Inferring entropy from structure

Gil Ariel

Department of Mathematics, Bar-Ilan University, 52000 Ramat Gan, Israel

Haim Diamant

Raymond and Beverly Sackler School of Chemistry, Tel Aviv University, 69978 Tel Aviv, Israel

The thermodynamic definition of entropy can be extended to nonequilibrium systems based on its relation to information. To apply this definition in practice requires access to the physical system's microstates, which may be prohibitively inefficient to sample or difficult to obtain experimentally. It is beneficial, therefore, to relate the entropy to other integrated properties which are accessible out of equilibrium. We focus on the structure factor, which describes the spatial correlations of density fluctuations and can be directly measured by scattering. The information gained by a given structure factor regarding an otherwise unknown system provides an upper bound for the system's entropy. We find that the maximum-entropy model corresponds to an equilibrium system with an effective pair interaction. Approximate closed-form relations for the effective pair potential and the resulting entropy in terms of the structure factor are obtained. As examples, the relations are used to estimate the entropy of an exactly solvable model and two simulated systems out of equilibrium. The focus is on low-dimensional examples, where our method, as well as a recently proposed compression-based one, can be tested against a rigorous direct-sampling technique. The entropy inferred from the structure factor is found to be consistent with the other methods, superior for larger system sizes, and accurate in identifying global transitions. Our approach allows for extensions of the theory to more complex systems and to higher-order correlations.

I. INTRODUCTION

After decades of slow progress, the theory of nonequilibrium thermodynamics has seen substantial advances in recent years. One outstanding example is the derivation of rigorous statistical relations between macroscopic thermodynamic variables out of equilibrium [1,2]. Another is the concise theoretical treatments of active matter [3]. The definition of thermodynamic state variables out of equilibrium remains a key issue. Consistent definitions of temperature [4,5] and pressure [6] have proved problematic and context-dependent. By contrast, the definition of entropy can readily be extended to nonequilibrium systems based on its relation to information [7,8]. Given the system's set of microstates $\{s\}$ and their probability distribution p_s, the physical Gibbs entropy is identified with the Shannon entropy,

$$H \equiv - \sum_s p_s \ln p_s, \tag{1a}$$

or its differential (continuous) analog,

$$H \equiv - \int p(s) \ln p(s) \, ds, \tag{1b}$$

where $p(s)$ is the probability density and the integration is over its support.

As H relates to (minus) the information encoded in p_s, it should hold regardless of whether p_s is an equilibrium distribution or not [9]. Equations (1) are also used in a range of fields and disciplines other than physics, including signal and image processing [10], machine learning [11,12], evaluation of independence [13], and independent component analysis [14].

Beyond the definition of entropy, we need ways to compute or measure it. In the case of physical systems this is useful, for example, for identifying changes in the strength of correlations such as the ones underlying global transitions. One may try to estimate p_s sufficiently well to obtain H directly from Eqs. (1). However, this essentially amounts to replacing a difficult problem with a more difficult one [15]. Moreover, sampling p_s also requires access to the physical system's microstates, which may be hard or impossible to get experimentally.

Despite a large number of suggested algorithms, estimating the entropy from independent sampling of the distributions remains a challenge, especially for physical applications which typically involve high-dimensional distributions in bounded domains [15–17]. In such cases, none of the commonly used direct estimation methods perform well [15,17,18]. It is beneficial, therefore, to relate the entropy to other integrated properties which are more directly accessible. At equilibrium such a relation exists between the entropy (rather, its temperature derivative) and the fluctuations of the total energy [19]. This relation assumes that p_s is the Boltzmann distribution, which is invalid out of equilibrium.

Density fluctuations are an integrated material property which is directly accessible, both in and out of equilibrium, e.g., via scattering. Deviations from the density fluctuations of a completely random particle distribution (an ideal gas)

encode spatial correlations and thus are related to the entropy associated with the spatial distribution, the structural entropy. Our aim is to establish this relation. For concreteness we restrict the discussion to two-point density correlations in a system of identical particles at steady state. The same approach, however, can be used to extend the theory to more complex systems and to higher-order correlations. We will discuss such extensions in Sec. V. Denoting the steady-state density of particles at position $\mathbf{r}$ by $\rho(\mathbf{r})$, the two-point correlation is $\langle \rho(\mathbf{r})\rho(\mathbf{r}')\rangle - \bar{\rho}^2$, where $\langle \cdot \rangle$ denotes an average over p_s and $\bar{\rho} = \langle \rho(\mathbf{r})\rangle$ is the mean density. Applying a Fourier transform, we obtain the structure factor,

$$S(\mathbf{q}) = N^{-1}\langle |\tilde{\rho}(\mathbf{q})|^2 \rangle, \tag{2}$$

where $\tilde{\rho}(\mathbf{q}) = \int d\mathbf{r} e^{-i\mathbf{q}\cdot\mathbf{r}} \rho(\mathbf{r})$, and N the number of particles.

Knowledge of the structure factor constrains the distribution of microstates and thus entails an entropy cost. The relation between entropy and structure factor is not unique; there may be many systems having the same $S(\mathbf{q})$ but different entropy. Still, we may ask the sharply defined question, how much information is gained by getting $S(\mathbf{q})$ of an otherwise unknown system. Out of those many systems sharing the same $S(\mathbf{q})$, the answer is given by the one which has the maximum entropy. Hence, for a given $S(\mathbf{q})$, we can get a bound for the system's entropy if we find the model which has the largest entropy out of all possible N-body joint distributions compatible with $S(\mathbf{q})$. In Sec. II A we show that this maximum-entropy model corresponds to an equilibrium system with an effective pair interaction dependent on the given $S(\mathbf{q})$. Finding such a pair potential seems to be possible for any physically realizable structure factor [20]. Additionally, a given equilibrium structure factor can be associated with only one pair interaction [21]. Hence, the effective pair potential corresponding to the maximum-entropy system should exist and is unique. The problem is thus reduced to (a) finding the effective pair potential and (b) calculating the resulting *equilibrium* entropy.

For step (b) we have the well-established arsenal of equilibrium statistical mechanics. It is the first step, of determining the effective pair potential, which poses the extra challenge. In the present work we prefer to remain on the level of analytic closed-form expressions. Section II B provides an approximate relation for the entropy, which is our central result,

$$h[S(\mathbf{q})] = \frac{1}{2(2\pi)^d \bar{\rho}} \int d\mathbf{q}(\ln S - S + 1). \tag{3}$$

Here $h \equiv (H - H_{\mathrm{id}})/N$ is the excess entropy per particle over the ideal-gas entropy, and d the spatial dimensionality.

Some of the points that we present appeared, from different perspectives, in earlier studies. Other relations between the entropy and spatial pair correlations were derived in the past based on equilibrium liquid theory [22–26]. We will compare Eq. (3) with those relations in subsequent sections. The study in Ref. [20] addressed the problem of finding a system that realizes a given structure factor. The results led to the conjecture that any realizable $S(\mathbf{q})$ can be obtained from an equilibrium system with an appropriate pair interaction, and a numerical scheme was developed to get the pair potential. Reference [27] studied universal features of correlation functions in the

limit of weak coupling and how they could be used to extract microscopic pair potentials from measurements. We will point more specifically to the correspondence between these earlier works and ours whenever relevant.

Recent works have proposed to overcome the sampling problem by measuring the entropy of physical systems from their information content, estimated through lossless data compression [28–31]. We will discuss this approach in detail later on. Information content was used also to identify elusive structures in amorphous materials using network analysis [32].

Entropy estimation based on the structure factor has several advantages over the other available methods, in particular, when applied to continuous physical systems. (a) Unlike sampling or data compression, it does not involve any extrinsic discretization. Such discretization is found to affect the results even when it is performed in a generic parameter-free way [33]. By contrast, the Fourier modes $\mathbf{q}$ are naturally discretized by the system size, which also allows for easy treatment of finite-size effects. In addition, unlike the alternative relations mentioned above, which are based on real-space pair correlations, Eq. (3) does not require binning of interparticle distances. We find weak dependence of Eq. (3) on the upper cutoff of $\mathbf{q}$. (b) Compression algorithms, while being very simple to implement, depend on additional extrinsic factors, such as transforming the data into a linear string, and scaling of compression output to entropy [33]. The former issue may hamper the identification of long-range correlations, whereas the structure factor encodes correlations of any range (albeit on the pairwise level only). See Ref. [33] for a recent investigation into these issues. Importantly, that study indicated that the compression method might miss qualitative entropy changes of certain driven systems which density correlations successfully picked up. (c) As to direct statistical estimation of entropy, the available algorithms are inaccurate and prohibitively inefficient even at moderate system sizes (100 particles and above). The structure-factor-based estimation is not limited by system size.

A key disadvantage of entropy estimation based on density correlations is that it is more particular. While sampling, and especially compression, can be applied "blindly" without prior knowledge of the system, relations such as Eq. (3) are limited to systems of a certain category. For example, Eq. (3) must be modified if it is to be applied to mixtures or to anisotropic particles. In addition, only approximate relations are available. Yet, as demonstrated below, Eq. (3) is found to correctly indicate qualitative features of strongly correlated as well as very small systems.

We will compare the entropy inferred from the structure factor with the results of two recently presented specific methods—one based on direct sampling [18] and the other on data compression [28]. These methods will be briefly reviewed in Sec. III. The direct-sampling method is particularly important for our purposes as it is guaranteed to yield accurate results for small systems. This will allow us to test the structure factor and compression methods against a rigorous benchmark out of equilibrium.

In Sec. IV we apply all three methods to specific systems. We have selected three examples. The first, treated in Sec. IV A, is a simple equilibrium model of an ideal gas

partitioned into single-particle cells. In this system, whose entropy is known exactly, density fluctuations are strongly suppressed due to *single-body* potentials. Thus, it subjects our effective pair-interaction approach to a stringent test. The second example (Sec. IV B) is a one-dimensional (1D) model inspired by periodically sheared suspensions [34]. This model undergoes a transition to an absorbing state involving a sharp change of entropy and strong density correlations. The third example, described in Sec. IV C, is a two-dimensional (2D)

model of random organization [35,36]. With a nonequilibrium transition involving a strongly correlated state, it allows us to test the methods in two dimensions. We focus on small system sizes (up to 100 particles), for which statistical estimation of entropy is feasible. Remarkably, the entropy reveals sharp transitions between absorbing and nonabsorbing steady-state dynamics even for systems containing as few as 10 particles. Finally, we summarize our conclusions and discuss further applications and extensions in Sec. V.

II. ENTROPY FROM STRUCTURE FACTOR

A. Maximum entropy model

In this section, we describe the maximum-entropy model for a given structure factor, i.e., out of all models that have a given $S(\mathbf{q})$, we find the one that has the largest possible entropy. A similar approach was used in Ref. [20] to find a system that realizes a given $S(\mathbf{q})$. Here we go further to derive an expression for the maximum entropy.

We consider N identical featureless particles in a volume V. To simplify notation, we assume a countable number of possible spatial configurations $\mathbf{R}^s = (\mathbf{R}_1^s, \ldots, \mathbf{R}_N^s) \in \mathbb{R}^{Nd}$. We denote the probability of configuration $\mathbf{R}^s$ by p_s. The density field corresponding to configuration $\mathbf{R}^s$ is $\rho^s(\mathbf{r}) = \sum_{n=1}^N \delta(\mathbf{r} - \mathbf{R}_n^s)$. Substitution in Eq. (2) gives the structure factor as

$$S(\mathbf{q}) = N^{-1} \sum_s p_s \sum_{m=1}^N \sum_{n=1}^N e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)}. \tag{4}$$

Different sets of p_s may have the same structure factor. Given a function $S(\mathbf{q})$, we are looking for the specific model which has the structure factor $S(\mathbf{q})$ and a maximum value of $H = -\sum_s p_s \ln p_s$. Note that $S(0) = N$ for any system; thus the mode $\mathbf{q} = 0$ is excluded from the constrained $S(\mathbf{q})$.

To find the maximum-entropy model, we rephrase the problem in terms of Lagrange multipliers. We look for a maximum of H under the infinite set of constraints (4) for all $\mathbf{q} \neq 0$, and the normalization condition $\sum_s p_s = 1$. This general strategy is not new; see, for example, Ref. [37]. For simplicity of notation, we use also a countable number of wave vectors $\mathbf{q}$. We define the Lagrangian,

$$\mathcal{L} = H - \sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \left[\sum_{m,n} \sum_s p_s e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - \sum_s p_s N S(\mathbf{q}) \right] - \mu \left[\sum_s p_s - 1 \right], \tag{5}$$

where $\lambda(\mathbf{q})$ are the Lagrange multipliers associated with each of the constraints (4), and μ is the Lagrange multiplier associated with the normalization. Then, the maximum-entropy model satisfies

$$\frac{\partial \mathcal{L}}{\partial p_s} = 0; \quad \frac{\partial \mathcal{L}}{\partial \lambda(\mathbf{q})} = 0; \quad \frac{\partial \mathcal{L}}{\partial \mu} = 0. \tag{6}$$

Differentiation yields

$$-\ln p_s - 1 - \sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - N S(\mathbf{q}) \right] - \mu = 0, \tag{7a}$$

$$\sum_{m,n} \sum_s p_s e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} = N S(\mathbf{q}), \tag{7b}$$

$$\sum_s p_s = 1. \tag{7c}$$

Solving Eq. (7a) for p_s, we have

$$p_s = e^{-(\mu+1)} \exp\left\{ -\sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - N S(\mathbf{q}) \right] \right\} \equiv e^{-(\mu+1)} r_s, \tag{8}$$

where r_s is the exponential term. Define

$$\mathcal{Z} = \sum_s r_s = \sum_s \exp\left\{ -\sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - N S(\mathbf{q}) \right] \right\}. \tag{9}$$

From Eq. (7c), $\mathcal{Z} = e^{\mu+1}$; hence,

$$p_s = r_s/\mathcal{Z}. \tag{10}$$

In addition,

$$\frac{\partial \ln \mathcal{Z}}{\partial \lambda(\mathbf{q})} = \mathcal{Z}^{-1} \frac{\partial \mathcal{Z}}{\partial \lambda(\mathbf{q})} = -\mathcal{Z}^{-1} \sum_s r_s \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - NS(\mathbf{q}) \right]$$

$$= -\sum_s p_s \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - NS(\mathbf{q}) \right] = -\sum_{m,n} \sum_s p_s e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} + NS(\mathbf{q}) = 0, \tag{11}$$

where the last equality is due to Eqs. (7b) and (7c). Thus, solving $\partial \ln \mathcal{Z}/\partial \lambda(\mathbf{q}) = 0$ is equivalent to enforcing the constraint on the structure factor.

Finally, we compute the entropy associated with the distribution p_s,

$$H = -\sum_s p_s \ln p_s = -\sum_s p_s \ln \frac{r_s}{\mathcal{Z}} = -\sum_s p_s \ln r_s + \ln \mathcal{Z}$$

$$= -\sum_s p_s \ln r_s + \ln \mathcal{Z}. \tag{12}$$

The first term satisfies

$$\sum_s p_s \ln r_s = \sum_s p_s \left\{ -\sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - NS(\mathbf{q}) \right] \right\}$$

$$= -\sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \sum_s p_s \left[\sum_{m,n} e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - NS(\mathbf{q}) \right]$$

$$= -\sum_{\mathbf{q}\neq 0} \lambda(\mathbf{q}) \left[\sum_{m,n} \sum_s p_s e^{i\mathbf{q}\cdot(\mathbf{R}_m^s - \mathbf{R}_n^s)} - NS(\mathbf{q}) \right]. \tag{13}$$

The square brackets vanish for all $\mathbf{q}$. We conclude that

$$H = \ln \mathcal{Z}. \tag{14}$$

In summary, we see that the maximum-entropy model has the equilibrium distribution of a system with an effective pair potential $\lambda(\mathbf{q})$. The probability of each configuration (8) is Boltzmann-like, and the entropy is simply the log of the effective partition function $\mathcal{Z}$.

B. Effective pair interaction and entropy

We now use the results of Sec. II A to obtain the effective pair interaction and the resulting entropy estimation for a given $S(\mathbf{q})$. We will outline the calculation and give its results while leaving the technical details to Appendix A. Since we are after the nominal entropy, we keep track of all prefactors.

Equation (9) is rewritten in continuous form as

$$\mathcal{Z} = \int \left(\prod_{n=1}^N d\mathbf{R}_n \right) \exp \left\{ -\int d\mathbf{q} \lambda(\mathbf{q}) \left[\sum_{m=1}^N \sum_{n=1}^N e^{i\mathbf{q}\cdot(\mathbf{R}_m - \mathbf{R}_n)} - NS(\mathbf{q}) \right] \right\}, \tag{15}$$

where we have absorbed the constant density of Fourier modes, $\nu \equiv V/(2\pi)^d$, into the definition of the Lagrange multipliers. In Appendix A we use standard methods to transform Eq. (15) into the following functional integral over a continuous field $\tilde{\psi}$:

$$\mathcal{Z} = Z_{\text{id}}(4\pi\nu^2)^{-\Omega_q/2} \exp \left\{ \int d\mathbf{q} \left[N\lambda S - \frac{\nu}{2} \ln(\lambda/\nu) \right] \right\} \int \mathcal{D}\tilde{\psi} e^{-\mathcal{H}_\psi},$$

$$\mathcal{H}_\psi = -\bar{\rho} \int d\mathbf{R} (e^{-i\int d\mathbf{q}\tilde{\psi}e^{-i\mathbf{q}\cdot\mathbf{R}}} - 1) + \int d\mathbf{q} \frac{|\tilde{\psi}|^2}{4\lambda}, \tag{16}$$

where Z_{id} is the ideal-gas partition function, and $\Omega_q = \nu \int d\mathbf{q}$ is the total number of Fourier modes.

Equation (16) is a useful starting point for finding the effective interaction and the entropy to the desired order of approximation. Here we confine ourselves to the lowest nontrivial order which yields the simple relation, Eq. (3). This is obtained

by expanding $\mathcal{H}_\psi$ to quadratic order in $\tilde{\psi}$ and performing the Gaussian integration. The result is

$$\mathcal{Z} = Z_{\text{id}} \exp\left\{ \int d\mathbf{q}\left(N\lambda S - \frac{\nu}{2}\ln\left[1 + \lambda/\alpha\right]\right)\right\}, \tag{17}$$

where $\alpha \equiv \nu/(2N) = [2(2\pi)^d \bar{\rho}]^{-1}$.

Applying the condition from Sec. II A, $\delta \ln \mathcal{Z}/\delta\lambda(\mathbf{q}) = 0$, gives the effective pair interaction,

$$\lambda(\mathbf{q}) = \alpha\left[\frac{1}{S(\mathbf{q})} - 1\right]. \tag{18}$$

This relation between structure factor and pair potential coincides with earlier results obtained in different contexts [20,27]. It is proportional also to (minus) the direct correlation function known from equilibrium liquids via the Ornstein-Zernike relation [38]. Indeed, in the high-temperature (i.e., weak-coupling) limit, the equilibrium direct correlation function is proportional to (minus) the pair potential [38]. The effective potential (18) vanishes, as expected, for an ideal gas ($S \equiv 1$). Substituting λ back in $H = \ln \mathcal{Z}$ leads to Eq. (3). Inspection of Eq. (3) readily confirms that the excess entropy h is nonpositive for any $S(\mathbf{q})$ and reaches its largest value of 0 for $S(\mathbf{q}) \equiv 1$, i.e., for an ideal gas.

C. Comparison with previous expressions

We now compare Eq. (3) with previous results from equilibrium liquid theory. At equilibrium, the excess entropy is commonly extracted from simulations using the thermodynamic identity [33,39],

$$h = \frac{u}{T} - \int_0^{\bar{\rho}} d\bar{\rho}' \frac{p}{T\bar{\rho}'^2}, \tag{19}$$

where T is the temperature (taking the Boltzmann constant equal to 1), u is the excess internal energy per particle over that of an ideal gas, and p the excess pressure over the ideal-gas pressure. More relevant to the present work, however, are relations between excess entropy and pair correlations.

The first relation between the entropy and pair correlations was derived by Green [22,23],

$$h^{\text{G}} = -\frac{\bar{\rho}}{2}\int d\mathbf{r}[g\ln g - (g-1)], \tag{20}$$

where $g(\mathbf{r})$ is the pair distribution function, related to the structure factor as

$$S(\mathbf{q}) = 1 + \bar{\rho}\int d\mathbf{r}e^{-i\mathbf{q}\cdot\mathbf{r}}[g(\mathbf{r}) - 1]. \tag{21}$$

A higher-order relation was presented by Hernando [25,26,40],

$$h^{\text{H}} = h^{\text{G}} + \alpha\int d\mathbf{q}\left[\ln S - S + 1 + \frac{1}{2}(S-1)^2\right]. \tag{22}$$

To compare Eq. (3) with these expressions we expand them around the ideal gas ($g = S = 1$). For Green's functional,

$$h^{\text{G}} \simeq -\frac{\bar{\rho}}{4}\int d\mathbf{r}(g-1)^2 + O(g-1)^3 = -\frac{1}{2}\alpha\int d\mathbf{q}(S-1)^2 + O(g-1)^3. \tag{23}$$

On the other hand, $\ln S - S + 1 = -\frac{1}{2}(S-1)^2 + O(S-1)^3$. We conclude that the three expressions agree to second order in the correlations. In Secs. IV A and IV C they will be further compared for specific examples.

III. ENTROPY ESTIMATIONS OUT OF EQUILIBRIUM

Despite a large number of suggested algorithms [15,17], the problem of estimating the entropy, in particular differential entropy, from independent sampling of the distributions remains a challenge in high dimensions [16,17]. Broadly speaking, traditional estimation methods can be classified into one of two approaches—binning and sample-spacing methods—or their multidimensional analogues—partitioning [41] and nearest-neighbor methods [42,43]. Traditional partitioning methods work well at low dimensions (typically two or three) and only if the support is known. In contrast, nearest-neighbor schemes perform well up to moderately high dimensions (typically 10–15) and for distributions with

unbounded support. However, they may fail completely when the density has a compact support and their time complexity grows exponentially with the dimension.

Overall, physical applications involving high-dimensional distributions in bounded domains pose a difficult challenge. Below, we review two recently proposed methods for estimating differential Shannon entropy of compact high-dimensional continuous distributions.

A. Direct statistical estimation

In this paper, we apply a recently suggested method that is based on decomposing the distribution into a product of the marginal distributions and the joint dependency, also known

as the *copula* [18]. The entropy of the marginals is estimated using 1D methods, while the entropy of the copula, which always has a compact support, is estimated recursively by splitting the data along statistically dependent dimensions. Below we present the method's principle, while the technical details of its application are given in Appendix B. Further information is found in Ref. [18].

Sklar's theorem [44,45] states that any continuous multi-dimensional probability density $p(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^D$, can be written uniquely as

$$p(\mathbf{x}) = p_1(x_1) \cdots p_D(x_D) c(F_1(x_1), \ldots, F_D(x_D)). \quad (24)$$

Here $\mathbf{x} = (x_1, \ldots, x_D)$, and $p_k(\cdot)$ denotes the marginal probability density of the kth dimension, with cumulative distribution function (CDF) $F_k(t) = \int_{-\infty}^{t} p_k(x) \, dx$. The density of the copula, $c(u_1, \ldots, u_D)$, is a probability density on the hypersquare $[0, 1]^D$, whose marginals are all uniform on [0,1],

$$\forall k, \; \left[\Pi_{j=1, \; j \neq k}^{D} \int du_j \right] c(u_1, \ldots, u_D) = 1. \quad (25)$$

Substituting Eq. (24) into the definition of entropy, Eq. (1), yields

$$H = \sum_{k=1}^{D} H_k + H_c, \quad (26)$$

where H_k is the entropy of the kth marginal, to be computed using appropriate 1D estimators, and H_c is the entropy of the copula, to be computed recursively by constructing a partition tree.

Splitting the overall estimation into marginal and copula contributions has several advantages. First, the support of the copula is compact, which is exactly the premise for which partitioning methods are most adequate. Second, since the entropy of the copula is nonpositive, adding up the entropy of marginals H_k across partition-tree levels provides an improving approximation (from above) of the entropy.

B. Compression-based estimation

Following the ideas of Refs. [28,29], we estimate the entropy using a lossless compression algorithm. Similar to Ref. [28], samples are binned into 256 equal bins in each dimension, and the data are converted into a matrix of 8-bit unsigned integers, in which each row constitutes an independent sample. The matrix is compressed using the LZW algorithm (implemented in Matlab's imwrite function) into a graphic interchange format (gif) file. In order to estimate the entropy, the file size is interpolated linearly between the file size of a compressed constant matrix (minimum entropy) and a random matrix with independent uniformly distributed values (maximum entropy), both of the same dimensions.

The algorithm is extremely simple, which is its main appeal. See Appendix B for implementation details.

IV. EXAMPLES

In this section we present the results of the three entropy estimation methods, based on the structure factor, direct sampling, and compression, for three simple models. The first is a 1D model of a partitioned gas, in which the particles are restricted to nonoverlapping identical segments. The second is a 1D model inspired by periodically sheared particle suspensions [34]. The third is a 2D random organization (RandOrg) model [35,36]. The selected models all have states of strongly suppressed density fluctuations (hyperuniformity [46]) to put our density-correlation approach to test [47]. The partitioned gas is an exactly solvable equilibrium model, while the other two systems are out of equilibrium, checking the ability of the methods to identify the entropy change due to a global dynamic transition [33].

A. Partitioned gas model

We begin with a very simple model which is particularly suitable for our purposes. It contains only single-particle constraints and thus is exactly solvable. The exactly calculated entropy and structure factor are far from those of an ideal gas due to strongly suppressed density fluctuations (hyperuniformity). We compare our estimation methods against these exact results. Since the entropy estimation based on structure factor uses effective *pair* interactions, this test is particularly stringent.

Consider N point particles on a line of length L. The particles are partitioned into N cells of length $a = L/N$ each, one particle per cell; see Fig. 1(a). Generalization to several particles per cell is straightforward. We set $a \equiv 1$, such that $L = N$, and the mean density is $\bar{\rho} = 1$. The entropy per particle is

$$H_{\text{PG}}/N = \ln a = 0, \quad (27)$$

compared to the ideal entropy per particle, $\ln a + 1 = 1$. Thus $h_{\text{PG}} = -1$.

To calculate the structure factor, we pick a random particle and place the origin $x = 0$ at the center of its cell. The particle's position, $x_0 \in (-1/2, 1/2)$, divides the cell into two segments. Denoting, $\ell_{\min} = \min(x_0 + 1/2, 1/2 - x_0)$ and $\ell_{\max} = \max(x_0 + 1/2, 1/2 - x_0)$, the probability per unit length to find a second particle at the position x is 0 if $|x - x_0| < \ell_{\min}$, $\bar{\rho}$ if $|x - x_0| > \ell_{\max}$, and $\bar{\rho}/2$ if $|x - x_0| \in (\ell_{\min}, \ell_{\max})$. Averaging over x_0, we get the pair distribution function,

$$g_{\text{PG}}(x) = \begin{cases} |x|, & |x| \leqslant 1 \\ 1, & |x| > 1 \end{cases}. \quad (28)$$

The structure factor is related to $g(x)$ as $S(q) = 1 + \bar{\rho} \int_{-\infty}^{\infty} dx e^{-iqx} [g(x) - 1]$, yielding

$$S_{\text{PG}}(q) = 1 - \frac{4 \sin^2(q/2)}{q^2}. \quad (29)$$

In the limit $q \ll 1$ we have $S_{\text{PG}} \simeq q^2/12$. The decay to zero for vanishing q reflects hyperuniformity [46], i.e., suppressed density fluctuations over distances containing many cells.

Substituting Eq. (29) in Eq. (18), we get the (approximate) effective pair interaction,

$$\lambda_{\text{PG}}(q) = \frac{1}{\pi q^2} \frac{\sin^2(q/2)}{1 - 4 \sin^2(q/2)/q^2}. \quad (30)$$

This is the fictitious Fourier-space pair potential which mimics the single-particle constraints. In the limit $q \ll 1$ it becomes
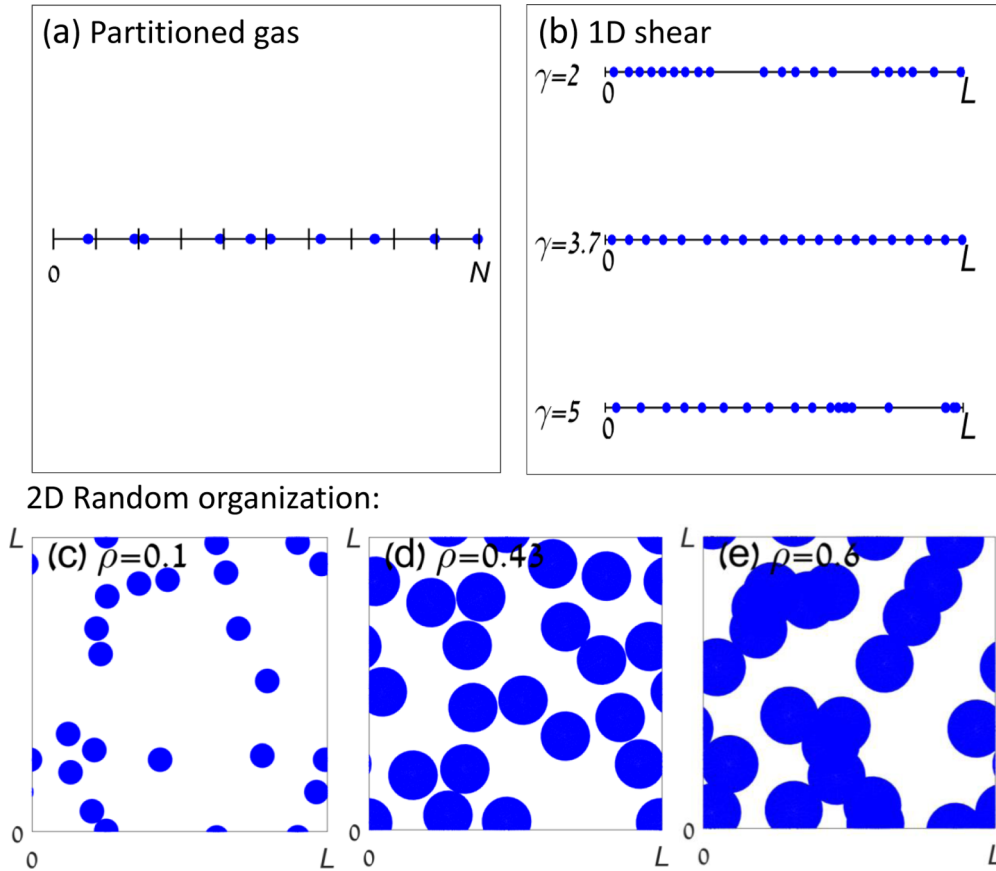
FIG. 1. Snapshots from simulations. (a) Partitioned gas model with $N = 10$ particles. (b) The 1D shear model with $N = 20$, below, at, and above the critical shear $\gamma_C = 3.7$. (c–e) The 2D RandOrg model with $N = 20$ below, at, and above the critical density $\rho_C = 0.43$.

$\lambda_{PG}(q) \simeq 3/(\pi q^2)$. Thus, over large distances, the particles are distributed as in a 1D Coulomb gas.

Substituting $S_{PG}(q)$ of Eq. (29) in Eq. (3) gives the entropy estimation $H/N = 0.22$, compared to the exact value of 0 (or $h = -0.78$ compared to -1). Substitution of $g_{PG}(x)$ and $S_{PG}(q)$ in the alternative relations due to Green [Eq. (20)] and Hernando [Eq. (22)], gives $H/N = 0.75$ and 0.14, respectively. These theoretical results correspond to an infinite system, $N \to \infty$.

We simulated the partitioned-gas model numerically for different numbers of particles, and applied the three entropy-estimation methods to the resulting configurations. Table I presents the results for $N = 10, 20, 50$, and 100 particles using 10^6 independent samples.

As anticipated, the copula method performs well, particularly at the lowest dimensions ($N = 10$). Theoretically, both the copula and compression methods must converge to the correct value for an infinite number of samples (rightmost column of the table). In practice, however, their accuracy is seen to deteriorate for larger systems due to insufficient sampling of the enlarged phase space.

By contrast, the methods based on the structure factor and pair-distribution function properly approach their individual infinite-system limits given above. Larger system size implies finer q discretization, leading to better approximation of the structure factor from the simulations. Note, on the other hand, that for the partitioned-gas model these methods cannot give

exact results, because the model does not have actual pair interactions. For this model we find that our relation performs much better than Green's and slightly worse than Hernando's. We recall that it has the advantage of using $S(\mathbf{q})$ alone, without the necessity of $g(\mathbf{r})$, as will be demonstrated further in Sec. IV C.

As proved in Sec. II A, if the actual model is unknown and only its $S(q)$ is given, then assuming an effective pair potential

TABLE I. Estimations of entropy per particle for the partitioned-gas model with N particles and K independent samples. The exact value for all systems is $H_{PG}/N = 0$ (compared to $H_{id}/N = 1$ for an ideal gas). The rightmost column (lower three rows) shows the analytical estimations, valid for $N \to \infty$, using our relation [Eq. (3)] and two alternative ones [Eqs. (20) and (22)]. In the limit of infinite sampling the copula and compression methods should become exact (rightmost column, upper two rows). The rest of the entries contain results from numerical simulations.

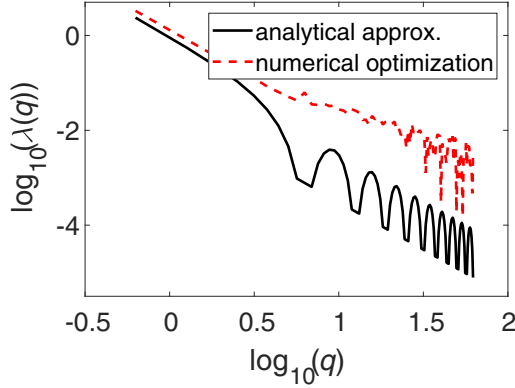| Method	$N = 10$ $K = 10^6$	$N = 20$ $K = 10^6$	$N = 50$ $K = 10^6$	$N = 100$ $K = 10^6$	$N \to \infty$ $K \to \infty$
Copula (III A)	4.8e-5	0.028	−0.0027	0.093	(0)
gif (III B)	0.28	0.17	0.01	−0.16	(0)
max entropy (3)	0.45	0.39	0.31	0.28	0.22
Green (20)	0.54	0.63	0.69	0.72	0.75
Hernando (22)	0.31	0.28	0.22	0.19	0.14

FIG. 2. Effective pair potential for the partitioned-gas model. The lower black curve shows the analytical approximate result [Eq. (30)]. The upper red curve is the numerically optimized pair potential for $N = 10$ particles, which yields a structure factor as close as possible to the exact $S_{\mathrm{PG}}(q)$ of Eq. (29). See more details in Appendix C.

is imperative. Yet, on top of replacing the actual partitioning by a fictitious pair potential, our expression for that pair potential in terms of $S(q)$ [Eq. (18)] is an approximation. We would like to isolate and assess the effect of this approximation. To do so we use a numerical optimization of the function $\lambda(q)$, similar in spirit to the algorithm of Ref. [20]. Simulated particle configurations $\{R_n\}$ and trial values for $\lambda(q)$ are used to construct a distribution p_s according to Eq. (8). This p_s is used to obtain a structure factor $S(q)$ according to Eq. (4). Then the distance between the obtained $S(q)$ and $S_{\mathrm{PG}}(q)$ of Eq. (29) is minimized over the trial values of $\lambda(q)$. Since any physical structure factor, including $S_{\mathrm{PG}}(q)$, can be tied to a unique pair potential $\lambda(q)$, the optimization in principle can get arbitrarily close to the exact $\lambda(q)$, i.e., to the exact maximum-entropy model. It is limited only by the number of the discrete trial λ. We note, however, that the inverse problem of obtaining the pair potential from correlations is known to be notoriously difficult [48]. More details of the numerical procedure are found in Appendix C. Figure 2 shows the optimized $\lambda(q)$ for $N = 10$ against the one found in Eq. (30).

Substituting the numerically optimal $S(q)$ in Eq. (3), we get the entropy of the maximum-entropy model while eliminating the additional approximation leading to Eq. (18). For $N = 10$ we find $H/N = 0.47$, to be compared with the value of 0.45 obtained when the structure factor from the simulations is used in Eq. (3) (see Table I). Thus, despite the apparent differences seen in Fig. 2 between the approximate pair potential (30) and the numerically optimized one, the ultimate effect on the entropy estimation is a mere 4.4%. We will comment on this point in Sec. V.

B. One-dimensional shear model

We now examine a 1D model proposed in Ref. [34] to schematically mimic periodically sheared suspensions. The model consists of N particles of diameter σ placed at positions $x_1, \ldots, x_N \in [0, L]$. Boundary conditions are periodic. At each simulation time step, particles are displaced a distance $\gamma\sigma$ to the right and back. Thus, a particle at position x_k will

collide with other particles positioned within the segment $[x_k - \sigma, x_k + \sigma(\gamma + 1)]$. At the end of each step, colliding particles are randomly displaced by an independent normally distributed shift with zero mean and variance $\epsilon\sigma$. See Fig. 1(b) for snapshots from simulations.

Varying γ while keeping all other parameters (N, $\bar{\rho} = N/L$, σ, and ϵ) fixed, one finds two distinct regimes. For γ smaller than a critical threshold γ_C, the system reaches an absorbing state in which no collisions occur. The time required to reach the final state diverges as $\gamma \to \gamma_C$. For $\gamma > \gamma_C$, the average rate of collisions per step does not vanish but tends to a constant. Interestingly, the critical threshold γ_C is smaller than the maximal density for which there exists a configuration with no collisions, $(\gamma + 2)\sigma/L$. This implies that the system does not reach an absorbing state even though there are configurations with no overlaps. For $\gamma = \gamma_C$, the distribution of particle positions is hyperuniform, i.e., strongly correlated, with a structure factor that vanishes as $q \to 0$ [34–36].

To estimate the entropy of the 1D shear model, 10^6 realizations with independent initial conditions were run with parameters $\sigma = 1$, $\epsilon = 1/2$, $\bar{\rho} = 0.2$, $N = 10, 20, 50, 100$, and γ in the range [0,6]. Steps were repeated until an absorbing state was reached (no collisions) or the collision rate changed by less than 1% in the second half of the simulation. Thus, for every N and γ, we obtain 10^6 independent samples from the steady-state $D = N$-dimensional distribution of $(x_1, \ldots, x_N)$. The samples are used to estimate the entropy by applying the three methods described above. To decrease the phase-space volume, all samples were sorted in increasing order along the x axis prior to estimation. We also show results for $N = 1000$ using 10^5 realizations.

Figure 3 shows the entropy estimations. For $N \leqslant 100$ all three methods yield a sharp change in entropy at the critical shear γ_C, as obtained independently by marking the largest value of γ for which an absorbing state is reached. Note that, due to the small system sizes, γ_C varies with N. The horizontal dashed line shows the ideal-gas entropy. At higher phase-space dimensions (larger N), estimations from the copula and compression methods become noisy and fluctuating, whereas the results of the structure-factor method become smoother. For $N = 1000$ the compression method no longer captures the transition. Overall, since the exact entropy is unknown, we cannot definitely say which method is more accurate. For small N, however, we expect the direct sampling (copula method) to give the most reliable results, as confirmed in Sec. IV A.

Thus, the entropy estimations using the structure factor and compression are in line with direct sampling. Our results demonstrate that entropy estimation is a robust method for detecting global transitions out of equilibrium even for small systems. Surprisingly, the sharp change in entropy is evident even for very small systems, containing just 10 particles.

C. Two-dimensional random organization model

Next, we check a 2D random organization (RandOrg) model. The model consists of N particles of diameter σ, placed in a rectangular domain $[0, L]^2$ with periodic boundaries. We implement a zero-shear version of the model
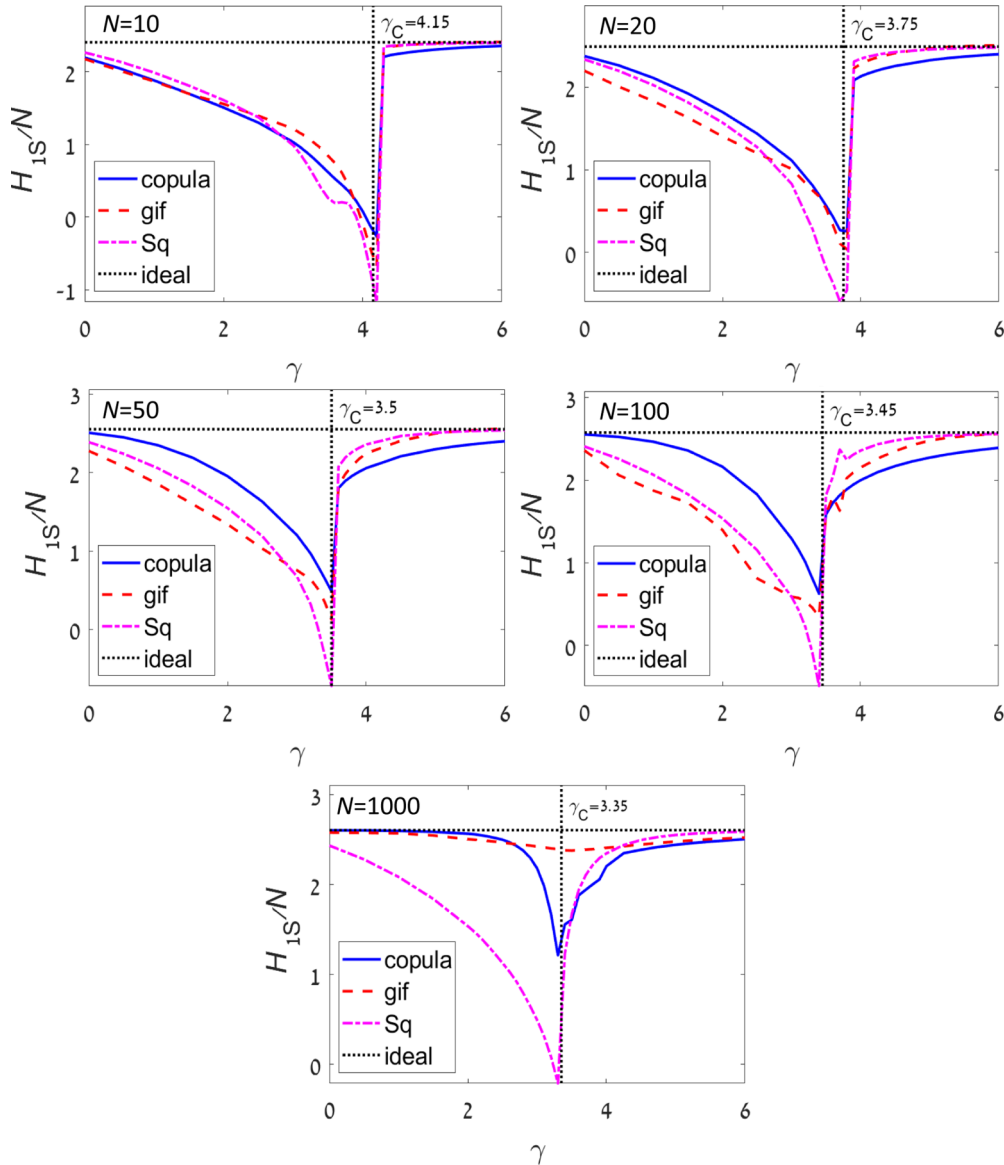
FIG. 3. 1D shear model: Entropy per particle with $N = 10, 20, 50, 100$, and 1000 particles. For the smaller systems all methods show a sharp change in entropy at the critical shear γ_C as inferred from the largest value of γ for which an absorbing state is reached (vertical dashed line). The horizontal dashed line shows the ideal-gas entropy.

[35,36], in which overlapping particles, termed "active," are displaced randomly. The displacements are uniformly distributed in a small circle with radius ϵ. When varying the area fraction $\rho = \pi\sigma^2/4L^2$ while keeping the rest of the parameters fixed, two regimes are observed, similar to the 1D shear model.

For ρ smaller than a critical threshold ρ_C, the system reaches an absorbing state in which no particles overlap. The time required to reach the final state diverges as $\rho \to \rho_C$. For $\rho > \rho_C$, the average rate of random displacements per step does not vanish but tends to a constant. Again, ρ_C is smaller than the maximum packing ratio for circles ($\sqrt{3}\pi/6 \simeq 0.907$). For $\rho = \rho_C$, the distribution of particle positions is hyperuniform, with a structure factor that vanishes as $q \to 0$ [35,36]. See Figs. 1(c)–1(e) for snapshots from simulations, below, at, and above the critical area fraction.

In order to estimate entropy in the 2D RandOrg model, 10^6 realizations with independent initial conditions were run with parameters $\sigma = 1, \epsilon = 1/2, N = 10, 20, 50$, and ρ in the range [0.01,0.6]. Steps were repeated until an absorbing state was reached (no collisions) or the collision rate changed by less than 1% in the second half of the simulation. Thus, for every N and ρ, we obtain 10^6 independent samples of the $D = 2N$-dimensional distribution of $(x_1, \ldots, x_N, y_1, \ldots, y_N)$. The samples are used to estimate the entropy by applying the three methods described above. Again, in order to decrease the phase space volume, all samples were sorted by their x component in increasing order prior to estimation. The order of the x and y coordinates, e.g. $(x_1, \ldots, x_N, y_1, \ldots, y_N)$ or $(x_1, y_1, \ldots, x_N, y_N)$, did not change the results.

Figure 4 shows the entropy estimations for the 2D RandOrg model. Similar to the 1D shear model, all methods yield a a sharp change in entropy at the critical density ρ_C;
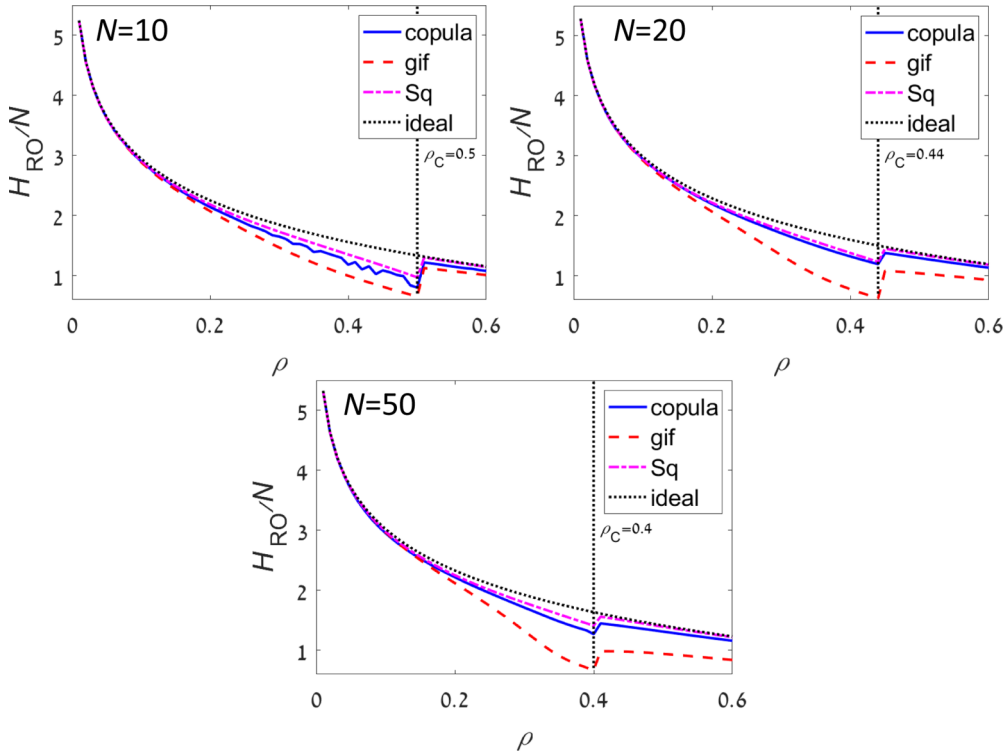
FIG. 4. 2D RandOrg model: Entropy per particle for $N = 10$, 20, and 50 particles (corresponding to a multidimensional phase space of 20, 40, and 100 dimensions, respectively). All methods show a small sharp change in entropy at the critical density ρ_C (vertical dashed line) inferred from the largest value of ρ for which an absorbing state is reached. The dashed curve shows the ideal-gas entropy.

yet, the change is smaller than in the 1D case. Here the structure-factor estimations lie closer to the copula results than the compression-based ones (which are lower). Note that the jumps seen in Fig. 4 are significantly larger than those reported in Ref. [29] for a similar model (see Fig. 4 there). In Ref. [29] the entropy was estimated using the lossless compression approach. The method was applied, however, to a single sample of a significantly larger system [about $(2$–$20) \times 10^3$ particles]. This difference demonstrates the advantage of considering systems with a relatively small number of particles when evaluating entropy based on data compression.

We further use the 2D RandOrg results, for which statistical estimation of the entropy is more challenging than in one dimension, to demonstrate some of the technical issues involved in the estimations. We begin with the copula method. Both the compression and structure-factor methods use the ideal-gas case as a reference value; they are, by definition, exact for an ideal gas. In contrast, direct estimation methods such as copula splitting are general-purpose estimators and do not make such assumptions. Hence, we treat the estimation of the copula method for the ideal-gas case as a systematic error, subtracting it from all the results. The difference between the raw results and the shifted ones is depicted in Fig. 5(a). Note that this error significantly increases with the dimension, growing from practically negligible at $N = 10$, to the values depicted for $N = 50$ in Fig. 5(a).

In the compression-based method, as explained in Sec. III B, the data are converted into a matrix of 8-bit unsigned integers, in which each row constitutes an independent sample. It is well known that the order of the compressed

data can have a significant effect on the compression ratio. For example, it is more efficient to compress text files by scanning columns rather than rows, because columns tend to be more repetitive. Similarly, in our application, transposing the data matrix yields significantly worse compression ratios because rows are independent, which implies that correlations among in-sample data are long-ranged. Figure 5(b) shows entropy estimates for the 2D RandOrg model obtained by transposing the data matrix. This small change causes the method to perform poorly, and the entropy at all surface fractions is practically indistinguishable from that of an ideal gas. We also tested ordering the samples along a Hilbert scan of the 2D square domain [29]. Our numerical examples show no advantage of the Hilbert scan compared to simple sorting along one dimension [see Figs. 5(a) and 5(b)].

Finally, Fig. 5(c) compares our relation, Eq. (3), with the alternative relations between entropy and pair correlations, Eqs. (20) and (22). We see that all expressions agree qualitatively (recall that Green's expression performed poorly in the partitioned-gas example). The expression due to Hernando slightly misses the location of the critical density. Equation (3) produces smoother and sharper results, partly due to the binning of interparticle distances required for the other two relations.

V. DISCUSSION

One of this work's aims has been to put the measurement and computation of entropy out of equilibrium on firmer ground. We have proved that spatial density correlations,
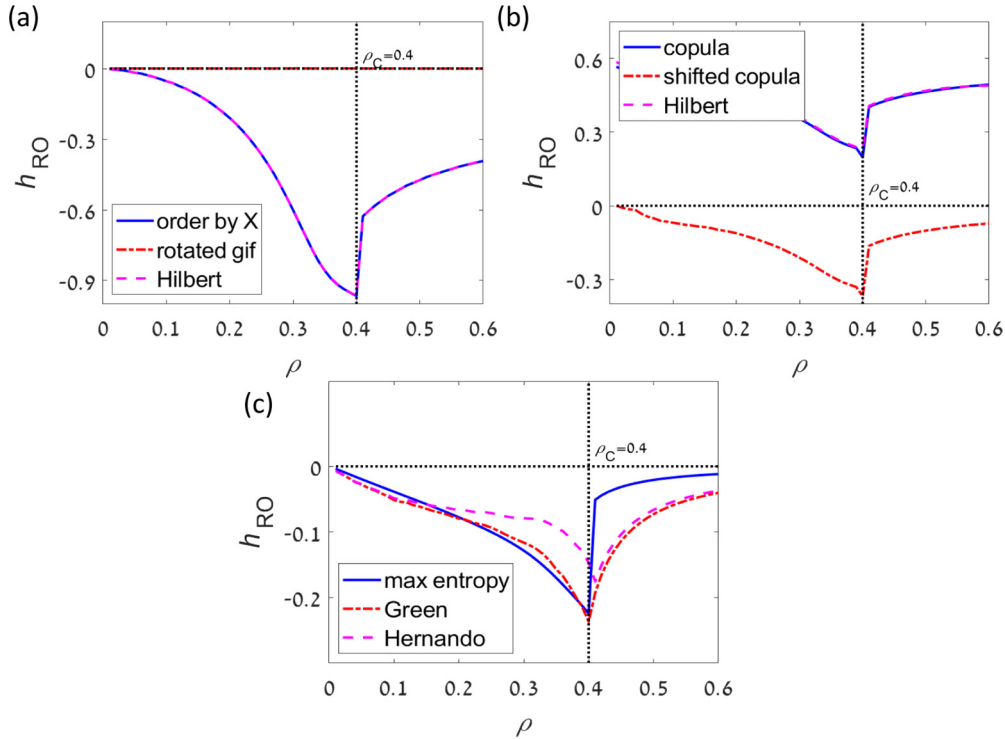
FIG. 5. Method issues. (a) Copula method: the raw results are shifted to match the ideal-gas limit. (b) Compression method: the results are highly sensitive to the order of independent samples in the compressed data matrix (red line), but not to the order of points within each sample (magenta line showing reordering by a Hilbert scan). (c) Structure factor: comparing different entropy–structure relations. All panels present the excess entropy per particle for the 2D RandOrg model with $N = 50$.

captured by the directly measurable structure factor, can be used to obtain a rigorous upper bound for the entropy of a particulate system. To infer the upper bound from structure, we have used an approximate relation [Eq. (3)] which, while giving comparable results to known relations, has the advantage of depending only on the Fourier-space structure factor [cf. Fig. 5(c)]. We have tested the structure-based entropy estimation, together with two additional methods based on either direct sampling or data compression, on three systems in and out of equilibrium. The direct-sampling (copula) method has offered a reliable benchmark for finding the entropy of small systems out of equilibrium (a feature that was missing in earlier studies).

Overall, the three methods have yielded qualitatively consistent estimations, and successfully identified nonequilibrium transitions. The statistical sampling and compression methods become less accurate for larger systems, because sampling high-dimensional distributions requires excessively large number of samples. The computational complexity of statistical sampling methods, such as the copula splitting applied here, severely limits their applicability even for moderate system sizes (up to 100 particles). Compression methods can be applied to much larger systems; yet, they do not bypass the sampling issue. In comparison, the structure-based method does not have this limitation. In the considered examples, for small systems, the results of the structure-based method are closer to the direct-sampling estimation, which may indicate better accuracy compared to data compression. We conclude that consistent entropy estimation out of equilibrium is achievable, with significant advantages to

doing so based on the structure factor whenever this method is relevant.

Two major approximations underlie Eq. (3). First, the relation takes into account only pair correlations. Information on higher-order correlations will necessarily reduce the entropy; thus, the upper bound considered here will not be violated but pushed further down. Incorporating such additional constraints into the construction in Sec. II A and the derivation in Appendix A is possible. Second, the partition function $\mathcal{Z}$ has been calculated within the simplest Gaussian approximation. Various methods have been developed in statistical physics to go beyond this approximation. [The exact Eq. (16) is a good starting point. See also Ref. [27].] However, unlike the usual case where the pair potential is given, here, having calculated a more accurate $\mathcal{Z}$, we will still need to find the effective interaction $\lambda(\mathbf{q})$. This will involve integral equations, requiring numerical solution or additional assumptions. The effects of the two approximations are hard to decouple. We have been able to do this for the partitioned gas, whose exact structure factor is known. Our numerical study indicates that the error in the entropy estimation (about 22%) has a small contribution from the Gaussian approximation (less than 5%); hence, the main source of error in the partition-gas example lies in the assumption of pair interactions. Improving on both approximations may be technically difficult, but the way to go about it is known.

In view of the above, Eq. (3) has performed remarkably well. It is in line with the copula method for small systems and accurately located critical points involving strong and long-ranged correlations. The reason might lie in the variational

nature of the calculation. There are well-known examples [e.g., the Self-Consistent Expansion [49]], where optimizing an inaccurate expression using free parameters [here $\lambda(\mathbf{q})$] to match a certain target, produces surprisingly accurate results. This is probably why, for the partitioned gas, the relation gives reasonable entropy estimates even though the model contains no actual pair interactions, and also why the Gaussian approximation introduces such a small error there despite the strong correlations built into the partitioned-gas model.

Another important property of the structure factor is that it implicitly takes into account symmetries of the system such as particle indistinguishability and translation invariance, to which the copula splitting and data compression are blind. This significantly reduces the relevant configuration space, thus improving the effectiveness of random sampling.

Hyperuniformity, the strong suppression of large-scale density fluctuations, has drawn significant attention lately [46]. The three examples chosen for the present study all exhibit hyperuniform states. This phenomenon is commonly defined and characterized through the structure factor, specifically, its decay to zero as $q \to 0$. Thus, Eq. (3) can readily be used to estimate the entropy cost of hyperuniformity. This may serve as an important ingredient in future modeling of hyperuniform systems.

Our reference state throughout this work has been the ideal gas [$H = 0$ for $S(\mathbf{q}) = 1$]. Certain systems may require a different reference state. For example, a fluctuating crystal or a "stealthy" hyperuniform system [46], whose $S(\mathbf{q})$ vanishes on a finite range of $\mathbf{q}$, have according to Eq. (3) an infinitely negative entropy. The natural reference in this case should be a perfect crystal, and a different ("low-temperature") calculation would be needed.

This work has been focused on entropy changes related to the spatial arrangement of particles. Materials exhibit a much wider variety of behaviors. There are phenomena (e.g., the glass transition) which leave little to no signature on the structure factor. Entropy changes associated with such phenomena, evidently, will not be captured by Eq. (3). Nevertheless, the approach taken here—specifying a constraint (including the proper symmetries) and obtaining an entropy bound from an effective equilibrium analogue—should be of more general applicability.

APPENDIX A: DERIVATION OF $\mathcal{Z}[S]$

This Appendix gives the detailed calculation leading to Eq. (17). Our starting point is the configurational integral (15),

$$\mathcal{Z} = \int \left(\prod_{n=1}^{N} d\mathbf{R}_n \right) \exp \left\{ - \int d\mathbf{q} \lambda(\mathbf{q}) \left[\sum_{m=1}^{N} \sum_{n=1}^{N} e^{i\mathbf{q} \cdot (\mathbf{R}_m - \mathbf{R}_n)} - NS(\mathbf{q}) \right] \right\}, \tag{A1}$$

and we apply a standard procedure to transform it into a functional integral of continuous fields.

The density field is $\rho(\mathbf{r}) = \sum_{n=1}^{N} \delta(\mathbf{r} - \mathbf{R}_n)$, and its Fourier transform is $\tilde{\rho}(\mathbf{q}) = \sum_{n=1}^{N} e^{-i\mathbf{q} \cdot \mathbf{R}_n}$. We define the conjugate field $\tilde{\psi}(\mathbf{q})$ according to

$$\prod_{\mathbf{q}} \delta \left[\tilde{\rho}(\mathbf{q}) - \sum_{n} e^{-i\mathbf{q} \cdot \mathbf{R}_n} \right] = (2\pi \nu)^{-\Omega_q} \int \mathcal{D}\tilde{\psi} \exp \left[i \int d\mathbf{q} \tilde{\psi} \left(\tilde{\rho} - \sum_{n} e^{-i\mathbf{q} \cdot \mathbf{R}_n} \right) \right].$$

Equation (A1) is then rewritten as

$$\mathcal{Z} = (2\pi \nu)^{-\Omega_q} e^{N \int d\mathbf{q} \lambda S} \int \left(\prod_{n} d\mathbf{R}_n \right) \mathcal{D}\tilde{\rho} \mathcal{D}\tilde{\psi} \exp \left[i \int d\mathbf{q} \tilde{\psi} \left(\tilde{\rho} - \sum_{n} e^{-i\mathbf{q} \cdot \mathbf{R}_n} \right) - \int d\mathbf{q} \lambda |\tilde{\rho}|^2 \right]. \tag{A2}$$

The integration over the discrete configuration $(\mathbf{R}_1, \dots, \mathbf{R}_N)$ is taken care of through

$$\int \left(\prod_{n} d\mathbf{R}_n \right) \exp \left(-i \int d\mathbf{q} \tilde{\psi} \sum_{n} e^{-i\mathbf{q} \cdot \mathbf{R}_n} \right) = \left[\int d\mathbf{R} \exp \left(-i \int d\mathbf{q} \tilde{\psi} e^{-i\mathbf{q} \cdot \mathbf{R}} \right) \right]^{N}$$

$$= \left\{ V + \int d\mathbf{R} \left[\exp \left(-i \int d\mathbf{q} \tilde{\psi} e^{-i\mathbf{q} \cdot \mathbf{r}} \right) - 1 \right] \right\}^{N} \xrightarrow{N, V \to \infty, \ N/V = \bar{\rho}} V^N \mathcal{I}[\tilde{\psi}(\mathbf{q})],$$

where

$$\mathcal{I}[\tilde{\psi}(\mathbf{q})] \equiv \exp \left[\bar{\rho} \int d\mathbf{R} (e^{-i \int d\mathbf{q} \tilde{\psi} e^{-i\mathbf{q} \cdot \mathbf{R}}} - 1) \right].$$

We are left now with continuous fields only,

$$\mathcal{Z} = Z_{\text{id}} (2\pi \nu)^{-\Omega_q} e^{N \int d\mathbf{q} \lambda S} \int \mathcal{D}\tilde{\rho} \mathcal{D}\tilde{\psi} \mathcal{I}[\tilde{\psi}] \exp \left[\int d\mathbf{q} (i \tilde{\psi} \tilde{\rho} - \lambda |\tilde{\rho}|^2) \right], \tag{A3}$$

where $Z_{\mathrm{id}} = V^N$ (or $V^N/N!$ if the particles are indistinguishable) is the uncorrelated (ideal gas) contribution.

We integrate over $\tilde{\rho}(\mathbf{q})$ to get

$$\mathcal{Z} = Z_{\mathrm{id}}(2\pi\nu)^{-\Omega_q} e^{N \int d\mathbf{q}\lambda S} \prod_{\mathbf{q}} [\pi\nu/\lambda(\mathbf{q})]^{1/2} \int \mathcal{D}\tilde{\psi}\mathcal{I}[\tilde{\psi}] \exp\left[-\int d\mathbf{q}\frac{|\tilde{\psi}|^2}{4\lambda}\right]$$

$$= Z_{\mathrm{id}}2^{-\Omega_q}\pi^{-\Omega_q/2}\exp\left\{\int d\mathbf{q}\left[N\lambda S - \frac{\nu}{2}\ln(\lambda/\nu)\right]\right\}Z_\psi, \tag{A4}$$

with

$$Z_\psi = \nu^{-\Omega_q}\int \mathcal{D}\tilde{\psi}e^{-\mathcal{H}_\psi}, \quad \mathcal{H}_\psi = -\bar{\rho}\int d\mathbf{R}(e^{-i\int d\mathbf{q}\tilde{\psi}e^{-i\mathbf{q}\cdot\mathbf{R}}} - 1) + \int d\mathbf{q}\frac{|\tilde{\psi}|^2}{4\lambda}. \tag{A5}$$

These expressions are exact.

We proceed with a Gaussian approximation for the integral (A5). To order $\tilde{\psi}^2$,

$$\mathcal{H}_\psi \simeq i(2\pi)^d\bar{\rho}\tilde{\psi}(0) + \int d\mathbf{q}\frac{|\tilde{\psi}|^2}{4f(\lambda)}, \quad f(\lambda) \equiv (\alpha^{-1} + \lambda^{-1})^{-1}, \tag{A6}$$

where $\alpha \equiv [2(2\pi)^d\bar{\rho}]^{-1}$. Omitting the $\mathbf{q} = 0$ term and integrating over $\tilde{\psi}$, we get

$$Z_\psi \simeq \prod_{\mathbf{q}}[(4\pi/\nu)f(\lambda)]^{1/2} = (4\pi)^{\Omega_q/2}\exp\left\{\frac{\nu}{2}\int d\mathbf{q}\ln[f(\lambda)/\nu]\right\}. \tag{A7}$$

Substitution in Eq. (A4) gives

$$\mathcal{Z} = Z_{\mathrm{id}}\exp\left(\int d\mathbf{q}\left\{N\lambda S + \frac{\nu}{2}\ln[f(\lambda)/\lambda]\right\}\right), \tag{A8}$$

which coincides with Eq. (17).

APPENDIX B: IMPLEMENTATION OF STATISTICAL ENTROPY ESTIMATIONS

1. Direct statistical estimation

As discussed in Sec. III A, we apply a recently suggested method that is based on decomposing the distribution into a product of the marginal distributions and the joint dependency, also known as the copula [18]. Recall Sklar's theorem [44,45], which states that any continuous multidimensional density $p(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^D$ can be written uniquely as

$$p(\mathbf{x}) = p_1(x_1) \cdot \dots \cdot p_D(x_D)c(F_1(x_1), \dots, F_D(x_D)), \tag{B1}$$

where, $\mathbf{x} = (x_1, \dots, x_D)$, $p_k(\cdot)$ denotes the marginal density of the kth dimension with CDF, $F_k(t) = \int_{-\infty}^{t} p_k(x)\,dx$, and $c(u_1, \dots, u_D)$ is the density of the copula, i.e., a probability density on the hyper-square $[0, 1]^D$ whose marginals are all uniform on $[0,1]$.

Analytically, the copula is obtained by a change of variables, $u_k = F_k(x_k)$. This motivates estimating the copula as follows. Let $\mathbf{x}^i = (x_1^i, \dots, x_D^i) \in \mathbb{R}^D$, where $i = 1 \dots N$ denote N independent samples from a real D-dimensional random variable with density $p(\mathbf{x})$. We would like to use the samples $\mathbf{x}^i$ in order to obtain samples from the copula density $c(u_1, \dots, u_D)$. This can be obtained by finding the rank (in increasing order) of samples along each dimension. Dividing by N, we obtain, $\mathbf{u}^i = (u_1^i, \dots, u_D^i) \in [0, 1]^D$, were $u_k^i = n/(N + 1)$ if x_k^i is the n smallest value out of the list $u_k^1, \dots, u_k^N$. Overall, we obtain N samples from the distribution $c(u_1, \dots, u_D)$.

In order to approximate the entropy, 1D entropy of marginals is estimated using either uniform binning or sample-spacing methods, depending on whether the support of the marginal is known to be compact (bins) or unbounded or unknown (spacing). The main challenge lies in evaluating the differential entropy of high-dimensional copulas [13,50]. Here, we compute it recursively, similar to the kDP approach [41]. Let $k \in \{1, \dots, D\}$ be one of the dimensions, to be chosen using any given order. The copula samples u^i are split into two equal parts (note that the median in each dimension is always $1/2$). Denote the two halves as $v_j^i = \{u_j^i|u_k^i \leqslant 1/2\}$ and $w_j^i = \{u_j^i|u_k^i > 1/2\}$. Scaling the halves as $2v_j^i$ and $2w_j^i - 1$ produces two sample sets for two copulas, each with $N/2$ points. A simple calculation shows that

$$H_c = \tfrac{1}{2}(H_{2v} + H_{2w-1}), \tag{B2}$$

where H_{2v} is the entropy estimate obtained using the set of points $2v_j^i$ and H_{2w-1} is the entropy estimate obtained using the set of points $2w_j^i - 1$. The marginals of each half may no longer be uniformly distributed in $[0,1]$, which suggests continuing recursively, i.e., the entropy of each half is a decomposed using Sklar's theorem, etc. See Fig. 6 for a schematic sketch of the method. See Ref. [18] for convergence analysis and numerical examples.

A Matlab code is available in MathWorks MATLAB Central.

2. Compression-based estimation

As discussed in Sec. III B, we estimate the entropy using a lossless compression algorithm. Similar to Ref. [28], samples are binned into 256 equal bins in each dimension, and the data are converted into a matrix of 8-bit unsigned integers,
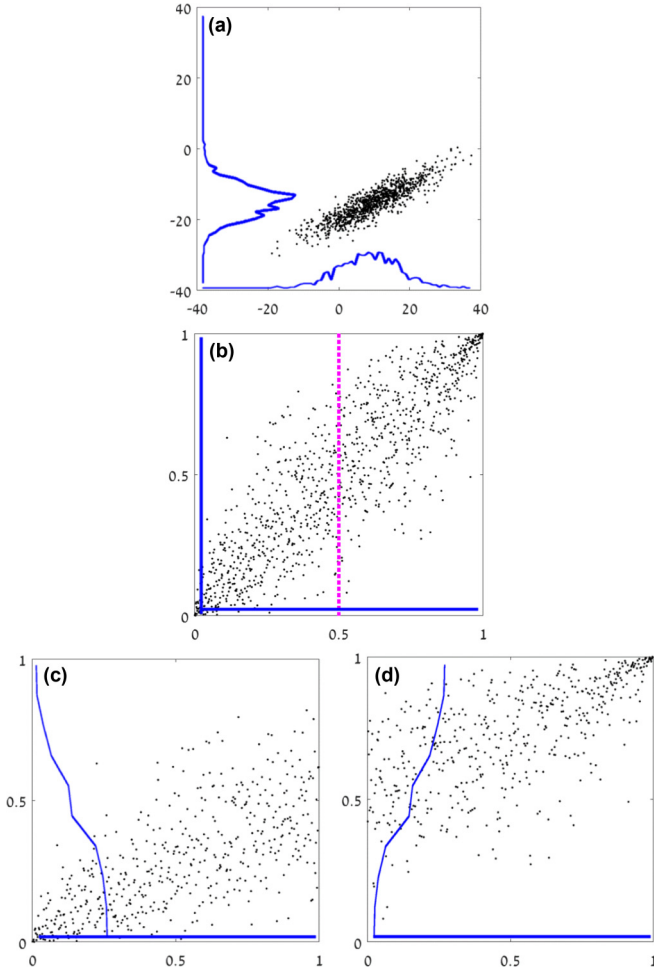
FIG. 6. A schematic sketch of the copula splitting method, reproduced from Ref. [18]. (a) A sample of 1000 points from a 2D Gaussian distribution. The blue lines depict the empirical density (obtained using uniform bins). (b) Enumerating the sorted data in each dimension (and dividing by N), the same data provides samples for the copula in $[0, 1]^2$. Splitting the data according to the median in one of the axes (always at 0.5) yields (c) (left half) and (d) (right half). The blue lines depict the empirical marginal densities in each half. These continue recursively.

in which each row constitutes an independent sample. The matrix is compressed using the LZW algorithm (implemented in Matlab's imwrite function) to a gif file.

Specifically, suppose we are given K samples, $x_1, \ldots, x_K$, in a D-dimensional box $[0, L]^D$. The algorithm we apply is as follows.

(1) Digitize the samples into 256 equally sized bins, $y_k = \lfloor 255 x_k / L \rfloor$, where $\lfloor \cdot \rfloor$ denotes rounding down to the nearest integer. Construct the $K \times D$ image matrix Y which has $y_1, \ldots, y_K$ as rows.

(2) Convert the image Y into a gif file. Denote the file size C_{data}.

(3) Convert an image Y_0, a $K \times D$ matrix of zeros (8-bit unsigned integers) into a gif file. Denote the file size C_0.

(4) Convert an image Y_1, a $K \times D$ matrix of random integers between 0 and 255 (8-bit unsigned integers) into a gif file. Denote the file size C_1.

(5) Interpolation: Let $\lambda = (C_{\text{data}} - C_0)/(C_1 - C_0)$ and take $H = \lambda H_0 + (1 - \lambda) H_{\text{ideal}}$. Here $H_0 = D(\ln L - \ln 255)$, the maximum entropy of a random variable that yields a sample matrix Y_0 (independent uniform distribution in $[0, L/255]$), and $H_{\text{ideal}} = -D \ln L$ is the entropy of an ideal gas of D particles in $[0, L]$.

APPENDIX C: NUMERICAL OPTIMIZATION OF $\lambda(q)$

This Appendix gives further details concerning the optimization results presented in Sec. IV A. A similar method was presented in Ref. [20]. To optimize $\lambda(q)$ we need to solve the system of equations resulting from the Lagrange multipliers, Eqs. (7). The main idea is that, given values $\lambda(\mathbf{q}_1), \ldots, \lambda(\mathbf{q}_M)$ for a subset of wave vectors $\mathbf{q}$, one can compute the structure factor (assuming all other λ are zero). Since the structure factor of a distribution of the form of Eq. (8) is unique [21], we numerically find the set of λ that minimize the distance from $S(q)$ given by (4) and the exactly known $S_{\text{PG}}(q)$. Then the entropy is approximated by numerically computing $\mathcal{Z}$ of Eq. (9). Thus, the numerical optimization method finds the entropy of the maximum-entropy model, eliminating the error incurred by the second-order expansion in Sec. II B and Appendix A. This numerical approach is applicable only for low-dimensional systems.

We write (8), the probability of a configuration s, as

$$p_s = \hat{\mathcal{Z}}^{-1} \exp\left[-\sum_{\mathbf{q} \neq 0} \lambda(\mathbf{q}) \sum_{m,n} e^{i\mathbf{q} \cdot (\mathbf{R}_m^s - \mathbf{R}_n^s)} \right], \quad \text{(C1)}$$

where we absorbed the exponential term including the structure factor into the normalization constant,

$$\hat{\mathcal{Z}} = \mathcal{Z} \exp\left[-N \sum_{\mathbf{q} \neq 0} \lambda(\mathbf{q}) S(\mathbf{q}) \right]. \quad \text{(C2)}$$

The sum (or integral) over configurations s is approximated using a naive Monte Carlo method as follows. First, approximate the infinite sum over $\mathbf{q}$ by a finite number, $\mathbf{q} = (q_0 = 0, q_1, 2q_1, \ldots \Omega_q q_1)$, where $q_1 = 2\pi/N$ is the smallest (nonzero) wave number. We generate K independent random configurations from an ideal-gas distribution. In the partitioned-gas case with N particles, R_n^s are independent, uniformly distributed in $[0, N]$, $s = 1 \ldots K$, $n = 1 \ldots N$. Denote

$$\alpha_s = \sum_{0 \neq \mathbf{q} \in \Omega_q} \lambda(\mathbf{q}) \sum_{m,n} e^{i\mathbf{q} \cdot (\mathbf{R}_m^s - \mathbf{R}_n^s)}, \quad \tilde{\alpha}_s = \alpha_s - \langle \alpha_s \rangle,$$

$$\tilde{p}_s = \tilde{\alpha}_s / \sum_s \tilde{\alpha}_s. \quad \text{(C3)}$$

Thus, $\tilde{p}_s$ is a numerical approximation of p_s. Note that we have centered the exponential factors α_s to avoid excessively large or small exponents. In the above, $\langle \cdot \rangle$ denotes averaging over all samples s. Substituting into the structure factor, Eq. (4), while replacing p_s with $\tilde{p}_s$, we obtain an approximation for the structure factor $S(\mathbf{q})$. Thus, given a set of Lagrange multipliers $\lambda(\mathbf{q})$, $q \in \Omega_q$, we can compute the resulting structure factor $\tilde{S}(\mathbf{q})$, assuming that $\lambda(\mathbf{q}) = 0$ for all $\mathbf{q} \notin \Omega_q$.

Our goal is to find a set of $\lambda(\mathbf{q})$, $q \in \Omega_q$, such that $\tilde{S}(\mathbf{q})$ is close to the required $S(\mathbf{q})$. Here we applied Matlab's nonlinear least-square fitting function lsqcurvefit using a zero initial guess. Figure 2 shows results with $N = 10$. Parameters are $\Omega_q = 10N$ and $K = 1000N\Omega_q$. The method works only for relatively small N (up to about 10). At higher dimensions, the naive Monte Carlo sampling fails and more sophisticated methods biasing the sampling toward high-probability regions are necessary.

Next, we approximate the entropy H. Here, care must be taken when converting integral to sums, as prefactors change the entropy. Recall the partition function, Eq. (15) or (A1). Several approximations are required: (a) Integration over configurations is replaced by Monte-Carlo sampling. (b) The integral over $\mathbf{q}$ is replaced with a sum over a finite number Ω_q of wave numbers. (c) The structure factor $S(\mathbf{q})$ is replaced with its approximation $\tilde{S}(\mathbf{q})$. Overall, $\mathcal{Z}$ is approximated as

$$\tilde{\mathcal{Z}} = K^{-1} \sum_{s=1}^{K} \exp \left\{ -\Delta q \sum_{0 \neq q \in \Omega_q} \lambda(\mathbf{q}) \left[\sum_{m=1}^{N} \sum_{n=1}^{N} e^{i\mathbf{q}\cdot(\mathbf{R}_m - \mathbf{R}_n)} - N\tilde{S}(\mathbf{q}) \right] \right\}, \tag{C4}$$

where $\Delta q = 2\pi/N$. Finally, the entropy $H = \ln \tilde{\mathcal{Z}}$ is computed as

$$\beta_s = \Delta q \sum_{0 \neq q \in \Omega_q} \lambda(\mathbf{q}) \left[\sum_{m=1}^{N} \sum_{n=1}^{N} e^{i\mathbf{q}\cdot(\mathbf{R}_m - \mathbf{R}_n)} - N\tilde{S}(\mathbf{q}) \right],$$

$$\tilde{\beta}_s = \beta_s - \langle \beta_s \rangle,$$

$$H = \ln \tilde{\mathcal{Z}} = \ln \left(\sum_s e^{-\tilde{\beta}_s} \right) - \ln K - \langle \beta_s \rangle. \tag{C5}$$

Once again, to reduce round-off errors, arguments are centered before taking the exponent.

[1] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, Ann. Rev. Condens. Matter Phys. **2**, 329 (2011).

[2] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[3] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).

[4] J. Casas-Vazquez and D. Jou, Temperature in non-equilibrium states: A review of open problems and current proposals, Rep. Progr. Phys. **66**, 1937 (2003).

[5] L. F. Cugliandolo, The effective temperature, J. Phys. A **44**, 483001 (2011).

[6] A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardar, and J. Tailleur, Pressure is not a state function for generic active fluids, Nat. Phys. **11**, 673 (2015).

[7] R. L. Stratonovich, The entropy of systems with a random number of particles, Sov. Phys. JETP **1**, 254 (1955).

[8] J. Parrondo, J. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

[9] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, Chapter 14: Gibbs and Boltzmann entropy in classical and quantum mechanics, in *Statistical Mechanics and Scientific Explanation*, edited by V. Allori (World Scientific Publishing Co., Singapore, 2020), pp. 519-581 (2020).

[10] M. A. Kerroum, A. Hammouch, and D. Aboutajdine, Textural feature selection by joint mutual information based on Gaussian mixture model for multispectral image classification, Pattern Recogn. Lett. **31**, 1168 (2010).

[11] N. Kwak and C. H. Choi, Input feature selection by mutual information based on Parzen window, IEEE Trans. Pattern Anal. Machine Intel. **24**, 1667 (2002).

[12] S. Zhu, D. Wang, K. Yu, T. Li, and Y. Gong, Feature selection for gene expression using model-based entropy, IEEE/ACM Trans. Comput. Biol. Bioinformatics **7**, 25 (2010).

[13] R. S. Calsaverini and R. Vicente, An information-theoretic approach to statistical dependence: Copula information, Europhys. Lett. **88**, 68003 (2009).

[14] L. Faivishevsky and J. Goldberger, Ica based on a smooth estimation of the differential entropy, in *Advances in Neural Information Processing Systems 21*, edited by D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou (Neural Information Processing Systems Foundation, Inc. (NIPS), La Jolla, CA, 2009), pp. 433–440.

[15] J. Beirlant, E. J. Dudewicz, L. Györfi, and E. C. Van der Meulen, Nonparametric entropy estimation: An overview, Intl. J. Math. Stat Sci. **6**, 17 (1997).

[16] G. A. Darbellay and I. Vajda, Estimation of the information by an adaptive partitioning of the observation space, IEEE Trans. Inf. Theory **45**, 1315 (1999).

[17] L. Paninski, Estimation of entropy and mutual information, Neural Comput. **15**, 1191 (2003).

[18] G. Ariel and Y. Louzoun, Estimating differential entropy using recursive copula splitting, Entropy **22**, 236 (2020).

[19] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, Singapore, 1985).

[20] G. Zhang and S. Torquato, Realizable hyperuniform and non-hyperuniform particle configurations with targeted spectral

functions via effective pair interactions, Phys. Rev. E **101**, 032124 (2020).

[21] R. L. Henderson, A uniqueness theorem for fluid pair correlation functions, Phys. Lett. A **49**, 197 (1974).

[22] H. S. Green, A general kinetic theory of liquids. II. Equilibrium properties, Proc. R. Soc. London A **189**, 103 (1947).

[23] H. S. Green, *The Molecular Theory of Fluids* (North-Holland, Amsterdam, 1952).

[24] R. E. Nettleton and M. S. Green, Expression in terms of molecular distribution functions for the entropy density in an infinite system, J. Chem. Phys. **29**, 1365 (1958).

[25] J. A. Hernando, Thermodynamic potentials and distribution functions. II. The HNC equation as an optimized superposition approximation, Mol. Phys. **69**, 327 (1990).

[26] B. B. Laird and A. D. J. Haymet, Calculation of the entropy from multiparticle correlation functions, Phys. Rev. A **45**, 5680 (1992).

[27] S. Chakrabarty and Z. Nussinov, High-temperature correlation functions: Universality, extraction of exchange interactions, divergent correlation lengths, and generalized Debye length scales, Phys. Rev. B **84**, 064124 (2011).

[28] R. Avinery, M. Kornreich, and R. Beck, Universal and accessible entropy estimation using a compression algorithm, Phys. Rev. Lett. **123**, 178102 (2019).

[29] S. Martiniani, P. M. Chaikin, and D. Levine, Quantifying hidden order out of equilibrium, Phys. Rev. X **9**, 011031 (2019).

[30] O. Melchert and A. K. Hartmann, Analysis of the phase transition in the two-dimensional Ising ferromagnet using a Lempel-Ziv string-parsing scheme and black-box data-compression utilities, Phys. Rev. E **91**, 023306 (2015).

[31] E. Estevez-Rams, R. Lora-Serrano, C. A. J. Nunes, and B. Aragon-Fernandez, Lempel-Ziv complexity analysis of one dimensional cellular automata, Chaos **25**, 123106 (2015).

[32] P. Ronhovde, S. Chakrabarty, D. Hu, M. Sahu, K. K. Sahu, K. F. Kelton, N. A. Mauro, and Z. Nussinov, Detecting hidden spatial and spatio-temporal structures in glasses and complex physical systems by multiresolution network clustering, Eur. Phys. J. E **34**, 105 (2011).

[33] M. Zu, A. Bupathy, D. Frenkel, and S. Sastry, Information density, structure and entropy in equilibrium and non-equilibrium systems, J. Stat. Mech. (2020) 023204.

[34] L. Corte, P. Chaikin, J. Gollub, and D. Pine, Random organization in periodically driven systems, Nat. Phys. **4**, 420 (2008).

[35] D. Hexner and D. Levine, Hyperuniformity of critical absorbing states, Phys. Rev. Lett. **114**, 110602 (2015).

[36] D. Hexner, P. M. Chaikin, and D. Levine, Enhanced hyperuniformity from random reorganization, Proc. Natl. Acad. Sci. U. S. A. **114**, 4294 (2017).

[37] E. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).

[38] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Elsevier, London, 1986).

[39] Y. D. Fomin, V. N. Ryzhov, and N. V. Gribova, Breakdown of excess entropy scaling for systems with thermodynamic anomalies, Phys. Rev. E **81**, 061201 (2010).

[40] The expression derived in Ref. [25] contains typos, which are corrected in Ref. [26]. Interestingly, this higher-order relation can be obtained directly from Green's 1947 article [Ref. [22], Eq. (6.7)], when the pair distribution function is taken in its lowest-order approximation, $g(r) = e^{-\phi(r)/(k_{\mathrm{B}}T)}$, where $\phi(r)$ is the pair potential.

[41] D. Stowell and M. Plumbley, Fast multidimensional entropy estimation by k-d partitioning, IEEE Signal Process. Lett. **16**, 537 (2009).

[42] L. F. Kozachenko and N. N. Leonenko, Sample estimate of the entropy of a random vector, Prob. Peredachi Inf. **23**, 9 (1987).

[43] W. Lord, J. Sun, and E. Bollt, Geometric k-nearest neighbor estimation of entropy and mutual information, Chaos **28**, 033114 (2018).

[44] P. Jaworski, F. Durante, W. K. Hardle, and T. Rychlik, *Copula Theory and Its Applications* (Springer, New York, 2010).

[45] F. Durante and C. Sempi, Copula theory: An introduction, in *Copula Theory and Its Applications*, edited by P. Jaworski, F. Durante, W. K. Hardle, and T. Rychlik (Springer-Verlag, Berlin Heidelberg, 2010), pp. 3–33.

[46] S. Torquato, Hyperuniform states of matter, Phys. Rep. **745**, 1 (2018).

[47] Note that, since in all examples $S(\mathbf{q})$ vanishes as $\mathbf{q} \to 0$, the integral in Eq. (3) is singular (yet convergent). Several numerical methods can integrate over a singular point. For simplicity, we use the trapezoidal rule and simply set the integrand at $\mathbf{q} = 0$ to zero.

[48] H. C. Nguyen, R. Zecchina, and J. Berg, Inverse statistical problems: From the inverse Ising problem to data science, Adv. Phys. **66**, 197 (2017).

[49] M. Schwartz and E. Katzav, The ideas behind self-consistent expansion, J. Stat. Mech. (2008) P04023.

[50] P. Embrechts and M. Hofert, Statistical inference for copulas in high dimensions: A simulation study, ASTIN Bull. **43**, 81 (2013).