - u_0\delta[i,j]$ in Eq. (B7). Let us note that under the shift Eq. (13), the first-order correction $[\mathbf{g}^{-1}[k]]_1$ shifts by $2nu_t + \frac{n}{2}u_J$. 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. μ derived objects

We next start with seed terms of $O(\lambda)$ calculated above and end with $[\mu]_2$ and $[\mathbf{g}^{-1}]_2$:

$$[\mu[i,m]]_1 = -\mathbf{g}^{(k)}[i^-,i]\delta[i,m]. \tag{B11}$$

Let us calculate the derived quantities from the above at the same level:

$$\begin{aligned} [\gamma[i]]_1 &= [\mu^{(k)}[\mathbf{a},i]]_1\cdot\mathbf{g}^{(k)}[i,\mathbf{a}] = -\mathbf{g}[i,i]\cdot\mathbf{g}^{(k)}[i,i], \\ [\gamma[i,m]]_1 &= [\mu^{(k)}[\mathbf{a},i]]_1\cdot\mathbf{g}^{(k)}[m,\mathbf{a}] = -\mathbf{g}[i,i]\cdot\mathbf{g}^{(k)}[m,i], \\ [Y[i,m]]_1 &= -t[i,m]\big(\mathbf{g}[i,i]\cdot\mathbf{g}^{(k)}[i,i] + \tfrac{1}{2}\mathbf{g}[m,m]\cdot\mathbf{g}^{(k)}[m,m]\big) \\ &\quad + \tfrac{1}{2}\delta[i,m]\big(J[i,\mathbf{j}]\mathbf{g}[\mathbf{j},\mathbf{j}]\cdot\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}] \\ &\quad - t[i,\mathbf{j}]\mathbf{g}[\mathbf{j},\mathbf{j}]\cdot\mathbf{g}^{(k)}[i,\mathbf{j}]\big). \end{aligned} \tag{B12}$$

Zero-source Fourier transforms:

$$\begin{aligned} ([\gamma[0]]_1)_{\mathcal{V}\to 0} &= -\frac{n^2}{4}, \\ ([\gamma[k]]_1)_{\mathcal{V}\to 0} &= -\frac{n}{2}\mathbf{g}[-k], \\ ([Y_1[k]]_1)_{\mathcal{V}\to 0} &= \frac{3n^2}{8}\varepsilon_k - \frac{n^2}{8}u_0 + \left(\frac{n^2}{8}J_0 + \frac{n}{4}\sum_q \varepsilon_q \mathbf{g}[q]\right). \end{aligned} \tag{B13}$$

Here we reinstated $J[i,j] \to J[i,j] - u_0\delta[i,j]$ in Eq. (B12) to obtain the $-\frac{n^2}{8}u_0$ term in Eq. (B13).

b. μ derived vertices

Next we calculate (using lowest-order functional derivatives)

$$\begin{aligned}\left[\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[i,m;a]\right]_1 &= -\delta[i,m]\sigma_1\sigma_2\mathbf{g}_{\bar{\sigma}_2\sigma_3}[i,a]\mathbf{g}_{\sigma_4\bar{\sigma}_1}[a,i],\\ \left[\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[i,m;a]\right]_{1\ \mathcal{V}\to 0} &= -\delta[i,m]\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\mathbf{g}[i,a]\mathbf{g}[a,i].\end{aligned} \tag{B14}$$

At zero sources so with $\mathcal{V}\to 0$

$$\begin{aligned}[\mathcal{U}(a)[i,m;a]]_1 &= -2\delta[i,m]\ \mathbf{g}[i,a]\mathbf{g}[a,i],\\ [\mathcal{U}^{(s)}[i,m;a]]_1 &= -\delta[i,m]\ \mathbf{g}[i,a]\mathbf{g}[a,i].\end{aligned} \tag{B15}$$

The zero-source Fourier transforms are as follows:

$$\left[\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p_1,p_2]\right]_1 = -\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\sum_q \mathbf{g}[q]\mathbf{g}[q+p_2-p_1],$$

$$\left[\mathcal{U}^{(a)}[p_1,p_2]\right]_1 = -2\sum_q \mathbf{g}[q]\mathbf{g}[q+p_2-p_1]. \tag{B16}$$

Similarly we find for the four-index vertices

$$\begin{aligned}&\left[\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[i,m;a,b]\right]_1\\ &\quad= -\delta[i,m]\ \sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\ \mathbf{g}[i,a]\mathbf{g}[b,i]\ \delta(\tau_a-\tau_b),\\ &[\mathcal{U}^{(a)}[i,m;a,b]]_{1\ \mathcal{V}\to 0}\\ &\quad= -2\delta[i,m]\mathbf{g}[i,a]\mathbf{g}[b,i]\delta(\tau_a-\tau_b).\end{aligned} \tag{B17}$$

The zero-source Fourier transforms are as follows:

$$\begin{aligned}\left[\mathcal{U}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p_1,p_2;p_3,p_4]\right]_1 &= -\sigma_1\sigma_2\delta_{\bar{\sigma}_2,\sigma_3}\delta_{\bar{\sigma}_4,\sigma_1}\delta_{p_1+p_4,p_2+p_3}\\ &\quad\times\mathbf{g}[p_3]\mathbf{g}[p_4],\\ [\mathcal{U}^{(a)}[p_1,p_2;p_3,p_4]]_1 &= -2\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4].\end{aligned} \tag{B18}$$

c. Ψ to $O(\lambda)$

We compute $[\Psi]_1$ from these:

$$\begin{aligned}[\Psi(k)]_1 &= \sum_p\left(\varepsilon_p+\frac{1}{2}\varepsilon_k+\frac{1}{2}J_{k-p}\right)\mathbf{g}[p][\mathcal{U}^{(a)}(p,k)]_1\\ &\quad+\sum_{pq}\frac{1}{2}\varepsilon_{q+p-k}\ \mathbf{g}[p][\mathcal{U}^{(a)}(p,k;q+p-k,q)]_1\\ &= -\sum_{p,q}(\varepsilon_p+\varepsilon_{k+q-p}+\varepsilon_k+\varepsilon_q+J_{k-p})\\ &\quad\times\mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[q+k-p]\\ &= \sum_{p,q}W(k,q;q+k-p,p)\mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[q+k-p].\end{aligned} \tag{B19}$$

d. Stepping up: μ to $O(\lambda^2)$

Stepping up, we calculate

$$[\mu[i,m]]_2 = -\delta[i,m]\mathbf{g}[i^-,i]\mathbf{g}^{(k)}[i^-,i]+[\Psi[i,m]]_1. \tag{B20}$$

Hence at zero sources, the Fourier transform reads

$$\begin{aligned}[\mu[k]]_2 &= \frac{n^2}{4}-\sum_{p,q}(\varepsilon_p+\varepsilon_{k+q-p}+\varepsilon_k+\varepsilon_q+J_{k-p})\\ &\quad\times\mathbf{g}[p]\mathbf{g}[q]\mathbf{g}[q+k-p].\end{aligned} \tag{B21}$$

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):

$$\begin{aligned}[\mathbf{g}^{-1}[i,m]]_1 &= \left[\mathbf{g}^{(k)}[i,i]\cdot\mathcal{V}_{i,m}+\delta[i,m]\mathbf{g}^{(k)}[\mathbf{a},i]\cdot\mathcal{V}^{(k)}_{i,\mathbf{a}}\right]\\ &\quad+\left[\delta[i,m]\tfrac{1}{2}J[i,\mathbf{j}]\mathbf{g}^{(k)}[\mathbf{j},\mathbf{j}]+\tfrac{1}{2}J[i,m]\mathbf{g}^{(k)}[i,m]\right]\\ &\quad\times[-t[i,m](\mathbf{g}^{(k)}[i,i]+\mathbf{g}^{(k)}[m,m])\\ &\quad-\delta[i,m]t[i,\mathbf{j}](\mathbf{g}^{(k)}[i,\mathbf{j}]+\mathbf{g}^{(k)}[\mathbf{j},i])].\end{aligned} \tag{B22}$$

f. Vertex functions to $O(\lambda)$

Three-point vertex:

$$\begin{aligned}\left[\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[i,m;a]\right]_1 &= -\left(\frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_a}\right)\left[\mathbf{g}^{-1}_{\sigma_1\sigma_2}[i,m]\right]_1\\ &= [(I)+(II)+(III)]^{\sigma_1\sigma_2}_{\sigma_3\sigma_4},\end{aligned} \tag{B23}$$

where the terms $(I),(II),(III)$ refer to the three square bracketed terms in Eq. (B22). For the first term we calculate

$$\begin{aligned}(I)^{\sigma_1\sigma_2}_{\sigma_3\sigma_4} &= -\left(\frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_a}\right)\left[(\sigma_1\sigma_a)\mathbf{g}_{\bar{\sigma}_a\bar{\sigma}_1}[i,i]\mathcal{V}^{\sigma_a\sigma_2}_{i,m}\right.\\ &\quad\left.+\delta[i,m]\ (\sigma_1\sigma_2)\mathbf{g}_{\bar{\sigma}_a\bar{\sigma}_1}[\mathbf{a},i]\mathcal{V}^{\bar{\sigma}_2\bar{\sigma}_a}_{i,\mathbf{a}}\right],\end{aligned} \tag{B24}$$

$$\mathcal{V}\to 0 = 0. \tag{B25}$$

For all other terms we can use a simple calculation:

$$\begin{aligned}\mathcal{X}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p,q;r] &= \left(\frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_r}\right)\mathbf{g}^{(k)}_{\sigma_1\sigma_2}[p,q]\\ &= (\sigma_1\sigma_2)\mathbf{g}_{\bar{\sigma}_2\sigma_3}[p,r]\mathbf{g}_{\sigma_4\bar{\sigma}_1}[r,q],\\ \mathcal{X}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p,q;r]_{\mathcal{V}\to 0} &= (\sigma_1\sigma_2)\delta_{\bar{\sigma}_2\sigma_3}\delta_{\bar{\sigma}_1\sigma_4}\mathbf{g}[p,r]\mathbf{g}[r,q],\\ \mathcal{X}^{(a)}[p,q;r]_{\mathcal{V}\to 0} &= 2\mathbf{g}[p,r]\mathbf{g}[r,q].\end{aligned} \tag{B26}$$

Therefore

$$\begin{aligned}(II)^{(a)}[i,m;a] &= -\delta[i,m]J[i,\mathbf{j}]\mathbf{g}[\mathbf{j},a]\mathbf{g}[a,\mathbf{j}]\\ &\quad-J[i,m]\mathbf{g}[i,a]\mathbf{g}[a,m]\end{aligned} \tag{B27}$$

and

$$\begin{aligned}&(III)^{(a)}[i,m;a]\\ &\quad= 2[t[i,m](\mathbf{g}[i,a]\mathbf{g}[a,i]+\mathbf{g}[m,a]\mathbf{g}[a,m])\\ &\qquad+\delta[i,m]t[i,\mathbf{j}](\mathbf{g}[i,a]\mathbf{g}[a,\mathbf{j}]+\mathbf{g}[\mathbf{j},a]\mathbf{g}[a,i])].\end{aligned} \tag{B28}$$

Zero-source Fourier transforms read as

$$\begin{aligned}(II)^{(a)}[p_1,p_2] &= -J_{p_2-p_1}\sum_q \mathbf{g}[q]\mathbf{g}[q+p_2-p_1] - \sum_q J_{p_1-q}\mathbf{g}[q]\mathbf{g}[q+p_2-p_1],\\ (III)^{(a)}[p_1,p_2] &= -2\sum_q \mathbf{g}[q]\mathbf{g}[q+p_2-p_1]\{\varepsilon_{p_2}+\varepsilon_{p_1}+\varepsilon_{q+p_2-p_1}+\varepsilon_q\}.\end{aligned} \tag{B29}$$

Hence adding up we obtain

$$\begin{aligned}\left[\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p_1,p_2]\right]_1 &= -(\sigma_1\sigma_2)\delta_{\bar\sigma_2\sigma_3}\delta_{\bar\sigma_1\sigma_4}\sum_{p_3,p_4}\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4]\left\{\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+\frac{1}{2}\left(J_{p_1-p_2}+J_{p_1-p_3}\right)\right\}\\ &= \frac{1}{2}(\sigma_1\sigma_2)\delta_{\bar\sigma_2\sigma_3}\delta_{\bar\sigma_1\sigma_4}\sum_{p_3,p_4}\mathbf{g}[p_3]\mathbf{g}[p_4]\left[W(p_2,p_3,p_4,p_1)+W(p_2,p_3,p_1,p_4)\right],\end{aligned} \tag{B30}$$

$$[\Lambda^{(a)}[p_1,p_2]]_1 = -\sum_{p_3,p_4}\mathbf{g}[p_3]\mathbf{g}[p_4][W(p_2,p_3,p_4,p_1)+W(p_2,p_3,p_1,p_4)]. \tag{B31}$$

Note that rotation invariance relations imply that since $[\Lambda^{(1)}[p_1,p_2]]_1 = 0$, we must have

$$[\Lambda^{(s)}[p_1,p_2]]_1 = -[\Lambda^{(t)}[p_1,p_2]]_1 = \frac{1}{2}[\Lambda^{(a)}[p_1,p_2]]_1. \tag{B32}$$

Four point vertex. The calculation proceeds similarly:

$$\left[\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[i,m;a,b]\right]_1 = -\left(\frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_{a,b}}\right)\left[\mathbf{g}^{-1}_{\sigma_1\sigma_2}[i,m]\right]_1 = [(IV)+(V)+(VI)]^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}. \tag{B33}$$

Here the terms (IV)–(VI) refer to the three square bracketed terms in Eq. (B22). For the first term we calculate with implicit $\tau_a = \tau_b$:

$$\begin{aligned}(IV)^{\sigma_1\sigma_2}_{\sigma_3\sigma_4} &= -\left(\frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_{a,b}}\right)\left[(\sigma_1\sigma_a)\mathbf{g}_{\bar\sigma_a\bar\sigma_1}[i,i]\mathcal{V}^{\sigma_a\sigma_2}_{i,m} + \delta[i,m]\,(\sigma_1\sigma_2)\mathbf{g}_{\bar\sigma_a\bar\sigma_1}[\mathbf{a},i]\mathcal{V}^{\bar\sigma_2\bar\sigma_a}_{i,\mathbf{a}}\right]\\ &= -\delta[i,a]\delta[m,b]\left[(\sigma_1\sigma_3)\delta_{\sigma_2\sigma_4}\mathbf{g}_{\bar\sigma_3\bar\sigma_1}[i,i]\right] - \delta[i,a]\delta[i,m]\left[(\sigma_1\sigma_2)\delta_{\bar\sigma_2,\sigma_3}\mathbf{g}_{\sigma_4\bar\sigma_1}[b,i]\right],\end{aligned} \tag{B34}$$

$$\begin{aligned}(IV)^{\sigma_1\sigma_2}_{\sigma_3\sigma_4} &\overset{\mathcal{V}\to 0}{=} -\delta[i,a]\delta[m,b]\delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4}\mathbf{g}[i,i] - \delta[i,a]\delta[i,m](\sigma_1\sigma_2)\delta_{\sigma_1\bar\sigma_4}\delta_{\sigma_2\bar\sigma_3}\mathbf{g}[b,i],\\ (IV)^{(a)} &= \underbrace{\delta[i,a]\delta[m,b]\mathbf{g}[i,i] - 2\delta[i,a]\delta[i,m]\mathbf{g}[b,i]}.\end{aligned} \tag{B35}$$

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:

$$\begin{aligned}\mathcal{Y}^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p,q;r,s] &= \left(\frac{\delta}{\delta\mathcal{V}^{\sigma_3\sigma_4}_{r,s}}\right)\mathbf{g}^{(k)}_{\sigma_1\sigma_2}[p,q],\\ \mathcal{Y}^{(a)}[p,q;r,s]_{\mathcal{V}\to 0} &= 2\mathbf{g}[p,r]\mathbf{g}[s,q]\delta(\tau_r-\tau_s).\end{aligned} \tag{B36}$$

Therefore with implicit $\tau_a = \tau_b$:

$$(V)^{(a)}[i,m;a,b] = -\delta[i,m]J[i,\mathbf{j}]\mathbf{g}[\mathbf{j},a]\mathbf{g}[b,\mathbf{j}] - J[i,m]\mathbf{g}[i,a]\mathbf{g}[b,m], \tag{B37}$$

$$\begin{aligned}(VI)^{(a)}[i,m;a,b] &= 2\left[t[i,m](\mathbf{g}[i,a]\mathbf{g}[b,i]+\mathbf{g}[m,a]\mathbf{g}[b,m]) + \delta[i,m]t[i,\mathbf{j}]\,(\mathbf{g}[i,a]\mathbf{g}[b,\mathbf{j}]+\mathbf{g}[\mathbf{j},a]\mathbf{g}[b,i])\right],\\ (IV)^{(a)}[i,m;a,b] &= \delta[i,a]\delta[m,b]\mathbf{g}[i,i] - 2\delta[i,a]\delta[i,m]\mathbf{g}[b,i],\\ (IV)^{(a)}[p_1,p_2,p_3,p_4] &= \underbrace{\delta_{p_1,p_3}\delta_{p_2,p_4}\mathbf{g}[0^-] - 2\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_4]},\end{aligned} \tag{B38}$$

$$\begin{aligned}(V)^{(a)}[i,m;a,b] &= -\delta[i,m]J[i,\mathbf{j}]\mathbf{g}[\mathbf{j},a]\mathbf{g}[b,\mathbf{j}] - J[i,m]\mathbf{g}[i,a]\mathbf{g}[b,m],\\ (V)^{(a)}[p_1,p_2,p_3,p_4] &= -\left(J_{p_2-p_1}+J_{p_1-p_3}\right)\mathbf{g}[p_3]\mathbf{g}[p_4]\delta_{p_1+p_4,p_2+p_3},\end{aligned} \tag{B39}$$

$$\begin{aligned}(VI)^{(a)}[i,m;a,b] &= 2\left[t[i,m](\mathbf{g}[i,a]\mathbf{g}[b,i]+\mathbf{g}[m,a]\mathbf{g}[b,m]) + \delta[i,m]t[i,\mathbf{j}]\,(\mathbf{g}[i,a]\mathbf{g}[b,\mathbf{j}]+\mathbf{g}[\mathbf{j},a]\mathbf{g}[b,i])\right],\\ (VI)^{(a)}[p_1,p_2,p_3,p_4] &= -2\{\varepsilon_{p_2}+\varepsilon_{p_1}+\varepsilon_{p_3}+\varepsilon_{p_4}\}\mathbf{g}[p_3]\mathbf{g}[p_4]\delta_{p_1+p_4,p_2+p_3}.\end{aligned} \tag{B40}$$

Hence

$$\left[\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}[p_1,p_2,p_3,p_4]\right]_1 = -(\sigma_1\sigma_2)\delta_{\bar{\sigma}_2\sigma_3}\delta_{\bar{\sigma}_1\sigma_4}\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4]\left\{\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+\tfrac{1}{2}\left(J_{p_1-p_2}+J_{p_1-p_3}\right)\right\}$$
$$\underbrace{-\delta_{p_1,p_3}\delta_{p_2,p_4}\delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4}\mathbf{g}[0^-]-2(\sigma_1\sigma_2)\delta_{\sigma_1\bar{\sigma}_4}\delta_{\sigma_2\bar{\sigma}_3}\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_4]}. \tag{B41}$$

Thus

$$[\Lambda^{(a)}[p_1,p_2,p_3,p_4]]_1 = -2\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_3]\mathbf{g}[p_4]\left\{\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+\tfrac{1}{2}\left(J_{p_1-p_2}+J_{p_1-p_3}\right)\right\}$$
$$+\underbrace{\delta_{p_1,p_3}\delta_{p_2,p_4}\mathbf{g}[0^-]-2\delta_{p_1+p_4,p_2+p_3}\mathbf{g}[p_4]}. \tag{B42}$$

Comparing Eq. (B30) and Eq. (B42), we see that other than the term with underbraces, these vertices satisfy Eq. (49) or Eq. (72).

g. Φ to $O(\lambda)$

We now assemble terms:

$$[\Phi(k)]_1 = \sum_p\left(\varepsilon_p+\frac{1}{2}\varepsilon_k+\frac{1}{2}J_{k-p}\right)\mathbf{g}[p][\Lambda^{(a)}(p,k)]_1+\sum_{pq}\frac{1}{2}\varepsilon_{q+p-k}\mathbf{g}[p][\Lambda^{(a)}(p,k;q+p-k,q)]_1. \tag{B43}$$

Let us rewrite this as ($k\to p_2, p\to p_1, q\to p_4$)

$$[\Phi(p_2)]_1 = \sum_{p_1}\left(\varepsilon_{p_1}+\frac{1}{2}\varepsilon_{p_2}+\frac{1}{2}J_{p_1-p_2}\right)\mathbf{g}[p_1][\Lambda^{(a)}(p_1,p_2)]_1+\sum_{p_1+p_4=p_2+p_3}\frac{1}{2}\varepsilon_{p_3}\mathbf{g}[p_1][\Lambda^{(a)}(p_1,p_2;p_3,p_4)]_1$$
$$=\underbrace{\frac{n}{4}\sum_{p_3}\varepsilon_{p_3}\mathbf{g}[p_3]-\sum_{p_1,p_4}\varepsilon_{p_1+p_4-p_2}\mathbf{g}[p_1]\mathbf{g}[p_4]}$$
$$-2\sum_{p_1+p_4=p_2+p_3}\mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4]\left(\varepsilon_{p_1}+\frac{1}{2}\varepsilon_{p_2}+\frac{1}{2}J_{p_1-p_2}\right)\left\{\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+\frac{1}{2}\left(J_{p_1-p_2}+J_{p_1-p_3}\right)\right\}$$
$$-\sum_{p_1+p_4=p_2+p_3}\mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4]\varepsilon_{p_3}\left\{\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+\frac{1}{2}\left(J_{p_1-p_2}+J_{p_1-p_3}\right)\right\}. \tag{B44}$$

The first line with underbraces arises from the term Λ 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

$$[\Phi(p_2)]_1 = -\sum_{p_1+p_4=p_2+p_3}\mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4]\left(\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+J_{p_1-p_2}\right)\left\{\varepsilon_{p_1}+\varepsilon_{p_2}+\varepsilon_{p_3}+\varepsilon_{p_4}+\frac{1}{2}\left(J_{p_1-p_2}+J_{p_1-p_3}\right)\right\}$$
$$=\frac{1}{2}\sum_{p_1,p_3,p_4}\mathbf{g}[p_1]\mathbf{g}[p_3]\mathbf{g}[p_4]W(p_2,p_3;p_4,p_1)\left[W(p_2,p_3;p_4,p_1)+W(p_2,p_3;p_1,p_4)\right], \tag{B45}$$

where in the first line we symmetrized further in $p_1\leftrightarrow p_4$.

We can bring this into standard notation by sending $p_2\to k, p_1\to p, p_3\to q, p_4\to k+q-p$:

$$[\Phi(k)]_1 = -\sum_{q,p}\mathbf{g}[q]\mathbf{g}[p]\mathbf{g}[k+q-p](\varepsilon_k+\varepsilon_p+\varepsilon_q+\varepsilon_{k+q-p}+J_{k-p})\left\{\varepsilon_k+\varepsilon_p+\varepsilon_q+\varepsilon_{k+q-p}+\frac{1}{2}(J_{k-p}+J_{p-q})\right\},$$
$$[\Phi(k)]_1 = \frac{1}{2}\sum_{q,p}\mathbf{g}[q]\mathbf{g}[p]\mathbf{g}[k+q-p]W(k,q;q+k-p,p)\left[W(k,q;q+k-p,p)+W(k,q;p,q+k-p)\right]. \tag{B46}$$

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 μ. 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:

$$\mathbf{g}^{-1}(k) = \mathbf{g}_0^{-1}(k)+\lambda[\mathbf{g}^{-1}(k)]_1+\lambda^2[\mathbf{g}^{-1}(k)]_2+O(\lambda^3),\quad [\mathbf{g}^{-1}[k]]_0 = i\omega_n+\boldsymbol{\mu}-\varepsilon_k-\frac{1}{4}J_0,$$
$$[\mathbf{g}^{-1}[k]]_1 = n\varepsilon_k-\frac{n}{2}u_0+\frac{1}{2}\sum_q J_{k-q}\mathbf{g}[q]+\left(\frac{1}{4}J_0n+2\sum_q\varepsilon_q\mathbf{g}[q]\right), \tag{B47}$$
$$[\mathbf{g}^{-1}[k]]_2 = -\frac{3n^2}{8}\varepsilon_k+\frac{n^2}{8}u_0-[\Phi(k)]_1-\left(\frac{n^2}{8}J_0+\frac{n}{4}\sum_q\varepsilon_q\mathbf{g}[q]\right).$$

We thus see that all the computed $[\mathbf{g}^{-1}]_j$ are invariant under the two shift theorems. Adding up terms to $O(\lambda^2)$,

$$\mathbf{g}^{-1}[k] = i\omega_n + \boldsymbol{\mu}' - \left(1 - \lambda n + \lambda^2 \frac{3n^2}{8}\right)\varepsilon_k + \lambda \sum_q \frac{1}{2} J_{k-q}\mathbf{g}[q] - \lambda^2 [\Phi(k)]_1 + O(\lambda^3), \tag{B48}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} - u_0 \frac{\lambda n}{2}\left(1 - \frac{\lambda n}{4}\right) + \left[J_0 \frac{\lambda n}{4}\left(1 - \frac{\lambda n}{2}\right) + 2\lambda\left(1 - \frac{\lambda n}{8}\right)\sum_q \varepsilon_q \mathbf{g}[q]\right], \tag{B49}$$

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 $J_0 \to 0$, while the second term involving $\sum_q \varepsilon_q \mathbf{g}[q]$ is expected to be weakly T dependent.

Similarly the caparison factor μ is found to $O(\lambda^2)$ as

$$\mu[k] = 1 + \lambda [\mu[k]]_1 + \lambda^2 [\mu[k]]_2 + O(\lambda^3),$$
$$[\mu[k]]_1 = -\frac{n}{2}, \quad [\mu[k]]_2 = \frac{n^2}{4} + [\Psi[k]]_1 . \tag{B50}$$

Adding up terms to $O(\lambda^2)$ we obtain

$$\mu[k] = 1 - \lambda \frac{n}{2} + \lambda^2 \frac{n^2}{4} + \lambda^2 [\Psi(k)]_1 + O(\lambda^3), \tag{B51}$$

along with the definition in Eq. (B19).

[1] D. Hansen and B. S. Shastry, arXiv:1211.0594 [cond-mat.str-el].

[2] B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011); **108**, 029702 (2012).

[3] B. Edegger, V. N. Muthukumar, and C. Gros, Adv. Phys. **56**, 927 (2007).

[4] P. W. Anderson, Science **235**, 1196 (1987).

[5] A. B. Harris and R. V. Lange, Phys. Rev. **157**, 295 (1967).

[6] F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).

[7] M. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).

[8] B. S. Shastry, Phys. Rev. B **84**, 165112 (2011).

[9] G.-H. Gweon, B. S. Shastry, and G. D. Gu, Phys. Rev. Lett. **107**, 056404 (2011).

[10] B. S. Shastry, Phys. Rev. Lett. **109**, 067004 (2012).

[11] B. S. Shastry, Phys. Rev. B **81**, 045121 (2010).

[12] A simple example $\mathcal{G} \sim c_0/(i\omega_n + \mu - \varepsilon_k)$ can be worked out in detail and illustrates this remark.

[13] A. A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

[14] G. D. Mahan, *Many-Particle Physics*, 3rd ed. (Kluwer-Plenum, 2000).

[15] 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* $X_i^{0\sigma}$ *with* H, e.g., in the expression Eq. (34), and the extra term in A_{i,σ_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.

[16] 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."

[17] J. M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960).

[18] W. Brinkman and T. M. Rice, Phys. Rev. B **2**, 4302 (1970).

[19] G. Baskaran, Z. Zhou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).

[20] G. Kotliar, Phys. Rev. B **37**, 3664 (1988).

Shastry Replies: The Letter [1] (L) presents a framework for calculating the properties of extremely correlated Fermi liquids (ECFL) with eliminated double occupancy, where the standard Feynman diagram formalism fails. By adapting Schwinger's method of source fields to generate exact functional differential equations, it overcomes the vexing absence of Wick's theorem for noncanonical electrons. Here the electronic Greens function $\mathcal{G}$ is found as the convolution of a canonical Greens function g involving a self energy Φ, and a dynamical spectral weight μ, given in terms of a second ''self energy'' Ψ. An expansion of the exact equations is given in a parameter λ argued to be equivalent to the particle density n. Equations for Ψ and Φ to first order in λ are written down [Eqs. L(18)-(20)], giving the first ω dependent corrections to the Fermi gas. These require detailed computations that are underway. Meanwhile for illustrating the formalism, a further analytical simplification to the above first order equations is made. This ''simplified ECFL theory'' thereby yields the explicit analytical approximation Eq. L(26) for the one electron spectral function $\rho_{\mathcal{G}}$ of the t-J model. In [2], this expression is successfully confronted with data on cuprate superconductors.

The preceding Comment (C) [3] questions the positivity of the spectral function of the simplified ECFL theory. As already recognized, and explicitly addressed in the Letter, maintaining positivity of the spectral function within various approximations of this framework can be challenging in general. This aspect of the formalism differs from the standard theory, and one must be prepared to make suitable physical approximations, e.g., by invoking hard or soft cutoffs. As written Eq. (26) is non-negative, thanks to the explicit cutoff in the numerator. On the other hand, the cutoff plays almost no role in the most interesting regime of this approximation relevant to angle-resolved photoemission, since for most occupied states the spectral function is already positive (see also L. Ref. [13]).

The Comment addresses the $\vec{k}$ dependence of the simplified ECFL theory [4], and compares it with a dynamical mean field theory calculation (C. Ref [3]). This exercise mistakenly supposes that the simplified ECFL theory is a proposed solution of the t-J model, in the limit of high dimensions. In clarification: the Letter does not claim to solve the t-J model in high dimensions by this approximation. The context of the simplified ECFL theory is discussed near Eq. L(24). In brief, this approximation is ''high dimensional'' only to the extent that the two self energies within the first order in λ equations, Eq. L(18)–(20), are proportional to each other, if their momentum dependence is neglected. Concerning higher order terms in λ, which must be addressed for a complete solution, we make no such claim. The term ''high dimensional approximation'' thus refers only to the relationship between the two self energies in the first order in λ theory Eq. L(18)–(20). Renaming the approximation as the ''simplified ECFL theory'' from our initial nomenclature makes its nature more clear. The ECFL solution of the t-J model for $d \to \infty$ is undoubtedly of interest. If found, it would clarify the role of the crucial variable $\frac{U}{d}$. In ECFL one assumes $\frac{U}{d} \to \infty$, while apparently $\frac{U}{d} \to 0$ in (C. Ref [3]).

The Comment remarks on the nature of the ECFL state and its purported inaccessibility from the Hubbard model. A physically meaningful and rigorous characterization of a quantum fluid is in terms of the symmetry of the ground state, and in the nature of its lowest excitations for small $(\vec{k} - \vec{k}_F, \omega)$ as reflected in $\mathcal{G}$, and as addressed by scaling theories [5]. The technique by which $\mathcal{G}$ is calculated seems peripheral to this issue. The Hubbard model has a Fermi liquid phase at suitable parameters with standard signatures [5]. The ECFL state also represents a Fermi liquid (FL) in the above sense, provided the energies $|\xi_k|, |\omega|$ are smaller than the energy scale $\Delta(\vec{k}_F, 0)$ [Eq. L(22)]. Generically, $\Delta(\vec{k}_F, 0)$ is nonvanishing in ECFL, as in the simplified ECFL theory. Thus, we asymptotically again arrive at a FL state, without a fundamental obstruction to adiabatic continuation in U between the two FL states. The theory interestingly also allows for a vanishing $\Delta(\vec{k}_F, 0)$ at special values of parameters, thereby locating possible quantum critical points. It also yields several useful results away from the asymptotic region [2].

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

B. Sriram Shastry
Physics Department, University of California, Santa Cruz, California 95064, USA

Received 14 November 2011; published 10 January 2012
DOI: 10.1103/PhysRevLett.108.029702
PACS numbers: 71.10.Fd

[1] B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011).
[2] G. H. Gweon, B. S. Shastry, and G. D. Gu, Phys. Rev. Lett. **107**, 056404 (2011).
[3] K. Matho, preceding Comment Phys. Rev. Lett. **108**, 029701 (2012).
[4] The ECFL theory may equally well be represented in terms of a Dyson self energy Σ, as shown in B. S. Shastry, Phys. Rev. B **84**, 165112 (2011). For the simplified ECFL theory, $\Im m\Sigma(\vec{k}, \omega)$ is displayed in Figs. (8,9) of this work. It is a function of $\vec{k}$ only through the single particle energy $\varepsilon_{\vec{k}}$. This seems to be the minimal $\vec{k}$ dependence required to reconcile data with theory [2].
[5] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994); D. Zanchi and H. J. Schulz, Phys. Rev. B **54**, 9509 (1996).

 © 2012 American Physical Society

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 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 $\boldsymbol{\mu}$ as

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

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 $O(\omega/\varepsilon_{\mathrm{F}})^3$ 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 ± 25 meV around $\boldsymbol{\mu}$, the theories and the fits of Ref. [9] (extrapolated to lower ω) 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 $\xi_k = \vec{\hat{k}} \cdot \vec{v}_{\vec{k}_{\mathrm{F}}}$) 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 ω. It also contains the Fermi function for occupied states $f_\omega = \{1 + \exp(\beta\omega)\}^{-1}$, a nonsymmetric function of ω. Therefore, we first formulate a Fermi symmetrized object:

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

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 $\mathcal{S}^{\mathrm{a-s}}_{\mathcal{G}}(\vec{k}_{\mathrm{F}}|\vec{\hat{k}}, \omega)$ and symmetric $\mathcal{S}^{\mathrm{s}}_{\mathcal{G}}(\vec{k}_{\mathrm{F}}|\vec{\hat{k}}, \omega)$ combinations

 © 2012 American Physical Society

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

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

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 $\mathcal{S}^{\rm s}_{\mathcal{G}}$ below for various theoretical models; $\mathcal{S}^{\rm a\text{-}s}_{\mathcal{G}}$ 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 $\rho^{\rm FL}_G(\vec{k},\omega)$ 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 $\hat{k}$ and ω. With $\vec{v}_{\vec{k}_{\rm F}}$, the Fermi velocity vector at $\vec{k}_{\rm F}$, the quasiparticle piece is specified by three parameters (i) renormalization factor $Z_{\vec{k}}$, with a linear dependence $Z_{\vec{k}} = Z_{\vec{k}_{\rm F}}[1 + c_1(\vec{\hat{k}}\cdot\vec{v}_{\vec{k}_{\rm F}})]$, (ii) the quasiparticle energy $E_{\vec{k}}$ vanishing linearly at the Fermi surface $E_{\vec{k}} = \frac{m}{m^*}(\vec{\hat{k}}\cdot\vec{v}_{\vec{k}_{\rm F}})$ with an effective mass renormalization $\frac{m}{m^*}$, and (iii) the line width $\gamma_{\vec{k}} \propto [E^2_{\vec{k}} + (\pi k_B T)^2]$ vanishes symmetrically at the Fermi surface. Thus near the Fermi surface:

$$\rho^{\rm FL}_G(\vec{k},\omega) \sim \rho^{(bg)}_G(\vec{k},\omega) + \frac{Z_{\vec{k}}}{\pi}\frac{\gamma_{\vec{k}}}{\gamma^2_{\vec{k}} + (\omega - E_{\vec{k}})^2}. \qquad (4)$$

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

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

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

$$\{\mathcal{S}^{\rm s}_{G_{\rm FL}}, \mathcal{R}_{G_{\rm FL}}\} = \{\mathcal{Q}(\vec{\hat{k}},\omega), c_1(\vec{\hat{k}}\cdot\vec{v}_{\vec{k}_{\rm F}})\}, \qquad (6)$$

where we retained only terms linear in $\hat{k}$, ω beyond the quasiparticle peak term $\mathcal{Q}(\vec{\hat{k}},\omega)$. Observe that to $O(\omega^2)$ the asymmetry ratio $\mathcal{R}$ is independent of ω. The requirement of neglecting the background is necessary, because it is hard to make a general statement about the (k,ω) 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 ω dependent but $\vec{k}$ independent. To the extent that we can experimentally identify a ω 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 ω, such as the marginal Fermi-liquid [11] and also various refinements of the RPA. Subleading corrections of the type $\omega \times T^2$ or ω^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 $\rho_{\mathbf{g}}(\vec{k},\omega)$ and a caparison factor $[\{1-\frac{n}{2}\} + \frac{\xi_k - \omega}{\Delta(\vec{k},\omega)} + \eta(\vec{k},\omega)]$, explicitly containing a linear dependence on the energy ω. 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 Δ is a constant determined by the number sum rule. In Refs. [9,12], the SECFL was tested against data on the high T_c cuprate $Bi_2Sr_2CaCu_2O_{8+\delta}$. 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 $\sim 2k_BT$, 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

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

Here the term d_0 arises from Taylor expanding $\eta(\vec{k}_{\rm F},0)$ 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 $\{\mathcal{S}^{s}_{\mathcal{G}_{\rm ECFL}}, \mathcal{R}_{\mathcal{G}_{\rm ECFL}}\}$ as

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

The asymmetry ratio $\mathcal{R}$ therefore has a linear ω and $\hat{k}$ 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 $\rho^{\rm Peak}_{\mathcal{G}_{\rm SECFL}}(\vec{k}_{\rm F} + \vec{k}, \omega)$ as

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

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

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

where $\mathcal{Q}(\hat{\vec{k}}, \omega)$ is obtained from Eq. (5) by replacing $m/m^* \to Z_{\rm F}$ and $\gamma_k \to \Gamma_k Z_k$. Note that, e.g., at $\hat{\vec{k}} = \vec{0}$ and any convenient ω_0, $|\mathcal{R}(0, \omega_0)| = \omega_0/\varepsilon_0$, and thus its magnitude yields the important energy scale Δ_0. We emphasize that Eq. (7) is more generally true within the ECFL approach. We display $\mathcal{S}^{\rm a\text{-}s}$ 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, $C_\Phi = 1\ (\text{eV})^{-1}$. Notice the distinctive increasing linear behavior with $\hat{\vec{k}}$ and a decreasing linear one with ω, 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 $\rho^{\rm (local)}_{\mathcal{G}}(\omega) = \sum_{\vec{k}}\rho_{\mathcal{G}}(\vec{k}; \omega)$. Its convolution with f_ω 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

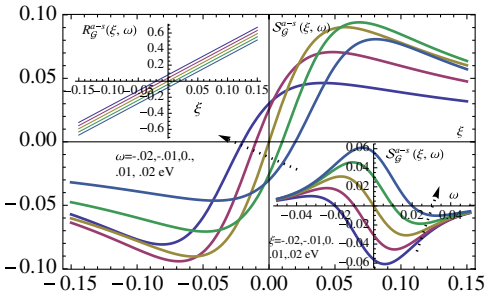

FIG. 1 (color online). Top inset shows the large predicted asymmetry $\mathcal{R}^{\rm a\text{-}s}_{\mathcal{G}_{\rm SECFL}}$ versus ξ in the small energy range of 150 meV. Similar magnitudes are found as functions of ω at various ξ. The figure shows $\mathcal{S}^{\rm a\text{-}s}_{\mathcal{G}_{\rm SECFL}}$ from Eq. (9) versus ξ (main), ω (inset) in electron volt at various ω (main), ξ (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 η. 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 ξ_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

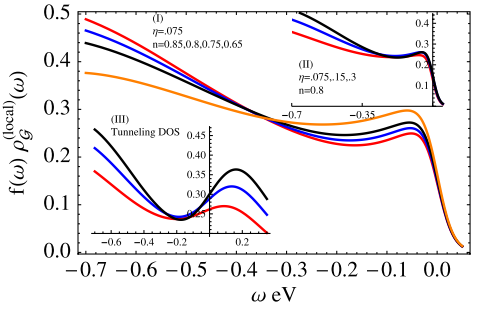

FIG. 2 (color online). (I) The predicted AIP spectrum showing a shallow minimum at $\omega \sim -0.2$ eV, and a rise as the binding energy $|\omega|$ increases. The rise is greater as the particle density n increases (bottom to top). Inset (II) reveals the role of elastic scattering width η (top to bottom). 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 $\delta(\vec{\hat{k}}\cdot\vec{v}_{\vec{k}_F}-\frac{m^*}{m}\omega)$. This yields the quasiparticle peak contribution:

$$\lim_{\omega\le\varepsilon_0}\rho^{(\mathrm{local})}_{\mathcal{G},\mathrm{Peak}}(\omega)\sim \mathrm{const}\times\left\{\varepsilon_0+\left(\frac{m^*}{m}-1\right)\omega\right\}. \quad (10)$$

Because $m\le m^*$, 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 $\Gamma_{\hat{k}}=A(k_BT)+Cv^2_{k_F}\hat{k}^2$, their expressions yield:

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

with $\mathcal{Q}'=\mathrm{const}\times\frac{\sin(q\pi/2)}{4\pi\cosh^2(\beta\omega/2)}/[\Gamma_0^2+(\omega-v_F\hat{k})^2]^{q/2}$. Therefore, this work also implies a nontrivial $\mathcal{R}$ with a linear ω, $\hat{k}$ 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 $n\to 0$. An important characteristic energy $\Delta^*(x,T)$, say the inverse of the slope of the linear in ω term in $\mathcal{R}$ contains much physics. In the CA theory $\Delta^*(x,T)\propto\Gamma_0$ vanishes at all densities x as $T\to 0$, thereby defining a line of quantum critical points. On the other hand in the ECFL calculations, the energy $\Delta^*(x,T\to 0)$ is nonzero but much smaller than the (bare) Fermi energy. However, it could vanish at a specific filling x_c: as $\Delta^*(x_c,T\to 0)\to 0$, 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 ω independent asymmetry ratio at small energies, as in the usual Fermi-liquid [6]. A similar ω 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 $\sim 2k_BT$ in ω 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 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 $\mathcal{S}^{\mathrm{s}}$. Specifically I propose that the search for a nontrivial (i.e., ω linear) asymmetry ratio $\mathcal{R}$ is important for identifying the correct underlying theoretical description of the cuprates.

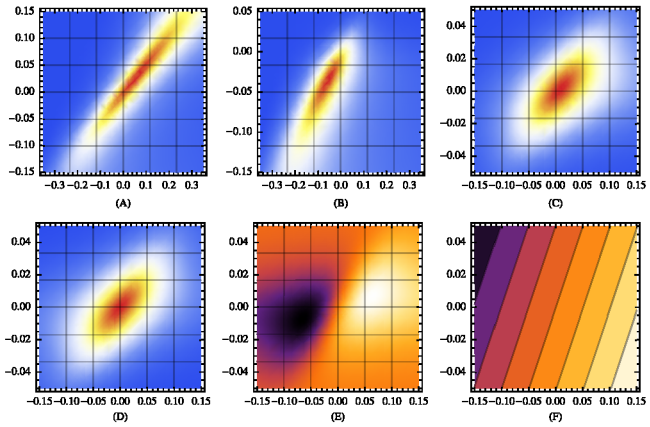

FIG. 3 (color online). Symmetry extraction illustrated for the simplified ECFL model. Here $n=0.85$, $\eta=0.05$, and $\Delta_0=0.0796$ with ω (ordinate) and ξ (abscts antisymmetricissa) in electron volt. (A) Shows the spectral function $\rho_{\mathcal{G}}$, (B) $\rho_{\mathcal{G}}f_\omega$, (C) $\rho_{\mathcal{G}}f_\omega\bar{f}_\omega$,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

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 $\hbar v_F\sim 5$ eV Å, the momentum resolution becomes critical. An error $\Delta\xi\sim 15$–20 meV can lead to quite incorrect conclusions. Thus, in order to draw unambiguous conclusions we require $\Delta k\sim 0.001$ $(\text{Å})^{-1}$, i.e., $\Delta\xi\sim 5$ 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 ω) correction to the usual ω^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 ± 25 meV around $\boldsymbol{\mu}$.

[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_ω, 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).

PHYSICAL REVIEW B **84**, 165112 (2011)

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 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,[1] one needs to deal with noncanonical electrons. A standard noncanonical problem involves the Hubbard operators[2] $X_j^{a,b}$ 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.[3]

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.[2] 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.

 ©2011 American Physical Society

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[10] of the Matsubara Greens function at finite temperatures given by

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

where G is the Greens function at a fixed wave vector k, ρ_G is its spectral density, and the integration range is $-\infty \leqslant x \leqslant \infty$. 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[2,10]

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

where $A(k)$ is the destruction operator, p_α is the Boltzmann probability of the state α given by $e^{-\beta\varepsilon_\alpha}/Z$, and ε_α 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 $\Psi(z)$:

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

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

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 ρ_G is a sum over say m delta functions located at the eigenenergies E_j (assumed distinct), so we can write the meromorphic representation

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

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

$$\sum_{j=1}^{m} a_j = a_G, \tag{6}$$

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 $\bar{G}(z) \to G(z)$. It is clear that for $z \gg \{E_j\}_{\max}$, we get the asymptotic behavior $\bar{G} \to \frac{a_G}{z}$, 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:

$$\bar{G}(z) = a_G \frac{P_{m-1}(z)}{Q_m(z)}, \quad Q(z) = \prod_{j=1}^{m}(z - E_j),$$
$$P(z) = \prod_{r=1}^{m-1}(z - \gamma_r), \tag{7}$$

where the roots γ_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:'

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

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

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

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.[12] Motivated by the structure of the theory of extremely correlated Fermi systems,[1] we next introduce the basic decomposition

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

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 $\Psi \to 0$). 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:

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

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 Ψ and rename $\Phi \to \Sigma$, then this gives the usual Dyson self-energy $\Sigma(z)$ determined uniquely using Eqs. (7) and (11) as

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

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:

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

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:

$$a_G P_{m-1} = a_G D_{m-1} + \psi_o K_{m-3},$$
$$Q_m = (z - \bar{E})D_{m-1} - \phi_0 L_{m-2}, \tag{14}$$

so that we may eliminate D and write an identity,

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

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 $K \to 0$ 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

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

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

$$\bar{G}(k,z) = \frac{\left(1 - \frac{n}{2}\right) + \Psi(k,z)}{z - E_k - \Phi(k,z)}, \quad \Psi(k,z) = \frac{B_k}{z + E_k},$$
$$\Phi(k,z) = \frac{A_k}{z + E_k}, \tag{17}$$

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 Ψ, Φ 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 $\overline{G}(k,z)$, 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 Ψ and Φ, 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

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

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 ρ_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:

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

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

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

We note the following standard result for completeness:

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

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

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

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 $Q \to \Sigma$ and conclude that

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

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

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

or more directly as

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

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 $\hat{E} \to \xi$ representing the single-particle energy measured from the chemical potential. We start with the expression:

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

and express it in terms of the two spectral functions ρ_Ψ and ρ_Φ.[13] We can write spectral function ρ_G:

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

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

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

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

The real part of G is also easily found as

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

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 Φ and Ψ from a given Σ 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 $E_k^{\rm FL}$ for a given ξ_k given as the root of

$$E_k^{\rm FL} - \xi_k - \mathcal{H}[\rho_\Phi]\left(\xi_k, E_k^{\rm FL}\right) = 0, \tag{31}$$

however, ρ_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:

$$\rho_\Psi\left(\xi_k, E_k^{\rm FL}\right) = -\rho_\Phi\left(\xi_k, E_k^{\rm FL}\right) \times \Re e\; G\left(\xi_k, E_k^{\rm FL}\right). \tag{32}$$

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[6] 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 Δ_0, an important physical parameter of the theory:

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

The physical meaning of Δ_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:[11]

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

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

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

where

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

With $\Gamma(x) = \pi\rho_\Phi(x)$, $\Re e\;\; \Phi(x+i0^+) = \mathcal{H}[\rho_\Phi](x)$ 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

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

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

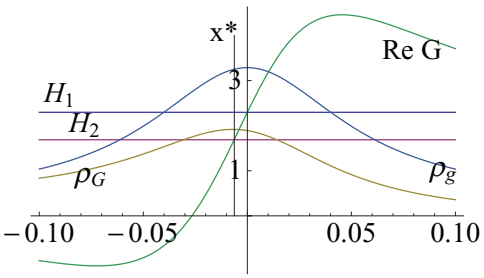

FIG. 1. (Color online) The density $n = 0.85$, temperature $T = 600$ K, $\Delta_0 = 0.0786$, and parameters are set I of Eq. (54). At this rather high temperature, we can see the details of the spectral shape clearly. The vertical line is at $x^* = E^*_{k_F}$, this energy is the location of the peak of the physical spectral function ρ_G as marked. 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

$$\rho_{\mathbf{g}}(\xi_k,x) = \frac{1}{\pi}\,\frac{\Gamma(x)}{\Gamma^2(x)+\varepsilon^2(\xi_k,x)}. \tag{39}$$

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 ξ_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 $x = E_k^{\rm FL}$ for a given ξ_k, as in standard FL theory from solving for the roots of Eq. (31). For $k \sim k_F$, we find $E_k^{\rm FL} = \xi_k Z_k$ with the momentum-independent self-energy $\Phi(z)$, where

$$Z_k = \lim_{x\to E_k^{\rm FL}} [1 - \partial\;\Re e\Phi(x)/\partial x]^{-1}. \tag{40}$$

Expanding around this solution, we write

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

We will also write $\Gamma_k \equiv \Gamma(x)/_{x\to E_k^{\rm FL}}$ at the FL quasiparticle location, where we expect for the Fermi liquid $\Gamma_k \sim c_1(k - k_F)^2 + c_2T^2$, with suitable values as described more fully in Sec. IV F). At this value, we have the identity $\Re e\; G(\xi_k, E_k^{\rm FL}) = n^2/(4\Delta_0)$ 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.

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

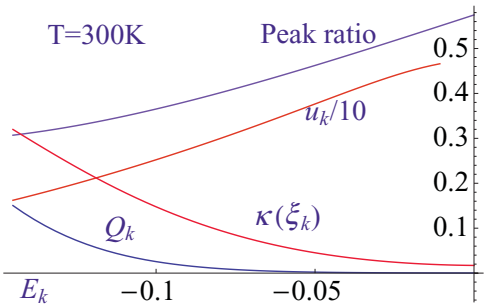

FIG. 2. (Color online) The density $n = 0.85$, $T = 300$ K, $\Delta_0 = 0.0678$, and parameters are from set I in Eq. (54). The various dimensionless variables shown against $E_k^{\rm FL}$ are the peak ratio from Eq. (50), the variable u_k (scaled by ten) from Eq. (44), the skew asymmetry factor $\kappa(\xi_k)$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

Similarly, the real part is found:

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

We introduce the following convenient positive variable u_k:

$$\sinh u_k \equiv \frac{\varepsilon_0 + \xi_k - E_k^{\rm FL}}{Z_k \Gamma_k}, \tag{44}$$

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., $\exp u_k \to 1/T^2$. 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 ξ_k, we introduce a dimensionless energy variable ϵ through the relation

$$x = E_k^{\rm FL} + Z_k \Gamma_k\, \epsilon, \tag{45}$$

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

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

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

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 $E_k^{\rm FL}(= Z_k\ \xi_k)$ as $E_k^* \equiv E_k^{\rm FL} - e^{-u_k} Z_k \Gamma_k$ at $\epsilon^* = -e^{-u_k}$. Simplifying, we find the EDC energy dispersion or spectrum

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

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

For a given ξ_k, the magnitude of the spectral function at this peak is given by

$$\Re e\, G^*(k) = \frac{n^2}{4\Delta_0}\left(1 - \frac{1}{2}Z_k\right),$$
$$\rho_G^*(k) = \frac{n^2\, Z_k}{8\pi\,\Delta_0} e^{u_k}. \tag{49}$$

The magnitude of $\Re e\, G^*(k)$ 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.[14]

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

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

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 ξ_k.[15] In the model of a $\vec{k}$-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:

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

Thus $\xi^*(x)$ is the peak position of ξ_k in constant-energy scans, whereas E_k^* in Eq. (48) represents peak position of energy at a fixed ξ_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

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

with $\tau = \pi k_B T$. We have added a scattering width η 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

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

where $\mathrm{Erfi}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}\,dt$ is the imaginary error function. A numerically small correction arising from η 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 $C_\Phi \to \frac{1}{\pi\Omega_0}$ and $\omega_c \to \frac{\omega_0^2}{\pi\Omega_0}$, in terms of the high- and low-frequency cutoff frequencies ω_0 and Ω_0. We use two sets of standard

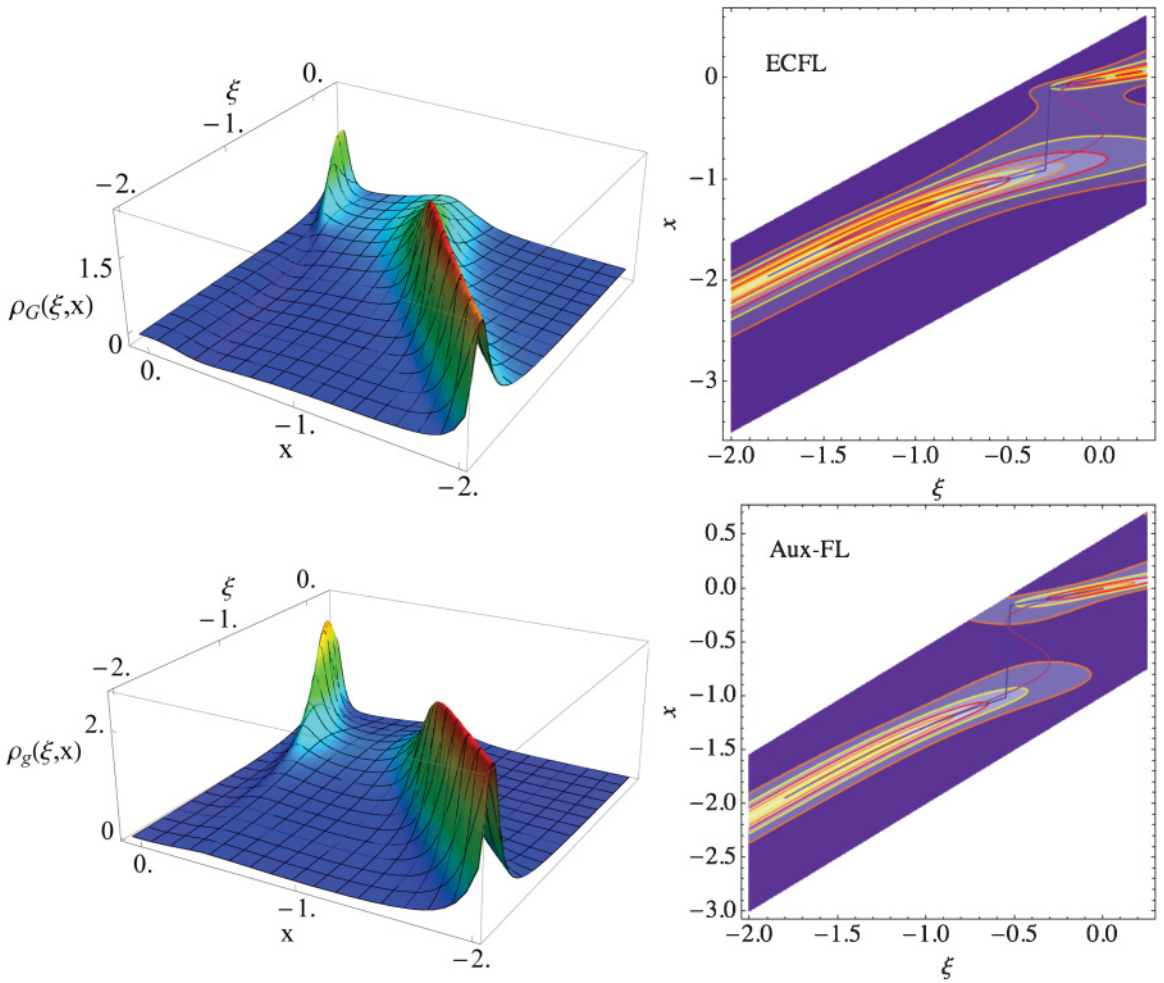

FIG. 3. (Color online) The ECFL (top left) and the auxiliary FL spectral functions (bottom left) at density $n = 0.85$, $T = 180$ K, $\Delta_0 = 0.12$ eV, $\eta = 0.12$ eV, and the other parameters are from Set II in Eq. (54). Here, ξ and x are in units of eV. In the ECFL curve on left, it is seen that the excitations near the Fermi energy become broad and dissolve into the continuum at an energy ~ -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;

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

Set I was used in Ref. 1 for schematic plots employing a simple band density of states $g_B(\varepsilon) = \frac{1}{2W}\Theta(W^2 - \varepsilon^2)$. 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 η is displayed in different plots. In Eq. (39), the spectral function ρ_g of the aux-FL is defined. The chemical potential is fixed by the number sum rule with $\xi = \epsilon - \mu$

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

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 Δ_0 is found from the sum rule:

$$\Delta_0 = \int_{-\infty}^{\infty} dx\ f(x) \int d\epsilon\ g_B(\epsilon)\, \rho_{\mathbf{g}}(\epsilon - \mu, x)\, (\epsilon - \mu - x). \tag{56}$$

Thus at a given density and temperature n, T, the model has only two parameters ω_c and C_Φ so that Δ_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

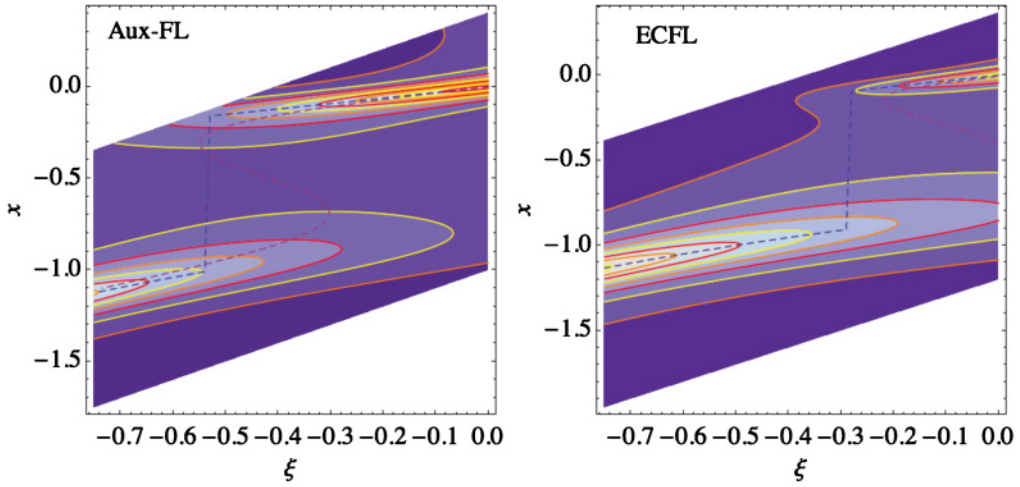

FIG. 4. (Color online) The contour plots of the aux-FL (left) and the ECFL model from Fig. 3 (right) with the same parameters as in Fig. 3 but over a smaller energy window. We superimpose the constant-wave-vector dispersion and MDC dispersion, with a value of $\eta = 0.12$ common to the contour plots. 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 ξ, found by numerical maximization), is plotted versus ξ along with the MDC spectrum Eq. (51). We recall that the latter expression is exact at all ξ 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 spectral function 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 ξ. 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 $-\Re e\Phi(x)$ occurs at an energy approximately $2\omega_0$, so that as x decreases from zero, $\xi^*_{\text{aux}-\text{FL}} = x - \Re e\Phi(x)$ goes back up for a certain range. In the case of the ECFL, Eq. (51) shows that the energy scale ε_0 enters the expression when $\Gamma(x)$ becomes comparable to $\varepsilon_0 - \Re e\Phi(x)$, and the net result is that the re-entrant behavior is pushed to lower binding energies.

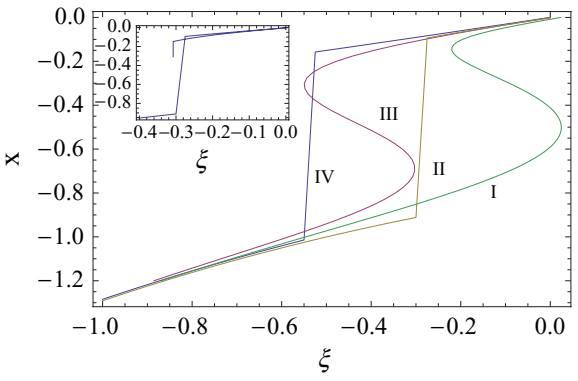

FIG. 5. (Color online) Energy dispersion curves in the ECFL and the aux-FL models. Here, the parameters are from set II in Eq. (54), with $n = 0.85$ and $T = 180$ K. With $\eta = 0.12$, curves I and II have the peaks in constant-wave-vector and constant-energy scans of the spectral function (37), and 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 $E_k^{\rm FL}$, 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

$$\Delta E_k = E_k^* - E_k^{\rm FL} = (1 - Z_k)\xi_k + \varepsilon_0 - \sqrt{[\varepsilon_0 + (1 - Z_k)\,\xi_k]^2 + Z_k^2\Gamma_k^2}, \tag{57}$$

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

$$\Delta E_k = -\frac{Z_{k_F}^2\,\Gamma_k^2}{2\varepsilon_0}. \tag{58}$$

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 ξ_k. This departure makes the dispersion in Eq. (48) depart significantly from the bare dispersion ξ_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 ε_0 and the Fermi-liquid parameter Γ_k. This energy shift is therefore also a useful method for extracting the fundamental parameter Δ_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 $[1 - \frac{n}{2} + \frac{n^2}{4\Delta_0}(\xi - x)]$, 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.

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 ξ. 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 ρ_G at various values of the energy ξ 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 ξ that shifts to the left along with the energy ξ, 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

$$\epsilon = \epsilon^* + \cosh(u_k)\bar{\epsilon}, \tag{59}$$

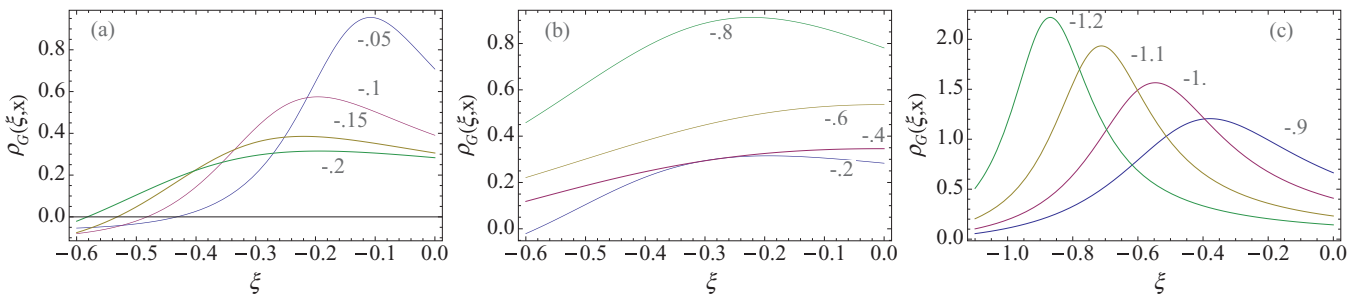

FIG. 6. (Color online) MDC line shapes at different values of energy x displayed in each curve. Here, the parameters are from set II in Eq. (54), with $n = 0.85$, $T = 180$ K, and $\eta = 0.12$. Panel (a) corresponds to x close to the chemical potential. It is interesting to note that the curves are skewed to the right, thus mirror imaging the leftward skew seen in the constant-ξ (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.

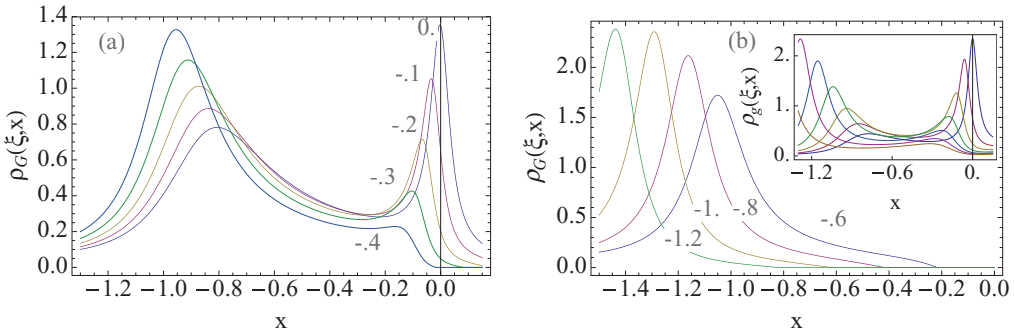

FIG. 7. (Color online) EDC line shapes at different values of energy ξ displayed in each curve. Here, the parameters are from set II in Eq. (54), with $n = 0.85$, $T = 180$ K, and $\eta = 0.12$. Panel (a) corresponds to ξ close to the chemical potential. Note that the curves are skewed to the left, i.e., 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 ξ. The inset in (b) shows the aux-FL constant-ξ scans for the same range; here, each ξ results in a pair of maxima, originating from the functional form of the self-energy in Eq. (37).

so that $\rho_G^{\rm Peak}(\xi_k,\bar{\epsilon}) = \rho_G^*(k)\gamma(Q_k,\bar{\epsilon})$, with a characteristic lineshape function γ given by

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

with

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

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

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 γ of Eq. (60) after normalizing to unit area. In order to have a well-defined quasiparticle-type peak in ρ_G for $\epsilon \sim O(1)$, Q_k must be small enough. This translates to the requirement of ξ_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

$$\varepsilon_0 + (1-Z_k)\xi_k = \frac{1}{\sqrt{3}}Z_k\Gamma_k, \tag{62}$$

beyond which it is meaningless to talk of quasiparticles. This gives ε_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):

$$\Delta_0 = \frac{n^2}{4-2n}\,\Re e\,\Phi\big(E_k^{\rm FL}\big). \tag{63}$$

Unless Δ_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 Δ_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

$$\bar{\epsilon}_{L,R} = -\tfrac{1}{2}Q_k \mp \sqrt{Q_k + \tfrac{1}{4}Q_k^2}. \tag{64}$$

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)]:

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

with the property that near the Fermi level when $u_k \to +\infty$ the variable $\kappa \to 0$, and we get a symmetric curve about the maximum. On the other hand, for deeper occupation, u_k decreases in magnitude toward zero, driving $\kappa \to 0.732$ and gives a curve that is increasingly biased to the left. The asymmetry κ 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 κ 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

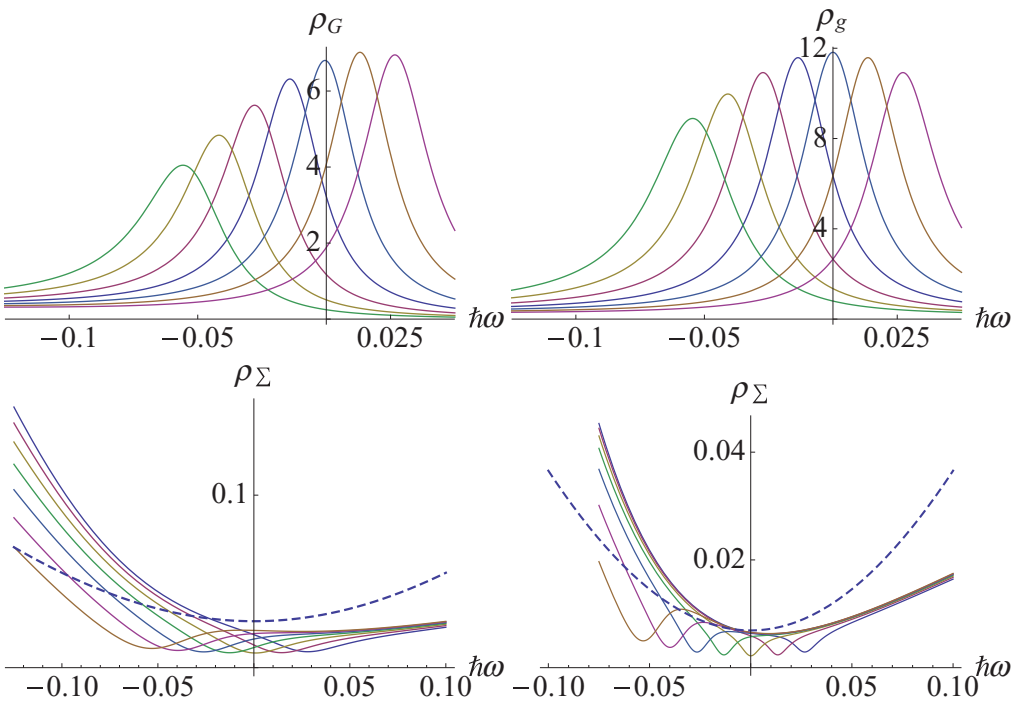

FIG. 8. (Color online) Top left panel: density $n = 0.85$, temperature $T = 300$ K, $\Delta_0 = 0.0678$, and other parameters are from set I in Eq. (54). From left to right $\rho_G(\xi,x)$ for energies in units of eV: $\xi = -0.1, -0.075, -0.05, -0.025, 0, 0.025, 0.05$. Top right panel: spectral function $\rho_g(\xi,x)$ from Eq. (39) corresponding to the same ξ 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:

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

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 conveniently determined in terms of the moments of the Greens function[7–9] in applications. The spectral function is given by

$$\rho_{\Sigma^{(1)}}(x) = \frac{a_\Sigma\ \rho_\Sigma(x)}{[\pi^2\rho_\Sigma(x)]^2 + [\Re e\ \Sigma(x)]^2}. \tag{67}$$

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

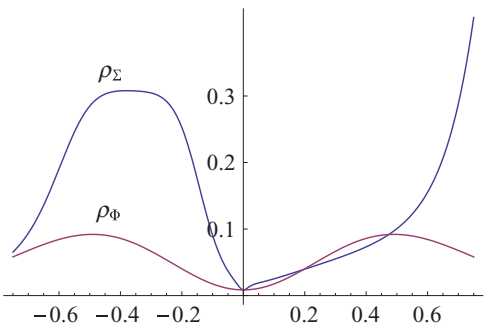

FIG. 9. (Color online) With $n = 0.85$ and $T = 300$ K, and the other parameters from set I in Eq. (54). The spectral function for the inferred Dyson self-energy $\rho_\Sigma(\xi = 0,x)$ using Eq. (23) and the the input Fermi-liquid spectral function $\rho_\Phi(x)$ over a larger energy range. Note the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

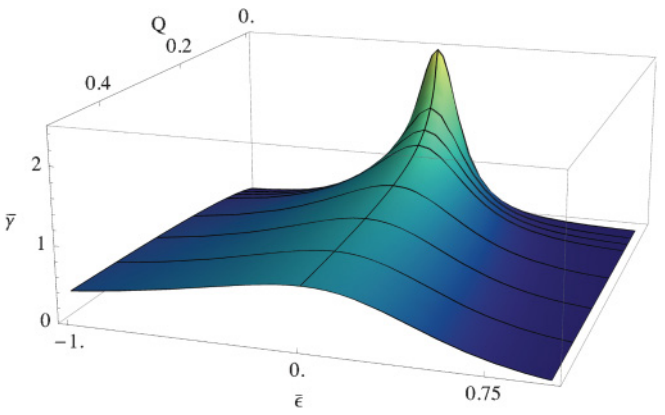

FIG. 10. (Color online) The spectral shapes possible are seen by plotting the shape function at different values of the parameter Q. In this curve, $\bar{\gamma}$ is the $\gamma(Q,\bar{\epsilon})$ of Eq. (60) normalized to unit area in the natural interval $[-1,1]$ for the variable $\bar{\epsilon}$.

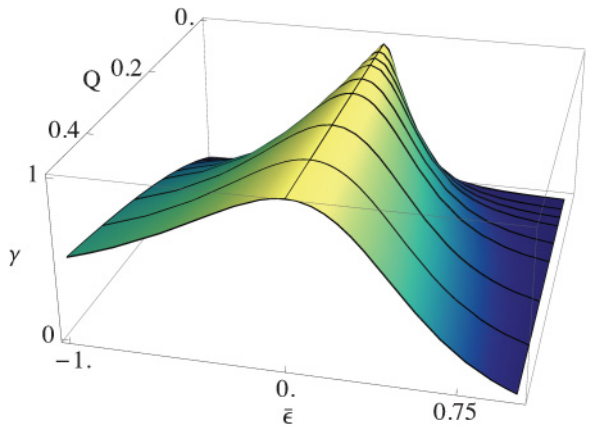

FIG. 11. (Color online) The same shapes as in Fig. 10 but now γ is normalized to unity at the peak as in Eq. (60). The sharp peaks for small $Q \leqslant 0.25$ flatten out as Q increases with a left skew asymmetry that is characteristic of this functional dependence.

next self-energy resembles closely the spectral function ρ_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.[7,8] 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:

$$C_{AB}(t) = \int_0^\beta d\tau\ \langle A(t - i\tau)B(0)\rangle,$$
$$\hat{C}_{AB}(s) = \int_0^\infty e^{-st}\ C_{AB}(t)\, dt. \tag{68}$$

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

$$\hat{C}_{AB}(s) = \int_{-\infty}^{\infty} d\nu\ \frac{\rho_{AB}(\nu)}{s - i\nu}, \quad \text{with a real density given by}$$
$$\rho_{AB}(\nu) = \sum_{nm} \frac{p_m - p_n}{\varepsilon_n - \varepsilon_m} \langle n|A|m\rangle\ \langle m|B|n\rangle\ \delta(\varepsilon_n - \varepsilon_m - \nu). \tag{69}$$

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

$$Z_0(t) \equiv \frac{C_{AB}(t)}{C_{AB}(0)}, \quad \hat{Z}_0(s) = \int_{-\infty}^{\infty} d\nu\ \frac{\rho_0(\nu)}{s - i\nu}. \tag{70}$$

Here, the real density $\rho_0(\nu) = \frac{1}{C(0)}\rho_{AB}(\nu)$, satisfies the normalization condition $\int_{-\infty}^{\infty} d\nu\ \rho_0(\nu) = 1$. 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 $\rho_0(\nu) = \frac{1}{\pi}\Re e\ \hat{Z}_0(0^+ + i\nu)$ follows. In order to find a Dyson-type representation for ρ_0, following Eqs. (8) and (9), we take the "equation of motion" by multiplying Eq. (70) by s and write

$$s\hat{Z}_0(s) = 1 + i \int_{-\infty}^{\infty} d\nu\ \frac{\nu\ \rho_0(\nu)}{s - i\nu} \equiv Y_0(s)\ \hat{Z}_0(s). \tag{71}$$

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

$$\hat{Z}_0(s) = \frac{1}{s - iY_0(s)}, \quad \text{with} \quad Y_0(s) = \frac{\int_{-\infty}^{\infty} d\nu\ \frac{\nu\ \rho_0(\nu)}{s - i\nu}}{\int_{-\infty}^{\infty} d\nu\ \frac{\rho_0(\nu)}{s - i\nu}}. \tag{72}$$

As $s \to \infty$, the function $Y_0(s)$ tends to ω_1, with a real frequency ω_1 given by

$$\omega_1 = \int_{-\infty}^{\infty} d\nu\ \nu\ \rho_0(\nu). \tag{73}$$

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

$$Y_0(s) - \ \omega_1 = i\ \alpha_1 \int_{-\infty}^{\infty} d\nu\ \frac{\rho_1(\nu)}{s - i\nu}, \tag{74}$$

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

$$Z_0(s) = \frac{1}{s - i\ \omega_1 + \alpha_1 \int\ d\nu\ \frac{\rho_1(\nu)}{s - i\nu}}. \tag{75}$$

The real number α_1 is found using the convention that $\rho_1(\nu)$ is normalized to unity. We may express ρ_1 solely in terms of the lower density $\rho_0(\nu)$ by using Eq. (70) as

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

We determine α_1 from Eq. (76) by integrating over ν 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 $\rho_0(\nu)$, the derived density $\rho_1(\nu)$ 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:

$$\hat{Z}_j(s) = \int_{-\infty}^{\infty} d\nu\ \frac{\rho_j(\nu)}{s - i\nu}, \quad \text{with normalization:}$$
$$\int_{-\infty}^{\infty} d\nu\ \rho_j(\nu) = 1. \tag{77}$$

These satisfy the recursion relation

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

where

$$\omega_{j+1} = \int_{-\infty}^{\infty} d\nu\ \nu\ \rho_j(\nu), \tag{79}$$

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

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

Note that the numbers α_j as well as ω_j are real, and for all j, the densities $\rho_j(\nu)$ are positive provided the the initial density $\rho_o(\nu)$ 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[7] 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,[3] 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.

[1] B. S. Shastry, Phys. Rev. Lett. **107**, 056403 (2011).
[2] B. S. Shastry, Phys. Rev. B **81**, 045121 (2010).
[3] G.-H. Gweon, B. S. Shastry, and G. D. Gu, Phys. Rev. Lett. **107**, 056404 (2011).
[4] D. E. Logan, M. P. Eastwood and M. A. Tusch, J. Phys. Condens. Matter **10**, 2673 (1998).
[5] Z. Wang, Y. Bang, and G. Kotliar, Phys. Rev. Lett. **67**, 2733 (1991).
[6] J. M. Luttinger and J. C. Ward, Phys. Rev **118**, 1417 (1960); J. M. Luttinger, *ibid*. **119**, 1153 (1960); **121**, 942 (1961).
[7] H. Mori, Prog. Theor. Phys. **33**, 423 (1965); **34**, 399 (1965).
[8] M. Dupuis, Prog. Theor. Phys. **37**, 502 (1967).
[9] J. J. Deisz, D. W. Hess, and J. W. Serene, Phys. Rev. **55**, 2089 (1997).
[10] 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.
[11] The positive part constraint on the right of the following equation can be often omitted, we found that it is violated very slightly ∼3% in many cases. We omit it for simplicity in the following.
[12] Shifting $\hat{E}$ by a constant is also possible but the optimal choice is the one made here.
[13] 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.
[14] At $T = 0$, this finite value persists and $\Re e G^{\rm Peak}[\xi_k, x^*(\xi_k)]$ 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 ξ_k, at the energy $x = E_k^{\rm FL}$ we find both a peak in the spectral function $\rho_g(\xi_k, E_k^{\rm FL})$ and a zero of the $\Re e g(\xi_k, E_k^{\rm FL})$, whereas at $T = 0$, we find a delta peak in the spectral function $\rho_g(\xi_k, E_k^{\rm FL})$ and a pole of the $\Re e g(\xi_k, E_k^{\rm FL})$. 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.
[15] The process of scanning ξ_k used here differs slightly from the true MDC's, where one scans the wave vector $\vec{k}$ rather than the energy ξ_k, but is more convenient here.
[16] 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 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.

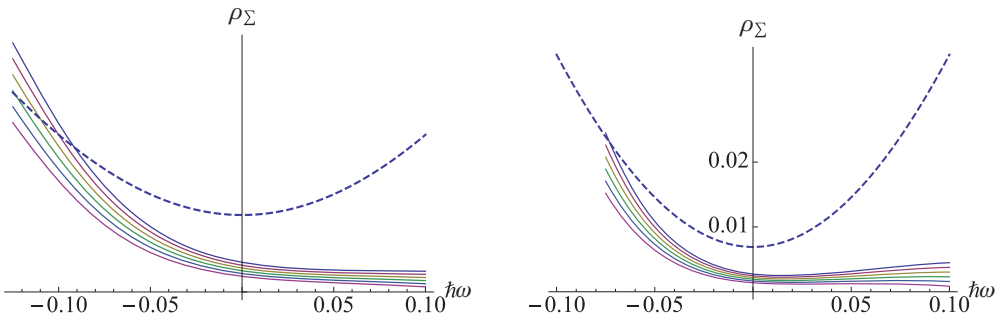

FIG. 1. (Color online) Replacement of lower panel in Fig. 8.

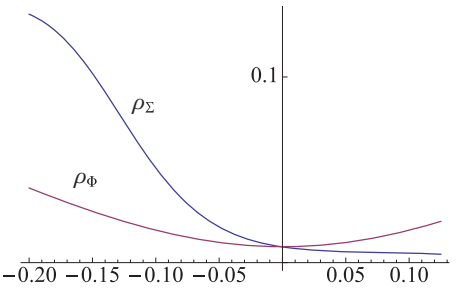

FIG. 2. (Color online) Replacement of Fig. 9.

 ©2012 American Physical Society

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

G.-H. Gweon,[1,*] B. S. Shastry,[1,†] and G. D. Gu[2]

[1]Physics Department, University of California, Santa Cruz, California 95064, USA

[2]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 $Bi_2Sr_2CaCu_2O_{8+\delta}$ and synchrotron data on $La_{1.85}Sr_{0.15}CuO_4$. 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 ($\gtrsim 15$ 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 $Bi_2Sr_2CaCu_2O_{8+\delta}$ (Bi2212) and $La_{1.85}Sr_{0.15}CuO_4$ (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° (angle).

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

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

where n is the number of electrons per 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, $A_{\rm FL}(\vec{k}, \omega) = \frac{1}{\pi}\mathrm{Im}\frac{1}{\omega - \xi_{\vec{k}} - \Phi(\omega)}$ with

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

where $\tau = \pi k_B T$, T is the temperature, and ω is to be understood as $\omega - i0^+$. Here, ω_0 is the AFL energy scale (i.e., high ω cutoff), and Ω_0 governs the lifetime, and, by causality, the quasiparticle weight (i.e., the wave function renormalization) of the AFL, $Z_{\rm FL} = (1 + \frac{\omega_0}{\sqrt{\pi}\Omega_0})^{-1}$, as identified from $\mathrm{Re}\Phi$ [6].

The ECFL energy scale Δ_0 measures the "average intrinsic inelasticity" of the AFL. It is given [4] as

 © 2011 American Physical Society

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

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 Δ_0, ω_0, $Z_{\rm FL}$, Ω_0. Of the last four parameters, only two are free parameters. For instance, ω_0 and $Z_{\rm FL}$ can be taken as free parameters, and Ω_0 and Δ_0 can be calculated using the equation for $Z_{\rm FL}$ and Eq. (3), respectively.

The parameter η 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 η. We therefore propose that this parameter summarizes the effective sample quality in different experiments. The difference in line shapes arising from these values of η 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 η to be determined separately for each class of data. The most time consuming part is the calculation of Δ_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 $Z_{\rm FL} = 1/3$, to account for the dispersion renormalization due to the high energy kink [10,12], which in this theory is caused by the energy scale ω_0 (cf. Fig. 5). (iv) Finally, in all simulations, we include the finite energy 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].

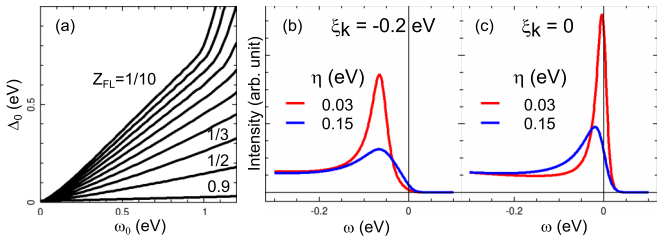

FIG. 1 (color online). (a) Δ_0 as a function of ω_0 for various $Z_{\rm FL}$. Other primary ECFL parameter values are $n = 0.85$, $T = 100$ K, and $\xi_{\vec{k}}$ as described in the text. A small η value, 0.010 eV, was used for this plot, which is used as a "lookup table" during the fit. (b),(c) Examples of the spectral function calculated with different values of the effective sample quality parameter η. See the caption of the next figure for parameter values used. The instrumental energy broadening of 10 meV (FWHM) is included.

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 ω_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 η are fixed, with one small exception in Fig. 4(d), where a slight change in ω_0 produces a much better fit over a larger energy range for LSCO.

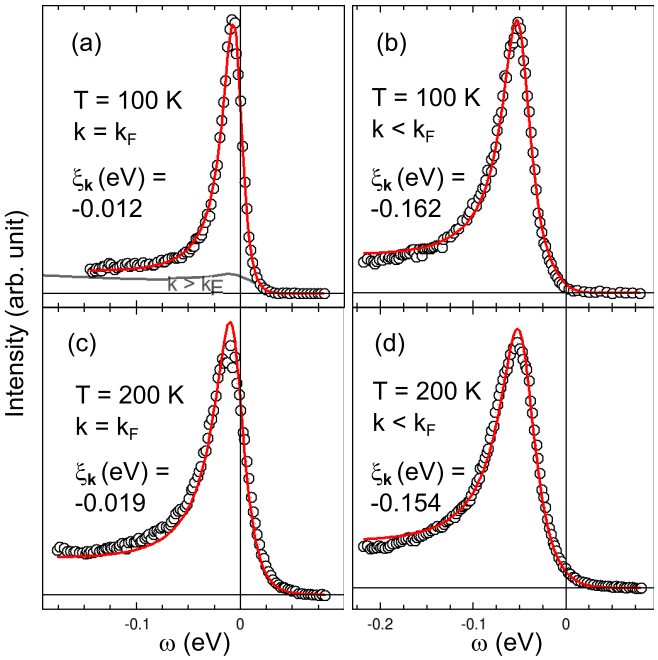

FIG. 2 (color online). Laser ARPES data (symbols, Bi2212) from Ref. [14] fit with the ECFL line shape (red lines). The free parameters of the fit were ω_0 (0.5 eV), η (0.032 eV), and $\xi_{\vec{k}}$ (shown). Fixed parameters: n (0.85), $Z_{\rm FL}$ (1/3). Derived parameters: Δ_0 (0.12 eV), Ω_0 (0.14 eV). Other than η 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 η. 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 η, 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.

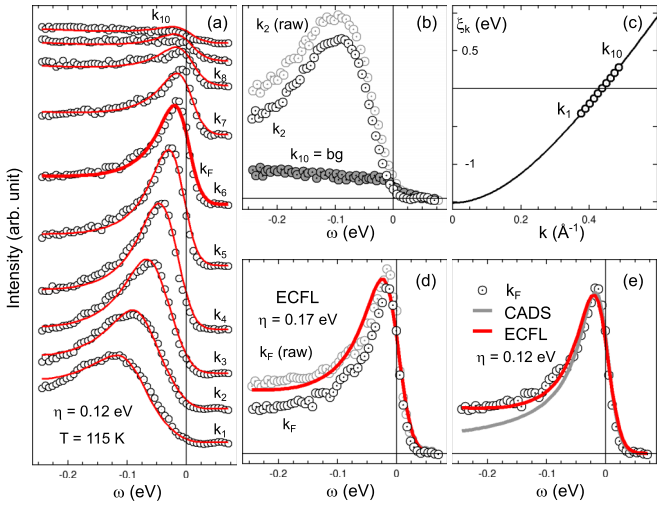

FIG. 3 (color online). Conventional ARPES data (Bi2212) fit with the ECFL line shape. The data are from Ref. [19] (T_c = 90 K). (a) The data (symbols) and the fit (red lines) are shifted vertically by the same amount for ease of view. (b) An example of the raw data and the fit data is shown for k_2. 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 $\xi_{\vec{k}}$ parameters used for the fit. Thus, in this figure, η is the only fit parameter (cf. Fig. 2 caption). (d) Raw data at $k = k_F$ fit with a somewhat greater η 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 η value (≈ 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 ω_0 and Δ_0 are determined from our fit as ~0.5 eV and ~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 ~0.3 eV to ~1 eV, as observed experimentally for the high energy kink [10,12]. As this feature

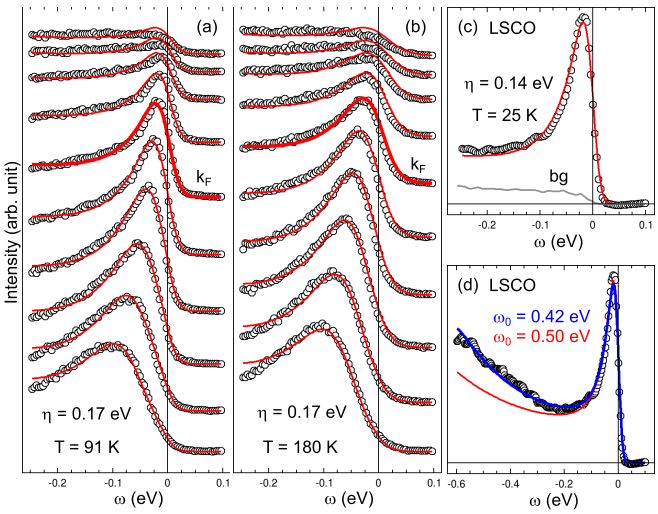

FIG. 4 (color online). Conventional ARPES data, including our own (a),(b), fit with the ECFL line shape. The procedure used to fit these data are identical with those of the previous figure, i.e., a fit with a single free parameter η, with (d) being a single exception. (a),(b), Optimally doped Bi2212 (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 ω_0 for the same data as in (c) but over a wider energy range. By changing ω_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.

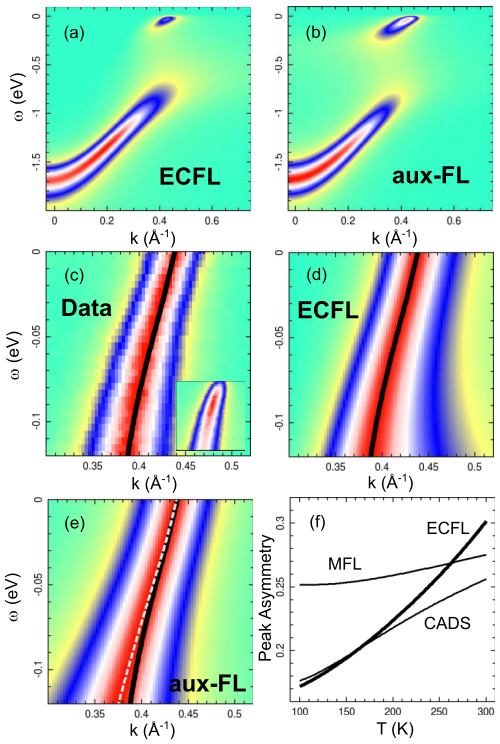

FIG. 5 (color online). Image plots of the spectral function for (a) the ECFL theory ($\eta = 0.17$ eV), and (b) that of the auxiliary FL. (c) The data of Fig. 4(a) before (inset) and after (main) MDC normalization, by which each MDC is scaled and shifted to have minimum 0 (green) and maxmimum 1 (red). 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 Δ_0 but rather with ω_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 ~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 ω value), an object discussed primarily for low energy kinks. Thus this feature originates from the scale Δ_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 $(\frac{n^2}{4}\frac{\xi_{\tilde{k}}-\omega}{\Delta_0})$ 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 $-\xi_{\tilde{k}}$ 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 (ω_0 and Δ_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. $A(\vec{k}, \omega) = \mathrm{Im}\, G(\vec{k}, \omega - i0^+)/\pi$ 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] $\mathrm{Re}\Phi(\omega) = -1/(\sqrt{\pi\Omega_0})\exp(-\frac{\tau^2}{\omega_0^2})[\omega_0\omega - 2(\omega^2+\tau^2)D(\frac{\omega}{\omega_0})]$, where $D(x) = (\sqrt{\pi}/2)e^{-x^2}\mathrm{erfi}(x)$ is the Dawson function. $C_\Phi = 1/(\pi\Omega_0)$ and $\omega_c = C_\Phi\omega_0^2$, where C_Φ and ω_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 (~ 0.02 Å) 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).

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 $U \to \infty$. 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 $U \to \infty$ 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

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

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,

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

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]:

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

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

 © 2011 American Physical Society

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 μ 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 $\Lambda^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p,q;r) = -\frac{\delta}{\delta\mathcal{V}_r^{\sigma_3\sigma_4}}\{\mathbf{g}^{-1}_{\sigma_1\sigma_2}(p,q)\}$. Thus $\mathbf{g}$, μ and $\mathbf{g}^{-1}$ are matrices in the spin space, and the vertex Λ has four indices. We also define a linear operator

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

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 ξ matrix to the source matrix $\mathcal{V}$ in the functional derivative. This notation used is illustrated in component form by $\cdots\xi^*_{\sigma_a\sigma_b}\cdots\delta/\delta\mathcal{V}^*_{\bar{\mathbf{j}}} = \cdots\sigma_a\sigma_b\cdots\delta/\delta\mathcal{V}^{\bar{\sigma}_a,\bar{\sigma}_b}_{\bar{\mathbf{j}}}$.

A useful chain rule for the functional derivative is noted

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

Using this chain rule, we see that

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

where

$$\Phi(i,m) = \mathbf{L}(i,\bar{\mathbf{i}})\cdot\mathbf{g}^{-1}(\bar{\mathbf{i}},m)$$
$$\Psi(i,m) = -\mathbf{L}(i,\bar{\mathbf{i}})\cdot\mu(\bar{\mathbf{i}},m). \tag{7}$$

Thus the two fundamental functions of this formalism Φ, Ψ 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

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

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

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

We have introduced the parameter λ above, with $0\le\lambda\le 1$, 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 λ connecting the two endpoints.

We now choose the hitherto undetermined function μ as

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

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

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

Notice that the right-hand side has a pure δ 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 μ. 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:

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

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 μ that form the core of this formalism. The hierarchy for $\mathbf{g}$ is essentially autonomous and drives that for μ. 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

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

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., $\Upsilon^{\sigma_1\sigma_2}_{\sigma_3\sigma_4}(p,q;r) = \frac{\delta}{\delta\mathcal{V}_c^{\sigma_3\sigma_4}}\{\mathcal{G}_{\sigma_1\sigma_2}(p,q)\}$.

Zero source limit in Fourier space.—When we turn off the source $\mathcal{V}$, the various matrix function $\mathcal{G}$, $\mathbf{g}$, μ 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:

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

where ε_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 Φ, and Ψ 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 $\sum_k \to \frac{1}{N_s\beta}\sum_{i\omega_k,\vec{k}}$ with N_s sites, the particle number sum rule is $\sum_k\mu(k)\mathbf{g}(k) = \frac{n}{2}$, i.e.,

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

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

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

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

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

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 $\rho_Q(x) = -\frac{1}{\pi}\mathrm{Im}\,Q(x + i0^+)$, with $x^+ \equiv x + i0^+$. The Matsubara frequency ω_Q is fermionic (bosonic) if Q is fermionic (bosonic). Proceeding further, at any order in λ, 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 μ to order $O(\lambda)^2$.—We next discuss a systematic expansion in powers of λ [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 λ, 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 λ, by taking the functional derivative of Eq. (10) and (12) we find after a brief calculation:

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

From Eq. (14) we note that these expressions Eq. (18) lead to a calculation of $\mathbf{g}^{-1}$ and μ correct up to $O(\lambda^2)$. Frequency dependent corrections arise only to second order in λ, 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

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

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

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

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 $\mathrm{Im}\,\Psi(k, x, T)$ and $\mathrm{Im}\,\bar{\Phi}(k, x, T)$. 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 $\rho_{\mathcal{G}}$:

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

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

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

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

The sign of the energy scale Δ in Eq. (22) is expected to be positive from Eq. (20). The dimensionless term η 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 η 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 $\rho_{\bar{\Psi}}(\vec{k}, x) \sim C_\Psi \sigma(x)$, as functions of frequency only. Here $\sigma(x)$ extends over energy range $\omega_c \sim O(2W)$, and C_Φ has dimensions of inverse energy and is positive due to $\rho_{\bar{\Phi}}$. Its Hilbert transform is called $h(x) \equiv \mathrm{P} \int dy \frac{\sigma(y)}{x-y}$. We use an analytically tractable Fermi liquid model [11] with $\tau = \pi k_B T$, where we set

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

The peak value of $C_\Phi \sigma(x)$ is of $O(1)$ and independent of C_Φ [12]. The other constant C_Ψ is dimensionless and negative. To complete the model, we note that the real parts are given in terms of $h(x)$ as $\mathrm{Re}\,\bar{\Phi}(x^+) = C_\Phi h(x)$ and $\mathrm{Re}\,\Psi(x^+) = C_\Psi h(x)$. 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 $\rho_{\mathbf{g}}(\xi, x) = \frac{1}{\pi} \frac{\Gamma(x)}{\Gamma^2(x)+\epsilon^2(\xi,x)}$ and $\mathrm{Re}\,\mathbf{g}(\xi, x) = \frac{\epsilon(\xi,x)}{\Gamma^2(x)+\epsilon^2(\xi,x)}$. 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 $\frac{n}{2} = \int_{-\infty}^{\infty} dx f(x) \langle \rho_{\mathbf{g}}(\xi, x)\rangle_\xi$.

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

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

The physically observable electronic spectral function reads

$$\rho_{\mathcal{G}}(\xi, x) = \frac{\Gamma(x)}{\pi} \frac{(\{1 - \frac{n}{2}\} + (\frac{n^2}{4})\{\frac{\xi - x}{\Delta_0}\})_+}{\Gamma^2(x) + \epsilon^2(\xi, x)}. \tag{26}$$

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 Δ_0 directly from the second level sum rule Eq. (16):

$$\Delta_0 = \int_{-\infty}^{\infty} dx f(x) \langle \rho_{\mathbf{g}}(\xi, x)\{\xi - x\}\rangle_\xi. \tag{27}$$

Thus $(2/n)\Delta_0$ is the average inelasticity $||(\xi - x)||$ of the FL Green's function over the entire occupied band. It vanishes if $\rho_{\mathbf{g}}$ 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 Δ_0 via the parameters C_Φ and ω_0 in the FL function $\sigma(x)$. For illustration we neglect the distinction between the band energy and the renormalized $\bar{\varepsilon}_k$, choose a flat band density of states per spin $\rho_0(\varepsilon) = \frac{1}{2W}\Theta(W^2 - \varepsilon^2)$ hence the band width is $2W$. Choose $C_\Phi = 1$ $W = 10^4$ K [14], this gives $\Delta_0 \sim 600$ 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 Δ_0 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 ω_0. In conclusion, we have presented essential ideas underlying the theory of extremely correlated Fermi liquids. We have

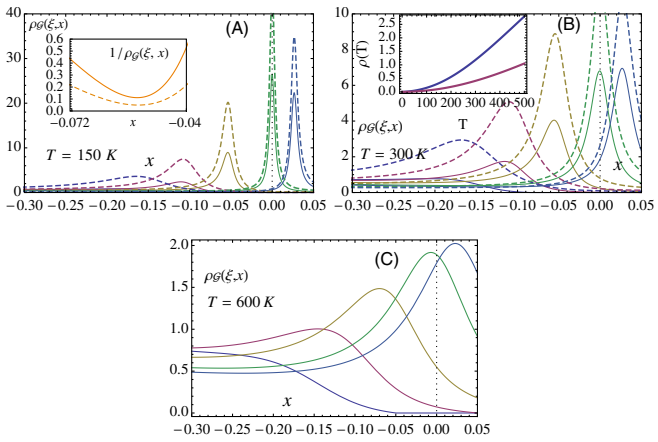

FIG. 1 (color online). The density $n = .85$ and $\omega_0 = 0.25$. From left to right $\rho_{\mathcal{G}}(x)$ for energies (in units of W) $\xi = -0.3, -0.2, -0.1, 0.,0.05$ for both the FL (dashed) and the ECFL(solid) theories. Inset in (A): provides an enlarged view of the $\xi = -0.1$ plots after *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 λ 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 $t, J, \lambda, \ldots$ are similar $\sim O(1)$, since the large parameter U of the Hubbard model has been set at ∞ at the outset. Thus λ 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 λ couples to $(\mathbf{g} \cdot \mu)^k \sim n$ in Eq. (10), an expansion in λ 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 $\bar{\varepsilon}_k = (1 - \lambda n)\varepsilon_k - \lambda \sum_p E(k, p)\mathbf{g}(p)e^{i\omega_p 0^+}$, thus writing $\mathbf{g}^{-1}(k) = i\omega_k + \boldsymbol{\mu} - \bar{\varepsilon}_k - \lambda\bar{\Phi}(k)$.

[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 $h_1(x) = \pi(x^2 + \tau^2)e^{-C_\Phi\{x^2+\tau^2\}/\omega_c}\mathrm{Erfi}(x\sqrt{C_\Phi/\omega_c})$ and $h_2(x) = x\sqrt{\pi\omega_c/C_\Phi}e^{-\tau^2 C_\Phi/\omega_c}$, where $\mathrm{Erfi}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}dt$ is the imaginary error function.

[13] In computing the various parameters self consistently, one finds that the positivity enforcing Θ function in Eq. (26) can be dropped with very little ($\sim 5\%$) error.

[14] This gives a bandwidth of ~ 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 Δ_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.