

This is the accepted manuscript made available via CHORUS. The article has been published as:

Characterizing pixel and point patterns with a hyperuniformity disorder length

A. T. Chieco, R. Dreyfus, and D. J. Durian

Phys. Rev. E **96**, 032909 — Published 15 September 2017

DOI: [10.1103/PhysRevE.96.032909](https://doi.org/10.1103/PhysRevE.96.032909)

Characterizing Pixel and Point Patterns with a Hyperuniformity Disorder Length

A.T. Chieco¹, R. Dreyfus² and D.J. Durian¹

¹*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104-6396, USA*

²*Complex Assemblies of Soft Matter, CNRS-Solvay-UPenn UMI 3254, Bristol, PA, 19007-3624, USA*

(Dated: September 1, 2017)

We introduce the concept of a “hyperuniformity disorder length” h that controls the variance of volume fraction fluctuations for randomly placed windows of fixed size. In particular, fluctuations are determined by the average number of particles within a distance h from the boundary of the window. We first compute special expectations and bounds in d dimensions, and then illustrate the range of behavior of h versus window size L by analyzing several different types of simulated two-dimensional pixel patterns – where particle positions are stored as a binary digital image in which pixels have value zero/one if empty/contain a particle. The first are random binomial patterns, where pixels are randomly flipped from zero to one with probability equal to area fraction. These have long-ranged density fluctuations, and simulations confirm the exact result $h = L/2$. Next we consider vacancy patterns, where a fraction f of particles on a lattice are randomly removed. These also display long-range density fluctuations, but with $h = (L/2)(f/d)$ for small f , and $h = L/2$ for $f \rightarrow 1$. And finally, for a hyperuniform system with no long-range density fluctuations, we consider “Einstein patterns” where each particle is independently displaced from a lattice site by a Gaussian-distributed amount. For these, at large L , h approaches a constant equal to about half the root-mean-square displacement in each dimension. Then we turn to grayscale pixel patterns that represent simulated arrangements of polydisperse particles, where the volume of a particle is encoded in the value of its central pixel. And we discuss the continuum limit of point patterns, where pixel size vanishes. In general, we thus propose to quantify particle configurations not just by the scaling of the density fluctuation spectrum but rather by the real-space spectrum of $h(L)$ versus L . We call this approach “Hyperuniformity Disorder Length Spectroscopy” (HUDLS).

The structure of crystals is straightforward to describe. By contrast, it remains a fundamental problem to identify and characterize the structural features of disordered many-body systems that relate to materials properties. In this regard the concept of hyperuniformity was developed for systems in which long-wavelength density fluctuations are suppressed to nearly the same extent as in crystals [1, 2]. Disordered systems that are hyperuniform accordingly possess some degree of long-ranged hidden order, not evident in the local structure, and this can give rise to special material properties. For example, hyperuniformity has been connected to jamming [3–12] and to novel photonic behavior [13–17] in amorphous materials. It is relevant for the spatial arrangement of certain biological [18] and astrophysical [19] objects. And it has been connected with irreversibility in seemingly-deterministic many-body systems subjected to periodic driving [20–22], as well as to dynamical behavior of Brownian particles subjected to steady driving [23]. Thus the importance of quantifying hyperuniformity is increasingly recognized.

According to the original definition, a d -dimensional system of point particles is said to be hyperuniform if the structure factor $S(q)$ vanishes at wavevector $q = 0$. [1]. Equivalently, it is hyperuniform if the variance in the number of particles enclosed by a set of randomly-placed measuring windows of volume $\propto L^d$ grows as the average number of particles on the window surface, $\sigma_N^2(L) \sim L^{d-1}$ [1]. Here L is the width or diameter of the window and d is dimensionality. By contrast, the number variance for liquid-like systems with long-range density fluctuations grows like the average number particles inside the entire window, $\sigma_N^2(L) \sim L^d$; this cor-

responds to $S(0) > 0$. For extended particles of nonzero volume, the quantities of interest are instead the spectral density $\chi(q)$ and the variance $\sigma_\phi^2(L)$ of volume-fraction fluctuations [2, 7, 8]. The usual focus is on the small- q / large- L scaling; i.e. on

$$\chi(q) \sim q^\epsilon, \quad (1)$$

$$\sigma_\phi^2(L) \sim 1/L^{d+\epsilon}, \quad (2)$$

and the value of the exponent. For $\epsilon = 0$, then there are long-ranged density fluctuations and the particles have a liquid-like arrangement that is significantly more random than in a crystal. For $\epsilon > 0$, the particles are arranged more uniformly throughout space and the system is said to be hyperuniform. The degree of hidden order increases with ϵ , and the upper limit $\epsilon = 1$ corresponds to a crystal-like uniformity. Logarithmic corrections to Eqs. (1-2) may be expected, and can be described by a single value of ϵ over a (perhaps experimentally-limited) range of q or L .

To relate to structure, it is helpful to translate measured quantities into a length scale. For example Hopkins *et al.* [24] define a correlation length from the value of $S(0)$. And Refs. [12, 20, 22] define other lengths based on the location of features in the form of $\chi(q)$ versus q or of $\sigma^2(L)$ versus L . Here, we introduce a completely length based on the value of $\sigma_\phi^2(L)$ that we call the “hyperuniformity disorder length”, $h(L)$. The size of $h(L)$ relates directly to the distance from the boundary over which fluctuations occur for a set of measuring windows, and its *value* – as well as its scaling with L – indicates the degree of uniformity.

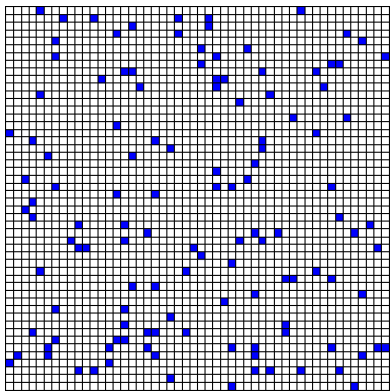


FIG. 1: (color online) Example binary pixel pattern, consisting of a 50×50 grid of pixels and 125 particles (solid blue) occupying an area fraction of $\phi = 125/(50 \times 50) = 0.05$.

Particle positions, whether from experiment or simulation, can be represented by a digital image where pixel size is determined by experimental or numerical precision. Accordingly, we develop our approach for d -dimensional pixel patterns. A binary 2-dimensional example is shown in Fig. 1, where space is represented by a square grid of pixels that have value $I(x, y) = 0$ if empty or $I(x, y) = 1$ if a particle is present. We begin by deriving the volume fraction variance for a few special cases. Using this as a guide, we discuss how to quantify hidden order from the measured variance, first by its ratio to that in a totally random arrangement of the same objects, and then by the hyperuniformity disorder length. The continuum limit of vanishing pixel size, i.e. of point particles, is also discussed. Then we illustrate this program for several classes of system, starting with binary pixel patterns and ending with grayscale pixel patterns representing a polydisperse collection of extended particles.

I. VARIANCE FOR PIXEL PATTERNS

In this section we define notation and compute the volume fraction variance exactly for four special arrangements of pixel particles (Poisson patterns, multinomial patterns, separated-particles, and cubic crystals). As in Fig. 1, the patterns to be characterized consist of a d -dimensional Cartesian grid of pixels with intensity values $I(x, y, z \dots)$ that specify the arrangement of pixel-sized particles. The “volume” of a particle of species i is considered to be $v_i = I_i p_o^d$, where p_o is the pixel width and p_o^d is the pixel volume. Then the volume fraction ϕ equals the average intensity in the pattern. For example, a system of extended particles with actual geometrical volumes v_i can be represented by a pixel pattern where the central pixel inside each particle is incremented by intensity $I_i = v_i/p_o^d$. In general, if species i has number density ρ_i , then it occupies volume fraction $\phi_i = v_i \rho_i = I_i p_o^d \rho_i$ and the total volume fraction

occupied by all particles is $\phi = \sum \phi_i$. The average probability for a pixel to contain a particle of species i is $q_i = p_o^d \rho_i = \phi_i/I_i$, to be used below.

The basic tool for quantifying the extent of hidden order and hyperuniformity is the volume fraction variance $\sigma_\phi^2(L)$ versus window size. Operationally, a large number of measuring windows of volume V_Ω are placed at random. In each window Ω , the volume fraction is computed as the average intensity. The results vary from window to window, but the average must converge to the total/true volume fraction ϕ if enough windows are sampled. The variance of the measured volume fraction values, on the other hand, depends on the particular pattern and on the window size/shape. The nature of the hidden order is to be analyzed from this behavior.

For analysis of simulated pixel patterns, we use cubic measuring windows of volume $V_\Omega = L^d$. This will be useful for future analyses of digital video data. For continuum space, it is more usual to use spherical measuring windows. However this is inconvenient for pixel patterns, because spheres cannot be constructed except as pixelated approximates. In this section, results for arbitrary window shapes are given in terms of V_Ω and results for hypercubic windows are given in terms of L . Continuum results for hyperspherical windows of diameter D are given in the conclusion. The behavior of $h(L)$ for hypercubic windows is almost the same as $h(D)$ for hyperspherical windows.

A. Random Patterns and Relative Variance

The volume fraction variance may be computed exactly, as follows, if the pixel particles are arranged at random. For a given L^d measuring window of $n = (L/p_o)^d$ pixels, the total volume of pixel particles is $V = \sum v_i N_i$ where N_i is the actual number of species i enclosed (i.e. the number of time enclosed pixels were incremented by $+I_i$). The corresponding average and variance over window locations are $\bar{V} = \sum v_i \bar{N}_i$ and $\sigma_V^2 = \sum v_i^2 \sigma_{N_i}^2 + \sum_{i \neq j} v_i v_j \sigma_{N_i N_j}$, where $\bar{N}_i = \rho_i L^d$ is the average number of species i enclosed, $\sigma_{N_i}^2$ is the variance, and $\sigma_{N_i N_j} = \bar{N}_i \bar{N}_j - \bar{N}_i \bar{N}_j$ is the covariance.

To evaluate the variance and covariance for the number of enclosed particles, we must now make a distinction between two different kinds of random configuration. If multiple particles can occupy the same pixel, then Poisson statistics hold, the number variance is $\sigma_{N_i}^2 = n q_i$ where the probability q_i was given above, and the covariance is zero. In this case, each particle’s location is chosen at random and multiple particles may freely occupy a pixel. Such configurations are called “Poisson” patterns. By contrast, if only one particle at a time is allowed on each pixel, then multinomial statistics hold: the number variance is $\sigma_{N_i}^2 = n q_i (1 - q_i)$ and the covariance is $\sigma_{N_i N_j} = -n q_i q_j$. This case corresponds to a pattern where the intensity of each pixel is randomly drawn from $\{0, I_1, I_2, \dots\}$ with respective probabilities

$\{1 - \sum q_i, q_1, q_2, \dots\}$. We call such configurations “multinomial” patterns.

Combining these ingredients, the volume fraction variance for the two different types of random pattern is computed to be

$$\sigma_\phi^2(L) = \frac{\langle I \rangle p_o^d}{L^d} \begin{cases} \phi & \text{Poisson,} \\ \phi[1 - \phi/\langle I \rangle] & \text{Multinomial,} \end{cases} \quad (3)$$

where $\langle I \rangle = \sum \phi_i I_i / \phi$ is the volume-fraction weighted average grayscale intensity of the particle species. The common prefactor may be re-written as $\langle v \rangle / V_\Omega$ where $\langle v \rangle = \langle I \rangle p_o^d$ is the volume-fraction weighted particle volume and V_Ω is the window volume – no matter what its shape. Thus, the variance scales as $1/L^d$ as expected for a system with long-range density fluctuations. And, importantly for later, the actual proportionality constants are now known. In principle, Eqs. (3-4) could be obtained from reciprocal-space results [2, 25–27].

To our knowledge, neither pixelated space nor the distinction between Poisson and multinomial statistics have been previously considered. This is important when a large fraction of pixels contain particles. For example, a random “binomial pattern” of non-overlapping binary particles, with $I = 1$, has variance $\sigma_\phi^2(L) = (\langle v \rangle / V_\Omega)[\phi(1 - \phi)]$; this holds at all ϕ , and is correctly unchanged by image inversion $I \rightarrow 1 - I$ and $\phi \rightarrow 1 - \phi$. Note that Poisson statistics are recovered to good approximation when the populated pixels are dilute and the q_i are small compared to one. This happens for a central-pixel representation when the image resolution is good, i.e. in the continuum limit where the pixel size is small enough that $\phi \ll \langle v \rangle / p_o^d = \langle I \rangle$ is true.

On the basis of Eqs. (3,4) we define the *relative variance* for general configurations as

$$\mathcal{V}(L) \equiv \begin{cases} \frac{\sigma_\phi^2(L)}{\phi} & \text{Poisson,} \\ \frac{\sigma_\phi^2(L)}{\phi[1 - \phi/\langle I \rangle]} & \text{Multinomial,} \end{cases} \quad (5)$$

where the two cases are respectively for particles that are, or are not, allowed to overlap. Then for both types of random pixel patterns, the relative variance is $\mathcal{V}(L) = \langle v \rangle / V_\Omega$ for any volume fraction. This equals $\langle I \rangle$ at the smallest window size, $V_\Omega = p_o^d$, and has an initial linear decay of $\mathcal{V}(L) = \langle I \rangle [1 - d(L - p_o)/p_o + \mathcal{O}(L - p_o)^2]$ for hypercubic windows. For patterns with hidden order, $\mathcal{V}(L)$ will be smaller than this upper bound [28].

B. Separated-Particles Limit

The volume fraction variance may also be computed exactly for small L , when all measuring windows con-

tain no more than one particle. For example if the pixel particles effectively repel each other, or if for another reason they have some minimum separation, then for a set of w randomly placed L^d measuring windows that are sufficiently small, the results will be some numbers $\{M_o, M_1, M_2, \dots\}$ of observed intensity values $\{0, I_1, I_2, \dots\}$. Then the measured volume fraction moments are $\phi_m^n = \sum (M_i/M)(I_i p_o^d / L^d)^n$, where $M = \sum M_i$, and the measured relative variance is

$$\mathcal{V}_m(L) = \frac{\sum \frac{M_i}{M\phi} (I_i p_o^d / L^d)^2 - \phi}{1 - \phi/\langle I \rangle}, \quad (7)$$

assuming multinomial statistics as appropriate for separated particles. The averages are set by $\bar{M}_i = M \rho_i L^d = M[\phi_i / (I_i p_o^d)] L^d$. This gives $\bar{\phi}_m^n = \sum \phi_i (I_i p_o^d / L^d)^{n-1}$, which leads to $\bar{\phi}_m = \sum \phi_i = \phi$ and a relative volume fraction variance of

$$\mathcal{V}(L) = \langle I \rangle \frac{(p_o/L)^d - \phi/\langle I \rangle}{1 - \phi/\langle I \rangle}, \quad (8)$$

Note that Eq. (8) is exact and holds for small enough L , for *any* arrangement of particles. If the arrangement is random, then it holds only for $L = p_o$ since there will be a nonzero number of adjacent pixel particles. If the particles are separated, then Eq. (8) holds up to some larger L that is set by the smallest particle-particle separation. For windows of arbitrary shape, Eq. (8) generalizes to $\mathcal{V}(L) = (\langle v \rangle / V_\Omega - \phi) / (1 - \phi/\langle I \rangle)$. To our knowledge, the small window limit of separated particles has not been previously considered.

C. Cubic Crystal

As a third example that may be computed exactly, we consider a pattern consisting of pixel particles of volume $v = I_o p_o^d$ on a d -dimensional cubic lattice of spacing b , using L^d hypercubic measuring windows, where both the lattice and the windows are aligned with the grid of pixels. For this situation there are only $(b/p_o)^d$ distinct locations for the windows; and there are only $d + 1$ distinct results for the number n of enclosed particles. These are all of form $n = \sum_{j=0}^d a_j [\text{floor}(L/b)]^j$, with $a_d = 1$. For example, in $d = 1$ there are $(b - \delta)/p_o$ ways for $n_1 = \text{floor}(L/b)$ particles to be enclosed, and δ/p_o ways for $n_1 + 1$ particles to be enclosed, where $\delta = L - b \text{floor}(L/b)$. The probability distribution and the moments for the number of enclosed particles may then be evaluated. Exact results for the average and variance of the measured volume fractions are found by direct summation to be $\phi = I_o (p_o/b)^d$, as expected, and

$$\sigma_{\phi^2}(L) = \phi^2 \left(\frac{b}{L} \right)^{2d} \left\{ \frac{L}{b} - \text{floor} \left(\frac{L}{b} \right) \left[1 - 2 \frac{L}{b} + \text{floor} \left(\frac{L}{b} \right) \right] \right\}^d - \phi^2. \quad (9)$$

We calculated Eq. (9) explicitly in 1, 2, and 3 dimensions, and we verified the $d = 2$ case by comparison with simulation. Though we are unaware of prior statement or proof, we suppose it is true in all dimensions. The corresponding number variance is $\sigma_N^2(L) = [\sigma_{\phi^2}(L)/\phi^2](L/b)^{2d}$. For $L < b$ the results simplify to $\sigma_{\phi^2}(L) = \phi^2[(b/L)^d - 1]$ and $\sigma_N^2(L) = (L/b)^d[1 - (L/b)^d]$. The former agrees with Eq. (8); the latter agrees with the $d = 1$ case considered in Ref. [1], where the number variance was computed from the structure factor and expressed in Eq. (83) as a Fourier series. For increasing L , the variance vanishes at integer values of L/b . The variance at half-way between the zeros gives the large- L decay envelope as $\sigma_{\phi^2}(L) = (d/4)(\phi b/L)^2$. This is pathological for $d > 1$, since crystals ought to be strongly hyperuniform with variance scaling of $1/L^{d+1}$. Our result is even more pathological than expected based on footnote 11 on p. 14 of Ref. [19], which indicates that $\sigma_{\phi^2}(L) \sim 1/L^{d-1}$ was found for cubic crystals with cubic measuring windows. The pathology arises because the distribution of measured ϕ_{Ω} values is highly non-Gaussian; hence, it may be removed by using spherical measuring windows [1, 19]. For pixelated space it is more convenient to tilt the lattice at multiple angles with respect to the pixel grid, as done in the simulations below.

D. Continuum Limits

The above expectations were all derived for pixelated space, where window widths and lattice spacings are an integer number of pixel widths p_o , and where each particle has intensity I_i and occupies a p_o^d voxel. But the results all also extend to continuous space for a point representation of particles of volume fraction ϕ and ϕ_i -weighted average particle volume $\langle v \rangle$. This is the limit of vanishing p_o and diverging I_i taken simultaneously such that $v_i = I_i p_o^d$ and ϕ_i are constant. Then multinomial statistics reduce to Poisson statistics, since $\phi/\langle I \rangle$ vanishes and there is zero probability for two point particles to lie on top of one another. For measuring windows of any shape, and volume V_{Ω} , the relative variance expectation of Eq. (3) for a random arrangement then becomes $\sigma_{\phi^2}(L)/\phi = \langle v \rangle/V_{\Omega}$. For an infinite system, where there is no minimum particle-particle separation, this holds for all V_{Ω} . For a finite system, however, it fails for small V_{Ω} where the windows all contain no more than one particle. For *any* pattern and window shape at such small V_{Ω} , the relative variance expectation from Eq. (8) is $\sigma_{\phi^2}(L)/\phi = \langle v \rangle/V_{\Omega} - \phi$. For cubic crystals, Eq. (9) holds in the continuum limit as written.

II. CHARACTERIZING HIDDEN ORDER

We now discuss how to measure the volume fraction variance for a given pixel pattern, then we propose two different ways to plot and interpret the results.

A. Measurements and Errors

Standard procedure is to measure the variance of the list of volume fractions found inside a large number w of randomly placed measuring windows Ω of desired size and shape [1]. For pixel patterns, the volume fraction inside a particular measuring window is $\phi_m = [\sum I(x, y, z, \dots) p_o^d]/V_{\Omega}$, which equals the total enclosed particle volume divided by the measuring window volume $V_{\Omega} \sim L^d$; this is the same as the average intensity of all the enclosed pixels. The average $\bar{\phi}_m$ and the variance $\sigma_m^2 = \bar{\phi}_m^2 - \phi^2$ of the measured volume fractions may then be computed. For a large number w of windows, $\bar{\phi}_m$ will converge to the actual total volume fraction ϕ of the whole pattern. But how large must w be, and what is the resulting expected statistical uncertainty between the measured variance and the true value? To our knowledge, this has not been addressed in prior work.

The error analysis for large measuring windows is most straightforward. Since then the distribution of measured volume fractions ought to be Gaussian, the uncertainty can be estimated from the standard error of the variance as $\Delta\sigma_{\phi^2}(L) = \sigma_{\phi^2}(L)\sqrt{2/(s-1)}$, where s is the number of independent samples. If the w measuring windows are small and the pattern is large, then the windows measure disjoint sets of particles and hence $s = w$ may be used. But for measuring windows that are sufficiently large in size or number, some of the particles will be sampled by more than one window and hence s will become smaller than w . In general, the number s of independent samples will be the total number of pixels covered by the entire set of measuring windows divided by the number of pixels per window. To account for this we create an auxiliary binary image where pixel values $I_{aux}(x, y)$ are flipped from zero to one if the corresponding pixel in the image lies within a measuring window. Then the number of independent samples is taken from the auxiliary image as $s = (\sum I_{aux})/(L/p_o)^d$.

To cover most of the sample with windows, there is no need for w to be more than on the order of V/V_{Ω} where V is the volume of the entire image. This gives a number of independent samples that is close to $s_{max} = V/V_{\Omega}$. As window size increases toward system size, s_{max} decreases

to one and causes the standard error of the variance to bloom. In our analyses of two-dimensional simulations, below, we take the number of $L \times L$ square windows to be $w = A/(2L^2)$, where A is the image area, subject to the constraint $10^2 \leq w \leq 10^4$. And the largest window size we use is typically $L = \sqrt{A}/2$. Results for larger measurement windows are contaminated by finite size artifacts, such that the area-fraction variance is artificially depressed toward zero as L approaches the image width [11]. Care must be taken that this systematic error not be mistakenly interpreted as a sign of hyperuniformity or lack thereof.

For small measuring windows, the standard error of the variance is not a good estimate because the volume fraction distribution is not Gaussian. Instead, the statistical uncertainty may be deduced from Eq. (7), which gives $(\Delta\mathcal{V})^2 \propto \sum (\Delta w_i)^2$ where w_i is the number of small measuring windows that contain one particle of species i . The key ingredient is $(\Delta w_i)^2 = \overline{w_i} = w \rho_i V_\Omega$. As before, w should be replaced with s . The resulting formula for the uncertainty is good only for small enough windows that contain no more than one particle. But we suppose it might serve as an estimate for intermediate window size, anyway. Combining this in quadrature with the standard error of the variance, which applies for large windows, gives the expected statistical uncertainty in the measured relative variance as

$$\Delta\mathcal{V} = \sqrt{\frac{\langle v^3 \rangle / V_\Omega^3}{s\phi[1 - \phi/\langle I \rangle]^2} + \frac{2\mathcal{V}^2}{s-1}}, \quad (10)$$

assuming multinomial statistics. For Poisson statistics, the continuum limit, or pixel particles with $\phi \ll \langle I \rangle$, the $-\phi/\langle I \rangle$ term is dropped in comparison with one. Note that the first term dominates for small windows, while the second dominates for large windows; therefore, it does no harm to simply add them together. We will use Eq. (10) to generate error bars in analysis plots, below. Note that the ϕ_i -weighted moments of the particle volume distribution may be written as $\langle v^n \rangle = \sum \phi_i v_i^n / \phi = (\sum N_i v_i^{n+1}) / (\sum N_i v_i)$ where the sums are over species and where N_i is the total number of particles of species i in the sample. This may also be written as $\langle v^n \rangle = \sum v_j^{n+1} / \sum v_j$ where the sums are over all particles in the sample.

B. Variance Ratio, $\mathcal{R}(L)$

The volume fraction variance is largest for a totally random pattern, and necessarily decreases with ordering. Therefore the presence of hidden order can be detected by comparing $\sigma_\phi^2(L)$ for the pattern in question with $[\sigma_\phi^2(L)]_{rand}$ for a totally random arrangement of the *same* set of particles. For non-overlapping pixel particles, the variance at $L = p_o$ merely measures the distribution of particle intensities independent of their arrangement; therefore, $\sigma_\phi^2(p_o) = [\sigma_\phi^2(p_o)]_{rand} = \phi \langle I \rangle [1 - \phi/\langle I \rangle]$ holds

and we may write $[\sigma_\phi^2(L)]_{rand} = \sigma_\phi^2(p_o)(p_o/L)^d$. Hence we define the variance ratio, and evaluate it without having to know ϕ and $\langle I \rangle$, as

$$\mathcal{R}(L) \equiv \frac{\sigma_\phi^2(L)}{[\sigma_\phi^2(L)]_{rand}} = \frac{\sigma_\phi^2(L)L^d}{\sigma_\phi^2(p_o)p_o^d}. \quad (11)$$

By construction, totally random patterns have $\mathcal{R}(L) = 1$ and this is an upper bound. Note, too, that the variance ratio is normalized to $\mathcal{R}(p_o) = 1$ for all pixel patterns. For the continuum limit of point particle and spherical measuring windows, simply replace L with diameter D , and p_o with some chosen smallest measuring diameter D_o that is less than the minimum particle-particle separation r_{min} in the sample. Either D_o should be vanishingly small compared to r_{min} , or else a correction should be made using the above results for separated-particle arrangements:

$$\mathcal{R}(D) = \frac{\sigma_\phi^2(D)D^d}{[\sigma_\phi^2(D_o) + \phi^2]D_o^d}. \quad (12)$$

This gives $\mathcal{R}(D) = 1 - \phi V_\Omega / \langle v \rangle$ exactly for $D < s$, and is normalized to $\mathcal{R}(D_o) = 1$ only as D_o vanishes and $\sigma_\phi^2(D_o)$ diverges. To interpret $\mathcal{R}(L)$, first note that the large- L scaling is $\mathcal{R}(L) \sim 1/L^\epsilon$ for $\sigma_\phi^2(L) \sim 1/L^{d+\epsilon}$. Therefore, $\mathcal{R}(L)$ goes to a constant for patterns with long-range density fluctuations ($\epsilon = 0$); and it goes to zero as a dimension-independent power-law if the pattern is hyperuniform, e.g. most notably $\mathcal{R}(L) \sim 1/L$ for strongly hyperuniform ($\epsilon = 1$). The utility of diagnosing hyperuniformity via the product $\sigma_\phi^2(L)L^d$ has been previously recognized [11, 12]. The advantage of additionally normalizing by $\sigma_\phi^2(p_o)p_o^d$ is that then the *value* (not just the scaling) has meaning. Smaller \mathcal{R} means more order, larger \mathcal{R} means more random, and $\mathcal{R} = 1$ means totally random. Thus, $\mathcal{R}(L)$ can be thought of as a randomness index, which decays from one as L increases and hidden order is detected. As shown below, $\mathcal{R}(L)$ can also be interpreted as the fraction f of space available for density fluctuations at wavelength L .

C. Hyperuniformity Disorder Length, $h(L)$

While $\mathcal{R}(L)$ is a useful new quantity, it does not connect to the original idea [1] that fluctuations in hyperuniform systems are governed by particles on the *surface* of the measuring windows. Since particles reside in a volume, we quantify this notion by introducing a hyperuniformity disorder length $h(L)$ that specifies the region near the window boundary where fluctuations are important. This concept is depicted in Fig. 2, which shows an $L \times L$ measuring window that is partitioned into a boundary region of thickness h and an $(L-2h) \times (L-2h)$ interior region. Intuitively, h is defined such that, if the system is ergodic and the time variation $\phi_\Omega(t)$ for one window has the same distribution as for a randomly placed set of windows, then only the boundary particles have opportunity

to temporarily leave the measuring window. For a totally random system where all enclosed particles participate, $h(L)$ would equal the maximum value $L/2$. For a “maximally” hyperuniform system, $h(L)$ would be a minimum and equal to $p_o/2$, which is the largest constant consistent with the $h \leq L/2$ requirement for all L ; in this case, the participating pixel particles are literally on the surface. Since $h(L)$ increases with disorder, we call it the hyperuniformity disorder length. In general, the nature of the fluctuations is thus easily visualized in terms of the value and form of the real-space spectrum of $h(L)$ versus L . We therefore call this analysis method “Hyperuniformity Disorder Length Spectroscopy” (HUDLS).

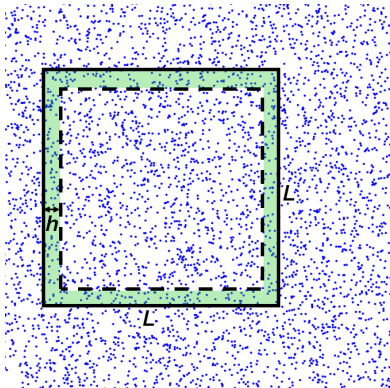


FIG. 2: (color online) Example “Einstein pattern” of pixel particles with a Gaussian-distributed random displacement from a triangular lattice. The lattice constant is 30 pixels, and the root-mean-square displacement is 150 pixels in each dimension. Since the particles are effectively bound to the lattice sites, the pattern is hyperuniform with no long-range density fluctuations – even though it appears very disordered to the eye. Randomly placed measurement windows hence show density fluctuations that are due only to particles within a distance h of the window boundary that is constant for large windows. Here this “hyperuniformity disorder length” is $h = 82$ pixels, as illustrated for an example 1100×1100 window. For clarity, the pixel particles are shown as dots that are 11 pixels across.

The technical definition of $h(L)$ is based on the measured variance compared to that for a totally random arrangement of the same set of particles. For a truly random pattern, the number of pixels with fluctuating particles is equal to the number $n = (L/p_o)^d$ of pixels in the entire L^d measuring window. For non-random patterns with smaller variance, we take it instead to be the number $n_B = [L^d - (L - 2h)^d]/p_o^d$ of pixels in the boundary region; this is where h enters. Intuitively, the particle number variance is set by the average number of particles in the boundary region. Repeating the arguments in the first three paragraphs of Section I A, but with n replaced by n_B as the only difference, gives the following **Funda-**

mental Equations of HUDLS for cubic windows:

$$\mathcal{V}(L) = \frac{\langle v \rangle}{L^d} \left[\frac{L^d - (L - 2h)^d}{L^d} \right], \quad (13)$$

$$\mathcal{R}(L) = 1 - (1 - 2h/L)^d, \quad (14)$$

$$h(L) = (L/2)\{1 - [1 - \mathcal{R}(L)]^{1/d}\}. \quad (15)$$

Note that the term in square brackets in Eq. (13) equals the right-hand side of Eq. (14), and is the ratio of boundary volume to window volume. As defined earlier, $\mathcal{V}(L)$ is shorthand for the relative volume fraction variance given by Eqs. (5-6) for either Poisson or multinomial statistics, as appropriate; $\langle v \rangle = \langle I \rangle p_o^d$ is the ϕ_i -weighted average particle volume; and $\mathcal{R}(L) \equiv \mathcal{V}(L)/[\langle v \rangle/L^d]$ is the ratio of the variance to that in a random arrangement of the same particles. We emphasize that Eqs. (13-15) are equivalent, and serve to define $h(L)$ in terms of the measured volume fraction variance. The same set of equations holds for point particles in continuum space, and also for spherical measuring windows if L is replaced by diameter.

Before putting this machinery into action, we examine special cases and bounds. First, note that $\mathcal{V}(L) = \langle v \rangle/L^d$ and $\mathcal{R}(L) = 1$ are recovered for totally random patterns, where h equals $L/2$ by construction. These are upper bounds [28]. For not-totally-random patterns with liquid-like long-range fluctuations, $\mathcal{V}(L)$ also scales as $1/L^d$; then $\mathcal{R}(L)$ goes to a constant and $h(L)$ scales like L but with a proportionality constant less than $1/2$. For separated particles at small L , where there is no more than one particle per window, the expectations are

$$\mathcal{V}(L) = \frac{\langle v \rangle/L^d - \phi}{1 - \phi/\langle I \rangle}, \quad (16)$$

$$\mathcal{R}(L) = \frac{1 - \phi L^d/\langle v \rangle}{1 - \phi/\langle I \rangle}, \quad (17)$$

$$h(L) = \frac{L}{2} - \frac{L}{2} \left[\frac{\phi L^d/\langle v \rangle - \phi/\langle I \rangle}{1 - \phi/\langle I \rangle} \right]^{1/d}. \quad (18)$$

These are lower bounds; however, at larger L where Eq. (18) falls below $p_o/2$, the lower bounds are given instead by the fundamental equations evaluated at $h = p_o/2$ (maximally hyperuniform). In the continuum limit, the $\phi/\langle I \rangle$ terms vanish. For patterns where $h(L) \ll L$ holds and fluctuations are hence only near the measuring window surfaces, then expansion of the fundamental equations gives

$$\mathcal{V}(L) = 2d \frac{\langle v \rangle h(L)}{L^{d+1}}, \quad (19)$$

$$\mathcal{R}(L) = 2d \frac{h(L)}{L}, \quad (20)$$

$$h(L) = \frac{1}{2d} \mathcal{R}(L) L. \quad (21)$$

The case of strong hyperuniformity is $\mathcal{V}(L) \sim \langle v \rangle / L^{d+1}$; this corresponds to a constant hyperuniformity disorder length $h(L) = h_e$. Note that the very form of the $\mathcal{V}(L) \sim \langle v \rangle / L^{d+1}$ scaling demands the existence of a new length scale, h_e , in order to become dimensionally correct: $\mathcal{V}(L) \propto \langle v \rangle h_e / L^{d+1}$.

For further intuition and direct connection with the surface coefficient Λ of Ref. [1], we repeat the derivation of Eq. (19) for the restricted case of monodisperse particles of volume v , Poisson statistics for boundary fluctuations, arbitrary window shapes, and $h \ll V_\Omega^{1/d}$. Then the boundary volume equals window surface area A_Ω times h , and the number variance equals the average number of particles in $A_\Omega h$:

$$\sigma_N^2 = \frac{A_\Omega h \phi}{v}. \quad (22)$$

The volume fraction variance is therefore

$$\sigma_\phi^2 \equiv \left(\frac{v}{V_\Omega} \right)^2 \sigma_N^2 = \frac{v A_\Omega h \phi}{V_\Omega^2}, \quad (23)$$

and the relative variance is

$$\mathcal{V}(L) \equiv \frac{\sigma_\phi^2}{\phi^2} = \frac{v A_\Omega h}{V_\Omega^2}. \quad (24)$$

This reduces to Eq. (19) for hypercubic windows, where $V_\Omega = L^d$ and $A_\Omega = 2dL^{d-1}$. In Ref. [1] a surface coefficient is defined by $\sigma_N^2 = \Lambda(R/b)^{d-1}$ and tabulated for hyperspherical windows of radius R and monodisperse

crystals with lattice spacing b . Thus Λ and h_e are directly related by $\Lambda(R/b)^{d-1} = A_\Omega h_e \phi / v$.

As another aside, it might be tempting to use the fundamental equations to analyze actual images of extended particles – rather than their central-pixel representation. We warn that this is correct only for sufficiently large windows. Ref. [29] generalizes our HUDLS analysis approach to extended particles, and corrects upon Eq. (13) at small L by accounting for particles that lie partially inside and partially outside the measuring windows.

D. Multinomial Expectations in $d = 2$

In the following sections we measure the area fraction variance for simulated two-dimensional patterns of non-overlapping pixel particles. So multinomial statistics are appropriate, and for reference the three quantities of interest are

$$\mathcal{V}(L) = \frac{\sigma_\phi^2(L)}{\phi[1 - \phi/\langle I \rangle]} = \frac{4\langle a \rangle(L - h)h}{L^4}, \quad (25)$$

$$\mathcal{R}(L) = \frac{\sigma_\phi^2(L)L^2}{\sigma_\phi^2(p_o)p_o^2}, \quad (26)$$

$$h(L) = (L/2)[1 - \sqrt{1 - \mathcal{R}(L)}], \quad (27)$$

where $\langle a \rangle = \langle I \rangle p_o^2$ is the area-fraction weighted average particle area. These equations will be used to deduce $\mathcal{R}(L)$ and $h(L)$ from measurements of $\mathcal{V}(L)$. Results will be compared with the following bounds:

$$\max \left\{ \frac{\langle a \rangle / L^2 - \phi}{1 - \phi/\langle I \rangle}, \frac{\langle I \rangle (2L/p_o - 1)}{(L/p_o)^4} \right\} \leq \mathcal{V}(L) \leq \frac{\langle a \rangle}{L^2}, \quad (28)$$

$$\max \left\{ \frac{1 - \phi L^2 / \langle a \rangle}{1 - \phi/\langle I \rangle}, \frac{\langle I \rangle (2L/p_o - 1)}{(L/p_o)^2} \right\} \leq \mathcal{R}(L) \leq 1, \quad (29)$$

$$\max \left\{ \frac{L}{2p_o} \left(1 - \sqrt{\frac{\phi L^2 / \langle a \rangle - \phi/\langle I \rangle}{1 - \phi/\langle I \rangle}} \right), \frac{1}{2} \right\} \leq \frac{h(L)}{p_o} \leq \frac{L}{2p_o}. \quad (30)$$

The upper bounds are for totally random patterns; these will be plotted as red dashed lines. The lower bounds are either the small- L expectation for separated particles, or the expectation for maximally-hyperuniform as defined by $h = p_o/2$; these will be plotted as green dotted curves. Note that, with substitution of $\langle a \rangle = \langle I \rangle p_o^2$, the bounds are all algebraic functions of the dimensionless variable $x = L/p_o$. For small L , and for very disordered patterns, it can happen that the measured variance exceeds the upper bound in accord with expected statistical uncertainty. Then \mathcal{R} is greater than one and h cannot be deduced from Eq. (27); instead, we take $h = (L/2)[1 + \sqrt{\mathcal{R} - 1}]$ as a reasonable way to show data points and error bars that

overlap the upper bound.

III. BINARY PIXEL PARTICLES

We now apply the above methods to characterize three different types of two-dimensional binary pixel patterns created by simulation. Here the image sizes are all $8600p_o \times 8600p_o$, somewhat larger than typical digital video images. Pixel values are either $I(x, y) = 0$ (empty) or $I(x, y) = 1$ (one particle); therefore multinomial statistics reduce to binomial statistics, all particle areas are $a = p_o^2$, and the area fraction equals the fraction of pix-

els that contain a particle. The area fraction variance $\sigma_\phi(L)^2$ is measured as described above using $L \times L$ measuring windows, where L is varied from p_o to at least $4300p_o$, half the image width.

A. Random Binomial Patterns

A simple algorithm is to flip each pixel from zero to one with probability equal to the desired area fraction, ϕ . This creates a random “binomial” pixel pattern, with an actual area fraction that is approximately ϕ . To create a binomial pattern with exactly the desired area fraction, we instead select a pixel at random, set its value to one, and repeat until the desired number of pixel particles is reached. Such binomial patterns are the pixelated analogue of random Poisson patterns for point particles. While points literally occupy zero volume in Poisson patterns, binomial patterns have $0 < \phi < 1$ and reduce to a Poisson pattern at both extremes.

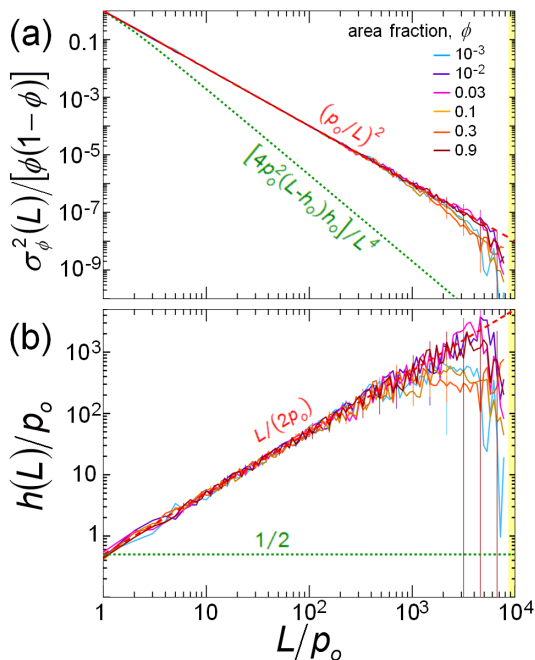


FIG. 3: (Color online) Area fraction variance (a) and hyperuniformity disorder length (b) versus measuring window size, for simulated two-dimensional random binomial pixel patterns with total area fraction ϕ as labeled. The system size is $8600p_o \times 8600p_o$ where p_o is the pixel width; this is indicated by yellow shading. The results agree well with expectation of (a) $\sigma_\phi^2(L)/[\phi(1-\phi)] = (p_o/L)^2$ and (b) $h = L/2$. However there are significant finite-size effects for large L , so all further simulated patterns will be analyzed only up to half the system size.

Simulation results for the relative variance $\mathcal{V}(L) = \sigma_\phi^2(L)/[\phi(1-\phi)]$ are plotted versus L/p_o in Fig. 3a for six different target area fractions, ranging widely from $\phi = 10^{-3}$ up to $\phi = 0.9$. All data collapse together

and match the expectation $(p_o/L)^2$ to within statistical uncertainty for L less than about half the image width, where finite-size effects are expected to become strong [11]. The corresponding hyperuniformity disorder lengths, deduced from Eq. (27), are plotted underneath in Fig. 3b. All results collapse together and match the expectation $h = L/2$ to within statistical uncertainty, as long as finite-size effects are absent. The good agreement between simulation and expectation validates our calculations as well as our image analysis procedures. Note in particular that the factor of $(1-\phi)$ in the relative variance is crucial for obtaining good collapse at high packing fractions.

B. Vacancy Patterns

Another type of disordered pattern can be created by randomly removing a fraction f of particles from a crystalline lattice. Here we study such vacancy patterns made from pixel particles on a two-dimensional triangular lattice that is rotated by 14° with respect to the image grid. The rotation, and the triangular pattern, help smooth out irregular features in the area fraction distributions that would be especially strong for a square lattice with square measuring windows. We take the lattice spacing to be $b = 30p_o$, which gives the area fraction of the base crystal as $\phi_c = \sqrt{4/3}(p_o/b)^2 = 0.00128$. When a fraction f of sites are vacant, the area fraction decreases to $\phi = (1-f)\phi_c$. These are small compared to 1, so we may neglect factors of $(1-\phi)$ and our simulations are effectively in the continuum limit with Poisson statistics.

Simulation results for the relative variance are shown in Fig. 4a for several values of f . Each curve represents an average of 20 independent runs. All data initially decay as $(p_o/L)^2$, just like for random patterns. As L increases toward the lattice spacing, the variance data fall below the $(p_o/L)^2$ upper bound, more quickly for smaller f . This behavior matches well the lower bound given by the small- L separated-particle expectation of Eq. (28). For $f > 0$, the final decay is not as steep and appears to scale as $\sim 1/L^2$. At small L there are prominent oscillations set by the lattice spacing. For large L and $f = 0$, the data appear to eventually approach and hug the maximally-hyperuniform bound. These trends are easier to see in terms of the corresponding variance ratios, $\mathcal{R}(L)$, plotted underneath in Fig. 4b. In particular, it's more apparent that the initial decay is in good agreement with the small- L separated-particles expectation of Eq. (29). And for large- L , before finite-size effects become strong, the $\mathcal{R}(L)$ data appear to approach a constant that equals the fraction f of vacancies. These features may all be readily understood, next, taking the relative variance of the perfect crystal as a given and using the binomial distribution to treat vacancies.

To model the volume fraction variance of vacancy patterns in d -dimensions, first note that the number of particles in a particular L^d measuring window is equal to the

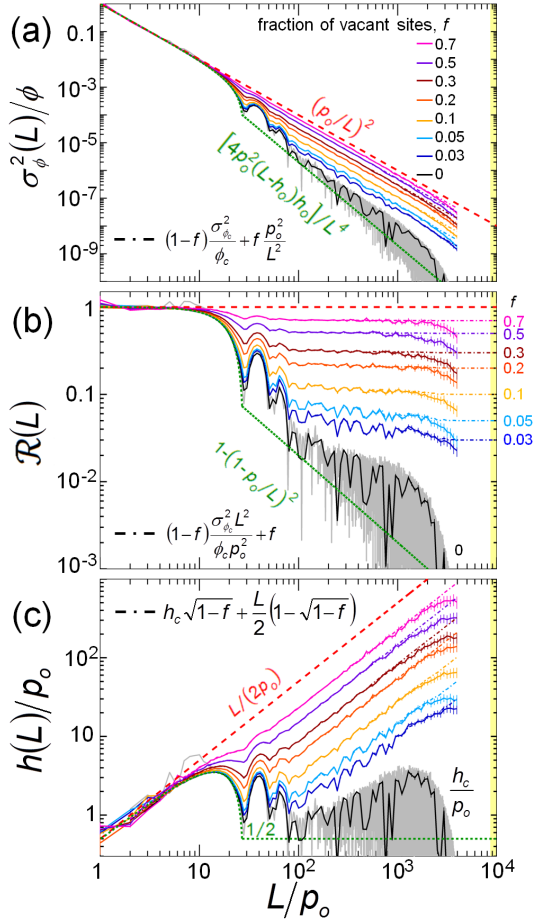


FIG. 4: (Color online) Area fraction variance (a), variance ratio (b), and hyperuniformity disorder length (c) versus measuring window size, for simulated two-dimensional triangular crystals with a specified fraction f of random vacancies. Here ϕ is the area fraction, p_o is the pixel width, $8600p_o$ is image width (yellow shading), and $b = 30p_o$ is the lattice spacing. Measurements are made at all integer L/p_o for $f = 0$ (gray curve), but only at a select subset for $f > 0$. Results for $f = 0$ at the same subset of L/p_o are shown by black curves. The dash-dotted curves correspond to our model, Eq. (31); these curves appear to stop below a certain L , but in fact they are present but entirely covered by the data curve. The red and green dashed curves represent the upper and lower bounds given by Eqs. (28-30). The separated-particle bound is shown only for the crystal.

number $N = N_c - N_v$ of enclosed crystal lattice sites minus the number of vacancies. Therefore, the average over a large set of L^d measuring windows is $\bar{N} = \bar{N}_c - \bar{N}_v$, and the variance is $\sigma_N^2 = \sigma_{N_c}^2 - 2(\bar{N}_c \bar{N}_v - \bar{N}_c \bar{N}_v) + \sigma_{N_v}^2$. For the crystal lattice sites, we may write $\bar{N}_c = \sum n p_n$ and $\bar{N}_c^2 = \sum n^2 p_n$ where p_n is the probability of finding n sites in a randomly placed L^d window. If there are n sites in a window, then the probability for k of them to be vacant is given by the binomial distribution as $q_k = \{n!/[k!(n-k)!]\} f^k (1-f)^{n-k}$. By direct summation, this gives $\bar{N}_v = \sum_n \sum_{k=0}^n (k q_k) p_n = f \bar{N}_c$ as expected.

Similar computation gives $\bar{N}_v^2 = \sum_n \sum_{k=0}^n (k^2 q_k) p_n = f(1-f) \bar{N}_c + f^2 \bar{N}_c^2$ and $\bar{N}_c \bar{N}_v = \sum_n \sum_{k=0}^n (n p_n) (k q_k) = f \bar{N}_c^2$. Plugging these into the expression for σ_N^2 , multiplying by $(p_o/L)^{2d}$ to convert to area fraction variance, and dividing left and right hand sides by $\phi = (1-f)\phi_c$, gives a final result that is surprisingly simple:

$$\frac{\sigma_\phi^2(L)}{\phi} = (1-f) \frac{\sigma_{\phi_c}^2(L)}{\phi_c} + f \left(\frac{p_o}{L}\right)^d. \quad (31)$$

Thus the relative variance of a vacancy pattern is the weighted average of that for a perfect crystal plus that for a random binomial pattern. The variance ratio and hyperuniformity length can then be expressed in terms of those quantities for the perfect crystal as

$$\mathcal{R}(L) = (1-f) \mathcal{R}_c(L) + f, \quad (32)$$

$$h(L) = h_c(L)(1-f)^{1/d} + \frac{L}{2}[1 - (1-f)^{1/d}]. \quad (33)$$

Again these are weighted averages, but for $h(L)$ the weighting is not linear in f . For $f \rightarrow 0$ the crystal results are recovered; and for $f \rightarrow 1$, as all particles are removed, the pattern becomes totally random with $h = L/2$.

Since crystals are hyperuniform, $\sigma_\phi^2(L)$ should decay faster than $1/L^d$ and therefore the second terms in Eqs. (31-33) will eventually dominate at large L . Then the large- L asymptotic behavior is

$$\frac{\sigma_\phi^2(L)}{\phi} = f \left(\frac{p_o}{L}\right)^d, \quad (34)$$

$$\mathcal{R}(L) = f, \quad (35)$$

$$h(L) = (L/2)[1 - (1-f)^{1/d}]. \quad (36)$$

For our $d = 2$ simulations, these predicted scalings with L are readily seen in Fig. 4. As a stronger test, $\mathcal{V}_c(L)$, $\mathcal{R}_c(L)$, and $h_c(L)$ were extracted from the $f = 0$ data, and the resulting expectations based on Eqs. (31-33) are plotted as dot-dashed curves. These match the $f > 0$ simulation data very well, until finite-size effects become strong. At our level of precision, such effects are noticeable even for L down to one tenth the sample width.

Based on the $\mathcal{R}(L) \rightarrow f$ result for vacancy patterns, we may interpret $\mathcal{R}(L)$ for general systems not just as a kind of randomness index but also more specifically as the fraction of space available for density fluctuations at wavelength L .

C. Einstein Patterns

Our previous examples all exhibited $h \sim L$ at long length scales, so now we study a type of disordered pattern that has no long-range density fluctuations and is unquestionably hyperuniform. In particular, we consider two-dimensional pixel particles that are individually displaced by a Gaussian-distributed amount from a crystalline lattice. We dub these ‘‘Einstein patterns’’

in honor of Einstein’s simplified model for the heat capacity of solids, where each atom is harmonically bound to a fixed lattice site. This may be compared with the “shuffled lattice” patterns of Ref. [19], where each object is placed randomly inside a cubic volume surrounding a lattice site. For our simulations we again use a triangular lattice with spacing $b = 30p_o$, which is large enough that factors of $(1 - \phi)$ may be dropped. Displacements in each dimension are randomly drawn from a Gaussian distribution specified by the dimensionless parameter $\delta = x_{rms}/b = y_{rms}/b$. For large δ , it is sometimes necessary to repeat a trial placement until an unoccupied pixel is found.

Example patterns are shown in Fig. 2 for one δ , and in Fig. 5 for a sequence of increasing δ . We include $\delta = 0$, although the crystal isn’t perfect due to pixilation effects that must be large for $x_{rms} < p_o$, i.e. here for $\delta < 0.02$. Note that the underlying crystal is evident for $\delta \ll 1$, including the case $\delta = 0.15$ that corresponds to the Lindemann criterions for melting. For $\delta > 0.5$, i.e. for displacements larger than about $1/2$ lattice spacing, it is difficult to detect the lattice by eye. Nevertheless, even for very large δ , the particles are still bound to a lattice and hence cannot have long-range density fluctuations. This will be reflected intuitively, as shown next, in terms of the behavior of the hyperuniformity disorder length.

We next measure the area fraction variance as above, and display the results in Fig. 6a. The corresponding variance ratios and hyperuniformity disorder lengths are shown underneath, in Figs. 6b-c. Each curve represents an average of fourteen Einstein patterns, all with different displacements, and where the underlying lattice was rotated by a different angle $\theta \in \{1^\circ, 2^\circ, \dots, 14^\circ\}$. This helps minimize artifacts from pixelation as well as from commensuration of the lattice with the square measuring windows. Such artifacts are still present for the “zero-temperature” crystal case, $\delta = 0$, where the relative variance is computed at each integer L/p_o in the range $1 - 4300$. Since these artifacts vanish for large enough δ , we then measure at fewer window sizes.

The observed behavior is as follows. At small L , the relative variances for all δ match well with the random binomial expectation with $h = L/2$. For increasing L , each data curve eventually falls below the binomial expectation – sooner for smaller δ , i.e. for less disorder. The $\delta = 0$ crystal results show pronounced oscillations with features located according to lattice spacing. At large enough L such features are less regular, though are still present, and the variance approaches the lower bound of maximally hyperuniform. For $\delta > 0$, the variance results similarly show a final decay of $1/L^3$ but with a numerical prefactor that grows with δ . Einstein patterns are therefore all hyperuniform, but not maximally so.

The extent of hyperuniformity may be diagnosed in terms of the behavior seen in Fig. 6c for the hyperuniformity disorder length at large L . In particular a signature of strong hyperuniformity is that h becomes a constant, and the degree of hyperuniformity may be judged by the

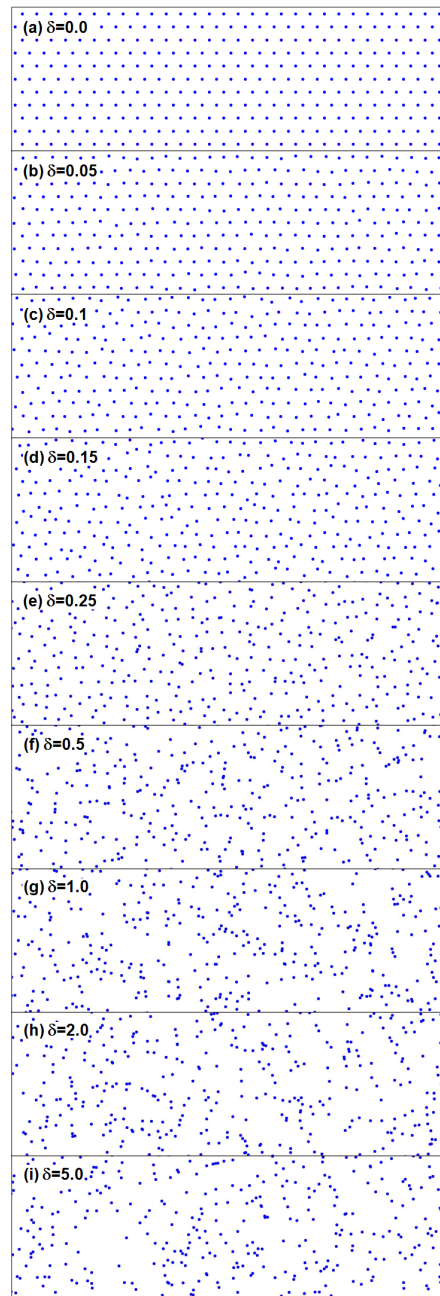


FIG. 5: (Color online) Example Einstein patterns for different root-mean-square displacements, labeled by the value of $\delta = x_{rms}/b = y_{rms}/b$. Here the lattice spacing b is 30 pixels, and for clarity the particles are shown as dots that are many pixels across. By eye, patterns (f)-(i) are nearly indistinguishable and appear quite random.

size h_e of this constant. To deduce h_e we either average the large- L results or we fit to

$$h(L) = \frac{h_e(e^{\alpha L/h_e} - 1)}{2\alpha + (e^{\alpha L/h_e} - 1)}, \quad (37)$$

where α and h_e are adjustable parameters. This empirical function has an exponential crossover from $h = L/2$

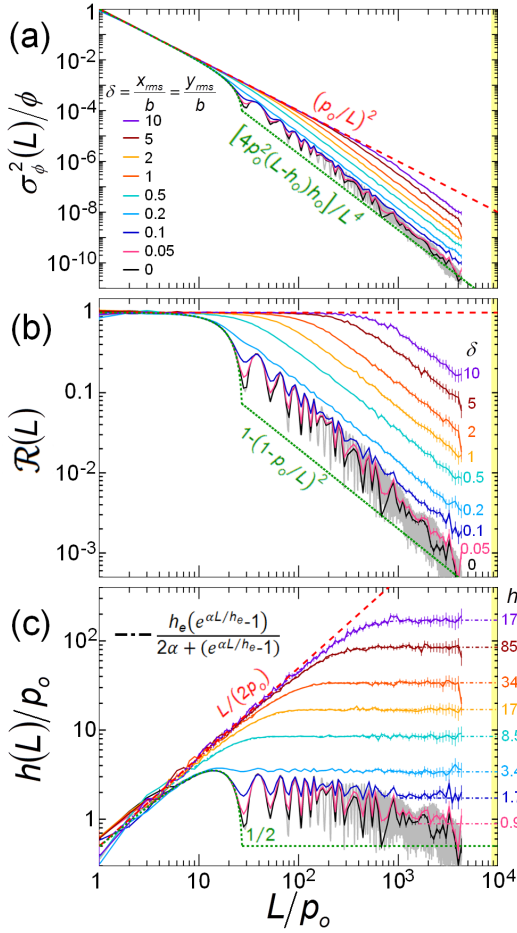


FIG. 6: (Color online) Area fraction variance (a), variance ratio (b), and hyperuniformity disorder length (c) versus measuring window size, for simulated two-dimensional Einstein patterns with specified dimensionless root-mean-square displacement, δ . Here ϕ is the area fraction, p_o is the pixel width, $8600p_o$ is image width (yellow shading), and $b = 30p_o$ is the lattice spacing. In (c) an empirical fitting function is shown, which may be used to estimate the constant value h_e of the hyperuniformity length at large L . Note that h_e is roughly one-half the root-mean-square displacement in each dimension. Measurements are made at all integer L/p_o for $\delta = 0$ (gray curve), but only at a select subset for $\delta > 0$. Results for $\delta = 0$ at the same subset of L/p_o are shown by black curves. The red and green dashed curves represent the upper and lower bounds given by Eqs. (28-30).

at small L to $h = h_e$ at large L . It is seen in Fig. 6c to match the data quite well, except for the oscillations coming from the crystal. The fitting results for α are constant: $\alpha = 0.77 \pm 0.02$. The fitting results for h_e are plotted versus $\delta = x_{rms}/b = y_{rms}/b$ in Fig. 7, and are well described by

$$h_e = \sqrt{(p_o/2)^2 + (cx_{rms})^2}, \quad (38)$$

with $c \approx 1/2$. Thus, h_e is approximately one-half the total root-mean-square displacement in each dimension,

which comes from the combination of pixelation and Gaussian kicks. This makes intuitive sense by thinking of a fixed measuring window and harmonically bound particles that independently oscillate in time with thermal energy. Number fluctuations are primarily due to the particles whose lattice sites lie within a distance h_e from the boundary of the measuring window. This degree of disorder is not evident to the eye when the root-mean-square displacement is larger than the lattice spacing. But it is readily detected by analyzing the relative variance in terms of the hyperuniformity disorder length.

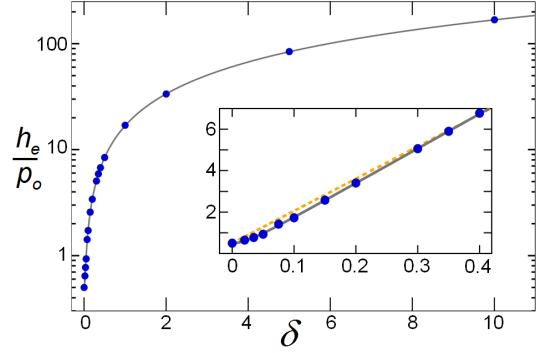


FIG. 7: (Color online) Large- L value of the hyperuniformity disorder length versus dimensionless root-mean square displacement, $\delta = x_{rms}/b = y_{rms}/b$, for the Einstein pattern results shown in Fig. 6b with lattice spacing $b = 30p_o$. The insert is a blow-up of the small- δ region. The fitting function $h_e = \sqrt{(p_o/2)^2 + (c\delta)^2}$ matches the data well with $c = 0.56 \pm 0.01$. By contrast a linear fit for all δ , shown by the dashed yellow line in the inset, fails for small δ .

D. Finite-size effects

The degree to which variance measurements are affected by the finite size of the system can be examined in terms of the ratio of measured to expected variance as a function of L/W where W is the width of the sample. This is shown in Fig. 8 for all three pattern types discussed above. For the binomial and vacancy patterns, we found no trend with area fraction and f , respectively. So results are averaged together. As seen in Fig. 8, finite size effects are generally largest for the binomial patterns and smallest for the Einstein patterns. In all cases, the variance is suppressed and the amount increases with disorder and the range of the density fluctuations. Thus Einstein patterns with small δ are least affected, and binomial patterns are most affected. For general guidance, systematic errors are no more than about 1% for $L < 0.05W$ and no more than about 10% for $L < 0.2W$. But for strongly hyperuniform systems the systematic errors can be considerably less.

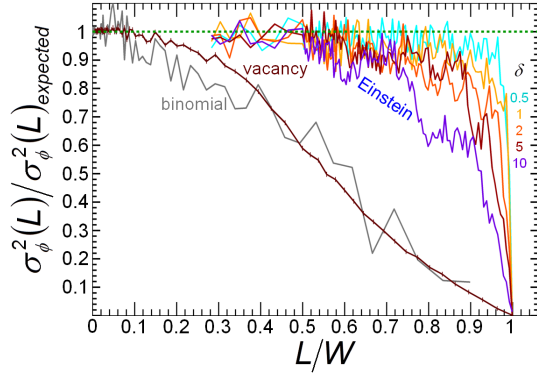


FIG. 8: (Color online) Ratio of simulated to expected area fraction variance versus L/W where L is the width of the measuring boxes and $W = 8300p_o$ is the width of the sample. Data are from previous figures. As seen, the effect of finite sample size is to suppress the variance at large L . The longer the range of the density fluctuations, the larger the effect.

IV. GRAYSCALE PIXEL PARTICLES

In this section we apply the same analysis approach to three different types of simulated two-dimensional patterns where the pixel particles are now polydisperse, with a range of intensity values I_i that differ from one. The particle areas are $a_i = I_i p_o^2$, and the expectations are determined by the weighted average $\langle a \rangle = \langle I \rangle p_o^2 = \sum \phi_i a_i / \phi$. Area fractions are given by the sum of intensities divided by number of pixels, and the variance $\sigma_\phi^2(L)$ is measured using $L \times L$ square windows.

A. Random Multinomial Patterns

The first test is for totally random multinomial patterns made either from equal *numbers* of three species of pixel particle with grayscale intensities $\{1, 2, 3\}$, or for patterns made from equal *area fractions* of four species of pixel particle with grayscale intensities $\{1, 2, 4, 8\}$. Two patterns are created for each mixture, with different total area fractions of 0.1 and 0.9, by randomly choosing pixels values as $\{0, I_1, I_2, \dots\}$ with appropriate probabilities $\{1 - \sum q_i, q_1, q_2, \dots\}$. The relative variance is expected to be $\mathcal{V}(L) = \sigma_\phi^2(L) / [\phi(1 - \phi/\langle I \rangle)] = \langle I \rangle (p_o/L)^2$, where the $\phi/\langle I \rangle$ term cannot be neglected. We therefore plot $\mathcal{V}(L)/\langle I \rangle$ in Fig. 9 and compare with $(p_o/L)^2$. As expected, the data for all four patterns perfectly collapse to this power law until finite size effects become noticeable at large L . This validates the multinomial statistics calculation, and serves to emphasize that the relevant average particle area is set by ϕ_i -weighting.

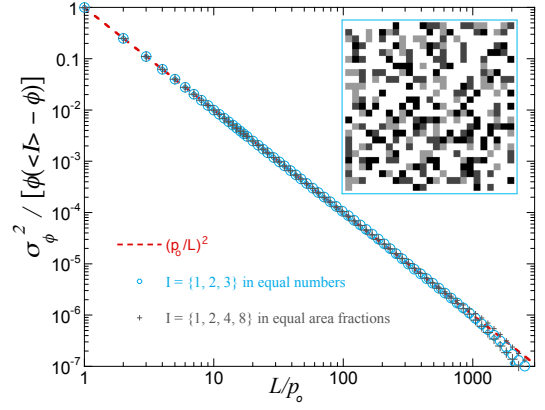


FIG. 9: (Color online) Normalized area fraction variance versus measuring window size for four random multinomial pixel patterns. Two mixtures of pixel particle species are used, as labeled, and patterns for each are created with total area fractions of 0.1 (smaller symbols) and 0.9 (larger symbols). A small portion of the $\{1, 2, 3\}$ system with $\phi = 0.9$ is shown in the inset. For all four patterns, the relative variance results agree well with the $(p_o/L)^2$ expectation shown by the dashed red line.

B. Bidisperse Squares: Overlapping

The second test is for a grayscale Poisson pattern made from of a 50:50 mixture of two different size square particles, $a_1 = (10p_o)^2$ and $a_2 = (20p_o)^2$, and the associated central pixel pattern. Extended particles are repeatedly added at random locations, such that each covered pixel is incremented by +1, until the respective area fractions reach $\phi_1 = 0.2$ and $\phi_2 = 0.8$. The total area fraction is then $\phi = 1$ and the ϕ_i -weighted average particle area is $\langle a \rangle = 340p_o^2$. Note that particle-particle overlaps are freely allowed, and that the value of each pixel is equal to the number of particles that cover it. In parallel we construct the central pixel pattern representation of the same configurations, where the center pixel for each particle of each species is incremented by $I_1 = 100$ or $I_2 = 400$. For both image types, the local area fraction is given by the sum of pixel values per unit area. Small examples of extended-particle and corresponding central-pixel images are shown in the insets of Fig. 10. Even though the total area fraction is one, the particles are large enough that the central pixel pattern is quite dilute and has intensity values of only $\{0, I_1, I_2\}$. Therefore, the distinction between Poisson and multinomial statistics for the central pixel representation can be neglected.

The main plot of Fig. 10 shows the relative area fraction variance for the two different representations of the same arrangement. The results for the central-pixel image are well described by the prediction $\sigma_\phi^2/\phi = \langle a \rangle/L^2 = 340(p_o/L)^2$. By contrast, the relative variance for the extended particle image merges onto $\langle a \rangle/L^2$ only for L much greater than the particle sizes; in this limit the two representations give identical results. At small- L the relative variance for the extended particle image has

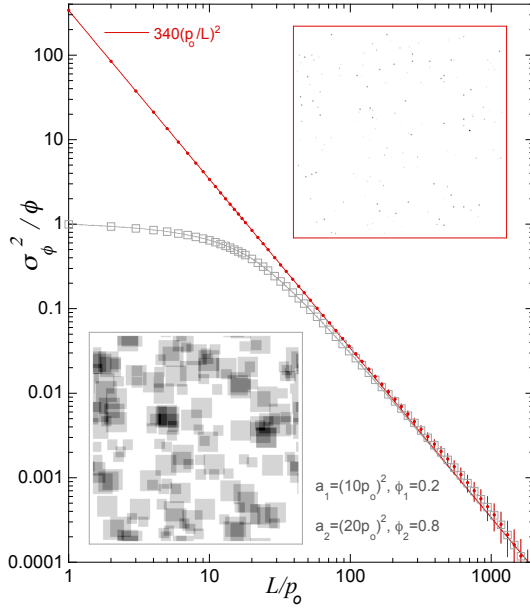


FIG. 10: (Color online) Relative variance versus measuring window size for a random arrangement of a 50:50 bidisperse mixture of extended square particles (gray squares) and the corresponding central-pixel representation (red dots). The actual image sizes are $4300p_o \times 4300p_o$; much smaller versions are shown as insets. Particle sizes and total area fractions are labeled. The ϕ_i -weighted average area of the two species is $\langle a \rangle = 340p_o^2$. The red line is the $\langle a \rangle/L^2$ expectation. The gray curve is from Ref. [29].

a slower decay starting from $\sigma_\phi^2/\phi = 1$ at $L = p_o$. This reflects the particle shape as well as the random arrangement [29], whereas the decay for central pixel patterns reflects only the nature of the arrangement.

C. Bidisperse Squares: Non-Overlapping

Lastly, we use a central-pixel representation to diagnose the uniformity of simulated 50:50 bidisperse mixtures of $a_1 = (10p_o)^2$ and $a_2 = (20p_o)^2$ square particles with $\langle a \rangle = 340p_o^2$ as a function of total area fraction. This system is like in Fig. 10, but with one major difference: Random trial locations are now rejected if any particle-particle overlap occurs. An example of the resulting binary pattern of extended particles is shown in the inset of Fig. 11. As before, central-pixel representations are simultaneously made and then analyzed.

Results for the relative variance, the variance ratio, and the corresponding hyperuniformity disorder lengths, are collected in Fig. 11. The general behavior is an amalgam of key features seen for the vacancy and Einstein patterns. Like the latter, the system matches the separated-particles expectations at small L followed by a developing plateau of constant $h(L)$ – but only out to intermediate L . At larger L , the scaling then crosses over to that seen for the vacancy patterns with long-range den-

sity fluctuations: $\mathcal{V}(L) \sim 1/L^2$, $\mathcal{R}(L) \sim \text{constant}$, and $h(L) \sim L$. With increasing ϕ , the plateau becomes more pronounced and the long-range density fluctuations decrease. If the length of the plateau were to diverge as ϕ increases toward random-close packing, then the onset of jamming would be accompanied by the development of strong hyperuniformity like in an Einstein pattern. This is the topic of a following paper [30], where different packing protocols are used to generate bigger systems with higher packing fractions than are accessible here. For now, Fig. 11 serves as demonstration of the HUDLS analysis method for polydisperse hard or soft particles and, hopefully, also whets the appetite for its larger scale uses.

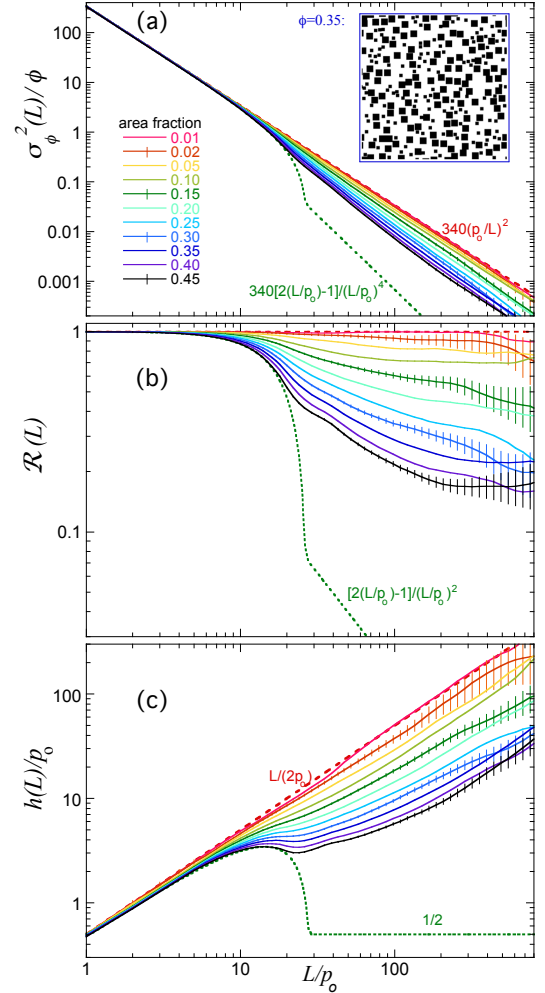


FIG. 11: (Color online) Relative variance (a), variance ratio (b), and hyperuniformity disorder length (c) versus measuring window size for the central-pixel representation of arrangements of non-overlapping $a_1 = (10p_o)^2$ and $a_2 = (20p_o)^2$ square particles, in equal numbers, like in the inset, for several area fractions. For clarity, error bars are plotted on only a few data as sets marked in the legend. The red and green broken curves represent the upper and lower bounds given by Eqs. (28-30). The separated-particle bound is shown only for the largest packing fraction.

V. CONCLUSION

In summary, we have introduced a powerful way to analyze $\sigma_\phi^2(L)$ versus L results and thereby diagnose the extent of hyperuniformity in terms of two real-space spectra: (a) the ratio $\mathcal{R}(L)$ of the variance for the pattern in question to that for a random arrangement of the same set of particles, and (b) the hyperuniformity disorder length $h(L)$. Whereas $\mathcal{R}(L)$ is a kind of randomness index that can also be interpreted as the fraction of space available for density fluctuations at wavelength L , the value of $h(L)$ specifies the distance from the boundary of $V_\Omega = L^d$ measuring windows over which fluctuations are important. This connects directly to the original idea that hyperuniform systems are controlled by particles on the surface of the measuring windows [1], and it makes the large- L scaling $\sigma_\phi^2(L) \propto \langle v \rangle h / L^{d+1}$ dimensionally correct. This applies equally well to liquid-like systems having long-range density fluctuations with $\mathcal{R}(L) \rightarrow \text{constant}$ and $h(L) \sim L$, as well as to strongly hyperuniform systems with $\mathcal{R}(L) \sim 1/L$ and $h(L) \rightarrow \text{constant}$, and to systems anywhere between.

Whereas prior work focuses on the form of the large- L scaling for diagnosing hyperuniformity, an important feature of our work is to bring meaning to the *value* of the variance, though intuitive interpretation of the corresponding values of $\mathcal{R}(L)$ and $h(L)$. These quantities are even further useful when compared to the exact bounds we computed for pixelated space, using Poisson and multinomial statistics. The resulting formulae and plots are expressed using the pixel length, p_o , explicitly, so as to be dimensionally correct.

While pixel patterns are natural for digital images coming from experiment, they might seem inappropriate for arrangements of *point* particles in continuous space. In fact they literally apply via the realization that limitations on numerical precision effectively set p_o . But furthermore, there is a well-defined continuum limit where the particle volume $v_i = I_i p_o^d$ remains constant while pixel length vanishes and the intensity diverges. The relevant formulae for a point representation of extended particles with ϕ_i -weighted average volume $\langle v \rangle$, and d -dimensional spherical measuring windows of volume V_Ω and diameter D , then become

$$\mathcal{V}(D) \equiv \sigma_\phi^2(D)/\phi \quad (39)$$

$$\Delta\mathcal{V} = \sqrt{\frac{\langle v^3 \rangle / V_\Omega^3}{S\phi} + \frac{2\mathcal{V}^2}{S-1}} \quad (40)$$

$$\mathcal{R}(D) = \frac{\mathcal{V}(D)D^d}{[\mathcal{V}(D_o) + \phi]D_o^d} \quad (41)$$

$$1 - \phi V_\Omega / \langle v \rangle \leq \mathcal{R}(D) \leq 1 \quad (42)$$

$$h(D) \equiv (D/2)\{1 - [1 - \mathcal{R}(D)]^{1/d}\} \quad (43)$$

where D_o is some smallest chosen measuring diameter that is less than the smallest particle-particle separation in the pattern, and S is the number of independent samples made by the randomly chosen set of measuring windows. The first and last of these are definitions; the others are new results. The upper and lower bounds are respectively set by the calculated variance for the separated-particles and the totally-random arrangements. The large- L scalings are $\mathcal{V}(L) \sim 1/L^{d+\epsilon}$, $\mathcal{R}(L) \sim 1/L^\epsilon$, and $h(L) \sim L^{1-\epsilon}$, where the exponent ϵ ranges from zero (liquid-like) to one (strongly hyperuniform).

Thus our general advances for analysis of particle arrangements include the concepts of \mathcal{R} and h , their intuitive meanings and bounds, and a systematic way to estimate the expected statistical uncertainty – for both pixelated and continuum space. It is unclear how this translates or could be done in terms of the $S(q)$ or $\chi(q)$ spectra. Therefore, we recommend that hyperuniformity be diagnosed from measurement of the real-space spectrum of σ_ϕ^2 versus window size. As standard procedure, essentially-raw variance data are to be plotted in terms of $\mathcal{R} \pm \Delta\mathcal{R}$ and then further analyzed in terms of $h \pm \Delta h$ versus measuring window size.

To verify expectations for random pixel patterns, and to build intuition and examine behavior relative to the bounds, we applied our Hyperuniformity Disorder Length Spectroscopy (HUDLS) approach to a variety of simulated two-dimensional patterns. The results are exactly as expected for both the random binomial and multinomial patterns, as long as L is small enough for finite-size effects to be small. The results for the vacancy and the Einstein patterns can be simply understood, and capture essential features that emerge in a small-scale simulation of a bidisperse mixture of extended particles as the packing fraction is increased. Thus the extent of hyperuniformity in polydisperse systems may now be diagnosed by similarly using HUDLS on larger-scale simulations and experiments. This opens up new lines of research, including study of foams [31] as well as the jamming transition [30], for both of which polydispersity is often present and acts to suppress crystallization.

Acknowledgments

We thank Jim Sethna for suggesting a crystal with vacancies as a model system exhibiting long-range Poissonian density fluctuations. This work was supported by NASA grant NNX14AM99G (primary) and by NSF grants MRSEC/DMR-1120901 and DMR-1305199.

[1] S. Torquato and F. H. Stillinger, Phys. Rev. E **68**, 041113 (2003).

[2] C. E. Zachary and S. Torquato, J. Stat. Mech.: Theory

- and Experiment p. P12015 (2009).
- [3] A. Donev, F. H. Stillinger, and S. Torquato, Phys. Rev. Lett. **95**, 090604 (2005).
 - [4] S. Torquato and F. H. Stillinger, Rev. Mod. Phys. **82**, 2633 (2010).
 - [5] R. Kurita and E. R. Weeks, Phys. Rev. E **82**, 011403 (2010).
 - [6] N. Xu and E. S. C. Ching, Soft Matter **6**, 2944 (2010).
 - [7] L. Berthier, P. Chaudhuri, C. Coulais, O. Dauchot, and P. Sollich, Phys. Rev. Lett. **106**, 120601 (2011).
 - [8] C. E. Zachary, Y. Jiao, and S. Torquato, Phys. Rev. Lett. **106**, 178001 (2011).
 - [9] R. Kurita and E. R. Weeks, Phys. Rev. E **84**, 030401 (2011).
 - [10] A. Ikeda and L. Berthier, Phys. Rev. E **92**, 012309 (2015).
 - [11] R. Dreyfus, Y. Xu, T. Still, L. A. Hough, A. G. Yodh, and S. Torquato, Phys. Rev. E **91**, 012302 (2015).
 - [12] Y. Wu, P. Olsson, and S. Teitel, Phys. Rev. E **92**, 052206 (2015).
 - [13] M. Florescu, S. Torquato, and P. J. Steinhardt, Proc. Nat. Acad. Sci. **106**, 20658 (2009).
 - [14] W. Man, M. Florescu, E. P. Williamson, Y. He, S. R. Hashemizad, B. Y. C. Leung, D. R. Liner, S. Torquato, P. M. Chaikin, and P. J. Steinhardt, Proc. Nat. Acad. Sci. **110**, 15886 (2013).
 - [15] N. Muller, J. Haberko, C. Marichy, and F. Scheffold, Adv. Optical Mater. **2**, 115 (2014).
 - [16] P. Ni, P. Zhang, X. Qi, J. Yang, Z. Chen, and W. Man, Optics Express **24**, 2420 (2016).
 - [17] L. S. Froufe-Pérez, M. Engel, P. F. Damasceno, N. Muller, J. Haberko, S. C. Glotzer, and F. Scheffold, Phys. Rev. Lett. **117**, 053902 (2016).
 - [18] Y. Jiao, T. Lau, H. Hatzikirou, M. Meyer-Hermann, J. C. Corbo, and S. Torquato, Phys. Rev. E **89**, 022721 (2014).
 - [19] A. Gabrielli, M. Joyce, and F. S. Labini, Phys. Rev. D **65**, 083523 (2002).
 - [20] D. Hexner and D. Levine, Phys. Rev. Lett. **114**, 110602 (2015).
 - [21] E. Tjhung and L. Berthier, Phys. Rev. Lett. **114**, 148301 (2015).
 - [22] J. H. Weijs, R. Jeanneret, R. Dreyfus, and D. Bartolo, Phys. Rev. Lett. **115**, 108301 (2015).
 - [23] R. L. Jack, I. R. Thompson, and P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015).
 - [24] A. B. Hopkins, F. H. Stillinger, and S. Torquato, Phys. Rev. E **86**, 021505 (2012).
 - [25] B. L. Lu and S. Torquato, J. Chem. Phys. **93**, 3452 (1990).
 - [26] J. Quintanilla and S. Torquato, J. Chem. Phys. **106**, 2741 (1997).
 - [27] J. Quintanilla and S. Torquato, J. Chem. Phys. **110**, 3215 (1999).
 - [28] As a caveat, the variance increases above this bound if pixel particles aggregate to create extended particles spanning more than one pixel. This can be seen from the Ref. [29] result $\sigma_\phi^2(L)/\phi = \langle v \rangle / L^d$ at large L for a Poisson pattern made from extended particles with ϕ -weighted average volume $\langle v \rangle$. In particular, aggregation causes $\langle v \rangle$ to increase while ϕ remains constant.
 - [29] D. J. Durian, “Hyperuniformity Disorder Length Spectroscopy for Extended Particles” (arXiv:1707.01524).
 - [30] A. T. Chieco, M. Zu, A. J. Liu, N. Xu, and D. J. Durian, “The Uniformity of Soft Disk Configurations Above and Below Jamming” (in preparation).
 - [31] A. T. Chieco, A. E. Roth, S. Torquato, and D. J. Durian, “Hyperuniformity of 2d Foam in Self-Similar Growth Regime” (in preparation).