Classical dynamical coarse-grained entropy and comparison with the quantum version

Dominik Šafránek [1,*] Anthony Aguirre,[1] and J. M. Deutsch[2]

[1]*SCIPP and Department of Physics, University of California, Santa Cruz, California 95064, USA*
[2]*Department of Physics, University of California, Santa Cruz, California 95064, USA*

We develop the framework of classical observational entropy, which is a mathematically rigorous and precise framework for nonequilibrium thermodynamics, explicitly defined in terms of a set of observables. Observational entropy can be seen as a generalization of Boltzmann entropy to systems with indeterminate initial conditions, and it describes the knowledge achievable about the system by a macroscopic observer with limited measurement capabilities; it becomes Gibbs entropy in the limit of perfectly fine-grained measurements. This quantity, while previously mentioned in the literature, has been investigated in detail only in the quantum case. We describe this framework reasonably pedagogically, then show that in this framework, certain choices of coarse-graining lead to an entropy that is well-defined out of equilibrium, additive on independent systems, and that grows toward thermodynamic entropy as the system reaches equilibrium, even for systems that are genuinely isolated. Choosing certain macroscopic regions, this dynamical thermodynamic entropy measures how close these regions are to thermal equilibrium. We also show that in the given formalism, the correspondence between classical entropy (defined on classical phase space) and quantum entropy (defined on Hilbert space) becomes surprisingly direct and transparent, while manifesting differences stemming from noncommutativity of coarse-grainings and from nonexistence of a direct classical analog of quantum energy eigenstates.

I. MOTIVATION

From the introductory pedagogical level to contemporary fundamental research, and in both classical and quantum contexts, the concept of entropy has caused a great deal of confusion. Not only are there many definitions of different types of entropies appropriate to different contexts [1–6], but, we would contend, there are two distinct fundamental *notions* as to what entropy is meant to signify.

On the one hand, per the definitions of Shannon, Gibbs, or von Neumann, entropy is an information-theoretic quantity associated with the probabilities attributed to states of a system. This entropy is preserved in a closed system undergoing evolution via the classical Liouville equation or a unitary quantum operator, reflecting the preservation of information in such systems, but the entropy may change (and will generally rise) if interactions with an external system are allowed.

On the other hand, entropy can measure how "generic" a particular state-of-affairs is, as described at a coarse-grained or macroscopic level. Thermodynamic entropy, and that defined by Boltzmann as the number of microstates associated with a given macrostate, share this character. This sort of entropy can (and tends to) increase in a closed system, satisfying some version of the Second Law of Thermodynamics.

These notions are often conflated because they tend to coincide in equilibrium systems, all converging to the logarithm of the number of states compatible with some set of fixed constraints—whether those states are cells in classical state space, energy levels in quantum theory, or defined otherwise. Yet if we wish to describe systems *out* of equilibrium, in which entropy can evolve, or to evaluate entropy for small numbers of particles, more conceptual and mathematical precision is necessary.

In this paper, we argue that there is a natural and rigorous definition of entropy at the classical level for which the following holds true:

(i) It is well-defined in any classical system with a fixed phase space and probability measure over that phase space; in particular, it is well-defined out of equilibrium and for small numbers of particles.

(ii) It constitutes a generalization of, and an interpolation between, classical Gibbs and Boltzmann entropies.

(iii) It is defined in terms of a *coarse-graining* that corresponds to a partitioning of phase space.

(iv) It evolves continuously over time, and generically toward larger values, corresponding to a Second Law.

(v) It can be cleanly transcribed into the quantum context—and in fact it is the classical version of the quantum "observational entropy" introduced by [7,8].

(vi) In addition, it can be used to describe the dynamics of classical systems. Specifically, it can be used to define "dynamical thermodynamic entropy" that changes with the evolution of a system.

The idea of coarse-graining has a very long history, going back to Boltzmann [9], and coarse-grained entropies or coarse-grained free energies have emerged in numerous applications, such as fluid dynamics and Navier-Stokes equations [10,11] (clearly present in the *continuum hypothesis* [12]), statistical mechanics of fields and renormalization group [13–16], in chemical engineering in computing the entropy of mixing [17–19], and in field theory in the guise of renormalization (leading to the 1982 Nobel Prize in Physics for work on critical phenomena using the renormalization

---

*dsafrane@ucsc.edu

group [20]). And there are definitions of entropy using some type of rigorously defined coarse-graining, such as entropy of partition [21], Kolmogorov-Sinai entropy [21–24], or topological entropy [21,22]. (These mostly apply to dynamical systems.)

Nonetheless, in many works coarse-graining is treated in a rather ad hoc and nonrigorous manner, with many subtleties ignored. The coarse-grained entropy that we are going to argue for here has also appeared previously in the literature. The quantum version of this entropy predates the classical; it was first introduced by von Neumann [25,26] as a resolution to the fact that the (von Neumann) entropy does not increase in isolated systems, and then it was briefly mentioned by Wehrl [2] as "coarse-grained entropy." It was further developed and generalized to include multiple coarse-grainings by the present authors in [7,8], and it was termed "observational entropy" because it can be interpreted as the amount of knowledge an observer has about the system when they have access only to a set of coarse-grained measurements (and each measurement corresponds to an observable). Quantum observational entropy has been found to dynamically describe thermalization of isolated quantum systems [8,27], it has been discussed in relationship with other types of entropies [6], it has been found to increase under Markovian stochastic maps [28], it has been argued for as a natural candidate for entropy production [29,30] because its definition does not need an explicit temperature dependence, it was found to follow an integral fluctuation theorem and to provide an alternative perspective on the conventional definition of mechanical work in quantum systems [30] (among other properties shown there), and it has been shown to illustratively model isolated systems out of equilibrium by studying its extreme fluctuations [31].

Classical observational entropy has appeared previously [2,23,32–35], but generally just as a mention in passing,[1] but never with a clear or comprehensive discussion of it, nor with any compelling treatment of applications to thermodynamics.

Thus the aim of this paper is to motivate and define a rigorous mathematical framework of classical observational entropy, which we believe gives an elegant unification of many coarse-graining techniques, and then to derive its various properties, interpret it both from an information-theoretic perspective and from a physical perspective (by connecting it to thermodynamic entropy), and finally spell out its relations to the quantum version. In other words, we will both argue that our definition of coarse-grained entropy satisfies the desiderata in points (i)–(vi) above, and also provide the reader with intuition about the behavior of this entropy by making the connection between the classical and quantum versions clear.

II. BOLTZMANN ENTROPY

Although there are variations, the type of "state counting" entropy defined by Boltzmann [9,36,37] generally attributes a number V of fundamental microstates to a given macrostate—

defined in some terms—attributing an entropy proportional to $\ln V$ to the macrostate. To be more precise, we make use of the following definitions.

The state space of a system can be partitioned into nonoverlapping subspaces that sum up to the full state space. This partition is called a "coarse-graining" and is denoted by $\mathcal{C}$. An element of this partition is a called a macrostate.

For a microstate m in a macrostate, we attach Boltzmann entropy of $S_B(m) = \ln(V)$ to this microstate (as well as to the macrostate of which it is an element), where V is the number of microstates contained in the macrostate.

This entropy typically rises, at least on average, in any classical dynamical system out of equilibrium. Consider, for example, the prototypical system of a small box of gas that is opened within a larger box, as depicted in Fig. 1, panel (a). In the Boltzmann view, we consider the phase space of gas particles in the full box, so that immediately postopening, the gas is in a low-entropy macrostate that might be described as "all particles in the small box." Under natural evolution, the microstate tends to wander out of this macrostate and into the much larger macrostate "particles spread throughout a large box." This is depicted in Fig. 2(a).

This scheme has the advantage of being defined *throughout* the process, not just in the equilibrium states. But it is problematic in that it changes discontinuously as the microstate transitions from one macrostate to another, and in that it requires perfect knowledge of the microstate, which is never available in practice. What if one wanted to describe a more realistic situation in which the observer has only a partial knowledge about the state of the system? We will show that in such situations, classical observational entropy extends the definition of Boltzmann, and leads to the continuously evolving quantity.

Another goal of this paper, as previously mentioned, is to use this generalized and smoothly varying definition of Boltzmann entropy to describe thermodynamic entropy as a *dynamical* quantity, as opposed to the standard definition, which yields a fixed value that is completely determined by the external parameters of a system. To elaborate, in equilibrium systems, thermodynamic entropy (or more precisely, microcanonical entropy) is defined in terms of the density of states near a given overall conserved energy. Illustrated in our example, thermodynamic entropy defined this way changes only during the sudden quench of the Hamiltonian, which changes discontinuously due to the removal of the barrier during stage $1 \rightarrow 2$ in Fig. 1(a). In other words, this prescription ascribes exactly the same entropy to states 2 and 4, as illustrated in Fig. 1(b) (top). This conflicts with the fact that the state 2 is a highly nonequilibrium state from which it should be possible in principle to extract a larger amount of work than from the equilibrium state 4. We will show that with a suitable choice of coarse-graining, it is indeed possible to define a thermodynamic entropy that does not depend on the barrier removal, but rather on the underlying state of a system, as shown in Fig. 1(b) (bottom). Moreover, this quantity will correspond to the standard thermodynamic entropy in its respective equilibrium states, increasing continuously from one equilibrium configuration (closed box) to the next (open box after a long time), making it a reasonable definition of a dynamical thermodynamic entropy.

---

[1]Equation (1.26) in [2], Eq. (1) in [23] (up to a minus sign and an additive constant), Eq. (A.1) in [32], Eq. (1.1) in [33], Def. 4.1 in [34], and Eq. (11) in [35].
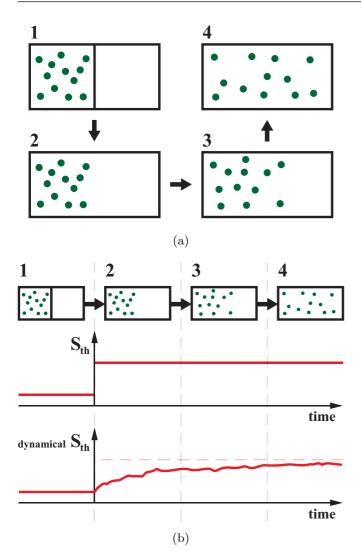
(a)

(b)

FIG. 1. (a) Evolution of an expanding gas. (b) Thermodynamic entropy S_{th}, calculated as a logarithm of energy density of states, discontinuously increases from $1 \rightarrow 2$, because the Hamiltonian (or equivalently, boundary conditions) discontinuously changes. Then it stays constant. One goal of this paper is to find an entropy measure that describes the *dynamical* process of equilibration, i.e., a measure that depends on the state of the system rather than on the boundary conditions. Such a measure is expected to stay constant as $1 \rightarrow 2$, to increase during $2 \rightarrow 4$, and to be approximately equal to thermodynamic entropy at points 1 and 4, when the system is in equilibrium. We would call such a measure a *dynamical thermodynamic entropy*.

III. CLASSICAL OBSERVATIONAL ENTROPY

We desire a generalization of Boltzmann entropy that can be applied to probability distributions given by phase-space density,

$$\rho(\boldsymbol{x}, \boldsymbol{p}; t), \tag{1}$$

over microstates. This dynamical entropy should act as a smoothing-out of the Boltzmann entropy, reducing to the actual Boltzmann entropy for a fully determined system [see Fig. 2(b)]. We also choose the phase-space density to be
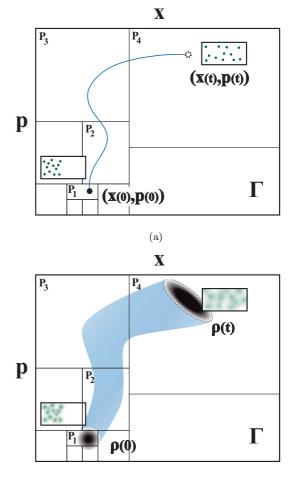

(a)

FIG. 2. Schematic picture of an evolution of a system through phase space. (a) The point in phase space naturally evolves from a small macrostate P_1 to a large macrostate P_4 and spends most of its time there. Part (b) shows the same evolution, but with indeterminate initial conditions signifying that the exact positions and momenta of particles are not known. In such situation, the state of the system can span over several macrostates at the same time.

normalized as

$$\int_{\Gamma} \rho(\boldsymbol{x}, \boldsymbol{p}; t) \, \mu \, d\boldsymbol{x} \, d\boldsymbol{p} = 1, \tag{2}$$

where μ is some normalization constant (measure) with units inverse of those of $d\boldsymbol{x} \, d\boldsymbol{p}$ so that $\rho(\boldsymbol{x}, \boldsymbol{p}; t)$ is dimensionless. By phase-space density $\rho(t) : (\boldsymbol{x}, \boldsymbol{p}) \rightarrow \rho(\boldsymbol{x}, \boldsymbol{p}; t)$ we mean a function parametrized by time t which attaches a non-negative-valued probability density to each point in phase space. Thus the probability that the state of the system is in the infinitesimal phase-space volume $d\boldsymbol{x} \, d\boldsymbol{p}$ at time t is given by $\rho(\boldsymbol{x}, \boldsymbol{p}; t)\mu \, d\boldsymbol{x} \, d\boldsymbol{p}$.

In the case of indistinguishable particles, the phase-space density is additionally required to be fully symmetric under the interchange of any two particles.

We assume that an observer—guided by some physical motivation—chooses a certain partition of this space, $\Gamma = \bigcup_i P_i$. We collect these disjoint subsets—regions of phase-space—into what we call a *coarse-graining*, and we denote it as $\mathcal{C} = \{P_i\}_i$.

Each region P_i, called a *macrostate*, is commonly defined by some inequality conditions on points of phase space $(\boldsymbol{x}, \boldsymbol{p})$. It usually represents a collection of points of phase space consistent with some macroscopically observed value or property.

The number of states in each macrostate, which we simply call its *volume*, is linearly proportional to its phase-space volume. The volume of each macrostate is computed as an integral over the macrostate with the constant[2] measure μ,

$$V_i = \int_{P_i} \mu \, d\boldsymbol{x} \, d\boldsymbol{p}. \tag{3}$$

With this definition, the constant measure can be interpreted as the inverse of the size of each microstate, and we will typically choose a conventional

$$\mu = \frac{1}{h^{Nd}}, \tag{4}$$

where d is the number of degrees of freedom of each particle ($d = 1$ for a particle moving on a line, and $d = 3$ for a particle in three-dimensional space), N is the number of particles, and h is Planck's constant. The advantage of this specific choice is that its dimensions render both ρ and V_i dimensionless (as required to take a logarithm), and because V_i then measures the actual number of quantum microstates within each macrostate, since each quantum microstate is considered to take a phase-space volume of h^{Nd}; see, e.g., [38,39].

In the case of indistinguishable particles, the microstate where the kth particle is at position and momenta (x, p), and the microstate where the lth particle is at the same point in phase space, are considered to be the same microstate. As a result, Eq. (3) is overcounting the number of microstates, and it has to be modified by a standard factor [40] to $V_i = \frac{1}{N!} \int_{P_i} \mu \, d\boldsymbol{x} \, d\boldsymbol{p}$.

To calculate the *probability of the state being in macrostate* P_i, we integrate the phase-space density over this region,

$$p_i(t) = \int_{P_i} \rho(\boldsymbol{x}, \boldsymbol{p}; t) \mu \, d\boldsymbol{x} \, d\boldsymbol{p}. \tag{5}$$

This gives a positive value, and all probabilities sum up to 1, since by definition the integral over the entire phase space is equal to 1, as shown by Eq. (2).

Defining *projectors* as window functions,

$$\hat{P}_i(\boldsymbol{x}, \boldsymbol{p}) = \begin{cases} 1, & (\boldsymbol{x}, \boldsymbol{p}) \in P_i, \\ 0, & (\boldsymbol{x}, \boldsymbol{p}) \notin P_i, \end{cases} \tag{6}$$

we can write

$$V_i = \int_{\Gamma} \hat{P}_i(\boldsymbol{x}, \boldsymbol{p}) \mu \, d\boldsymbol{x} \, d\boldsymbol{p}, \tag{7a}$$

$$p_i(t) = \int_{\Gamma} \rho(\boldsymbol{x}, \boldsymbol{p}; t) \hat{P}_i(\boldsymbol{x}, \boldsymbol{p}) \mu \, d\boldsymbol{x} \, d\boldsymbol{p}, \tag{7b}$$

where the integral now goes over the entire phase space. This can be further compacted using the L^2 inner product on real-valued functions,

$$(f, g) \equiv \int_{\Gamma} f(\boldsymbol{x}, \boldsymbol{p}) g(\boldsymbol{x}, \boldsymbol{p}) \mu \, d\boldsymbol{x} \, d\boldsymbol{p}, \tag{8}$$

so that

$$V_i = (\hat{P}_i, 1), \tag{9a}$$

$$p_i(t) = (\hat{P}_i, \rho(t)). \tag{9b}$$

[$V_i = \frac{1}{N!}(\hat{P}_i, 1)$ for indistinguishable particles.]

Since V_i denotes the number of microstates in macrostate i, and p_i denotes the probability of being in that macrostate, then assuming an observer cannot distinguish between any two microstates within the same macrostate, they associate equal probability $p_i^{(k)} := p_i/V_i$ to each microstate k within that macrostate. The Shannon entropy that this observer assigns to a system, given their inability to distinguish microstates within the same macrostate, is therefore $-\sum_{i,k} p_i^{(k)} \ln p_i^{(k)} = -\sum_i V_i \frac{p_i}{V_i} \ln \frac{p_i}{V_i}$. This motivates the definition of *classical observational entropy with coarse-graining $\mathcal{C}$* as[3]

$$S_{O(\mathcal{C})}(t) = -\sum_i p_i(t) \ln \frac{p_i(t)}{V_i}. \tag{10}$$

Since the coarse-graining $\mathcal{C}$ can be uniquely defined either by a set of macrostates P_i, or by a set of corresponding projectors $\hat{P}_i$, we can identify these two otherwise mathematically distinct objects and write

$$\mathcal{C} = \{P_i\}_i \equiv \{\hat{P}_i\}_i. \tag{11}$$

So far, our definition of coarse-grained entropy is mathematically identical to those used by Refs. [2,32–35]; the next section introduces an important generalization to multiple coarse-grainings.

IV. MULTIPLE COARSE-GRAININGS

The above definition applies for any coarse-graining, but some coarse-grainings are more relevant than others. To define a thermodynamically relevant specification of observational entropy, and for other purposes, it is necessary to generalize this entropy to include multiple coarse-grainings. This is done as follows.

Suppose we have several coarse-grainings of the phase-space, $(\mathcal{C}_1, \ldots, \mathcal{C}_n)$. We define a *joint coarse-graining*,

$$\mathcal{C}_{1,\ldots,n} = \{P_{i_1,\ldots,i_n}\}_{i_1,\ldots,i_n}, \tag{12}$$

---

[2]We demand the constant measure for the following reason: Due to Liouville's evolution, the phase-space volume of any state described by a phase-space density ρ is constant. If measure μ varied depending on $(\boldsymbol{x}, \boldsymbol{p})$, then the volume $V_\rho = \int_{\{(\boldsymbol{x}, \boldsymbol{p}) | \rho(\boldsymbol{x}, \boldsymbol{p}; t) \neq 0\}} \mu \, d\boldsymbol{x} \, d\boldsymbol{p}$ associated with phase-space density would change in time. Since this volume is supposed to represent the number of microstates within it, and since it is reasonable to demand that this number stay the same during the evolution, we must demand a constant measure μ. More generally, Liouville's equation and the property of time-independent phase-space volume is one of the primary reasons for using phase space to describe thermodynamics, as opposed to, for example, configuration space, where the volume of states does not stay constant during time evolution.

[3]This construction uses coarse-graining of Shannon entropy of measurement. We could define a slightly more general quantity by coarse-graining the Tsallis entropy of parameter q, which would lead to Tsallis observational entropy $S_{O(\mathcal{C})}^q = \frac{1}{q-1}(1 - \sum_i p_i^q V_i^{1-q})$. Observational entropy is then recovered by the limit $q \to 1$.

where the corresponding multi-macrostates are given by over-lap of the previous macrostates, and the projectors are given as a multiple of projectors,

$$P_{\boldsymbol{i}} \equiv P_{i_1,\ldots,i_n} \equiv P_{i_1} \cap \cdots \cap P_{i_n}, \tag{13}$$

$$\hat{P}_{\boldsymbol{i}} \equiv \hat{P}_{i_1,\ldots,i_n} \equiv \hat{P}_{i_1} \cdots \hat{P}_{i_n}. \tag{14}$$

In the above, we have also employed multi-index $\boldsymbol{i} = (i_1, \ldots, i_n)$.

Inserting the joint coarse-graining into the definition, Eq. (10), motivates the definition of the observational entropy with multiple coarse-grainings as

$$S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(t) \equiv -\sum_{\boldsymbol{i}} p_{\boldsymbol{i}}(t) \ln \frac{p_{\boldsymbol{i}}(t)}{V_{\boldsymbol{i}}}, \tag{15}$$

where

$$V_{\boldsymbol{i}} = (\hat{P}_{\boldsymbol{i}}, 1), \tag{16}$$

$$p_{\boldsymbol{i}}(t) = (\hat{P}_{\boldsymbol{i}}, \rho(t)). \tag{17}$$

[$V_{\boldsymbol{i}} = \frac{1}{N!}(\hat{P}_{\boldsymbol{i}}, 1)$ for indistinguishable particles.] Indeed, from the definition it follows that $S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)} = S_{O(\mathcal{C}_{1,\ldots,n})}$. (Note the subtle notational difference.) That is, a set of coarse-grainings can also be considered as a single-composite coarse-graining.[4]

V. PROPERTIES

We briefly mention some properties of classical observational entropy. These are classical equivalents of theorems proved for the quantum observational entropy [8]. The proofs are mostly analogous to those quantum mechanics (with some important subtleties), although we were able to slightly simplify some of them due to a simpler structure of the phase space as compared to the Hilbert space (see Appendix A).

Theorem 1. Observational entropy is a generalization of the Boltzmann entropy. For a single point in phase space $(\boldsymbol{x}, \boldsymbol{p}) \in P_{\boldsymbol{i}}$, equivalent to a delta function $\rho(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}}) = \delta(\tilde{\boldsymbol{x}} - \boldsymbol{x}, \tilde{\boldsymbol{p}} - \boldsymbol{p})$, we have

$$S_{O(\mathcal{C})}(\rho) = S_B(\rho) = \ln V_{\boldsymbol{i}}. \tag{18}$$

Definition 1. Finer coarse-graining. We say that coarse-graining $\mathcal{C}_2$ is *finer* than coarse-graining $\mathcal{C}_1$ (and we denote $\mathcal{C}_1 \hookrightarrow \mathcal{C}_2$) when for every $P_{i_1} \in \mathcal{C}_1$ there exists an index set $I^{(i_1)}$ such that $P_{i_1} = \bigcup_{i_2 \in I^{(i_1)}} P_{i_2}$, $P_{i_2} \in \mathcal{C}_2$. (That is, each element of $\mathcal{C}_1$ can be partitioned using elements of $\mathcal{C}_2$.)

When $\mathcal{C}_1 \hookrightarrow \mathcal{C}_2$, we can also write $\hat{P}_{i_1} = \sum_{i_2 \in I^{(i_1)}} \hat{P}_{i_2}$.

Theorem 2. Observational entropy is a monotonic function of the coarse-graining. If $\mathcal{C}_1 \hookrightarrow \mathcal{C}_2$, then

$$S_{O(\mathcal{C}_1)}(\rho) \geqslant S_{O(\mathcal{C}_2)}(\rho). \tag{19}$$

Definition 2. Coarse-graining given by an observable. Let $A : (\boldsymbol{x}, \boldsymbol{p}) \to a$ be a classical observable that assigns value a (property) to each point in phase space. We define macrostates associated with value a as $P_a = \{(\boldsymbol{x}, \boldsymbol{p})|A(\boldsymbol{x}, \boldsymbol{p}) = a\}$ in the case of an observable with discrete values, or as

$P_a = \{(\boldsymbol{x}, \boldsymbol{p})|a \leqslant A(\boldsymbol{x}, \boldsymbol{p}) < a + da\}$ in the case of a continuous observable.[5] We define coarse-graining given by the observable[6] A as $\mathcal{C}_A = \{P_a\}_a$.

Theorem 3. Observational entropy with multiple coarse-grainings is bounded,

$$S_G(\rho) \leqslant S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\rho) \leqslant \ln V, \tag{20}$$

where $S_G(\rho) \equiv -\int_\Gamma \rho(\boldsymbol{x}, \boldsymbol{p}) \ln \rho(\boldsymbol{x}, \boldsymbol{p})\mu \, d\boldsymbol{x} \, d\boldsymbol{p}$, and $V \equiv \int_\Gamma \mu \, d\boldsymbol{x} \, d\boldsymbol{p}$ is the total volume of the phase space (the total number of microstates). $S_G(\rho) = S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\rho)$ if and only if $\mathcal{C}_\rho \hookrightarrow \mathcal{C}_{1,\ldots,n}$, i.e., if the joint coarse-graining is fine enough to distinguish between points in phase space that have different assigned probabilities.

$S_G(\rho)$ represents Gibbs entropy.[7] Gibbs entropy is invariant under Liouville's evolution. It is zero for a single point in phase space, and it is a property of a state, and not of a coarse-graining. This quantity also appears as the functional H in the classical H-theorem, as interpreted by Tolman [42]. Since μ is the inverse of the phase-space volume of a single microstate, $p_i = \rho(\boldsymbol{x}, \boldsymbol{p})$ is the probability of being in a microstate (of phase-space volume μ^{-1}, or equivalently of volume $V_i = 1$), so we can also write $S_G(\rho) \equiv -\sum_i p_i \ln p_i$, i.e., the Shannon entropy of these probabilities. The quantum-mechanical equivalent of Gibbs entropy is the von Neumann entropy, which is invariant under unitary evolution, is zero for pure states, and which is also a property of a state and not of a coarse-graining.

Theorem 4. Observational entropy is nonincreasing with each added coarse-graining,

$$S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\rho) \leqslant S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_{n-1})}(\rho) \tag{21}$$

for any set of coarse-grainings $(\mathcal{C}_1, \ldots, \mathcal{C}_n)$ and any phase-space density ρ.

These theorems show that observational entropy can be elegantly interpreted as the amount of knowledge an observer would obtain if he or she were to measure the macroscopic observables that define the coarse-grainings. While Theorem 2 says that an observer with better resolution will get to know more about the system, Theorem 3 says that no matter what coarse-grained measurements they choose to perform, their knowledge will still be limited by an inherent uncertainty in the system given by the Gibbs entropy. On the other hand, regardless of the coarse-graining, their knowledge cannot be

---

[4]As discussed below, this statement does not transfer to quantum mechanics unless the coarse-grainings commute.

---

[5]The infinitesimal increment da plays the role of the resolution in measuring the observable A.

[6]Note 1: Although ρ is not usually considered an observable, it fulfills our definition and we will treat it as such in the theorem that follows. Note 2: we can also define the spectral decomposition of an observable as $A = \sum_a a\hat{P}_a$, or $A = \int_{-\infty}^\infty a\hat{P}_a \, da$, where eigenvalues a are considered to be different from each other. This spectral decomposition is unique, and it plays an identical role in phase space with inner product (8) to the role of the spectral decomposition of a quantum observable in Hilbert space.

[7]The definition of Gibbs entropy is not consistent in the literature. Our definition seems to be the most common, and it is used, for example, in [1]. Sometimes, Gibbs entropy is defined as Shannon entropy of probabilities of energy distribution (which has a quantum equivalent called diagonal entropy [41]), and what is called Gibbs entropy in [35] is actually our observational entropy.

worse than that measured by the maximal entropy, which signifies complete uncertainty about the system's state. Theorem 4 then shows an intuitive statement that each additional macroscopic measurement will provide better knowledge of the system, at least on average.

In addition, rewriting Eq. (15) as

$$S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(t) \equiv -\sum_i p_i(t) \ln p_i(t) + \sum_i p_i(t) \ln V_i \quad (22)$$

provides an intuitive information-theoretic interpretation. The first term denotes an uncertainty as measured by the Shannon entropy regarding to which macrostate the microstate of the system belongs. In other words, if one were to make a coarse-grained measurement at time t given by the coarse-graining (for example, a measurement determining the system's energy), the first term measures the uncertainty in the measurement outcomes in such a coarse-grained measurement. The second term measures the average remaining uncertainty about the microstate after this coarse-grained measurement was done. Put together, observational entropy measures the average amount of uncertainty about a microstate of a system, from the point of view of an observer that can track only certain macroscopic properties by his/her ability to perform coarse-grained measurement.

VI. THERMODYNAMICALLY RELEVANT NONEQUILIBRIUM ENTROPIES: INTRODUCTION

The treatment thus far has pertained to any possible coarse-graining. However, even if entropy increase is generic, there is no reason to expect that an arbitrary coarse-graining will be closely connected with *thermodynamics*, which in particular concerns temperature, energy, and entropy.

In the following sections, we introduce two types of (composite) coarse-grainings that each lead to an entropy that is relevant for describing the dynamics of isolated quantum systems: S_{xE} (observational entropy with local particle coarse-graining and global energy coarse-grainings) and S_F (observational entropy with local particle and local energy coarse-grainings).[8] Both will be perfectly defined for systems outside of equilibrium, both will grow to the equilibrium thermodynamic value when the system thermalizes, and both will describe physical regions coming into equilibrium with each other.

However, there are some important distinctions in their properties closely related to distinctions in their meanings: as explained in the following sections, S_{xE} can be interpreted as the value of entropy the system would attain in the long-time limit if the regions were allowed to exchange energy, but not particles. S_F, on the other hand, represents the value of entropy associated with the system if regions equilibrate but cannot exchange either energy or particles.

This interpretation of S_{xE} shows that it is not in general additive on subsystems. For example, given two equally sized regions consisting of the same number of particles but with very different temperature, S_{xE} assigns an entropy value appropriate to the two systems' temperatures having already equalized through interaction between the regions. S_F, on the other hand, is additive and equal to the sum of the thermodynamic entropies of the two regions. Additionally, since measuring local energies also determines the global value of energy, in classical physics $S_F \leqslant S_{xE}$ always.

Since in real systems both particles and energy are exchanged between the regions, both S_{xE} and S_F are inherently time-dependent quantities outside of equilibrium.

In the following sections, we focus on S_{xE} (Secs. VII and VIII), not because it is of more interest but because its definition requires only the global Hamiltonian, making it more transparent. We will derive its dynamical properties and show a simulation illustrating its behavior. Defining S_F requires a notion of local Hamiltonians. Despite this, its properties are derived in close analogy to those of S_{xE} (Sec. IX).[9]

Both of these entropies, S_{xE} and S_F, originated as fully quantum-mechanical versions published in [7,8]. However, there are some important differences. For example, unlike classical S_{xE}, the quantum S_{xE} is (quite surprisingly) an additive quantity. These will be discussed in detail in Sec. X.

VII. S_{xE}: ENTROPY OF MEASURING LOCAL PARTICLE NUMBERS AND GLOBAL ENERGY

In this section, we introduce a (composite) coarse-graining that defines one of the thermodynamically relevant entropies, S_{xE}.

Let us consider a system of N particles contained in a one-dimensional box from position L_1 to L_2, so of size $L = L_2 - L_1$. We coarse-grain this box into m physical regions (*bins*) of size $\Delta x = \frac{L}{m}$. Considering the vector of positions as $\boldsymbol{x} = (x_1, \ldots, x_N)$, and the vector of the number of particles in each part of the box as $\boldsymbol{n} = (n_1, \ldots, n_m)$ (where $n_1 + \cdots + n_m = N$), we define *local particle number (configuration) macrostates* as[10]

$P_{\boldsymbol{n}} \equiv \{(\boldsymbol{x}, \boldsymbol{p}) | n_1$ particles with position $x \in [L_1, L_1 + \Delta x),$

$\ldots, n_m$ particles with position $x \in [L_1 + (m-1)\Delta x, L_2)\}.$ (23)

Clearly, we can easily generalize this to any number d of spatial dimensions. This coarse-graining corresponds to mea-

---

[8]In the quantum context [8], S_F is also named "factorized observational entropy," or FOE for short, due to the tensor-product form of its local energy coarse-graining.

[9]While both definitions can be written in the standard Hamiltonian formalism, which assumes there is a fixed number of particles in the entire system, the description is much more elegant when one develops a formalism of the classical analog of Fock space—the *Fock phase space*—which allows for describing classical systems with a variable number of particles. After all, each subsystem (both for S_{xE} and S_F) has a variable number of particles in it, and therefore it should be possible to describe each subsystem independently. Observables of local particle number and local energies will then be seen as observables on the Fock phase space of a subsystem. The coarse-graining defining S_F is then created as a Cartesian product of local coarse-grainings. This will be done elsewhere [43].

[10]A more mathematical description of this definition uses permutations and can be inferred from Eq. (D3) by ignoring the local energy part.

suring the number of particles in each one of the m bins, i.e., to a coarse-grained measurement of local particle numbers, and we denote it as $C_X = \{P_n\}_n$.[11] For example, $P_{(3,0,1)}$ corresponds to a macrostate with three particles in the first bin, zero in the second, and one in the third.

We also define *energy macrostates* with width ΔE as

$$P_{E,\Delta E} \equiv \{(\boldsymbol{x}, \boldsymbol{p}) | E \leqslant H(\boldsymbol{x}, \boldsymbol{p}) < E + \Delta E\}, \qquad (24)$$

where H denotes the Hamiltonian. When $\Delta E \equiv dE$ is an infinitesimal increment, we simply denote $P_E \equiv P_{E,dE}$, its corresponding projector as $\hat{P}_E \equiv \hat{P}_{E,dE}$, and volume as $V_E \equiv V_{E,dE}$. We call the coarse-graining with such an infinitesimal energy increment a "fine-graining" in energy, and we denote it $C_E = \{P_E\}_E$ (considering definition 2, $C_E \equiv C_H$). It is important to emphasize here that dE is fixed to be the same for all energies E.[12]

The *local particle number-energy macrostates* are then defined as $P_{nE} \equiv P_n \cap P_E$, and the respective projectors are $\hat{P}_{nE} = \hat{P}_n \hat{P}_E$. With macrostate volumes $V_{nE} = (\hat{P}_{nE}, 1)$, and probabilities $p_{nE} = (\hat{P}_{nE}, \rho_t)$, we define observational entropy corresponding to measuring local particle numbers and the global energy as

$$S_{xE}(t) \equiv S_{O(C_X, C_H)}(t) = -\sum_{n,E} p_{nE}(t) \ln \frac{p_{nE}(t)}{V_{nE}}. \qquad (25)$$

Figure 3 gives an example of these quantities for a single particle in a harmonic potential.

VIII. THERMODYNAMIC BEHAVIOR OF S_{xE}

We now take a closer look at the thermodynamic behavior of S_{xE} using the prototypical example of a gas at first contained in half of a box, then expanding into its full volume,

---

[11]We denote coarse-graining C_X, with a subscript X signifying position mostly for historical reasons, because in the case of indistinguishable particles focused on in the quantum version of observational entropy [7,8], measuring coarse-grained positions of particles is the same as measuring the local number of particles. Thus an equivalent name for C_X is positional coarse-graining.

[12]For example, one could also think of an energy coarse-graining $C_E^{(\text{eigen})} = \{P_E^{(\text{eigen})}\}_E$, consisting of projectors onto "eigenstates" of the classical Hamiltonian, where we fit the width of each energy shell dE so that $V_E^{(\text{eigen})} = 1$, meaning that each macrostate corresponds to a single microstate. In that case, $dE = dE(E)$ will generally depend on the energy. Projectors (window functions) $\hat{P}_E^{(\text{eigen})}$ are then a classical equivalent of projectors onto quantum energy eigenstates $|E\rangle\langle E|$. However, this would not give a desirable long-time behavior for S_{xE}; in particular, the entropy assigned to a microstate in a given energy macrostate P_E would not be the correct microcanonical entropy, which is defined as a logarithm of the number of microstates in energy shell $[E, E + dE]$. Projectors $\hat{P}_E \in C_E$ with fixed dE correspond to the sum $\sum_{E \leqslant \tilde{E} < E + dE} |\tilde{E}\rangle\langle\tilde{E}|$ in quantum mechanics. We also note that correspondence between projectors onto eigenstates of the Hamiltonian $|E\rangle\langle E|$ (called *stationary Liouville eigenstates*) and $\hat{P}_E^{(\text{eigen})}$ have been explored in [44] and references therein. It was shown that the Wigner function of $|E\rangle\langle E|$ converges to a distribution that can be viewed as a highly peaked $\hat{P}_E^{(\text{eigen})}$ in the $h \to 0$ limit [see Eq. (24) in [44]].
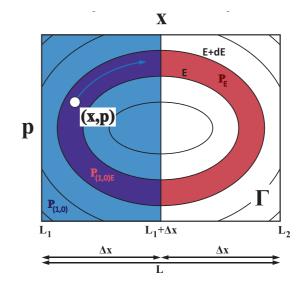

FIG. 3. Phase space of a single particle in a harmonic potential, with a local particle number (configuration) coarse-graining, and energy coarse-graining. The blue local particle number macrostate $P_{(1,0)}$ corresponds to the statement, "the particle is on the left side of the box," while the red energy macrostate P_E corresponds to "the particle has energy between E and $E + dE$." Overlap of these two macrostates $P_{(1,0)E}$ (purple) is a local particle number-energy macrostate that corresponds to "the particle is on the left side of the box and has energy between E and $E + dE$." Energy macrostates—shells of constant energy of width dE—form ellipses in this example because they are given by equations $E = \frac{1}{2m}p^2 + \frac{1}{2}kx^2 = \text{const}$. The volume of the energy macrostate V_E is proportional to the red area (including purple) in the picture, and it defines the microcanonical entropy $S_{\text{micro}}(E) \equiv \ln V_E$. In an isolated system, a particle will never jump out of its energy shell, it will only rotate through it (blue arrow). In the situation depicted, where the volume of the macrostate with particles on the left is the same as the volume of the macrostate with particles on the right, $V_{(1,0)E} = V_{(0,1)E} = \frac{1}{2}V_E$, the dynamical thermodynamic entropy has a constant value $S_{xE}(t) = \ln(\frac{1}{2}V_E) = S_{\text{micro}}(E) - \ln 2$, which is the same as the microcanonical entropy of a particle in half of the box.

as shown in Fig. 1. We will show that for the cases in which all particles are only in one of the regions, S_{xE} is equal to the thermodynamic entropy of that region, and that as the system thermalizes by letting the particles spread, S_{xE} grows to the thermodynamic entropy of the full system. This entropy is therefore enough to effectively model, for example, the expansion of an ideal gas.

We model our system by a Hamiltonian undergoing quench (a sudden discontinuous change) at time $t = 0$,

$$H(\boldsymbol{x}, \boldsymbol{p}, t) = \begin{cases} H_1(\boldsymbol{x}, \boldsymbol{p}) = H_0(\boldsymbol{x}, \boldsymbol{p}) + U_1(\boldsymbol{x}), & t < 0, \\ H_2(\boldsymbol{x}, \boldsymbol{p}) = H_0(\boldsymbol{x}, \boldsymbol{p}) + U_2(\boldsymbol{x}), & t \geqslant 0, \end{cases} \qquad (26)$$

where H_0 denotes the Hamiltonian of the particles themselves (which may or may not contain interaction between the particles), and

$$U_1(\boldsymbol{x}) = \begin{cases} 0, & \forall i, \ 0 < x_i < \frac{L}{2}, \\ +\infty, & \text{otherwise}, \end{cases} \qquad (27)$$

$$U_2(\boldsymbol{x}) = \begin{cases} 0, & \forall i, \ 0 < x_i < L, \\ +\infty, & \text{otherwise} \end{cases} \qquad (28)$$

denote two infinite potential wells.

For Hamiltonian H, the microcanonical entropy is defined as the logarithm of the number of microstates in each energy shell [45], which we can write in several ways as

$$S_{\text{micro}}^{(H)}(E) = \ln V_{E,\Delta E}^{(H)} = \ln (\rho(E)\Delta E) = S_{\text{micro}}(E, \mathcal{V}, N). \quad (29)$$

$V_{E,\Delta E}^{(H)}$ is the volume of an energy macrostate, and $\rho(E)$ denotes the density of energy states. ΔE is the width of the energy macrostate, which can be fixed to some small but nonzero value[13] that is considered to be independent of E. $\mathcal{V}$ is the physical volume occupied by N particles of total energy E. More generally, we can define thermodynamic entropy as

$$S_{\text{th}}^{(H)} \equiv S_{O(\mathcal{C}_H)}, \quad (30)$$

which coincides with the definition of microcanonical entropy for a single point in phase space, and with canonical entropy for a canonical distribution $\rho(\boldsymbol{x}, \boldsymbol{p}) = \frac{1}{Z} e^{-\beta H(\boldsymbol{x},\boldsymbol{p})}$, where β denotes inverse temperature, and $Z = \int_\Gamma e^{-\beta H(\boldsymbol{x},\boldsymbol{p})} \mu \, d\boldsymbol{x} \, d\boldsymbol{p}$ is the partition function.

In the following, we consider $S_{xE} \equiv S_{O(\mathcal{C}_X, \mathcal{C}_{H_2})}$ with a local particle number coarse-graining $\mathcal{C}_X$ that halves the box ($\boldsymbol{n} = (n_1, n_2)$, $L_1 = 0$, $L_2 = L$, $\Delta x = \frac{L}{2}$), and with energy coarse-graining $\mathcal{C}_{H_2}$ given by the Hamiltonian after the quench.

From Theorem 4, we can immediately see that thermodynamic entropy of the entire box bounds the S_{xE},[14]

$$S_{xE} = S_{O(\mathcal{C}_X, \mathcal{C}_{H_2})} = S_{O(\mathcal{C}_{H_2}, \mathcal{C}_X)} \leqslant S_{O(\mathcal{C}_{H_2})} = S_{\text{th}}^{(H_2)}. \quad (31)$$

Now we move to studying the actual dynamics of S_{xE}.

A. Dynamics of S_{xE} for determinate initial conditions

First we assume a fully determined state of N particles contained in the left part of the box, described by a state $(\boldsymbol{x}_0, \boldsymbol{p}_0) \in P_{(N,0)E}$. We have

$$S_{xE}(\boldsymbol{x}_0, \boldsymbol{p}_0) = \ln V_{(N,0)E} = \ln V_E^{(H_1)} = S_{\text{micro}}^{(H_1)}(E) = S. \quad (32)$$

The first equality comes from Theorem 1, and the second comes from the fact that the macrostate of N particles in the left side of the box with energy E is identical to the energy macrostate of the initial Hamiltonian, $P_{(N,0)E} = P_E^{(H_1)}$. Therefore, for initial particles contained in the left side of the box, S_{xE} gives the microcanonical entropy of this side of the box.

As the particles evolve, some of them will go to the right side of the box and some will remain. After some time, the state of the system can be described as, "about half of the particles are on the left, and about half of the particles are on the right side of the box."

We can say the same by saying that the point in phase space will wander around, and it will most likely end up in one of the largest macrostates. In the current situation, the largest macrostate corresponds to the statement, "half of the particles are on the left side of the box, half are on the right, and the total energy is E." This is schematically depicted as stages $1 \to 2$ in Fig. 4.

Due to the slow growth of the logarithm, the entropy associated with either of those large macrostates will not differ much: there will be some corrections, but they will become irrelevant in the thermodynamic limit. So in order to find the long-time behavior of S_{xE}, it is therefore enough to calculate the entropy for the largest of the macrostates, which is what we do in Appendix C. Assuming indistinguishable particles, we find that for m bins,[15]

$$\lim_{t \to +\infty} S_{xE}(t) = S_{\text{micro}}(E, L, N) + \frac{1}{2} \ln(2\pi N) - \frac{m}{2} \ln \frac{2\pi N}{m}, \quad (33)$$

where $S_{\text{micro}}(E, L, N) = S_{\text{micro}}^{(H_2)}(E)$ denotes the total microcanonical entropy of the full system. This means that in the case of perfect knowledge of the system, S_{xE} matches microcanonical entropy of the full system, up to some small corrections. The second term becomes negligible compared to the first in the thermodynamic limit of particle number $N \to \infty$. The last term represents a finite-size effect, which for a small number of bins m (or equivalently, large bin sizes) is also small in comparison with the first term. We will see in the next section that these corrections are an artifact of taking the determinate initial conditions: they disappear in the case of indeterminate initial conditions.

B. Dynamics of S_{xE} for indeterminate initial conditions

Let us define the phase-space density of a microcanonical ensemble as

$$\rho_{E,\Delta E}^{(\text{micro})}(\boldsymbol{x}, \boldsymbol{p}) \equiv \begin{cases} \frac{1}{V_{E,\Delta E}}, & (\boldsymbol{x}, \boldsymbol{p}) \in P_{E,\Delta E}, \\ 0, & (\boldsymbol{x}, \boldsymbol{p}) \notin P_{E,\Delta E}. \end{cases} \quad (34)$$

Equivalently, we can write

$$\rho_{E,\Delta E}^{(\text{micro})} = \frac{\hat{P}_{E,\Delta E}}{V_{E,\Delta E}}. \quad (35)$$

It is easy to see that for this microcanonical ensemble, as long as $dE \ll \Delta E$ (i.e., the energy coarse-graining is fine enough to determine the width of the ensemble), S_{xE} gives the microcanonical entropy:

$$S_{xE}(\rho_{E,\Delta E}^{(\text{micro})}) = \ln V_{E,\Delta E} = S_{\text{micro}}(E), \quad (36)$$

where $\rho(E)$ denotes the energy density of states.

---

[13]Since $\rho(E)$ rises exponentially with E and $\ln (\rho(E)\Delta E) = \ln \rho(E) + \ln \Delta E$, the choice of ΔE ultimately does not matter, because $\ln \Delta E$ acts only as an additive constant that is small compared to the first term. Moreover, since we are usually interested in the changes in entropy, rather than in absolute values, in such situations the value of this additive term becomes completely irrelevant.

[14]Symmetry $S_{O(\mathcal{C}_X, \mathcal{C}_{H_2})} = S_{O(\mathcal{C}_{H_2}, \mathcal{C}_X)}$ holds for classical observational entropy, but not for quantum observational entropy [8].

---

[15]The proof of an equivalent long-time limit behavior for quantum S_{xE} is very different from the present classical case. Instead of computing the entropy of the largest macrostate, it uses an ansatz of a random Hamiltonian, and the assumption that in the long-time limit the phases in an energy basis can be considered to be drawn randomly from a uniform ensemble (which is equivalent to saying that the differences of any two energy eigenvalues are irrational) [8].
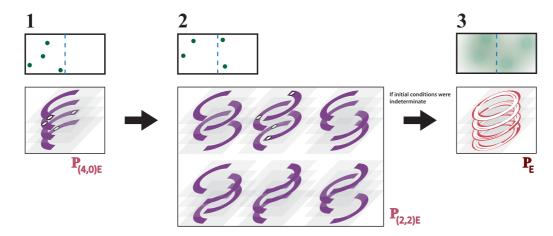
FIG. 4. Evolution of $N = 4$ indistinguishable particles in phase space, with local particle number-energy coarse-graining taken from Fig. 3. We plot the 16-dimensional phase space as four single-particle phase-spaces stacked onto each other. (1) Particles start in the left side of the box, which corresponds to macrostate $P_{(4,0)E}$ ("four particles on the left, zero on the right"), where E denotes the total energy of the system. This configuration gives microcanonical entropy of the left half of the box, $S_{xE} = \ln V_{(4,0)E} = S_{\text{micro}}^{(H_1)}(E)$. (2) As particles evolve, they wander into the largest macrostate $P_{(2,2)E}$ ("two particles on the left, two on the right") allowed by the given energy, which is six times bigger than the initial macrostate, and they spend most of their time there. The entropy of this largest macrostate is equal to the microcanonical entropy of the entire box up to some correction, $S_{xE} = \ln V_{(2,2)E} = S_{\text{micro}}^{(H_2)}(E) - \frac{1}{2}\ln(2\pi)$, $m = 2$ in Eq. (33). (3) If the initial state was not fully determined, then after some time the particle positions become uncertain, and the phase-space density becomes quite uniformly smeared over the entire energy shell $P_E = \bigcup_{n, n_1+n_2=4} P_{nE}$, which is the effect known as mixing. This erases corrections to the entropy, which then exactly equals the microcanonical entropy of the entire box: $S_{xE} = \ln V_E = S_{\text{micro}}^{(H_2)}(E)$.

More generally, in Appendix A we prove a theorem that gives meaning to all stationary states: For phase-space densities that are mixtures of energy macrostates (such as the microcanonical and canonical ensembles), S_{xE} gives the thermodynamic entropy:

Theorem 5. For phase-space density of the form $\rho = \sum_E f(E)\hat{P}_E$, where $f(E)$ is any function of energy E normalized as $\sum_E f(E)V_E = 1$,

$$S_{xE}(\rho) = S_{O(\mathcal{C}_H)} = S_{\text{th}}(\rho). \quad (37)$$

Now we address the dynamics. We consider an initial state at some time $t < 0$ that is a microcanonical state of N particles contained in the left part of the box, and none in the other half, which we can denote as a microcanonical ensemble of the initial Hamiltonian,

$$\rho_0 \equiv \rho_{E,\Delta E}^{(\text{micro};H_1)}. \quad (38)$$

Since this is a stationary state of H_1, $\rho(t) = \rho_0$ for $t < 0$.

Similar to Eq. (32), we find that for $t < 0$,

$$S_{xE}(\rho_t) = \ln V_{E,\Delta E}^{(H_1)} \equiv S_{\text{micro}}^{(H_1)}(E) = S_{\text{micro}}\left(E, \frac{L}{2}, N\right). \quad (39)$$

(See Appendix B for details.) In other words, S_{xE} of the initial state is equal to the microcanonical entropy of the first half of the box.

At time $t = 0$, the available phase space suddenly expands, and phase-space density $\rho(t)$ starts to explore the full extent of it. That first leads to a quick increase of entropy for the same reasons as for the case of determinate conditions, i.e., because the phase-space density will wander into the largest macrostate. However, due to mixing [46], with time the positions of the particles become increasingly uncertain, and the phase-space density becomes uniformly smeared over all

points in each energy shell. This is depicted as stages $2 \rightarrow 3$ in Fig. 4.

Mathematically, this means that the phase-space density converges to the microcanonical state of the second Hamiltonian,

$$\lim_{t \to +\infty} \rho(t) = \rho_{E,\Delta E}^{(\text{micro};H_2)}. \quad (40)$$

According to Theorem 5, S_{xE} of this state must be equal to the microcanonical entropy of the entire box,[16]

$$\lim_{t \to +\infty} S_{xE}(t) = \ln V_{E,\Delta E}^{(H_2)} \equiv S_{\text{micro}}^{(H_2)}(E) = S_{\text{micro}}(E, L, N). \quad (41)$$

Clearly, one can generalize this to any initial ensembles, since due to mixing, any initial phase-space density will become a stationary state of Hamiltonian H_2, in a sense of Theorem 5.

---

[16]For this convergence to hold, we must take the width of energy coarse-graining to be small, but nonzero, $0 < dE \ll \Delta E$. $0 < dE$ comes from the fact that Eq. (40) does not hold in a strict mathematical sense (there is, however, a way to write an exact mathematical statement, but we will not elaborate on this here [46,47]). This is because the phase-space density behaves like an incompressible fluid, and as such it never uniformly fills up the energy shells of phase space. However, it becomes dense in each shell, meaning that after waiting a long time, from the coarse-grained description given by nonzero dE there will be no observable difference between the real microcanonical state and a state that is dense in an energy shell. And this is enough for the entropy S_{xE} not to register a difference between such states. We are taking $dE \ll \Delta E$ because we want Eq. (36) to hold.
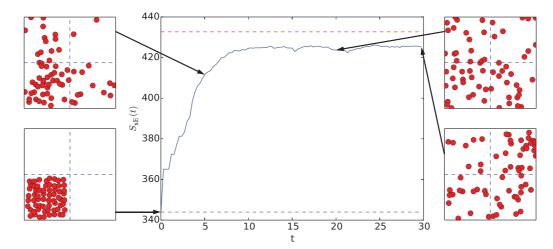
FIG. 5. Simulation of the evolution of S_{xE} of a system of $N = 64$ particles in a two-dimensional box coarse-grained into $m = 4$ physical regions, evolving through a Hamiltonian including weak interparticle interactions, and taking periodic boundary conditions. The green and red dashed lines represent thermodynamic entropy of the bottom left quarter and the full system, respectively. As the system evolves, the particles spread throughout the regions, and the S_{xE} grows to the thermodynamic entropy of the full system, up to some finite-size corrections, as expected from Eq. (33). Illustrations of the particle spread are the real snapshots of the system at different times of evolution in our simulation.

Having indeterminate initial conditions therefore results in S_{xE} converging to thermodynamic entropy exactly, without the corrections of Eq. (33).

C. Simulations

To support our analytical arguments, we have performed a simulation of a thermodynamic system of gas in $d = 2$ spatial dimension, and for 4 and 16 partitions. The case of 4 partitions is depicted in Fig. 5.

We take $N = 64$ particles of identical mass $m = 1$, and we initialize them in the lower left corner of size 8×8, within the full box of size 16×16 with periodic boundary conditions. The velocity of each particle has been randomly drawn from the normal distribution. Particles interact via a Lennard-Jones potential

$$u(r) = \varepsilon \left[\left(\frac{r_m}{r} \right)^{12} - 2 \left(\frac{r_m}{r} \right)^{6} \right], \tag{42}$$

where we took $\varepsilon = 1/120$ and $r_m = 1$ as parameters of the model, and r denotes the distance between each particle. Particles are then evolved via a velocity Verlet algorithm with time step 10^{-4}.

As particles and heat spread from one region to the other, entropy S_{xE} grows from the thermodynamic entropy of the first bin to the thermodynamic entropy of the full system, up to some finite-size corrections as expected from Eq. (33), effectively modeling thermalization of an expanding gas.

D. Interpretation

We have shown that S_{xE}, which is well-defined out of equilibrium, corresponds to the thermodynamic entropy of the initial region if all the particles are contained within this region, and it grows to thermodynamic entropy of the full system for cases of both determinate and indeterminate initial conditions.

Physically, we can interpret S_{xE} as a measure of equilibrium between the different regions defined by the coarse-graining in local particle numbers, but not necessarily as a measure of thermal equilibrium. If particles are uniformly distributed between the regions (the number of particles in each region being proportional to the size of the region), S_{xE} is equal to thermodynamic entropy. When S_{xE} is low, it means that many particles are contained in one or a few small regions, and the system is therefore in a highly nonequilibrium state. S_{xE} is therefore a measure of uniformity of particle density.

To illustrate this mathematically, for a single point in phase space, $(\boldsymbol{x}, \boldsymbol{p})(t) \in P_{(n_1, \ldots, n_m)E}$, S_{xE} gives the value

$$S_{xE}(t) = \ln V_{(n_1, \ldots, n_m)E}, \tag{43}$$

which is exactly the thermodynamic entropy that one would assign to a system of N particles distributed into m bins, in the situation in which the energy is allowed to be exchanged between the bins, but the particle number in each bin is fixed.

To imagine this quantity operationally, imagine a situation in which an observer/experimenter inserts elastic membranes in between the bins at time t (infinitely quickly), which allow for energy transfer but not for particle transfer, and then waits until the system relaxes. The value of thermodynamic entropy he or she would assign to that system after it relaxes to thermal equilibrium is $S_{xE}(t)$.

This brings out a problematic feature of this entropy, which is that the coarse-graining in global energy does not distinguish any microstates that differ in energy locally. This means that the same entropy will be associated with any microstate that has $E_1 + E_2 = E$, with E_1 and E_2 being the energy of the left region and of the right region, respectively. As an example, imagine a situation with 10 particles in the left region and 10 particles in the right region. Situations when the left part is really hot and the right part is really cold, and when both parts have equal temperatures, will both have the same

associated $S_{xE} = \ln V_{(10,10)E}$ as long as the total energy E is the same. This also means that S_{xE} is not in general additive.[17]

As we will see, factorized observational entropy S_F solves this problem by building on the definition of S_{xE} but using the *local* energy coarse-grainings instead of the global energy coarse-graining.

IX. S_F: ENTROPY OF MEASURING LOCAL PARTICLE NUMBERS AND LOCAL ENERGIES

In the same 1D system introduced in the preceding section, we define local particle number-local energy macrostates as

$$P_{nE} \equiv \{(\boldsymbol{x}, \boldsymbol{p}) | n_1 \text{ particles with position } x \in [L_1, L_1 + \Delta x),$$

$$\dots, n_m \text{ particles with position } x \in [L_1 + (m-1)\Delta x, L_2),$$

energy of the bin $[L_1, L_1 + \Delta x)$ is between $[E_1, E_1 + dE)$,

$\dots$, energy of the bin $[L_1 + (m-1)\Delta x, L_2)$ is between

$[E_m, E_m + dE)\},$

(44)

where $\boldsymbol{n} = (n_1, \dots, n_m)$ and $\boldsymbol{E} = (E_1, \dots, E_m)$ are vectors representing the local particle numbers and local energies. $N = n_1 + \dots + n_m$, $E = E_1 + \dots + E_m$, and $L = L_2 - L_1$ are the total particle number, the total energy, and the total spatial volume (length of the box). This can be generalized to any spatial dimension d and spatial volume $\mathcal{V}$, with the number of bins defined as $m = \frac{\mathcal{V}}{\Delta x^d}$.

Observational entropy with coarse-graining $\mathcal{C}_{XE} = \{P_{nE}\}_{nE}$ corresponding to measuring local particle numbers and local energies

$$S_F(t) \equiv S_{O(\mathcal{C}_{XE})}(t) = -\sum_{\boldsymbol{n}, \boldsymbol{E}} p_{nE}(t) \ln \frac{p_{nE}(t)}{V_{nE}} \quad (45)$$

satisfies all properties required of *dynamical thermodynamic entropy*.

Assuming indistinguishable particles of a sufficiently dilute and weakly interacting gas, the volume of macrostate P_{nE} is equal to the product of local volumes, $V_{nE} = V_{n_1 E_1} \cdots V_{n_m E_m}$ (see Appendix D). As a result, for a single point in phase space, $(\boldsymbol{x}, \boldsymbol{p})(t) \in P_{nE}$,

$$S_F(t) = \ln V_{nE} = \sum_{k=1}^{m} S_{\text{micro}}(E_k, \Delta x^d, n_k), \quad (46)$$

where $S_{\text{micro}}(E_k, \Delta x^d, n_k) = \ln V_{n_k E_k}$ is the microcanonical entropy of the kth bin. In other words, unlike S_{xE}, for a single point in phase space, S_F is equal to the sum of local microcanonical entropies, and it is therefore also additive on independent systems.

---

[17]The studied example of an expanding gas was a special case because there were no particles in the right part of the box, therefore this part did not contribute to the total entropy (explained in our operational view: no energy could be transferred if there were no particles on the right when the elastic membrane was inserted).

In the long-time limit, this entropy converges to the total microcanonical entropy,

$$\lim_{t \to +\infty} S_F(t) = S_{\text{micro}}(E, \mathcal{V}, N) + \ln(2\pi N) - m \ln \frac{2\pi N}{m}$$
$$- \frac{m-1}{2} \ln \frac{d}{2}, \quad (47)$$

up to some corrections that are about a factor of 2 larger than those for S_{xE}. (For $d = 1$, $\mathcal{V} = L$.) These corrections are due to determinate initial conditions (starting as a single point in phase space), and they disappear in the long-time limit when starting with indeterminate conditions due to mixing. The proof is very similar to the proof of the same property for S_{xE}, Eqs. (33) and (41), and it can also be found in Appendix D.

Similar to S_{xE}, S_F is always defined even for systems out of equilibrium, and for systems with indeterminate initial conditions, given by a phase-space density ρ.

Since the coarse-graining used by S_F is finer than that of S_{xE}, from Theorem 2 it follows that

$$S_F(t) \leqslant S_{xE}(t). \quad (48)$$

This holds up to some finite-size corrections, which comes from the fact that S_F uses local Hamiltonians[18] that ignore interaction between the bins.

S_F is a measure of thermal equilibrium between the regions. If the value of S_F is high, it means that both particles and energy are uniformly distributed across the entire system. If the value of S_F is low, it means that there are regions with a much higher density of particles and energy compared to other regions, which corresponds to a highly nonequilibrium state. S_F is therefore a measure of uniformity of both particle and energy density.

Operationally, $S_F(t)$ corresponds to a situation when an observer/experimenter inserts walls in between the bins at time t (infinitely quickly), which do not allow for either energy or particle transfer between the bins. Waiting until the system relaxes means that each region is in a thermal equilibrium with itself, but not with other regions. The thermodynamic entropy that the experimenter would assign to that system after this type of relaxation is $S_F(t)$.

X. COMPARISON OF CLASSICAL AND QUANTUM OBSERVATIONAL ENTROPY

Having investigated in detail particular coarse-grainings that connect classical observational entropy to thermodynamics, we compare observational entropy, and these particular coarse-grainings, across the classical and quantum case.

In the same way that a point in phase space describes a classical system, a vector in a Hilbert space describes a quantum system. This state vector (or wave function) encodes every property of a quantum system and is evolved using the Schrödinger equation. More generally, a quantum system with an indeterminate initial condition is described by a density matrix $\hat{\rho}$, which is a positive-semidefinite operator acting on

---

[18]For the definition of local Hamiltonians, see Appendix D.

the Hilbert space. The density matrix is a quantum equivalent of phase-space density, and it is evolved through the von Neumann equation, as compared to Liouville's equation in a classical system.

In classical systems, the coarse-graining is defined as a partitioning of phase space. These partitions/regions—subsets of phase space—are called macrostates. To define coarse-graining in quantum physics, we have to partition a Hilbert space. However, partitions of Hilbert space are not composed of subsets but of sub*spaces* combined in a direct sum: $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots$. We can form these subspaces, for example, by choosing a basis of the Hilbert space and putting every basis vector into a group. A subspace is then created by taking all the linear combinations of vectors in the group. For example, real three-dimensional space could be partitioned as $\mathbb{R}^3 = \mathbb{R}^2 \oplus \mathbb{R}$, where subspace $\mathbb{R}^2 = \text{span}\{(1,0,0),(0,1,0)\}$ describes the x-y plane, and $\mathbb{R} = \text{span}\{(0,0,1)\}$ describes the z-axis. We call these subspaces macrostates, and their collection is called a coarse-graining, $\mathcal{C} = \{\mathcal{H}_i\}_i$.

As we did with the classical system in Eq. (6), here we can also switch from describing the coarse-graining as a collection of subspaces, and we describe it as a collection of projectors that project onto these subspaces instead. In other words, for each subspace $\mathcal{H}_i$ there exists a unique projector $\hat{P}_i$ that projects onto this subspace, and the coarse-graining $\mathcal{C} = \{\hat{P}_i\}_i$ is then defined as a complete set of orthogonal projectors ($\hat{P}_i^2 = \hat{P}_i$, $\hat{P}^\dagger = \hat{P}$, $\hat{P}_i \hat{P}_j = \delta_{ij} \hat{P}_i$, $\sum_i \hat{P}_i = \hat{I}$).

With an ordered set of coarse-grainings $(\mathcal{C}_1, \ldots, \mathcal{C}_n)$, we define the volume of a multi-macrostate $\boldsymbol{i} = (i_1, \ldots, i_n)$ as

$$V_{\boldsymbol{i}} = \text{tr}\big[\hat{P}_{i_n} \cdots \hat{P}_{i_1} \cdots \hat{P}_{i_n}\big], \tag{49}$$

and the probability of being in a macrostate is defined as

$$p_{\boldsymbol{i}}(\hat{\rho}) = \text{tr}\big[\hat{P}_{i_n} \cdots \hat{P}_{i_1} \hat{\rho} \hat{P}_{i_1} \cdots \hat{P}_{i_n}\big]. \tag{50}$$

This probability can also be interpreted as a probability of obtaining the sequence of outcomes $\boldsymbol{i}$ when performing a sequence of measurements in measurement bases $(\mathcal{C}_1, \ldots, \mathcal{C}_n)$ on the system.

The (quantum) observational entropy is defined identically to the classical case, Eq. (15), but with use of the above definitions for $V_{\boldsymbol{i}}$ and $p_{\boldsymbol{i}}(\hat{\rho})$:

$$S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\hat{\rho}) = -\sum_{\boldsymbol{i}} p_{\boldsymbol{i}}(\hat{\rho}) \ln \frac{p_{\boldsymbol{i}}(\hat{\rho})}{V_{\boldsymbol{i}}}. \tag{51}$$

We compare descriptions of classical and quantum systems in Table I. Notably, coarse-grainings in the classical scenario always commute, while in the quantum scenario they need not. As a consequence, a joint coarse-graining may not exist in the quantum case, and switching the order of noncommuting coarse-grainings leads to different observational entropies. Surprisingly, this does not affect many other properties, as theorems equivalent to 1–5 still hold.

More stark differences appear when defining thermodynamically relevant coarse-grainings. Due to the noncommutativity of coarse-grainings, there are two quantum versions of S_{xE}, one that corresponds to first measuring the local particle numbers, and then energy (denoted S_{xE}), and one that corresponds to first measuring energy and then local particle numbers (denoted S_{Ex}). The behavior of quantum

S_{xE} or S_{Ex} further depends heavily on whether the energy resolution is finite or infinite, i.e., whether it is assumed that the experimenter cannot distinguish between two energy eigenstates with energies $|E_1 - E_2| < \Delta E$. Since classical energy eigenstates do not exist, the classical S_{xE} inherently entails such finite resolution in its definition.

This leads to a complex and rather subtle set of correspondences. When ΔE is non-negligible (relative, say, to the temperature of the system) and the temperature is high, these differences are negligible, and classical and quantum S_{xE} and S_{Ex} all behave in the same way. (*However*, it should be noted that in this case—as touched upon in Sec. VIII D and detailed in Appendix E—S_{xE} is not always additive.) For $\Delta E > 0$ at *low* temperatures and/or *small* bin sizes Δx, S_{xE} and S_{Ex} can differ appreciably due to the large effects of noncommutativity. Finally, for $\Delta E = 0$ (requiring the quantum case), S_{xE} and S_{Ex} are quite distinct: S_{xE} behaves much like quantum S_F (and it is additive; see Appendix E) while S_{Ex} is constant in time, equal to thermodynamic entropy S_E, and zero on any global energy eigenstate.

Turning to S_F, there is no issue of commutation, but the quantum case again gives an option of $\Delta E = 0$ or finite ΔE. The $\Delta E = 0$ case defined S_F in [7,8], and was shown to be very similar to quantum S_{xE}. One could also define a quantum S_F with $\Delta E > 0$ to closely correspond to the classical case.[19] In all cases, S_F is additive.

Generally, one could expect differences between classical and quantum definitions of entropies for systems in which the value of Planck's constant is comparable to quantities with the same dimensions that appear in the problem [48].

XI. CONCLUSION AND OUTLOOK

This paper has discussed in detail the framework of classical "observational entropy," an idea previously defined in the quantum context [7,8]. This quantity is defined precisely in or out of equilibrium, is generically nondecreasing, and corresponds to thermodynamic entropy in equilibrium. Our treatment has aimed to define this quantity clearly and rigorously, while also exhibiting in detail three core sets of relations.

First, the treatment shows how observation entropy generalizes and interpolates between classical "Gibbs" and "Boltzmann" entropies. The latter is often thought of as Gibbs entropy, where equal probability is attributed to all microstates compatible with a given set of macroscopic constraints. Here, we see Boltzmann entropy as a limit of observational entropy in which all probability is attributed to a single microstate, while Gibbs entropy appears in the limit in which the coarse-graining is as fine as possible so that each microstate constitutes a macrostate.

---

[19]For $\Delta E > 0$ it is necessary to include also coarse-graining in local particle numbers in order for this entropy to give the correct thermodynamic values. The definition then reads $S_F \equiv S_{O(\mathcal{C}_{\hat{N}_1} \otimes \cdots \otimes \mathcal{C}_{\hat{N}_m}, \mathcal{C}_{\hat{H}_1^{(\Delta E)}} \otimes \cdots \otimes \mathcal{C}_{\hat{H}_m^{(\Delta E)}})} = -\sum_{\boldsymbol{n},E} p_{\boldsymbol{n}E} \ln \frac{p_{\boldsymbol{n}E}}{V_{\boldsymbol{n}E}}$. For $\Delta E = 0$ this was not necessary, because measuring local energy eigenstates uniquely determined the local particle numbers for particle-conserving Hamiltonians.

TABLE I. Classical and quantum descriptions of an isolated physical system.

Classical	**Quantum**
Phase-space Γ	Hilbert space $\mathcal{H}$
defines a classical system,	defines a quantum system,
is a space of all possible classical states,	is a space of all possible quantum states,
all states are orthogonal,	employs Hilbert space inner product,
employs L^2-inner product for observables.	employs Hilbert-Schmidt inner product for observables.
Point in phase-space $(\boldsymbol{x}, \boldsymbol{p})$	Vector in Hilbert space $|\psi\rangle$ – wave-function
describes a state of a classical system,	describes a state of a quantum system,
is evolved through Hamilton's equations of motion,	is evolved through Schrödinger equation,
is also called a microstate.	is also called a microstate.
Phase-space density ρ	Density matrix $\hat{\rho}$
describes a classical system of indeterminate state,	describes a quantum system of indeterminate state,
is evolved through Liouville's equation.	is evolved through von Neumann equation.
Point in phase-space is described by a δ-function,	Vector in Hilbert space is described by a rank-1
$\rho(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}}) = \delta(\tilde{\boldsymbol{x}} - \boldsymbol{x}, \tilde{\boldsymbol{p}} - \boldsymbol{p})$.	density matrix, $\hat{\rho}^2 = \hat{\rho}$.
Coarse-graining $\mathcal{C} = \{P_i\}_i$	Coarse-graining $\mathcal{C} = \{\mathcal{H}_i\}_i$
is a partition of phase-space, $\Gamma = \bigcup_i P_i$.	is a partition of Hilbert space, $\mathcal{H} = \bigoplus_i \mathcal{H}_i$.
Equivalently, $\mathcal{C} = \{\hat{P}_i\}_i$	Equivalently, $\mathcal{C} = \{\hat{P}_i\}_i$
is a complete set of orthogonal window functions, $\sum_i \hat{P}_i = 1$.	is a complete set of orthogonal projectors, $\sum_i \hat{P}_i = \hat{I}$.
P_i or equivalently $\hat{P}_i$ is called a macrostate.	$\mathcal{H}_i$ or equivalently $\hat{P}_i$ is called a macrostate.
A joint coarse-graining of multiple coarse-grainings always exists.	A joint coarse-graining exists only if the coarse-grainings commute.
Observable A	Observable $\hat{A}$
is a real-valued function $A : (\boldsymbol{x}, \boldsymbol{p}) \to a$ acting on a phase space,	is a Hermitian operator acting on a Hilbert space,
admitting spectral decomposition $A = \sum_a a\hat{P}_a$,	admitting spectral decomposition $\hat{A} = \sum_a a\hat{P}_a$,
with expectation value $\langle A \rangle = (A, \rho)$,	with expectation value $\langle \hat{A} \rangle = \mathrm{tr}[\hat{A}\hat{\rho}]$,
where the probability of observing a is $p_a = (\hat{P}_a, \rho)$,	where the probability of observing a is $p_a = \mathrm{tr}[\hat{P}_a\hat{\rho}]$,
and coarse-graining given by an observable is $\mathcal{C}_A = \{\hat{P}_a\}_a$.	and coarse-graining given by an observable is $\mathcal{C}_{\hat{A}} = \{\hat{P}_a\}_a$.
Observational entropy $S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\rho) = -\sum_i p_i \ln \frac{p_i}{V_i}$,	Observational entropy $S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\hat{\rho}) = -\sum_i p_i \ln \frac{p_i}{V_i}$,
where $p_i = \int_\Gamma \rho \hat{P}_{i_1} \cdots \hat{P}_{i_n} \mu \, d\boldsymbol{x} \, d\boldsymbol{p}$ is a probability of belonging to multi-macrostate $\boldsymbol{i} = (i_1, \ldots, i_n)$,	where $p_i = \mathrm{tr}[\hat{P}_{i_n} \cdots \hat{P}_{i_1} \hat{\rho}\hat{P}_{i_1} \cdots \hat{P}_{i_n}]$ is a probability of belonging to multi-macrostate $\boldsymbol{i} = (i_1, \ldots, i_n)$,
and $V_i = \int_\Gamma \hat{P}_{i_1} \cdots \hat{P}_{i_n} \mu \, d\boldsymbol{x} \, d\boldsymbol{p}$ is volume of multi-macrostate $\boldsymbol{i}$.	and $V_i = \mathrm{tr}[\hat{P}_{i_n} \cdots \hat{P}_{i_1} \cdots \hat{P}_{i_n}]$ is volume of multi-macrostate $\boldsymbol{i}$.
A point in phase-space belongs only into a single macrostate.	A wave function can span over several macrostates.
$S_{O(\mathcal{C}_1,\mathcal{C}_2)} = S_{O(\mathcal{C}_2,\mathcal{C}_1)}$ always holds.	$S_{O(\mathcal{C}_1,\mathcal{C}_2)} = S_{O(\mathcal{C}_2,\mathcal{C}_1)}$ holds for commuting coarse-grainings, but does not hold for noncommuting coarse-grainings.
$S_G(\rho) \leqslant S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\rho) \leqslant \ln V$	$S_{VN}(\hat{\rho}) \leqslant S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\hat{\rho}) \leqslant \ln \dim \mathcal{H}$
$S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\rho) \leqslant S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_{n-1})}(\rho)$	$S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_n)}(\hat{\rho}) \leqslant S_{O(\mathcal{C}_1,\ldots,\mathcal{C}_{n-1})}(\hat{\rho})$
Dynamical thermodynamic entropy	Dynamical thermodynamic entropy
$\boldsymbol{S_{xE}} \equiv S_{O(\mathcal{C}_X,\mathcal{C}_H)} = -\sum_{\boldsymbol{n},E} p_{\boldsymbol{n}E} \ln \frac{p_{\boldsymbol{n}E}}{V_{\boldsymbol{n}E}}$,	$\boldsymbol{S_{xE}} \equiv S_{O(\mathcal{C}_{\hat{X}},\mathcal{C}_{\hat{H}})} = -\sum_{\boldsymbol{n},E} p_{\boldsymbol{n}E} \ln \frac{p_{\boldsymbol{n}E}}{V_{\boldsymbol{n}E}}$,
where $\mathcal{C}_X = \{P_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ is a set of a local particle number macrostates, and $\mathcal{C}_H = \{P_E\}_E$ is a set of energy macrostates, which are defined as energy shells of the Hamiltonian,	where $\mathcal{C}_{\hat{X}} = \{\hat{P}_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ is a set of a local particle number macrostates, and $\mathcal{C}_{\hat{H}} = \{\hat{P}_E\}_E$ is a set of energy macrostates, which are defined as projectors onto energy eigenstates,
$P_E = \{(\boldsymbol{x}, \boldsymbol{p}) | E \leqslant H(\boldsymbol{x}, \boldsymbol{p}) < E + dE\}$ (dE independent of E).	$\hat{P}_E = |E\rangle\langle E|$, $\hat{H}|E\rangle = E|E\rangle$.
Converges to thermodynamic entropy. Is not additive.	Converges to thermodynamic entropy. Is additive.
	Is not additive if a finite resolution in energy ΔE is used.
	$\boldsymbol{S_{Ex}} \equiv S_{O(\mathcal{C}_{\hat{H}(\Delta E)},\mathcal{C}_{\hat{X}})} = -\sum_{\boldsymbol{n},E} p_{E\boldsymbol{n}} \ln \frac{p_{E\boldsymbol{n}}}{V_{E\boldsymbol{n}}}$,
	where $\mathcal{C}_{\hat{H}(\Delta E)} = \{\hat{P}_{E,\Delta E}\}_E$ is a set of energy macrostates, which are defined as projectors onto coarse-grained energy macrostates, $\hat{P}_{E,\Delta E} = \sum_{E \leqslant \tilde{E} < E + \Delta E} |\tilde{E}\rangle\langle \tilde{E}|$.
Dynamical thermodynamic entropy	Dynamical thermodynamic entropy
$\boldsymbol{S_F} \equiv S_{O(\mathcal{C}_{XE})} = -\sum_{\boldsymbol{n},E} p_{\boldsymbol{n}E} \ln \frac{p_{\boldsymbol{n}E}}{V_{\boldsymbol{n}E}}$,	$\boldsymbol{S_F} \equiv S_{O(\mathcal{C}_{\hat{H}_1} \otimes \cdots \otimes \mathcal{C}_{\hat{H}_m})} = -\sum_E p_E \ln \frac{p_E}{V_E}$,
where $\mathcal{C}_{XE} = \{P_{\boldsymbol{n}E}\}_{\boldsymbol{n}E}$ is a set of a local particle number and local energy macrostates.	where $\mathcal{C}_{\hat{H}_1} \otimes \cdots \otimes \mathcal{C}_{\hat{H}_m}$ is a set of local energy macrostates.
Converges to thermodynamic entropy. Is additive.	Converges to thermodynamic entropy. Is additive.
	Is modified to include coarse-graining in local particle numbers if a finite resolution in energy ΔE is used.[19]

Second, observational entropy has both an information-theoretic and a thermodynamic interpretation. Every additional measurement on a system—corresponding to an additional coarse-graining—that better pins down a particular subset of phase space yields more information in a clear and quantifiable way. At the same time, here [Eq. (33)] and in previous work [7,8] we have shown that with an appropriate choice of coarse-graining, observational entropy describes dynamically the thermalization of a gas, converging to thermodynamic entropy when the system equilibrates. These coarse-grainings also have an elegant interpretation out of equilibrium, corresponding to the thermodynamic entropy the system would attain in equilibrium *if* particular constraints were imposed on the sharing of particles and energy across spatial regions.

Third, this work demonstrates a clear and well-defined correspondence between the quantum and classical cases. By defining projection operators that on Hilbert space and classical phase space, respectively, carefully defining volume units for classical phase space, and suitably defining quantum and classical density operators, we show that formulas and most theorems concerning observational entropy carry over directly between the classical and quantum case. The only key fundamental differences arise in the noncommutation of coarse-grainings in the quantum case, and from the nonexistence of a direct classical analog of quantum energy eigenstates. In contrast, quantum entanglement has no direct classical analog, so entanglement entropy lacks—to our knowledge—a clear correspondence with a classical quantity that behaves similarly.

Thus we view observational entropy as a framework that relates and unifies a number of disparate views and definitions of entropy, as well as providing useful new perspectives on the connection between quantum and classical treatments of phenomena. We hope that this treatment proves both illuminating and useful in understanding how entropy manifests and operates in a wide range of physical systems.

ACKNOWLEDGMENTS

APPENDIX A: PROOFS

We are going to use Jensen's inequality: Let f be a strictly concave function, $0 \leqslant a_i \leqslant 1$, $\sum_i a_i = 1$. Then for any $b_i \in \mathbb{R}$,

$$f\left(\sum_i a_i b_i\right) \geqslant \sum_i a_i f(b_i). \qquad (A1)$$

$f(\sum_i a_i b_i) = \sum_i a_i f(b_i)$ if and only if $(\forall i, j | a_i \neq 0, a_j \neq 0)(b_i = b_j)$.

Now we proceed with proofs of the paper's theorems.

Proof (Theorem 1: Observational entropy is a generalization of the Boltzmann entropy). Clearly, for a single point in phase space, $(\mathbf{x}, \mathbf{p}) \in P_i$, we have $p_i = 1$ and $p_j = 0$ for $j \neq i$.

Therefore,

$$S_{O(\mathcal{C})}(\rho) = \ln V_i = S_B(\rho). \qquad (A2)$$

Proof (Theorem 2: Observational entropy is a monotonic function of the coarse-graining). $\mathcal{C}_1 \hookrightarrow \mathcal{C}_2$ means that $P_{i_1} = \bigcup_{i_2 \in I^{(i_1)}} P_{i_2}$ for each i_1, therefore $p_{i_1} = \sum_{i_2 \in I^{(i_1)}} p_{i_2}$ and $V_{i_1} = \sum_{i_2 \in I^{(i_1)}} V_{i_2}$. Then from Jensen's inequality applied on the concave function $f(x) = -x \ln x$, taking $a_{i_2} = \frac{V_{i_2}}{V_{i_1}}$ and $b_{i_2} = \frac{p_{i_2}}{V_{i_2}}$, we have

$$\begin{aligned}
S_{O(\mathcal{C}_1)}(\rho) &= -\sum_{i_1} p_{i_1} \ln \frac{p_{i_1}}{V_{i_1}} \\
&= -\sum_{i_1} \sum_{i_2 \in I^{(i_1)}} p_{i_2} \ln \frac{\sum_{i_2 \in I^{(i_1)}} p_{i_2}}{V_{i_1}} \\
&= \sum_{i_1} V_{i_1} \left(-\sum_{i_2 \in I^{(i_1)}} \frac{V_{i_2}}{V_{i_1}} \frac{p_{i_2}}{V_{i_2}} \ln \sum_{i_2 \in I^{(i_1)}} \frac{V_{i_2}}{V_{i_1}} \frac{p_{i_2}}{V_{i_2}} \right) \\
&\geqslant \sum_{i_1} V_{i_1} \left(-\sum_{i_2 \in I^{(i_1)}} \frac{V_{i_2}}{V_{i_1}} \frac{p_{i_2}}{V_{i_2}} \ln \frac{p_{i_2}}{V_{i_2}} \right) \\
&= -\sum_{i_1} \sum_{i_2 \in I^{(i_1)}} p_{i_2} \ln \frac{p_{i_2}}{V_{i_2}} = S_{O(\mathcal{C}_2)}(\rho). \qquad (A3)
\end{aligned}$$

The equality conditions from Jensen's inequality show that $S_{O(\mathcal{C}_1)}(\hat{\rho}) = S_{O(\mathcal{C}_2)}(\hat{\rho})$ if and only if

$$\begin{aligned}
&\big(\forall i_1 | V_{i_1} \neq 0\big)\big(\forall i_2, \tilde{i}_2 \in I^{(i_1)} | V_{i_2} \neq 0, V_{\tilde{i}_2} \neq 0\big) \\
&\times \left(\frac{p_{i_2}}{V_{i_2}} = \frac{p_{\tilde{i}_2}}{V_{\tilde{i}_2}} = c^{(i_1)} \right). \qquad (A4)
\end{aligned}$$

To determine the constant $c^{(i_1)}$, we multiply the equation by V_{i_2} and sum over all $\forall i_2 \in I^{(i_1)}$, which gives

$$c^{(i_1)} = \frac{p_{i_1}}{V_{i_1}}. \qquad (A5)$$

And considering that for all $V_{i_2} = 0$ also $p_{i_2} = 0$, we can simplify Eq. (A4), and we obtain that $S_{O(\mathcal{C}_1)}(\hat{\rho}) = S_{O(\mathcal{C}_2)}(\hat{\rho})$ if and only if

$$\big(\forall i_1 | V_{i_1} \neq 0\big)\big(\forall i_2 \in I^{(i_1)}\big)\left(p_{i_2} = \frac{V_{i_2}}{V_{i_1}} p_{i_1} \right). \qquad (A6)$$

Proof (Theorem 3: Observational entropy with multiple coarse-grainings is bounded). Since observational entropy with multiple coarse-grainings can be rewritten as observational entropy with a joint coarse-graining,

$$S_{O(\mathcal{C}_1, \ldots, \mathcal{C}_n)}(\rho) = S_{O(\mathcal{C}_{1,\ldots,n})}(\rho), \qquad (A7)$$

it is enough to prove the inequalities just for a single coarse-graining $\mathcal{C}$,

$$S_G(\rho) \leqslant S_{O(\mathcal{C})}(\rho) \leqslant \ln V. \qquad (A8)$$

For the second inequality, we define coarse-graining with a single element—the entire phase-space—$\mathcal{C}_\Gamma = \{\Gamma\}$. Clearly, this coarse-graining is coarser than any other coarse-graining, therefore from Theorem 2 we have

$$S_{O(\mathcal{C})}(\rho) \leqslant S_{O(\mathcal{C}_\Gamma)}(\rho) = -p_\Gamma \ln \frac{p_\Gamma}{V_\Gamma} = -\ln \frac{1}{V} = \ln V. \qquad (A9)$$

The first inequality comes from the fact that one can choose a coarse-graining where macrostates are single points in phase space. This coarse-graining is finer than any other coarse-graining, and one can easily derive that the observational entropy is then equal to the Gibbs entropy. By the same argument, we can therefore obtain the first inequality. However, this argument does not give us the equality conditions, for which we will have to make a more elaborate derivation as follows.

To prove the first inequality and the equality conditions, we define the spectral decomposition of the phase-space density in its eigenvector projectors $\rho = \sum_{(x,p)} \rho(x,p)\hat{P}_{(x,p)}$ [meaning $\rho(\tilde{x}, \tilde{p}) = \sum_{(x,p)} \rho(x,p)\hat{P}_{(x,p)}(\tilde{x}, \tilde{p})$, where eigenvalues $\rho(x,p)$ do not have to be different for different (x,p)]. The eigenvector projectors projected onto infinitesimal regions surrounding (x,p) are defined as

$$\hat{P}_{(x,p)}(\tilde{x}, \tilde{p}) = \begin{cases} 1, & (\tilde{x}, \tilde{p}) \in [x, x+dx) \times [p, p+dp), \\ 0 & \text{otherwise.} \end{cases} \tag{A10}$$

We also define spectral decomposition of the density matrix in a form where the eigenvector projectors associated with the same eigenvalue are now grouped together,

$$\rho = \sum_\lambda \lambda \hat{P}_\lambda, \tag{A11}$$

where eigenvalues λ are now different from each other. This decomposition is unique. It follows that for each (x,p) there exists λ such that $\rho(x,p) = \lambda$.

Now recall Eq. (9): the probability of a state being in a macrostate i of volume $V_i = (\hat{P}_i, 1)$ is $p_i = (\hat{P}_i, \rho)$.

Defining

$$a_{(x,p)}^{(i)} \equiv \frac{(\hat{P}_i, \hat{P}_{(x,p)})}{V_i} \tag{A12}$$

for $V_i \neq 0$ and $a_{(x,p)}^{(i)} \equiv 0$ for $V_i = 0$, and then using the spectral decomposition of ρ, we have

$$\frac{p_i}{V_i} = \frac{\left(\hat{P}_i, \sum_{(x,p)} \rho(x,p)\hat{P}_{(x,p)}\right)}{V_i} = \sum_{(x,p)} \rho(x,p)a_{(x,p)}^{(i)}. \tag{A13}$$

Since $\sum_{(x,p)} \hat{P}_{(x,p)} = 1$,

$$\sum_{(x,p)} a_{(x,p)}^{(i)} = 1, \tag{A14}$$

and since $\sum_i \hat{P}_i = 1$, also

$$\sum_i V_i a_{(x,p)}^{(i)} = (1, \hat{P}_{(x,p)}) = \mu \, dx \, dp. \tag{A15}$$

A series of equalities and inequalities follow:

$$S_{O(\mathcal{C}_1,...,\mathcal{C}_n)}(\hat{\rho}) = -\sum_i p_i \ln \frac{p_i}{V_i} = -\sum_i V_i \frac{p_i}{V_i} \ln \frac{p_i}{V_i}$$

$$= \sum_i V_i \left(-\sum_{(x,p)} \rho(x,p)a_{(x,p)}^{(i)} \ln \right.$$

$$\left. \times \sum_{(x,p)} \rho(x,p)a_{(x,p)}^{(i)} \right)$$

$$\geqslant \sum_i V_i \left(-\sum_{(x,p)} \rho(x,p)a_{(x,p)}^{(i)} \ln \rho(x,p) \right)$$

$$= -\sum_{(x,p)} \left(\sum_i V_i a_{(x,p)}^{(i)} \right) \rho(x,p) \ln \rho(x,p)$$

$$= -\int_\Gamma \rho(x,p) \ln \rho(x,p) \mu \, dx \, dp \equiv S_G(\rho). \tag{A16}$$

We have used Eqs. (A13) and (A15) for the equalities, and we applied Jensen's theorem to the function $f(x) = -x \ln x$ to derive the inequality. We have chosen $a_{(x,p)} \equiv a_{(x,p)}^{(i)}$ and $b_{(x,p)} \equiv \rho(x,p)$ for the theorem.

According to Jensen's theorem, inequality becomes equality if and only if

$$(\forall i | V_i \neq 0)(\forall (x,p), (\tilde{x}, \tilde{p}) | (\hat{P}_i, \hat{P}_{(x,p)}) \neq 0, (\hat{P}_i, \hat{P}_{(\tilde{x},\tilde{p})}) \neq 0)$$

$$\times (\rho(x,p) = \rho(\tilde{x}, \tilde{p})). \tag{A17}$$

That is, the inequality becomes equality when for a given index i, all eigenvector projectors $\hat{P}_{(\tilde{x},\tilde{p})}$ of the phase-space density such that $(\hat{P}_i, \hat{P}_{(x,p)}) \neq 0$ have the same associated eigenvalue $\rho(x,p)$ with them. In other words, we can associate this unique eigenvalue to the index i itself, $\rho_i \equiv \rho(x,p)$, where $\rho(x,p)$ is given by any representative (x,p) such that $(\hat{P}_i, \hat{P}_{(x,p)}) \neq 0$. Realizing that $(\hat{P}_i, \hat{P}_{(x,p)}) \neq 0$ is the same as saying $P_i \cap P_{(x,p)} \neq \oslash$, or $\hat{P}_{(x,p)}\hat{P}_i \neq 0$, we can rewrite Eq. (A17) as

$$(\forall i | V_i \neq 0)(\forall (x,p), (\tilde{x}, \tilde{p}) | \hat{P}_{(x,p)}\hat{P}_i \neq 0, \hat{P}_{(\tilde{x},\tilde{p})}\hat{P}_i \neq 0)$$

$$\times (\rho(x,p) = \rho(\tilde{x}, \tilde{p}) = \rho_i). \tag{A18}$$

Assuming that this holds, we can write

$$\rho\hat{P}_i = \sum_{(x,p)} \rho(x,p)\hat{P}_{(x,p)}\hat{P}_i = \rho_i \sum_{(x,p)} \hat{P}_{(x,p)}\hat{P}_i = \rho_i\hat{P}_i, \tag{A19}$$

where we have used $\sum_{(x,p)} \hat{P}_{(x,p)} = 1$. Summing the above equation over i, and using $\sum_i \hat{P}_i = 1$, we obtain

$$\rho = \sum_i \rho_i\hat{P}_i, \tag{A20}$$

i.e., ρ can be decomposed using coarse-graining $\mathcal{C} = \{\hat{P}_i\}$. Defining sets $I^{(\lambda)} = \{i | \rho_i = \lambda\}$, we can rewrite this equation as

$$\rho = \sum_\lambda \lambda \sum_{i \in I^{(\lambda)}} \hat{P}_i, \tag{A21}$$

and since decomposition Eq. (A11) is unique, it must be that

$$\hat{P}_\lambda = \sum_{i \in I^{(\lambda)}} \hat{P}_i, \tag{A22}$$

which by definition means $\mathcal{C}_\rho \hookrightarrow \mathcal{C}$. For multiple coarse-graining, this then means that $\mathcal{C}_\rho \hookrightarrow \mathcal{C}_{1,...,n}$.

Conversely, we assume that Eq. (A22) holds. Points $(x,p), (\tilde{x}, \tilde{p})$ such that $\hat{P}_{(x,p)}\hat{P}_i \neq 0$, $\hat{P}_{(\tilde{x},\tilde{p})}\hat{P}_i \neq 0$ belong in the

same macrostate $\hat{P}_i$, and therefore by Eq. (A22) they must have the same associated eigenvalue,

$$\rho(\boldsymbol{x}, \boldsymbol{p}) = \rho_i = \rho(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}}), \tag{A23}$$

which means that Eq. (A18) holds, thus inequality in (A16) becomes equality.

Proof (Theorem 4: Observational entropy is nonincreasing with each added coarse-graining). Since joint coarse-graining $\mathcal{C}_{1,\dots,n}$ is finer than joint coarse-graining $\mathcal{C}_{1,\dots,n-1}$, from Theorem 2 we have

$$S_{O(\mathcal{C}_{1,\dots,n})}(\rho) \leqslant S_{O(\mathcal{C}_{1,\dots,n-1})}(\rho). \tag{A24}$$

Statement of the theorem then follows from Eq. (A7).

Proof (Theorem 5). Using $\hat{P}_E \hat{P}_{\boldsymbol{n}\tilde{E}} = \delta_{E,\tilde{E}} \hat{P}_{\boldsymbol{n}\tilde{E}}$, we have

$$
\begin{aligned}
p_{\boldsymbol{n}\tilde{E}}(\rho) &= \int_{(\boldsymbol{x},\boldsymbol{p})} \sum_E f(E) \hat{P}_E(\boldsymbol{x},\boldsymbol{p}) \hat{P}_{\boldsymbol{n}\tilde{E}}(\boldsymbol{x},\boldsymbol{p}) \mu \, d\boldsymbol{x} \, d\boldsymbol{p} \\
&= \int_{(\boldsymbol{x},\boldsymbol{p})} \left(\sum_E f(E) \delta_{E,\tilde{E}} \right) \hat{P}_{\boldsymbol{n}\tilde{E}}(\boldsymbol{x},\boldsymbol{p}) \mu \, d\boldsymbol{x} \, d\boldsymbol{p} \\
&= f(\tilde{E}) \int_{(\boldsymbol{x},\boldsymbol{p})} \hat{P}_{\boldsymbol{n}\tilde{E}}(\boldsymbol{x},\boldsymbol{p}) \mu \, d\boldsymbol{x} \, d\boldsymbol{p} \\
&= f(\tilde{E}) V_{\boldsymbol{n}\tilde{E}},
\end{aligned}
\tag{A25}
$$

from which

$$
\begin{aligned}
S_{xE}(t) &= -\sum_{\boldsymbol{n},E} f(E) V_{\boldsymbol{n}E} \ln \frac{f(E) V_{\boldsymbol{n}E}}{V_{\boldsymbol{n}E}} \\
&= -\sum_E \left(\sum_{\boldsymbol{n}} V_{\boldsymbol{n}E} \right) f(E) \ln f(E) \\
&= -\sum_E V_E f(E) \ln f(E).
\end{aligned}
\tag{A26}
$$

APPENDIX B: DETAILS FOR THE DYNAMICS OF S_{xE} FOR A MICROCANONICAL ENSEMBLE

Here we provide details for the calculation done in Sec. VIII B, showing that $S_{xE} = S_{O(\mathcal{C}_x, \mathcal{C}_{H_2})}$ for a microcanonical state of particles confined in the left part of the box is equal to the microcanonical entropy of this left part.

The initial state is a microcanonical state of N particles contained in the left part of the box. It can be written in two ways as

$$\rho_0 \equiv \rho_{E,\Delta E}^{(\text{micro};H_1)} = \frac{\hat{P}_{E,\Delta E}^{(H_1)}}{V_{E,\Delta E}^{(H_1)}} = \frac{\hat{P}_{(N,0)} \hat{P}_{E,\Delta E}^{(H_2)}}{V_{E,\Delta E}^{(H_1)}}. \tag{B1}$$

$\hat{P}_{E,\Delta E}^{(H_1)}$ and $\hat{P}_{E,\Delta E}^{(H_2)}$ denote projectors onto energy macrostates given by the first and the second Hamiltonian, and $\hat{P}_{(N,0)}$ is the projector onto a local particle number macrostate, corresponding to a statement that N particles are on the left side of the box, and zero particles are on the right.

We take $S_{xE} = S_{O(\mathcal{C}_x, \mathcal{C}_{H_2})}$ with local particle number coarse-graining that halves the box [$\boldsymbol{n} = (n_1, n_2)$, $L_1 = 0$, $\Delta x = \frac{L}{2}$], and that has energy macrostates given by the second Hamiltonian, H_2. Using

$$\hat{P}_{\boldsymbol{n}} \hat{P}_{(N,0)} = \delta_{n_1,N} \hat{P}_{(N,0)}, \tag{B2}$$

$$\hat{P}_{\tilde{E}}^{(H_2)} \hat{P}_{E,\Delta E}^{(H_2)} = \begin{cases} \hat{P}_{(N,0)} \hat{P}_{\tilde{E}}^{(H_2)}, & \tilde{E} \in [E, E+\Delta E) \\ 0 & \text{otherwise}, \end{cases} \tag{B3}$$

where $\delta_{n_1,N}$ denotes Kronecker's delta, we have

$$p_{\boldsymbol{n}\tilde{E}}(\rho_0) = \begin{cases} \dfrac{V_{\tilde{E}}^{(H_1)}}{V_{E,\Delta E}^{(H_1)}}, & n_1 = N \;\wedge\; \tilde{E} \in [E, E+\Delta E) \\ 0 & \text{otherwise} \end{cases} \tag{B4}$$

from which we obtain

$$S_{xE}(\rho_0) = \ln V_{E,\Delta E}^{(H_1)} \equiv S_{\text{micro}}^{(H_1)}(E). \tag{B5}$$

In other words, S_{xE} of the initial state is equal to the microcanonical entropy of the first half of the box.

APPENDIX C: DYNAMICS OF S_{xE} FOR A SINGLE POINT IN PHASE SPACE

Here we provide calculations on the long-time limit of S_{xE} in the case of an initial single point in phase space. We note that the derivation of S_{xE} of an initial state follows exactly the same pattern as in Appendix B, with the only difference being that we have dE instead of ΔE.

We now focus on calculating microcanonical entropy of the largest macrostate of phase space jointly coarse-grained by $\mathcal{C}_X$ and $\mathcal{C}_{H_2}$ (we will further denote $H \equiv H_2$).

Consider N identical classical particles in d dimensions evolving through a Hamiltonian H. Their positions and momenta are described by a point in phase space, $\Gamma \equiv (\boldsymbol{x}_1, \dots, \boldsymbol{x}_N, \boldsymbol{p}_1, \dots, \boldsymbol{p}_N)$. They are confined to a box of spatial volume $\mathcal{V} = L^d$, where L is a linear dimension. We subdivide the box into smaller ones of linear dimensions Δx, so that there are $m \equiv (\frac{L}{\Delta x})^d$ boxes. An arbitrary Γ will be in some coarse-grained region $P_{\boldsymbol{n}E}$, and this will have n_1 particles in the first box, n_2 in the second, and in general n_i in the ith box. We would like to know how many distinct ways, $\mathcal{N}$, there are of arranging the N particles having precisely these $\boldsymbol{n} = (n_1, \dots, n_m)$. The answer is the multinomial distribution

$$\mathcal{N}(n_1, \dots, n_m) = \frac{N!}{\prod_{i=1}^m n_i!}, \tag{C1}$$

where $\sum_{i=1}^m n_i = N$.

$\mathcal{N}$ is maximized when the n_i's are uniformly distributed, that is, for large $\frac{N}{m}$, $n_i = \frac{N}{m}$.

The spatial volume of a coarse-grained region is $\Delta x^{Nd} \mathcal{N}(\{n_1, \dots, n_m\})$. If we fine-grain in energy, we can write the (phase-space) volume of the ith region that Γ may be in,

$$V_{\boldsymbol{n}E} = \mu \int dp^{Nd} \int_{x \in P_{\boldsymbol{n}E}} dx^{Nd} \delta(E - H(\{\boldsymbol{x}_i\}_i, \{\boldsymbol{p}_i\}_i)), \tag{C2}$$

where $\hat{P}_E(\{\boldsymbol{x}_i\}_i, \{\boldsymbol{p}_i\}_i) = \delta(E - H(\{\boldsymbol{x}_i\}_i, \{\boldsymbol{p}_i\}_i))$ is the projector onto an energy shell. $P_{\boldsymbol{n}E}$ contains all permutations of particles. This means that if we consider a single Γ in this region, $P_{\boldsymbol{n}E}$ will also contain other Γ's with all of particle permutations consistent with the same particle numbers, $\boldsymbol{n} = (n_1, \dots, n_m)$, and the same energy E.

1. Case of dilute gas

Consider the case in which the interaction between the particles is small. We will also take the particles to be indistinguishable, meaning all phase-space volumes are divided by $N!$. In the limit of small interactions, the Hamiltonian only depends on the $\boldsymbol{p}$'s, and we can integrate over x and write

$$V_{\boldsymbol{n}E} = \frac{1}{N!} \frac{N!}{\prod_{i=1}^{m} n_i!} \Delta x^{Nd} \mu \int dp^{Nd} \delta(E - H(\{\boldsymbol{p}_i\}_i)). \quad (C3)$$

Defining the number of microstates of N indistinguishable particles in energy shell E occupying spatial volume $\mathcal{V}$ as

$$\Omega(E, \mathcal{V}, N) = \mathcal{V}^N \mu \int dp^{Nd} \frac{1}{N!} \delta(E - H(\{\boldsymbol{p}_i\}_i)), \quad (C4)$$

we can write

$$V_{\boldsymbol{n}E} = \frac{N!}{\prod_{i=1}^{m} n_i!} \Omega(E, \Delta x^d, N). \quad (C5)$$

$\Omega(E, \mathcal{V}, N)$ defines the microcanonical entropy,[20] $S_{\text{micro}}(E, \mathcal{V}, N) = \ln[\Omega(E, \mathcal{V}, N)]$ (taking $k_B = 1$). For small interactions considered here, we can approximate [Eq. (2.40) in [49]][21]

$$\Omega(E, \mathcal{V}, N) = \frac{1}{N!} \frac{\mathcal{V}^N (2\pi M E)^{\frac{Nd}{2}}}{\left(\frac{Nd}{2}\right)!}, \quad (C6)$$

where M is the mass of a gas particle.[22]

$\mathcal{N}$ is maximized with $n_i = \frac{N}{m}$. In this case, using Stirling's approximation,

$$\ln(\mathcal{N}_{\text{max}}) = \ln \frac{N!}{\left(\frac{N}{m}!\right)^m}$$

$$\approx N \ln(m) + \frac{1}{2} \ln(2\pi N) - \frac{m}{2} \ln\left(2\pi \frac{N}{m}\right). \quad (C7)$$

Combining Eqs. (C5), (C6), and (C7), and realizing that $\Delta x^d = \frac{\mathcal{V}}{m}$ (where $\mathcal{V}$ denotes the full spatial volume), we obtain S_{xE} for the point in phase space Γ that wandered into the largest macrostate,

$$S_{xE} = \ln\left(V_{\boldsymbol{n}E}^{(\text{max})}\right) = \ln(\mathcal{N}_{\text{max}}) + \ln \Omega(E, \Delta x^d, N)$$

$$= \ln(\mathcal{N}_{\text{max}}) + \ln\left(m^{-N} \Omega(E, \mathcal{V}, N)\right)$$

$$\approx N \ln(m) + \frac{1}{2} \ln(2\pi N) - \frac{m}{2} \ln\left(2\pi \frac{N}{m}\right)$$

$$- N \ln(m) + \ln \Omega(E, \mathcal{V}, N)$$

$$= S_{\text{micro}}(E, \mathcal{V}, N) + \frac{1}{2} \ln(2\pi N) - \frac{m}{2} \ln\left(2\pi \frac{N}{m}\right) \quad (C8)$$

If the point in phase space did not wander into the largest macrostate, but rather, let us say, the second largest (given by $n_1 = \frac{N+1}{m} - 1$, $n_i = \frac{N+1}{m}$ for $i > 1$), there would be minor modifications to the above formula, which would be negligible compared to the first term—the microcanonical entropy. In the end, it does not matter into which of the large macrostates the particle wanders—the leading term will always be the microcanonical entropy.

To understand to correction term in more detail, we can look at the entropy per particle. Then the relevant quantity to consider here is $\ln(\mathcal{N}_{\text{max}})/N$,

$$\frac{S_{xE}}{N} = \frac{S_{\text{micro}}(E, \mathcal{V}, N)}{N} + \frac{1}{2N} \ln(2\pi N) - \frac{m}{2N} \ln\left(2\pi \frac{N}{m}\right). \quad (C9)$$

We take a limiting process in which we fix the particle density and send N to infinity while keeping the particle density constant. For large N we can ignore the second term on the right-hand side. $N_m \equiv \frac{N}{m}$ is the average number of particles per box. In the limit where we fix the density average number of particles and increase the box length so that N_m becomes large, the last term on the right-hand side, $-(1/2N_m) \ln(2\pi N_m)$, becomes small, and it vanishes as $N_m \to \infty$. This correction term, therefore, represents a finite-size effect.

2. General classical systems

The case of a dilute gas can be extended to a general homogeneous classical system that is extensive. The system with maximum entropy will be the one where the number in each box is the same. This follows from the fact that all boxes must have the same chemical potential in equilibrium. The system will most likely find itself in this coarse-grained region. The logarithm of Eq. (C2) represents the entropy of a system that has been partitioned. As compared to the entropy of the entire system, each particle is confined to a box of width Δx. It is as if barriers had been added to prevent the exchange of particles between regions. However, energy can still be exchanged, therefore the temperature of each subsystem is the same. If the system has a density ρ and temperature T, corresponding to its total energy, then the entropy per particles $s(E, T)$ will not depend on system size in the thermodynamic limit. There are corrections to this due to finite-size effects, and they are particularly pronounced at critical points [51]. They have the general behavior as above, becoming negligible in the limit of large box size. The actual error will depend on the universality class of the system being studied.

---

[20]We remind the reader that the microcanonical entropy is defined by the right-hand side of Eq. (36), in which $V_E \equiv \Omega(E, \mathcal{V}, N) \equiv \rho(E)\Delta E$. Here we use $\Omega(E, \mathcal{V}, N)$ instead of the other two to spell out its dependence on the three variables.

[21]One can still include μ if needed, however, because due to the logarithm, this constant does not have much of a role apart from giving the correct units, and it appears below just as an additive constant. It disappears from the discussion entirely when studying only differences in entropies, which seems to be the only relevant physical quantity in physical scenarios being directly connected to heat. (The third thermodynamic law defining zero entropy at absolute zero appears to be a conceptual choice rather than having any physical relevance.)

[22]This expression leads to the famous Sackur-Tetrode equation [50], $S_{\text{micro}}(E, \mathcal{V}, N) = N\{\ln[\frac{\mathcal{V}}{N}(\frac{4\pi ME}{Nd})^{\frac{d}{2}}] + 1 + \frac{d}{2}\}$, but we will not use it here directly.

APPENDIX D: PROPERTIES OF S_F

First we rephrase the coarse-graining $\mathcal{C}_{XE} = \{P_{nE}\}_{n,E}$, Eq. (44), in mathematically precise terms. For simplicity, we set $L_1 = 0$ and $L_2 = L$.

The local particle number–local energy macrostates are then

$P_{nE} \equiv \{(\boldsymbol{x}, \boldsymbol{p}) | n_1 \text{ particles with position } x \in [0, \Delta x),$

$\ldots, n_m \text{ particles with position } x \in [(m-1)\Delta x, L),$

energy of the bin $[0, \Delta x)$ is between $[E_1, E_1 + dE)$,

$\ldots$, energy of the bin $[(m-1)\Delta x, L)$ is between

$[E_m, E_m + dE)\}.$ (D1)

To define this in mathematical terms, we need to define the notion of a *local Hamiltonian*. While the general theory that applies also to a phase-space density will be done elsewhere [43], here we can still treat the case of a single point of phase space. At every time t, there is some number of particles in each of the m bins. Let us say we have n_1 particles between $[0, \Delta x)$. For this particular time, the local Hamiltonian H_1 will be a function of phase-space points of n_1 particles,[23] which maps a point in phase space to the energy

$$H_1 : (x_1, \ldots, x_{n_1}, p_1, \ldots, p_{n_1}) \to E_1, \quad (D2)$$

where $x_i \in [0, \Delta x)$, $i = 1, \ldots, n_1$ (note that the lower index here is just a label denoting variables—we are not saying that exactly the first n_1 particles are in the first bin; in the case of indistinguishable particles, this label does not even matter). For instance, in the case of noninteracting particles, we have $H_1 = \sum_{i=1}^{n_1} \frac{p_i^2}{2M}$. If there is some interaction, the interaction between particles within the bin is accounted for in the local Hamiltonian, while the interaction between different bins is ignored.

We denote S_N as the set of permutations on the set $\{1, 2, \ldots, N\}$. With these tools, we can write

$P_{nE} \equiv \{(\boldsymbol{x}, \boldsymbol{p}) | (\exists \pi \in S_N) \text{ such that}$

$(\forall x_{\pi(i)} | i = 1, \ldots, n_1)(x_{\pi(i)} \in [0, \Delta x)), \ldots,$

$(\forall x_{\pi(i)} | i = N - n_m + 1, \ldots, N)(x_{\pi(i)} \in [(m-1)\Delta x, L)),$

$E_1 \leqslant H_1(x_{\pi(i)}, p_{\pi(i)}) < E_1 + dE, \ldots,$

$E_m \leqslant H_m(x_{\pi(i)}, p_{\pi(i)}) < E_m + dE\}.$ (D3)

Clearly, independently of the permutation, the phase-space volume will always be the same for a specific permutation π. The question, then, is how many permutations lead to a specific distribution of particles $\boldsymbol{n}$ between the bins. The answer is again

$$\mathcal{N}(n_1, \ldots, n_m) = \frac{N!}{\prod_{i=1}^m n_i!}, \quad (D4)$$

which is the total number of permutations divided by the possible number of permutations within the first bin, within

the second bin, and so on. We can therefore write the volume of macrostate P_{nE} as

$$V_{nE} = \mathcal{N}(n_1, \ldots, n_m) \frac{\mu}{N!} \Pi_{i=1}^{n_1} \int_{x_i \in [0, \Delta x)} dx_i \int_{p_i} dp_i$$

$$\times \cdots \times \Pi_{i=N-n_m+1}^N \int_{x_i \in [(m-1)\Delta x, L)} dx_i \int_{p_i} dp_i$$

$$\times \delta\big(E_1 - H_1(x_1, \ldots, x_{n_1}, p_1, \ldots, p_{n_1})\big) \times \cdots$$

$$\times \delta\big(E_m - H_m(x_{N-n_m+1}, \ldots, x_N, p_{N-n_m+1}, \ldots, p_N)\big)$$

$$= \Pi_{k=1}^m \Omega(E_k, \Delta x, n_k). \quad (D5)$$

Assuming generalization to spatial dimension d, we similarly derive $V_{nE} = \Pi_{k=1}^m \Omega(E_k, \Delta x^d, n_k)$. This means that for a single point in phase space, $\gamma \in P_{nE}$,

$$S_F = \ln V_{nE} = \sum_{k=1}^m S_{\text{micro}}(E_k, \Delta x^d, n_k), \quad (D6)$$

and therefore S_F is additive.

Now, let us study the long-time limit. After a long time, assuming identical particles of a dilute and weakly interacting gas, the largest macrostate is given by $E_k = \frac{E}{m}$, $n_k = \frac{N}{m}$ (where $m = \frac{\mathcal{V}}{\Delta x^d}$), giving

$$S_F = \ln\big(V_{n\epsilon}^{(\text{max})}\big) = m S_{\text{micro}}\left(\frac{E}{m}, \Delta x^d, \frac{N}{m}\right)$$

$$= m \ln \frac{1}{\frac{N}{m}!} \frac{\left(\frac{\mathcal{V}}{m}\right)^{\frac{N}{m}} \left(2\pi M \frac{E}{m}\right)^{\frac{Nd}{2m}}}{\left(\frac{Nd}{2m}\right)!}$$

$$= S_{\text{micro}}(E, \mathcal{V}, N) + \ln\left(\frac{m^{-N-\frac{Nd}{2}} N! \frac{Nd}{2}!}{\left(\frac{N}{m}!\right)^m \left(\frac{Nd}{2m}!\right)^m}\right)$$

$$= S_{\text{micro}}(E, \mathcal{V}, N) + \ln(2\pi N) - m \ln\left(2\pi \frac{N}{m}\right)$$

$$- \frac{m-1}{2} \ln \frac{d}{2}. \quad (D7)$$

Here we have used Eq. (C6) and Stirling's approximation for the factorials. The correction term $\frac{m-1}{2} \ln \frac{d}{2}$ represents a finite-size effect, which result in an order-1 correction in the thermodynamic limit.

APPENDIX E: ADDITIVITY OF QUANTUM S_{xE} AND NONADDITIVITY OF CLASSICAL S_{xE}

As explained in Sec. VIII D, the classical S_{xE} is not additive. This is because the classical global energy macrostate is defined as

$$P_E \equiv \{(\boldsymbol{x}, \boldsymbol{p}) | E \leqslant H(\boldsymbol{x}, \boldsymbol{p}) < E + \Delta E\}. \quad (E1)$$

This means that any microstate $(\boldsymbol{x}, \boldsymbol{p})$ that has energy E_1 in the left region and energy E_2 in the right region, $E_1 + E_2 = E$, will fall into the same macrostate P_E with any other states given by $\tilde{E}_1 + \tilde{E}_2 = E$.

Quantum mechanically, for infinite resolution in energy ($\Delta E = 0$), this is not true. This is because in a generic system, even states with a very similar sum of energies $E \approx E_1 +$

---

[23]For a different time, it can be a function of a different number of particles—this hints at the construction for the general case, which requires the construction of a Fock phase space [43].
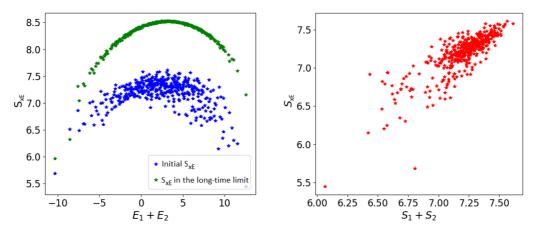
FIG. 6. Simulation showing additivity of quantum S_{xE} with infinite resolution in energy ($S_1 + S_2 = 0$). The system studied is of spinless fermions on a one-dimensional lattice evolved by a nonintegrable Hamiltonian used in [7,8]. The parameters are a lattice of $L = 16$ sites with hard-wall boundary conditions, with coarse-graining size $\Delta x = 4$ sites and the total number of particles $N = 6$. The state is initialized as a product state of random pure thermal states [7,8]), $|\psi\rangle = |\psi_{E_1}\rangle \otimes |\psi_{E_2}\rangle$, where $|\psi_{E_1}\rangle$ is a state of three particles occupying the eight sites on the left of the lattice with mean energy E_1, and $|\psi_{E_2}\rangle$ is a state of three particles occupying the eight sites on the right with mean energy E_2. The left figure represents a comparison of the S_{xE} of the initial state, $S_{xE}(|\psi\rangle)$, with the S_{xE} in the long-time limit, $S_{xE}(e^{-i\hat{H}t}|\psi\rangle)$ (with a large t). Classically, these two values would be the same and the curves would overlap, because both would correspond to $S_{xE} = \ln V_{(3,3)E}$ (where $E = E_1 + E_2$). As we can see, they do not overlap in a quantum-mechanical system when $\Delta E = 0$. The right figure shows the additivity of S_{xE} directly by comparing S_{xE} and the sum of thermodynamic entropies of the left eight sites (S_1) and of the right eight sites (S_2), for various product states $|\psi\rangle = |\psi_{E_1}\rangle \otimes |\psi_{E_2}\rangle$ such that $E = E_1 + E_2$.

$E_2 \approx \tilde{E}_1 + \tilde{E}_2$ are not exactly the same, so energy eigenstates of the global Hamiltonian $|E_1 + E_2\rangle$ and $|\tilde{E}_1 + \tilde{E}_2\rangle$ constitute their own macrostates, $\hat{P}_{E_1+E_2} = |E_1 + E_2\rangle\langle E_1 + E_2|$ and $\hat{P}_{\tilde{E}_1+\tilde{E}_2} = |\tilde{E}_1 + \tilde{E}_2\rangle\langle \tilde{E}_1 + \tilde{E}_2|$. (We ignored corrections coming from the interaction terms between the two regions that also contribute to the total Hamiltonian and therefore also to local energy eigenstates.)

If a finite resolution in energy is used ($\Delta E > 0$), then the situation for quantum S_{xE} is very similar to the classical case, because then the quantum macrostates are constituted by

$$\hat{P}_{E,\Delta E} \approx \sum_{E \leqslant E_1+E_2 \leqslant E+\Delta E} \hat{P}_{E_1+E_2}. \qquad (E2)$$

Put simply, classical energy macrostates correspond to quantum macrostates with $\Delta E > 0$, but for $\Delta E = 0$ the quantum macrostates are much smaller—actually consisting of a single microstate. This means that for $\Delta E = 0$, the global energy macrostates are actually small enough to distinguish between different local energies, which is why the quantum S_{xE} is additive. We have confirmed this numerically in some particular cases, as exhibited in Fig. 6.

Making the analytical argument for additivity precise would require similar arguments to those employed in the Appendixes of Ref. [8]. That is, we would need to identify the interaction part of the Hamiltonian with a random matrix, and then study the overlaps between energy eigenstates, the full Hamiltonian of the system, and those of the Hamiltonian system without the interaction, which gives the probability $p_E = |\langle E_1, E_2|E\rangle|^2$ that goes into S_{xE}. We leave this for future work.

[1] E. T. Jaynes, Am. J. Phys. **33**, 391 (1965).
[2] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[3] R. H. Swendsen, Am. J. Phys. **79**, 342 (2011).
[4] F. Balibrea, J. Mod. Phys. **7**, 219 (2016).
[5] S. Goldstein, D. A. Huse, J. L. Lebowitz, and P. Sartori, in *International Workshop on Stochastic Dynamics out of Equilibrium* (Springer, 2017) pp. 581596.
[6] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, arXiv:1903.11870 [cond-mat.stat-mech].
[7] D. Šafránek, J. M. Deutsch, and A. Aguirre, Phys. Rev. A **99**, 010101(R) (2019).
[8] D. Šafránek, J. M. Deutsch, and A. Aguirre, Phys. Rev. A **99**, 012103 (2019).
[9] L. Boltzmann, Entropy **17**, 1971 (2015).
[10] P. Español, M. Serrano, and I. Zuñiga, Int. J. Mod. Phys. C **8**, 899 (1997).
[11] T. Gao, M. D. Betterton, A.-S. Jhang, and M. J. Shelley, Phys. Rev. Fluids **2**, 093302 (2017).
[12] C. K. Batchelor and G. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge University Press, Cambridge, 1967).
[13] M. E. Fisher, Rev. Mod. Phys. **70**, 653 (1998).
[14] M. Kardar, *Statistical Physics of Particles* (Cambridge University Press, Cambridge, 2007).
[15] M. Kardar, *Statistical Physics of Fields* (Cambridge University Press, Cambridge, 2007).
[16] S.-K. Ma, *Modern Theory of Critical Phenomena* (Routledge, New York, 2018).
[17] J. M. Smith, *Introduction to Chemical Engineering Thermodynamics* (ACS Publications, New York, 1950).
[18] E. A. Guggenheim, *Statistical Thermodynamics; A Version of Statistical Mechanics for Students of Physics and Chemistry* (Cambridge University Press, Cambridge, 1956).

[19] H. B. Callen, *Thermodynamics and an Introduction to Thermo-statistics* (AAPT, 1998).

[20] K. G. Wilson, Phys. Rev. B **4**, 3174 (1971).

[21] J. Jost, *Dynamical Systems: Examples of Complex Behaviour* (Springer Science & Business Media, Berlin, 2006).

[22] J. D. Farmer, Z. Naturforsch., A **37**, 1304 (1982).

[23] V. Latora and M. Baranger, Phys. Rev. Lett. **82**, 520 (1999).

[24] R. Frigg, Br. J. Philos. Sci. **55**, 411 (2004).

[25] J. von Neumann, Eur. Phys. J. H **35**, 201 (2010).

[26] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955), pp. 410–416.

[27] C. S. Lent, Phys. Rev. E **100**, 012101 (2019).

[28] J. Gemmer and R. Steinigeweg, Phys. Rev. E **89**, 042113 (2014).

[29] P. Strasberg, arXiv:1906.09933 [cond-mat.stat-mech] .

[30] P. Strasberg and A. Winter, arXiv:2002.08817 [quant-ph] .

[31] D. Faiez, D. Šafránek, J. Deutsch, and A. Aguirre, Phys. Rev. A **101**, 052101 (2020).

[32] M. Nauenberg, Am. J. Phys. **72**, 313 (2004).

[33] V. V. Kozlov and D. Treshchev, Theor. Math. Phys. **151**, 539 (2007).

[34] G. Piftankin and D. Treschev, Chaos **18**, 023116 (2008).

[35] P. Županović and D. Kuić, J. Phys. Commun. **2**, 045002 (2018).

[36] L. Boltzmann, in *The Kinetic Theory of Gases: An Anthology of Classic Papers with Historical Commentary* (World Scientific, Amsterdam, 2003), pp. 262–349.

[37] S. G. Brush, *History of the Kinetic Theory of Gases* (Istituto della Enciclopedia, Italiana, 2004), p. 1.

[38] R. Balian, *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics* (Springer Science & Business Media, Berlin, 2007), Chap. 2.3.2.

[39] D. L. Goodstein, *States of Matter* (Courier, New York, 2014).

[40] R. Balian, *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics* (Springer Science & Business Media, Berlin, 2007), Vol. 2.

[41] A. Polkovnikov, Ann. Phys. **326**, 486 (2011).

[42] R. C. Tolman, *The Principles of Statistical Mechanics* (Courier, New York,1979).

[43] D. Šafránek, A. Aguirre, and J. M. Deutsch (unpublished).

[44] J. Wilkie and P. Brumer, Phys. Rev. A **55**, 27 (1997).

[45] J. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity* (Oxford University Press, Oxford, 2006), Vol. 14 .

[46] V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).

[47] I. V. Volovich, arXiv:0907.2445 [cond-mat.stat-mech].

[48] R. Balian, *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics* (Springer Science & Business Media, Berlin, 2007), Chap. 3.2.4.

[49] D. V. Schroeder, *An Introduction to Thermal Physics* (American Association of Physics Teachers, Washington D.C., 1999).

[50] S. K. Ma, *Statistical Mechanics* (World Scientific, Singapore, 1985).

[51] V. Privman and M. E. Fisher, Phys. Rev. B **30**, 322 (1984).