Modeling heterogeneous materials via two-point correlation functions: Basic principles

Y. Jiao
Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey 08544, USA

F. H. Stillinger
Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA

S. Torquato*
Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA;
Princeton Institute for the Science and Technology of Materials, Princeton University, Princeton, New Jersey 08544, USA;
Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey 08544, USA;
Princeton Center for Theoretical Physics, Princeton University, Princeton, New Jersey 08544, USA
(Received 9 May 2007; published 11 September 2007)

Heterogeneous materials abound in nature and man-made situations. Examples include porous media, biological materials, and composite materials. Diverse and interesting properties exhibited by these materials result from their complex microstructures, which also make it difficult to model the materials. Yeong and Torquato [Phys. Rev. E **57**, 495 (1998)] introduced a stochastic optimization technique that enables one to generate realizations of heterogeneous materials from a prescribed set of correlation functions. In this first part of a series of two papers, we collect the known necessary conditions on the standard two-point correlation function $S_2(\mathbf{r})$ and formulate a conjecture. In particular, we argue that given a complete two-point correlation function space, $S_2(\mathbf{r})$ of any statistically homogeneous material can be expressed through a map on a selected set of bases of the function space. We provide examples of realizable two-point correlation functions and suggest a set of analytical basis functions. We also discuss an exact mathematical formulation of the (re)construction problem and prove that $S_2(\mathbf{r})$ cannot completely specify a two-phase heterogeneous material alone. Moreover, we devise an efficient and isotropy-preserving construction algorithm, namely, the lattice-point algorithm to generate realizations of materials from their two-point correlation functions based on the Yeong-Torquato technique. Subsequent analysis can be performed on the generated images to obtain desired macroscopic properties. These developments are integrated here into a general scheme that enables one to model and categorize heterogeneous materials via two-point correlation functions. We will mainly focus on basic principles in this paper. The algorithmic details and applications of the general scheme are given in the second part of this series of two papers.

DOI: 10.1103/PhysRevE.76.031110 PACS number(s): 05.20.−y, 61.43.−j

I. INTRODUCTION

A heterogeneous material (medium) is one that is composed of domains of different materials or phases (e.g., a composite) or the same material in different states (e.g., a polycrystal). Such materials are ubiquitous; examples include sandstones, granular media, animal and plant tissue, gels, foams, and concrete. The microstructures of heterogeneous materials can only be characterized statistically via various types of n-point correlation functions [1]. The effective transport, mechanical, and electromagnetic properties of heterogeneous materials are known to be dependent on an infinite set of correlation functions that statistically characterize the microstructure [1].

Reconstruction of heterogeneous materials from a knowledge of limited microstructural information (a set of lower-order correlation functions) is an intriguing inverse problem [2–5]. An effective reconstruction procedure enables one to generate accurate structures and subsequent analysis can be performed on the image to obtain macroscopic properties of the materials; see, e.g., Ref. [6]. This provides a nondestructive means of estimating the macroscopic properties: a problem of important technological relevance. Another useful application is reconstruction of a three-dimensional structure of the heterogeneous material using information extracted from two-dimensional planar cuts through the material [3]. Such reconstructions are of great value in a wide variety of fields, including petroleum engineering, biology, and medicine, because in many cases one only has two-dimensional information such as a micrograph or image. Generating realizations of heterogeneous materials from a set of hypothetical correlation functions is often referred to as a "construction" problem [2]. A successful means of construction enables one to identify and categorize materials based on their correlation functions. One can also determine how much information is contained in the correlation functions and test realizability of various types of hypothetical correlation functions. Furthermore, an effective (re)construction procedure can be employed to investigate any physical phenomena where the understanding of spatiotemporal patterns is fundamental, such as in turbulence [1,7].

A popular (re)construction procedure is based on the use of Gaussian random fields: successively passing a normalized uncorrelated random Gaussian field through a linear and

*torquato@electron.princeton.edu

 ©2007 The American Physical Society

then a nonlinear filter to yield the discrete values representing the phases of the structure. The mathematical background used in the statistical topography of Gaussian random fields was originally established in the work of Rice [8,9]. Many variations of this method have been developed and applied since then [10–13]. The Gaussian-field approach assumes that the spatial statistics of a two-phase random medium can be completely described by specifying only the volume fraction and standard two-point correlation function $S_2(\mathbf{r})$, which gives the probability of finding two points separated by vector distance $\mathbf{r}$ in one of the phases [1]. However, to reproduce Gaussian statistics it is not enough to impose conditions on the first two cumulants only, but also to simultaneously ensure that higher-order cumulants vanish [14]. In addition, the method is not suitable for extension to non-Gaussian statistics, and hence is model dependent.

Recently, Torquato and co-workers have introduced another stochastic (re)construction technique [2–5,15]. In this method, one starts with a given, arbitrarily chosen, initial configuration of a random medium and a set of target functions. The medium can be a dispersion of particlelike building blocks [15] or, more generally, a digitized image [2–5]. The target functions describe the desirable statistical properties of the medium of interest, which can be various correlation functions taken either from experiments or theoretical considerations. The method proceeds to find a realization (configuration) in which calculated correlation functions best match the target functions. This is achieved by minimizing the sum of squared differences between the calculated and target functions via stochastic optimization techniques, such as the simulated annealing method [16]. This method is applicable to multidimensional and multiphase media, and is flexible enough to include any type and number of correlation functions as microstructural information. It is both a generalization and simplification of the aforementioned Gaussian-field (re)construction technique. Moreover, it does not depend on any particular statistics [1].

There are many different types of statistical descriptors that can be chosen as target functions [1]; the most basic one is the aforementioned two-point correlation function $S_2(\mathbf{r})$, which is obtainable from small-angle x-ray scattering [17]. However, not every hypothetical two-point correlation function corresponds to a realizable two-phase medium [1]. Therefore, it is of great fundamental and practical importance to determine the necessary conditions that realizable two-point correlation functions must possess [18,19]. Shepp showed that convex combinations and products of two scaled autocovariance functions of one-dimensional media (equivalent to two-point correlation functions; see definition below) satisfy all known necessary conditions for a realizable scaled autocovariance function [20]. More generally, we will see that a hypothetical function obtained by a particular combination of a set of realizable scaled autocovariance functions corresponding to d-dimensional media is also realizable.

In this paper, we generalize Shepp's work and argue that given a complete two-point correlation function space, $S_2(\mathbf{r})$ of any statistically homogeneous material can be expressed through a map on a selected set of bases of the function space. We collect all known necessary conditions of realizable two-point correlation functions and formulate a conjecture. We also provide examples of realizable two-point correlation functions and suggest a set of analytical basis functions. We further discuss an exact mathematical formulation of the (re)construction problem and show that $S_2(\mathbf{r})$ cannot completely specify a two-phase heterogeneous material alone, apart from the issue of chirality. Moreover, we devise an efficient and isotropy-preserving construction algorithm to generate realizations of materials from their two-point correlation functions. Subsequent analysis can be performed on the generated images to estimate desired macroscopic properties that depend on $S_2(\mathbf{r})$, including both linear [1,21–26] and nonlinear [27,28] behavior. These developments are integrated here into a general scheme that enables one to model and categorize heterogeneous materials via two-point correlation functions. Although the general scheme is applicable in any space dimension, we will mainly focus on two-dimensional media here. In the second part of this series of two papers [29], we will provide algorithmic details and applications of our general scheme.

The rest of this paper is organized as follows. In Sec. II, we briefly introduce the basic quantities used in the description of two-phase random media. In Sec. III, we gather all the known necessary conditions for realizable two-point correlation functions and make a conjecture on a possible necessary condition based on simulation results. In Sec. IV, we propose a general form through which the scaled autocovariance functions can be expressed by a set of chosen basis functions and discuss the choice of basis functions. In Sec. V, we formulate the (re)construction problem using rigorous mathematics and show that $S_2(\mathbf{r})$ alone cannot completely specify a two-phase random medium. Thus, it is natural to solve the problem by stochastic optimization methods (e.g., simulated annealing). The optimization procedure and the lattice-point algorithm are also discussed. In Sec. VI, we provide several illustrative examples. In Sec. VII, we make concluding remarks.

II. DEFINITIONS OF n-POINT CORRELATION FUNCTIONS

The ensuing discussion leading to the definitions of the n-point correlation functions follows closely the one given by Torquato [1]. Consider a realization of a two-phase random heterogeneous material within d-dimensional Euclidean space $\mathbb{R}^d$. To characterize this binary system, in which each phase has volume fraction ϕ_i $(i=1,2)$, it is customary to introduce the indicator function $I^{(i)}(\mathbf{x})$ defined as

$$I^{(i)}(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in V_i, \\ 0, & \mathbf{x} \in \overline{V_i}, \end{cases} \tag{1}$$

where $V_i \in \mathbb{R}^d$ is the region occupied by phase i and $\overline{V_i} \in \mathbb{R}^d$ is the region occupied by the other phase. The statistical characterization of the spatial variations of the binary system involves the calculation of n-point correlation functions

$$S_n^{(i)}(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n) = \langle I^{(i)}(\mathbf{x}_1)I^{(i)}(\mathbf{x}_2)\cdots I^{(i)}(\mathbf{x}_n)\rangle, \tag{2}$$

where the angular brackets $\langle\cdots\rangle$ denote ensemble averaging over independent realizations of the random medium.

For statistically homogeneous media, the n-point correlation function depends not on the absolute positions but on their relative displacements, i.e.,

$$S_n^{(i)}(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n) = S_n^{(i)}(\mathbf{x}_{12},\ldots,\mathbf{x}_{1n}), \tag{3}$$

for all $n \geqslant 1$, where $\mathbf{x}_{ij}=\mathbf{x}_j-\mathbf{x}_i$. Thus, there is no preferred origin in the system, which in Eq. (3) we have chosen to be the point $\mathbf{x}_1$. In particular, the one-point correlation function is a constant everywhere, namely, the volume fraction ϕ_i of phase i, i.e.,

$$S_1^{(i)} = \langle I^{(i)}(\mathbf{x})\rangle = \phi_i, \tag{4}$$

and it is the probability that a randomly chosen point in the medium belongs to phase i. For statistically isotropic media, the n-point correlation function is invariant under rigid-body rotation of the spatial coordinates. For $n \leqslant d$, this implies that $S_n^{(i)}$ depends only on the distances $x_{ij}=|\mathbf{x}_{ij}|$ $(1 \leqslant i < j \leqslant n)$. For $n \geqslant d+1$, it is generally necessary to retain vector variables because of chirality of the medium.

The two-point correlation function $S_2^{(i)}(\mathbf{x}_1,\mathbf{x}_2)$ defined as

$$S_2^{(i)}(\mathbf{x}_1,\mathbf{x}_2) = \langle I^{(i)}(\mathbf{x}_1)I^{(i)}(\mathbf{x}_2)\rangle, \tag{5}$$

is one of the most important statistical descriptors of random media. It also is as the probability that two randomly chosen points $\mathbf{x}_1$ and $\mathbf{x}_2$ both lie in phase i. For statistically homogeneous media, $S_2^{(i)}$ only depends on relative displacements, i.e.,

$$S_2^{(i)}(\mathbf{x}_1,\mathbf{x}_2) = S_2^{(i)}(\mathbf{r}), \tag{6}$$

where $\mathbf{r}=\mathbf{x}_{12}$, For statistically homogeneous and isotropic media, the two-point function only depends on scalar distances, i.e., $S_2^{(i)}(r)$, where $r=|\mathbf{r}|$.

Global information about the surface of the ith phase may be obtained by ensemble averaging the gradient of $I^{(i)}(\mathbf{x})$. Since $\nabla I^{(i)}(\mathbf{x})$ is different from zero only on the interfaces of the ith phase, the corresponding specific surface s_i defined as the total area of the interfaces divided by the volume of the medium is given by [1]

$$s_i = \langle |\nabla I^{(i)}(\mathbf{x})|\rangle. \tag{7}$$

Note that there are other higher-order surface correlation functions which are discussed in detail by Torquato [1].

The calculation of higher-order correlation functions encounters both analytical and numerical difficulties, and very few experimental results needed for comparison purposes are available so far. However, their importance in the description of collective phenomena is indisputable. A possible pragmatic approach is to study more complex lower-order correlation functions; for instance, the two-point cluster function $C^{(i)}(\mathbf{x}_1,\mathbf{x}_2)$ defined as the probability that two randomly chosen points $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to the same cluster of phase i [30]; or the lineal-path function $L^{(i)}(\mathbf{x}_1,\mathbf{x}_2)$ defined as the probability that the entire line segment between points $\mathbf{x}_1$ and $\mathbf{x}_2$ lies in phase i [31]. $C^{(i)}(\mathbf{x}_1,\mathbf{x}_2)$ and $L^{(i)}(\mathbf{x}_1,\mathbf{x}_2)$ of the reconstructed media are sometimes computed to study the nonuniqueness issue of the reconstruction [2–5].

III. NECESSARY CONDITIONS ON THE TWO-POINT CORRELATION FUNCTION

The task of determining the necessary and sufficient conditions that $S_2^{(i)}(\mathbf{r})$ must possess is very complex. In the context of stochastic processes in time (one-dimensional processes), it has been shown that the autocovariance functions must not only meet all the necessary conditions we will present in this section but another condition on "corner-positive" matrices [32]. Since little is known about corner-positive matrices, this theorem is very difficult to apply in practice. Thus, when determining whether a hypothetical function is realizable or not, we will first check all the necessary conditions collected here and then use the construction technique to generate realizations of the random medium associated with the hypothetical function as further verification.

A. Known necessary conditions

Here we collect all of the known necessary conditions on S_2 [1,18–20]. For a two-phase statistically homogeneous medium, the two-point correlation function for phase 2 is simply related to the corresponding function for phase 1 via the expression

$$S_2^{(2)}(\mathbf{r}) = S_2^{(1)}(\mathbf{r}) - 2\phi_1 + 1, \tag{8}$$

and the autocovariance function

$$\chi(\mathbf{r}) \equiv S_2^{(1)}(\mathbf{r}) - {\phi_1}^2 = S_2^{(2)}(\mathbf{r}) - {\phi_2}^2, \tag{9}$$

for phase 1 is equal to that for phase 2. Generally, for $\mathbf{r}=\mathbf{0}$,

$$S_2^{(i)}(\mathbf{0}) = \phi_i, \tag{10}$$

and in the absence of any long-range order

$$\lim_{|\mathbf{r}|\to\infty} S_2^{(i)}(\mathbf{r}) \to {\phi_i}^2. \tag{11}$$

An important necessary condition of realizable $S_2^{(i)}(\mathbf{r})$ for a two-phase statistically homogeneous medium in $\mathbb{R}^d$ is that the d-dimensional Fourier transform of the autocovariance function $\chi(\mathbf{r})$, denoted by $\tilde{\chi}(\mathbf{k})$ must be non-negative for all wave vectors [1], i.e., for all $\mathbf{k}$

$$\tilde{\chi}(\mathbf{k}) = \int_{\mathbb{R}^d} \chi(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \geqslant 0. \tag{12}$$

This non-negativity result is sometimes called the Wiener-Khintchine condition, which physically results since $\tilde{\chi}(\mathbf{k})$ is proportional to the scattered radiation intensity. The two-point correlation function must satisfy the following bounds for all $\mathbf{r}$:

$$0 \leqslant S_2^{(i)}(\mathbf{r}) \leqslant \phi_i, \tag{13}$$

and the corresponding bounds on the autocovariance function are given by

$$-\min({\phi_1}^2, {\phi_2}^2) \leqslant \chi(\mathbf{r}) \leqslant \phi_1\phi_2. \tag{14}$$

A corollary of Eq. (14) recently derived by Torquato [19] states that the infimum of any two-point correlation function

of a statistically homogeneous medium must satisfy the inequalities

$$\max(0,2\phi_i-1) \leqslant \inf[S_2^{(i)}(\mathbf{r})] \leqslant \phi_i^2. \tag{15}$$

Another necessary condition on $S_2^{(i)}(\mathbf{r})$ in the case of statistically homogeneous and isotropic media, i.e., when $S_2^{(i)}(\mathbf{r})$ is dependent only the distance $r \equiv |\mathbf{r}|$, is that its derivative at $r=0$ is strictly negative for all $0<\phi_i<1$:

$$\left.\frac{dS_2^{(i)}}{dr}\right|_{r=0} = \left.\frac{d\chi}{dr}\right|_{r=0} < 0. \tag{16}$$

This is a consequence of the fact that slope at $r=0$ is proportional to the negative of the specific surface [1]. Taking that it is axiomatic that $S_2^{(i)}(|\mathbf{r}|)$ is an even function, i.e., $S_2^{(i)}(|\mathbf{r}|) = S_2^{(i)}(-|\mathbf{r}|)$, then it is nonanalytic at the origin.

A lesser-known necessary condition for statistically homogeneous media is the so-called "triangular inequality" that was first derived by Shepp [20] and later rediscovered by Matheron [33]

$$S_2^{(i)}(\mathbf{r}) \geqslant S_2^{(i)}(\mathbf{s}) + S_2^{(i)}(\mathbf{t}) - \phi_i, \tag{17}$$

where $\mathbf{r}=\mathbf{t}-\mathbf{s}$. Note that if the autocovariance $\chi(\mathbf{r})$ of a statistically homogeneous and isotropic medium is monotonically decreasing, nonnegative and convex (i.e., $d^2\chi/d^2r \geqslant 0$), then it satisfies the triangular inequality (17). The triangular inequality implies several pointwise conditions on the two-point correlation function. For example, for statistically homogeneous and isotropic media, the triangular inequality implies the condition given by Eq. (16), the fact that the steepest descent of the two-point correlation function occurs at the origin [20]:

$$\left|\left.\frac{dS_2^{(i)}(r)}{dr}\right|_{r=0}\right| \geqslant \left|\frac{dS_2^{(i)}(r)}{dr}\right|, \tag{18}$$

and the fact that $S_2^{(i)}(r)$ must be convex at the origin [34]:

$$\left.\frac{d^2S_2^{(i)}}{dr^2}\right|_{r=0} = \left.\frac{d^2\chi}{dr^2}\right|_{r=0} \geqslant 0. \tag{19}$$

Torquato [19] showed that the triangular inequality is actually a special case of the more general condition

$$\sum_{i=1}^{m}\sum_{j=1}^{m} \varepsilon_i\varepsilon_j\chi(\mathbf{r}_i-\mathbf{r}_j) \geqslant 1, \tag{20}$$

where $\varepsilon_i=\pm 1$ ($i=1,\ldots,m$ and m is odd). Note that by choosing $m=3$; $\varepsilon_1\varepsilon_2=1$, $\varepsilon_1\varepsilon_3=\varepsilon_2\varepsilon_3=-1$, Eq. (17) can be rediscovered. If $m=3$; $\varepsilon_1\varepsilon_2=\varepsilon_1\varepsilon_3=\varepsilon_2\varepsilon_3=1$ are chosen instead, another "triangular inequality" can be obtained, i.e.,

$$S_2^{(i)}(\mathbf{r}) \geqslant -S_2^{(i)}(\mathbf{s}) - S_2^{(i)}(\mathbf{t}) + (4\phi_i^2-\phi_i), \tag{21}$$

where $\mathbf{r}=\mathbf{t}-\mathbf{s}$. Equation (21) was first derived by Quintanilla [35].

Equation (20) is a much stronger necessary condition that implies that there are other necessary conditions beyond those identified thus far. However, Eq. (20) is difficult to check in practice, because it does not have a simple spectral analog. One possible method is to randomly generate a set of m points and compute the value of $\chi_{ij}=\chi(\mathbf{r}_i-\mathbf{r}_j)$. Among these values of χ_{ij}, select the largest m ones and set their coefficients $\varepsilon_i\varepsilon_j$ equal to -1. Thus, we have m equations for m ε_i's. Then we can substitute the solved ε_i's into Eq. (20) and check the inequality. If the inequality holds, then we can generate several different sets of random points and test the inequality in the same way.

B. Conjecture on a necessary condition

In addition to the aforementioned explicit necessary conditions, we find in two-dimensional simulations that the value of the second peak of a nonmonotonic $S_2(r)$ for a statistically homogeneous and isotropic two-dimensional medium is never larger than that of the medium composed of circular disks on a triangular lattice at fixed volume fractions, i.e.,

$$S_2^{(i)}(r_p;\phi_i) \leqslant S_2^{(i)}(r_p';\phi_i)^{\text{tri}}, \tag{22}$$

where r_p and r_p' denote the positions of the second peak for the two media, respectively and the superscript "tri" denotes the medium composed of disks on a triangular lattice. A hypothetical damped-oscillating two-point correlation function with an artificially higher second peak than that of the medium composed of disks on a triangular lattice at fixed volume fractions was tested by the construction algorithm. The results in Fig. 1 show that the structure for which S_2^B ("B" denotes the black phase) best matches the target function indeed has its "particles" arranged on triangular lattice while the second peak of the target function still cannot be reached.

Here we make the conjecture that for any d-dimensional statistically homogeneous and isotropic medium with a two-point correlation function $S_2(r)$ that is nonmonotonic in r, the value of the first peak of its $S_2^{(i)}(r)$ away from the origin is bounded from above by the value of the first peak of the two-point correlation function associated with the densest packings of d-dimensional identical hard spheres at fixed a volume fraction, i.e.,

$$S_2^{(i)}(r_p;\phi_i) \leqslant S_2^{(i)}(r_p';\phi_i)^{\text{CPS}}, \tag{23}$$

where the superscript "CPS" denotes closest packings of spheres, for the first three dimensions. They are the regular array of hard rods, hard disks on triangular lattice, and hard spheres on face-centered cubic lattice, respectively. For $d=4$ and $d=5$, the densest packings are believed to be four- and five-dimensional checkerboard lattice packings, respectively [36]. Note that this conjectured condition could be a corollary of Eq. (20) or some other unknown necessary conditions.

IV. MODELING TWO-POINT CORRELATION FUNCTION VIA BASIS FUNCTIONS

A. Combination of realizable two-point correlation functions

It is first shown by Shepp [20] that the convex combination and product of two realizable scaled autocovariance functions for one-dimensional statistically homogeneous media satisfy all known necessary conditions, i.e.,

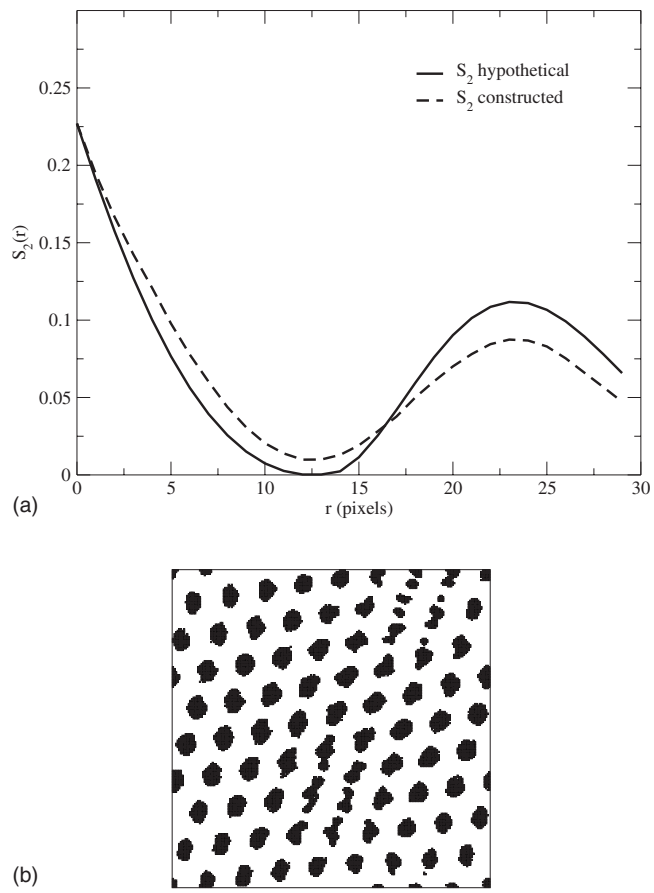

FIG. 1. Numerical support of the conjecture: (a) Two-point correlation functions of black phase for hypothetical and constructed medium. (b) Constructed medium for which S_2^B best matches the target one. The linear size of the system $N=200$ (pixels), volume fraction of black pixels $\phi_1=0.227$.

$$f^c(r) = \alpha_1 f_1(r) + \alpha_2 f_2(r), \quad f^p(r) = f_1(r) f_2(r), \tag{24}$$

where $0 \leqslant \alpha_i \leqslant 1$ $(i=1,2)$, $\alpha_1+\alpha_2=1$ and the superscripts "c" and "p" denote "combination" and "product," respectively. The scaled autocovariance function $f(\mathbf{r})$ of a statistically homogeneous material is defined as [19]

$$f(\mathbf{r}) \equiv \frac{\chi(\mathbf{r})}{\phi_1 \phi_2} = \frac{S_2^{(i)}(\mathbf{r}) - \phi_i^2}{\phi_1 \phi_2}. \tag{25}$$

The necessary conditions for realizable scaled autocovariance function $f(\mathbf{r})$ can be obtained from Eq. (25) and the equations by which the necessary conditions for realizable two-point correlation function $S_2^{(i)}(\mathbf{r})$ are given. From Eqs. (17) and (21), we can obtain the triangular inequalities for $f(\mathbf{r})$, respectively,

$$f(\mathbf{r}) \geqslant f(\mathbf{s}) + f(\mathbf{t}) - 1, \tag{26}$$

$$f(\mathbf{r}) \geqslant -f(\mathbf{s}) - f(\mathbf{t}) - 1. \tag{27}$$

Moreover, the bounds of $f(\mathbf{r})$ become

$$-\min\left[\frac{\phi_1}{\phi_2}, \frac{\phi_2}{\phi_1}\right] \leqslant f(\mathbf{r}) \leqslant 1, \tag{28}$$

and the corollary Eq. (15) is equivalent to

$$-\min\left[\frac{\phi_1}{\phi_2}, \frac{\phi_2}{\phi_1}\right] \leqslant f_{\text{inf}} \leqslant 0, \tag{29}$$

where f_{inf} is the infimum of $f(\mathbf{r})$. Our focus in this paper will be hypothetical continuous functions $f(r)$ that are dependent only on the scalar distance $r=|\mathbf{r}|$ which could potentially correspond to statistically homogeneous and isotropic media without long range order, i.e.,

$$f(0) = 1, \quad \lim_{r\to\infty} f(r) \to 0. \tag{30}$$

$f(r)$ is also absolutely integrable so that the Fourier transform of $f(r)$ exists and is given by

$$\tilde{f}(k) = (2\pi)^{d/2} \int_0^\infty r^{d-1} f(r) \frac{J_{(d/2)-1}(kr)}{(kr)^{(d/2)-1}} dr \geqslant 0, \tag{31}$$

where $k=|\mathbf{k}|$ and $J_v(x)$ is the Bessel function of order v.

Generalization of Eq. (24) to higher dimensions is straightforward. Suppose $f_i(r)$ $(i=1,\ldots,m)$ are the scaled autocovariance functions for d-dimensional statistically homogeneous and isotropic media, then the convex combination $f^c(r)$ and product $f^p(r)$ defined as

$$f^c(r) = \sum_{i=1}^m \alpha_i f_i(r), \quad f^p(r) = \prod_{i=1}^m f_i(r), \tag{32}$$

satisfy all known necessary conditions, where $0 \leqslant \alpha_i \leqslant 1$ $(i=1,\ldots,m)$ and $\Sigma_{i=1}^m \alpha_i = 1$. Equation (32) is of great fundamental and practical importance. On the one hand, it enables us to construct realizable two-point correlation functions with properties of interest, corresponding to structures of interest, from a set of known functions. Thus, one can categorize microstructures with the set of known functions and the proper combinations. On the other hand, suppose that we can find a "full" set of those *basis* scaled autocovariance functions $\{f_i(r)\}_{i=1}^m$, then the scaled autocovariance function of any statistically homogeneous and isotropic medium can be expressed in term of the combinations of the basis functions, i.e.,

$$f(r) = \wp[\{f_i(r)\}_{i=1}^m] \equiv \wp[f_1(r), f_2(r), \ldots, f_m(r)], \tag{33}$$

where $\wp$ denotes a map composed of convex combinations and products of $\{f_i(r)\}_{i=1}^m$. For example, for $m=5$, a possible explicit form for $\wp$ is

$$\wp[\{f_i(r)\}_{i=1}^5] = \alpha_1 f_1(r) + \alpha_2[\beta_1 f_2(r) + \beta_2 f_3(r)] + \alpha_3[f_4(r) f_5(r)], \tag{34}$$

where $0 \leqslant \alpha_i$, $\beta_j \leqslant 1$ and $\Sigma_i \alpha_i = \Sigma_j \beta_j = 1$ $(i=1,2,3; j=1,2)$. Once the scaled autocovariance function (or equivalently the two-point correlation function) of a medium is known, an effective reconstruction procedure enables one to generate accurate structures at will, and subsequent analysis can be performed on the image to obtain desired macroscopic prop-

erties of the medium. In other words, the medium is actually modeled by a set of basis scaled autocovariance functions $\{f_i(r)\}_{i=1}^m$ and a particular map $\wp[\{f_i(r)\}_{i=1}^m]$. There could be different choices of the basis functions (like different basis choices of a Hilbert space), and we would like the basis functions to have nice mathematical properties, such as simple analytical forms. Let $\{f_i^0(r)\}_{i=1}^m$ denote our choice of the basis functions. Thus, the media can be represented merely by different maps $\wp^0$'s. Note that a hypothetical two-point correlation function corresponds to a hypothetical map $\wp_h^0$ and effective construction algorithms can be used to test the realizability of $\wp_h^0$.

B. Choice of basis functions

A systematic way of determining the basis functions $\{f_i^0\}_{i=1}^m$ is not available yet. Here we take the first step to determine the bases by considering certain known realizable analytical two-point correlation functions and the corresponding scaled autocovariance functions. For convenience, we categorize these functions into three families: (i) monotonically decreasing functions, (ii) damped-oscillating functions, and (iii) functions of known constructions.

The family of monotonically decreasing functions includes the simple exponentially decreasing function introduced by Debye [17] and polynomial functions. The former is given by

$$f_D(r) = \exp(-r/a), \quad r \geqslant 0, \tag{35}$$

where a is a correlation length, corresponding to structures in which one phase consists of "random shapes and sizes" [17,37] (shown in Fig. 2). It is now known that certain types of space tessellations have autocovariance functions given by Eq. (35) [38]. We have referred to this class of structures as Debye random media [1,2].

Another example of monotonically decreasing functions is the family of polynomials of order n $(n \geqslant 1)$ given by

$$f_P^n(r) = \begin{cases} (1-r/a)^n, & 0 \leqslant r \leqslant a, \\ 0, & r > a, \end{cases} \tag{36}$$

where a is the correlation length. The polynomial function of order 1 is shown to be realizable only for a statistically homogeneous two-phase medium in one dimension [19,20]. We have constructed realizations of random media in dimensions $d \leqslant n$ that correspond to the polynomial function of order n with very high numerical precision using the Yeong-Torquato construction technique [2]. However, for dimensions higher than n, Eq. (36) violates Eq. (31). We will henceforth assume that the polynomial functions of order n are realizable in dimensions $d \leqslant n$.

An example of the family of damped-oscillating functions is given by [2,19]

$$f_O(r) = \sum_i A_i \exp(-r/a_i)\cos(q_i r + \psi_i), \quad r \geqslant 0, \tag{37}$$

where the parameters A_i and a_i $(i=1,2,\ldots)$ control the amplitude of the f_O profile, q_i is the wavenumber, and ψ_i is the phase angle The form of (37) could be further generalized by

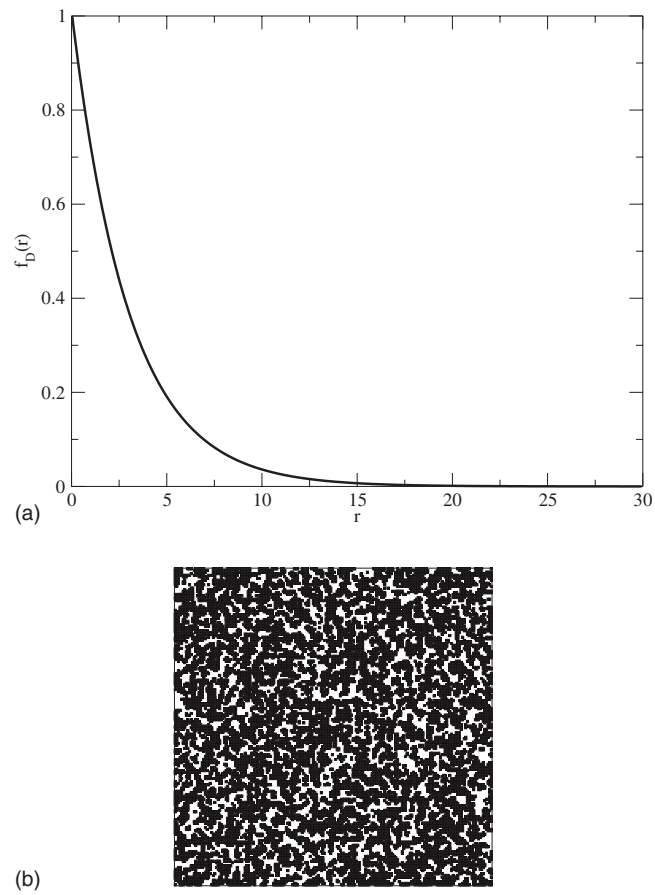

FIG. 2. (a) Debye random medium function $f_D(r)$ with $a=5$. (b) A realization of Debye random media with the volume fractions $\phi_1=0.68$, $\phi_2=0.32$.

changing the variable of cosine functions from a linear function of r to a more general monotonically increasing function of r in order to capture more complex local microstructural features. Note that q_i and ψ_i need to be carefully chosen such that $f_O(r)$ satisfies all known necessary conditions.

The family of functions of known constructions includes scaled autocovariance functions of d-dimensional identical overlapping spheres [1,39] and symmetric-cell materials (see Fig. 3) [1,40]. For overlapping spheres of radius R, the scaled autocovariance function for the particle phase (spheres) is given by

$$f_S(r) = \frac{\exp[-\rho v_2(r;R)] - \phi_1^2}{\phi_1\phi_2}, \tag{38}$$

where ϕ_1 and ϕ_2 are volume fractions of the spheres and matrix, respectively, $\rho = N/V$ is the number density of spheres and $v_2(r;R)$ is the union volume of two spheres of radius R whose centers are separated by r. For the first three space dimensions, the latter is, respectively, given by

$$\frac{v_2(r;R)}{v_1(R)} = 2\Theta(r-2R) + \left(1 + \frac{r}{2R}\right)\Theta(2R-r), \tag{39}$$

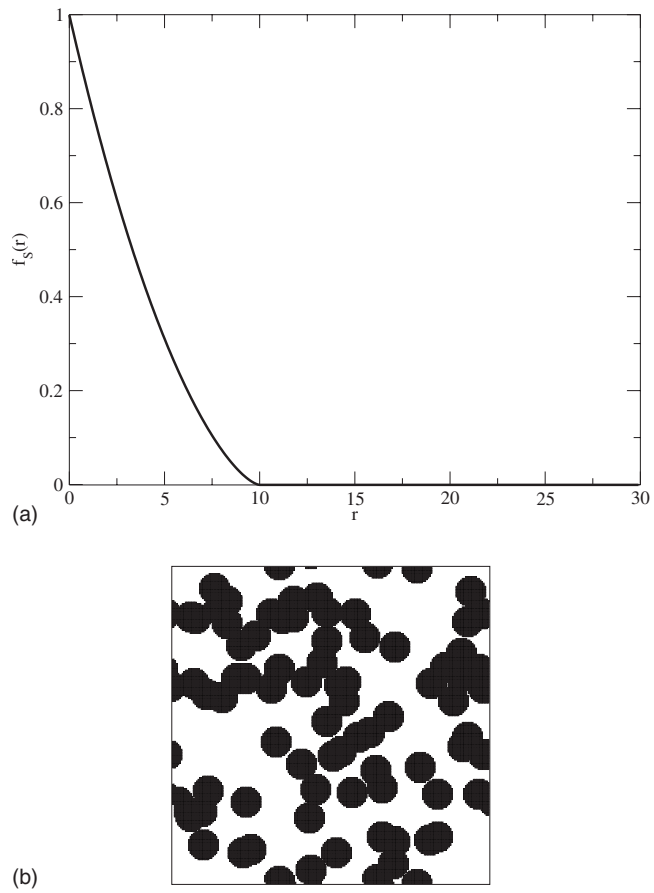

FIG. 3. (a) Scaled autocovariance function $f_S(r)$ of two-dimensional identical overlapping disks with volume fractions $\phi_1=0.45$, $\phi_2=0.55$. (b) A realization of two-dimensional identical overlapping disks with volume fractions $\phi_1=0.45$, $\phi_2=0.55$. The radius of disks $R=5$.

$$\frac{v_2(r;R)}{v_1(R)} = 2\Theta(r-2R) + \frac{2}{\pi}\left[\pi + \frac{r}{2R}\left(1-\frac{r^2}{4R^2}\right)^{1/2} - \cos^{-1}\left(\frac{r}{2R}\right)\right]\Theta(2R-r), \tag{40}$$

$$\frac{v_2(r;R)}{v_1(R)} = 2\Theta(r-2R) + \left[1+\frac{3r}{4R}-\frac{1}{16}\left(\frac{r}{R}\right)^3\right]\Theta(2R-r), \tag{41}$$

where $\Theta(x)$ is the Heaviside function and $v_1(R)$ is the volume of a d-dimensional sphere of radius R given by

$$v_1(R) = \frac{\pi^{d/2}}{\Gamma(1+d/2)}R^d, \tag{42}$$

where $\Gamma(x)$ is the gamma function. For $d=1$, 2, and 3, $v_1(R)=2R$, πR^2, and $4\pi R^3/3$, respectively.

Two-phase symmetric-cell materials are constructed by partitioning space into cells of arbitrary shapes and sizes, with cells being randomly designated as phase 1 and 2 with probability ϕ_1 and ϕ_2 from a uniform distribution [1]. For such a statistically homogeneous and isotropic medium, the scaled autocovariance function is given by

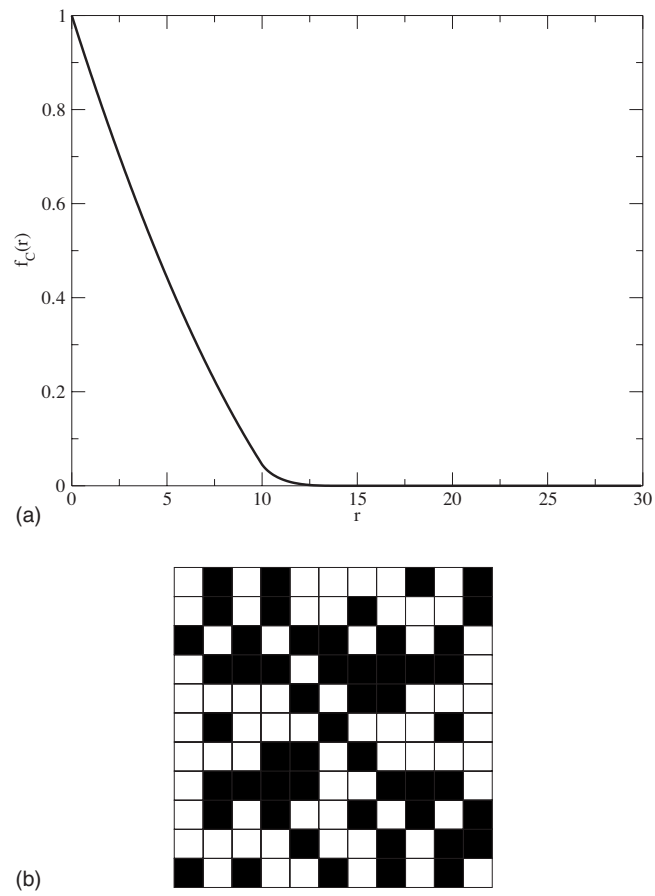

FIG. 4. (a) Scaled autocovariance function $f_C(r)$ of two-dimensional random checkerboard. (b) A realization of two-dimensional random checkerboard. The length of the cells $a=10$.

$$f_C(r) = W_2^{(1)}(r), \tag{43}$$

where $W_2^{(1)}(r)$ is the probability that two points separated by distance r are in the same cell. This quantity is only a function of cell shapes and sizes, depending on a dimensionless size-averaged intersection volume of two cells defined by

$$W_2^{(1)}(r) = \frac{\langle v_2^{\text{int}}(r;R)\rangle_R}{\langle v_1(r)\rangle_R}, \tag{44}$$

where R is the size parameter for each cell, $\langle v_2^{\text{int}}(r;R)\rangle_R$ is the size-averaged intersection volume of two cells whose centers are separated by r, and $\langle v_1(r)\rangle_R$ is the size-averaged single-cell volume. The random checkerboard is a very useful model of symmetric-cell material because its $f_C(r)$ is known analytically for $d=1$ and 2 [1] (see Fig. 4). For $d=1$, it is easy to verify that the probability of finding two points in the

same one-dimensional cell is given by

$$W_2^{(1)}(r) = \begin{cases} 1 - r/a, & 0 \leqslant r \leqslant a, \\ 0, & r \geqslant a, \end{cases} \tag{45}$$

where a is the length of the side of a square cell. Note that Eq. (45) is just the polynomial function of order one [see Eq. (36)] for one-dimensional statistically homogeneous media. For $d=2$, $W_2^{(1)}(r)$ is given by

$$W_2^{(1)}(r) = \begin{cases} 1 + \frac{1}{\pi}\left[\left(\frac{r}{a}\right)^2 - 4\left(\frac{r}{a}\right)\right], & 0 \leqslant r \leqslant a, \\ 1 - \frac{1}{\pi}\left[2 + \left(\frac{r}{a}\right)^2\right] + \frac{4}{\pi}\left[\sqrt{\left(\frac{r}{a}\right)^2 - 1} - \cos^{-1}\left(\frac{a}{r}\right)\right], & a \leqslant r \leqslant \sqrt{2}a, \\ 0, & r \geqslant \sqrt{2}a. \end{cases} \tag{46}$$

For $d=3$, $W_2^{(1)}(r)$ is given by

$$W_2^{(1)}(r) = \frac{2}{\pi}\int_0^{\pi/2}\int_0^{\pi/2} W(r,\theta,\phi)\sin\theta d\theta d\phi, \tag{47}$$

where

$$\begin{aligned} W(r,\theta,\phi) = {} & [1 - r\cos\theta][1 - r\sin\theta\sin\phi][1 - r\sin\theta\cos\phi] \\ & \times \Theta(1 - r\cos\theta)\Theta(1 - r\sin\theta\sin\phi) \\ & \times \Theta(1 - r\sin\theta\cos\phi). \end{aligned} \tag{48}$$

Note that $W_2^{(1)}(r)$ for $d=3$ does not have a simple analytical form.

There are some other known scaled autocovariance functions for different types of materials (e.g., d-dimensional identical hard spheres [1]) and realizable hypothetical functions (e.g., complementary error function, see the Appendix). However we do not include these functions in our basis function set because they do not have simple analytical mathematical forms. Note that some of the basis functions are dimension-dependent (e.g., f_O and f_C), and thus the proper forms of basis functions should be used for different dimensions. In the following discussion, without further specification, we will focus on two-dimensional cases.

V. GENERATING REALIZATIONS OF HETEROGENEOUS MATERIALS

Consider a digitized (i.e., pixelized) representation of a heterogeneous material. Different colored pixels (in a discrete coloring scheme) may have numerous interpretations. The image can reflect different properties, such as the geometry captured by a photographic image, topology of temperature and scalar velocity fields in fluids, distribution of magnitudes of electric and magnetic fields in the medium, or variations in chemophysical properties of the medium. In the last case, typical examples are composite materials in which the different phases may have different thermal, elastic or electromagnetic properties, to name a few.

A. Exact equations for digitized media

Our focus in this paper is two-dimensional, two-phase statistically homogeneous and isotropic random media composed of black and white pixels. Such a system can be represented as a two-dimensional array, i.e.,

$$\mathbf{I} = \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1N} \\ I_{21} & I_{22} & \cdots & I_{2N} \\ \vdots & \ddots & \vdots & \\ I_{N1} & I_{N2} & \cdots & I_{NN} \end{bmatrix}, \tag{49}$$

where the integer N categorizes the linear size of the system (N^2 is the total number of pixels in the system) and the entries I_{ij} $(i,j=1,\ldots,N)$ can only take the value of 0 or 1, which correspond to the white and black phases, respectively. Note that Eq. (49) is only an abstract representation and the real morphological configuration of the medium also depends on the choice of *lattices*. For example, as shown in Fig. 5, the isotropic medium composed of overlapping disks generated on a square lattice and the anisotropic medium composed of orientated ellipses generated on a triangular lattice have the same array presentation. A vector distance in the digitized medium can be uniquely expressed as

$$\mathbf{r} = n_1\mathbf{e}_1 + n_2\mathbf{e}_2, \tag{50}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are lattice vectors for the particular lattice and n_1, n_2 are integers. For example, for square lattice, $\mathbf{e}_1=\mathbf{i}$, $\mathbf{e}_2=\mathbf{j}$, where $\mathbf{i}$, $\mathbf{j}$ are unit vectors along horizontal and vertical directions, respectively; while for triangular lattice $\mathbf{e}_1=\sqrt{3}\mathbf{i}+\mathbf{j}/2$, $\mathbf{e}_2=\mathbf{j}$. Note that other lattices can be used whose symmetry is consistent with the particular anisotropy of the media.

Without loss of generality, we choose the black phase to be the phase of interest and assume that a periodic boundary condition is applied, which is commonly used in computer simulations. The two-point correlation function $S_2(\mathbf{r})$ of the black phase can be calculated based on its probabilistic nature, i.e., the probability of finding two points separated by

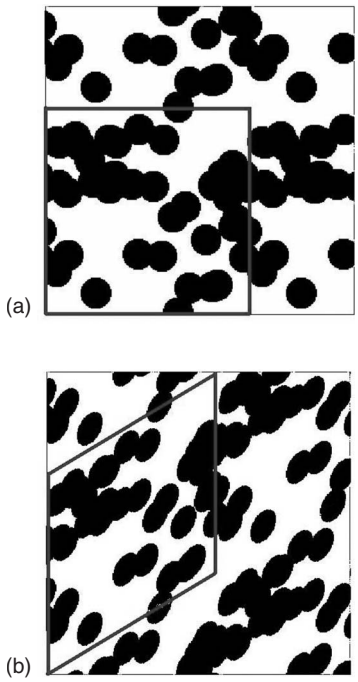

FIG. 5. Different digitized media with the same array representation. (a) Overlapping disks generated on square lattice with a square unit cell. (b) Orientated overlapping ellipses generated on triangular lattice with a rhomboidal unit cell.

the vector distance $\mathbf{r}$ in the same phase. The value of two-point correlation function for a particular $\mathbf{r}=n_1\mathbf{e}_1+n_2\mathbf{e}_2$ is given by

$$S_2(\mathbf{r}) \equiv S_2(n_1,n_2) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} I_{ij}I_{(i+n_1)(j+n_2)}}{N^2}, \tag{51}$$

where I_{ij} are entries of $\mathbf{I}$ defined in Eq. (49), n_1 and n_2 are integers satisfying $|n_1|,|n_2|\leqslant[N/2]$ due to minimum image distance convention. For statistically isotropic media, the two-point correlation function only depends on the magnitude of $\mathbf{r}$, i.e., $r\equiv|\mathbf{r}|$, thus we have

$$S_2(r) = \frac{\sum_{(m,n)\in\Omega}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} I_{ij}I_{(i+m)(j+n)}\right]}{\omega N^2}, \tag{52}$$

where

$$\Omega = \{(m,n)|m^2+n^2=r^2, r\leqslant[N/2]\}, \tag{53}$$

and ω is the number of elements of set Ω.

It is well known that the two-point correlation function cannot completely specify a two-phase heterogeneous material alone. Here, we provide a proof of this statement for two-dimensional statistically homogeneous digitized media. Note that the proof trivially extends to any dimension. Suppose we already know the value of $S_2(\mathbf{r})$ for every vector distance $\mathbf{r}$, using Eq. (51), we can obtain a set for equations of I_{ij}, i.e., for each $\mathbf{r}=n_1\mathbf{e}_1+n_2\mathbf{e}_2$

$$\sum_{i=1}^{N}\sum_{j=1}^{N} I_{ij}I_{(i+n_1)(j+n_2)} - N^2S_2(n_1,n_2) = 0. \tag{54}$$

Since the digitized medium is represented by $\mathbf{I}$, once we obtain all the entries I_{ij} [the unknowns in Eq. (54)], the medium is (re)constructed. In Eq. (54), the number of unknowns N_u equals N^2 and the number of equations N_e equals the number of all possible vector distances in the digitized medium. To calculate N_e, all possible different combinations of integers n_1 and n_2 subjected to $|n_1|,|n_2|\leqslant[N/2]$ need to be considered. This quantity is given by $N_e=4[N/2]^2-8[N/2]+6$, which is smaller than the number of unknowns $N_u=N^2$ for normal-sized systems (i.e., $N=10-10^3$).

Similar proof is applied to the case when we average over the angles of vector $\mathbf{r}$ to yield $S_2(r)$ that depends only on the radial distance r. The angle-averaged equations of I_{ij} are given by

$$\sum_{(m,n)\in\Omega}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} I_{ij}I_{(i+m)(j+n)}\right] - \omega N^2S_2(r) = 0, \tag{55}$$

where Ω is given by Eq. (53). The number of equations is even smaller for the angle-averaged case, while the number of unknowns does not change. The analysis shows that one could never find a unique solution of I_{ij} from either Eq. (54) or Eq. (55) unless some assumptions have been made that reduce N_u such that $N_u=N_e$. For example, in an interesting model to study spatial distribution of algae, the algae are put on top of each other in order to reduce the unknowns [41]. In general cases, one could use the stochastic optimization procedure (i.e., simulated annealing) [2,16] to find solutions of Eqs. (54) and (55). Note that although the aforementioned proofs focus on two-dimensional media, they trivially extend to any dimensions (e.g., $d=1,3$).

B. Stochastic optimization procedure

Generally, consider a given set of correlation functions $f_n^\alpha(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)$ of the phase of interest that provides partial information on the random medium. The index α is used to denote the type of correlation functions. Note that the set of f_n^α should not be confused with the basis function set $\wp$; the former contains correlation functions of different type, i.e., two-point correlation function, lineal-path function, two-point cluster function, etc., while the latter contains basis functions through which the scaled autocovariance function of the medium of interest can be expressed. The information contained in f_n^α could be obtained either from experiments or it could represent a hypothetical medium based on simple models. In both cases we would like to generate the underlying microstructure with a specified set of correlation functions. In the former case, the formulated inverse problem is frequently referred to as a "reconstruction" procedure, and in the latter case as a "construction."

As we have noted earlier, it is natural to formulate the construction or reconstruction problem as an optimization problem [2–5]. The discrepancies between the statistical properties of the best generated structure and the imposed ones is minimized. This can be readily achieved by introduc-

ing the "energy" function E defined as a sum of squared differences between target correlation functions, which we denote by $\hat{f}_n^\alpha$, and those calculated from generated structures, i.e.,

$$E = \sum_{\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n} \sum_\alpha [f_n^\alpha(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n) - \hat{f}_n^\alpha(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)]^2. \tag{56}$$

Note that for every generated structure (configuration), there is a set of corresponding f_n^α. If we consider every structure (configuration) as a "state" of the system, E can be considered as a function of the states. The optimization technique suitable for the problem at hand is the method of simulated annealing [16]. It is a popular method for the optimization of large-scale problems, especially those where a global minimum is hidden among many local extrema. The concept of finding the lowest energy state by simulated annealing is based on a well-known physical fact: If a system is heated to a high temperature T and then slowly cooled down to absolute zero, the system equilibrates to its ground state. At a given temperature T, the probability of being in a state with energy E is given by the Boltzmann distribution $P(E) \sim \exp(-E/T)$. At each annealing step k, the system is allowed to evolve long enough to thermalize at $T(k)$. The temperature is then lowered according to a prescribed annealing schedule $T(k)$ until the energy of the system approaches its ground state value within an acceptable tolerance. It is important to keep the annealing rate slow enough in order to avoid trapping in some metastable states.

In our problem, the discrete configuration space includes the states of all possible pixel allocations. Starting from a given state (current configuration), a new state (new configuration) can be obtained by interchanging two arbitrarily selected pixels of different phases. This simple evolving procedure preserves the volume fraction of all involved phases and guarantees ergodicity in the sense that each state is accessible from any other state by a finite number of interchange steps. However, in the later stage of the procedure, biased and more sophisticated interchange rules, i.e., surface optimization, could be used to improve the efficiency. We choose the Metropolis algorithm as the acceptance criterion: the acceptance probability P for the pixel interchange is given by

$$P(E_{\text{old}} \to E_{\text{new}}) = \begin{cases} 1, & \Delta E < 0, \\ \exp(-\Delta E/T), & \Delta E \geqslant 0, \end{cases} \tag{57}$$

where $\Delta E = E_{\text{new}} - E_{\text{old}}$. The temperature T is initially chosen so that the initial acceptance probability for a pixel interchange with $\Delta E \geqslant 0$ averages approximately 0.5. An inverse logarithmic annealing schedule which decreases the temperature according to $T(k) \sim 1/\ln(k)$ would in principle evolve the system to its ground state. However, such a slow annealing schedule is difficult to achieve in practice. Thus, we will adopt the more popular and faster annealing schedule $T(k)/T(0) = \lambda^k$, where constant λ must be less than but close to one. This may yield suboptimal results, but, for practical purposes, will be sufficient. The convergence to an optimum is no longer guaranteed, and the system is likely to freeze in one of the local minima if the thermalization and annealing rate are not adequately chosen.

The two-point correlation function of a statistically homogeneous and isotropic medium is the focus of this paper. In this case, Eq. (56) reduces to

$$E = \sum_i [S_2(r_i) - \hat{S}_2(r_i)]^2. \tag{58}$$

Since for every configuration (structure), the corresponding two-point correlation function needs to be computed, the efficiency of the construction or reconstruction is mainly determined by the efficiency of the S_2-sampling algorithm. Furthermore, the properties of generated configurations (structures), i.e., isotropy of the medium, are also affected by the S_2-sampling algorithm. One of the most commonly used and efficient S_2-sampling algorithms is the orthogonal-sampling algorithm introduced by Yeong and Torquato [2,3]. Due to the isotropic nature of the medium, every sampling direction should be equivalent. For simplicity, two orthogonal directions (usually the horizontal and vertical directions of a square lattice) are chosen and the two-point correlation function is sampled along these directions and averaged. At each pixel interchange, only the values of $S_2(r)$ sampled along the rows and columns that contain the interchange pixels are changed. Thus, the complexity of the algorithm is $O(N)$, where N is the linear size of the system. However, for certain media with long-range correlations, the generated media have microstructures with two orthogonal anisotropic directions due to the biased sampling. Modifications of the orthogonal-sampling algorithm to preserve the isotropy of the underlying medium have been proposed, such as adding more sampling directions and using more isotropic lattices [4,5]. Cule and Torquato introduced a new isotropy-preserving fast Fourier transform (FFT) algorithm [4]. At each pixel interchange step, the two-point correlation function $S_2(\mathbf{r})$ containing angle information is calculated in momentum space using an efficient FFT algorithm. Since information of all directions is considered, the generated media always have the required isotropy structures. However, since the complexity of FFT is $O(N \log_2 N)$, the algorithm is relatively time consuming.

We have developed an efficient and isotropy-preserving algorithm, namely, the lattice-point algorithm by considering the black pixels as hard "particles" on a particular lattice. The two-point correlation function is then computed in a similar way of obtaining the pair correlation function $g_2(r)$ for an isotropic point process [1]. At each Monte Carlo step, the randomly selected "particle" (black pixel) is given a random displacement subjected to the nonoverlapping constraint and the distances between the moved "particle" and all the other "particles" need to be recomputed. Thus the complexity of the algorithm is $O(N)$. Since all directions are effectively sampled, constructions based on the angle-averaged $S_2(r)$ will preserve isotropy of the media. A detailed discussion and applications of the algorithm will be included in the second paper of this series [29]. In this paper, we only pro-

FIG. 6. A realization of square array of nonoverlapping disks. The linear size of the system $N=200$ pixels, volume fraction of black pixels $\phi_1=0.326$.

vide several illustrative examples generated from this algorithm.

VI. ILLUSTRATIVE EXAMPLES

As illustrative examples, we use the aforementioned (re)construction techniques to investigate both deterministic, crystal-like structures, and random systems. We also study a hypothetical medium with the two-point correlation function obtained from convex combination of known ones. In the case of completely deterministic structures, the algorithm produces almost perfect reconstructions. However, the optimization of disordered structures is significantly harder. Furthermore, we will see that for the media with long-range correlations, e.g., a damped-oscillating $S_2(r)$, the orthogonal-sampling algorithm may produce unexpected anisotropy, while the lattice-point algorithm well preserves isotropy of the media.

A. Regular array of nonoverlapping disks

First, we consider specific two-dimensional and two-phase structure composed of a square array of nonoverlapping disks, as shown in Fig. 6. This morphology may be viewed as a cross section of two-phase materials containing rodlike or fiberlike inclusions. Various transport properties of these materials have been well explored because of their practical and theoretical importance in materials science [42].

The regular structure is discretized by introducing an $N \times N$ square lattice. The volume fractions of black and white phases are $\phi_1=0.326$, $\phi_2=0.674$, respectively. The target two-point correlation of the digitized medium is sampled using both the orthogonal and the lattice-point algorithm for comparison purpose. The simulations start from random initial configurations (i.e., random collections of black and white pixels), at some initial temperature T_0, with fixed volume fractions ϕ_i. At each Monte Carlo (MC) step, when an attempt to exchange two randomly chosen pixels with different colors (or to randomly displace a chosen black pixel) is made, $S_2(r)$ is efficiently recomputed by using the orthogonal-sampling algorithm (or the lattice-point algorithm). The set of constants $\{\lambda_{\mathrm{MC}},\lambda_{\mathrm{tot}},\lambda\}$ specifies the annealing schedule: At each temperature, the system is thermalized until either $\lambda_{\mathrm{MC}}N^2$ MC moves are accepted or the total number of attempts to change the original configurations reaches the value $\lambda_{\mathrm{tot}}N^2$. Subsequently, the system temperature is decreased by the reduction factor λ, i.e., $T_{\mathrm{new}}=\lambda T_{\mathrm{old}}$.

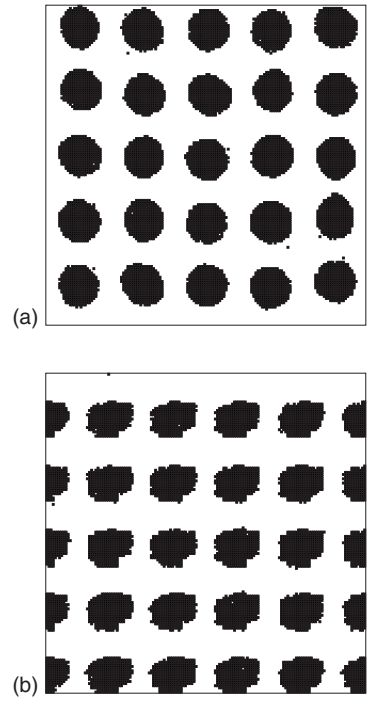

FIG. 7. Reconstructed structures. (a) Square array of almost circular particles generated by the lattice-point algorithm. (b) Square array of particles generated by the orthogonal-sampling algorithm.

The reconstruction results are shown in Fig. 7. Both of the algorithms are able to reproduce the exact global square-array arrangement of clusters of black pixels. This implies that the two-point correlation function of regular configurations contains enough structural information to properly characterize the long-range correlations. However, it is clear that the structure generated by the lattice-point algorithm has a better local arrangement of the pixels (i.e., the shape of the particles) than that generated by the orthogonal-sampling algorithm. This is because the orthogonal algorithm only uses structural information along two directions, which is not sufficient to reproduce detailed local structures, while the lattice-point algorithm efficiently uses information along all possible directions.

B. Hypothetical random media with long-range correlations

In this example, we will generate two-dimensional statistically homogeneous and isotropic random media with long-range correlations (i.e., nontrivial interparticle interactions). Examples of this type of media include low-density fluids and amorphous materials (i.e., porous media, randomly polymerized plastics, glass, etc.). A meaningful, yet nontrivial, two-point correlation function capturing these features is [4,5,19]

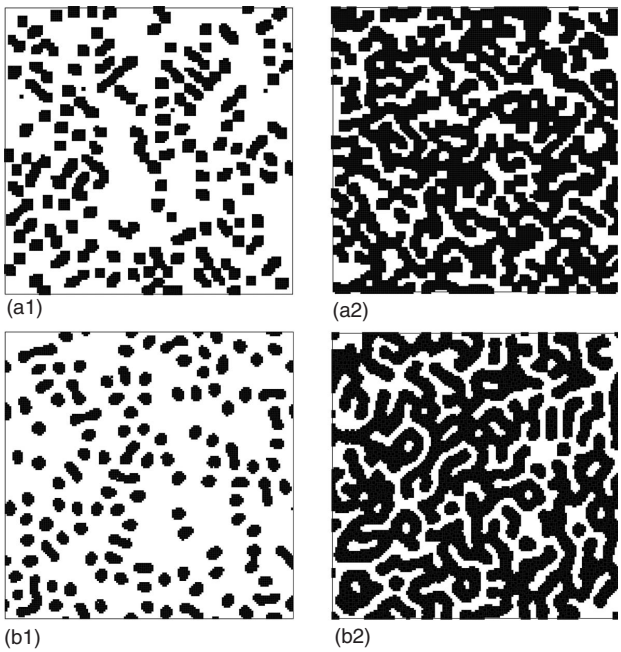

FIG. 8. (a) Media with $\hat{S}_2(r)$ given by Eq. (59) generated by the orthogonal-sampling algorithm. Left panel, volume fraction of black pixels $\phi_1=0.2$. Right panel, volume fraction of black pixels $\phi_1=0.5$. The linear size of the systems $N=200$. (b) Media with $\hat{S}_2(r)$ given by Eq. (59) generated by the lattice-point algorithm. Left panel, volume fraction of black pixels $\phi_1=0.2$. Right panel, volume fraction of black pixels $\phi_1=0.5$. The linear size of the systems $N=200$.

$$\hat{S}_2(r) = \phi_1^2 + \phi_1\phi_2 e^{-r/r_0}\frac{\sin(kr)}{kr}, \tag{59}$$

where $k=2\pi/a_0$. Here r_0 and a_0 are two characteristic length scales. The overall exponential damping is controlled by the correlation length r_0, determining the maximum correlations in the system. The constant a_0 determines oscillations in the term $\sin(kr)/(kr)$ which also decays with increasing r, such that a_0 can reduce the effective range of r_0. Interestingly, this hypothetical function is *not* exactly realizable, because it violates the convexity condition Eq. (19) at the origin, or more generally the triangular inequality Eq. (17). However, we mainly focus on its damped-oscillating property and we will see that the algorithms are robust enough to detect violation of the convexity condition.

For comparison purposes, both the orthogonal-sampling algorithm and the lattice-point algorithm are used in the construction, the results are shown in Fig. 8. At a lower density of the black phase ϕ_1, a_0 is manifested as a characteristic repulsion among different elements with diameter of order a_0. The repulsion vanishes beyond the length scale r_0. At a higher density, both length scales a_0 and r_0 are clearly noticeable in the distribution of the black and white phases. Note that the structures generated by the orthogonal-sampling algorithm exhibit some anisotropy features, i.e., containing stripes along ±45° directions, which implies that the orthogonal-sampling algorithm should be used with care in the case where the medium has long-range correlations.

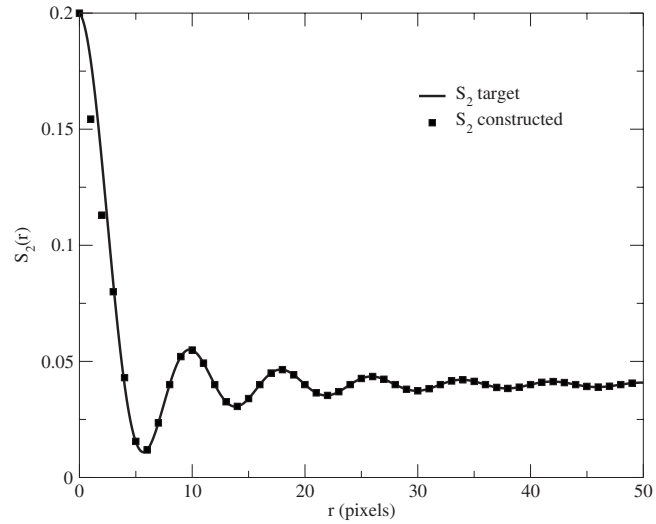

FIG. 9. Target two-point correlation function given by Eq. (59) and that of constructed media with volume fraction $\phi_1=0.2$.

The target two-point correlation function $\hat{S}_2(r)$ for $\phi_1=0.2$ and $S_2(r)$ sampled from generated structures are shown in Fig. 9. It can be seen clearly that $\hat{S}_2(r)$ is nonconvex at the origin and the largest discrepancies between $S_2(r)$ and $\hat{S}_2$ occur around the origin because $S_2(r)$ satisfies Eq. (19). This implies that our algorithms are robust enough and can be used to test realizability of hypothetical functions. Note that the following classes of functions [19]:

$$f(r) = \exp\left[-\left(\frac{r}{a}\right)^{\alpha}\right], \quad \alpha > 1 \tag{60}$$

and

$$f(r) = \frac{1}{[1+(r/a)^2]^{\beta-1}}, \quad \beta \geqslant d \tag{61}$$

cannot correspond to a two-phase medium in d dimensions also because they also violate the triangular inequality (26).

C. Hypothetical random media with realizable correlation functions

In this last example, we study a hypothetical statistically homogeneous and isotropic medium whose scaled autocovariance function is a convex combination of Debye random medium function $f_D(r)$ [cf. Eq. (35)] and damped-oscillating function $f_O(r)$ [cf. Eq. (37)], i.e.,

$$f(r) = \alpha_1 f_D(r) + \alpha_2 f_O(r), \tag{62}$$

where $\alpha_1+\alpha_2=1$. We use the one-term version of $f_O(r)$ here; the parameter a in f_D is $a=20$ (pixels) and the parameters in f_0 are $A_1=1$, $A_i=0$ $(i>1)$, $a_1=5$ (pixels) and $q_1=10$ (pixels). Since both $f_D(r)$ and $f_O(r)$ are independent of volume fractions, the medium with scaled autocovariance function $f(r)$ has phase-inversion symmetry [1], i.e., the morphology of phase 1 at the volume fraction ϕ_1 is statistically identical to

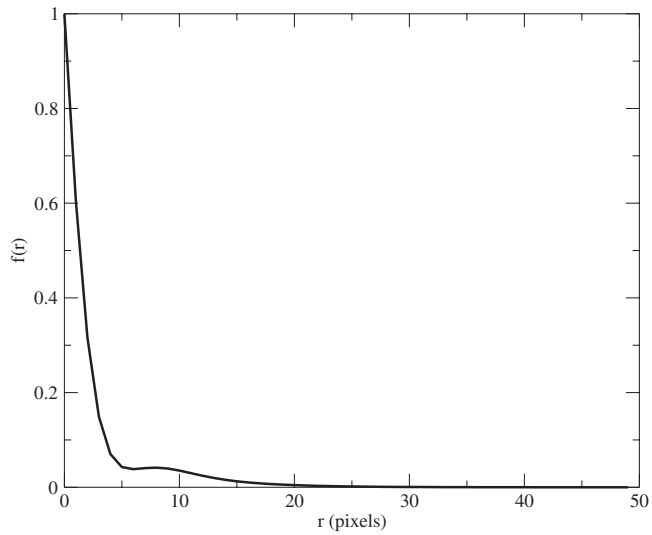

FIG. 10. Combined scaled autocovariance function $f(r)$ according to Eq. (62) with coefficients $\alpha_1=0.25$, $\alpha_2=0.75$. Here $a=20$ (pixels), $A_1=1$, $A_i=0(i>1)$, $a_1=5$ (pixels), $q_1=10$ (pixels).

that of phase 2 when the volume fraction of phase 1 is $1-\phi_1$. (For such media, "phase-interchange" relations for the effective properties [43] can prove to be useful.)

Different constant pairs (α_1,α_2) can be used to construct $f(r)$ with required properties. In particular, we choose two pairs (0.25, 0.75) and (0.75, 0.25). The construction results obtained by application of the lattice-point algorithm are shown in Figs. 10 and 11 and Figs. 12 and 13. For $\alpha_1=0.25$, $\alpha_2=0.75$, $f_O(r)$ is dominant in the combination. At lower densities, the generated structures resemble those with "pure" damped-oscillating two-point functions, i.e., dispersions of particles, although they contain more clusters. At higher densities, some stripe-like structures and several (almost equal-sized) clusters can be identified. For $\alpha_1=0.75$, $\alpha_2=0.25$, $f_D(r)$ is dominant in the combined $f(r)$. Clusters with considerable sizes form even at low densities, which is a consequence of the large effective correlation length in $f_D(r)$. However, no stripelike structures can be identified in the generated structures, since the oscillating feature of $f_O(r)$ is significantly suppressed by its exponentially damping part and the small value of α_1. The results imply that even a simple combination of two basis functions enables one to obtain scaled autocovariance functions with properties of interest and to generate a variety of structures with controllable morphological features, e.g., local "particle" shape and cluster size.

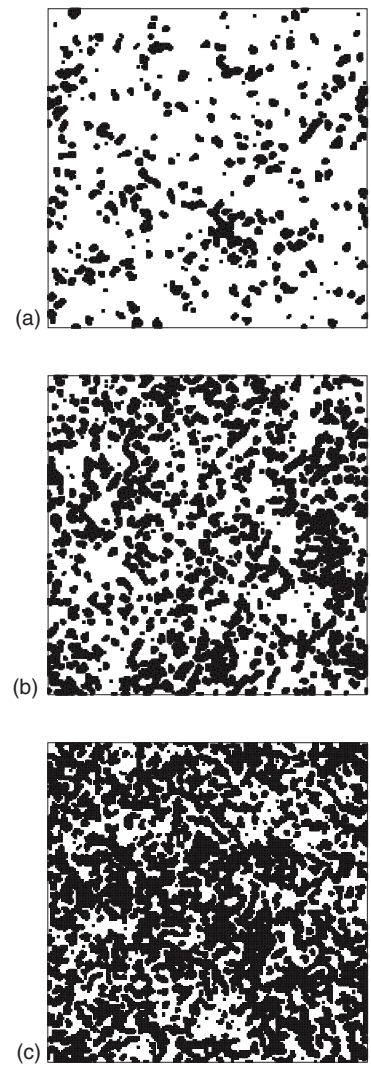

FIG. 11. Constructed media with scaled autocovariance function shown in Fig. 10. (a) Volume fraction of black pixels $\phi_1=0.1$. (b) Volume fraction of black pixels $\phi_1=0.3$. (c) Volume fraction of black pixels $\phi_1=0.5$. The linear size of the systems $N=200$.

VII. CONCLUSIONS

In this paper, we have provided a general rigorous scheme to model and categorize two-phase statistically homogeneous and isotropic media. In particular, given a set of basis functions, we have shown that the medium can be modeled by a map $\wp$ composed of convex combination and product operations. The basis functions should be realizable but, if they are not, they should at least satisfy all the known necessary conditions for a realizable autocovariance function. We have gathered all the known necessary conditions and made a conjecture on a possible new condition based on simulation results. A systematic way of determining basis functions is not available yet. We proposed a set of basis functions with simple analytical forms that capture salient microstructural features of two-phase random media.

We gave a rigorous mathematical formulation of the (re)construction problem and showed that the two-point correlation function alone cannot completely specify a two-phase heterogeneous material. Moreover, we devised an efficient and isotropy-preserving (re)construction algorithm, namely, the lattice-point algorithm to generate realizations of materials based on the Yeong-Torquato technique. We also provided an example of nonrealizable yet nontrivial two-point correlation function and showed that our algorithm can be used to test realizability of hypothetical functions. An

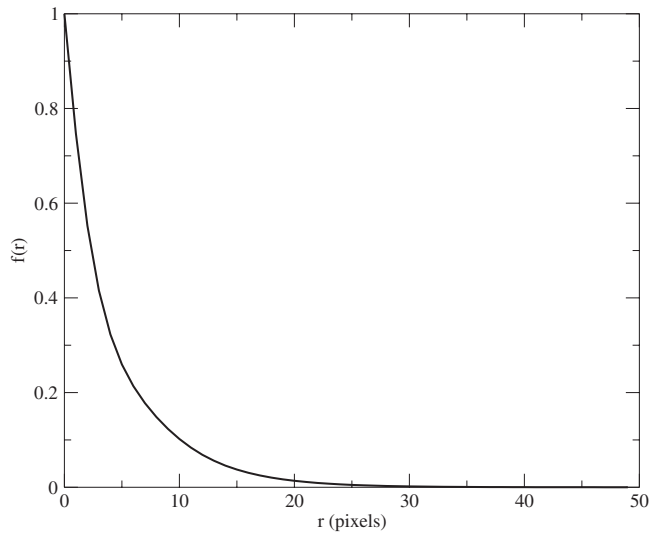

FIG. 12. Combined scaled autocovariance function $f(r)$ according to Eq. (62) with coefficients $\alpha_1=0.75$, $\alpha_2=0.25$. Here $a=20$ (pixels), $A_1=1$, $A_i=0(i>1)$, $a_1=5$ (pixels), $q_1=10$ (pixels).

example of generating hypothetical random media with combined realizable correlation functions was given as an application of our general scheme. We showed that even a simple combination of two basis functions enables one to produce media with a variety of microstructures of interest and therefore a means of categorizing microstructures.

We are investigating applications of our general scheme in order to model real materials. We are also developing more efficient (re)construction algorithms. There is a need for a theoretical and numerical analysis of the energy threshold of the algorithm, which is the aforementioned "acceptable tolerance." This quantity provides an indication of the extent to which the algorithms have reproduced the target structure and it is directly related to the nonuniqueness issue of reconstructions [2–5]. Moreover, additional realizable basis functions are needed to construct a complete basis set. Such work will be reported in our future publications.

ACKNOWLEDGMENTS

Acknowledgment is made to the Donors of the American Chemical Society Petroleum Research Fund for support of this research.

APPENDIX

The complementary error function $f_{\mathrm{CE}}(r)$ is defined as

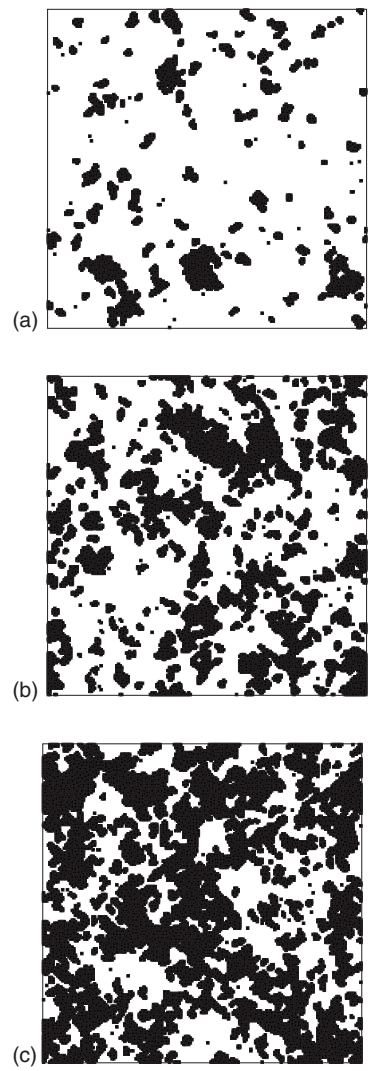

FIG. 13. Constructed media with scaled autocovariance function shown in Fig. 12. (a) Volume fraction of black pixels $\phi_1=0.1$. (b) Volume fraction of black pixels $\phi_1=0.3$. (c) Volume fraction of black pixels $\phi_1=0.5$. The linear size of the systems $N=200$.

$$f_{\mathrm{CE}}(r) = \frac{2}{\sqrt{\pi}}\int_{r/a}^{\infty} e^{-t^2}dt, \tag{A1}$$

where $r\geqslant 0$ and a is the effective correlation length. It is easy to check that $f_{\mathrm{CE}}(r)$ satisfies the known necessary conditions collected in this paper except for Eq. (20), which can only be checked for a finite number of cases. Realizations of the random medium associated with $f_{\mathrm{CE}}(r)$ have been constructed using the Yeong-Torquato technique with very high numerical precision. Thus, we believe $f_{\mathrm{CE}}(r)$ to be a valid candidate for a realizable scaled autocovariance function.

[1] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer-Verlag, New York, 2002).
[2] C. L. Y. Yeong and S. Torquato, Phys. Rev. E **57**, 495 (1998).
[3] C. L. Y. Yeong and S. Torquato, Phys. Rev. E **58**, 224 (1998).
[4] D. Cule and S. Torquato, J. Appl. Phys. **86**, 3428 (1999).
[5] N. Sheehan and S. Torquato, J. Appl. Phys. **89**, 53 (2001).
[6] H. Kumar, C. L. Briant, and W. A. Curtin, Mech. Mater. **38**, 818 (2006).
[7] G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge University Press, New York, 1982).
[8] S. O. Rice, Bell Syst. Tech. J. **23**, 282 (1944).
[9] R. J. Adler, *The Geometry of Random Fields* (Wiley, New York, 1981).
[10] J. A. Quiblier, J. Colloid Interface Sci. **98**, 84 (1984).
[11] N. F. Berk, Phys. Rev. Lett. **58**, 2718 (1987).
[12] M. Teubner, Europhys. Lett. **14**, 403 (1991).
[13] A. P. Roberts and M. A. Knackstedt, Phys. Rev. E **54**, 2313 (1996).
[14] N. G. van Kampen, J. Stat. Phys. **24**, 175 (1981).
[15] M. D. Rintoul and S. Torquato, J. Colloid Interface Sci. **186**, 467 (1997);J. Crawford, S. Torquato, and F. H. Stillinger, J. Chem. Phys. **119**, 7065 (2003).
[16] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
[17] P. Debye and A. M. Bueche, J. Appl. Phys. **20**, 518 (1949).
[18] S. Torquato, J. Chem. Phys. **111**, 8832 (1999).
[19] S. Torquato, Ind. Eng. Chem. Res. **45**, 6923 (2006).
[20] L. A. Shepp (unpublished).
[21] S. Prager, Chem. Eng. Sci. **18**, 227 (1963).
[22] D. R. S. Talbot and J. R. Willis, IMA J. Appl. Math. **39**, 215 (1987).
[23] S. Prager, Phys. Fluids **4**, 1477 (1961).
[24] J. G. Berryman and G. W. Milton, J. Chem. Phys. **83**, 754 (1985).
[25] S. Torquato and D. C. Pham, Phys. Rev. Lett. **92**, 255505 (2004).
[26] D. C. Pham and S. Torquato, J. Appl. Phys. **97**, 013535 (2005).
[27] P. Ponte Castaneda and P. Suquet, Adv. Appl. Mech. **34**, 171 (1998).
[28] L. V. Gibiansky and S. Torquato, J. Appl. Phys. **84**, 5969 (1998).
[29] Y. Jiao, F. H. Stillinger, and S. Torquato (in preparation).
[30] S. Torquato, J. D. Beasley, and Y. C. Chiew, J. Chem. Phys. **88**, 6540 (1988).
[31] B. Lu and S. Torquato, Phys. Rev. A **45**, 922 (1992).
[32] B. McMillan, J. Soc. Ind. Appl. Math. **3**, 119 (1955).
[33] G. Matheron, Cahi. Geostatistique **107**, 107 (1993).
[34] K. Markov (unpublished).
[35] J. Quintanilla (in preparation).
[36] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups* (Springer, Berlin, 1987).
[37] P. Debye, J. H. R. Anderson, and H. Brumberger, J. Appl. Phys. **28**, 679 (1957).
[38] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Application* (Wiley, New York, 1987).
[39] S. Torquato and G. Stell, J. Chem. Phys. **79**, 1505 (1983).
[40] B. Lu and S. Torquato, Phys. Rev. B **42**, 4453 (1990).
[41] C. Deutsch, J. L. Sarmiento, D. M. Sigman, N. Gruber, and J. P. Dunne, Nature (London) **445**, 163 (2007).
[42] J. B. S. Tokarzewski, J. Blawzdziewicz, and I. Andrianov, Appl. Phys. A **59**, 601 (1994).
[43] V. Gibiansky and S. Torquato, Int. J. Eng. Sci. **34**, 739 (1996).