Precise algorithms to compute surface correlation functions of two-phase heterogeneous media and their applications

Zheng Ma and Salvatore Torquato

Precise algorithms to compute surface correlation functions of two-phase heterogeneous media and their applications

Zheng Ma

Department of Physics, Princeton University, Princeton, New Jersey 08544, USA

Salvatore Torquato

Department of Chemistry, Department of Physics,
Princeton Institute for the Science and Technology of Materials,
and Program in Applied and Computational Mathematics,
Princeton University, Princeton, New Jersey 08544, USA

The quantitative characterization of the microstructure of random heterogeneous and porous media in d-dimensional Euclidean space $\mathbb{R}^d$ via a variety of n-point correlation functions is of great importance, since the respective infinite set determine the effective physical properties of the media. In particular, surface-surface F_{ss} and surface-void F_{sv} correlation functions (obtainable from radiation scattering experiments) contain crucial interfacial information that determine transport properties of the media (e.g., the mean survival time and fluid permeability) and complement the information content of the conventional two-point correlation function. However, the current technical difficulty involved in sampling surface correlation functions has been a stumbling block in their widespread use. We first present a concise derivation of the small-r behaviors of these functions, which are linked to the *mean curvature* of the system. Then we demonstrate that one can reduce the computational complexity of the problem, without sacrificing accuracy, by extracting the necessary interfacial information from a cut of the d-dimensional statistically homogeneous and isotropic system with an infinitely long line. Accordingly, we devise algorithms based on this idea and test them for two-phase media in continuous and discrete spaces. Specifically for the exact benchmark model of overlapping spheres, we find excellent agreement between numerical and exact results. We compute surface correlation functions and corresponding local surface-area variances for a variety of systems, including overlapping spheres, hard spheres in equilibrium, decorated "stealthy" patterns, as well as snapshots of evolving pattern formation processes (e.g., spinodal decomposition). It is demonstrated that the precise determination of surface correlation functions provides a powerful means to characterize a wide class of complex multiphase microstructures.

I. INTRODUCTION

Random heterogeneous media are ubiquitous and arise in many applications in physics, materials science, biology and geophysics. Examples of such media include composites [1], porous materials [2–4], biological tissues [5, 6] and even cosmological structures [7]. The quantitative characterization of the structure via higher-order correlation functions of these complex media is of importance in many fields [1, 8]. In general, an infinite set of correlation functions are required to exactly determine the effective physical properties of the media [1–3]. However, such complete structural information about the medium is generally not available and hence one must settle for reduced information in the form of lower-order correlation functions. The study of these descriptors has proved fruitful and new applications involving these descriptors are constantly coming up, including in reconstructions using state-of-the-art techniques such as neural networks [9, 10].

There are a variety of two-point structural descriptors including the two-point correlation function S_2 [11, 12], two-point cluster function C_2 [13], surface-surface correlation function F_{ss} [14], surface-void correlation function F_{sv} [14] and pore-size density function $P(\delta)$ [1] that are practically accessible via computer simulations or imaging techniques. Among them, the most well-known descriptor is the standard two-point correlation function S_2, which can be obtainable from scattering experiments [15, 16]. This quantity has been employed to characterize the microstructure and physical properties of heterogeneous materials [1], reconstruct the microstructure of heterogeneous materials [11, 12] and recently, to quantify the hyperuniformity of two-phase systems [17–19]. Although knowledge of the two-point correlation function S_2 has proved to be extremely useful, the corresponding correlation functions that characterize the interface of two-phase media such as the specific surface s, surface-surface correlation function $F_{ss}(r)$ and surface-void correlation function $F_{sv}(r)$, which contain crucial structural information, have received considerably less attention, especially F_{ss} and F_{sv}. This is due partly to the fact that these surface correlation functions are not as easy to sample as S_2, which we remedy in this paper, as described in Sec. IV.

While the two-point correlation function $S_2(r)$ contains important structural information, it is usually insufficient to determine both the structure and physical
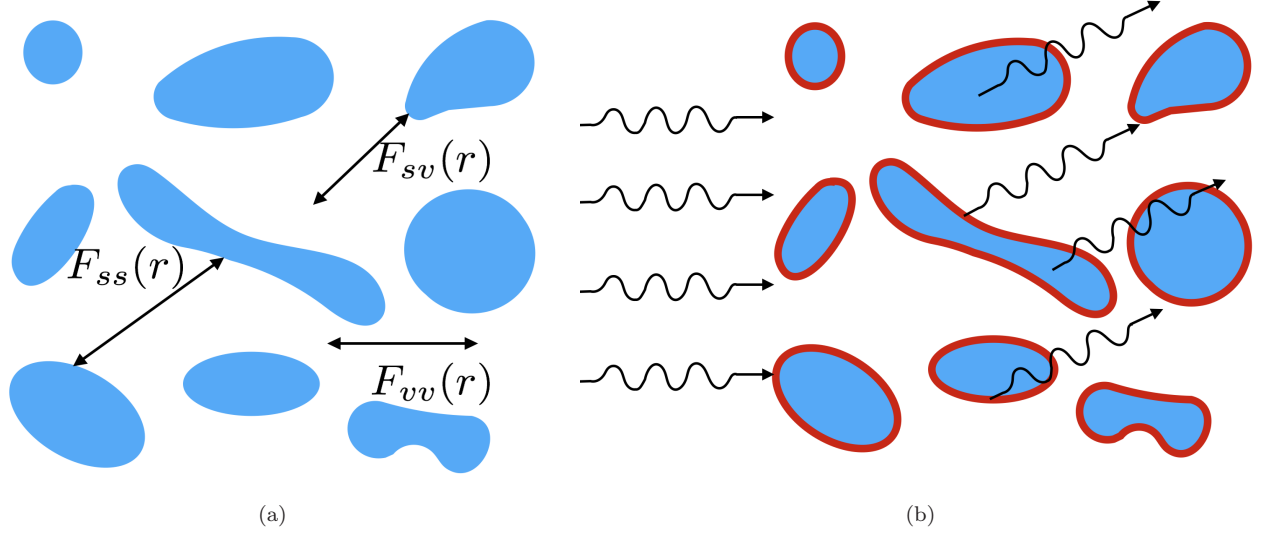
FIG. 1. (Color online.) (a) A schematic two-phase medium showing the surface-surface correlation function $F_{ss}(r)$, surface-void correlation function $F_{sv}(r)$ and void-void correlation function $F_{vv}(r)$ (or $S_2(r)$), where the blue phase is the "solid" phase. (b) Corresponding schematic showing the scattering of radiation by a two-phase medium, where the blue phase indicates the bulk and the red region indicates the interface.

properties of heterogeneous media [1, 13, 20]. It has been shown that supplementing S_2 with surface correlation functions can lead to improved reconstructions of two-phase media [1, 21, 22].

These surface correlation functions determine rigorous upper bounds on the fluid permeability k of porous media and mean survival time τ associated with diffusion-controlled reactions among traps. These two-point "interfacial-surface" bounds have been shown to be much sharper than the so-called two-point "void" bound involving S_2 alone, reflecting the importance of surface correlation functions. For isotropic media in three-dimensional Euclidean space, these involve the following two key integrals:

$$I_1 = \int_0^\infty [\frac{\phi_1^2}{s^2}F_{ss}(r) - \frac{2\phi_1}{s}F_{sv}(r) + F_{vv}(r)]rdr, \quad (1)$$

$$I_2 = \int_0^\infty [F_{vv}(r) - \phi_1^2]rdr. \quad (2)$$

where F_{vv} is just another way to write S_2 for the void phase if we focus on porous media, i.e., $F_{vv} \equiv S_2$, and ϕ_1 is the volume fraction of the void phase. A schematic plot of the correlation functions is shown in Fig. 1(a). For statistically isotropic media, the two-point "interfacial-surface" upper bound for the fluid permeability is given by [1]

$$k \le \frac{2}{3}I_1, \quad (3)$$

while the two-point "void" bound is

$$k \le \frac{2}{3\phi_2^2}I_2, \quad (4)$$

where ϕ_2 is the volume fraction of the solid phase. Similarly, the analogous bounds on the mean survival time are given by [1]

$$\tau \le \frac{I_1}{\phi_1 \mathcal{D}}, \quad (5)$$

and

$$\tau \le \frac{I_2}{\phi_1 \phi_2^2 \mathcal{D}}. \quad (6)$$

where $\mathcal{D}$ is the diffusion coefficient of the reactant. The fact that the key integrals in these bounds on the fluid permeability are the same as those for the mean survival time is more than a coincidence. Indeed, k is rigorously bounded from above in terms τ for general media [23].

Moreover, two-point correlation functions determine local volume-fraction and local surface-area fluctuations as measured by the relevant variances. These variances enable one to generalize the concept of hyperuniformity [17], which was originally conceived in the context of point configurations, namely, it refers to the anomalous suppression of density fluctuations on large length scales [24, 25]. Notably, it is proven that sphere packings will inherit the hyperuniformity of the underlying point pattern [17]. The local volume-fraction variance $\sigma_V^2(R)$ within a spherical observation window of radius R in d-dimensional Euclidean space $\mathbb{R}^d$ is given by [26]

$$\sigma_V^2(R) = \frac{1}{v_1(R)}\int_{\mathbb{R}^d}\chi_V(\mathbf{r})\alpha(r;R)d\mathbf{r}, \quad (7)$$

where

$$\chi_V(\mathbf{r}) = F_{vv}(\mathbf{r}) - \phi_1^2 \quad (8)$$

is the autocovariance function associated with $S_2(\mathbf{r})$, and $v_1(R)$ is the volume of a d-dimensional sphere of radius R, and $\alpha(r; R)$ is the scaled intersection volume, the ratio of the intersection volume of two spherical windows of radius R whose centers are separated by a distance r to the volume of a spherical window. A two-phase system is hyperuniform with respect to volume-fraction variances if $\sigma_V^2(R)$ decreases more rapidly than R^{-d} for large R [17], or equivalently

$$\lim_{|\mathbf{k}| \to 0} \tilde{\chi}_V(\mathbf{k}) = 0, \qquad (9)$$

where $\tilde{\chi}_V(\mathbf{k})$ is the Fourier transform of $\chi_V(\mathbf{r})$. Similarly, the local surface-area variance $\sigma_S^2(R)$ has been defined by [26]

$$\sigma_S^2(R) = \frac{1}{s^2 v_1(R)} \int_{\mathbb{R}^d} \chi_S(\mathbf{r})\alpha(r; R)d\mathbf{r}, \qquad (10)$$

where

$$\chi_S(\mathbf{r}) = F_{ss}(\mathbf{r}) - s^2 \qquad (11)$$

is the autocovariance function associated with $F_{ss}(\mathbf{r})$. A two-phase system is hyperuniform with respect to surface-area variances if $\sigma_S^2(R)$ decreases more rapidly than R^{-d} for large R [17], or equivalently

$$\lim_{|\mathbf{k}| \to 0} \tilde{\chi}_S(\mathbf{k}) = 0, \qquad (12)$$

where $\tilde{\chi}_S(\mathbf{k})$ is the Fourier transform of $\chi_S(\mathbf{r})$. It has been suggested that surface-area fluctuations are more sensitive microstructural measures for heterogeneous media than corresponding volume-fraction fluctuations in some cases [17, 18]. Results obtained in this paper further support this conclusion.

Similar to the two-point correlation function S_2, surface correlation functions can be related to and obtained from the scattering intensity as well [27, 28]. In the most general case that involves scattering from both the bulk and the surface (see Fig. 1(b) for a schematic plot), the scattering intensity can be written as

$$I(k) = c_1\tilde{\mathcal{F}}(S_2) + c_2\tilde{\mathcal{F}}(F_{ss}) + c_3\tilde{\mathcal{F}}(F_{sv}) \qquad (13)$$

where c_1, c_2, c_3 are certain coefficients. When the scattering from the surface is comparable to the bulk, one must consider all these three terms to determine the scattering intensity, while if only bulk or surface scattering is dominant, then one should only care about the corresponding correlation function, this interpretation can potentially provide a general way to understand hyperuniformity in two-phase media.

The rest of the paper is organized as follows: in Sec. II, we provide necessary definitions and background. In Sec. III, we present a concise and simple derivation of the small-r behavior of the two-point surface correlation function, which involves the *mean curvature* of the entire system. In Sec. IV, we introduce and describe a general algorithm that enables the efficient computation

of F_{ss} and F_{sv}. We verify the accuracy of our algorithm by applying it to overlapping spheres for which we have exact results [1]. In Sec. V, we show how to apply the algorithm to treat digitized two-phase media, which is of practical importance. Using Gaussian random fields as an example, we will demonstrate that the image resolution and some drop-out in sampling are crucial in order to obtain reliable results. In Sec. VI we explicitly show results of overlapping spheres, hard spheres in equilibrium and decorated stealthy point patterns. In Sec. VII we explicitly show results of patterns from spinodal decomposition and patterns from the Swift-Hohenberg equation. Using these examples, we demonstrate how surface correlation functions will be very useful for microstructural characterization and can be superior to S_2 in certain cases. Finally, in Sec. IX, we make concluding remarks and discuss the implications of our findings.

II. BACKGROUND AND DEFINITIONS

A two-phase random medium is a domain of space $\mathcal{V} \subseteq \mathbb{R}^d$ that is partitioned into two disjoint regions that make up $\mathcal{V}$: a phase 1 region $\mathcal{V}_1$ of volume fraction ϕ_1 and a phase 2 region $\mathcal{V}_2$ of volume fraction ϕ_2 [1]. The phase indicator function $\mathcal{I}^{(i)}(\mathbf{x})$ for a given realization is defined as

$$\mathcal{I}^{(i)}(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \mathcal{V}_i, \\ 0, & \mathbf{x} \notin \mathcal{V}_i. \end{cases} \qquad (14)$$

For statistically homogeneous media, the volume fraction for phase i

$$\phi_i = \left\langle \mathcal{I}^{(i)}(\mathbf{x}) \right\rangle \qquad (15)$$

is a constant. The two-point correlation function is defined as

$$S_2^{(i)}(\mathbf{x}_1, \mathbf{x}_2) = \left\langle \mathcal{I}^{(i)}(\mathbf{x}_1)\mathcal{I}^{(i)}(\mathbf{x}_2) \right\rangle. \qquad (16)$$

For homogeneous media, this quantity only depends on the relative displacement vector $\mathbf{r}_i \equiv \mathbf{x}_i - \mathbf{x}_1$. The two-point correlation function simplifies as $S_2(\mathbf{x}_1, \mathbf{x}_2) = S_2(\mathbf{r})$, where $\mathbf{r} \equiv \mathbf{x}_2 - \mathbf{x}_1$. If the system is also statistically isotropic, then $S_2(r)$ depends only on the radial distance $r = |\mathbf{r}|$.

The interface indicator function is defined as [1]

$$\mathcal{M}(\mathbf{x}) = |\nabla \mathcal{I}^{(1)}(\mathbf{x})| = |\nabla \mathcal{I}^{(2)}(\mathbf{x})|. \qquad (17)$$

The specific surface is the expected area of the interface per unit volume, and for homogeneous media is simply the ensemble average of the interface indicator function, i.e.,

$$s = \langle \mathcal{M}(\mathbf{x}) \rangle. \qquad (18)$$

The surface-surface correlation function measures the correlation between two points on the interface, and for homogeneous media is defined as

$$F_{ss}(\mathbf{r}) = \langle \mathcal{M}(\mathbf{x})\mathcal{M}(\mathbf{x} + \mathbf{r}) \rangle. \qquad (19)$$

The surface-void correlation function measures the correlation between one point on the interface and the other in the void phase, and for homogeneous media is defined as

$$F_{sv}(\mathbf{r}) = \left\langle \mathcal{M}(\mathbf{x})\mathcal{I}^{(void)}(\mathbf{x} + \mathbf{r}) \right\rangle. \qquad (20)$$

Higher-order surface correlation functions are similarly defined [1], but the focus in this paper will be the two-point varieties.

Closed-form expressions for the two-point surface correlation functions are very limited. The most notable one is for the model of overlapping spheres [1, 8], which is generated by circumscribing spheres of radius a around each point in a Poisson point process with density ρ. The space interior to the spheres is the solid phase and the space exterior is the void phase [1]. For statistically homogeneous overlapping spheres in three dimensions, we have

$$F_{ss}(r) = S_2(r) \qquad (21)$$
$$\left\{ \frac{9\eta^2}{a^2}[1 - (\frac{1}{2} - \frac{r}{4a})\Theta(2a - r)]^2 + \frac{3\eta}{2ra}\Theta(2a - r) \right\},$$

and

$$F_{sv}(r) = \frac{3\eta}{a}[1 - (\frac{1}{2} - \frac{r}{4a})\Theta(2a - r)]S_2(r), \qquad (22)$$

where $r = |\mathbf{r}|$ is a radial distance, $\eta = \rho v_1(a)$ is a reduced density and $\Theta(x)$ is the Heaviside step function. Here,

$$S_2(r) = \exp(\frac{-\eta v_2(r; a)}{v_1(a)}) \qquad (23)$$

is the two-point correlation function for the "void" phase. There relations were first given by Doi [29].

Using the canonical function H_n [1], Torquato derived the following expressions for surface correlation functions for hard spheres

$$F_{ss}(r) = \frac{s}{2r}\Theta(2a - r) + s^2 + \rho^2\delta \otimes \delta \otimes h, \qquad (24)$$

and

$$F_{sv}(r) = s - \frac{s}{2}(1 - \frac{r}{2a})\Theta(2a - r) - s\eta - \rho^2 m \otimes \delta \otimes h, \qquad (25)$$

where $s = d\eta/a$ is the specific surface, δ is the radial Dirac delta function and m is the sphere indicator function. The quantity $h(\mathbf{r})$ is the total correlation function defined as $h(\mathbf{r}) = g_2(\mathbf{r}) - 1$, where $g_2(\mathbf{r})$ is the pair correlation function, and $\otimes$ denotes the convolution of two functions.

The impenetrability constraint alone is not sufficient to specify the hard-sphere model; a hard-sphere system can be in equilibrium or be derived from an infinite number of nonequilibrium ensembles [1]. The pair correlation function is generally not known for nontrivial hard-sphere models for all densities, an exception being the "ghost" random sequential addition packing model

[30]. For $d = 3$, Torquato used the Percus-Yevick approximation and the Verlet-Weis correction to evaluate these functions for statistically isotropic systems of hard spheres in equilibrium [14]. These are useful benchmark results that will be used in Sec VI. To date, numerical evaluations of the surface correlation functions have been limited to hard spheres in equilibrium [14, 31] and maximally random jammed sphere packings [32].

III. SOME THEORETICAL REMARKS ON SURFACE CORRELATION FUNCTIONS

A. The small-r behavior of F_{ss} and F_{sv} in general

Debye and co-workers [15, 16] showed that the slope of $S_2(r)$ at the origin (r=0) is directly proportional to the specific surface s, which enables people to obtain the surface area of the whole system by measuring the tail of a scattering profile. The small-r behavior of the two-point surface correlation functions have been derived previously by taking higher-order derivatives of $S_2(r)$ of a dilated interface and then letting the thickness go to zero [33, 34]. Here we present a much simpler derivation based on a probabilistic interpretation of the surface correlation functions.

We restrict ourselves to the discussion of systems with interfaces that are differentiable everywhere. This assumption enables us to approximate the vicinity of a point on the interface with planes or spheres in the following discussion.

The small-r behavior of the surface-surface correlation function is straightforward to obtain. First, randomly pick a reference point p_0 on the interface (with specific surface s for the entire system). Second, consider a concentric shell with radius r to $r + dr$ around the reference point, then a local specific surface of the shell can be defined as $dr \to 0$. The quantity $F_{ss}(r)$ is then the product of the specific surface s of the system and the average local specific surface over the interfaces (the local specific surface is defined at every point on interfaces, thus can be integrated to compute the average). We present a schematic plot that elucidates the derivation in Fig. 2(a) in three dimensions. When r is very small, the vicinity of p_0 is basically flat for the zeroth-order approximation (see the quadrangle in Fig. 2(a)) and there is no other interface intersecting with the shell. As shown in Fig. 2(a), the area of interface contained in the shell is $2\pi rdr$, and the volume of the shell is $4\pi r^2 dr$, so

$$F_{ss}(r) \sim s \times \left\langle \frac{2\pi rdr}{4\pi r^2 dr} \right\rangle = \frac{s}{2r}, \quad r \to 0. \qquad (26)$$

For $d = 2$, follow the same method we have

$$F_{ss}(r) \sim s \times \left\langle \frac{2dr}{2\pi rdr} \right\rangle = \frac{s}{\pi r}, \quad r \to 0. \qquad (27)$$

Since the zeroth-order approximation of $F_{ss}(r)$ is divergent as $r \to 0$, we will not discuss higher-order finite

correction terms here [33].

The determination of the small-r behavior of the surface-void correlation function is more involved. Again, we randomly pick a reference point p_0 on the interface. Next consider a "test" sphere of radius r centered at the reference point. We denote by $P_{sv}(p_0)$ the following conditional probability: given a point p_0 on the interface, the probability that a uniformly and randomly placed vector $\mathbf{r}$ emanating from p_0 lands in the void phase. The quantity $F_{sv}(r)$ is then the product of the specific surface of the system s and the average of $P_{sv}(p_0)$ over the interfaces, i.e., $F_{sv} = s \langle P_{sv}(p_0) \rangle$. Since it is more convenient to illustrate the basic idea behind the computation in two dimensions and the result can be easily generalized to three dimensions, a schematic plot that elucidates our approach is illustrated in Fig. 2(b) in two dimensions, where the shaded area is the part of the small "test" sphere contained in the solid phase. In $d = 2$, under aforementioned assumptions, we can approximate the interface with a circular arc of radius of curvature $r_c(p_0)$. Then we can work out the probability $P_{sv}(p_0)$ up to the first-order approximation with respect to r, which writes as

$$P_{sv}(p_0) = \frac{1}{2} + \frac{r}{2\pi r_c(p_0)}. \tag{28}$$

Average out this quantity on interfaces, we get the final result

$$F_{sv}(r) = s\left(\frac{1}{2} + \frac{r}{2\pi}\left\langle \frac{1}{r_c(p_0)} \right\rangle\right), \tag{29}$$

where $\langle 1/r_c(p_0) \rangle$ is the average of $1/r_c(p_0)$ on interfaces.

Following the same procedure in three dimensions, we find

$$P_{sv}(p_0) = \frac{1}{2} + \frac{r}{4r_c(p_0)}. \tag{30}$$

Thus we have

$$F_{sv}(r) = s\left(\frac{1}{2} + \frac{r}{4}\left\langle \frac{1}{r_c(p_0)} \right\rangle\right). \tag{31}$$

However, in three dimensions, the curvature varies when the normal plane rotates, and hence here $1/r_c(p_0)$ is to be interpreted to be the *mean curvature* at the point. One should also notice that $1/r_c(p_0)$ is a signed quantity in general, although we only illustrate the positive situation in Fig. 2(b) for the sake of simplicity. Note our simple approach can be easily extended to derive the small-r behavior in higher dimensions. In any d dimension, we find

$$F_{sv}(r) = s\left(\frac{1}{2} + \frac{r}{2B(\frac{d-1}{2}, \frac{1}{2})}\left\langle \frac{1}{r_c(p_0)} \right\rangle\right), \tag{32}$$

where $B(\frac{d-1}{2}, \frac{1}{2})$ is the beta function. Using this approach, the connection between the small-r behavior of $F_{sv}(r)$ and mean-curvature interfacial growth problems

[35] is intuitively clear.

Remarks: Note that when the "test" sphere of p_0 intersects with the non-differentiable singularities, such as edges or corners, the derivation above breaks down. Thus, formulas Eq. (29) and Eq. (31) do not hold in general for interfaces that have singularities, even though the integrated mean curvature may still be defined and computed in these systems [36]. Using the same approach, we obtain in Appendix A some results for certain systems in which the interfaces have singularities. A discussion of these issues can be found in Ref. [37].

B. Phase-interchange relations for F_{sv}

Here we remark on phase-interchange relations involving the surface-void correlation function F_{sv}. For a two-phase medium, since the sum of indicator functions for phase 1 and phase 2 is unity everywhere, we have

$$\left\langle \mathcal{M}(\mathbf{x})(\mathcal{I}^{(1)}(\mathbf{x}+\mathbf{r}) + \mathcal{I}^{(2)}(\mathbf{x}+\mathbf{r})) \right\rangle = \langle \mathcal{M}(\mathbf{x}) \rangle, \tag{33}$$

implying that the sum of the two surface-void correlation functions for phases 1 and 2 equals the specific surface, i.e.,

$$F_{sv}^{(1)}(\mathbf{r}) + F_{sv}^{(2)}(\mathbf{r}) = s. \tag{34}$$

Furthermore, if the two phases are statistically the same, these surface-void correlation functions are constants [33], namely,

$$F_{sv}^{(1)}(\mathbf{r}) = F_{sv}^{(2)}(\mathbf{r}) = \frac{s}{2}, \tag{35}$$

which is a remarkable relation given that it applies to complex microstructures with such symmetries. This will be verified in Sec. VII.

IV. PRECISE ALGORITHMS TO COMPUTE BOTH F_{ss} AND F_{sv}

Despite the fact that surface correlation functions contain crucial microstructural information, the technical difficulty involved in computing them has been a stumbling block in their widespread use. Methods have been devised to compute the surface correlation functions for dispersions of spheres that rely on dilating the interfaces [31, 32]. A schematic illustration of how the algorithm works for F_{ss} is presented in Fig. 3, where ϵ is the dilation thickness. The algorithm simply measures the two-point probability function $S_2(r; \epsilon)$ of the dilated phase and then one takes the appropriate limit of ϵ. Surface-surface correlation functions have also been used as input information to reconstruct two-phase digitized materials by Jiao, Stillinger and Torquato [13]. Since reconstruction algorithms require numerous evaluations of evolving microstructures, the
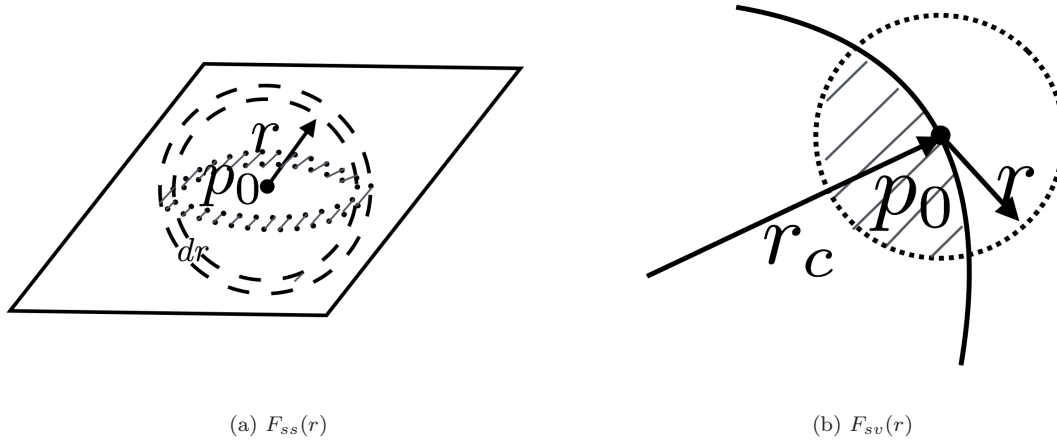
(a) $F_{ss}(r)$

(b) $F_{sv}(r)$

FIG. 2. (Color online.) (a) Schematic that illustrates the small-r asymptotic behavior of the surface-surface correlation function in three dimensions in which the vicinity of the p_0 is approximated by a plane, where the area of interface contained in the shell is shaded. (b) Schematic that illustrates the small-r asymptotic behavior of the surface-void correlation function in two dimensions, where r_c is the local radius of curvature of the interface.

surface correlation functions were approximated to improve computational speed.
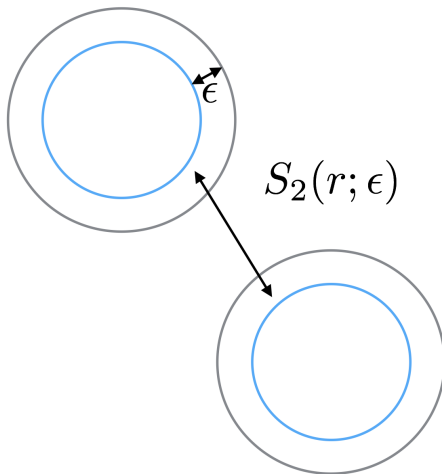
FIG. 3. (Color online.) An illustration of the previous algorithm that computes $F_{ss}(r)$, where ϵ is the dilation thickness. Then $F_{ss}(r)$ is computed by $S_2(r; \epsilon)/\epsilon^2$ as $\epsilon \to 0$.

A. Algorithmic Details

Here we describe efficient general algorithms that enable the precise determination of the surface-surface correlation function $F_{ss}(r)$ and the surface-void correlation function $F_{sv}(r)$ for most situations that one may encounter in simulations and experiments. We consider d-dimensional statistically homogeneous and isotropic two-phase systems within a cubic fundamental simulation cell of side length L under periodic boundary conditions. We also assume that the interfaces are differentiable almost everywhere with exceptions for corners and edges only.

The idea behind the algorithm is that one can reduce the complexity of the problem by extracting information from a cut of the d-dimensional statistically homogeneous and isotropic system with a m-dimensional subspace ($m = 1, 2, \ldots, d - 1$) [1]. For example, the fully three-dimensional two-point correlation function $S_2(r)$ of such a two-phase system is the same as the one-dimensional two-phase system formed from the cut of the original system with an infinitely long line. Similar ideas can be exploited to compute surface correlation functions as well. A straight line intersects with the interface in $\mathbb{R}^d$ and leaves infinitely many intersection points, in principle. We can recover the fully three-dimensional surface correlation functions of the system by analyzing these intersections, but here we need to weight the points in accordance with the fact that the line cuts through the interface at different angles at each intersection point. In particular, because the interface projects to the line differently, each intersection point carries the weight $1/\cos\theta$, where θ is the acute angle between the straight line and the normal vector at the intersection point. From a "dilation" point of view, the straight line will cut through the dilated phase and leave line segments with lengths $\epsilon/\cos\theta$, then $S_2(r; \epsilon)/\epsilon^2$ will reduce to the pair correlation function of intersection points with weights $1/\cos\theta$ in the limit of $\epsilon \to 0$.

Using this simple observation, the calculation of the surface-surface correlation function $F_{ss}(r)$ consists of the following steps:

1. Generate a straight line parallel to one of the edges of the box at a random position.

2. Find all the intersection points $(P_1, P_2,...P_n)$ with interfaces of the system along this straight line. Store their positions $x_1, x_2...x_n$.

3. Find the normal vectors at each intersection point and the angles (the acute one) between the straight line and these norm vectors $\theta_1, \theta_2...\theta_n$. Compute $1/\cos\theta_1$, $1/\cos\theta_2...1/\cos\theta_n$.

4. Bin the distance between every pair of intersection points (suppose the size of each bin is L_{bin}). Add $1/(\cos\theta_i \cos\theta_j)$ to the corresponding bin.

5. Normalize the value in each bin by dividing $2LL_{bin}$.

6. Repeat the process from the beginning.

7. Compute the average of the results.
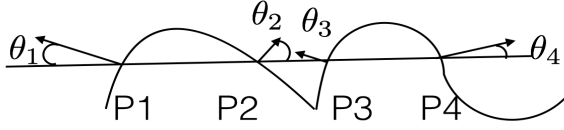
FIG. 4. (Color online.) A schematic plot that elucidates our new algorithm that computes surface correlation functions. Here the sampling straight line intersects with the interface at the points P_1, P_2, P_3 and P_4.

The calculation of the surface-void correlation function $F_{sv}(r)$ consists of the following steps:

1. Generate a straight line parallel to one of the edges of the box at a random position.

2. Find all the intersection points $(P_1, P_2,...P_n)$ with interfaces of the system alone this straight line. Store their positions $x_1, x_2...x_n$.

3. Find the normal vectors at each intersection point and the angles (the acute one) between the straight line and these norm vectors $\theta_1, \theta_2...\theta_n$. Compute $1/\cos\theta_1$, $1/\cos\theta_2...1/\cos\theta_n$.

4. Generate t random points along the straight line. Determine whether each point is in the void phase or not. Suppose $Q_1, Q_2,...Q_m$ are the points in the void phase, store their positions $y_1, y_2...y_m$.

5. Bin the distance between every pair of P_i and Q_j (suppose the size of each bin is L_{bin}). Add $1/\cos\theta_i$ to

the corresponding bin.

6. Normalize the value in each bin by dividing $2tL_{bin}$.

7. Repeat the process from the beginning.

8. Compute the average of the results.

A schematic plot that elucidates our new algorithm is shown in Fig. 4. For systems with hard-wall boundary conditions (the usual case for experimental images), the value in the kth bin should be multiplied by a factor $L/(L - kL_{bin})$ due to the fact that fewer pairs can be formed near both ends of the boundaries. One can also easily generalize the algorithm to anisotropic media and to higher-order correlation functions such as F_{ssv} and F_{svv} [1].

For a d-dimensional system consisting of N particles (or voxels), the complexity for generating a single sampling line is $\mathcal{O}(N^{1/d})$. Computing each pair of intersection points on the line requires $\mathcal{O}(N^{2/d})$. The number of sampling lines is usually a preset number, and thus the overall complexity for computing surface-surface correlation function is $\mathcal{O}(N^{2/d})$. By a similar analysis, we know the complexity for computing surface-void correlation function is $\mathcal{O}(N^{1/d})$. Note that our algorithms are as efficient as the approximation method used in reconstructions [13], but with much better accuracy. Actually, as N increases, fewer lines are needed, since each line contains more intersection points. If the total number of pairs we want to sample is fixed, both algorithms can give constant time complexity.

B. Testing Against the Benchmark of Overlapping Spheres

Three-dimensional overlapping sphere systems provide an excellent benchmark to test our algorithm, since the surface correlation functions are known exactly; see Eq. (21) and Eq. (22). We generate a single but large configuration consisting of 250,000 overlapping spheres with a reduced density $\eta = 1.047$ and particle-phase volume fraction $\phi = 0.649$. We generate one million straight lines at random locations and on each line we generate 1000 random points ($t = 1000$) in the case of computing F_{sv}. As we can see from Fig. 5, the theoretical and simulation results for the surface correlation functions are in excellent agreement with one another, even at the non-differentiable point $r = D$, indicating that the algorithm works remarkably well. As we discussed in Sec. III, the surface-surface correlation function F_{ss} diverges at the origin. We also ran our algorithm at other particle-phase volume fractions and again find excellent agreement with the corresponding theoretical results.
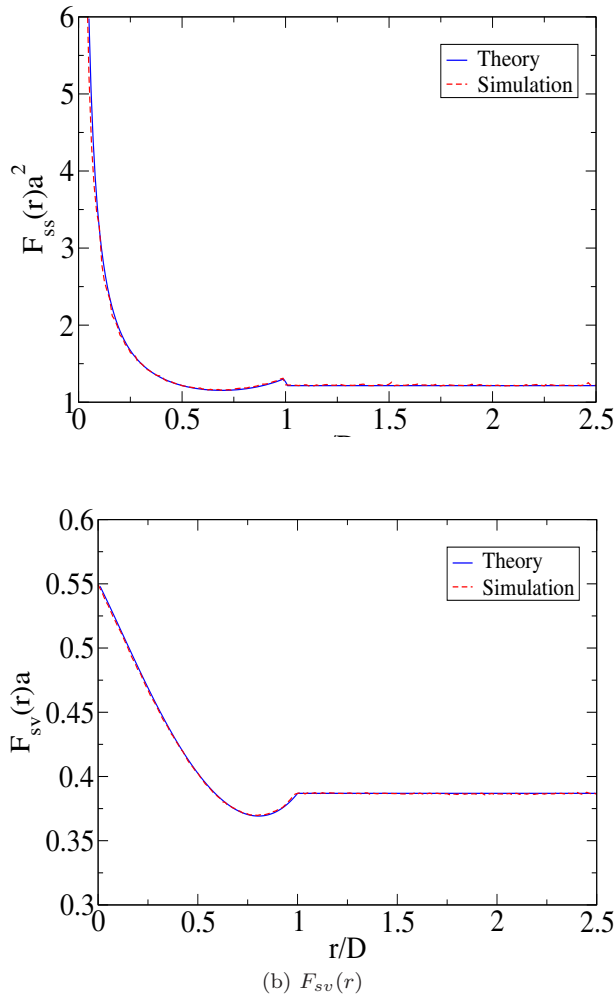
FIG. 5. (Color online.) Comparison of theoretical and simulation results of surface-surface correlation function $F_{ss}(r)$ and surface-void correlation function $F_{sv}(r)$ for overlapping spheres in three dimensions, where D is the diameter of the sphere. The simulations are carried out using 250,000 overlapping spheres in a cubic box under periodic boundary conditions using 1000,000 line samples. The particle-phase volume fraction ϕ is 0.649.

V. COMPUTING SURFACE CORRELATION FUNCTIONS FOR DIGITIZED TWO-PHASE MEDIA

Unlike continuous-space microstructures (e.g., overlapping spheres), where we know surfaces exactly, images of heterogeneous materials are necessarily digitized, which presents algorithmic challenges to identify surfaces and normal vectors. We devote this section to the discussion of how to apply the aforementioned algorithm to this practical setting. Considering that experimental images are generally gray scale, we first discuss the case in which the two-phase medium is obtained from a level cut of a digitized scalar field $F(\mathbf{x})$ in $\mathbb{R}^d$ [1]. This common way to produce a two-phase medium enables us to identify interface normal vectors by the gradient of the scalar field. We then apply this idea to black and white images by first converting the given two-phase medium to a scalar field.

A. Two-phase Media Obtained From Level Cuts of Scalar Fields

Suppose we set a threshold F_0 to convert a scalar field $F(\mathbf{x})$ to a two-phase medium: regions that satisfy $F(\mathbf{x}) > F_0$ constitute phase 1, and regions that satisfy $F(\mathbf{x}) < F_0$ constitute phase 2. The phase indicator function $\mathcal{I}(\mathbf{x})$ for phase 1 is given by

$$\mathcal{I}(\mathbf{x}) = \Theta[F(\mathbf{x}) - F_0], \tag{36}$$

where $\Theta(x)$ is the Heaviside step function. The interface between two phases is simply the contour defined by $F(\mathbf{x}) = F_0$. For any point on the contour, the normal vector is defined by the gradient of the scalar field, i.e. $\nabla F(\mathbf{x})$.

The algorithm can be implemented in essentially the same way as discussed in Sec. IV, but must be specialized to digitized two-phase media. In order to locate points of intersection of the line with interfaces, we need to find where $F(\mathbf{x}) - F_0$ changes sign along a straight line, and then interpolate the position of the point. The gradient at the point can be computed approximately by the finite differences of its neighboring pixels. However, the most significant difference between dealing with continuous models and digitized media is that the number of sampling straight lines one can afford is bounded by the resolution in the later case. Indeed, for an $n \times n$ image, one can only sample at most $\mathcal{O}(n)$ times if the sampling straight lines are lined up with the grid. Thus the resolution of the image is crucial to obtain reliable results.

Here we use Gaussian random field [38] as an example to demonstrate the importance of resolution. The field is generated by a superposition of 10000 plane waves, as we employed elsewhere [19], to give a rather disordered structure. The results for the surface-surface correlation function are summarized in Fig. 6. Here we considered the field within a fixed square region but with different resolutions 1000×1000, 2000×2000, 4000×4000 and 10000×10000. We also include a continuum result which is calculated by directly solving the contour and computing the gradient analytically. It can be seen that as the resolution increases, the numerical results rapidly converge to the continuum result.
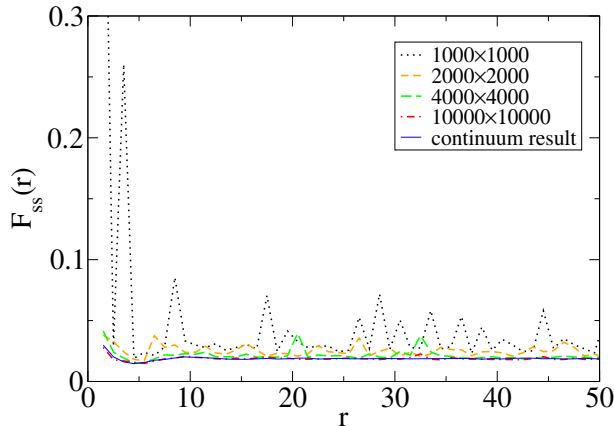
FIG. 6. (Color online.) Simulation results of $F_{ss}(r)$ of a digitized Gaussian random field with level cut $F_0 = 0$ computed under different resolutions. The resolutions shown here are 1000×1000, 2000×2000, 4000×4000 and 10,000×10,000. The continuum result is computed directly using the analytic expression of the scalar field. One can see that as the resolution increases, the numerical result is closer to the continuum result.

B. Significance of the $1/\cos\theta$ Threshold

Figure 6 shows that the computed F_{ss} fluctuates widely when the resolution is low. We discuss the origin of this behavior and how to deal with it in this subsection.

To begin, consider the simple situation illustrated in Fig. 7, which involves a straight line sampling the boundary of a unit circle. The discussion of this case is instructive because the vicinity of the intersection point can be approximated by sphere surfaces in most cases, and when r is large enough, $F_{ss}(r)$ is proportional to $\langle 1/\cos\theta\rangle^2$. So for simplicity, the aim here is to estimate $\langle 1/\cos\theta\rangle$ for the lower left quarter of the circle. Suppose the straight line samples from $x = 0$ to $x = 1$ uniformly along the direction that is perpendicular to itself; then we have

$$\int_0^1 \frac{dx}{\cos\theta} = \int_{\frac{\pi}{2}}^0 \frac{d(1-\sin\theta)}{\cos\theta} = \int_0^{\frac{\pi}{2}} d\theta = \frac{\pi}{2}. \quad (37)$$
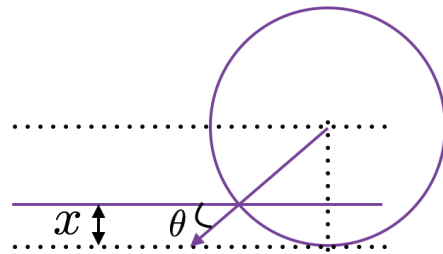

FIG. 7. (Color online.) An illustration of the sampling scenario. The unit circle is the interface and the solid straight line samples along the direction that is perpendicular to itself by varying x.

The integral correctly gives the surface area of the lower left quarter of the circle. Notice that although the integrand $1/\cos\theta$ is divergent at $x = 0$, it is still integrable because the probability of hitting the vicinity of the singularity is proportionally infinitesimally small. However, it is easy to see that the variance of $1/\cos\theta$ is divergent since the integral of $1/\cos^2\theta$ diverges, which implies that large deviations can result when estimating the mean of $1/\cos\theta$. However, the simulation results suggest that increasing the number of sampling lines still reduces the fluctuations from the expected value, and a large sampling number yields good estimates, as one can see in Fig. 5. This suggests that the probability of getting a large deviation diminishes when the sampling number is increased. This is indeed the case, as we show in Appendix B.

However, in the case of digitized media, one cannot increase the sampling number arbitrarily. On the other hand, the probability of hitting the vicinity where $\theta \approx \pi/2$ can be rounded to a relatively large fraction due to the finite resolution. For example, a curved interface can align parallel to the sampling line after the digitization. The consequence is that we are more likely to encounter large deviations, as one can see in Fig. 6, where the abnormal peaks (as well as the universal trend of overestimating $F_{ss}(r)$) are due to certain very large values of $1/\cos\theta$ encountered in the sampling. Although both problems can be alleviated by simply increasing the resolution, it is generally not known a prior that what resolution is required. Furthermore, obtaining high-resolution representations can also be computationally or economically costly, or simply beyond access due to the limitation of experimental techniques or available memory for a simulation. These restrictions force us to come up with a more efficient way to bypass the problems of digitized media. A straightforward way to remove this effect is to simply discard samples when they are

larger than a certain threshold δ, i.e., $1/\cos\theta > \delta$. The bias induced by this method is usually small and insignificant, but with this small compromise, one can significantly reduce fluctuations (see a detailed analysis in Appendix C). To demonstrate the effect of applying thresholds to digitized media, we take the lowest resolution representation (1000×1000) of the Gaussian random field in the last subsection and recompute F_{ss} with a threshold $\delta = 100$. The result is shown in Fig. 8. Note that after applying a threshold, the fluctuations are dramatically suppressed and the result is much closer to the continuum result, even comparable to the ones with much higher resolutions in Fig. 6.
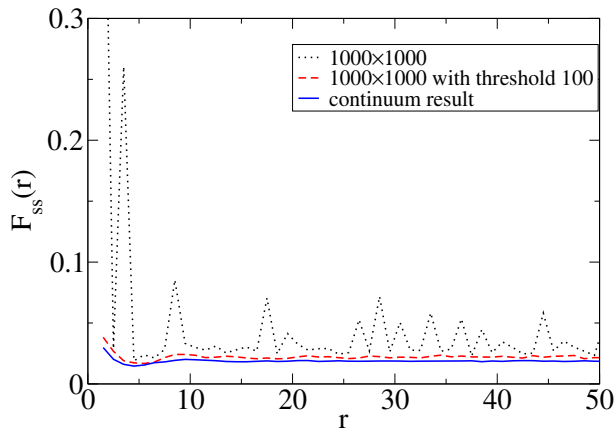
FIG. 8. (Color online.) A comparison of simulation results of $F_{ss}(r)$ of a digitized Gaussian random field with level cut $F_0 = 0$ computed with and without applying a threshold δ. The resolution is 1000×1000, and the threshold δ is 100. The continuum result is computed directly using the analytic expression of the scalar field. By applying a threshold, the fluctuations are largely suppressed and the result is much closer to the continuum result, even comparable to the ones with much higher resolutions.

C. Converting Digitized Two-phase Media into Scalar Fields

We complete our discussion of two-phase media by discussing the case in which all of the information provided about the system is a binary digitized medium. Due to the jagged interface geometry, the transition from one phase to another is sharp and there is no easy way to estimate the direction of the norm vector as we did in the case of Gaussian random fields by computing the gradient of the scalar field.
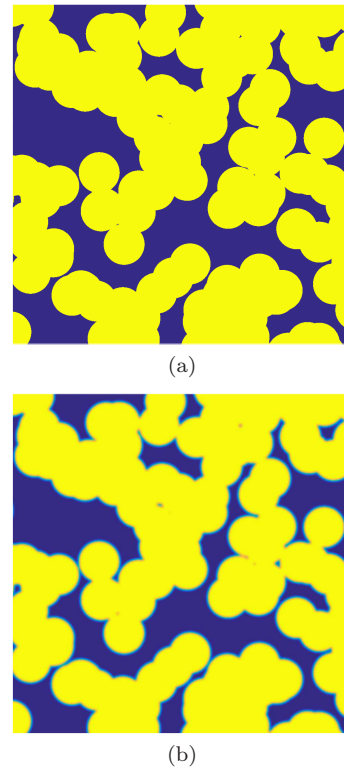

FIG. 9. (Color online.) (a) Digitized random overlapping disks with particle-phase volume fraction 0.677. (b) Corresponding scalar field of (a) by using Gaussian kernels with $b = 0.042D$.
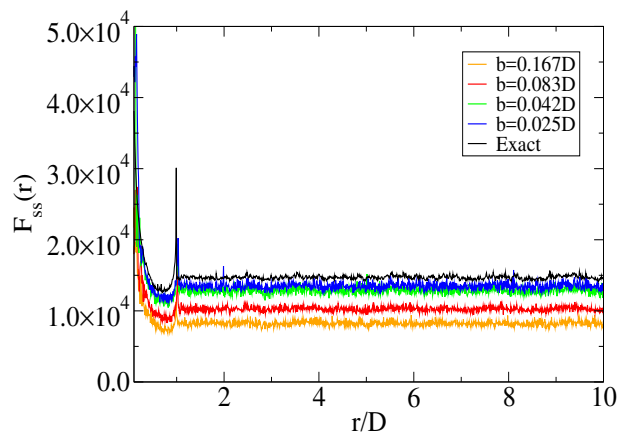

FIG. 10. (Color online.) Simulation results of $F_{ss}(r)$ of a digitized overlapping disk configuration by applying Gaussian kernels for different values of b.

Here we propose a straightforward method to deal with this situation. We first convert the two-phase medium to a coarse-grained scalar field, then convert it back to a two-phase medium by thresholding. In this way we can again use the algorithm introduced in Sec. V. A. We follow the procedure described in Refs. [1] and [39]. By taking pixels in phase i as source points, we can convert the two-phase medium into a scalar field

$F(i, j)$ by convolving the indicator function $\mathcal{I}^{(i)}$ with a kernel or filter $K(\mathbf{x})$. Then, the scalar field is

$$F(i, j, \{C\}) = \sum_k \sum_l \mathcal{I}^{(i)}(i+k, j+l) K(k, l, \{C\}), \quad (38)$$

where $\{C\}$ represents the parameters of the kernel. One of the most common choices of kernels is the Gaussian filter,

$$K(\mathbf{x}; a) = \exp(-\frac{|\mathbf{x}|^2}{b^2}), \quad (39)$$

where a is a length parameter that controls the size of the "influence" region of the filter. By taking a level cut of the scalar field at a threshold F_0, we can then convert the scalar field back into a two-phase medium. The threshold F_0 is chosen to retain the original volume fraction of phase i.

We include an example of overlapping spheres in two dimensions processed by this method. We prepare a digitized realization of 10,000 overlapping disks under periodic boundary conditions at a particle-phase volume fraction $\phi = 0.677$. The resolution is chosen to be that the side length of a pixel is $1/120$ of the diameter D of the disk. We apply the Gaussian filter mentioned above for different values of b and compute the corresponding F_{ss} of the converted scalar fields. A comparison of a portion of the system before and after applying the filter ($b = 0.042D$) is shown in Fig. 9, one can see that the structure of the system is maintained while there is a transition region between two phases. The comparison of F_{ss} computed with different filters is shown in Fig. 10 along with the exact result computed from the continuum model. It is noteworthy that although all the surface-surface correlation functions computed capture the shape of the exact one, they all tend to underestimate the actual function. The possible explanation is that the digitized version loses detailed interfacial information and hence the interface appears to be less curved, which leads to smaller surface areas. However, as b decreases and the filter becomes more localized, the difference between the computed F_{ss} and the exact one monotonically diminishes. The smallest value of b shown in Fig. 10 is three times of the pixel width, one may expect that when the resolution is high enough, the curve of the digitized version will finally converge to the exact one of the underlying pattern. Although images obtained in experiments may not necessarily be of high resolution, generally they are grey-scale images, which means one can simply use the algorithm described for scalar fields directly.

VI. RESULTS FOR OVERLAPPING AND NONOVERLAPPING SPHERE PACKINGS

In this section, we compute the surface correlation functions of several particle systems, including overlapping spheres, hard-spheres in equilibrium and decorated "stealthy" patterns. Given our abilities to compute the surface-surface correlation function, we can calculate local surface-area variances through Eq. (10), and compare them with local volume-fraction variances in these systems.
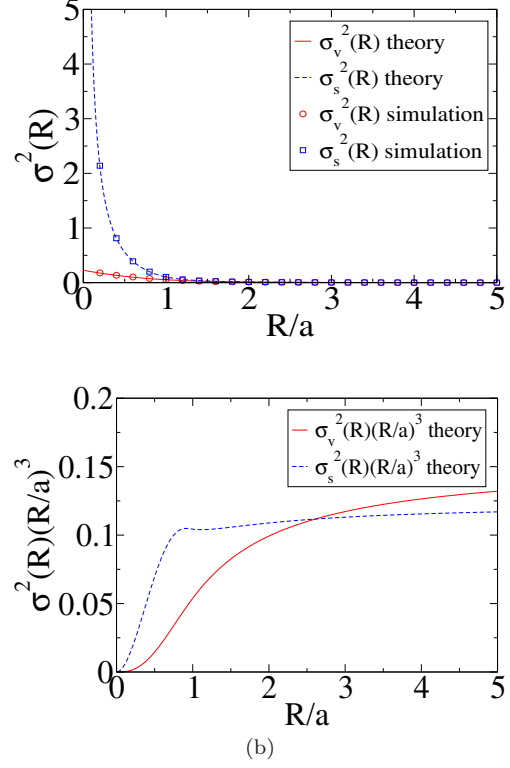

(b)

FIG. 11. (Color online.) (a) A comparison of the local surface-area variances $\sigma_S^2(R)$ with the volume-fraction variances $\sigma_V^2(R)$ for three-dimensional overlapping spheres of radius a as functions of window radius R at particle-phase volume fraction $\phi = 0.649$. Note the surface-area fluctuation is much larger at small R, suggesting that it is a more sensitive microstructure descriptor. (b) A comparison of rescaled local surface-area variances with the volume-fraction variances from (a).

We start by analyzing overlapping spheres in three dimensions. The volume-fraction variance $\sigma_V^2(R)$ and surface-area variance $\sigma_S^2(R)$ for the same system studied in Sec. IV are presented in Fig. 11(a), where R is the radius of the spherical window and a is the radius of particles. These quantities are computed by numerically computing the integral in Eq. (10) as well as through Monte Carlo simulations. In the later method, we generate windows at random positions and calculate the volume-fraction and surface-area variances directly. Since in this model spheres can form very complex clusters, we evaluate the volume fraction and surface area inside each window by generating random points uniformly in the window or on the surface of spheres and counting their fractions inside or on the surface of the clusters correspondingly. The theoretical prediction and

simulation results agree very well, as one can see in Fig. 11(a). It is also noteworthy that in Fig. 11(a) the surface-area variance is much larger compared to the volume-fraction variance. However, one can see that in Fig. 11(b), after multiplied by R^3 (in order to show the large-R behavior of fluctuations), it is clear that there is a crossover of function values around $R = 2.6a$. Further numerical experiments show that the crossover only happens when the particle-phase volume fraction is between 0.57 and 0.7, outside this interval the surface-area variance is always larger than the volume-fraction variance, suggesting that the surface-area variance is a more sensitive descriptor. We further compare the surface-area variance $\sigma_s^2(R)$ with the volume-fraction variance $\sigma_v^2(R)$ of hard spheres in equilibrium in three dimensions at different packing fractions. The variances are again computed using Eq. (10), however the autocovariance functions are not known analytically in this case. As mentioned previously, we use the results included in Ref. [14], which was computed using the Percus-Yevick approximation and the Verlet-Weis correction. We include our results in Fig. 12. Note that the surface-area variance is always larger than the volume-fraction variance across a large span of packing fractions. In Fig. 13, the surface-area variance and the volume-fraction variance are compared respectively at different packing fractions. As the packing fraction increases, the hard-sphere system becomes more short-range ordered [40], thus the variances are expected to drop. The overall trend of $\sigma_s^2(R)$ and $\sigma_v^2(R)$ is consistent with this intuition. However, the volume-fraction variances experience another crossover at small R, while the surface-area variances drop monotonically and larger gaps can be seen between the curves for different packing fractions. These results strongly suggest that the surface-area variance is a more sensitive measure of microstructures of the system compared to the volume-fraction variance.
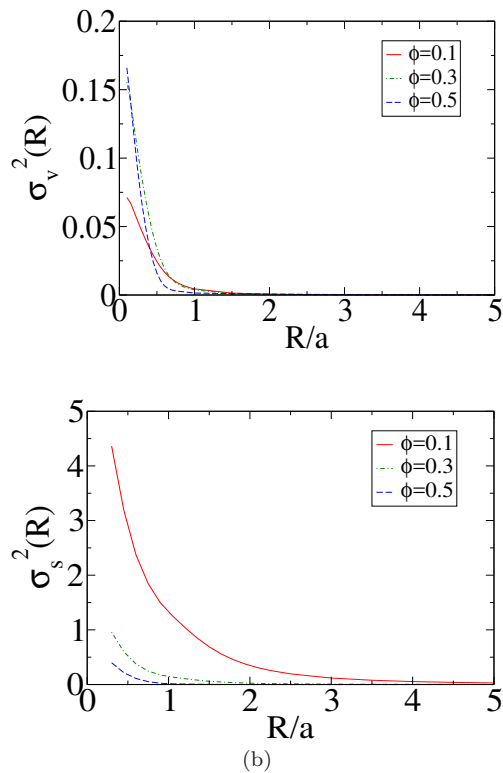

FIG. 13. (Color online.) (a) A comparison of volume-fraction variances of hard spheres at equilibrium at different packing fractions, it is noteworthy that there is a crossover at small R. (b) A comparison of surface-area variances of hard spheres at equilibrium at different packing fractions, which clearly reflect the increase of short-range order as packing fraction increases.

Finally, we compare surface-area variances of hard spheres in equilibrium and overlapping spheres at different volume fractions in Fig. 14. The fact that the hard-sphere systems always suppress surface-area fluctuations variances to a greater degree than those of overlapping spheres, which reflects the stronger pair correlations in the former system.
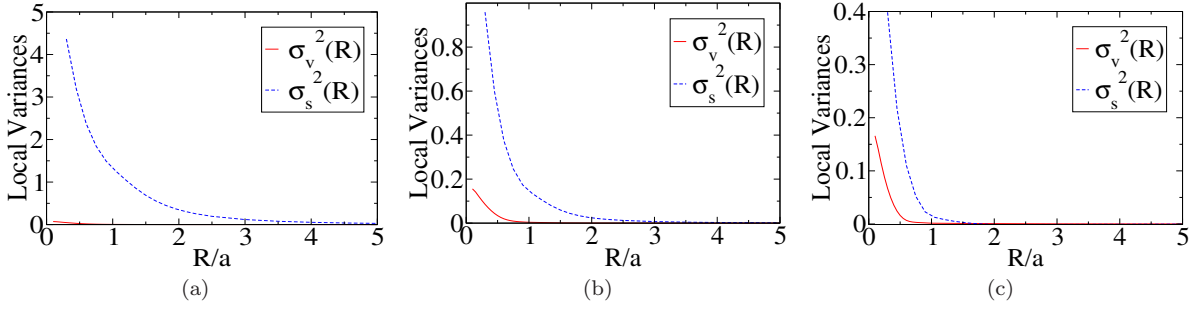
FIG. 12. (Color online.) A comparison of local surface-area variances $\sigma_S^2(R)$ with volume-fraction variances $\sigma_V^2(R)$ for three-dimensional hard spheres of radius a as functions of window radius R in equilibrium at different packing fractions ϕ; (a) $\phi = 0.1$. (b) $\phi = 0.3$. (c) $\phi = 0.5$.
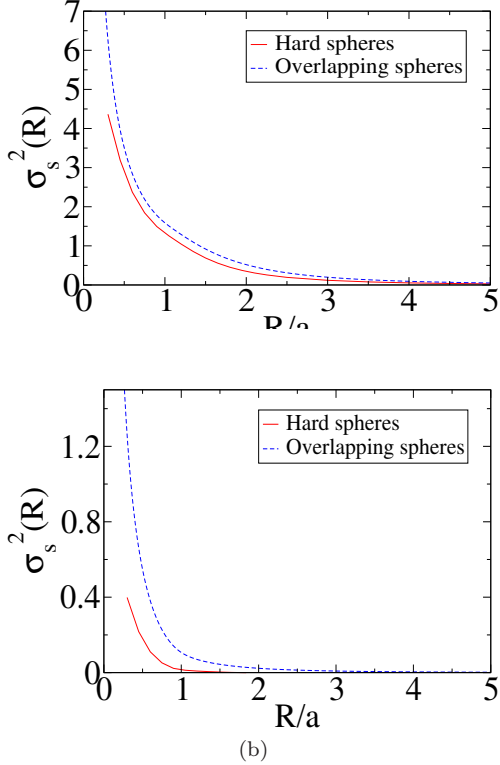

FIG. 14. (Color online.) A comparison of surface-area variances $\sigma_S^2(R)$ of three-dimensional hard spheres in equilibrium and overlapping spheres of radius a as functions of window radius R at different volume fractions. (a) $\phi = 0.1$. (b) $\phi = 0.5$.

Besides using F_{ss} to compute surface-area variances, it can itself be used as a "fingerprint" to detect important structural information, such as short-range order or the hyperuniformity of the system. Here we consider a special hyperuniform point patterns called "stealthy" point patterns that were studied in a recent paper [41], and follow the procedure of circumscribing each point with a sphere to make the system a two-phase medium. The "stealthy" patterns are generated in a simulation box with basis vectors $(c, 0, 0)$, $(0, c, 0)$ and $(c/2, c/2, c/2)$. We compare two sets of stealthy point

patterns, with the parameter $\chi = 0.08$ and 0.46 (In general, the system with larger χ will have more short-range order). Each point is decorated with a variable-sized sphere. We include our results for F_{ss} in Fig. 15. When the decorated spheres are very small, such as the case in Fig. 15(a), they do not overlap with one another, and thus F_{ss} should reveal structural features of the underlying point pattern. Clearly, the curve corresponding to $\chi = 0.46$ in Fig. 15(a) exhibits stronger features, which is consistent with the fact that the pattern is more short-range ordered than that for $\chi = 0.08$. As stated in Ref. [41], the system loses its hyperuniformity when spheres begin to touch each other. From Fig. 15, it is seen that as the diameter D increases, the correlation function begins to lose its features, and ultimately these two "stealthy" cases becomes indistinguishable from each other as well as the corresponding correlation function for overlapping spheres, shown in Fig. 5(a). The dramatic decrease around $r = D$ corresponds to the fact that the correlation between any two points on the same sphere cannot contribute to the function value beyond $r = D$, and thus reveals a characteristic length scale of the system. Moreover, although the two systems start almost at the same specific surface, the gap between two curves continues to increase, and in the end the system with $\chi = 0.46$ has a much larger specific surface. This suggests that the system with $\chi = 0.46$ has greater short-range order that keeps the decorated spheres from overlapping with one another, and thus leads to a larger specific surface.

VII. RESULTS FOR SNAPSHOTS OF EVOLVING SPATIAL PATTERNS

In this section, we go beyond the analysis of well-known sphere models and extend the application of our algorithm to other important disordered patterns encountered in the physical and biological sciences. Specifically, we focus on time-dependent pattern formation processes that are governed by the Cahn-Hilliard equation and the Swift-Hohenberg equation. These patterns have recently been shown to be hyperuniform and could
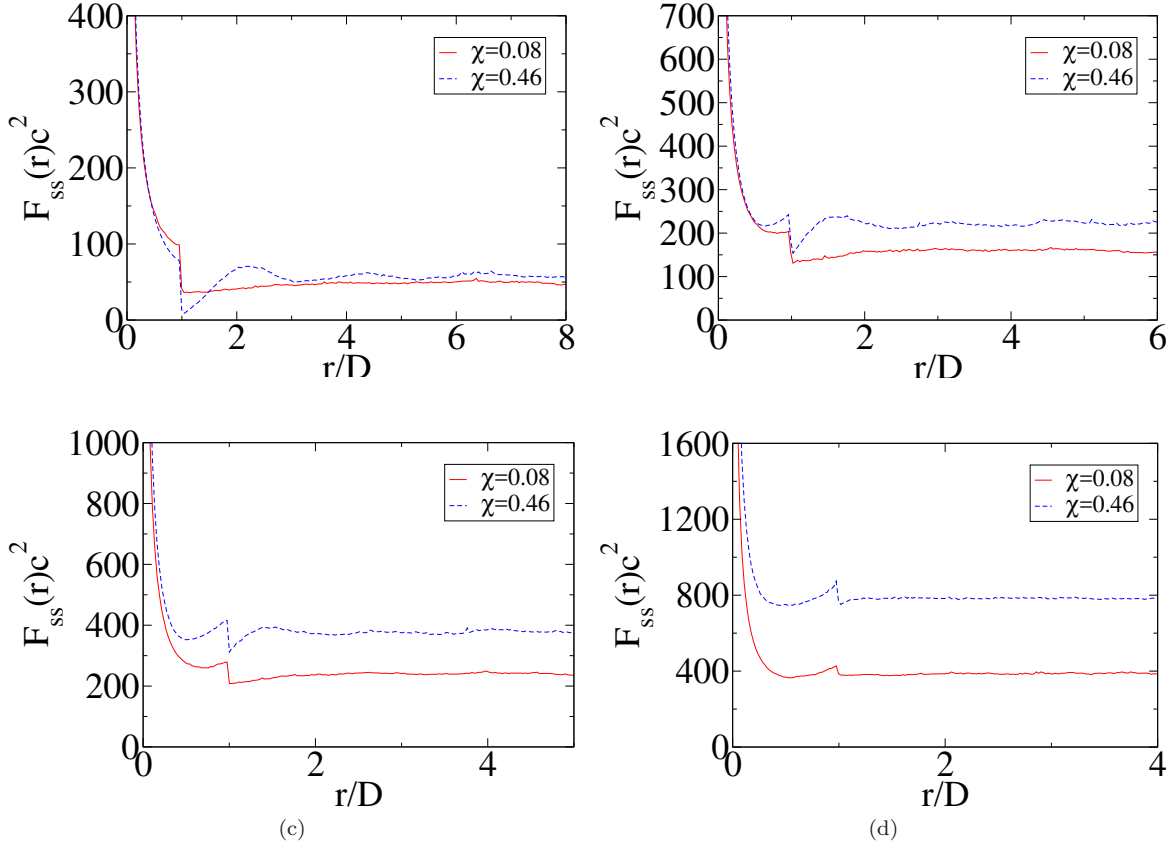
FIG. 15. (Color online.) Surface-surface correlation functions F_{ss} of stealthy patterns decorated with spheres with different diameters: (a) $D = 0.05c$. (b) $D = 0.07c$. (c) $D = 0.08c$. (d) $D = 0.1c$.

have important applications in material science [19]. We determine correlation functions of snapshots of these patterns here.

A. Spinodal decomposition patterns from the Cahn-Hilliard equation

The Cahn-Hilliard equation was introduced to describe phase separation by spinodal decomposition [42] and has been applied to model alloys [43], polymer blends [44] and even pattern formations in ecology [45].

In Fig. 16, we show two typical patterns generated by the equation. The left one is at critical quench, in which case the volume fraction ratio for two phases is 1:1, while the right one is off critical quench and has volume fraction ratio 2:8. The interface between the two phases is highlighted.

One important feature of the Cahn-Hilliard equation is that the system will enter a "scaling regime" after some time, and the system will remain statistically the same after scaled by a growing characteristic length. This provides an indirect way to check our algorithm on digitized media. We can compute the surface correlation functions at different times and then an appropriate scaling enables them to collapse onto a single curve.

The rescaled surface-surface and surface-void correlation functions at different times are shown in Fig. 18 and Fig. 19 for two systems. The curves do collapse onto each other as we expect, further justify the applicability of our algorithm. In Fig. 19(a), one can see an interesting flat correlation function, which is an example of our argument in Sec. III. B.

Note that although these two systems appear to be structurally different, the standard autocovariance functions $\chi_V(r)$ in Fig. 17 almost look the same. The inability to distinguish the structures of these two systems is easily overcome by complementing $\chi_V(r)$ with the information of $F_{ss}(r)$ and $F_{sv}(r)$, as they differ greatly for these two systems. Specifically, one can see that in Fig. 19, the surface-void correlation function for the critical quench has a flat slope at the origin, while the one for the off critical quench has a downward slope. This can be well explained by the small-r behavior of F_{sv} we derived in Sec. III. A.. From Eq. (29) we know that the slope of F_{sv} at origin is proportional to the *mean curvature* of the system. In the case of critical quench, the surface consists of both concave and convex parts, whose contributions cancel each other out, and thus the *mean curvature* is zero. In the case of off critical quench, where the matrix is the solid phase and the

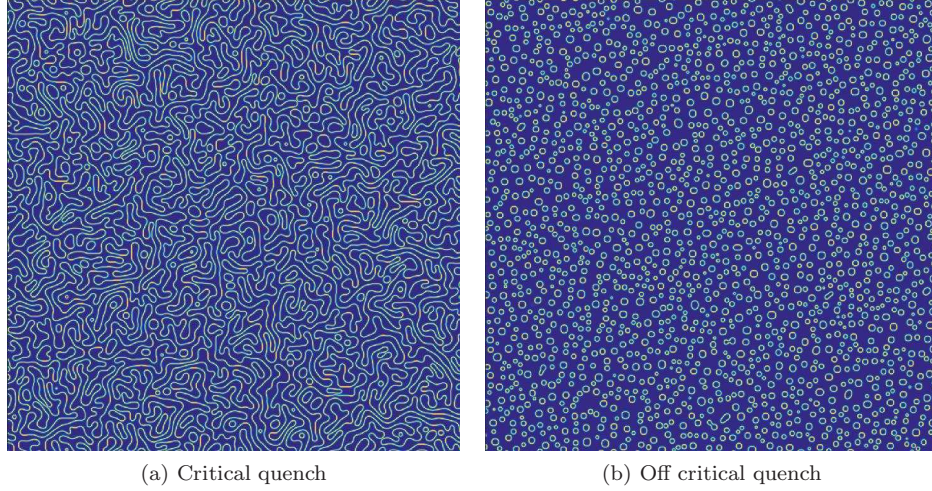(a) Critical quench (b) Off critical quench

FIG. 16. (Color online.) The interfaces of two binary mixtures undergoing a phase separation at critical quench (volume fraction ratio of two phases is 1:1) and off critical quench (volume fraction ratio of two phases is 2:8) respectively. The system size i
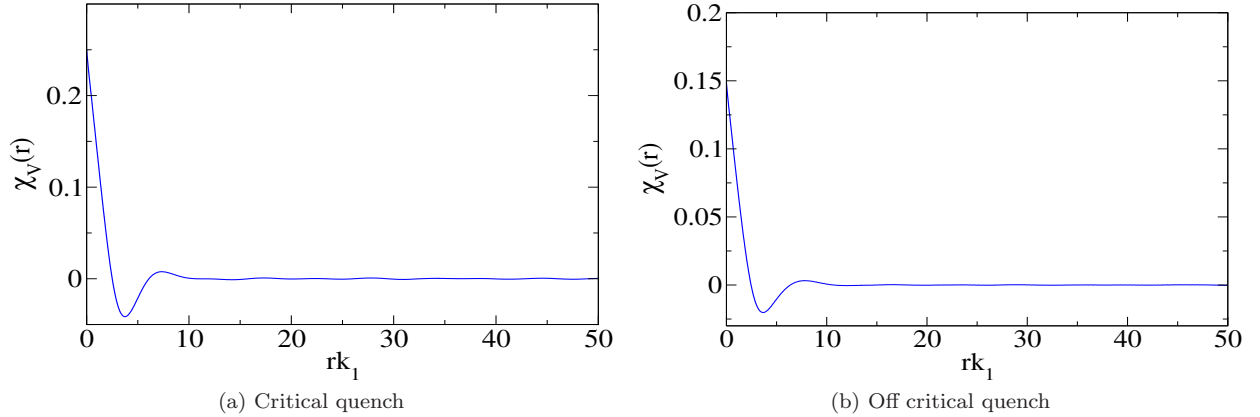

(a) Critical quench (b) Off critical quench

FIG. 17. (Color online.) (a) The autocovariance function $\chi_V(r)$ versus rk_1 at $t = 100,000$ associated with the spinodal decomposition pattern shown in the left panel of Fig. 16, where r is scaled by the characteristic wavenumber k_1. (b) The autocovariance function $\chi_V(r)$ versus rk_1 at $t = 100,000$ associated with the spinodal decomposition pattern shown in the right panel of Fig. 16. There is no significant difference between these two curves in (a) and (b).

droplets are taken to be the void phase, the *mean curvature* is apparently negative, and thus in Fig. 19(b) we see the curve slopes down initially. This example again demonstrates the value of surface correlation functions in characterizing complex patterns.

B. Patterns from the Swift-Hohenberg equation

The Swift-Hohenberg equation was developed to study Rayleigh-Bénard (RB) convection in hydrodynamics and later it became a subject of interest on its own in pattern formations [46]. The pattern produced by this equation is usually labyrinth-like, and the width of the "channel" is determined by a pre-selected wavenumber k_0. It has been shown that the patterns can have different degree of hyperuniformity [19] when some

tuning parameters are changed, although they may appear to be alike.

Here we compute and compare two surface-surface correlation functions for two patterns generated under different k_0, namely $k_0 = 0.7$ and $k_0 = 0.32\pi$ in the same way in the authors' previous paper [19]. It has been shown that the later one is more long-range ordered, which is also justified in our plot of F_{ss} in Fig. 20. It is evident that the F_{ss} for $k_0 = 0.32\pi$ is much more long-ranged than the one for $k_0 = 0.7$. Both curves have sharp spikes when rk_0 is integer times of π, which corresponds to the fact that the underlying patterns consist of stripes with width of π/k_0, leaving roughly parallel interfaces with the same spacing at short scales. Note that spikes in $F_{ss}(r)$ also occur in sphere systems but only at the single location $r = D$ (see Fig. 5 and Fig. 15 for examples), while the corresponding S_2 for these systems are smooth functions without sharp transitions

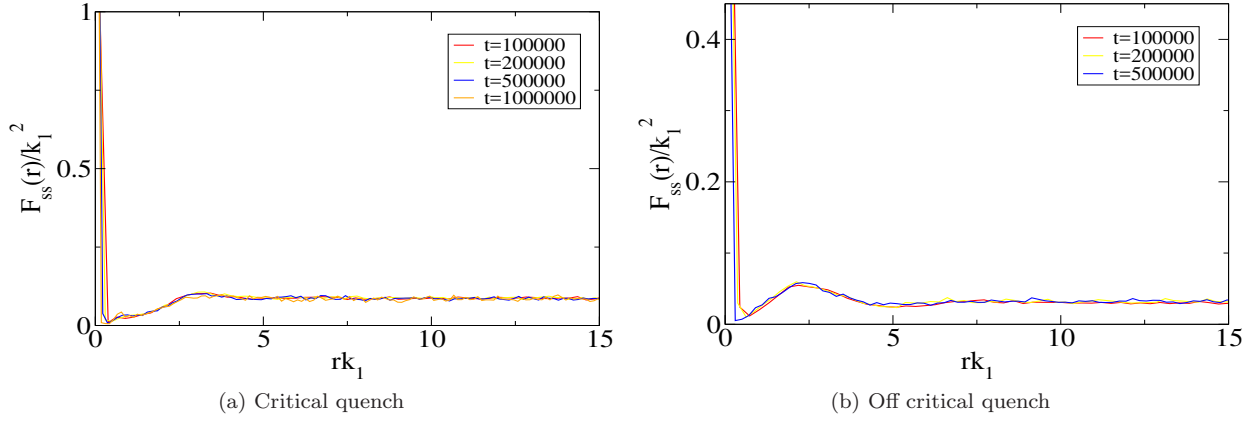(a) Critical quench   (b) Off critical quench

FIG. 18. (Color online.) (a) The scaled surface-surface correlation function $F_{ss}(r)/k_1(t)^2$ versus $rk_1(t)$ at different time stages associated with the spinodal decomposition pattern shown in the left panel of Fig. 16. (b) The scaled surface-surface correlation function $F_{ss}(r)/k_1(t)^2$ versus $rk_1(t)$ at different time stages associated with the spinodal decomposition pattern shown in the right panel of Fig. 16. One can see that they both collapse onto a single curve respectively after the rescaling. The s

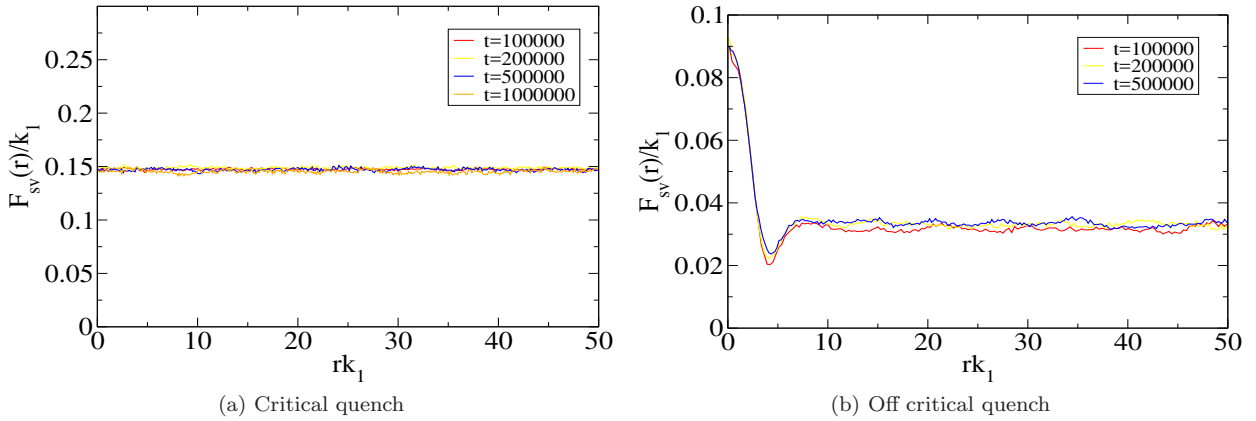(a) Critical quench   (b) Off critical quench

FIG. 19. (Color online.) (a) The scaled surface-void correlation function $F_{sv}(r)/k_1(t)$ versus $rk_1(t)$ at different time stages associated with the spinodal decomposition pattern shown in the left panel of Fig. 16. (b) The scaled surface-void correlation function $F_{sv}(r)/k_1(t)$ versus $rk_1(t)$ at different time stages associated with the spinodal decomposition pattern shown in the right panel of Fig. 16. One can see that they both collapse onto a single curve respectively after the rescaling. The shapes of two curves are significantly different.

(see Refs. [1] and [20] for plots). This again shows that surface-surface correlations can be superior in detecting short-scale microstructural features compared to that of the standard two-point correlation function $S_2(r)$.

We also evaluate the local surface-area variances in these systems using F_{ss} and Eq. (10). The results are shown in Fig. 20. Note that the surface-area variance for $k_0 = 0.7$ scales like R^{-3}, implying hyperuniformity [17]. However, the variance for $k_0 = 0.32\pi$ scales even slower than R^{-2}. The explanation is that in the case of $k_0 = 0.32\pi$, the corresponding wavelength is too small compared to the pixel size, which makes the numerical integration in Eq. (10) unreliable.
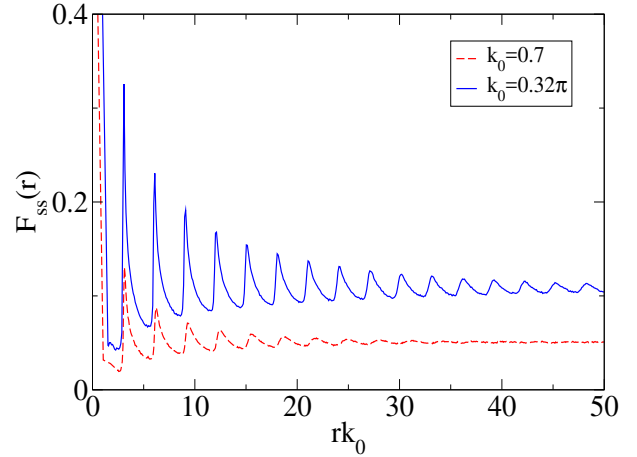

FIG. 20. (Color online.) Surface-surface correlation function $F_{ss}(r)$ of patterns generated from the Swift-Hohenberg equation with two different values of the parameter k_0.
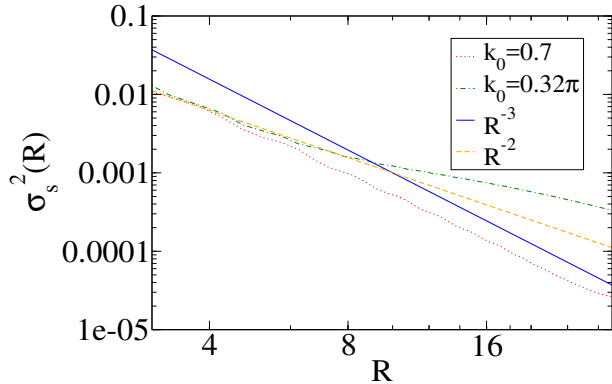
FIG. 21. (Color online.) Local surface-area variances $\sigma_s^2(R)$ as functions of window radius R computed from Eq. (10) for patterns generated from the Swift-Hohenberg equation with two different values of the parameter k_0, with comparison of different scalings.

VIII. CONCLUSIONS AND DISCUSSION

In this paper, we developed efficient general algorithms that enable the sampling of the surface-surface correlation function $F_{ss}(r)$ and the surface-void correlation function $F_{sv}(r)$ with heretofore unattained precision. Our algorithms have advantages over the traditional "dilation" method [31, 32]. First, the dilation method can only be easily implemented when the interfaces are relatively smooth and easy to be parameterized (e.g., packings of spheres and ellipsoids). However, our algorithms can be easily adapted to general complex interfaces. Second, the dilation method is difficult to implement for digitized media, which greatly limits its application to experimental data. By contrast, we have shown that our algorithms can be straightforwardly applied to digitized media. Third, as the dilation thickness ϵ approaches to zero, the probability of hitting the dilated phase will proportionally decrease, which requires a large number of samplings to ensure the accuracy, and hence greater computational time. However, in the extreme situation that the information of a large but single system is available, our algorithms can yield accurate results from a single sample, since it is possible for the straight line to penetrate the interface a sufficiently large number of times. Moreover, our algorithms can be generalized to compute three-point surface correlation functions straightforwardly. Application of our algorithms to a variety of model disordered microstructures reveals that surface-surface correlation function $F_{ss}(r)$ is a sensitive descriptor of small-scale structural features, especially compared to the information content of the standard two-point correlation function $S_2(r)$.

We also showed that surface correlation functions can be used to compute the surface-area variance, a quantity that can be a more sensitive measure of the microstructure compared to the volume-fraction variance. Through examples of spinodal decomposition patterns, we showed that surface correlation functions contain in-

formation that supplements that of S_2, and the small-r behavior of $F_{sv}(r)$, which is determined by the *mean curvature* of the system. In two dimensions, the *total curvature* of a closed simple curve is a constant 2π, implying that when the system approaches a percolation threshold, the absolute value of the *mean curvature* will drop dramatically due the formation of large clusters. This observation suggests that the surface-void correlation function $F_{sv}(r)$ may aid in detecting the onset of continuum percolation, which is an interesting topic for future exploration. We also showed how surface-surface correlation functions can be used to determine the hyperuniformity of two-phase media using patterns generated by the Swift-Hohenberg equation as examples.

Lower-order correlation functions have been successfully used to infer the physical properties of random media as well as to reconstruct them. This bodes well for their use in machine learning in the area of material optimization [47, 48]. We expect that the algorithms to compute precisely the surface-surface correlation function $F_{ss}(r)$ and the surface-void correlation function $F_{sv}(r)$ presented in this paper will equip the community with new powerful computational tools to characterize the structure and physical properties of multiphase media, especially with respect to those physical processes that are intimately linked to the interfaces. In particular, our algorithms can be adapted in reconstruction algorithms [13, 20] with heretofore unattained accuracy without sacrificing computational speed.

Appendix A: The small-r behavior of surface correlation functions of systems with singularities

Since the surface-surface correlation function discussed in Sec. III is only approximated to the zero order, we focus our attention here to the surface-void correlation function of systems with singularities. We include results for certain specific cases, namely, two-dimensional systems that all singularities are corners and three-dimensional overlapping spheres.

For two-dimensional systems that all singularities are corners, suppose the angle formed by a corner to the solid phase is θ, one can show that the $F_{sv}(r)$ for small r can be written as

$$F_{sv}(r) = s\left(\frac{1}{2} + \frac{r}{2\pi}\left\langle\frac{1}{r_c}\right\rangle_\Omega\right) + 2\rho_c\left\langle\cot\frac{\theta}{2}\right\rangle r, \quad \text{(A1)}$$

where Ω is the set of all the points on the interfaces that are differentiable, and ρ_c is the density of corners in the system. One interesting implication of this formula is that the expression of the surface-void correlation function for packings of equilateral polygons is the same as the expression for packings of their inscribed circles.

Singularities in three dimensions are much more complex to analyze. Here we only take overlapping spheres as an example. Surprisingly, this is a nontrivial model of a heterogeneous material, since the lack of spatial correlation implies that the particles may overlap to form complex clusters, and leave many non-differentiable edges in the system. Using the same geometric approach in Sec. III, by naively plugging in a as r_c in the formula (31), we should have

$$F_{sv}(r) = s(\frac{1}{2} + \frac{r}{4a}), \qquad (A2)$$

where a is the radius of spheres and $s = 3\eta e^{-\eta}/a$ is the specific surface of overlapping spheres [1]. However, expanding Eq. (22) to the first order directly gives us

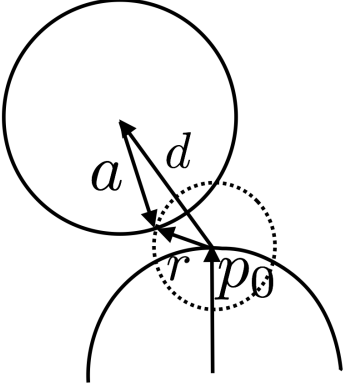$$F_{sv}(r) = s(\frac{1}{2} + \frac{r}{4a} - \frac{3\eta r}{8a}). \qquad (A3)$$

FIG. 22. An illustration of evaluating the two-body correction to the small r behavior of $F_{sv}(r)$, where the upper left sphere is the "invading" sphere and the sphere in dotted line is the "test" sphere.

The extra negative term $-3s\eta r/8a$ implies that we have overestimated the probability of falling into the void phase P_{sv} by neglecting the fact that another sphere (see the upper left sphere in Fig. 22, which we henceforth call the "invading" sphere) can approach to and intersect with our "test" sphere (see the "dotted" sphere in Fig. 22) and reduce its fraction of surface area covered in the void phase. Indeed, any sphere whose centroid lies in the concentric shell with radius a to $a + r$ around the reference point p_0 will intersect with the "test" sphere (we do not consider spheres that are closer than a since then the reference point would

no longer be on the interface). Here we evaluate the reduced fraction of surface area in the void phase of the "test" sphere.

When the "invading" sphere overlaps with both the "test" sphere and the interface it can be difficult to evaluate the extra surface area of the "test" sphere covered by the "invading" sphere. Luckily, since the volume of the shell is $4\pi a^2 r$, which is already first order, we can approximate the interface around the reference point as a flat plane that divides the "test" sphere into two hemispheres. By symmetry we know the extra surface area covered on average is just $1/2$ of the total surface area covered by the "invading" sphere on average. Suppose the distance between the center of the "invading" sphere and the reference point is d, then the surface area of the spherical crown that is covered is

$$S = 2\pi r^2 (1 - \frac{d^2 + r^2 - a^2}{2dr}). \qquad (A4)$$

Letting $d = a + x$, the fraction of surface area of the "test" sphere that is covered by the "invading" sphere is

$$\frac{S}{4\pi r^2} = \frac{1}{2}(1 - \frac{(a+x)^2 + r^2 - a^2}{2(a+x)r}) \approx \frac{1}{2}(1 - \frac{x}{r}). \quad (A5)$$

Then on average the total fraction that is covered by the "invading" sphere is

$$\int_0^r \frac{1}{2}(1 - \frac{x}{r}) \times 4\pi(a+x)^2 \rho dx = \pi a^2 r \rho + \mathcal{O}(r^2). \quad (A6)$$

Finally, by symmetry, we know the correction term to $F_{sv}(r)$ is $-s\pi a^2 r\rho/2$ or $-3s\eta r/8a$, as in Eq. (A3). This correction term will disappear in the dilute limit, since spheres will not overlap with one another.

One can carry out the same analysis for impenetrable spheres, but the calculation will be much more involved. In this case, the other sphere can only approach from the void phase, and the non-overlapping condition will restrict its direction to a small range, which will in the end make the correction term of the order $\mathcal{O}(r^2)$ as long as $g_2(D^+)$ is not a delta function, where D is the diameter of a sphere. Thus our general formula will apply to hard spheres in equilibrium or random sequential addition (RSA) packings [49].

Appendix B: The probability of getting an abnormal peak

It is instructive to estimate the probability $P\{\frac{1}{N}\sum_{i=1}^{N} \frac{1}{\cos\theta_i} > \frac{\pi}{2} + e\}$, where N is the number of sampling and e stands for a given error. To do so, first we can sort $1/\cos\theta$ such that $1/\cos\theta_i \geq 1/\cos\theta_{i+1}$ for $i = 1, 2, ..., N - 1$. Define

$$T = \min\{j| \sum_{i=1}^{j} \frac{1}{\cos\theta_i} > N(\frac{\pi}{2} + e)\}, \qquad (B1)$$
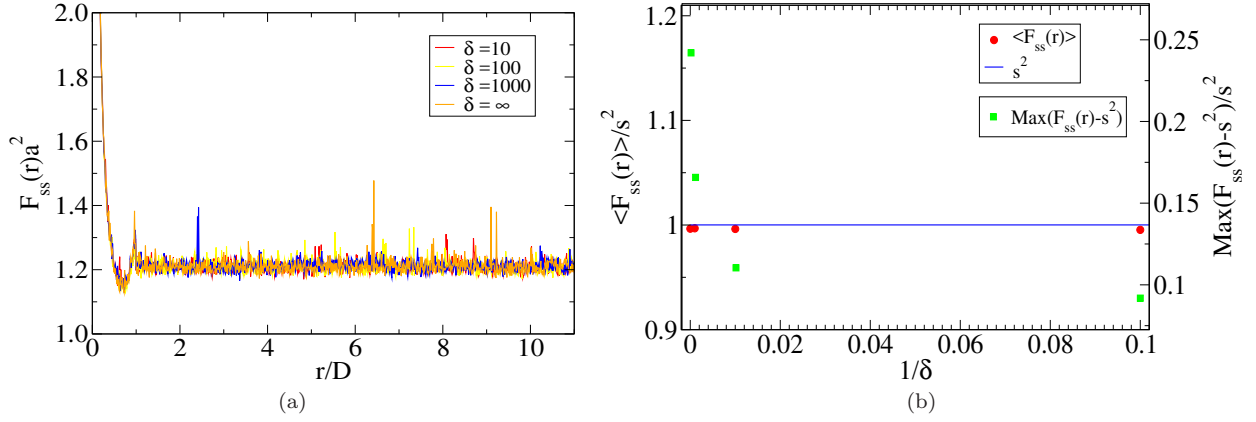
FIG. 23. (Color online.) (a) Simulation results of $F_{ss}(r)$ for systems of three-dimensional overlapping spheres computed with different thresholds 10, 100, 1000 and without threshold. We use the same system in Sec. IV, except that only 10,000 sampling lines are used. Clearly, applying thresholds does not change the overall shape of the curve, but significantly reduce the fluctuations. (b) Red circles: Average of $F_{ss}(r)$ in the interval $[D, 11D]$ for different thresholds. The blue line is s^2, which is the theoretical value of $F_{ss}(r)$ when $r > D$. Green squares: the largest deviation of $F_{ss}(r)$ from s^2 in the interval $[D, 11D]$ for different thresholds. The average of $F_{ss}(r)$ is always very close to the expected value (within 1% for all cases). However, by applying more restricted threshold, there is a trend to reduce the abnormal peaks significantly.

which is the smallest number of elements needed to make the inequality hold. Then we can decompose the probability by conditioning on T, i.e.,

$$P\{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\cos\theta_i} > \frac{\pi}{2} + e\} = \sum_{i=1}^{N}P\{T=i\}. \quad \text{(B2)}$$

We can write down each term explicitly. When N is large enough, we have

$$P\{T=1\} = \int_0^{\arcsin\frac{1}{N(\frac{\pi}{2}+e)}} \sin\theta\cos^{N-1}\theta d\theta \propto \frac{1}{N^2}, \quad \text{(B3)}$$

$$P\{T=2\} = \int_{\arcsin\frac{1}{N(\frac{\pi}{2}+e)}}^{\arcsin\frac{2}{N(\frac{\pi}{2}+e)}} \sin\theta d\theta$$

$$\int_\theta^{\arcsin\min\{1,\frac{1}{N(\frac{\pi}{2}+e)-\frac{1}{\sin\theta}}\}} \sin\phi\cos^{N-2}\phi d\phi \propto \frac{1}{N^3} \quad \text{(B4)}$$

$$\vdots$$

The reason why Eq. (B4) has this scaling behavior is because $\cos^{N-2}\phi$ is effectively zero when ϕ is larger than $1/\sqrt{N}$. One can continue this process and it is easy to see that the major contribution to the summation in Eq. (B2) comes from the first a few terms. Thus one can see that by increasing N one can significantly reduce the chance of getting an abnormal peak in the simulation, which is why the results in Fig. 5 are very smooth, especially compared to the "$\delta = \infty$" one in Fig. 23(a), which has a much smaller sampling number.

Appendix C: The effect of setting a threshold

We still take Fig. 7 as an example. By setting a threshold δ, the new expected value of $1/\cos\theta$ is

$$\langle\frac{1}{\cos\theta}\rangle = \int_{\arccos\frac{1}{\delta}}^0 \frac{d(1-\sin\theta)}{\cos\theta} = \frac{\pi}{2} - \arcsin\frac{1}{\delta}. \quad \text{(C1)}$$

The relative error is then $2/\pi\arcsin 1/\delta$. When $\delta \gg 1$, it is approximately $2(\pi\delta)^{-1}$. We can easily control this error by using a moderate threshold. For example, by setting $\delta = 100$, the error is already below 0.7%. However, with this little compromise we can have a finite second moment, i.e.,

$$\langle\frac{1}{\cos^2\theta}\rangle = \int_{\arccos\frac{1}{\delta}}^0 \frac{d(1-\sin\theta)}{\cos^2\theta} = \ln\delta(1 + \sqrt{1-\frac{1}{\delta^2}}). \quad \text{(C2)}$$

Thus, one can reduce the fluctuation to any level simply by adding more samples. When $\delta \gg 1$, the variance is approximated by $\ln\delta$. In the case of $\delta = 100$, one can reduce the relative standard deviation to 1% by using around 12000 samples. It is noteworthy that the mean and variance have different asymptotic behaviors, the error diminishes as $1/\delta$, while the variance only grows as $\ln\delta$, which gives us a great flexibility to choose δ.

The suppression of abnormal peaks by using a threshold δ can also be deduced from Eq. (B2). Since now $1/\cos\theta$ is bounded, it requires at least $N(\pi/2+e)/\delta$ terms to have the inequality in Eq. (B1), which means the leading $N(\pi/2+e)/\delta$ terms in Eq. (B2) vanish, leaving the probability significantly smaller than the case without a threshold (actually by Hoeffding's inequality the probability will decrease exponentially).

We test this idea on two models: overlapping spheres and Gaussian random fields. In the former case, we

consider the same system in Sec. IV, while restricting ourselves to a relatively small amount of sampling lines (10,000) and compare the computed F_{ss} with different thresholds $\delta = 10, 100, 1000$ and ∞ (no threshold). The result is shown in Fig. 23(a). Clearly, in all these cases, F_{ss} fluctuates around a common curve, while as the threshold is tightened, fluctuations are suppressed. To show this point quantitatively, we focus on the behavior of F_{ss} in the interval $[D, 11D]$. As noted in Sec. IV, when $r > D$ the surface-surface correlation function F_{ss} is a constant s^2 in theory and should be a flat line in the plot; while in the simulations F_{ss}, fluctuates around a baseline. To compare the simulation and theoretical results, we compute the average of surface-surface correlation function $\langle F_{ss}(r) \rangle$ and the largest deviation from s^2 in this interval. As shown in Fig. 23(b), the average of $F_{ss}(r)$ is always very close to the expected value (within 1% error for all cases). However, by applying a more restrictive threshold, there is a trend to reduce the abnormal peaks significantly. We indeed get much smoother curves by making a very minor sacrifice of accuracy.

[1] S. Torquato, *Random heterogeneous materials: microstructure and macroscopic properties*, (Springer, New York , 2002).

[2] T. I. Zohdi and P. Wriggers, *An introduction to computational micromechanics* (Springer, New York, 2008).

[3] M. Sahimi, *Heterogeneous Materials I: Linear transport and optical properties*, (Springer, New York, 2003).

[4] J. Vasseur and F. B. Wadsworth, Bull. Volcanol **79**, 77 (2017).

[5] L. J. Gibson and M. F. Ashby, *Cellular solids: structure and properties* (Cambridge university press, 1999).

[6] J. L. Gevertz and S. Torquato, PLOS Comput. Biol **4**, e1000152 (2008).

[7] P. J. E. Peebles, *Principles of physical cosmology* (Princeton University Press, 1993).

[8] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications* (John Wiley & Sons, 2013).

[9] L. Mosser, O. Dubrule, and M. J. Blunt, Phys. Rev. E **96**, 043309 (2017).

[10] N. Lubbers, T. Lookman, and K. Barros, Phys. Rev. E **96**, 052111 (2017).

[11] Y. Jiao, F. Stillinger, and S. Torquato, Phys. Rev. E **76**, 031110 (2007).

[12] Y. Jiao, F. Stillinger, and S. Torquato, Phys. Rev. E **77**, 031135 (2008).

[13] Y. Jiao, F. Stillinger, and S. Torquato, Proc. Natl. Acad. Sci. USA **106**, 17634 (2009).

[14] S. Torquato, J. Chem. Phys. **85**, 4622 (1986).

[15] P. Debye and A. Bueche, J. Appl. Phys. **20**, 518 (1949).

[16] P. Debye, H. Anderson Jr, and H. Brumberger, J. Appl. Phys. **28**, 679 (1957).

[17] S. Torquato, Phys. Rev. E **94**, 022122 (2016).

[18] S. Torquato, J. Phys.: Condens. Matter **28**, 414012 (2016).

[19] Z. Ma and S. Torquato, J. Appl. Phys. **121**, 244904 (2017).

[20] C. Yeong and S. Torquato, Phys. Rev. E **57**, 495 (1998).

[21] C. E. Zachary and S. Torquato, Phys. Rev. E **84**, 056102 (2011).

[22] E.-Y. Guo, N. Chawla, T. Jing, S. Torquato, and Y. Jiao, Mater Charact **89**, 33 (2014).

[23] S. Torquato, Phys. Rev. Lett **64**, 2644 (1990).

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

[25] C. E. Zachary and S. Torquato, J. Stat. Mech.: Theory Exp. **2009**, P12015 (2009).

[26] B. Lu and S. Torquato, J. Chem. Phys. **93**, 3452 (1990).

[27] R. Strey, Colloid Polym. Sci. **272**, 1005 (1994).

[28] S. Dietrich and A. Haase, Phys. Rep. **260**, 1 (1995).

[29] M. Doi, J. Phys. Soc. Jpn **40**, 567 (1976).

[30] S. Torquato and F. Stillinger, Phys. Rev. E **73**, 031106 (2006).

[31] N. Seaton and E. Glandt, J. Chem. Phys. **85**, 5262 (1986).

[32] M. A. Klatt and S. Torquato, Phys. Rev. E **94**, 022152 (2016).

[33] M. Teubner, J. Chem. Phys. **92**, 4501 (1990).

[34] S. Ciccariello, G. Cocco, A. Benedetti, and S. Enzo, Phys. Rev. B **23**, 6474 (1981).

[35] J. Von Neumann, American Society for Metals, Cleveland **108** (1952).

[36] K. R. Mecke, Int. J. Mod. Phys. B **12**, 861 (1998).

[37] S. Ciccariello and A. Benedetti, Phys. Rev. B **26**, 6384 (1982).

[38] R. J. Adler and J. E. Taylor, *Random fields and geometry* (Springer Science & Business Media, 2009).

[39] R. Blumenfeld and S. Torquato, Phys. Rev. E **48**, 4492 (1993).

[40] S. Torquato, T. M. Truskett, and P. G. Debenedetti, Phys. Rev. Lett. **84**, 2064 (2000).

[41] G. Zhang, F. Stillinger, and S. Torquato, J. Chem. Phys. **145**, 244109 (2016).

[42] J. W. Cahn and J. E. Hilliard, J. Chem. Phys. **28**, 258 (1958).

[43] K. Rundman and J. Hilliard, Acta Metall **15**, 1025 (1967).

[44] C. Smolders, J. Van Aartsen, and A. Steenbergen, Kollold. Z. Z. Polym. **243**, 14 (1971).

[45] Q.-X. Liu, A. Doelman, V. Rottschäfer, M. de Jager, P. M. Herman, M. Rietkerk, and J. van de Koppel, Proc. Natl. Acad. Sci. USA **110**, 11905 (2013).

[46] M. Cross and H. Greenside, *Pattern formation and dynamics in nonequilibrium systems* (Cambridge University Press, 2009).

[47] O. Stenzel, O. Pecho, L. Holzer, M. Neumann, and V. Schmidt, AIChE J. (2017).

[48] M. Röding, P. Svensson, and N. Lorén, Comput. Mater. Sci **134**, 126 (2017).

[49] G. Zhang and S. Torquato, Phys. Rev. E **88**, 053312 (2013).