# Resolving entropy contributions in nonequilibrium transitions

Benjamin Sorkin [ORCID],[1] Joshua Ricouvier [ORCID],[2] Haim Diamant [ORCID],[1] and Gil Ariel [ORCID][3,*]

[1]*School of Chemistry and Center for Physics and Chemistry of Living Systems, Tel Aviv University, 69978 Tel Aviv, Israel*
[2]*Department of Chemical and Biological Physics, Weizmann Institute of Science, 76100 Rehovot, Israel*
[3]*Department of Mathematics, Bar-Ilan University, 52000 Ramat Gan, Israel*

We derive a functional for the entropy contributed by any microscopic degrees of freedom as arising from their measurable pair correlations. Applicable both in and out of equilibrium, this functional yields the maximum entropy which a system can have given a certain correlation function. When applied to different correlations, the method allows us to identify the degrees of freedom governing a certain physical regime, thus capturing and characterizing dynamic transitions. The formalism applies also to systems whose translational invariance is broken by external forces and whose number of particles may vary. We apply it to experimental results for jammed bidisperse emulsions, capturing the crossover of this nonequilibrium system from crystalline to disordered hyperuniform structures as a function of mixture composition. We discover that the cross-correlations between the positions and sizes of droplets in the emulsion play the central role in the formation of the disordered hyperuniform states. We discuss implications of the approach for entropy estimation out of equilibrium and for characterizing transitions in disordered systems.

## I. INTRODUCTION

Macroscopic states of matter differ by the statistics of their microscopic degrees of freedom (DOFs). Moreover, the physical mechanism underlying a change of state is reflected in the roles played by the different DOFs. One way to quantify the contributions of particular DOFs (or the interactions between them) to the change of state is through their effect on the entropy. Unlike temperature and pressure, whose consistent definitions, and thus measurement, are problematic out of equilibrium [1–3], the relation between entropy and information content is believed to hold for any steady state, regardless of whether the system is at thermodynamic equilibrium or not.

According to Shannon's definition [4], the information entropy is given by

$$H[P_s] = -\sum_s P_s \ln P_s, \tag{1a}$$

$P_s$ being the probability distribution of finding the system in the discrete microstate $s$. The continuous form, relevant in many physical systems, is

$$H[P(\mathbf{X})] = -\int d\mathbf{X} \, P(\mathbf{X}) \ln P(\mathbf{X}), \tag{1b}$$

with $P(\mathbf{X})$ being the probability density function of finding the system around the phase-space point $\mathbf{X}$.

Applying Eqs. (1) directly and, furthermore, resolving the contributions of the different DOFs, is usually impractical; out of equilibrium, the required knowledge of $P_s$ [and even more so, $P(\mathbf{X})$], is prohibitively hard to acquire by sampling [5–11].

Recently, several algorithms aiming to estimate entropy of physical systems have been demonstrated, based on data compression [12–16] or recursive mutual-information estimation [17]. Since such methods still rely on sampling of microscopic configurations, they are bound to suffer from undersampling for real materials due to the high dimensionality and continuity of phase space.

Given that the detailed $P(\mathbf{X})$ is practically unattainable, we propose to follow the trusted path of statistical mechanics and consider instead lower-dimensional marginal distributions, namely, correlation functions (CFs) [18]. Although it may seem like a severe compromise, this coarse graining has three important advantages. First, spatial CFs are robustly measurable in experiments and simulations, without extensive sampling of microscopic states. Second, they provide a useful characterization of the structure of matter, both in and out of equilibrium [18–22]. Third, by supplying different CFs, we will be able to discern their contributions to the entropy and determine the dominant factor underlying the changes in the system.

An illustrative example is the isotropic-to-nematic transition in a system of hard rods [23,24]. In this transition the physical mechanism is known, being the competition between the translational and rotational DOFs over the available phase-space volume. Resolving the different entropy contributions from the CF of each DOF, and possibly also their cross-correlation, would clarify this mechanism were it unknown, and indicate the transition point at which the rotational DOFs lose in the entropy competition and become ordered. The contribution of the orientations CF would change considerably between the two phases, while that of the density CF (i.e., the structure factor) would hardly change. Later we demonstrate this idea in a system whose underlying physics is much less

*\*arielg@math.biu.ac.il

TABLE I. Examples of possible fields and their corresponding correlation functions, $\mathcal{C}(\mathbf{q})$. In all examples, $\mathbf{r}_{nm} = \mathbf{r}_n - \mathbf{r}_m$ is the inter-particle separation. The $n, m$ summations are over all $N$ particles. Greek indices correspond to the $d$ spatial axes. Einstein's summation rule is applied. The normalization matrices $\mathcal{A}$ are computed using Eq. (3b). The unit vector $\hat{\mathbf{n}}$ is the solid angle. The nematic tensor is $\mathbf{Q} = (d\hat{\mathbf{n}}\hat{\mathbf{n}}^\dagger - \mathbf{I})/(d - 1)$. Since $\mathbf{Q}$ is symmetric, we rearrange its $M = d(d + 1)/2$ unique components in a column vector. For binary mixtures, $t_n = 1, 2$ denotes the type of the $n$th particle, and the structure factor adopts a matrix form, $\mathcal{S}_{ij}(\mathbf{q})$, where $i, j = 1, 2$ are particle-type labels and $N_i$ is the number of $i$-type particles. In the volume-fraction field, $a_n$ is the $n$th particle volume, and $\overline{a^2} = \int da \, \mathrm{Pr}(a)a^2$ is the second moment of the single-particle volume distribution. The correlation function $\chi(\mathbf{q})$ is known as the spectral density.

| Field | $\mathbf{j}_n$ | $\underline{f}(\mathbf{j})$ | $\mathcal{C}(\mathbf{q})$ |
|---|---|---|---|
| Density | None | 1 | $\mathcal{C} \equiv S(\mathbf{q}) = \langle \frac{1}{N} \sum_{n,m} e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}} \rangle$ |
| Velocity | $\mathbf{v}_n$ | $\mathbf{v}$ | $\mathcal{C}_{\alpha\beta} = (\mathcal{A}^{-1})_{\alpha\gamma} \langle \frac{1}{N} \sum_{n,m} e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}} v_{\gamma,n} v_{\beta,m} \rangle$ |
| Orientation | $\hat{\mathbf{n}}_n$ | $\hat{\mathbf{n}}$ | $\mathcal{C}_{\alpha\beta} = \langle \frac{d}{N} \sum_{n,m} e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}} \hat{n}_{\alpha,n} \hat{n}_{\beta,m} \rangle$ |
| Nematic tensor | $\hat{\mathbf{n}}_n$ | $\{Q_{\alpha\geqslant\beta}(\hat{\mathbf{n}})\}$ | $\mathcal{C}_{ij} = (\mathcal{A}^{-1})_{ik} \langle \frac{1}{N} \sum_{n,m} e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}} Q_k(\hat{\mathbf{n}}_n) Q_j(\hat{\mathbf{n}}_m) \rangle$ |
| Mixture composition | $t_n$ | $(\delta_{t1}, \delta_{t2})^T$ | $\mathcal{C}_{ij} \equiv \mathcal{S}_{ij}(\mathbf{q}) = \langle \frac{1}{N_i} \sum_{n=1}^{N_i} \sum_{m=1}^{N_j} e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}} \rangle$ |
| Volume fraction | $a_n$ | $a$ | $\mathcal{C} \equiv \chi(\mathbf{q}) = \frac{1}{\overline{a^2}} \langle \frac{1}{N} \sum_{n,m} e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}} a_n a_m \rangle$ |

understood, and whose transitions are less evident in the CFs *per se*.

The goal of the present work is to establish a theoretical framework for the above. We wish to estimate the entropy associated with a two-point CF of arbitrary nature. Our strategy is as follows. (a) We look for the information content of a given CF, i.e., the "entropy cost" of knowing that CF. We do so by finding the distribution $P(\mathbf{X})$ that maximizes the Shannon entropy, Eq. (1), under the constraint that it should yield the given two-point CF. Substituting it back into Eq. (1) gives the information content, i.e., the entropy bound. Obtaining this expression involves a Gaussian approximation (see Appendix B). A proof of concept for this step in the simplest case, where the DOFs are the positions of pointlike particles and the corresponding two-point CF is the structure factor, was presented in Ref. [25]. Here we extend the formalism to the general case of any DOFs and their CFs. (b) Imposing a set of measured CFs, we obtain a series of entropy upper bounds. Comparison of the different bounds points at the DOFs which dominate the physical behavior under consideration, and their interaction. The DOFs are completely arbitrary, and may include labels of particle species, their orientation, velocities, sizes, etc. (see Table I). (c) A tighter upper bound can be obtained by imposing several CF constraints at once. This allows us to isolate the contributions of cross-correlations to the underlying physics.

The formalism can be generalized further to include additional constraints, such as imposed inhomogeneity and fluctuations in the number of particles (see Appendix A). In addition, as a byproduct, we provide an alternative route for entropy estimation. By integrating out the unavailable statistics we bypass the undersampling problem.

Indeed maximum-entropy methods have proved useful for obtaining effective interaction models for complex, out-of-equilibrium systems, in both experiments (e.g., [26–28]) and simulations (e.g., [26,29,30]). To date, such methods are applicable only for maximization over a small set of effective parameters [26,27,29,30] and in the particular case of the structure factor [25,31,32].

The paper is constructed as follows. Section II gives the main result—the entropy functional of a given CF—and describes how it is implemented. In Sec. III we apply the

formalism to experimental results obtained for a confined emulsion of polydisperse droplets [33]. In Sec. IV we summarize our findings and discuss the many potential applications of the method. We defer more technical information to the Appendixes and Supplemental Material (SM) [34]. In Appendix A we present the required adjustments to the central result when translational symmetry is broken, e.g., by an external field. The detailed derivation of the main result is given in Appendix B, for the case where the CF is the only constraint. In the SM we derive the adjusted entropy bound in the presence of imposed inhomogeneity, and the treatment of a variable number of particles (grand-canonical ensemble). We also provide the "recipe" for how the CFs in the experiments were constructed, along with more experimental snapshots and figures of the CFs. Finally, the SM presents a toy model demonstrating the failure of various computational methods for entropy estimation in the case of a variable number of particles.

## II. GENERAL FORMALISM

The system of interest contains the locations of $N$ particles [their centers of mass (CMs)] in $d$ dimensions, $\{\mathbf{r}_{n=1,\ldots,N} \in \mathbb{R}^d\}$, and possibly an additional set of DOFs that we denote by $\{j_{k,n=1,\ldots,N}\}$. The index $k$ marks the scalar components of all additional DOFs, and we abbreviate $\mathbf{j}_n = (j_{1,n}, j_{2,n}, \ldots)$. For example, in the case of rodlike particles, $\mathbf{j}_n$ may include the components of the solid angle, CM velocity, and angular velocity of the $n$th rod. The indices may also mark the type of particle in a many-component mixture. Another important example are particles, such as polyatomic molecules, which are made of subparticles (atoms) arranged in certain spatial conformations [35]. See Table I for examples of correlations between such DOFs.

To keep the analysis as general as possible, we define the field

$$\underline{\phi}(\mathbf{r}) = \sum_{n=1}^{N} \delta(\mathbf{r} - \mathbf{r}_n)\underline{f}(\mathbf{j}_n), \tag{2a}$$

where $\underline{f}$ is any function (for simplicity, in the form of a column vector) of the additional DOFs $\mathbf{j}_n$. For example, for the particle-density field $\underline{f} = 1$, for the orientation vector field

$f = \hat{\mathbf{n}}$, and for the nematic tensor field [24], $\mathbf{Q} = (d\hat{\mathbf{n}}\hat{\mathbf{n}}^T - \bar{\mathbf{I}})/(d-1)$, $f$ consists of the independent components of $\mathbf{Q}$. We denote the total number of components within $f$ and $\phi(\mathbf{r})$ by $M$ [i.e., $f, \phi(\mathbf{r}) \in \mathbb{C}^M$].

We now construct the two-point CF associated with the field of Eq. (2a). First, for convenience, the field is Fourier-transformed into

$$\phi(\mathbf{q}) = \sum_{n=1}^{N} e^{-i\mathbf{q}\cdot\mathbf{r}_n} f(\mathbf{j}_n), \tag{2b}$$

where the uniform ($\mathbf{q} = \mathbf{0}$) mode is excluded [36]. The CF is a $M \times M$ matrix, defined as

$$\mathcal{C}(\mathbf{q}) = \mathcal{A}^{-1}\langle N^{-1}\phi(\mathbf{q})\phi^{\dagger}(\mathbf{q})\rangle, \tag{3a}$$

where $\langle\cdot\rangle$ denotes an ensemble average. Note that we include $N$ within the brackets in case the number of particles is not fixed. The normalization matrix $\mathcal{A}$ ensures that, for uncorrelated fields (ideal gas), $\mathcal{C}(\mathbf{q}) = \mathcal{I}$ for all $\mathbf{q}$, where $\mathcal{I}$ is the $M \times M$ unit matrix. Explicitly, $\mathcal{A}$ is the second-moment matrix of $f(\mathbf{j})$, given by

$$\mathcal{A} = \langle f(\mathbf{j})f^{\dagger}(\mathbf{j})\rangle = \int d\mathbf{j}\,\Pr(\mathbf{j})f(\mathbf{j})f^{\dagger}(\mathbf{j}), \tag{3b}$$

where $\Pr(\mathbf{j})$ is the single-particle marginal distribution of the DOFs $\mathbf{j}$. Note that the full distribution is not required (only the second moment of $f$ is) and that $\mathbf{j}$ may have either annealed or quenched statistics. The definitions above hold in more general settings in which $f$ depends explicitly on $\mathbf{q}$. In fact, the only condition on $f$ is that $\mathcal{A}$ of Eq. (3b) is invertible.

Since the CFs are measurable, e.g., by scattering or simulations, the $\mathbf{q}$ modes are naturally discretized by the system size. While their number is infinite in principle, only a finite number of $\mathbf{q}$ modes, $\Omega = \sum_{\mathbf{q}\neq\mathbf{0}} 1$, are measured in practice. Constraining such a subset of modes will merely loosen the entropy bound. The limit of continuous $\mathbf{q}$ will be addressed below as well.

Let us clarify again the notation used above. The bold symbols are physical vectors, tensors, etc. (e.g., $\mathbf{r}$, $\mathbf{Q}$). The underlined objects [$f$ and $\phi(\mathbf{q})$] are column vectors in $\mathbb{C}^M$. Lastly, the calligraphic objects [$\mathcal{C}(\mathbf{q})$, $\mathcal{A}$, and $\mathcal{I}$] are $M \times M$ matrices. See again Table I for a few specific examples.

Now that the mathematical structure of the constraints has been defined, we give our central result—an upper bound on the thermodynamic entropy $S$,

$$\frac{S - S^{\mathrm{id}}}{\langle N\rangle} \leqslant h_{\mathrm{ex}}. \tag{4}$$

Here $\langle N\rangle$ is the mean number of particles, $S^{\mathrm{id}}$ is the entropy of the ideal gas, and $h_{\mathrm{ex}}$ is the excess entropy (per particle) of the measured two-point CF, $\mathcal{C}(\mathbf{q})$ (its information content), which we find to be approximately

$$h_{\mathrm{ex}}[\mathcal{C}] = \frac{1}{2\langle N\rangle}\sum_{\mathbf{q}\neq\mathbf{0}} \mathrm{tr}\{\ln[\mathcal{C}(\mathbf{q})] + \mathcal{I} - \mathcal{C}(\mathbf{q})\}. \tag{5}$$

The matrix logarithm appearing in Eq. (5) is handled according to $\mathrm{tr}[\ln(\cdot)] = \ln[\det(\cdot)]$. This expression is also applicable to cases where the number of particles is not fixed. The inequality in Eq. (4) comes from the fact that the CF may

not constitute the entire information encoded in the system's state (as in systems with many-body interactions or out of equilibrium). The bound is approximate due to a leading-order (Gaussian) approximation which has been employed (see Appendix B).

The transition to continuous $\mathbf{q}$ is done through $\sum_{\mathbf{q}}(\cdot) = \nu\int d\mathbf{q}(\cdot)$, where $\nu = V/(2\pi)^d$ is the density of $\mathbf{q}$ modes, and $V = \int d\mathbf{r}$ is the system's volume. The continuous version of Eq. (5) is then

$$h_{\mathrm{ex}}[\mathcal{C}] = \frac{1}{2(2\pi)^d\bar{\rho}}\int_{\mathbf{q}\neq\mathbf{0}} d\mathbf{q}\,\mathrm{tr}\{\ln[\mathcal{C}(\mathbf{q})] + \mathcal{I} - \mathcal{C}(\mathbf{q})\}, \tag{6}$$

where $\bar{\rho} = \langle N\rangle/V$ is the mean density [37].

The detailed derivation of Eq. (5) is given in Appendix B. Although the setup above may seem complicated, it reflects the wide applicability of the approach. In practice, once the commonly used CFs (e.g., structure factor, orientations or velocity CFs; see Table I) are found, they can simply be inserted into Eq. (5) to obtain useful entropy bounds. This will become clearer below when we apply Eq. (5) to a specific system.

Equation (5) can be extended to cases where an external field is present. For example, with $f = 1$ (density field), the steady-state density of particles may be nonuniform (e.g., due to gravity). We can include such a nonuniform steady profile, $\underline{G}(\mathbf{q}) = \langle N^{-1/2}\phi(\mathbf{q})\rangle$, as another measurable constraint. In addition, due to the breaking of translational symmetry, the CF defined in Eq. (3a) would cease to be $\mathbf{q}$ diagonal and become $\mathcal{C}(\mathbf{q}, \mathbf{q}') = \mathcal{A}^{-1}\langle N^{-1}\phi(\mathbf{q})\phi^{\dagger}(\mathbf{q}')\rangle$. These modifications are treated in Appendix A. The resulting bound, extending Eq. (5), is given in Eq. (A3).

## III. APPLICATION TO BIDISPERSE MIXTURES

Dense random packings of particles have been a central issue in statistical physics for years [38]. Recently, disordered random packings close to the jamming point, referred to as maximally random-jammed, have been found to exhibit hyperuniformity [33,39–41]—strong suppression of density fluctuations over large distances [42]. In two dimensions, densely packed monodisperse disks tend to locally arrange into hexagonal crystals. However, defects are usually present in the crystal structure, randomly distributed over space. A small degree of polydispersity limits the size of the crystalline domains and reduces the presence of rattlers. Hence, a densely packed mixture of bidisperse disks is expected to transition, as a function of the number fraction (NF) and size ratio (SR) of the two species, between mostly crystalline structures and disordered structures.

Experiments on jammed quasi-2D emulsions of bidisperse droplets, confined and flattened between two surfaces in a microfluidic channel, have demonstrated this optimization [33]. They showed, moreover, that the increase in disorder came together with stronger hyperuniformity. In this section we revisit the data from those experiments in light of the present formalism, to obtain the entropy contributions of the droplets' different DOFs (position and size), and examine how these contributions change along the transition.

First, we briefly describe the experimental setup; an extended description of the experimental devices and conditions can be found in Ref. [33]. Standard soft photolithography and
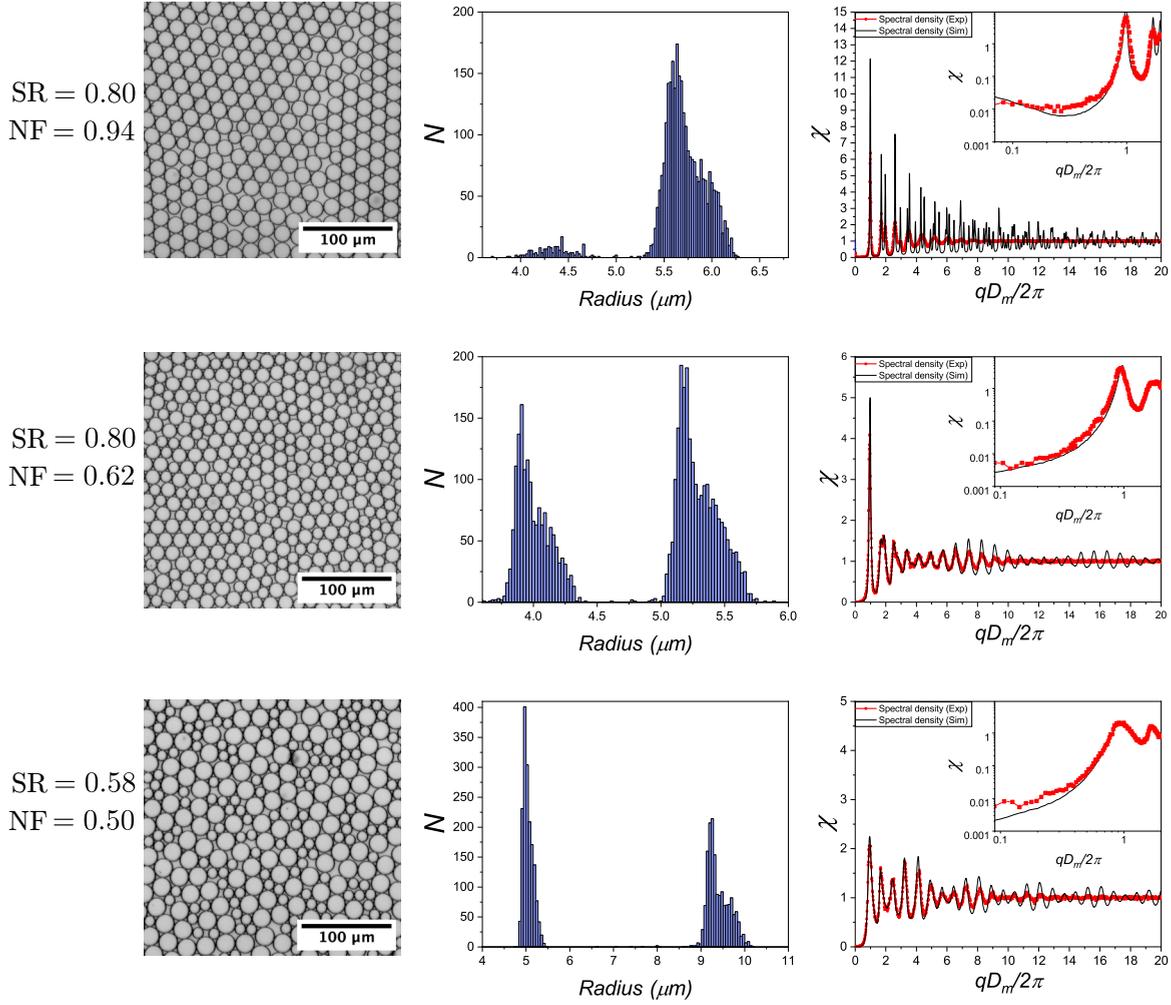
FIG. 1. Bidisperse mixtures. Left: Experimental closeup snapshots of the quasi-2D emulsion for several size ratios SR and number fractions NF, as indicated. The values of SR and NF are changed using different pressures at the inlet of the microfluidic chip. Center: Distribution of droplet radii, demonstrating approximate bidispersity. Right: Spectral densities as obtained from experiment (red connected dots) and simulation (black line). The results from experiment and simulation match qualitatively, the oscillations in the experiment being more damped. Insets: Closeup on the spectral densities at small $q$, indicating the presence or absence of hyperuniformity. Data adapted from Ref. [33].

replica-molding techniques are used to create the microfluidic device [43]. The pressures applied at the inlets of two T junctions allow a precise control over the droplet sizes and their production rates. Downstream, a microfluidic mixer randomizes the entrance of the droplets into the collection chamber. The chamber height (10 μm) forces the droplets to adopt a pancake-like shape, and they assemble into 2D configurations. Their deformability and the pressure ensure that the assembly is close to a jammed configuration. Images are taken inside the observation chamber every 30 seconds in order for all the droplets to be renewed. Thus, the snapshots depict uncorrelated configurations. A total of 30 such images are taken for every set of parameters. For each set we verify that (a) the two populations of droplets are well mixed, (b) the droplets maintain a circular shape, and (c) the Fourier transform of each image has a circular symmetry.

Here we focus on two sets of data. In the first, the size ratio is kept fixed at SR $\equiv R_{\text{small}}/R_{\text{big}} \simeq 0.8$ (0.7–0.82), while the number fraction, NF $\equiv N_{\text{big}}/(N_{\text{big}} + N_{\text{small}})$, is varied. In

the second set, the number fraction is kept fixed at NF $\simeq$ 0.5 (0.4–0.6), while the size ratio is varied. For comparison, we also study by simulation the corresponding 2D bidisperse jammed configurations of hard disks. The simulations apply the freely available code based on the Lubachevsky-Stillinger algorithm [44,45].

Figure 1 shows three typical bidisperse mixtures. The top panels correspond to SR = 0.80 and NF = 0.94, such that most droplets are big. For this high NF, the droplets form a polycrystalline structure. The crystalline domains are clearly visible, outlined by line defects which host the small droplets. Some small droplets can also be found within the crystalline lattice. The size distribution shows two peaks corresponding to the two populations of droplets. The standard deviation for each population is smaller than 0.05, justifying the bidisperse approximation. The figure also shows the spectral density,

$$\chi(\mathbf{q}) = \frac{1}{a^2}\left\langle\frac{1}{N}\sum_{n,m}e^{-i\mathbf{q}\cdot\mathbf{r}_{nm}}a_n a_m\right\rangle, \qquad (7)$$
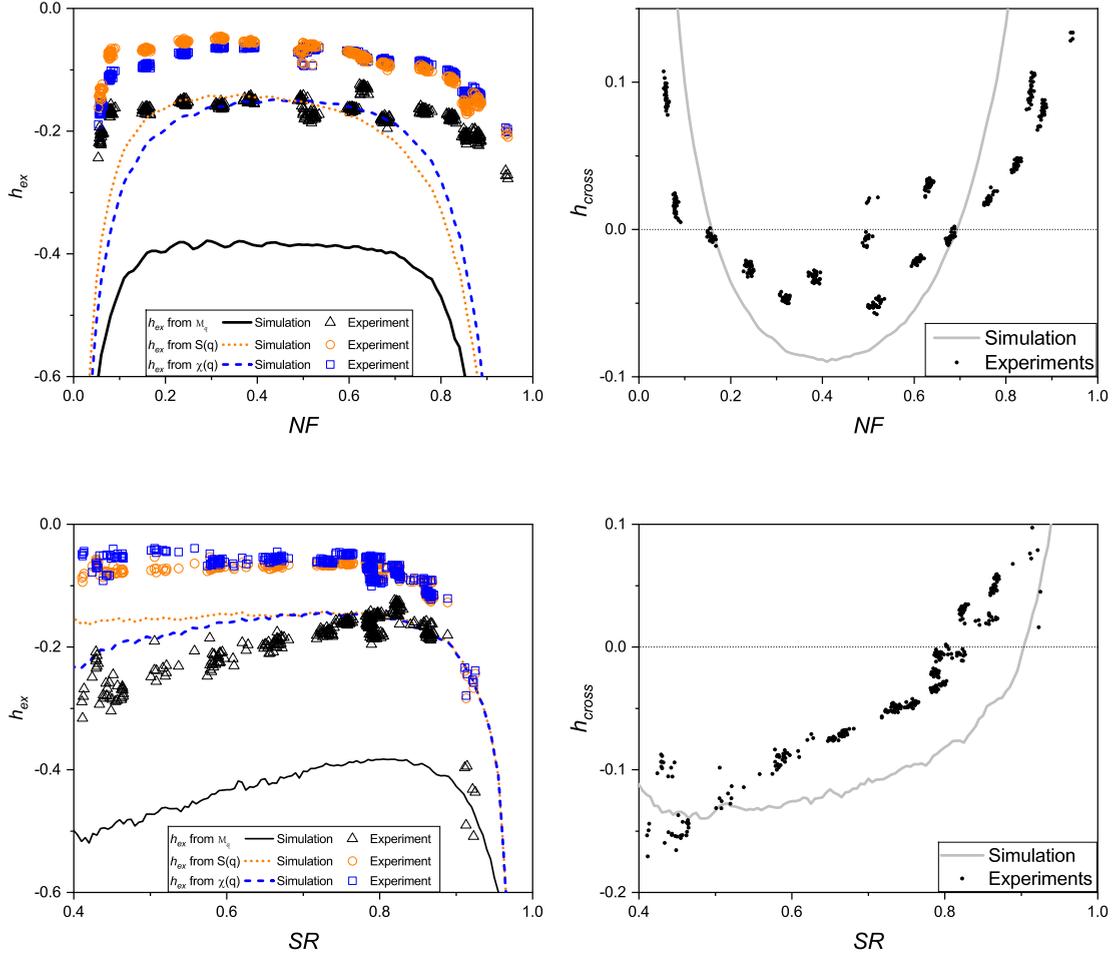
FIG. 2. Bidisperse mixtures. Entropy bounds as a function of number fraction NF (top panels, for SR $\simeq 0.8$) and size ratio SR (bottom panels, for NF $\simeq 0.5$). Left: Excess entropy per particle, $h_{ex}$, as obtained from the structure factor $S_{\mathbf{q}}$ (experiment in orange $\bigcirc$ and simulation in dotted orange line), spectral density $\chi_{\mathbf{q}}$ (blue $\square$ and dashed blue line), and mixed correlation function $\mathcal{M}_{\mathbf{q}}$ (black $\triangle$ and black solid line). Right: Net entropy contribution of position-area cross-correlations, $h_{\mathrm{cross}}$ (experiment in dots, simulations in solid line). The crossover from positive to negative values relates to the crossover between polycrystalline and disordered hyperuniform states.

where $a_n$ is the area of droplet $n$, and $\overline{a^2}$ is the second moment of the droplet area distribution. In the terminology of Sec. II, $\chi(\mathbf{q})$ is the CF with $f(a) = a$; see Table I. The spectral density shows damped Bragg peaks which are characteristic of a 2D hexagonal polycrystal. The first peak corresponds to the most common interdroplet distance, denoted $D_m$.

The middle panels of Fig. 1 show the results for SR = 0.8 and NF = 0.62. This mixture is disordered and does not show extended crystal-like domains. The size distribution shows once again approximate bidispersity. The spectral density is more damped than for NF = 0.94 and does not exhibit the typical Bragg peaks of hexagonal lattices. At small $q$ it tends (linearly) toward zero, which is a signature of hyperuniformity [33,38]. Qualitatively similar behavior is seen for SR = 0.58 and NF = 0.5 in the bottom panels of Fig. 1.

The spectral density obtained in the experiment is more damped than in the simulations. This is due to residual polydispersity and small deformations present in the experimental mixture of droplets, which randomize the correlations for large $q$, whereas the simulations are carried out for strictly

bidisperse, nondeformable hard disks. The stronger disorder in the experimental system will translate below into a positive offset in the excess entropy.

The microscopic DOFs in the experiments and simulations are the locations of droplet centers, $\{\mathbf{r}_n\} = \{x_n, y_n\}$, and their areas, $\{a_n\}$. The two-point CFs that we consider are the following. (a) The structure factor, $S(\mathbf{q})$, defined with $f(a) = 1$ in Eqs. (2a). (b) The spectral density, $\chi(\mathbf{q})$, with $\overline{f(a)} = a$, which takes into account also the areas of the droplets [18,46]; see Eq. (7). (c) We consider also the mixed CF, $\mathcal{M}(\mathbf{q})$, defined with $\underline{f(a)} = (1, a)$, imposing $S(\mathbf{q})$ and $\chi(\mathbf{q})$ simultaneously along with the position-area cross-correlations. See the SM [34] for a detailed description of how these CFs are constructed and the way the excess entropies are calculated.

The excess-entropy bounds $h_{ex}$, obtained from the measured CFs using Eq. (5), are shown in the left panels of Fig. 2. Data points and lines correspond to results from experiments and simulations, respectively. To demonstrate the robustness of the method, we have not ensemble-averaged the CFs; hence the many data points for each parameters set, showing a small

spread. Note that by using $h_{ex}$ (the excess entropy per particle relative to the uncorrelated system), the trivial dependencies of the entropy on the number of particles and field of view are removed. Note also that the field of view may contain different numbers of droplets (see, e.g., the snapshots in Fig. 1), which highlights the ability of the formalism to treat open systems.

All three bounds in the left panels of Fig. 2 show the expected simple trend—as the system gets away from the ordered crystalline state (i.e., away from NF = 0, 1 in the top panel, and SR = 1 in the bottom one), the entropy increases. In the most ordered state, most droplets have six neighboring droplets, with a sixfold symmetry. Thus positions and cluster orientations are strongly correlated. When SR decreases or NF gets further away from 0 or 1, more possibilities open up for the immediate environment of a droplet. Therefore, the local orientation and the number of neighbors become less correlated, which increases the excess entropy. The results from experiments and simulations agree qualitatively. As explained above, additional disorder in the experimental system leads to a positive entropy difference with respect to the simulation. While the distinction between the two phases is seen clearly also in the CF (the spectral density) of Fig. 1, it is less evident for close values of NFs and SRs (Figs. S1 and S2 [34]). The entropy $h_{ex}$, however, and especially $h_{cross}$, sharply change in the crossover regions, indicating much more clearly the trends described above. For instance, for SR > 0.76 (where a droplet may have six neighbors; see Figs. S1 and S2), while the spectral density does not seem to change much (the Bragg peaks still persist), $h_{cross}$ shows a significant drop, suggesting that the structural information does exist in the CF, but is subtly buried.

The entropy bounds nicely capture the asymmetry on the two sides of the symmetric composition NF = 0.5, i.e., NF and $(1 - NF)$ are not statistically equivalent (top-left panel of Fig. 2). For example, the measured entropy for NF = 0.2 is higher than the one for NF = 0.8. Indeed, there are more possibilities to disperse small particles in a polycrystalline arrangement of big particles than the other way around. The asymmetry is most prominent for the $S(\mathbf{q})$ bound, which is "blind" to the difference in particle sizes and reflects the difference in the number of positional arrangements as if the difference originated from some underlying interaction. This asymmetric behavior on the two sides of NF = 0.5 is a feature concealed in subtle differences between the CFs for different NFs (see SM [34]). In contrast, the asymmetry is clearly revealed by the entropy functional.

The mixed CF $\mathcal{M}(\mathbf{q})$ gives the lowest bound in both left panels. It constrains the structure factor and the spectral density simultaneously (along with cross-correlations), thus containing more information than each of the other two CFs. The bounds associated with $S(\mathbf{q})$ and $\chi(\mathbf{q})$ cross around NF $\simeq$ 0.5 in both experiment and simulation (top-left panel). This is because the spectral density is a "reweighting" of the structure factor in favor of the big particles [see Eq. (7)]. For small NF (majority of small droplets) this enhances the correlations as characterized by $\chi(\mathbf{q})$, leading to a lower entropy, and the opposite happens for large NF (majority of large droplets). Comparing the bounds from $S(\mathbf{q})$ and $\chi(\mathbf{q})$ in the bottom-left panel, we find that the $S(\mathbf{q})$ one is mostly higher in the simula-

tion but slightly lower (better) in the experiment. This too may be a result of the residual polydispersity in the experiment, which weakens the correlations as measured by $\chi(\mathbf{q})$.

The right panels of Fig. 2 show the net entropy contribution of the position-area cross-correlations, defined as $h_{cross} \equiv h_{ex}[\mathcal{M}(\mathbf{q})] - h_{ex}[S(\mathbf{q})] - h_{ex}[\chi(\mathbf{q})]$. In the absence of cross-correlations, it is easy to see from Eq. (5) that $h_{ex}[\mathcal{M}(\mathbf{q})]$ is equal to $h_{ex}[S(\mathbf{q})] + h_{ex}[\chi(\mathbf{q})]$, leading to $h_{cross} = 0$. Thus $h_{cross}$ quantifies the additional information stored in the cross-correlations. However, $h_{cross}$ as defined above may be negative or positive [47]. This is because $S(\mathbf{q})$ and $\chi(\mathbf{q})$ are CFs of DOFs which are not independent (both CFs include position correlations). For example, in the limit of a monodisperse system, $\chi(\mathbf{q})$ and $S(\mathbf{q})$ coincide, and we have $h_{ex}[\mathcal{M}(\mathbf{q})] = h_{ex}[S(\mathbf{q})] = h_{ex}[\chi(\mathbf{q})]$, leading to $h_{cross} = -h_{ex}[S(\mathbf{q})] > 0$. See Fig. 2 at small and large NF, and large SR. Position-area cross-correlations become strong when the spatial arrangement of nearby droplets is affected by their sizes. For instance, a droplet can be surrounded by six others when SR is close to 1, and by seven smaller ones if SR < 0.76. This accounts for the sharp decrease of $h_{cross}$ with decreasing SR (bottom-right panel).

Overall, the key observation in both right panels of Fig. 2 is the dominant contribution of the cross-correlations between position and area to the entropy of the heterogeneous, disordered and hyperuniform system.

## IV. DISCUSSION

Any measurement provides information on the system of interest and thus lowers the upper bound on its entropy. Equations (5), (6), and (A3) quantify this entropy reduction quite generally when the measurement is of a pair-correlation function. The extent to which these results are useful has been examined in Sec. III, where we have used the entropy bounds to investigate subtle features of a transition out of equilibrium.

The central advantage of the approach presented here is the ability to not only estimate the entropy but also separate it into distinct contributions coming from different DOFs and their intercorrelations. This "finer resolution" intimately connects the method to the physics of the system under study. When applied to jammed mixtures (Sec. III), the entropy bounds have revealed the crucial role played by cross-correlations between the particles' positions and sizes. Previous works associated the formation of disordered hyperuniform structures solely with strong positional correlations [42]. Our results suggest that, in polydisperse systems, the entropy reduction due to hyperuniformity is dominated by strong position-size cross-correlations.

Another key feature of the formalism is that, by choosing measurable, concrete constraints (i.e., CFs), we essentially bypass the undersampling problem of entropy estimation, integrating out the unconstrained information. Importantly, any symmetry (or symmetry breaking) reflected in the CFs is by construction taken into account. The cost of these advantages is that the method is not a "black box" which can be blindly applied to any data, but requires knowledge of the system's DOFs.

A few important technical advantages should be noted as well. (a) The computational cost of the method does not scale

with system size. (b) There is no artificial discretization of phase space. (c) Generalization to include other observables is achieved by including the corresponding constraints. This flexibility, for example, has allowed us to include the effects of varying particle number and external fields (see Appendix A).

The formalism is restricted in two main respects. (a) It is limited to upper bounds for the entropy, which might not always be sufficiently tight to be useful. In such cases one may improve the bound, e.g., by including higher-order correlations (see, for example, the useful high-order correlations measured in Ref. [48].) (b) The main results, Eqs. (5), (6), and (A3), involve a Gaussian approximation. One may improve them using standard perturbation techniques. Indeed, some constraints may give rise to equations which cannot be solved analytically, which will require numerical solutions [32].

The possibilities to apply entropy-bound functionals to additional systems are vast. We mention two examples which seem particularly appealing. Indications of entropy changes occurring at the glass transition [49,50] may be checked in more detail, similar to the case of the jammed mixtures in Sec. III. If such changes are found, one might be able to point at the dominant DOFs, or their interactions, underlying the transition. This particular application is likely to require the inclusion of higher-order correlations. Active matter exhibits dynamic transitions, such as bacterial swarming [51] and motility-induced phase separation [52]. Resolving different contributions to the entropy may provide new insights into these far-from-equilibrium phenomena.

Finally, it may be possible, and highly desirable, to go beyond steady states and use a similar approach for time-dependent phenomena. Thus, for example, the contributions of different DOFs to the entropy production might be resolved by exploiting temporal CFs as dynamical constraints.

## ACKNOWLEDGMENTS

## APPENDIX A: ENTROPY BOUND FOR INHOMOGENEOUS SYSTEMS

The purpose of this Appendix is to extend the formalism described in Sec. II to systems whose translational symmetry is broken by some externally applied force. The adjusted bound is given in Eq. (A3) below. In addition, this Appendix will facilitate the detailed derivation of Eq. (5) in Appendix B.

As before, the key observable is a two-point CF. We rely on the setup presented in Sec. II. First, we define the (Fourier transformed) field of interest following Eqs. (2a). Now, due to the inhomogeneity, two distinct **q** modes might have nonzero correlations. Hence, the CF of Eq. (3a) is extended to include off-diagonal terms in **q**,

$$\mathcal{C}(\mathbf{q}, \mathbf{q}') = \mathcal{A}^{-1} \langle N^{-1} \underline{\phi}(\mathbf{q}) \underline{\phi}^{\dagger}(\mathbf{q}') \rangle, \tag{A1a}$$

where $\mathcal{A}$ is the same normalization matrix defined in Eq. (3b).

In addition to the more complex CF, the average field $\langle \underline{\phi}(\mathbf{q}) \rangle$ may be nonuniform. For example, the steady-state density of particles (using $\underline{f} = 1$) may be inhomogeneous due to gravity. We characterize these effects by introducing the average profile,

$$\underline{G}(\mathbf{q}) = \langle N^{-1/2} \underline{\phi}(\mathbf{q}) \rangle, \tag{A1b}$$

as another constraint. The CFs and profiles may be available from experiments (e.g., scattering and absorption, respectively) and simulations. As before, we exclude the uniform $\mathbf{q} = \mathbf{0}$ mode.

A few additional definitions are required. We recall that only a finite set $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_\Omega\}$ of $\mathbf{q}$ modes is considered. It is useful to consider $\mathbf{q}$ as an additional index in an extended matrix that treats the $\Omega$ modes $\mathbf{q}$ and the $M$ components of $\underline{f}$ together. We define the $(\Omega M) \times (\Omega M)$ matrices $\hat{\mathcal{C}}, \hat{\mathcal{A}}$, and the unit matrix $\hat{\mathcal{I}}$, and similarly the $(\Omega M)$-long vector $\underline{\hat{G}}$. These are just $\Omega \times \Omega$ blocks of the $M \times M$ matrices $\mathcal{C}(\mathbf{q}, \mathbf{q}'), \mathcal{I}$, and $\mathcal{A}$, and $\Omega$-times stacked $M$-long vector $\underline{G}(\mathbf{q})$,

$$\hat{\mathcal{C}} = \begin{pmatrix} \mathcal{C}(\mathbf{q}_1, \mathbf{q}_1) & \mathcal{C}(\mathbf{q}_1, \mathbf{q}_2) & \cdots & \mathcal{C}(\mathbf{q}_1, \mathbf{q}_\Omega) \\ \mathcal{C}(\mathbf{q}_2, \mathbf{q}_1) & \mathcal{C}(\mathbf{q}_2, \mathbf{q}_2) & \cdots & \mathcal{C}(\mathbf{q}_2, \mathbf{q}_\Omega) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}(\mathbf{q}_\Omega, \mathbf{q}_1) & \mathcal{C}(\mathbf{q}_\Omega, \mathbf{q}_2) & \cdots & \mathcal{C}(\mathbf{q}_\Omega, \mathbf{q}_\Omega) \end{pmatrix},$$

$$\hat{\mathcal{A}}^{-1} = \begin{pmatrix} \mathcal{A}^{-1} & \mathcal{O} & \cdots & \mathcal{O} \\ \mathcal{O} & \mathcal{A}^{-1} & \cdots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \cdots & \mathcal{A}^{-1} \end{pmatrix},$$

$$\underline{\hat{G}} = \left( \underline{G}^T(\mathbf{q}_1) \mid \underline{G}^T(\mathbf{q}_2) \mid \cdots \mid \underline{G}^T(\mathbf{q}_\Omega) \right)^T. \tag{A2}$$

Now that the constraints have been defined and rearranged, the entropy bound obtained for a two-point CF $\mathcal{C}(\mathbf{q}, \mathbf{q}')$ and a steady profile $\underline{G}(\mathbf{q})$ can be written as

$$h_{\text{ex}}[\mathcal{C}, \underline{G}] = \frac{1}{2\langle N \rangle} \text{tr}_{\Omega M}[\ln_{\Omega M}(\hat{\mathcal{C}} - \hat{\mathcal{A}}^{-1} \underline{\hat{G}} \underline{\hat{G}}^{\dagger}) + \hat{\mathcal{I}} - \hat{\mathcal{C}}]. \tag{A3}$$

The detailed derivation is given in the SM [34]. The matrix logarithm and trace are applied to the objects of Eq. (A2), and the subscripts $\Omega M$ are a reminder that they are trace and matrix-log operations performed on $(\Omega M) \times (\Omega M)$ matrices. The number of particles is allowed to vary, with a given average number $\langle N \rangle$, and with a CF $\mathcal{C}(\mathbf{q}, \mathbf{q}')$ and steady profile $\underline{G}(\mathbf{q})$ which take into account the varying $N$ through Eqs. (A1).

We may interpret Eq. (A3) by considering two opposite limits. In systems of single-particle interactions (i.e., in the presence of a strong external field), the two-point CF becomes unity, $\mathcal{C}(\mathbf{q}, \mathbf{q}') = \mathcal{I}\delta(\mathbf{q} - \mathbf{q}')$ (i.e., $\hat{\mathcal{C}} = \hat{\mathcal{I}}$), and the bound reduces to

$$h_{\text{ex}}[\underline{G}] = \frac{1}{2\langle N \rangle} \text{tr}[\ln(\hat{\mathcal{I}} - \hat{\mathcal{A}}^{-1} \underline{\hat{G}} \underline{\hat{G}}^{\dagger})]$$

$$\simeq -\frac{1}{2\langle N \rangle} \sum_{\mathbf{q} \neq \mathbf{0}} \underline{G}^{\dagger}(\mathbf{q}) \mathcal{A}^{-1} \underline{G}(\mathbf{q}). \tag{A4}$$

This agrees to second order (due to the Gaussian approximation in the derivation) with the trivial result $h_{\text{ex}} =$

$-\int d\mathbf{r}\,d\mathbf{j}\,\Pr(\mathbf{r},\mathbf{j})\ln[\upsilon\,\Pr(\mathbf{r},\mathbf{j})]$, where $\Pr(\mathbf{r},\mathbf{j})$ is the joint single-particle distribution of $\mathbf{r}$ and $\mathbf{j}$, and $\upsilon = \int d\mathbf{r}\,d\mathbf{j}$. In the other limit of homogeneous systems, the steady profile is absent, $\underline{G} = \underline{0}$, and the CF is $\mathbf{q}$-diagonal [as in Eq. (3a)]. The expression for the entropy bound reduces to Eq. (5),

$$h_{\text{ex}}[\mathcal{C}] = \frac{1}{2\langle N\rangle}\sum_{\mathbf{q}\neq\mathbf{0}}\text{tr}[\ln\mathcal{C}(\mathbf{q},\mathbf{q})+\mathcal{I}-\mathcal{C}(\mathbf{q},\mathbf{q})].$$

Here the matrix logarithm and trace are of $M \times M$ matrices only, and the summation over $\mathbf{q}$ is explicitly written. Comparing to Eq. (A3), we see that, in the absence of a steady profile ($\underline{G} = \underline{0}$), the only component in the field's "variance," $\hat{\mathcal{C}} - \hat{\mathcal{A}}^{-1}\hat{\underline{G}}\hat{\underline{G}}^\dagger$, that survives is the second-moment matrix (its $\mathbf{q}$-diagonal terms), $\mathcal{C}(\mathbf{q},\mathbf{q})$. Together, Eqs. (A4) and (5) demonstrate how Eq. (A3) could be separated into single-particle and two-particle contributions. Also, they do not require the more cumbersome and computationally costly objects of Eq. (A2).

## APPENDIX B: DERIVATION OF THE CENTRAL RESULT

Here we give a detailed derivation of the main result—an upper bound for the entropy given a general two-point CF. A steady-state average field and a grad-canonical ensemble are considered in the SM [34].

Following the main text, we deal with systems whose particles possess more DOFs than just locations. A complete microscopic configuration is therefore described by $\{\mathbf{x}_n\} = \{\mathbf{r}_n, \mathbf{j}_n\}$, where $\mathbf{r}_n$ is the spatial position of the $n$th particle and the vector $\mathbf{j}_n$ lists the scalar components of its additional DOFs.

We consider an arbitrary vectorial function of the DOF $\mathbf{j}$, to obtain a field as written in Eq. (2a), $\underline{\phi}(\mathbf{r}) = \sum_n \delta(\mathbf{r} - \mathbf{r}_n)\underline{f}(\mathbf{j}_n)$, where $\underline{f} \in \mathbb{C}^M$ is a column vector. A simple example is the case of 3D orientations, $\mathbf{j} = (\vartheta, \varphi)$, where the function is the decomposition into spherical harmonics, arranged in a column vector $\underline{f}_{l^2+(m+l)+1}(\mathbf{j}) = Y_l^m(\vartheta, \varphi)$. Next, a Fourier transform is taken to obtain Eq. (2b), $\underline{\phi}(\mathbf{q}) = \sum_n e^{-i\mathbf{q}\cdot\mathbf{r}_n}\underline{f}(\mathbf{j}_n)$. For shorthand, we replace $e^{-i\mathbf{q}\cdot\mathbf{r}_n}\underline{f}(\mathbf{j}_n)$ with a simple $\hat{\underline{g}}(\mathbf{x}_n)$ [the discrete $\mathbf{q}$ absorbed as an index, as in the objects of Eq. (A2)], such that the transformed field is

$$\hat{\underline{\phi}} = \sum_n \hat{\underline{g}}(\mathbf{x}_n).$$ (B1)

Proceeding with this notation, where the position DOF is not written explicitly, highlights the fact that the formalism [and the bound appearing in Eq. (A3)] applies for nonspatial CFs as well.

The two constraints to be imposed are the following. (a) The CF [Eq. (A1a)],

$$\hat{\mathcal{C}} = \langle\hat{\mathcal{B}}^{-1}\hat{\underline{\phi}}\,\hat{\underline{\phi}}^\dagger\rangle,$$ (B2a)

where $\hat{\mathcal{B}} \equiv N\hat{\mathcal{A}} = N\int d\mathbf{x}\,\Pr(\mathbf{x})\hat{\underline{g}}(\mathbf{x})\hat{\underline{g}}^\dagger(\mathbf{x})$ is a the normalization matrix, and $\Pr(\mathbf{x})$ the single-particle distribution. (b) If the system is inhomogeneous, we additionally constrain the steady-state profile [Eq. (A1b)],

$$\hat{\underline{G}} = \langle N^{-1/2}\hat{\underline{\phi}}\rangle.$$ (B2b)

Recall that we exclude the uniform mode (for the spatial field, the $\mathbf{q} = \mathbf{0}$ mode); hence, $\underline{g}$ is not a complete basis (as one component is missing). This will come into play in the calculation performed below (at the saddle-point approximation step).

As mentioned in the main text, experimentally, a finite number ($\Omega$) of $\mathbf{q}$-modes are considered, which only loosens the entropy bound. Similarly, including only a finite number ($M$) of the transform's components, $\underline{f}$, would merely loosen the bound, too. In the examples of 3D orientations, taking $\underline{f} = \hat{\mathbf{n}}$ implies the inclusion of only the $M = 3$ terms $Y_1^{0,\pm1}$, and taking $\underline{f}$ as the independent components of the nematic tensor implies the inclusion of the $M = 6$ terms $Y_0^0, Y_2^{0,\pm1,\pm2}$. Thus, as established in Appendix A, $\hat{\underline{G}}, \hat{\underline{\phi}}, \hat{\underline{g}}(\mathbf{x}_n) \in \mathbb{C}^{M\Omega}$, and $\hat{\mathcal{C}}, \hat{\mathcal{B}}$, and $\hat{\mathcal{I}}$ are $M\Omega \times M\Omega$ matrices. While $M\Omega \equiv \Xi$ is finite in practice, for the sake of the derivation we assume that $\hat{\underline{g}}(\mathbf{x})$ (up to adding the uniform mode) is a complete basis for $\mathbf{x}$. (This subtlety is required because later in the derivation we rely on the existence of the inverse transform.) We will drop the hats of the $\Xi$-sized column vectors and $\Xi \times \Xi$ matrices; the "hatless" $M$-sized column vectors and $M \times M$ matrices (shown in Sec. II and Appendix A) do not appear here. The connection to actual physical observables is more evident from the notation adopted in the main text. The more elaborate notation used here and in the SM [34] makes the detailed derivation more concise.

We maximize the entropy [Eqs. (1)] over the probability density function $p(\mathbf{X})$ of microstates $\mathbf{X} = \{\mathbf{x}_n\}$, under two constraints: (a) normalization, $\int \overline{d\mathbf{X}}p(\mathbf{X}) = 1$, and (b) the key constraint—a given CF [Eq. (B2a)],

$$\mathcal{C} = \mathcal{B}^{-1}\int \overline{d\mathbf{X}}p(\mathbf{X})\sum_{n=1}^{N}\sum_{m=1}^{N}\underline{g}(\mathbf{x}_n)\underline{g}^\dagger(\mathbf{x}_m).$$ (B3)

We denote $\overline{d\mathbf{X}} = d\mathbf{X}\prod_{n=1}^{N}\Pr(\mathbf{x}_n)$ to allow the case of quenched DOFs, whereby the single-particle distribution $\Pr(\mathbf{x})$ serves as a "weight function." Similarly, $\overline{d\mathbf{x}} = d\mathbf{x}\,\Pr(\mathbf{x})$. In this case $p(\mathbf{X})$ may be interpreted either as the $N$-particle distribution function [18], or, up to a change of variables [$\mathbf{x} \to \text{CDF}(\mathbf{x})$, the cumulative distribution function], the copula density [53]. Therefore,

$p(\mathbf{X})$ is related to the probability density function to find the system in microstate $\mathbf{X}$, $P(\mathbf{X})$ [the one appearing in Eq. (1b)], through $P(\mathbf{X}) = p(\mathbf{X}) \prod_{n=1}^{N} \Pr(\mathbf{x}_n)$. In the simpler case of annealed DOFs, $\Pr(\mathbf{x}) = 1/\upsilon$, where $\upsilon = \int d\mathbf{x}$ is the phase-space volume.

The thermodynamic entropy, following Eqs. (1), is

$$S = -\int d\mathbf{X} \prod_{n=1}^{N} \Pr(\mathbf{x}_n) p(\mathbf{X}) \ln\left[\prod_{n=1}^{N} \Pr(\mathbf{x}_n) p(\mathbf{X})\right] = -\int d\mathbf{X} \prod_{n=1}^{N} \Pr(\mathbf{x}_n) p(\mathbf{X}) \ln p(\mathbf{X}) + N s_1, \tag{B4}$$

where $s_1 = -\int d\mathbf{x} \Pr(\mathbf{x}) \ln \Pr(\mathbf{x})$ is the entropy of a single isolated particle (an *a priori* known constant). In the absence of interparticle coupling, $N s_1$ is the only surviving contribution to the entropy; hence, we refer to it as the ideal gas contribution, $S^{\mathrm{id}} = N s_1$, and define the excess entropy as

$$S - S^{\mathrm{id}} = -\int \overline{d\mathbf{X}} p(\mathbf{X}) \ln p(\mathbf{X}). \tag{B5}$$

In order to find the bound $H_{\mathrm{ex}} \geqslant S - S^{\mathrm{id}}$, we seek to maximize the Lagrangian

$$\mathcal{L} = -\int \overline{d\mathbf{X}} p(\mathbf{X}) \ln p(\mathbf{X}) - \eta\left(\int \overline{d\mathbf{X}} p(\mathbf{X}) - 1\right) - \mathrm{tr}\left[\mathcal{U}^\dagger \int \overline{d\mathbf{X}} p(\mathbf{X}) \left(\sum_{n,m} \underline{g}(\mathbf{x}_n)\underline{g}^\dagger(\mathbf{x}_m) - \mathcal{BC}\right)\right], \tag{B6}$$

such that $\mathcal{C}$ is the given CF, and find the entropy-maximizing $p(\mathbf{X})$ which reproduces that particular $\mathcal{C}$. Here $\mathcal{U}$ (a $\Xi \times \Xi$ matrix) and $\eta$ (scalar) are the Lagrange multipliers imposing the constraints, and the trace is of $\Xi \times \Xi$ matrices. The trace notation is a convenient shorthand arising from the matrix identity $\mathrm{tr}\mathcal{A}^\dagger \mathcal{B} = \sum_{ii'} \mathcal{A}_{ii'}^* \mathcal{B}_{i'i}$ (the Frobenius matrix inner product). Equation (B6) is equivalent to imposing every component of $\mathcal{BC}$ as a separate constraint. We should solve the equations

$$\frac{\delta \mathcal{L}}{\delta p(\mathbf{X})} = 0, \qquad \frac{\partial \mathcal{L}}{\partial \eta} = 0, \qquad \frac{\partial \mathcal{L}}{\partial \mathcal{U}^\dagger} = \mathcal{O}, \tag{B7}$$

where $(\partial/\partial \mathcal{U}^\dagger)$ denotes $\Xi^2$ partial derivatives with respect to each component of $\mathcal{U}^\dagger$.

Taking the variation with respect to $p(\mathbf{X})$ yields

$$p(\mathbf{X}) = e^{-1-\eta} \Gamma(\mathbf{X}), \quad \Gamma(\mathbf{X}) = \exp\left\{-\mathrm{tr}\left[\mathcal{U}^\dagger\left(\sum_{n,m} \underline{g}(\mathbf{x}_n)\underline{g}^\dagger(\mathbf{x}_m) - \mathcal{BC}\right)\right]\right\}. \tag{B8}$$

Substituting into the normalization condition for $p(\mathbf{X})$, we find

$$\eta = \ln Z - 1, \qquad Z = \int \overline{d\mathbf{X}} \Gamma(\mathbf{X}), \tag{B9}$$

leading to $p(\mathbf{X}) = Z^{-1} \Gamma(\mathbf{X})$. We see that $Z$ and $\mathcal{U}$ play the analogous roles of the partition function (relative to the ideal gas) and pair interaction, inside a Boltzmann-like factor. This analogy is a result of the variational procedure, despite the nonequilibrium setting of the formalism. The main difference from conventional equilibrium calculations is that we do not know the effective pair interaction, $\mathcal{U}$, which is yet to be found.

Next, consider the derivative of the partition function with respect to the components of $\mathcal{U}^\dagger$,

$$\frac{\partial \ln Z}{\partial \mathcal{U}^\dagger} = Z^{-1} \frac{\partial}{\partial \mathcal{U}^\dagger} \int \overline{d\mathbf{X}} \Gamma(\mathbf{X}) = \int \overline{d\mathbf{X}} p(\mathbf{X}) \sum_{n,m} \underline{g}(\mathbf{x}_n)\underline{g}^\dagger(\mathbf{x}_m) - \mathcal{BC} = \mathcal{O}. \tag{B10}$$

Therefore, solving $\partial \ln Z/\partial \mathcal{U}^\dagger = \mathcal{O}$ is equivalent to forcing the constraints on the CF. The maximum entropy is obtained as

$$H_{\mathrm{ex}} = -\int \overline{d\mathbf{X}} p(\mathbf{X}) \ln p(\mathbf{X}) = \int \overline{d\mathbf{X}} p(\mathbf{X}) \ln Z - \int \overline{d\mathbf{X}} p(\mathbf{X}) \ln \Gamma(\mathbf{X}) = \ln Z, \tag{B11}$$

where we used $\int \overline{d\mathbf{X}} p(\mathbf{X}) \ln \Gamma(\mathbf{X}) = 0$ upon enforcement of the constraint.

The derivation above shows that $p(\mathbf{X})$ is a critical point for the entropy. We now show that $p(\mathbf{X})$ is indeed a maximum. Assume that the true distribution is $p_0(\mathbf{X})$, which is normalized and satisfies Eq. (B3). Using the inequality $\ln x - 1 + 1/x \geqslant 0$ for any $x > 0$, we first notice that $\int \overline{d\mathbf{X}} p_0(\mathbf{X}) \ln[p_0(\mathbf{X})/p(\mathbf{X})] \geqslant \int \overline{d\mathbf{X}} p_0(\mathbf{X})[1 - p(\mathbf{X})/p_0(\mathbf{X})] = 0$. Thus, the thermodynamic excess entropy is bound according to

$$S - S^{\mathrm{id}} = -\int \overline{d\mathbf{X}} p_0(\mathbf{X}) \ln p_0(\mathbf{X}) \leqslant -\int \overline{d\mathbf{X}} p_0(\mathbf{X}) \ln p(\mathbf{X}) = \int \overline{d\mathbf{X}} p_0(\mathbf{X}) \ln Z - \int \overline{d\mathbf{X}} p_0(\mathbf{X}) \ln \Gamma(\mathbf{X}) = \ln Z = H_{\mathrm{ex}}. \tag{B12}$$

where we have used $\int \overline{d\mathbf{X}} p_0(\mathbf{X}) \ln \Gamma(\mathbf{X}) = 0$ since $p_0(\mathbf{X})$ gives rise to the same constrained CF. This proof holds for any additional constraints (including those which we introduce in the SM [34]).

Since $\mathcal{BC}$ is Hermitian by construction, and the Lagrangian is real, only the Hermitian part of $\mathcal{U}$ survives. In addition, transposing $\mathcal{U}$ amounts to relabeling the constraints. Henceforth, we will drop the †.

Thus we have found the relation between the bound on the excess entropy $H_{\text{ex}}$ and the effective partition function $Z$, and between $Z$ and the Lagrange multiplier $\mathcal{U}$ and the constraint $\mathcal{C}$. The procedure continues as follows. (a) Perform the integral of Eq. (B9) explicitly and find $Z$ as a functional of $\mathcal{C}$ and $\mathcal{U}$. (b) Obtain $\mathcal{U}$ as a function of $\mathcal{C}$ using Eq. (B10) and, consequently, find $Z$ in terms of $\mathcal{C}$ alone. (c) Obtain $H$ [Eq. (B11)] as a functional of $\mathcal{C}$. A similar calculation was carried out in Ref. [25]. There, $\mathcal{C}$ was limited to the number density field, as the only DOF was the particles' locations, and the CF was the structure factor. Here $\mathcal{C}$ is the field of a general DOF **j**. In the SM [34] we also include a steady inhomogeneous field and a varying number of particles.

Similar to Ref. [25] we transform to integration over fields, $\underline{\phi} = \sum_n \underline{g}(\mathbf{x}_n)$, which we introduce through a functional delta,

$$\delta\left[\underline{\phi} - \sum_n \underline{g}(\mathbf{x}_n)\right] = (2\pi)^{-\Xi} \int D\underline{\psi} \exp\left[i\underline{\psi}^{\dagger}\left(\underline{\phi} - \sum_n \underline{g}(\mathbf{x}_n)\right)\right], \tag{B13}$$

where $\underline{\psi}$ (a $\Xi$-component column vector) is the conjugate field. These fields are no longer functions of individual particles' positions, but rather functions of the modes. For later convenience, we add and subtract the uniform mode, which by construction satisfies $\underline{\phi}_0 = \sum_n \underline{g}_0(\mathbf{x}_n) = N$ for any $\mathbf{x}_n$. Some arbitrary value for the corresponding $\underline{\psi}_0$ may be assumed. The partition function becomes

$$Z = (2\pi)^{-\Xi} e^{\text{tr}\,\mathcal{UBC}} \int \overline{d\mathbf{X}} D\underline{\phi} D\underline{\psi} \exp\left[i\underline{\tilde{\psi}}^{\dagger}\left(\underline{\tilde{\phi}} - \sum_n \underline{\tilde{g}}(\mathbf{x}_n)\right) - \underline{\phi}^{\dagger}\mathcal{U}\underline{\phi}\right], \tag{B14}$$

where the tilde indicates that $\tilde{g}(\mathbf{x})$ is the complete transform (i.e., including the uniform mode which we have been excluding thus far). Therefore, $\underline{\tilde{g}}(\mathbf{x})$, $\underline{\tilde{\phi}}$, and $\underline{\tilde{\psi}}$ have $\Xi + 1$ components. The "tildeless" $\underline{g}$, $\underline{\phi}$, and $\underline{\psi}$ still exclude the uniform mode, and as before are of size $\Xi$. The integral over the configurations can be rewritten as

$$\int \overline{d\mathbf{X}} \exp\left(-i\underline{\tilde{\psi}}^{\dagger}\sum_{n=1}^N \underline{\tilde{g}}(\mathbf{x}_n)\right) = \left[\int \overline{d\mathbf{x}} e^{-i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{g}}(\mathbf{x})}\right]^N$$

$$= \left[1 + \int \overline{d\mathbf{x}}(e^{-i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{g}}(\mathbf{x})} - 1)\right]^N \xrightarrow{N\to\infty,\,\text{Pr}(\mathbf{x})\sim\upsilon^{-1}\to 0,\,N\overline{d\mathbf{x}}\sim 1;\,\underline{\tilde{\psi}}\to 0} \exp\left[N\int \overline{d\mathbf{x}}(e^{-i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{g}}(\mathbf{x})} - 1)\right]. \tag{B15}$$

We are left with the field integrals,

$$Z = (2\pi)^{-\Xi} e^{\text{tr}\,\mathcal{UBC}} \int D\underline{\phi} D\underline{\psi} \exp\left[N\int \overline{d\mathbf{x}}\left(e^{-i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{g}}(\mathbf{x})} - 1\right) + i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{\phi}} - \underline{\phi}^{\dagger}\mathcal{U}\underline{\phi}\right]. \tag{B16}$$

One course of action is to first perform the exact Gaussian integral over $\underline{\phi}$ and remain with a non-Gaussian integral over $\underline{\psi}$. The next step in Ref. [25] was to take the Gaussian approximation for this integral, thus including the leading-order fluctuations. An alternative course of action (involving the same level of approximation) is to first perform a saddle-point approximation for $\underline{\psi}$, and then take a Gaussian approximation for $\underline{\phi}$. For the sake of this Appendix, both methods are equally convenient, but in the SM [34] (including the constraint on a steady profile $\underline{G}$), the latter turns out to be much easier. The former method could be found in Ref. [25], and the latter is shown below. These approximations—saddle point in $\underline{\psi}$ and Gaussian in $\underline{\phi}$—are the only ones employed.

We rewrite the partition function as

$$Z = (2\pi)^{-\Xi} e^{\text{tr}\,\mathcal{UBC}} \int D\underline{\phi}\, e^{-\underline{\phi}^{\dagger}\mathcal{U}\underline{\phi}} \int D\underline{\psi}\, e^{-F[\underline{\tilde{\psi}};\underline{\tilde{\phi}}]}, \tag{B17a}$$

$$F[\underline{\tilde{\psi}};\underline{\tilde{\phi}}] = -N\int \overline{d\mathbf{x}}(e^{-i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{g}}(\mathbf{x})} - 1) - i\underline{\tilde{\psi}}^{\dagger}\underline{\tilde{\phi}}. \tag{B17b}$$

In the saddle point approximation, we seek $\underline{\tilde{\psi}} = \underline{\tilde{\psi}}_{\text{sp}}(\underline{\tilde{\phi}})$ that minimizes the effective free energy, $F[\underline{\tilde{\psi}}_{\text{sp}}(\underline{\tilde{\phi}});\underline{\tilde{\phi}}]$, for each $\underline{\tilde{\phi}}$, and then take leading order corrections. The equation for $\underline{\tilde{\psi}}_{\text{sp}}$ is

$$i\frac{\partial F}{\partial \underline{\tilde{\psi}}^{\dagger}} = \underline{0} = \underline{\tilde{\phi}} - N\int \overline{d\mathbf{x}}\,\underline{\tilde{g}}(\mathbf{x})e^{-i\underline{\tilde{\psi}}_{\text{sp}}^{\dagger}\underline{\tilde{g}}(\mathbf{x})}. \tag{B18}$$

Since the transform $\underline{\tilde{g}}(\mathbf{x})$ is a complete basis, there exists the dual object $\underline{\tilde{b}}(\mathbf{x})$ such that $\underline{\tilde{b}}^{\dagger}(\mathbf{x})\underline{\tilde{g}}(\mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}')/\text{Pr}(\mathbf{x})$ [and $\int \overline{d\mathbf{x}}\,\underline{\tilde{g}}(\mathbf{x})\underline{\tilde{b}}^{\dagger}(\mathbf{x}) = \tilde{\mathcal{I}}$, being the $(\Xi + 1) \times (\Xi + 1)$ unit matrix], where again $\text{Pr}(\mathbf{x})$ serves as the weight function. As an example, consider the simple particle-density field, where $\underline{\tilde{g}}_{\mathbf{q}}(\mathbf{r}) = e^{-i\mathbf{q}\cdot\mathbf{r}}$, such that $\underline{\tilde{b}}_{\mathbf{q}}(\mathbf{r}) = e^{-i\mathbf{q}\cdot\mathbf{r}}$ and $\text{Pr}(\mathbf{r}) = 1/V$, and, indeed,

$\sum_{\mathbf{q}} \underline{\tilde{b}}_{\mathbf{q}}^*(\mathbf{r}) \underline{\tilde{g}}_{\mathbf{q}}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')/\Pr(\mathbf{r})$ [and $\int d\mathbf{r} \Pr(\mathbf{r}) \underline{\tilde{g}}_{\mathbf{q}}(\mathbf{r}) \underline{\tilde{b}}_{\mathbf{q}'}^*(\mathbf{r}) = \delta_{\mathbf{qq}'}$]. With the help of $\underline{\tilde{b}}(\mathbf{x})$, we can solve Eq. (B18) as

$$\underline{\tilde{\psi}}_{\mathrm{sp}}^\dagger(\underline{\tilde{\phi}}) = i \int \overline{d\mathbf{x}} \ln[N^{-1} \underline{\tilde{b}}^\dagger(\mathbf{x}) \underline{\tilde{\phi}}] \underline{\tilde{b}}^\dagger(\mathbf{x}). \tag{B19}$$

We now expand the free energy around the saddle point, $\underline{\tilde{\psi}}_{\mathrm{sp}}$,

$$F[\underline{\tilde{\psi}}_{\mathrm{sp}}(\underline{\tilde{\phi}}); \underline{\tilde{\phi}}] = N \int \overline{d\mathbf{x}} \left[ \frac{\underline{\tilde{b}}^\dagger(\mathbf{x})\underline{\tilde{\phi}}}{N} \ln\left( \frac{\underline{\tilde{b}}^\dagger(\mathbf{x})\underline{\tilde{\phi}}}{N} \right) + 1 - \frac{\underline{\tilde{b}}^\dagger(\mathbf{x})\underline{\tilde{\phi}}}{N} \right] \simeq \frac{1}{2} \underline{\phi}^\dagger \mathcal{B}^{-1} \underline{\phi}, \tag{B20a}$$

$$\frac{\delta^2 F}{\delta \underline{\psi}^\dagger \delta \underline{\psi}} = \int \overline{d\mathbf{x}} \underline{\tilde{g}}(\mathbf{x}) \underline{\tilde{g}}^\dagger(\mathbf{x}) \underline{\tilde{b}}^\dagger(\mathbf{x}) \underline{\tilde{\phi}} \simeq \mathcal{B}. \tag{B20b}$$

We have kept terms to quadratic order in $\underline{\psi}$ and $\underline{\phi}$, $\underline{\phi}_0$ has naturally disappeared, and fluctuations in $\underline{\psi}_0$ are irrelevant. The approximate free energy reads

$$F[\underline{\psi}; \underline{\phi}] \simeq \tfrac{1}{2} \underline{\phi}^\dagger \mathcal{B}^{-1} \underline{\phi} + \tfrac{1}{2} \underline{\psi}^\dagger \mathcal{B} \underline{\psi}. \tag{B21}$$

Substituting Eq. (B21) in Eq. (B17a), we obtain the partition function

$$Z \simeq (2\pi)^{-\Xi} e^{\mathrm{tr}\, \mathcal{U}\mathcal{B}\mathcal{C}} \int D\underline{\phi}\, e^{-\frac{1}{2}\underline{\phi}^\dagger(2\mathcal{U}+\mathcal{B}^{-1})\underline{\phi}} \int D\underline{\psi}\, e^{-\frac{1}{2}\underline{\psi}^\dagger \mathcal{B}\underline{\psi}} = \exp\left( \mathrm{tr}\left[ \mathcal{U}\mathcal{B}\mathcal{C} - \frac{1}{2}\ln\left(2\mathcal{U}\mathcal{B} + \mathcal{I}\right) \right] \right), \tag{B22}$$

where we have used the matrix property, $\det[\exp(\cdot)] = \exp[\mathrm{tr}(\cdot)]$.

Thus we have found $Z$ as a functional of $\mathcal{U}$ and $\mathcal{C}$. From Eq. (B10), we obtain the effective pair potential,

$$\mathcal{U} \simeq \tfrac{1}{2}(\mathcal{C}^{-1} - \mathcal{I})\mathcal{B}^{-1}. \tag{B23}$$

Substitution of Eq. (B23) in Eq. (B22), and the result in Eq. (B11) gives

$$h_{\mathrm{ex}}[\mathcal{C}] \simeq \frac{1}{2N} \mathrm{tr}[\ln \mathcal{C} + \mathcal{I} - \mathcal{C}], \tag{B24}$$

which is the desired bound, Eq. (A3) (with $\underline{G} = \underline{0}$ and $\langle N \rangle = N$). In the SM [34], we derive the bound for a nonzero $\underline{G}$, and for the grand-canonical ensemble.

[1] J. Casas-Vazquez and D. Jou, Rep. Prog. Phys. **66**, 1937 (2003).

[2] L. F. Cugliandolo, J. Phys. A: Math. Theor. **44**, 483001 (2011).

[3] A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardar, and J. Tailleur, Nat. Phys. **11**, 673 (2015).

[4] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[5] L. F. Kozachenko and N. N. Leonenko, Probl. Peredachi Inf. **23**, 9 (1987).

[6] J. Beirlant, E. J. Dudewicz, L. Györfi, and E. C. Van der Meulen, Int. J. Math. Stat. Sci. **6**, 17 (1997).

[7] G. A. Darbellay and I. Vajda, IEEE Trans. Inf. Theory **45**, 1315 (1999).

[8] L. Paninski, Neural Comput. **15**, 1191 (2003).

[9] D. Stowell and M. Plumbley, IEEE Signal Process. Lett. **16**, 537 (2009).

[10] W. Lord, J. Sun, and E. Bollt, Chaos **28**, 033114 (2018).

[11] G. Ariel and Y. Louzoun, Entropy **22**, 236 (2020).

[12] R. Avinery, M. Kornreich, and R. Beck, Phys. Rev. Lett. **123**, 178102 (2019).

[13] S. Martiniani, P. M. Chaikin, and D. Levine, Phys. Rev. X **9**, 011031 (2019).

[14] M. Zu, A. Bupathy, D. Frenkel, and S. Sastry, J. Stat. Mech.: Theory Exp. (2020) 023204.

[15] D. P. Feldman and J. P. Crutchfield, Phys. Rev. E **67**, 051104 (2003).

[16] O. Melchert and A. K. Hartmann, Phys. Rev. E **91**, 023306 (2015).

[17] A. Nir, E. Sela, R. Beck, and Y. Bar-Sinai, Proc. Natl. Acad. Sci. USA **117**, 30234 (2020).

[18] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 3rd ed. (Elsevier, London, 1986).

[19] A. Cavagna, D. Conti, I. Giardina, T. S. Grigera, S. Melillo, and M. Viale, Phys. Biol. **13**, 065001 (2016).

[20] S. Ro, Y. Kafri, M. Kardar, and J. Tailleur, Phys. Rev. Lett. **126**, 048003 (2021).

[21] D. J. G. Pearce, J. Nambisan, P. Ellis, A. Fernandez-Nieves, and L. Giomi, Phys. Rev. Lett. **127**, 197801 (2021).

[22] T. Sanchez, D. T. Chen, S. J. DeCamp, M. Heymann, and Z. Dogic, Nature **491**, 431 (2012).

[23] L. Onsager, Ann. N.Y. Acad. Sci. **51**, 627 (1949).

[24] P.-G. De Gennes and J. Prost, *The Physics of Liquid Crystals* (Oxford University Press, Oxford, 1993).

[25] G. Ariel and H. Diamant, Phys. Rev. E **102**, 022110 (2020).

[26] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, Proc. Natl. Acad. Sci. USA **109**, 4786 (2012).

[27] Y. Shemesh, Y. Sztainberg, O. Forkosh, T. Shlapobersky, A. Chen, and E. Schneidman, eLife **2**, e00759 (2013).

[28] R. Asor, L. Selzer, C. J. Schlicksup, Z. Zhao, A. Zlotnick, and U. Raviv, ACS Nano **13**, 7610 (2019).

[29] E. Schneidman, S. Still, M. J. Berry, and W. Bialek, Phys. Rev. Lett. **91**, 238701 (2003).

[30] A. Cavagna, I. Giardina, F. Ginelli, T. Mora, D. Piovani, R. Tavarone, and A. M. Walczak, Phys. Rev. E **89**, 042707 (2014).

[31] S. Chakrabarty and Z. Nussinov, Phys. Rev. B **84**, 064124 (2011).

[32] G. Zhang and S. Torquato, Phys. Rev. E **101**, 032124 (2020).

[33] J. Ricouvier, R. Pierrat, R. Carminati, P. Tabeling, and P. Yazhgur, Phys. Rev. Lett. **119**, 208001 (2017).

[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.107.014138 for further experimental data, detailed derivation of the extensions, and entropy estimation with variable number of particles [54].

[35] A. J. Leadbetter, *Structural Studies of Nematic, Smectic A and Smectic C phases*, Chap. 13 (Academic Press, London, 1979).

[36] Equations (2a) define position-dependent CFs, but our formalism is not restricted to spatial coordinates. For example, the DOFs may be only the orientations, and the Fourier transform would be replaced by a decomposition into spherical harmonics.

[37] In certain cases, as in the case of the structure factor, the CF includes a contribution of order $N$ at $\mathbf{q} = \mathbf{0}$, which is why this mode is removed from the summation or integration. They must begin from $|\mathbf{q}| \to 0^+$.

[38] S. Torquato, *Random Heterogeneous Materials* (Springer, New York, 2001).

[39] C. E. Zachary, Y. Jiao, and S. Torquato, Phys. Rev. Lett. **106**, 178001 (2011).

[40] Y. Jiao and S. Torquato, Phys. Rev. E **84**, 041309 (2011).

[41] S. Wilken, R. E. Guerra, D. Levine, and P. M. Chaikin, Phys. Rev. Lett. **127**, 038002 (2021).

[42] S. Torquato, Phys. Rep. **745**, 1 (2018).

[43] D. C. Duffy, J. C. McDonald, O. J. Schueller, and G. M. Whitesides, Anal. Chem. **70**, 4974 (1998).

[44] M. Skoge, A. Donev, F. H. Stillinger, and S. Torquato, Phys. Rev. E **74**, 041127 (2006).

[45] S. Atkinson, F. H. Stillinger, and S. Torquato, Phys. Rev. E **94**, 032902 (2016).

[46] J. K. Percus and G. J. Yevick, Phys. Rev. **110**, 1 (1958).

[47] For orthogonal bases $\underline{f}$ [i.e., $\mathcal{A}$ of Eq. (3b) is diagonal], $h_{\mathrm{cross}} \leqslant 0$ always, which is not the case here. For example, $h_{\mathrm{cross}}$ from a CF that mixes position and orientation would be nonpositive, since the basis is a subset of circular and spherical harmonics.

[48] P. Wochner, C. Gutt, T. Autenrieth, T. Demmer, V. Bugaev, A. D. Ortiz, A. Duri, F. Zontone, G. Grübel, and H. Dosch, Proc. Natl. Acad. Sci. USA **106**, 11511 (2009).

[49] T. M. Truskett, S. Torquato, and P. G. Debenedetti, Phys. Rev. E **62**, 993 (2000).

[50] J. Mittal, J. R. Errington, and T. M. Truskett, J. Chem. Phys. **125**, 076102 (2006).

[51] A. Be'er, B. Ilkanaiv, R. Gross, D. B. Kearns, S. Heidenreich, M. Bär, and G. Ariel, Commun. Phys. **3**, 66 (2020).

[52] M. E. Cates and J. Tailleur, Annu. Rev. Condens. Matter Phys. **6**, 219 (2015).

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

[54] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, Cambridge, 2001).