Progress in Materials Science 95 (2018) 1–41

Contents lists available at ScienceDirect

Progress in Materials Science

journal homepage: www.elsevier.com/locate/pmatsci

Computational microstructure characterization and reconstruction: Review of the state-of-the-art techniques

Ramin Bostanabad [a], Yichi Zhang [a], Xiaolin Li [a], Tucker Kearney [a], L. Catherine Brinson [a,b,c], Daniel W. Apley [d], Wing Kam Liu [a,c], Wei Chen [a,b,*]

[a] Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA
[b] Department of Materials Science & Engineering, Northwestern University, Evanston, IL 60208, USA
[c] Theoretical & Applied Mechanics, Northwestern University, Evanston, IL 60208, USA
[d] Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA

ARTICLE INFO

Article history:
Received 9 July 2017
Accepted 18 January 2018
Available online 9 February 2018

Keywords:
Microstructure
Characterization and reconstruction
Processing-structure-property links
Computational materials design
Spectral methods
Correlation functions
Texture synthesis
Supervised and unsupervised learning
Statistical equivalency

ABSTRACT

Building sensible processing-structure-property (PSP) links to gain fundamental insights and understanding of materials behavior has been the focus of many works in computational materials science. Microstructure characterization and reconstruction (MCR), coupled with machine learning techniques and materials modeling and simulation, is an important component of discovering PSP relations and inverse material design in the era of high-throughput computational materials science. In this article, we provide a comprehensive review of representative approaches for MCR and elaborate on their algorithmic details, computational costs, and how they fit into the PSP mapping problems. Multiple categories of MCR methods relying on statistical functions (such as n-point correlation functions), physical descriptors, spectral density function, texture synthesis, and supervised/unsupervised learning are reviewed. As no MCR method is applicable to the analysis and (inverse) design of all material systems, our goal is to provide the scientific community with a close examination of the state-of-the-art techniques for MCR, as well as useful guidance on which MCR method to choose and how to systematically apply it to a problem at hand. We illustrate applications of MCR on materials modeling and building structure-property relations via two examples: One on learning the materials law of a class of composite microstructures, and the second on relating the permittivity and dielectric loss to a structural parameter in nanodielectrics.

Contents

* Corresponding author at: Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA.
E-mail address: weichen@northwestern.edu (W. Chen).

https://doi.org/10.1016/j.pmatsci.2018.01.005
0079-6425/

1. Introduction

The heart of computational materials science lies in providing fundamental insights and understanding of materials behavior and properties across different scales, which further enables cost-effective design of materials with targeted properties. The significance of this task is highlighted by the Materials Genome Initiative [1–4] and the emergence of computational tools and frameworks such as materials by design [5,6], microstructure sensitive design [7], and integrated computational materials engineering (ICME) [8]. As a material's morphology heavily affects its properties [9,10], the central theme of these frameworks is on *inverse materials design* by elucidating the link between processing, structure, and properties (aka PSP links) [5,11]. The fact that such inverse mappings are not unique (see Fig. 1a), introduces both opportunities and challenges for developing PSP mapping techniques.

Evidently, experimental (aka Edisonian or cause-effect [6,12]) analysis, design, and discovery of materials is costly and time-consuming. This noticeably limits the technological and industrial sectors as they are heavily connected with material systems and their growth and long-lasting success hinges on efficient development and deployment of superior materials [1–3,6]. Such a strong connection between materials and industry's progress, demands that new material systems consistently demonstrate superior properties because significant difficulties and expenses are incurred when the materials of an established technology must be altered to, e.g., address environmental concerns or meet industry or government specifications. Materials design is additionally hampered by the fact that it must be, to some extent, compatible with existing processing technologies.

The so-called high-throughput computational materials science [13] provides an appealing alternative to the expensive experimental approach for building PSP links and materials design (see Fig. 1b). Here, the central concept is to first create a massive database that stores microstructural characteristics and properties of materials and then intelligently interrogate this dataset to gain insights and use those new insights to discover materials with desired properties. To this end, many existing works apply predictive models including neural networks [14] or support vector machines [15] to material databases for building PSP links and, subsequently, estimate the material properties where no data (either experimental or computational) are available [16–20]. In the high-throughput fueled frameworks, it is highly desirable to quantitatively **characterize** the microstructure to (*i*) increase the predictive power of such models, and (*ii*) achieve a deeper understanding on how the microstructure, on the one hand, is formed by composition and processing history [21,22] and, on the other hand, affects the material properties. Realizing that in most material systems the corresponding microstructure embodies some degree of randomness (e.g., grain size distribution in crystals or particle dispersion in fibrous composites), it is evident that

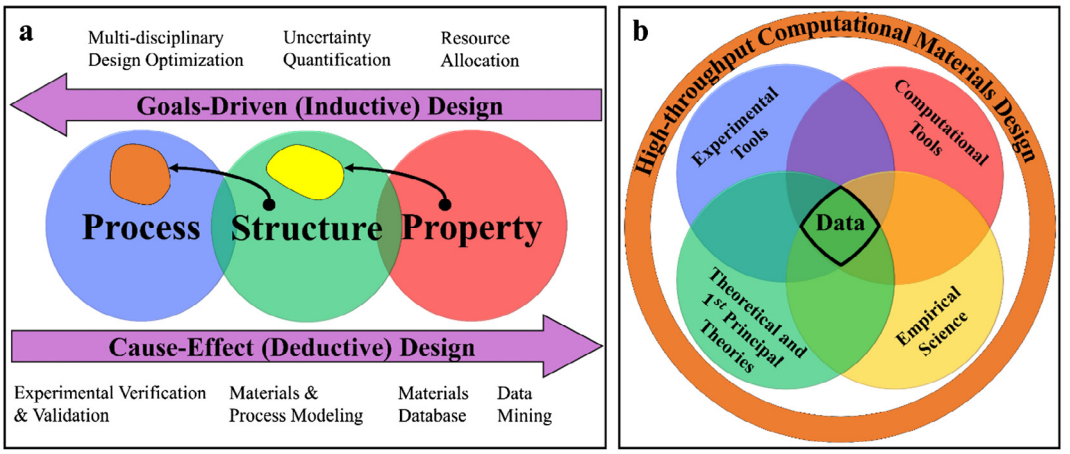

Fig. 1. (a) Forward and inverse PSP links in materials science: Unlike the direct cause-effect links, the links in the goals-driven materials design are not one-to-one. Such a feature in the indirect approach, while providing design flexibility, requires identification of the set of processes (structures) resulting in a particular structure (property). For these reasons, the methods and computational tools used in the two approaches are different. In this illustrative figure, a point in the process, structure, or property disk corresponds to a particular, respectively, processing procedure, microstructure morphology, and set of properties. (b) Data-driven materials design: The massive amount of data on various material systems, built through application of experimental and computational tools in theoretical and empirical sciences, has enabled high-throughput materials design.

the characterization must, in general, be *statistical*. The characterized information, in turn, can be used to computationally **reconstruct** new microstructures to augment the available imaging data or even guide future imaging experiments (e.g., to determine the required imaging scale and resolution).

Computational microstructure characterization and reconstruction (MCR) consists of statistical methods to (*i*) quantitatively represent the microstructure (and its possible inherent randomness), and (*ii*) build an ensemble of statistically equivalent microstructure [23] samples. MCR allows one to systematically go beyond the limits where the empirical data are available and build forward and inverse PSP links to find the processes (structures) that result in a particular structure (property), see Fig. 1a.

In this article, we report on recent advances in MCR of random heterogeneous materials, an essential component in the computational and data-driven paradigms of materials science. The focus is specifically on methods developed to investigate the PSP links. In this context, while **characterization** refers to the statistical representation of a material's morphology via a (finite set of) characteristic function(s) and/or feature(s), **reconstruction** is the process of generating a (set of) microstructure(s) whose morphology embodies a (set of) prescribed characteristic(s). Characterization and reconstruction are complementary and, together, provide the means to delve deeply into the detailed morphological characteristics at the various length scales of a material's hierarchical structure through imaging data. According to [24], such data "offers considerably more than merely the illustration of a material system; in fact, it contains quantitative structural and functional information. This information is spatially distributed and often has a complex multidimensional nature". MCR is of common use in:

- Predictive materials modeling [23,25–28], i.e., finding the materials law that describes the homogenized response of a heterogeneous material (see Section 5.1 for an example).
- Design of material systems with targeted properties [29–33] (see Section 5.2 for an example).
- Transformation and preparation of experimental data for computer simulations [34,35]. Examples include building a representative microstructure image from a set of small micrographs [36], reconstructing 3*D* structures from 2*D* images [37,38], and recovering the original microstructure image using the results of X-ray diffraction experiments [39].
- Determination of feasibility and/or sufficiency of a given set of characteristic functions [40–42]. Quantifying the high-dimensional morphology of a microstructure with parametric functions is particularly useful in that it, e.g., enables building PSP links or helps to identify materials with the same underlying randomness (see Section 3.1.1 and Fig. 3). However, a given parametric function might not sufficiently characterize a microstructure or even necessarily correspond to a feasible one. In such cases, a reconstruction algorithm can be used to determine the feasibility and/or sufficiency (see Section 3.1 for more details).
- Microstructure-induced uncertainty quantification and propagation [23,43–46]: Microstructure randomness manifests itself in probabilistic material properties and performance; necessitating the study of, e.g., mean and variability behavior and, more importantly, rare events [47–49]. To enable such computational studies, one must reconstruct an ensemble of microstructure samples (that capture the randomness), estimate their properties (e.g., via the finite element method, FEM), and then proceed to build PSP links and conduct statistical analyses (such as sensitivity analysis [50], model calibration and bias correction [45], and uncertainty quantification and propagation [33,43,46,51,52]).

From a recent historical standpoint, statistical microstructure characterization can be traced back to the early works of Debye and co-workers [53,54] and then Corson [55–58], where the scattering power for electromagnetic radiation (resulting in the *two-point correlation function*) was used to characterize the inhomogeneities in porous media. Later, reconstructing such media by level-cutting *random fields* (to realize the microstructures corresponding to the characterized information) was proposed by Joshi [59] and then extended by Quiblier [60]. To address the shortcomings associated with the use of random fields in reconstruction (detailed in Section 3.3), Hazlett [61] and then Rintoul and Torquato [62] used an optimization approach based on simulated annealing. To enable the use of statistical functions other than the two-point correlation in characterization, this optimization-based reconstruction scheme was subsequently generalized by Yeong and Torquato [39,63] to the stochastic reconstruction (aka YT) method. Because the optimization process in the YT method is expensive and the reconstruction quality heavily depends on, among other things, the choice of the statistical functions used in the characterization, considerable research has been conducted to improve the YT method (detailed in Section 3.1.2.2). Due to these and other shortcoming of the YT method (see Section 3.1.2), other approaches have been developed for efficient MCR in different material systems. These approaches commonly use *physical descriptors* (especially for the analysis and design of crystalline or particulate microstructures), *supervised/unsupervised learning techniques* (which are computationally efficient and applicable to a wide range of materials), *spectral methods* (for modeling the microstructure as a realization of a random field), and *texture synthesis* (for synthesizing a microstructure based on an exemplar). In Section 3, we elaborate on the details of these approaches.

As described in Section 3, several methods have been developed for MCR and the optimal choice for a given problem depends on the material system, properties of interest, the length scale, as well as a balance between accuracy and cost. As none of these methods are applicable to all problems (see above and Section 5 for some examples), a researcher should make an informed decision for his/her MCR task. To aid in this end, we review the prevailing methods, comment on their advantages and disadvantages, and provide some examples and comparisons. In Section 2, we define the technical terms used in the literature and throughout this manuscript. In Section 3 we describe various MCR methods and elaborate on their algorithmic details, computational costs, and applicability. We compare the methods in Section 4 and provide two material examples in Section 5. Conclusions are provided in Section 6.

2. Definition of technical terms

Several terms are frequently used in the computational MCR literature. Before diving into the technical descriptions in Section 3, in this section we define these terms and provide the relevant mathematical equations.

Characterization: Statistical quantification of the morphology of a microstructure. The quantification needs to be statistical to account for the (possible) inherent morphological randomness. Complete characterization is often very difficult, if not impossible, and might not actually be necessary (see Section 5.1 for an example). Correlation functions (e.g., two-point correlation function) and physical descriptors (e.g., volume fraction and average particle or grain size) are commonly used in characterization.

Ensemble: A collection of samples that are statistically equivalent. The ensemble members might have different morphological (fine) details but are identical from a coarse-scale point of view and have similar effective properties [64]. Consider the sample space S of all statistically equivalent microstructures where the individual members and their probability density functions are represented by, respectively, α and $p(\alpha)$. The ensemble average of the function $q(\boldsymbol{x},\alpha)$ at $\boldsymbol{x}$ over this sample space equals [65]:

$$\bar{q}(\boldsymbol{x}) = \int_S q(\boldsymbol{x},\alpha)p(\alpha)d\alpha, \tag{2-1}$$

q is, basically, a property or quantity of interest (such as the effective Young's modulus) that depends on the microstructure as well as some other parameters collectively denoted by $\boldsymbol{x}$. An ensemble of statistically equivalent samples corresponds to, for instance, a collection of samples taken from a process line operating under stable processing conditions or a number of small samples taken from a large homogeneous one [64].

Ergodic hypothesis: A hypothesis implying that monitoring a stochastic system (e.g., an evolving microstructure) over a long period of time would provide the same statistical measures as taking many independent realizations of that system over time [7]. More formally, it states that all the microstates available to an ensemble of samples are available to every sample of the ensemble [64]. This hypothesis allows investigating one arbitrarily chosen, *sufficiently large* ensemble member (an RVE, see below) to calculate the average of a function over the ensemble space by equating it to the volumetric average [66]. This averaged function would be independent of α and is given by:

$$\bar{q}(\boldsymbol{x}) = \langle q(\boldsymbol{x},\alpha)\rangle \sim \lim_{V\to\infty}\frac{1}{|V|}\int_V q(\boldsymbol{x}+\boldsymbol{y},\alpha)d\boldsymbol{y}, \tag{2-2}$$

where $\langle\cdot\rangle$ denotes the expectation operator. One way to fulfill this condition is to presume that the coarse-scale structure can be modeled by a self-repeating unit cell (Ω, see Fig. 2) with periodic boundary conditions (PBC) [66]:

$$\lim_{V\to\infty}\frac{1}{|V|}\int_V q(\boldsymbol{x}+\boldsymbol{y},\alpha)d\boldsymbol{y} \sim \frac{1}{|\Omega|}\int_\Omega q(\boldsymbol{x}+\boldsymbol{y},\alpha)d\boldsymbol{y}, \tag{2-3}$$

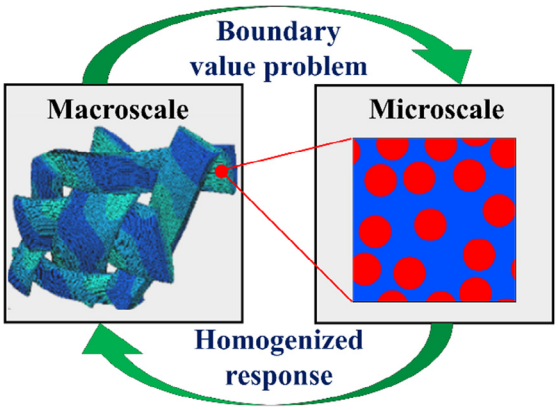

Fig. 2. Multiscale finite element simulation of a two-scale structure: Each material (aka integration or Gauss) point of the macrostructure (left) is a realization of a microstructure with fine details (right). Oftentimes the same microstructure is assigned to all the material points but in more realistic cases, the microstructures might have to be different across the macroscale domain.

Often only a single realization of the process can be observed (i.e., a single microstructure sample is available), and so the ergodic hypothesis is assumed to hold rather than tested.

Microstructure[1]*:* The structure of a material whose scale is (*i*) large enough to contain sufficient local morphological details, but (*ii*) much smaller than the characteristic length of the sample [9]. Microstructure is a general term that describes the structural features and topological arrangements at a particular length-scale (e.g., nanoscale or macroscale). For instance, in multiscale finite element analysis (FEA) of a material system, each material (aka integration or Gauss) point at the coarse scale (e.g., macroscale) is a realization of a structure with finer resolution (e.g., microstructure) that passes the constitutive behavior to the coarse scale (see Fig. 2). In continuum mechanics literature, a microstructure is called, depending on its size, either an RVE or an SVE (see below) [23,67,68].

Random field: An infinite collection of random variables which are indexed in time or space and share a common probability model [69]. Modeling a material system as a random field (aka stochastic or random process), would constitute that the microstructure images corresponding to that material system can be thought of as independent realizations from that random field.

Random heterogeneous materials*:* Materials that are comprised of different constituents (e.g., a composite) or a single constituent but in different arrangements (e.g., a polycrystal with multiple crystallographic orientations) [9,62]. The local morphological details of such materials often indicate some degree of randomness when it can only be characterized statistically (hence the word random) [70]. Composites, concrete, block co-polymers, polycrystals, and sandstone are examples of synthetic and natural random heterogeneous materials.

It is noted that, to include ordered microstructures (such as highly textured metallic polycrystals or ordered block co-polymers) in this definition as well, the statistical characterization must identify their randomness degree as zero or negligible.

Reconstruction*:* Construction of a (set of) statistically equivalent microstructure image(s) given some statistical characterization. Such statistical characterizations are either derived from an actual material sample or defined by the user. In the latter case, the user must ensure that the statistics correspond to a feasible structure.

Representative volume element (RVE)*:* A finite region of a material sample which is (*i*) large enough to be representative of the entire sample's characteristics, and (*ii*) has (nearly) the same effective properties as the entire sample [71,72]. With this definition, the size of the RVE clearly depends on the material system as well as the effective property under consideration [65,73–82].

Statistical equivalency*:* The state where statistical measures of two (or more) microstructures match. This condition, when met, would imply having similar effective properties (the opposite of this statement generally does not hold).

Statistical homogeneity*:* The condition when the calculated statistics for a microstructure are independent of the coordinate system translation.

Statistical isotropy: The condition when the calculated statistics for a microstructure are independent of the coordinate system rotation.

Statistical volume element (SVE): A finite region of a material that is smaller than an RVE. SVEs should be large enough so that an ensemble of them contains at least as much information as a single RVE. Small SVEs necessitate the use of a large ensemble for accurate estimation of effective properties. When analyzing SVEs instead of RVEs, the ergodic hypothesis is invoked.

[1] As our focus is on computational studies, the term microstructure actually refers to the image of a microstructure.

Texture: A realization of a stationary random field (in computer graphics community). In materials science, the crystallographic orientation map in a polycrystalline structure is also referred to as a texture. Herein, we use the former description.

3. MCR methods

To achieve statistical equivalency, the reconstruction algorithm should be *compatible* with the characterization method. This way, the reconstructed microstructures will possess the statistical features obtained from an original microstructure sample (e.g., by image analysis) or defined by the user (e.g., for design purposes, see also Section 5.1). Following this compatibility requirement, we have categorized several MCR methods based on the *characterization* scheme and review them, along with the appropriate reconstruction methods, below.

The depth of details on each method reflects the proportion of papers covering them. For example, while there are a large number of papers on the use of statistical functions for MCR, there are only a few recent works investigating the applications of supervised/unsupervised learning in MCR.

Hereafter, we primarily focus on two-phase microstructures and always assign the subscript/superscript 1 to the primary phase, i.e., the phase which must be characterized (e.g., the voids in a porous material or the fibers in a fibrous composite). The discussions can be extended to multi-phase materials in a straightforward manner. We also note that experimental imaging techniques such as scanning electron microscopy (SEM) and transmission electron microscopy (TEM) often result in grayscale images but herein it is assumed that the images are appropriately denoised (e.g., with a Gaussian filter) and thresholded (e.g., via either matching the volume fraction [83] or Otsu's method [84]). Fig. 3 illustrates how MCR is generally used in computational materials science: Once the available microstructure sample is prepared, it is characterized to represent its morphology (a very high dimensional space) quantitatively. This characterized information is then fed into the reconstruction algorithm to build an ensemble of samples with as many statistically equivalent microstructures as required. Afterwards, the properties of the ensemble members are predicted via computer simulations to quantify their variability. If a finite set of parameters are identified during characterization and property evaluation steps (denoted by, respectively, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in Fig. 3), by changing the characterization parameters (i.e. $\boldsymbol{\alpha}$) in a feasible space and repeating the above process, a training dataset can be created to relate the microstructure (through $\boldsymbol{\alpha}$) to the properties, $\boldsymbol{\beta}$, via machine learning.

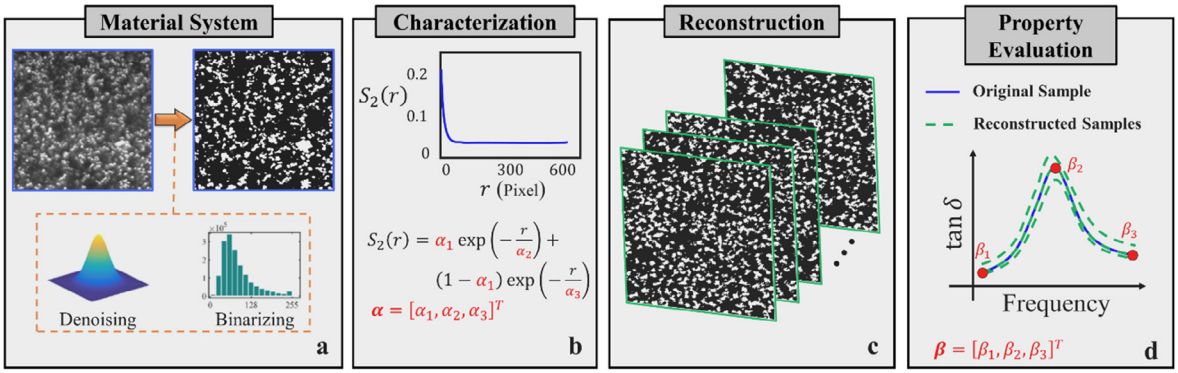

Fig. 3. MCR in computational materials science: (a) The material system of interest is determined by the given sample. If the original image is grayscale, it needs to be processed (e.g., denoised and binarized) first. (b) The binarized image is characterized using one of the methods described in Sections 3.1–3.4 (here, the two-point correlation function is used). (c) A (set of) statistically equivalent microstructure(s) is reconstructed. The choice of the method depends on (and must be compatible with) the characterization scheme. (d) The properties of the reconstructed sample(s) are calculated via, e.g., FEA. Here, the changes in the *tanδ* curves of different microstructures are due to dispersion status. If the characterized features as well as the properties are parametrized (here, with $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$), structure-property links can be established by reconstructing an ensemble of samples, evaluating their properties, and subsequently using machine learning techniques.

3.1. Statistical functions

In this section, we start by describing the functions commonly used for statistical characterization of random heterogeneous materials, and then introduce the well-known Yeong and Torquato (aka YT) method for microstructure reconstruction.

3.1.1. Characterization

A microstructure image may be quantitatively characterized via some statistical functions which essentially capture the degree of spatial correlation among different locations in a probabilistic sense. As shown in [85–90] and discussed below, most microstructures cannot be sufficiently and uniquely characterized with only a single statistical function as these functions quantify spatial correlations of different natures. Some of the most widely used functions in the literature are reviewed below. The interested reader may consult [7,9,70,91–93] for other functions and more detailed discussions.

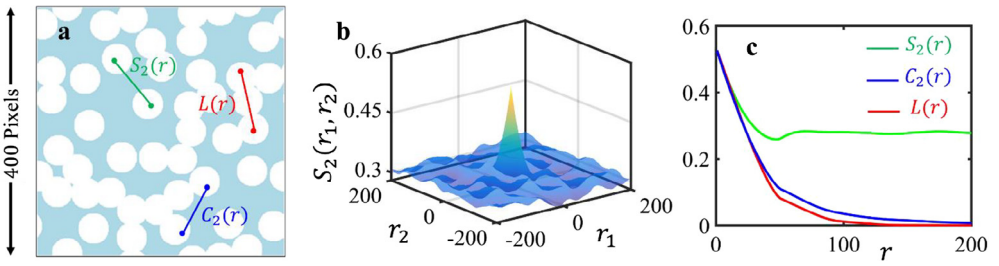

Fig. 4. Statistical functions: (a) A binary microstructure consisting of equal-size overlapping disks with radius of 25 pixels (the maximum allowable overlap is 12 pixels). (b) The two-point correlation function of microstructure in (a), see Eq. (3-2). (c) Radially averaged two-point, two-point cluster, and lineal-path functions of the microstructure in (a).

3.1.1.1. Two-point correlation function. Let $\boldsymbol{X}$ denote the microstructure image on a square lattice. The collection of pixels (voxels in 3*D*) of $\boldsymbol{X}$ are indexed and their brightness represents different material phases. For instance, a two-phase material (see Fig. 4a) may be represented with a binary (here, blue and white) image. An image might also be represented on a triangular lattice, especially to better characterize highly anisotropic structures with long range correlations [37]. As this is rarely the case, herein we only focus on microstructure images on a square lattice:

$$\boldsymbol{X}_{uv} = \begin{cases} 1 & \text{if } uv \in \text{phase } 1 \\ 0 & \text{otherwise} \end{cases}, \tag{3-1}$$

where uv is the pixel index and determines its location within the image. Denoting this location by the vector $\boldsymbol{r}$, the two-point (auto)correlation function [55,56,94] for phase i is defined as:

$$S_2^i(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle \boldsymbol{X}_{\boldsymbol{r}_1}, \boldsymbol{X}_{\boldsymbol{r}_2} \rangle, \tag{3-2}$$

where $\langle \cdot \rangle$ denotes the expectation operator. In simple words, this function provides information on pair distributions regarding phase i within the microstructure image. This definition can be generalized to higher order functions such as three and four-point correlation functions (denoted by $S_3^i(\boldsymbol{r}_1, \boldsymbol{r}_2)$ and $S_4^i(\boldsymbol{r}_1, \boldsymbol{r}_2)$, respectively).

If $\boldsymbol{X}$ is statistically stationary and isotropic, S_2^i will only depend on the distance between the two points: $S_2^i(\boldsymbol{r}_1, \boldsymbol{r}_2) = S_2^i(\Delta \boldsymbol{r}_{12}) = S_2^i(|\Delta \boldsymbol{r}_{12}|) = S_2^i(r)$. Hence, for a stationary and isotropic material, S_2^i has a simplified formulation and can be efficiently calculated (e.g., via FFT [95–97], Monte Carlo [90], orthogonal sampling [39,63], or lattice-point algorithm [42]). $S_2^i(r)$ is perhaps the most widely used and well-known statistical function and can be thought of as the probability of tossing a line of length r on a microstructure image and having both its ends land on phase i. $S_2(r)$ belongs to the family of n-point correlation functions which have been widely used in cosmology to investigate dark matter halo [98,99] as well as materials science to predict the bounds on effective (i.e., macroscopic) properties such as electrical permittivity [56,58,100] and bulk modulus [57,101]. Another attractive feature of $S_2(r)$ is that many experimental characterization techniques including small angle X-ray scattering [53], time dynamics of energy transfer [102], and distant dipolar field nuclear magnetic resonance [103] can provide structural information in the form of $S_2(r)$. For a stationary material consisting of two phases with volume fractions (V_f)[2] of ϕ_1 and ϕ_2, $S_2^i(r)$ of a digitized microstructure (i.e., an image) must have the following properties [9,40,41,104,105]:

- Upper and lower bounds: $0 \leqslant S_2^i(r) \leqslant \phi_i$
- Non-negative Fourier transform
- Equality to ϕ_i and proportionality to specific surface (s_s^i) at $r = 0$: $S_2^i(0) = \phi_i$ and $\left.\frac{dS_2^i(r)}{dr}\right|_{r=0} = -\frac{s_s^i}{2D}$, where $D \in \{1, 2, 3\}$ is the dimensionality of the image
- Asymptotic value (for a large enough image of size l and in the absence of long range correlations): $\lim_{|r| \to l} S_2^i(r) \to \phi_i^2$
- Satisfying the triangular inequality: $S_2^i(r) \geqslant S_2^i(z) + S_2^i(t) - \phi_i$, $r = t - z$

As mentioned above, $S_2(r)$ quantifies the microstructure in a probabilistic sense and so cannot be readily used to build sensible structure-property maps. To provide a physical understanding of the quantified information and enable the identification of microstructures that have two-point correlation functions with similar features (e.g., the existence of fluctuations), parametric basis functions have been used. Sample functions include monotonically decreasing functions (applied especially to porous materials having voids of random shape and size) [42,54,55], damped oscillating functions

[2] Area fraction in 2*D*. For simplicity, the term volume fraction is also used in 2*D* but the meaning is clear from the context.

[30,39,42,106,107], and functions with known structures corresponding to certain material systems, such as randomly packed overlapping spheres [9,42,99,106].

Although it is commonly believed that there are many different microstructures compatible with a given $S_2(r)$ [39,108], Rozman et al. [40] argue that periodic microstructures with smooth boundaries are completely specified by their two-point correlation functions. In their language, "smooth" refers to the condition where the length scales of $S_2(r)$ have both an upper bound and a lower bound determined by, respectively, the periodicity and the shortest wavelength present in the structure. Unfortunately, most materials do not satisfy these conditions and hence $S_2(r)$ cannot uniquely characterize them (e.g., it does not include information about percolating clusters within the system [85]).

Higher order correlation functions (e.g., $S_3(r)$ and $S_4(r)$) can also be used for characterization but their computation is more demanding and it is illustrated in [85,88] that more efficient functions exist (see below).

3.1.1.2. Lineal-path function. As discussed above, $S_2(r)$ cannot uniquely characterize the majority of material systems and therefor additional statistical functions are required for effective characterization. The lineal-path function [86], $L(r)$, is one such function that quantifies the amount of *clusteredness* along straight lines within a microstructure (see Fig. 4). $L^i(r)$ is a two-point quantity and from a probabilistic point of view, gives the probability of throwing a line on the microstructure image and having the entire line land on phase *i*. It can be calculated either with a Monte Carlo approach [90] or by building the histogram of chords [89].

Similar to $S_2(r)$, analytical expressions for $L(r)$ have also been derived for certain material systems such as 1*D* digitized medium of periodic/random rods [39], overlapping equally-sized particles [109], and equilibrium hard disks [39]. Unlike $S_2(r)$, the lineal-path functions of different phases within a single microstructure are not linearly dependent [39]. In other words, they are phase-distinguishing and can all be used in characterization (e.g., for an isotropic binary microstructure $S_2^1(r) = S_2^2(r) + \phi_1^2 - \phi_2^2$ where ϕ_i is the volume fraction of phase *i*).

The lineal-path function may be related to other statistical functions as well. It is a rigorous lower bound on the two-point cluster correlation function [109], which is not available analytically for overlapping particle models for spatial dimension $D \geqslant 2$. In addition, the chord-length distribution function [9,86,109] can be evaluated by taking the second derivative of $L(r)$.

$L(r)$ generally underestimates the clusteredness since the connectivity is only measured along straight lines (see Fig. 4c where $L(r)$ is below $C_2(r)$). In other words, this measure would render two connected sites within a microstructure disconnected if the connectivity is not along a straight line. In addition, it is generally calculated along certain directions (e.g., horizontal, vertical, and diagonal) within the microstructure rather than along all the possible directions (see the recent work of Turner et al. [110] for a computationally efficient method known as scan line method [111,112] to simultaneously calculate $L(r)$ along various directions). Amid these points, multiple studies [36,39,63,89] have shown that $L(r)$ is useful for the characterization of clusters in a microstructure.

3.1.1.3. Two-point cluster correlation function. The two-point cluster correlation function [9,85,88], $C_2(r)$, embodies important topological connectivity statistics and is perhaps the most informative "two-point" statistical function. $C_2^i(r)$ quantifies the probability of finding two sites within phase *i* that are at distance *r* and in the same cluster. Although it is intrinsically a 3*D* descriptor, $C_2(r)$ has been widely used to characterize 2*D* images as well. Given the digitized image of an arbitrary microstructure, $C_2(r)$ can be calculated with either Monte Carlo [113] or building the histogram (i.e., binning) of the pair-distances for all the clusters (normalized by the total number of pair-distances).

Decomposition of the two-point correlation function discussed above gives some insight as to why $C_2(r)$ is more informative than other two-point (and possibly higher order) statistical functions. Noting that in calculating the probabilities of $S_2(r)$, the points which are in different clusters are not distinguished, the decomposition for a binary image reads [88]:

$$S_2(r) = C_2(r) + D_2(r), \tag{3-3}$$

where $D_2(r)$ measures the probability of tossing a line on the microstructure and having its ends land on different phases. As Eq. (3-3) shows, unlike $S_2(r)$, $C_2(r)$ is sensitive to clustering and percolation and in fact its volume integral diverges at the percolation threshold [9,85]. As schematically illustrated in Fig. 5, $C_2(r)$ is more informative than $S_3(r)$ for microstructures characterization.

3.1.2. Reconstruction

Having chosen statistical functions for characterization, reconstructing a statistically equivalent digitized microstructure can be cast as an optimization problem by adjusting an initial digitized image $\boldsymbol{Y}$ (with the same V_f for each phase as the reference image) to minimize some cost (aka energy) function, hereafter denoted by *E*. The adjustment is done by swapping the pixels (voxel in 3*D*) of $\boldsymbol{Y}$ to minimize *E*, which measures the differences between the statistical function(s) of the original image ($\boldsymbol{X}$) and those of $\boldsymbol{Y}$. As noted in Section 3.1.1, one specific function cannot solely characterize most microstructures so usually multiple functions are incorporated into *E* as:

$$E = \sum_{j=1}^{m}\sum_{r=0}^{l}\alpha_j|\hat{f}_j(r) - f_j(r)|^2 \tag{3-4}$$

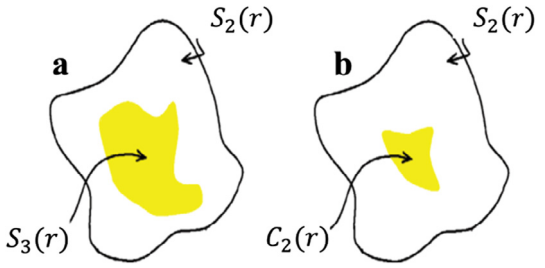

Fig. 5. Effectiveness of $C_2(r)$ in microstructure characterization: (a) The region enclosed by the solid line schematically represents the set of all microstructures compatible with a particular $S_2(r)$. The microstructures in the shaded yellow region within this contour illustrate those with a particular $S_3(r)$. (b) Unlike (a), here the shaded yellow region corresponds to microstructures with a particular $C_2(r)$. The shaded yellow region in (b) is smaller than that in (a). The image is reproduced with kind permission from [88].

where m is the total number of the functions considered (e.g., $m = 2$ if $f_1(r) = S_2(r)$ and $f_2 = L(r)$), α_j's are the weights that quantify the importance of each function, l is the maximum spatial distance at which the functions are compared at, and $\hat{f}_j(r)$ are the target functions (e.g., those of $\boldsymbol{X}$). We will elaborate on the choice of the functions in Eq. (3-4) below but note that they generally depend on the available computational resources and tools as well as the material system of interest (e.g., because percolation is important in porous materials, appropriate functions must be used to characterize the clusteredness).

3.1.2.1. The Yeong and Torquato (YT) reconstruction method. The profile of the energy function in Eq. (3-4) has multiple local optima [39,63,106,114] and therefore a heuristic method such as simulated annealing (*SA*) [39,63,88,115] or genetic algorithm [116–118] is required to escape them while swapping the pixels. Hazlett [61] used *SA* to reconstruct a 3*D* image of Berea sandstone by incorporating the variogram (the counterpart of $S_2(r)$) into E. Similarly, Rintoul and Torquato [62] used *SA* and the radial distribution function (as opposed to the variogram) to reconstruct structure of dispersions. Yeong and Torquato [39,63] generalized this approach to the stochastic reconstruction (aka YT) method in the late 1990s for reconstruction of general random media. Over the past two decades, the YT method has been significantly improved and successfully applied to MCR of various materials including particulate structures [62,116,117,119,120], chalk [112], porous media such as soil [121] and sandstone [37,38,114,122,123], materials with labyrinth patterns [124], and filamentary [125] or multiconnected [126] structures. In its most basic form, the procedures of the YT method are as follows (see Fig. 6):

- An initial image ($\boldsymbol{Y}$) with the same V_f as the target image ($\boldsymbol{X}$) is randomly generated and its statistical functions ($f_j(r)$'s) are compared to those of $\boldsymbol{X}$ ($\hat{f}_j(r)$'s) to calculate E.
- Two pixels of $\boldsymbol{Y}$ with different phases are swapped, and $f_j(r)$'s are updated to obtain E'. If $E' < E$, i.e. if the discrepancy between the statistical functions of $\boldsymbol{Y}$ and $\boldsymbol{X}$ decreases, the swap is accepted. Otherwise, some adaptive probabilistic rule (which, e.g., can be controlled by the cooling schedule in simulated annealing) is used to determine the acceptance or rejection of the swap. The updated image would then serve as the initial image for the next iteration and this process is continued until $E < \varepsilon$, where ε is the tolerance (a small number).

The YT method has also been used for reconstructing 3*D* samples given one or multiple 2*D* images [37,38,116,125,127–131] where it is assumed that the statistical functions of the original 3*D* image are the same as those of the given 2*D* sample (s), see Fig. 7.

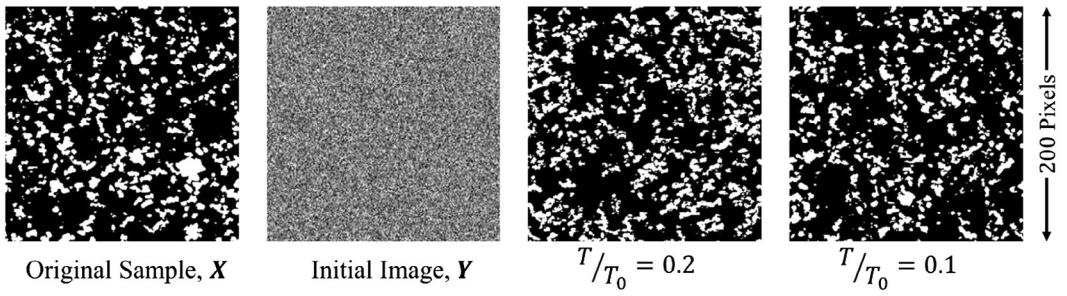

Fig. 6. Illustrative example of the YT reconstruction method: The initial binary image is adjusted (optimized) via simulated annealing. As the temperature drops (i.e., the microstructure system cools down), the energy (i.e., E) decreases. Here, only $S_2(r)$ is incorporated into E in Eq. (3-4), and T and T_0 refer to, respectively, the current and initial temperatures in the *SA* algorithm.

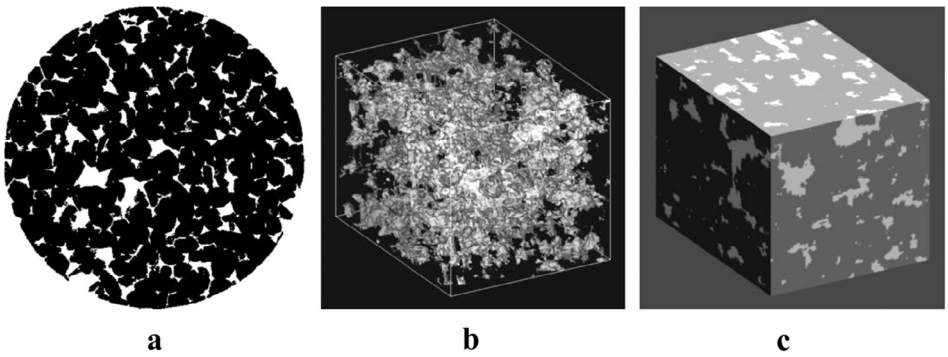

Fig. 7. 3*D* reconstruction from 2*D* images: (a) The sample 2*D* image of a Fontainebleau sandstone, (b) pore space image and (c) perspective image of the reconstructed 3*D* microstructure. The reconstruction is achieved by incorporating both $S_2(r)$ and $L(r)$ into the cost function. Images are reproduced with kind permission from [63].

3.1.2.2. Improvements of the YT method. The YT method is computationally expensive (especially if the microstructure image is anisotropic, multiphase, or has a large number of pixels) and so significant research is conducted to increase the speed as well as the accuracy (i.e., statistical equivalency) by investigating (*i*) what statistical functions to use in E as the reconstruction process progresses, (*ii*) how to efficiently *update* the $f_j(r)$'s after each swap, (*iii*) how to determine the acceptance or rejection of a swap, (*iv*) how to initialize $\boldsymbol{Y}$, and (*v*) how image analysis techniques may be incorporated into the above procedure. In what follows, we discuss these points one by one.

3.1.2.2.1. Choice of statistical functions and their updating procedure. Rozman et al. [40] argue that for certain structures (see Section 3.1.1.1), evaluating E solely via $S_2(r)$ would result in accurate enough reconstructions (see Fig. 8). This scheme has been used, with acceptable accuracy, in the past for MCR of soil pore space [121], checkerboard pattern and core-shell structure [106,132], overlapping disks [106], and sandstone [38]. However, as described in Section 3.1.1, for complex media that exhibit both short and long range spatial correlations (such as porous media) and filamentary or multiconnected structures, multiple functions must be incorporated into E to ensure statistical equivalency. Prior work [63,89,114,119,123] has demonstrated that the use of $L(r)$ (of either the primary phase or the secondary phase or both) in addition to $S_2(r)$ increases the accuracy by preserving the clustering statistics of $\boldsymbol{X}$ in $\boldsymbol{Y}$. In these works, as noted in Section 3.1.1, $L(r)$ is only calculated in certain directions rather than all possible directions due to computational costs.

Use of other functions such as surface-surface correlation and chord-length distribution [88] has also been shown to improve the accuracy (although the latter function can be directly derived from $L(r)$ and Pant et al. [89] recommend the use of $L(r)$ as its profile is smoother). The extensive study of Jiao et al. [88] and the reviews in [36,133,134] provide strong arguments that incorporation of the two-point cluster correlation function into E (besides $S_2(r)$) would best preserve the statistical equivalency between $\boldsymbol{X}$ and $\boldsymbol{Y}$ (see Fig. 9). For this reason, it is a common practice to treat $C_2(r)$ as a test measure, i.e., putting a combination of various functions such as $S_2(r)$ and $L(r)$ into E in MCR and subsequently evaluating the performance by comparing the two-point cluster correlation functions of $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Computational expenses are the primary reasons that one tries to minimize the number of statistical functions in E while ensuring the statistical equivalency: the cost of evaluating such functions at each optimization iteration (i.e., updating E after a pixel/voxel swap) is not negligible, especially because the number of iterations is often in the order of thousands, if not millions. This is particularly true for anisotropic or multi-phase microstructures. To address this issue, it has been noticed that:

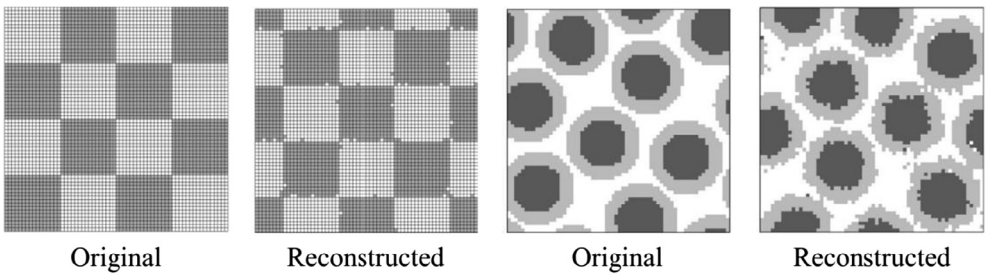

Fig. 8. Reconstruction of random media with $S_2(r)$: A two-phase checkerboard and a three-phase core-shell structures are successfully reconstructed via $S_2(r)$. Images are reproduced with kind permission from [132].

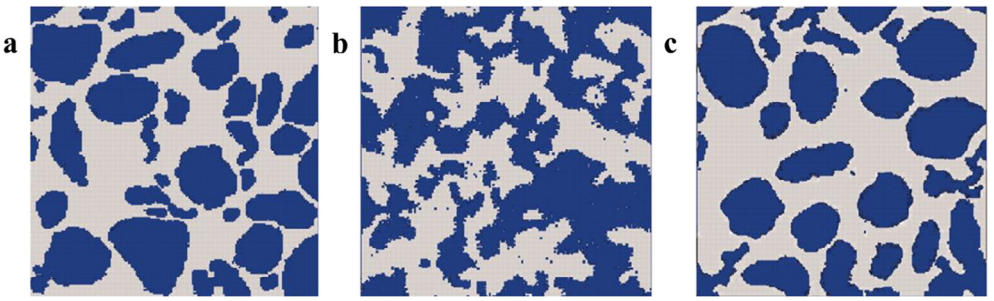

Fig. 9. Reconstruction of the two-phase microstructure of a concrete sample via correlation functions: (a) The original microstructure, (b) reconstruction via $S_2(r)$ alone, (c) reconstruction via $S_2(r)$ and $C(r)$. Images are reproduced with kind permission from [88].

- In the early stages of the optimization, most of the swaps are successful [37,38,89] and so one can start the heuristic optimization and achieve quick convergence by evaluating E only via $S_2(r)$ and incorporate other functions in the later stages of optimization.
- For certain materials (e.g., isotropic), the functions may be evaluated in only a few directions (e.g., horizontal and vertical) rather than all possible directions. Caution must be exercised because such implementations might introduce unwanted anisotropy [106] if there are considerable short range correlations in the structure.
- To update the $f_j(r)$'s of $\boldsymbol{Y}$, one does not need to calculate them anew as the swaps only slightly perturb $f_j(r)$'s (see e.g. [37,89]).
- The accumulative cost of generating random numbers for each optimization iteration of the *SA* algorithm is not negligible. This cost could be reduced by generating a large array of random numbers prior to optimization [37].
- *Biased* adjustment at the later stages of optimization (i.e., when the convergence rate with pure random selection is slow) increases the convergence speed while reducing the noise (i.e., isolated black or white pixels). Rather than swapping the phases of two randomly chosen pixels, Rozman [132] illustrated that exclusive selection of *interfacial pixels* (with at least one neighboring pixel of different phase) substantially increases the convergence rate. Zhao et al. [123] proposed a similar method and required the selected pixels to have a user-defined connectivity measure. Along the same line, Jiao et al. [37] introduced the surface optimization algorithm where the pixels of the primary phase are grouped into two sets of high and low energy and the random swaps are restricted to the former set (the high energy pixels are those that are *not* connected to any pixel with the same phase). Finally, Tang et al. [38] introduced the different-phase-neighbor (DPN) pixel selection rule and illustrated its superior performance. In the DPN method, the selection probability of any pixel is proportional to the number of neighboring pixels with a different phase (see Fig. 10). As pointed out by Jiao et al. [37], the idea of restricting the pixel selections to the interfacial pixels at the later stages of optimization is similar to the solidification process: when a large enough cluster is formed, it attracts the free particles to its surface in order to reduce the free energy of the system. See [38,89] for comparative studies on the effect of pixel selection rule on convergence rate.

3.1.2.2.2. Optimization procedure via simulated annealing. As mentioned, the existence of multiple local optima in the profile of E necessitates the use of a heuristic optimization algorithm. Simulated annealing and its variants have been widely used for this purpose. *SA* was first introduced by Kirkpatrick et al. [135] who applied it to VLSI layout and graph partitioning. Global minimization of E via *SA* is based on the physical annealing process: If a system is heated up to a high temperature and then slowly cooled down to zero Kelvin, the ground state of the system will be achieved. Within the context of our reconstruction problem, the initial image ($\boldsymbol{Y}$) and the original microstructure ($\boldsymbol{X}$) are, respectively, the system at high temperature and zero Kelvin, while the pixel swaps resemble the system stabilization at each annealing step. Because annealing steps must reduce the energy of the system over time, not every pixel swap is acceptable. Variants of the *SA* algorithm essentially differ on what makes a pixel swap acceptable.

Assuming $k-1$ iterations are done, the *SA* algorithm used by Yeong and Torquato [39,63] and many others [10,38,42,88,106,114,119,121,125,126,129,130,136,137] assigns the probability P to a pixel swap at iteration k based on the Metropolis rule:

$$P = \begin{cases} 1 & E' \leqslant E \\ \exp\left(-\frac{E'-E}{T(k)}\right) & E' > E \end{cases}, \tag{3-5}$$

where $T(k)$ is the so-called temperature at the current iteration and its functional form determines the cooling schedule. Eq. (3-5) has close connections with the Boltzmann distribution: The probability of having a material system at a particular state with temperature T and energy E is given by $P \sim \exp\left(-\frac{E}{T}\right)$. Once P is calculated, the swap will be accepted/rejected by drawing a Bernoulli random variable with event probability P. An inverse logarithmic schedule $T(k) \sim 1/\ln(k)$ would in principle drive the system to its ground state [39]. Such a schedule is too slow for practical applications and so alternative formulations such as $T(k) = T_0\lambda^k$ (where T_0 is the initial temperature and λ is a constant less than unity) are oftentimes practiced.

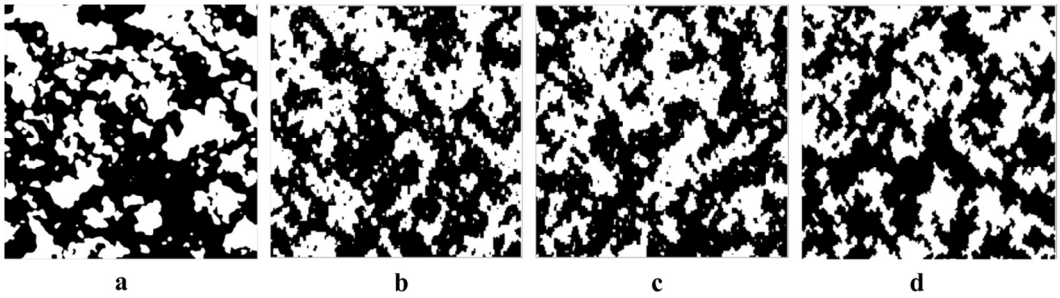

Fig. 10. Effect of acceptance and pixel selection rules on reconstruction via $S_2(r)$: (a) The original microstructure of a PEFC catalyst layer, (b) and (c) reconstructed microstructures with random pixel swapping via, respectively, simulated annealing (SA) and threshold acceptance (TA). The TA algorithm slightly outperforms SA. (d) Reconstructed microstructure with biased pixel swapping (DPN method) via TA. Biased pixel selection is clearly superior to random pixel selection. In all the reconstructions, the accuracy will be improved by incorporating more statistical functions (such as lineal path) into the cost function. Images are reproduced with kind permission from [89].

Dueck et al. [138] developed a variant of *SA* known as threshold acceptance (*TA*) and illustrated its computational superiority. The advantages of using *TA* over the original *SA* algorithm for microstructure reconstruction was first suggested by Cule and Torquato [106] and later investigated by Pantil et al. [89] (see Fig. 10). The main difference between *SA* and *TA* is that while in *SA* the acceptance of a bad swap is decided based on some probability measure, in *TA* it is done based on a threshold value. Specifically, a bad swap at iteration k is accepted if $E' - E \leqslant E^{th}$ where E^{th} is the threshold energy and is calculated as $E^{th} = E_0\lambda^k$. E_0 and λ are, respectively, the initial energy and a constant less than unity.

The simpler version of *TA* known as the Great Deluge (*GD*) algorithm [139] has also been applied to the problem of reconstruction [106,123,132]. In the *GD* algorithm, E^{th} is set to 0. This choice of E^{th} leads to a rapid reduction in the energy of the system but convergence may not be achieved because the acceptance rate becomes impractically low at the final stages of the optimization. To address this issue, Rozman et al. [132] suggest to allow a bad swap (i.e., one that would increase the energy) after a large number of rejections.

In both the *SA* and *TA* algorithms, the optimization parameters such as T_0 or E_0, λ, maximum number of iterations, and number of swaps in each iteration (like the annealing process, in both *SA* and *TA*, each iteration consists of multiple swaps to allow for the stabilization of the system) need to be carefully tuned. For this reason, the *GD* algorithm might appear more appealing as it only has one adjustable input. However, further research must be conducted to investigate the efficiency of *GD* for complex energy profiles (e.g., when multiple correlation functions are incorporated in E).

3.1.2.2.3. Further improvements. A few works [136,140,141] have demonstrated that efficient initialization (e.g., via random fields, see also Section 3.3.2.1) of $\boldsymbol{Y}$ can decrease the computational costs of the YT method. To decrease the computational costs of the YT method, the use of mathematical morphology has also been practiced. Zachary et al. [126] and later Gau et al. [125] used the dilation and erosion (see Appendix A.1) technique to enable effective MCR of, respectively, doubly connected donut-like and filamentary media solely via $S_2(r)$. Here, MCR is basically done on the dilated (and hence more connected) primary phase and then erosion is applied to the reconstructed image to adjust both the V_f and the connectedness. In this approach, the shape and size of the dilation/erosion element must be chosen by the user and one has to make sure that the characteristic features of the original image are not lost through dilation. In other attempts, the hierarchical annealing [142] and the closely related stable phase [130] methods are developed where the reconstruction process is conducted in multiple resolutions to speed up the YT method (see Section 3.5.1 and Appendix A.2). Here, the basic idea is much like that in the Gaussian and Laplacian pyramids [143,144] which are used to compress digital images through a multi-resolution decomposition procedure (see Appendix A.2).

3.2. Physical descriptors

Similar to the previous section, we begin this section by elaborating on what physical descriptors are and how they can be used for microstructure characterization in different material systems (Section 3.2.1). As the main advantage of such descriptors lies in building PSP links, we provide some examples on this line of work in Section 3.2.1.1. We then describe the reconstruction techniques (Section 3.2.2) and conclude the section with some notes on how to statistically choose the best (sub) set of descriptors for MCR (Section 3.2.3).

3.2.1. Characterization

Characterization of finite size objects with physically meaningful descriptors in one-, two-, three-, and higher-dimensional spaces has drawn notable attention in many disciplines including spatial statistics [145], materials science [30,35,93,146–148], forestry [149], and biology [150]. In the realm of materials science, enabling direct elucidation of processing-structure and structure-property relations in crystalline structures and composites (see Fig. 11) is perhaps the most advantageous feature of this characterization method. The study of physical descriptors is particularly motivated by

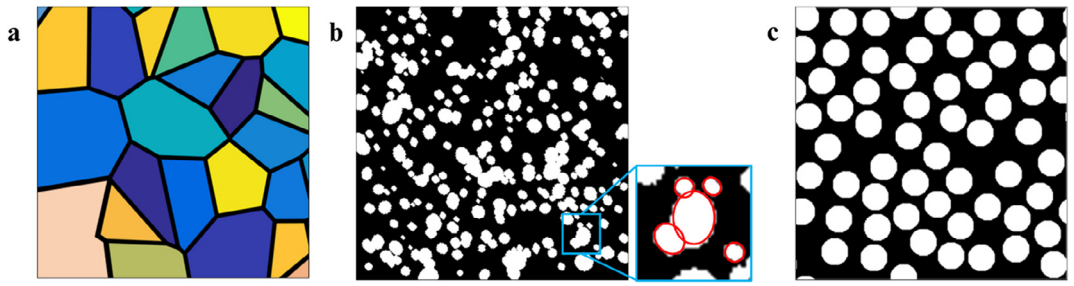

Fig. 11. Sample material systems: (a) Crystalline structure illustrated by Voronoi cells. The colors distinguish different crystallographic orientations. (b) Polymer matrix with randomly-shaped inclusions. The inclusions can be modeled with randomly-shaped overlapping ellipses. (c) Composite with equally-sized fibers.

(*i*) the existence of a large number of descriptors and the strong correlation among them [151], and (*ii*) the fact that the choice of descriptors for efficient characterization depends on the material system as well as, for building structure-property links, on the properties of interest.

Physical descriptors may be classified as either *deterministic* such as V_f or *statistical* like nearest-neighbor distances or area/volume of the Voronoi cells. In the former case, a single value is sufficient to characterize the descriptor and will be assigned to the entire microstructure image (e.g., a binary microstructure with inclusion V_f of 40%). In the latter case, however, a cumulative distribution function is required for descriptor characterization (e.g., the set of inclusions' diameter or nearest neighbor in a fibrous composite will follow a particular distribution. See, e.g., [35]). Table 1 lists some descriptors that have been successfully used for building structure-property links in composite materials. Descriptors can also be categorized based on the *information scale* that they characterize [30,83]: for composites, at the highest scale is the V_f of the individual constituents, followed by, respectively, the descriptors describing the dispersion status and the geometry of the inclusions (in other words, a high-scale descriptor is assigned to the entire microstructure while a low-scale descriptor to the individual features). In alloys, on the other hand, the descriptors characterizing the alloy texture are at the highest scale and are followed by those characterizing the individual grains. It should be noted that the above classifications depend, to some extent, on the material system as well as on the reconstruction procedure (see the example in Section 5.1). For example, in a polymer nanocomposite with equally-sized non-overlapping circular inclusions, both the V_f and the number of clusters (or equivalently the number of inclusions) may be categorized as deterministic descriptors at the highest length-scale. If the inclusions are of different size and allowed to overlap, however, the number of clusters will describe the dispersion status and the number of particles will be a statistical descriptor. For more discussions on this, see [30,151,152].

Table 1
Characterization with physical descriptors: List of physical descriptors commonly used for characterizing the primary phase in microstructures consisting of a matrix and overlapping inclusions with random size and orientation. While some of the descriptors are deterministic (i.e., a unique number can be assigned to the entire microstructure or all the inclusions), others are statistical and therefore identified via a distribution. It must be highlighted that the descriptor type, to some extent, depends on the material system of interest. See the text for more details.

Descriptor symbol	Description	Type
Scale I: Composition		
V_f	Volume fraction	Deterministic
Scale II: Dispersion		
r_{ncd}	Cluster's nearest center distance	Statistical
θ	Orientation angle of a cluster's principle axis [153]	Statistical
I_{filler}	Total surface area	Deterministic
N_c	Number of clusters	Deterministic
Scale III: Geometry		
r_p	Pore sizes (inscribed circle's radius) [154]	Statistical
A	Area (of individual particles)	Statistical
r_c	Equivalent radius, $r_c = \sqrt{A/\pi}$	Statistical
δ_{cmp}	Compactness [151]	Statistical
δ_{rnd}	Roundness [155]	Statistical
δ_{ecc}	Eccentricity [155]	Statistical
δ_{asp}	Aspect ratio [156,157]	Statistical
δ_{rec}	Rectangularity [155]	Statistical
δ_{tor}	Tortuosity [155]	Statistical

To characterize $\boldsymbol{X}$ (the digitized image of a microstructure) with physical descriptors, first various image segmentation techniques [158], such as edge or cluster detection, are applied to identify all the objects (i.e., lines, edges, clusters, etc.) with the primary phase. Then, the objects are analyzed to find the values and distributional characteristics (such as the mean and variance) of, respectively, deterministic and statistical descriptors. Unless the set of descriptors is chosen a priori (e.g., based on experience), these steps generally result in many descriptors, some of which might be strongly correlated or have negligible effect on either the material property or the quality of characterization. To address this issue, Xu et al. [151] developed a machine learning based approach where after identifying the descriptors of an ensemble of nanocomposite samples (prior to constructing structure-property relations), the number of descriptors is reduced by excluding both the correlated ones and those which affect the statistical functions (introduced in Section 3.1.1) the least. In Section 3.2.3, some dimension reduction techniques are introduced that can be used to identify the most effective descriptors for microstructure characterization.

In computational materials modeling [23,159] or design [30,83,137,151,152,160], rather than learning the descriptors from a microstructure image, one determines the values of a set of *pre-selected* descriptors (via, e.g., design of experiments [161–165]) and then proceeds to reconstruct the corresponding virtual microstructures and estimate their properties (via e.g., FEA). In this application, it must be ensured that the set of descriptors (*i*) sufficiently characterizes the material system of interest to eliminate (or at least minimize) the sources of uncertainty in building structure-property relations (see Section 5.1 for an example), and (*ii*) all the points within the design hull (defined by the set of descriptors) correspond to realizable structures.

To enable the reconstruction of a 3*D* sample given 2*D* image(s) from it, stereological techniques can be used. Such techniques commonly make some assumptions on the 3*D* geometrical features and then estimate various attributes of 3*D* objects (such as number density, particle surface area, size, etc.) from those of the corresponding 2*D* objects [83,166]. For example, the ST method introduced in [166,167] can be used to construct 3*D* particle distributions from 2*D* micrographs with the assumption that the particles are spherical and have circular cross-section. For another example, the 3*D* aspect ratio in an isotropic material system with elliptical inclusions may be estimated from a 2*D* image by [83]:

$$\delta_{3D} = \frac{1}{\theta^*}\int_0^{\theta^*} f^{-1}(\delta_{2D})d\theta, \tag{3-6}$$

where f^{-1} is the inverse function for mapping 2*D* aspect ratio to 3*D* aspect ratio, and θ^* is the upper bound on the rotation angle given the thickness of the sample. θ^* is defined as:

$$Max\left\{ y \left| \frac{(x\cdot\cos\theta + y\cdot\sin\theta)^2}{f^{-1}(\delta_{2D})^2} + (-x\cdot\sin\theta + y\cdot\cos\theta)^2 = a_{2D}^2 \right.\right\} = \tau, \tag{3-7}$$

where τ, the distance between inclusion centers and the nearest surface of the layer, follows the uniform distribution $uni(0,H)$. We note that, there are many inclusions that, after projection to 2*D*, will have the same aspect ratio (see Fig. 12b) and in Eq. (3-6) it is assumed that all such inclusions contribute equally to the integral (i.e., they are equally probable to be found in the sample). Assuming the minor radii are the same in 2*D* and 3*D* ($a_{2D} = a_{3D}$), f can be formulated as:

$$\delta_{2D} = f(\delta_{3D}) = \frac{max\left\{ x \left| \frac{(x\cdot\cos\theta + y\cdot\sin\theta)^2}{\delta_{3D}^2} + (-x\cdot\sin\theta + y\cdot\cos\theta)^2 = a_{2D}^2 \right.\right\}}{a_{2D}}, \tag{3-8}$$

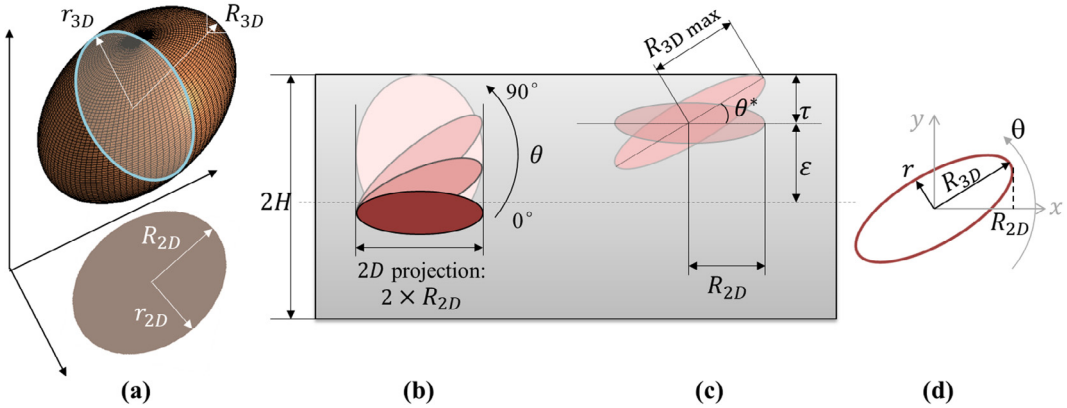

Fig. 12. Effect of thin-layer observation on the 2*D* projection of a 3*D* inclusion: (a) The 3*D* ellipsoidal inclusion and its 2*D* projection, (b) possible 3*D* structures and the correspondent rotation angles for a single 2*D* projection, (c) upper bound on the rotation angle given the limited thickness, and (d) $x-y$ coordinate system for deducing the projection relation. Images are reproduced with kind permission from [83].

3.2.1.1. Building structure-property maps via physical descriptors. As noted above, perhaps the primary advantage of characterization with physical descriptors lies in unraveling the direct effect of microstructure features on material property. In what follows, we summarize some of the findings to highlight the significance and potentials of this line of work.

The family of nearest-neighbor distances have been rigorously investigated as they play an important role in transport processes in particulate heterogeneous systems [9], microstructural evolution during recrystallization [168], particle coarsening [169], and liquid phase sintering [168]. Refs. [170–172] provide analytical expressions for relating the mean of nearest neighbor distances to the *n*-point correlation functions of the uniform random microstructures of impenetrable spherical particles. In [146] an expression for the mean values of first-, second-, and higher-order (up to sixth) nearest-neighbor distances for systems of equally-sized hard-core spheres in a 3*D* space is derived.

It is well accepted in the literature [29,30,83,151,152,173–177] that the degree of property enhancement of composites (with a metal or polymer matrix), in addition to the constituent material and interface properties, depends on morphological factors such as volume fraction, size, shape, and spatial distribution of the reinforcements. In metal-matrix composites, Brockenbrough et al. [178] illustrated that the effect of fiber distribution is significant at lower volume fractions and Christman et al. [179] concluded that clustering of the fibers significantly reduces the flow stress and strain hardening. The effect of various morphologies on the damage initiation and evolution process in ductile matrix composites has also been investigated via the Voronoi cell finite element [180,181]. These findings have been in line with the experimental studies of, e.g., Lewandowski and co-workers [182,183] and Mummery et al. [184].

Similar studies have been conducted on polymer composites [175]. For example, it is found that wide distribution of particle size increases the electrical conductivity of rubber composites with carbon black fillers because the probability of agglomeration increases and consequently the average gap width between conducting particles reduces.

In crystalline structures, perhaps the most critical features correspond to grains. The effect of shape, size, and crystallographic orientation of the grains on the properties has been extensively studied [185,186] and it is found, e.g., that intergranular corrosion is sensitive to grain boundaries [78]. In addition, grain size has an inverse relation with yield strength and toughness and thus its refinement is often used to improve mechanical properties.

3.2.2. Reconstruction

Once the set of physical descriptors is defined and their values/distributional characteristics are calculated (from an original microstructure sample) or prescribed (by the user), reconstruction of a microstructure sample can be achieved by adjusting an initial structure to match its descriptors to the target ones. The adjustment procedure is similar to that described in Section 3.1.2 in that it commonly involves optimization for matching some of the descriptors.

Xu. et al. [30] developed a hierarchical reconstruction methodology for reconstructing composites with overlapping elliptical inclusions of random size (see Fig. 13). Briefly, this method starts the reconstruction with the descriptors at the highest scale level. Depending on the identified descriptors in the characterization step, this may involve, e.g., determining the V_f and number of inclusions in a particulate structure. The method then proceeds to adjust the dispersion status of the inclusions that often requires a heuristic optimization algorithm such as *SA*, e.g., to match the nearest neighbor distances or Voronoi cell area/volume to the target values. Finally, local geometrical characteristics such as particle orientation and shape are assigned. Zhang et al. [152] integrated a decomposition and reassembly strategy into this method to enable effective reconstruction of composite microstructures with very low V_f and uneven distribution of cluster size. The hierarchical reconstruction method has been successfully used in materials design [29,30,83,151,152,160] and modeling [23] where the physical descriptors are used as input variables in, respectively, the optimization process and the learned constitutive law.

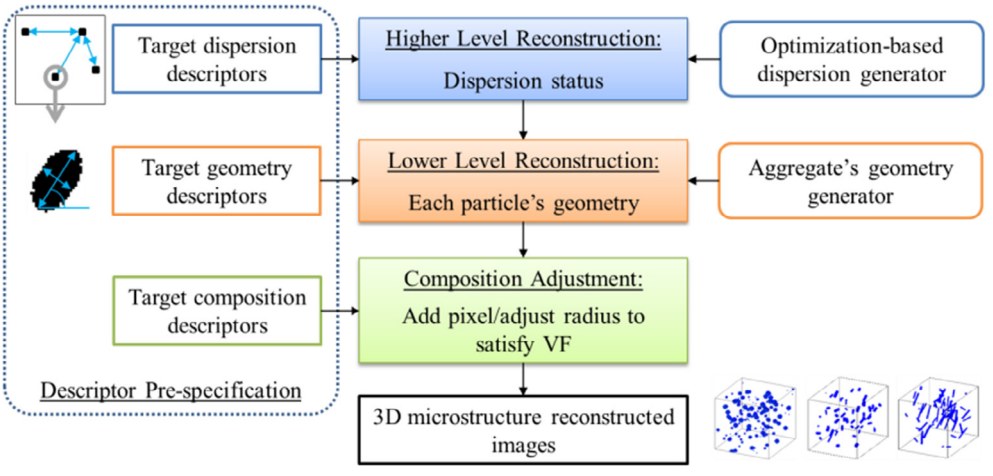

Fig. 13. Flowchart of the hierarchical descriptor-based reconstruction algorithm. Image is reproduced with kind permission from [30].

In crystalline structures, most approaches rely on ellipsoid packing and Voronoi tessellations to locate the grains in the reconstructed microstructure [186–190]. The crystallographic orientations are then assigned either randomly or following a certain distribution. As an example, the microstructure builder software package [78,187] has been developed to reconstruct 3*D* polycrystalline materials (see Fig. 14). It starts with an overpacked structure and swaps ellipsoids in and out with simulated annealing to achieve an optimal arrangement. Then, crystallographic orientations are assigned. The input is typically grain size and shape data and the output is a 3*D* voxel microstructure. This software has been greatly improved over the past few years and evolved to DREAM.3D [191].

Physics-based evolving algorithms such as Johnson–Mehl [192] and cellular algorithms [192,193] or the Voronoi polyhedral model [194] use a similar approach. In such methods, first a prescribed number of points representing the individual grains are positioned via, e.g., Poisson process and then grown to result into a crystalline microstructure (see Fig. 2 in [192]). The nucleation (i.e., the growth) is produced in the kinetic limit of simultaneous activation of the available nucleation sites and its rate is assumed to be isotropic and constant over all the sites. For a recent improvement on such algorithms see the work of Barbe et al. [195].

3.2.3. Dimension reduction

As discussed above, physical descriptors are well suited for building PSP relationships and materials design. In many cases, however, many descriptors may be identified in the characterization step when the "curse of dimensionality" (i.e., the large number of descriptors) would make reconstruction, PSP analyses, and materials design difficult. In this section, we elaborate on some dimensionality reduction approaches to determine the set of descriptors that best characterize a material system. Hereafter, we use the terms "feature" and "descriptor" interchangeably as the former term is more frequently used in the machine learning and statistics community where most dimensionality reduction techniques originate from.

3.2.3.1. Feature selection. In general, the objective of feature selection is threefold: improving predictive performance, providing more cost-effective predictors, and facilitating the discovery of underlying probabilistic principle of data generation [196].

Variable ranking is one of the most common techniques for feature selection where a scoring function is used to determine the score (i.e., significance) of each feature [196]. With higher scores indicating more important features, variable ranking enables the identification of the most informative features for building parsimonious predictive models. For example, Xu et al. [151] employed a two-step feature selection process using descriptor pairwise correlation analysis and RReliefF variable ranking approach [197] to select the physical descriptors that best control the damping property of polymer composites. In another work, Hassinger et al. [198] used feature selection and identified the interfacial area as the most significant physical descriptor that can be correlated via two heuristic processing energy descriptors to establish a predictive processing-structure model for polymer extrusion.

Exploratory factor analysis [199,200] is another technique for identifying the important features by grouping the correlated descriptors together to build a set of latent common factors. Zhang et al. [152] employed factor analysis into a structural equation-modeling approach for the design of dielectric polymer composites.

In short, feature selection approaches are generally applied in MCR when the identified descriptors are too many to work with and/or believed to be highly correlated. With feature selection, redundant statistical features can be dropped before further analyses are conducted.

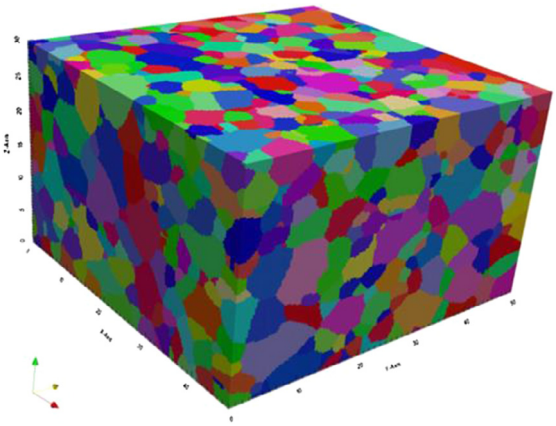

Fig. 14. Reconstruction with microstructure builder: Visualization of a synthetic polycrystalline microstructure with DREAM.3D. Image is reproduced with kind permission from [191].

3.2.3.2. Feature extraction. Different from feature selection, in this set of approaches the feature space is transformed into a lower dimensional one where the physical interpretations are diminished. In this section, we survey two major categories of feature extraction methods, namely, linear and nonlinear embedding.

3.2.3.2.1. Linear embedding. Among the feature extraction approaches that have been applied in materials science, linear embedding is the simplest but most powerful technique. Principle Component Analysis (PCA) is perhaps the most famous linear dimensionality reduction method. In short, PCA consists of an orthogonal transformation that converts a set of possibly correlated variables into a set of linearly uncorrelated variables known as the principle components (see [201] for more details).

A model reduction procedure is established in [202] with PCA and Karhunun-Loeve expansion to convert the high-dimensional feature space of 3*D* microstructure images to lower-dimensional approximations. This low-dimensional representation is then demonstrated to be very effective when taken as the inputs in the authors' multi-scale analysis framework of diffusion behavior. The application of such schema is limited to binary microstructures with simple spatial correlations. In another work, Choudhury et al. [203] quantify similarity/difference between microstructures in ternary eutectic alloys by first converting them to some spatial-correlation functions, and subsequently employing PCA to obtain an objective low-dimensional representation of those functions. These representations are then taken as the predictors for classifying different alloys in the later steps. To model the variability of workpiece properties in a deformation process, Wen et al. [204] study the microstructure variability by applying a two-level bi-orthogonal Karhunen-Loeve decomposition on the high-dimensional microstructure feature space. The set of heuristic features consists of the grain size and axis-angle in FCC nickel alloy and is shown to effectively capture the variability of the grains in the microstructures.

3.2.3.2.2. Nonlinear embedding. Two major sets of nonlinear embedding approaches have been investigated in the literature [205]. One set is the bottom-up approach where it is assumed that a nonlinear manifold (embedded in the original feature space) governs the data distribution. Here, one attempts to preserve the local geometry of data on the manifold, i.e., to ensure the nearby points in the original feature space would also be close in the lower-dimensional representations. Locally linear embedding [206], Laplacian eigenmaps [207], and Hessian eigenmaps [208] are examples of this set of nonlinear embedding approaches.

The second major set is the top-down approach where one attempts to preserve the geometric relations at all scales (i.e. nearby and distant points on the manifold map to, respectively, nearby and distant points in the lower-dimensional space). Examples of this set of approaches include Isomap and its variants [209,210].

The top-down approach is used in Li et al.'s work [205] for creating low-dimensional surrogates that characterize polycrystalline microstructures. In particular, the Isomap method is first used to reduce the dimensionality of the feature sets (grain size and its orientation distribution) into lower-dimensional surrogates ($\mathcal{M} \rightarrow$ Å) and then reconstruction is achieved by sampling from the surrogate space via an inverse mapping (Å $\rightarrow \mathcal{M}$). While the $\mathcal{M} \rightarrow$ Å process is straightforward and inexpensive, the inverse mappings Å $\rightarrow \mathcal{M}$ requires expensive step-wised iterative calculations for each of the statistical moments.

While not preserving as many physical interpretations as feature selection methods do, feature extraction techniques are advantageous in lowering the dimensionality of the design space and are trained easier to achieve a higher predictive accuracy. Hence, it is generally suggested to apply these methods when a very accurate and robust prediction model is desirable. It should be noted that, the choice of feature extraction algorithms requires in-depth knowledge and investigation of the complexity of the material systems.

3.3. Spectral density function

The Fourier transform (FT) is ubiquitously used across broad and diverse areas of science and engineering including biomedical engineering [211], spectroscopy [212], numerical methods [213], radar [214], acoustics [215], image processing [216], structural analysis [217], and mechanics [218–220]. In essence, the FT decomposes a waveform (e.g., a 1*D* signal or a 2*D* image) into a sum of sinusoids of different frequencies [221]. In the realm of MCR, such a decomposition would represent the dominant microstructural features in the frequency space. Previous research suggests that the spectral density function (SDF) (defined in Section 3.3.1) is sufficient to characterize some complex heterogeneous microstructures and, given an SDF, the corresponding original sample can be reconstructed using, e.g., phase recovery techniques [97]. SDF generally takes a simple parametric form; enabling materials design by considerably reducing the microstructure design space (aka microstructure hull [7]). In addition to these capabilities, previous studies have shown that SDF is closely related to some microstructural properties, e.g., in optical applications where it has been used for design representation in the optimization of light trapping performance of solar cells [222]. Many researchers have demonstrated the use of SDF for characterizing quasi-random structures made by bottom-up manufacturing processes such as nanoparticle self-assembly and nano-wrinkling [223,224]. For instance, Yu and Zhang [32] recently developed an SDF-based concurrent design approach to bridge the gap between structure-property and process-structure links in light-trapping nanostructures made of thin-film solar cells. In this section, we first introduce microstructure characterization using SDF and then summarize two different reconstruction techniques, namely, random field (RF) and disk packing.

Although SDF is closely related to the two-point correlation function (in that it describes the structure correlations in a reciprocal frequency space rather than the real space), we devote an entire section to it because, as argued in Section 3.3.1,

SDF provides a more physically meaningful representation than $S_2(r)$ and the associated reconstruction techniques are largely different.

3.3.1. Characterization

To obtain the SDF of a microstructure, we start with its Fourier transform $F(\boldsymbol{k})$, which is a complex function of frequency:

$$F(\boldsymbol{k}) = \mathsf{F}\{\boldsymbol{X}\} = A_{\boldsymbol{k}} e^{i\varnothing_{\boldsymbol{k}}}, \tag{3-9}$$

where $\boldsymbol{X}$ represents the binary microstructure image, $\mathsf{F}\{\cdot\}$ denotes the Fourier transform operator, $\boldsymbol{k}$ is a vector denoting the frequency, $A_{\boldsymbol{k}}$ and $\varnothing_{\boldsymbol{k}}$ represent, respectively, the magnitude and phase angle of the FT, and $i = \sqrt{-1}$. A microstructure's SDF, $\rho(\boldsymbol{k})$, is then defined as the squared magnitude of its FT:

$$\rho(\boldsymbol{k}) = |\mathsf{F}\{\boldsymbol{X}\}|^2 = A_{\boldsymbol{k}}^2, \tag{3-10}$$

For an isotropic microstructure, the vector $\boldsymbol{k}$ can be reduced to a scalar k through radial averaging to convert $\rho(\boldsymbol{k})$ to a one-dimensional function satisfying $\rho(k) = \rho(|\boldsymbol{k}| = k)$.

The SDF essentially describes the distribution of the Fourier components over the frequency range where different frequencies correspond to real space features at different length scales. According to Winner-Khinchin theorem [225], the two-point correlation function is equivalent to the inverse FT of SDF:

$$S_2(\mathrm{r}) = \mathsf{F}^{-1}\{\rho(\mathrm{k})\}, \tag{3-11}$$

Nevertheless, compared to $S_2(r)$, SDF provides a more convenient representation for designing microstructures. To illustrate this, we compare the SDF and $S_2(r)$ characterization of four *quasi-random* microstructures in Fig. 15. These structures mainly differ in feature scale and randomness degree and so it is highly desired for the characterization method to sensibly quantify such differences. The key points of the studies demonstrated in Fig. 15 are as follows:

- In microstructures with a single dominant feature scale (Fig. 15a and b), unlike $S_2(r)$, the SDF clearly demonstrates the increase (or decrease) in the feature scale and can be parameterized parsimoniously. In Fig. 15e and f, the SDF's can be simply approximated via step functions (the red-dashed curves, which can be fully characterized with only 2 parameters) while the corresponding $S_2(r)$ curves (Fig. 15i and j) require complex parameterization.

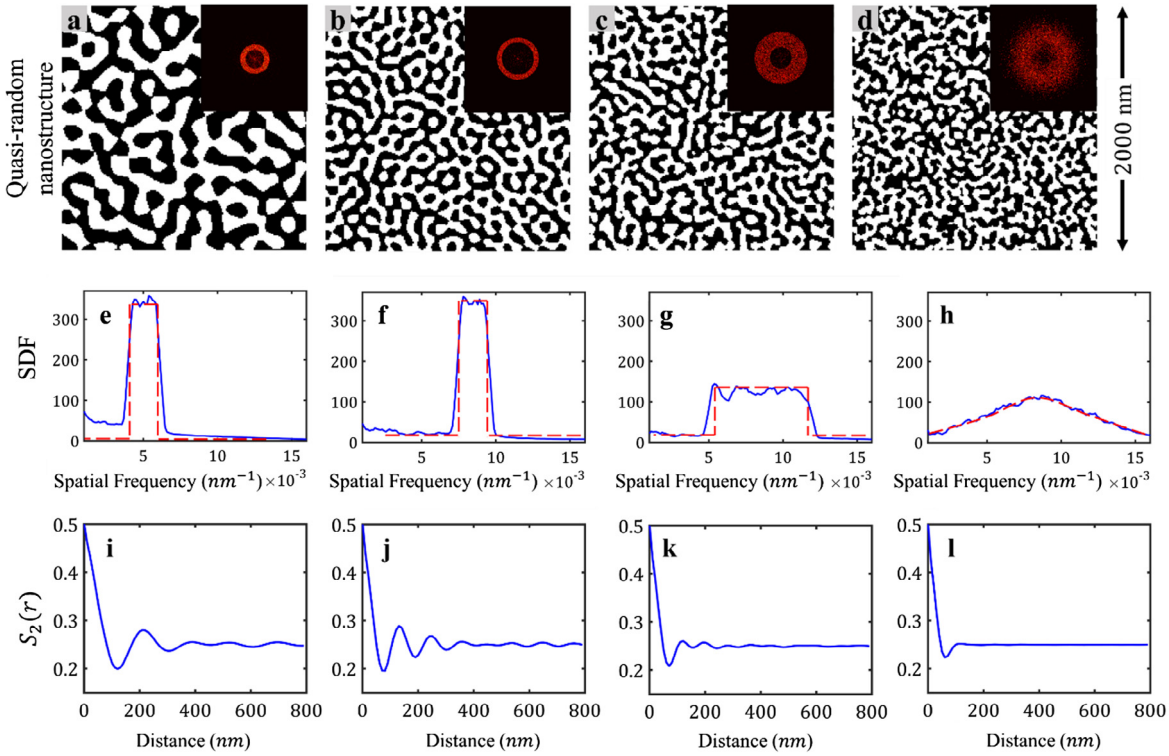

Fig. 15. Comparison between SDF and $S_2(r)$ characterization of four quasi-random microstructures: The differences in the microstructures such as feature scale and degree of randomness are better represented via SDF rather than $S_2(r)$. The insets in a through d are the 2*D* SDF's of the microstructures and due to isotropy, these 2*D* SDF's reduce to the 1*D* SDF's in e through h. See the text for more details.

- SDF significantly reduces design dimensionality and is very efficient in highlighting the effect of randomness degree in disordered media through widening the effective power spectrum range (see Fig. 15a–c and compare their corresponding SDF representations) while the same phenomenon is captured by difficult-to-characterize damped oscillations in $S_2(r)$ curves (see Fig. 15i–k).
- In microstructures with a spread feature scale (Fig. 15d), the SDF is more illustrative than $S_2(r)$ (compare Fig. 15h and l) and can be simply formulated as a Gaussian distribution function.
- The key characteristics in SDF have direct associations with material functionalities. For example, the step SDF in Fig. 15g and the spread SDF in Fig. 15h corresponds to morphologies which are appropriate for, respectively, light-trapping in solar cells and producing angularly independent white color coating.

In short, SDF provides a convenient and informative representation for a wide range of microstructures and enables more straightforward mathematical modeling as compared with $S_2(r)$. Identifying the appropriate parametric form of the SDF is achieved through both physical understanding of the material processing and image analysis of the physical samples. For example, a formulation of the SDF conforming to the Landau free-energy theory has been derived [226] to describe the quasi-random materials born through the spinodal-decomposition process:

$$\rho(k) = \frac{A}{(k^2 + k_0^2)(k^4 + 2k^2b^2\cos(2t) + b^4)}, \tag{3-12}$$

where A is a normalization factor ensuring the integral of SDF $\rho(k)$ over $(0, +\infty)$ equals 1, and (k_0, b, t) are the three governing parameters that serve as the microstructure design variables in, e.g., an optimization process for achieving desired property.

3.3.2. Reconstruction

In this section, we illustrate two efficient analytical reconstruction methods that use the SDF representation as input to reconstruct microstructures with different morphological characteristics; including channel-type and disk-type. For the channel-type microstructures, we adopt the level-cutting random field approach for reconstruction, while for the disk-type we employ a disk-packing algorithm. Both approaches are analytical and hence very efficient. Though demonstrated for 2*D* microstructures, the approaches can be easily extended to 3*D* reconstructions.

3.3.2.1. Level-cutting random field. Realizing the microstructure image as a multi-variate sample taken from an underlying random field (RF, e.g., Poisson [227,228] and, esp., Gaussian [60,136,140,141,229]), one can deduce that the reconstruction of statistically equivalent microstructures can be achieved by finding that RF. Once the RF is found, reconstructed microstructures are realized by simply taking realizations from it, and subsequently level-cutting them. To find the RF that shares the same characteristics (i.e., SDF) with the original microstructure sample, two types of simulations are developed in the literature: on-lattice (discontinuous) [59,60] and off-lattice (continuous) [136,229–231].

The on-lattice simulation of 3*D* Gaussian random field (GRF) using linear filtering [60] adapts well to the problems dealing with numerical partial differential equations [229]. However, this method requires solving a set of linear equations and the discontinuous description of the pore network is problematic for, e.g., molecular dynamics computation, simulation of Brownian motion, or excitation relaxation inside disordered porous medium [229]. For this reason, Cahn's scheme [231] is widely used for efficiently generating continuous GRF and is discussed in detail below. Reconstruction using GRF is in general efficient and accurate for random porous media but the performance deteriorates for particulate structures. To address this issue, hybrid approaches are developed [140,141,227,228] where the results from GRF reconstruction are refined through an optimization post-process via *SA* (detailed in Section 3.1.2). In [227,228], the author proposed to use level-cut filtered Poisson fields based on Monte Carlo simulation to generate two-phase microstructures. It is shown that the widely used level-cut Gaussian fields are special cases of level-cut filtered Poisson fields [227]. The merit of using filtered Poisson field is that, unlike GRF, various geometries of inclusions such as high aspect ratio ellipses can be realized by adjusting the parameters of the distribution (although the adjustment is not done in a straightforward manner).

3.3.2.1.1. Cahn's scheme. Cahn's Scheme [226,230,231] is essentially an analytical approach for generating realizations (i.e., reconstructed microstructure samples) from a GRF that has the same SDF as the original microstructure. This is achieved through:

$$Y(\boldsymbol{r}) = \left(\frac{2}{N}\right)^{1/2} \sum_{i=1}^{N} cos(k_i \hat{\boldsymbol{k}}_i \cdot \boldsymbol{r} + \phi_i), \tag{3-13}$$

where ϕ_i and $\hat{\boldsymbol{k}}_i$ are uniformly distributed on, respectively, $[0, 2\pi]$ and a unit circle. k_i is a random variable whose probability density function, $P(k)$, is determined by the SDF: $P(k) = \rho(k)k$ for a 2*D* GRF and $P(k) = \rho(k)k^2$ for a 3*D* GRF. After generating the GRF, a level cutting process is often implemented to obtain a binary structure (see Fig. 16a and b) by making the volume fraction equal to a desired value (e.g. 50%). As shown in Fig. 16c, $\rho(k)$ in this example resembles a ring shape in 2*D* or, once radially averaged, a Dirac delta function in 1*D*.

To explain why Cahn's scheme works, we note that a standard GRF with zero mean is completely characterized by its two-point correlation function [69]. In addition, as defined in Eq. (3-11), $S_2(r)$ is the inverse Fourier transform of $\rho(k)$ and can be

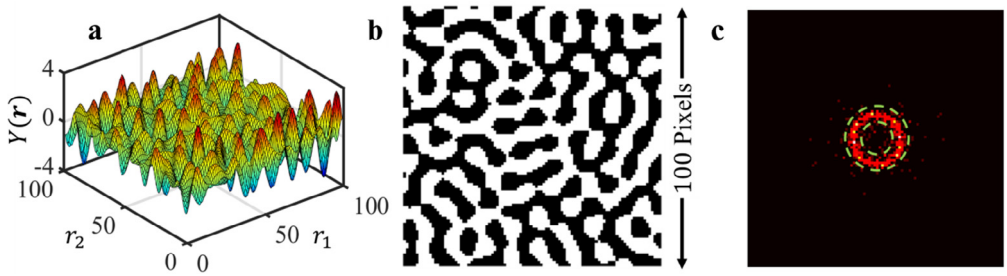

Fig. 16. Reconstruction of a binary microstructure via Cahn's Scheme given a target SDF: (a) A zero mean GRF generated by Cahn's scheme. (b) The binary microstructure after level cutting the GRF in (a) with $V_f = 50\%$. (c) Thin ring shape 2*D* SDF of the binary microstructure in (b). The target SDF is uniformly distributed within the narrow band marked by the dashed green donut.

represented by an n-dimensional Fourier integral. Using the Wiener-Khintchine theorem, this *n*-dimensional integral takes a simpler form and is replaced by a one-dimensional Hankel transform:

$$S_2(r) = 2^{\frac{n-2}{2}}\Gamma(\mathrm{n}/2)\int_0^{\infty}\frac{J_{\frac{n-2}{2}}(kr)}{(kr)^{\frac{n-2}{2}}}k^{n-1}\rho(k)dk, \tag{3-14}$$

where *J* is the Bessel function of the first kind, *n* is the dimension of the GRF, and $\rho(k)$ is the one-dimensional SDF of the GRF. For a 2*D* GRF, the above reduces to:

$$S_2(r) = \int_0^{\infty} J_0(kr)\rho(k)kdk = \int_0^{\infty} J_0(kr)P(k)dk, \tag{3-15}$$

Essentially, the above formulas illustrate that $S_2(r)$ and SDF can both be used for characterizing a GRF.

3.3.2.2. Disk-packing. Disk-type structures are another important class of microstructures that can be fabricated by bottom-up processes such as self-assembly of polymer nanospheres and cuttlefish ink nanoparticles [232]. Existing works have investigated the possibility of packing mono-sized disk-shape particles to achieve certain types of SDF by adjusting the center distances and number of disks. Most of these structures are loosely packed and possess a thin ring shape SDF [233]. Yu and Zhang recently established a semi-analytical relationship between the distribution of disk numbers and disk radius for targeted SDF profiles. The approach is based on a dense disk packing algorithm with a simple strategy: Two neighboring disks repel each other if they overlap and vice versa. See [32] for the algorithmic details.

The SDF of mono-size disk packing structures resembles a ring-shape in 2*D* (or, after radially averaging it, Dirac Delta function in 1*D*) with a simplified relationship between the peak frequency location k_{peak} of the SDF and disk radius *r* [32]:

$$r \approx \frac{L}{1.69}k_{peak}, \tag{3-16}$$

where *L* is the side length of the structure. Complex SDF profiles can be achieved by packing disks with different sizes. Our empirical studies in [32] suggest that the intensity of SDF at a certain frequency k_i is proportional to the total area occupied by the corresponding disks of radius r_i, with k_i and r_i satisfying the relationship in Eq. (3-16): $\rho(k_i) \propto N_i r_i^2$, where N_i is the number of disks of radius r_i. In addition, when the total areas of disks for each size are equal, the corresponding $\rho(k_i)$ are approximately equal. Thus, one can estimate the distribution of disk sizes for arbitrary SDF. As an example, the following condition should hold to achieve a uniform SDF:

$$N_1 r_2^2 = N_2 r_2^2 = \cdots = N_n r_n^2, \tag{3-17}$$

Three examples are shown in Fig. 17 to illustrate the effectiveness and flexibility of this approach. Fig. 17b and c are microstructures with very similar underlying (radially averaged) SDF (Fig. 17a) that consists of two narrow bands. Although the particle geometries are distinct (circle vs. triangle), the underlying spatial correlation and feature scales are close, resulting in similar SDFs. The SDF in Fig. 17d follows a broad Gaussian distribution and the corresponding microstructure in Fig. 17e possesses multiple feature scales (as opposed to those in Fig. 17b and c with only two apparent feature scales).

3.4. Machine learning

Recently, some machine learning techniques have been successfully applied to MCR. The main motivation of these techniques lies in their speed and flexibility in that they are generally, as opposed to the reviewed techniques so far, applicable to a broader range of material systems. We have categorized such works into three groups and review them in the following sections.

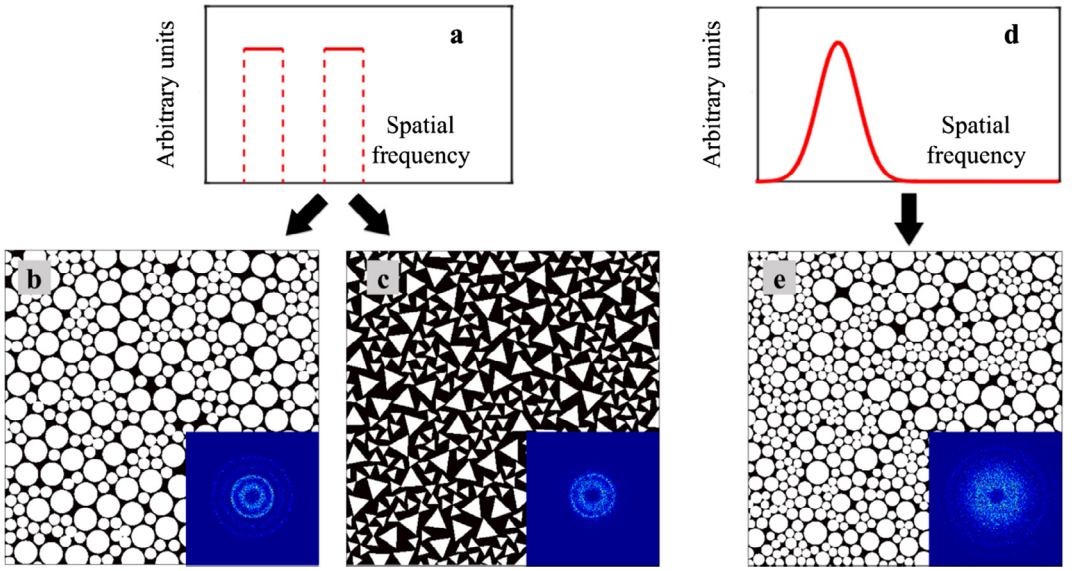

Fig. 17. Reconstruction of quasi-random nanostructures via the particle packing algorithm: The SDF (a) of the microstructures in (b) and (c) consists of two narrow bands that represent the two distinct feature scales, while the SDF (d) of microstructure in (e) has a Gaussian distribution. The insets in b, c, and e are the 2*D* SDFs of the corresponding microstructures and due to isotropy, these 2*D* SDFs reduce to the 1*D* SDFs in a and d.

3.4.1. Unsupervised deep learning

The emerging research area of deep learning has stimulated a plethora of applications across a variety of disciplines [234]. In general, deep learning models are hierarchies of stacked neural network layers that encode data into hyper-dimensional spaces through linear multiplications and non-linear transformations. While widely recognized nowadays, neural networks dropped to near pseudoscience status during the "AI winter" in the 1990 s because training large-scale deep neural networks suffered from the so called vanishing gradient problem. The introduction of the layer-by-layer pre-training strategy via unsupervised learning by Hinton et al. [235] addressed this issue and revitalized the use of such models in the field of artificial intelligence (AI) [236,237].

While unsupervised deep learning is widely used in the computer science community, its application in materials science is still scarce. Recently, Cang et al. [238,239] employed the convolutional deep belief network (CDBN) [240] to hierarchically extract implicit features and realize reconstructions for multi-scale anisotropic alloy microstructure. Their model consists of three convolutional restricted Boltzmann machine (CRBM) layers, with a pooling layer between the second and the third CRBM layers for further dimensionality reduction (see Fig. 18). While this model preforms satisfactorily in extracting features from complex morphologies such as anisotropic alloy microstructures, it relies on a heuristic post-processing step, flow-based difference-of-Gaussian (fDoG) filtering, which is specifically designed for the alloy system studied. Additionally, unlike other deep learning models used in applications outside of MCR that are trained on millions of data, the model is trained on only a hundred microstructure images. Insufficient training data may lead to an over-fitted model that is only

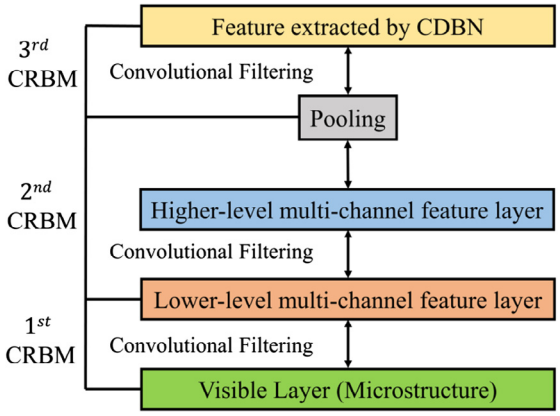

Fig. 18. The network configuration of Cang et al. [241] CDBN model for MCR: The model has three layers with a pooling one for further dimensionality reduction.

applicable to a specific material (sub)system (from which the data is collected) and difficult to be generalized. Since training deep learning models that generalize well requires a significantly large amount of data, this model is suggested to be applied when a cheap microstructure evolution simulation is available or a better strategy for parameter initialization is found. It should be also noted that the potential of such deep learning model in materials design has not been fully exploited. For example, the features extracted by CDBN or other unsupervised deep learning models have not been incorporated with processing conditions and materials properties to establish predictive models. More investigations on this direction are needed.

3.4.2. Instance-based learning

Different from the characterization approaches that aim to quantify microstructure using statistical models, there is another category of approaches that perform instance similarity search in databases of microstructures [242]. Briefly, the essence of such approaches lies in three major steps: (*i*) Creating a database that contains representative sample instances, (*ii*) Defining the feature space and similarity metrics applicable to all the samples in the database, and (*iii*) Conducting brute-force or algorithmic search to find the most similar instance in the database upon query. In general, there are two major challenges associated with these instance-based approaches: data sufficiency and search efficiency. As for data sufficiency, not only the database should be large enough, but also the feasible space must be well covered. To this end, a sophisticated data evolution approach is essential as experimental data in materials science is oftentimes limited quantitatively and qualitatively. With a data evolution process, the size of database increases exponentially and so it is essential to develop a rapid searching mechanism that can efficiently locate the most similar instance in the database upon a query.

Sundararaghavan et al. [15] addressed these two challenges by employing a Monte Carlo based microstructure evolution approach and hierarchical classification models to reconstruct a 3*D* sample given a 2*D* microstructure image. Specifically, a large database of 3*D* microstructure instances created computationally and experimentally is employed. Lower-level features of these instances, as illustrated in Fig. 19, are then extracted for similarity analysis to create a hierarchical support vector machine classification model. Once a 2*D* microstructure example is provided, the classification model performs a fast search through the instance database and finds the most likely equivalent 3*D* microstructure as the reconstructed microstructure. The retrieved 3*D* microstructures by this approach, are then validated by comparing their point-correlation functions to those of the original 2*D* image. While a good match between the statistical functions suggests the validity of this approach, the hierarchical classification model highly depends on the lower-level features (which are selected heuristically) and it is yet unknown how this approach would perform if alternative features are chosen.

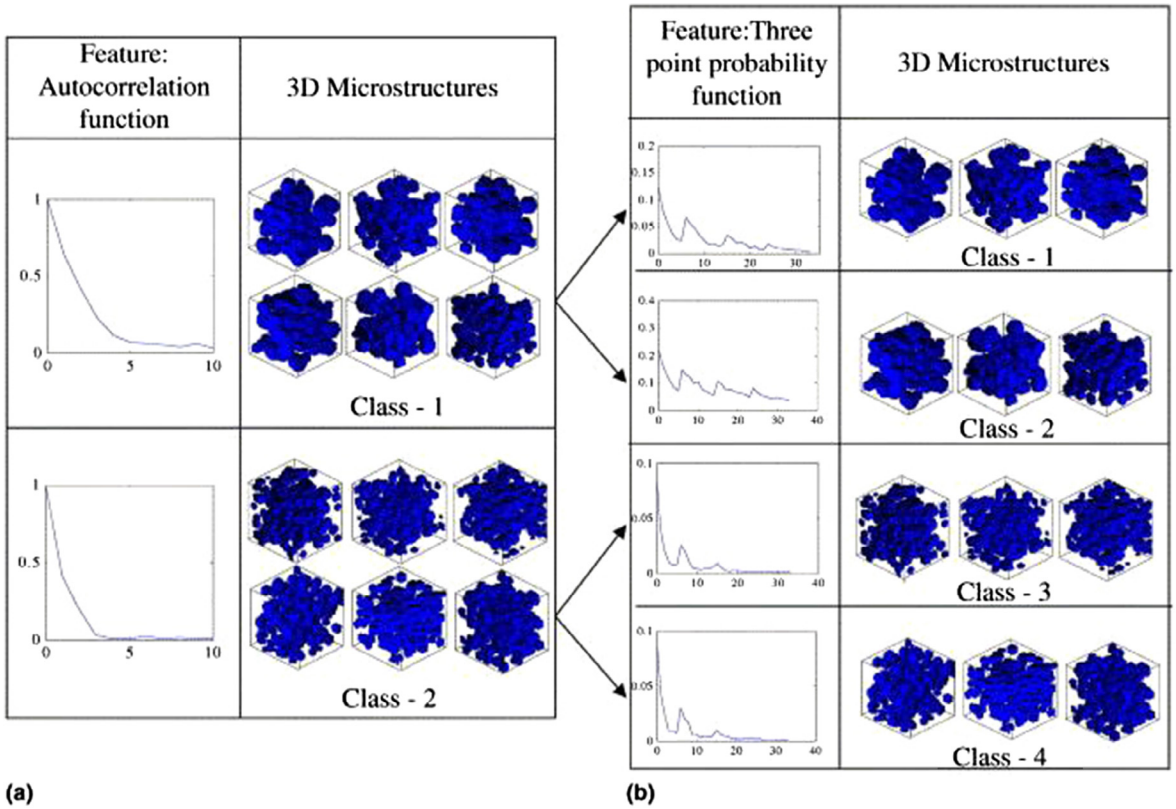

Fig. 19. Hierarchical microstructure classification: (a) Classification based on $S_2(r)$, and (b) classification based on $S_3(r)$. Image is reproduced with kind permission from [15].

The reconstruction procedure in instance-based learning is essentially a sampling process where only the stored instances (i.e., microstructures) in the database can be realized. This sampling strategy limits the application of this approach to the cases where a huge database of microstructure samples with great coverage is available.

3.4.3. Supervised learning

Recently, Bostanabad et al. [90] developed a versatile approach for MCR based on supervised learning for general material systems. In this approach, characterization involves fitting an appropriate supervised learner to the original microstructure image, $\boldsymbol{X}$, and reconstruction consists of taking realizations from that supervised learner; fitting the model is like coding while reconstruction is like decoding.

3.4.3.1. Characterization. In short, the fundamental idea here is to convert $\boldsymbol{X}$ into a training dataset, $\boldsymbol{D}$, where the phase of each pixel is represented as a function of its surrounding pixel phases (see Fig. 20). Once $\boldsymbol{D}$ is built, a supervised leaning model is fitted to it to learn the conditional probability distribution of any individual pixel's value given its surrounding pixel phases. This fitted model, rather than storing the data (i.e., the image), efficiently summarizes it and can be subsequently employed for fast and efficient reconstruction of statistically equivalent microstructures with arbitrary size.

Realizing $\boldsymbol{X}$ as a random sample from its underlying full joint distribution, $f(\boldsymbol{X})$, reconstruction of a statistically equivalent image $\boldsymbol{Y}$ of any size requires the (implicit) estimation of $f(\boldsymbol{X})$. Bostanabad et al. [90] proposed two methods to characterize $f(\boldsymbol{X})$ and subsequently reconstruct an image: A *non-causal* one based on Gibbs sampling [243] and a *causal* one. As the former approach is expensive and requires careful use of the Gibbs sampler, the latter is preferred [90,113] and discussed hereafter.

To learn $f(\boldsymbol{X})$ for causal reconstruction, one can decompose it as:

$$\begin{aligned} f(\boldsymbol{X}) &= f(X_{11}) f(X_{12}|X_{11}) f(X_{13}|X_{11}, X_{12}) \cdots f(X_{n_1 n_2}|X_{11}, X_{12}, \ldots, X_{n_1(n_2-1)}) \\ &= f(X_{11}|\boldsymbol{X}^{(<11)}) f(X_{12}|\boldsymbol{X}^{(<12)}) f(X_{13}|\boldsymbol{X}^{(<13)}) \cdots f(X_{n_1 n_2}|\boldsymbol{X}^{(<n_1 n_2)}), \end{aligned} \tag{3-18}$$

where $\boldsymbol{X}^{(<ij)}$ denotes the set of all the pixels in $\boldsymbol{X}$ ordered before X_{ij} (assuming a raster-scan order, see Fig. 21). The above decomposition illustrates that given a set of models that approximate the various conditional distributions on the right-hand side (and hence $f(\boldsymbol{X})$), $\boldsymbol{Y}$ can be reconstructed pixel-by-pixel via sampling from those models.

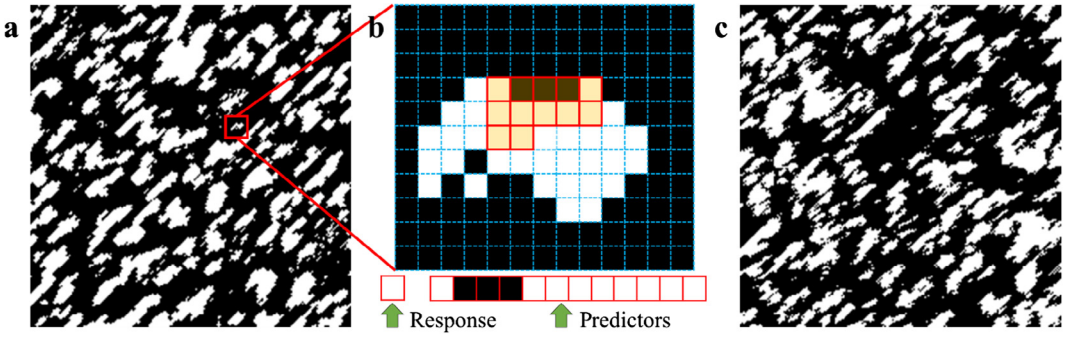

Fig. 20. MCR via supervised learning: (a) The original binary image $\boldsymbol{X}$. The white and black colors represent phase values. (b) The magnified view of a small portion of $\boldsymbol{X}$. The highlighted region (pixels within red boxes) represent an instance of a causal neighborhood with size $w = 2$. These pixels are rearranged into a row vector for building a training dataset. This procedure is repeated for all the pixels in $\boldsymbol{X}$ to convert it to a training dataset. (c) The reconstructed microstructure. The image is reproduced with kind permission from [113].

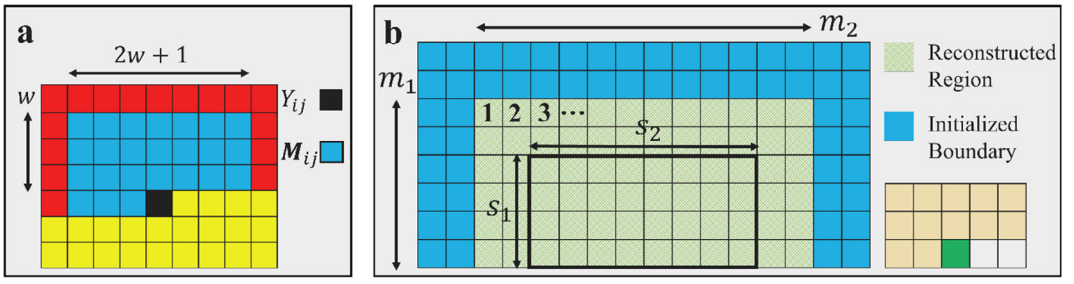

Fig. 21. Neighborhood and boundary effects in MCR via supervised learning: (a) The subset of the pixels indexed before Y_{ij} is the causal neighborhood ($\boldsymbol{M}_{ij}$) of that pixel (here, $\boldsymbol{M}_{ij}$ is of size $w = 3$). The response pixel and those within $\boldsymbol{M}_{ij}$ are color-coded as, respectively, black and blue. The red and blue pixels are indexed before Y_{ij} while the yellow pixels are indexed after Y_{ij}. (b) The green region represents the reconstructed image (from the pseudocode above) and the blue pixels are added to its exterior so as its boundary pixels would not have missing data in their neighborhood. Here, the thickness of the blue region equals the neighborhood size ($w = 2$). The central part of the striped green region of size $s_1 \times s_2$ is chosen as the final reconstructed image.

To make the approximation of $f(\boldsymbol{X})$ tractable and use only a *single* model for estimating all the conditional distributions in Eq. (3-18), Bostanabad et al. [90] assumed that $\boldsymbol{X}$ can be modeled as a form of stationary Markov random field (MRF) with the following properties:

- Locality: $f(X_{ij}|\boldsymbol{X}^{(<ij)}) = f(X_{ij}|\boldsymbol{M}_{ij})$ for a sufficiently large neighborhood $\boldsymbol{M}_{ij}$.
- Stationarity: $f(X_{ij}|\boldsymbol{M}_{ij})$ does not depend on pixel location (i,j).

In the above, $\boldsymbol{M}_{ij}$ contains a sufficiently large group of pixels surrounding one side of X_{ij} (see Fig. 21) and the conditional probability $f(X_{ij}|\boldsymbol{M}_{ij})$ has a Bernoulli distribution with an event probability that depends on the pixel values in $\boldsymbol{M}_{ij}$ (for a binary microstructure image where the phase of interest is represented with 1, e.g., this event is defined as $X_{ij} = 1$). $\boldsymbol{M}_{ij}$ is called the causal [244] neighborhood of X_{ij} because it only includes some of the pixels indexed (in a raster-scan order) before X_{ij}.

The procedures for determining the size of $\boldsymbol{M}$ in a data-driven way via cross-validation (CV) is outlined in [90,113]. Briefly, the optimum neighborhood size is typically on the order of the largest topological feature in $\boldsymbol{X}$ and can be estimated by starting with a relatively large neighborhood and shrinking the size down until the cross-validation (CV) error is minimized. An attractive feature of the supervised learning approach is that it is almost insensitive to the chosen neighborhood size as long as it is not considerably small [90,113].

Once the size of $\boldsymbol{M}$ is chosen and $\boldsymbol{D}$ is built, any off the shelf supervised learning algorithm can be used to fit a model to it. Bostanabad et al. [90,113] used non-parametric classification trees [245–247] because they are particularly well suited for handling categorical variables and very computationally efficient to either fit or make predictions with.

Bostanabad et al. [113] improved the original methodology by (*i*) incorporating user-defined predictors into the supervised learning model, (*ii*) characterizing and reconstructing 3*D* microstructures, and (*iii*) introducing a reduced order model to reduce the size of $\boldsymbol{D}$ and avoid potential computational issues in 3*D* characterization.

MCR with supervised learning is applicable to a wide range of material systems (e.g. isotropic, anisotropic, and porous with high/low volume fraction). In this framework, the characterization has the following features:

- It does not necessitate any preprocessing procedure (e.g., choosing features and then conducting image analysis to calculate their statistics) and requires only one parameter (i.e., the neighborhood size).
- It is fast and only done once (unlike texture synthesis and multiple-point statistics reviewed in Section 3.5). If the size of $\boldsymbol{D}$ is considerably large, the potential computational issues (e.g., storing $\boldsymbol{D}$ or fitting cost) can be addressed via using a reduced order model [113].
- It would benefit from incorporation of user-defined predictors [113] especially those that characterize some physical features such as local V_f or connectivity.
- Cannot be readily used to arrive at the characterization of a 3*D* microstructure given some 2*D* samples as the joint probability distribution in 3*D* cannot be estimated from that in 2*D*.
- Other than some general trends, does not explicitly provide sensible physical features that can be used for material design. That is, the applicability of the model to various material systems comes at a price in that it is not straightforward to use this method and directly study the effect of some features (such as particle distribution on the effective properties) as there is no explicit descriptor or feature in the model.

3.4.3.2. Reconstruction. Once $\boldsymbol{X}$ is used to learn a model with an $\boldsymbol{M}$ of size w, reconstruction of an image with arbitrary size $s_1 \times s_2$ in 2*D* is achieved pixel-by-pixel (voxel in 3*D*) by sampling from that model (see Fig. 21). The following steps summarize the procedure (the steps for 3*D* reconstruction are similar, see [113]):

- Build an initial image $\boldsymbol{Y}$ of size $(m_1 + w) \times (m_2 + 2w)$, with $m_1 > s_1$, and $m_2 > s_2$.
- For $i = w+1, w+2, \ldots, w+m_1$ (row)
  For $j = w+1, w+2, \ldots, w+m_2$ (column)
  a. Use the fitted model to predict the Bernoulli parameter $p_{ij} = f(Y_{ij}|\boldsymbol{M}_{ij})$
  b. Generate $Y_{ij} \sim \text{Beroulli}(p_{ij})$ and update the corresponding pixel in $\boldsymbol{Y}$.
- Pick the central part of $\boldsymbol{Y}$ with size $s_1 \times s_2$ as the reconstructed image.

At the boundary pixels of $\boldsymbol{Y}$ not all the predictors are available in $\boldsymbol{M}$ (see Fig. 21) and so the predictions thereof will be noisy. To address this issue, Bostanabad et al. [90,113] recommend to reconstruct a larger image than required and choose the central part because the noise dies out as one moves away from the boundary regions. The size of the initial image (i.e., the differences $m_1 - s_1$, and $m_2 - s_2$) depends on how its boundaries are initialized (the blue region in Fig. 21). The studies in [90,113] indicate that putting spliced copies of the original 2*D* image on the boundaries perhaps produces the best results where the differences can be set to $(m_1 - s_1, m_2 - s_2) \sim (w, 2w)$.

An appealing feature of reconstructing random media with supervised learning lies in the negligible computational costs; it only takes a few seconds to use the fitted model to reconstruct a batch of samples. Another attractive attribute is that the reconstruction procedure is independent of the material system and does not entail any calibration parameters. The

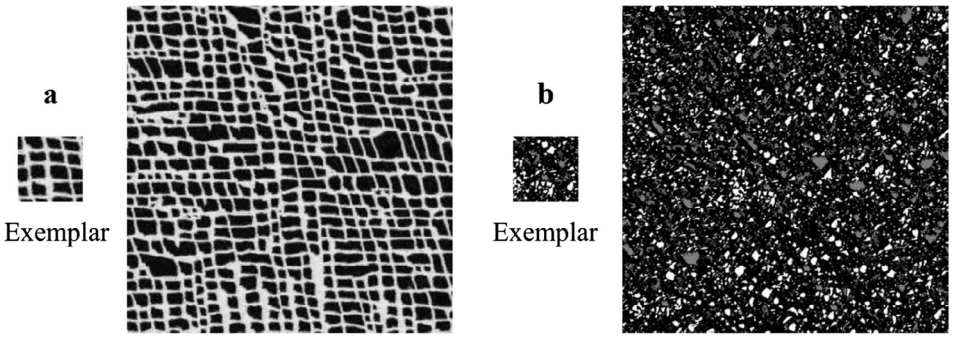

Fig. 22. MCR via texture synthesis: Synthesizing (a) a large binary texture, and (b) a three-phase microstructure from small exemplars. The exemplars are taken from [253] and [107], respectively.

reconstruction quality, of course, depends on how well the supervised learner (*i*) is sampled from during reconstruction, and (*ii*) models the original microstructure. The results in [90,113] illustrate that the procedures summarized above preserves the learned features very well, and so it can be concluded that the degree of statistical equivalency between the original microstructure image and the reconstructed ones (as well as the range of applications) depends on the quality of the fitted model in the characterization step.

3.5. Texture synthesis and multiple-point statistics[3]

Texture synthesis originated in computer graphics and has a wide range of applications including occlusion fill-in, lossy image, video compression, and foreground removal. The basic idea here is to synthesize a (large) texture (i.e., a stationary image) given a small sample of the same texture (see Fig. 22). The early works in [248–250], and particularly that of Popat and Picard [251], established the general framework of texture synthesis, but it was the pioneering and highly recognized works of Efros and Leung [252] and then Wei and Levoy [253] and Efros and Freeman [254] that reformulated the problem of texture synthesis to an *exemplar-based* framework and provided more practical solutions. In short, these works addressed the ill-posed nature of the problem by assuming that the given texture sample comes from a stationary MRF and has the spatial locality property. Such assumptions, as discussed in Section 3.4.3.1, imply that the conditional probability distribution of brightness value of any pixel, given the brightness values of its neighboring pixels, is independent of the rest of the image.

Unlike the supervised learning method introduced in Section 3.4.3, the conditional probability distribution in Eq. (3-18) is not explicitly modeled in texture synthesis and an image is reconstructed by repeatedly querying the given sample. Specifically, the reconstruction is usually done pixel-by-pixel (voxel in 3D) in a specific order (e.g. raster scan), where each pixel's value in the reconstructed image is found by searching for the pixel (or set of pixels) in the original image whose neighboring pixels best match the neighbors of the pixel to be generated. The methods largely differ in their choice of neighborhood geometry, definition of similarity, handling boundaries, and search method. For example, the *basic* procedure described in [253] to use a small exemplar ($\boldsymbol{X}$) of size $m_1 \times m_2$ and synthesize an image ($\boldsymbol{Y}$) of arbitrary size $s_1 \times s_2$ is summarized as follows:

- Build an initial (random) image $\boldsymbol{Y}$ of size $s_1 \times s_2$.
- Determine the size (w) of the L-shaped causal neighborhood (see Fig. 21a).
- For $i = 1, 2, \ldots, s_1$ (row)
  For $j = 1, 2, \ldots, s_2$ (column)
  c. Build $\boldsymbol{M}_{ij}$, the causal neighborhood of the pixel Y_{ij}.
  d. Compare $\boldsymbol{M}_{ij}$ to the causal neighborhood of all the $(m_1 - w) \times (m_2 - 2w)$ non-boundary pixels in $\boldsymbol{X}$.
  e. Choose the pixel in $\boldsymbol{X}$ whose neighborhood best matches $\boldsymbol{M}_{ij}$. Set the value of that pixel to Y_{ij}.

Recognizing the microstructure image as a texture, the above methodology can be readily applied to MCR. Sundararaghavan [255] was perhaps the first to directly apply texture synthesis methods to MCR where he reconstructed 3*D* microstructures using three orthogonal 2*D* sectional images taken along the *x*-, *y*-, and *z*-planes (see Fig. 23). Following the procedures in [253], Liu and Shapiro [107] reconstructed various binary and grayscale microstructures in 2*D*/3*D* and illustrated the effectiveness of texture synthesis methods in preserving statistical equivalency.

Recently, texture synthesis methods have also been successfully applied to the reconstruction of the equilibrium state (i.e., a single image) [256] as well as the temporal evolution (i.e., a series of correlated images) [257] of polycrystalline

[3] In these methods, characterization and reconstruction are best described together so we do not devote separate sections to them.

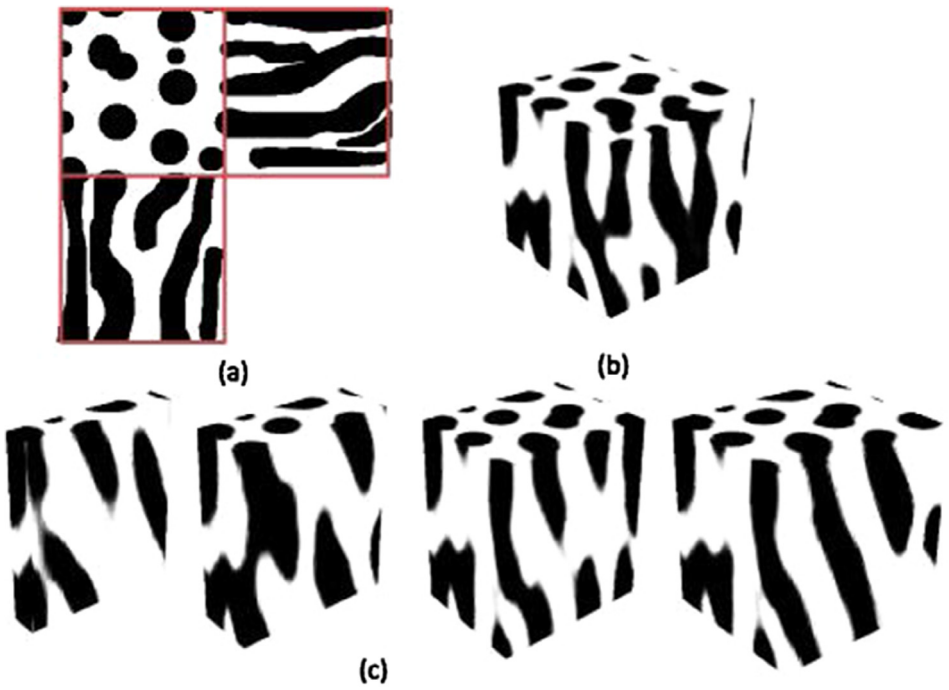

Fig. 23. 2*D* to 3*D* reconstruction with texture synthesis: (a) Three 2*D* sectional images from the *x*-, *y*-, and *z*-planes, (b) reconstructed 3*D* microstructure, (c) 3*D* sectional views of the microstructure. The image is reproduced with kind permission from [255].

microstructures. In the latter case, for example, one can conduct expensive phase field simulations for modeling polycrystalline grain growth [258,259] on a relatively small domain and then use texture synthesis to enlarge the domain (see Fig. 24).

Texture synthesis based methods are applicable to various 2*D*/3*D* material systems (see e.g., Table 1 and Figs. 12 through 14 in [107]) and can also be used to synthesize a 3*D* solid texture from a 2*D* exemplar [107,255,260], see also Fig. 23. They involve no fitted model for characterizing the microstructure and reconstruction is done by exhaustively searching for similar neighborhoods in the original image. This exhaustive search makes these methods computationally prohibitive for microstructures with long-range correlations or high resolution samples where a relatively large neighborhood is required in the search process. We note that, some texture synthesis methods [254] use patches instead of pixels and hence are computationally more efficient (see [261,262] for similar works in geostatistics). Also, use of Gaussian pyramids for multi-resolution synthesis (introduced in Section 3.5.1) is a common practice in texture synthesis.

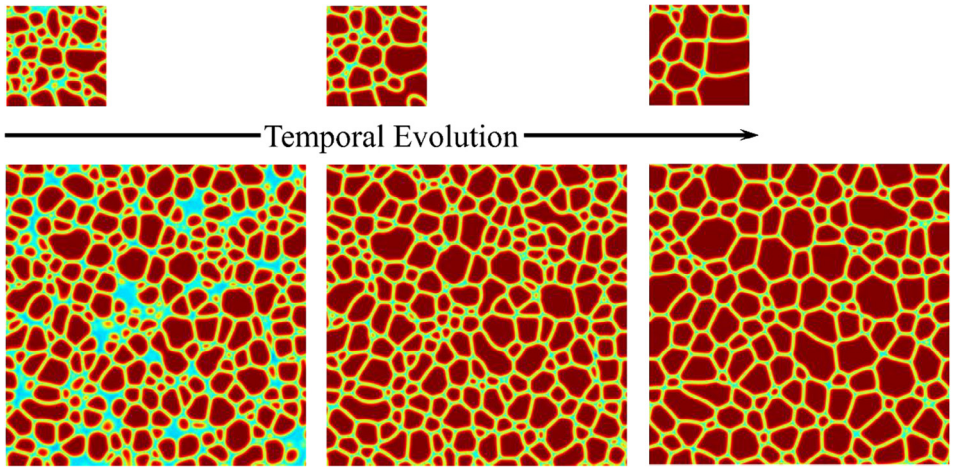

Fig. 24. Modeling polycrystalline grain growth with texture synthesis: The series of images illustrating grain coarsening on a small domain can be obtained via expensive phase field simulations (top row). MCR via texture synthesis can be used to enlarge this domain (i.e., the microstructure size) while preserving the statistics of the microstructure (bottom row).

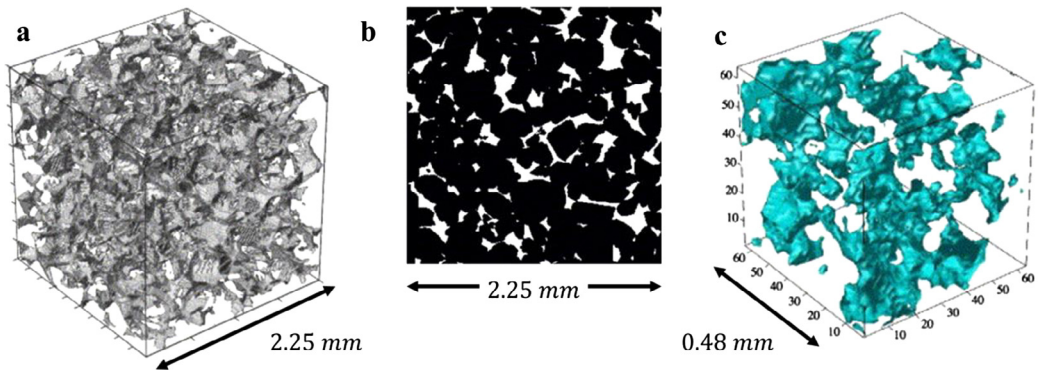

Fig. 25. Microstructure reconstruction with multiple-point statistics: (a) Pore-space illustration in a Fontainebleau sandstone sample, (b) a 2*D* cross-section of the 3*D* sample, and (c) sub-grid of the reconstructed sample (the light blue regions represent the pore space). The images are reproduced with kind permission from [267].

The multiple-point statistics method is very similar to texture synthesis [263] in that a structure is modeled as a stationary MRF and then characterized by calculating and storing the conditional probabilities of finding a specific phase at a pixel, given the phases of a particular configuration of neighboring pixels. Thus, characterization is done implicitly; through exhaustive enumeration of all possible phase combinations for a variety of neighborhood shapes observed in the original sample. Reconstruction, similar to texture synthesis, is accomplished pixel-by-pixel through searching in the stored conditional probabilities for the training neighborhood that best matches that of the pixel being reconstructed, and subsequently sampling from the conditional probability associated with that neighborhood. Different methods primarily vary in the reconstruction order (e.g., random or raster scan), choice of neighborhood geometry, and search method. This method is mainly used in the geostatistics literature [264–270] where the study of porosity in geological structures (such as sandstone) is of particular interest (see Fig. 25).

In texture synthesis and multiple-point statistics methods, perhaps the most important parameter is the neighborhood size. This parameter greatly affects the results [107,252,255–257] and is oftentimes chosen manually via ad hoc methods, e.g., to arrive at visually appealing results or minimize some predefined cost measure. For instance, Liu and Shapiro [107] show that the chosen neighborhood size affects the quality (e.g., the preserved randomness) of the results and recommend to set it to the range after which the two-point correlation function dies out.[4] In addition, Harrison [271] illustrates that the reconstruction order (e.g. a bottom-up raster-scan vs. a top-down raster-scan) usually affects the results in texture synthesis (and so in multiple-point statistics) in that the spatial correlations might not be preserved. Addressing these issues in 3*D* MCR is particularly challenging because on the one hand, 3*D* microstructures are much more complex than 2*D* ones, and on the other hand the associated computational costs are significantly higher. Although applicable to a wide variety of material systems in 2*D* and 3*D*, texture synthesis and multiple points statistics cannot be readily applied to material design or building PSP links because the characterization is implicit and does not explicitly quantify the topological features.

We note that, herein we focused on exemplar-based texture synthesis and multiple-point statistics methods as they provide more practical solutions and are evidently more appealing to the materials science community. Texture synthesis in the computer graphics community has also been done via parametric MRF models [272,273] (where, similar to Section 3.3.1, the model is fitted to the microstructure image, and then used to reconstruct new microstructures) and statistics matching [248,249,274,275] (the statistics being, e.g., the pixels' brightness histogram or distribution of filter responses) approaches. Similarly, in geosciences, variogram-based [276] and parametric [277] approaches have also previously been used to model (spatial) correlations in data.

3.5.1. Multi-resolution reconstruction via Gaussian pyramids

The exhaustive search for similar neighborhoods makes the methods introduced in Section 3.5 computationally demanding. To address this issue (esp. when a large neighborhood is required to capture the long-range correlations), and ensure convergence to a statistically equivalent image, Gaussian[5] pyramids have been used in many works. Essentially, such pyramids are constructed by stacking instances of an image at different resolutions upon one another (see Fig. 26), with the largest (i.e., the one with the highest resolution) and smallest images located at, respectively, the base and top of the pyramid.

One strategy to leverage Gaussian pyramids for speeding up an MCR technique (e.g., texture synthesis or the YT method discussed in Section 3.1.2) is as follows:

[4] Perhaps a better estimate might be made using the SDF because it characterizes a structure more clearly as illustrated in Section 3.3.

[5] Due to simplicity and greater spatial localization, Gaussian pyramids are commonly preferred to other pyramids such as Laplacian, steerable, and feature-based.

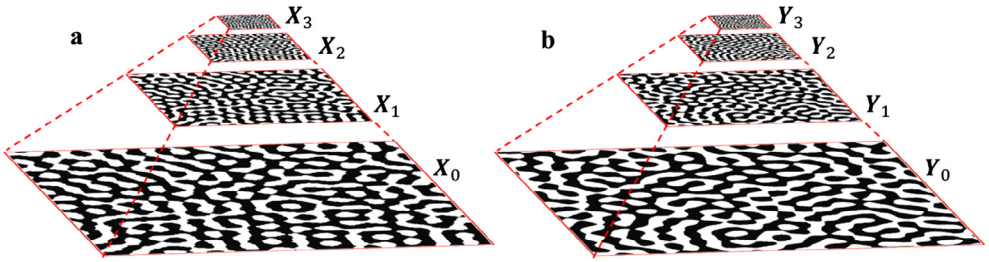

Fig. 26. Gaussian pyramids: The four-level pyramids of the original microstructure image (a) and the reconstructed one (b). The microstructures at the same level (e.g., $\boldsymbol{X}_1$ and $\boldsymbol{Y}_1$) are statistically equivalent.

- Build the pyramid (see Appendix A.2) of the original image, $\boldsymbol{X}$, with n-levels. Denote the image at level i ($i = 0, 1, \ldots, n-1$) with $\boldsymbol{X}_i$.
- Reconstruct a microstructure, $\boldsymbol{Y}_{n-1}$, that is statistically equivalent to $\boldsymbol{X}_{n-1}$.
- For $k = 1, 2, \ldots, n-1$ (go through levels)
  a. Generate $\boldsymbol{Y}_{n-k-1}$ by up-sampling $\boldsymbol{Y}_{n-k}$ (see Appendix A.2).
  b. Optimize $\boldsymbol{Y}_{n-k-1}$ (e.g., via the YT method or texture synthesis) to ensure its statistical equivalency to $\boldsymbol{X}_{n-k-1}$.

This hierarchical process results in a Gaussian pyramid for the reconstructed image where each level can be considered as a statistically equivalent microstructure image to the corresponding level of the original image's Gaussian pyramid (see Fig. 26). Generally, more speed-up is gained by increasing the number of levels, but caution must be taken because if $\boldsymbol{X}_{n-1}$ (the image at the highest pyramid level) does not entail the essential morphological details of $\boldsymbol{X}_0$, either $\boldsymbol{Y}_0$ might not converge (to a statistically equivalent image to $\boldsymbol{X}_0$), or the computational costs might start to increase (due to increased optimization costs in step b).

4. Remarks

Having elaborated on various MCR methods in Section 3, in this section we summarize the key findings and provide some guidance as to which methods can be applied to a particular problem. The discussions are focused around the methods that utilize the characterized information from a given sample to reconstruct *statistically equivalent* microstructures (rather than those which aim to recover the original microstructure). We also comment on the relative *computational cost* and *accuracy* (in terms of ensuring the statistical equivalency between the original and reconstructed microstructures) of each method.

4.1. Statistical functions

MCR via statistical functions provides a very general framework for the characterization and reconstruction of various material systems (such as anisotropic, porous, and multiphase) in 2*D* or 3*D*. Assuming the statistical functions of an unknown 3*D* structure are the same as those of the given 2*D* image(s), this approach can also be used to reconstruct 3*D* structures given a (set of) 2*D* samples.

While multiple functions can be used for characterization, prior research has demonstrated that two-point correlation, lineal-path, and two-point cluster correlation functions (denoted, respectively, with $S_2(r)$, $L(r)$, and $C_2(r)$ for an isotropic material) sufficiently characterize a wide range of material systems. For anisotropic or multi-phase (i.e., more than two phases) microstructures, the selected functions must take into account, respectively, the direction of anisotropy and all the necessary phase-to-phase correlations. See Section 3.1.1 for some detailed discussions on $S_2(r)$, $L(r)$, and $C_2(r)$.

With statistical functions, the reconstruction is usually done via the stochastic optimization (aka Yeong and Torquato or YT) method reviewed in Section 3.1.2. The YT method ensures (by minimizing Eq. (3-4)) that the *selected* statistical functions of the reconstructed sample are similar to those of the original microstructure.

In this framework, the accuracy depends on the choice of the statistical functions (more functions are required to accurately characterize complex microstructures) as well as the efficiency of the reconstruction step in preserving those functions in the reconstructed sample. The overall computational cost, however, mainly depends on the reconstruction step: while the statistical functions of the original sample are only evaluated once during characterization, those of the sample being reconstructed are evaluated many times in reconstruction. Although incorporating more functions in Eq. (3-4) provides more accuracy, this would notably increase the reconstruction cost. Hence, based on the available resources and desired accuracy, a compromise has to be devised. In Section 3.1.2.2, we elaborated on various techniques that can significantly speed up the reconstruction with the YT method.

In the YT method, one has to carefully implement the *SA* algorithm to achieve sufficient accuracy (i.e., minimize E in Eq. (3-4)) with a manageable computational cost. The adjustment procedure might not be straightforward and even with the best parameters the reconstruction cost using this method is generally expected to be high (hours to days, depending on

the problem). The *GD* algorithm, as opposed to *SA*, only has one adjustable input and has been shown to minimize E (in Eq. (3-4)) faster for simple structures that are completely characterized by $S_2(r)$. Further research must be conducted to investigate the efficiency of *GD* for complex microstructures.

Finally, we note that this framework is generally not applicable to building PSP links or materials design because (*i*) the statistical functions do not characterize the microstructure with physically meaningful parameters (even if the functions are approximated with some parametric form, such parameters cannot be easily related to sensible microstructure features), and (*ii*) the YT method is too expensive for an iterative microstructure design process.

4.2. Physical descriptors

Descriptor-based approaches are applicable to the MCR of crystalline and particulate structures. Such methods may also be used when the morphology can be approximated with some configuration of certain descriptors (see, e.g., Fig. 11b). Using the stereological techniques and with some mild assumptions, this method enables reconstruction of 3*D* microstructures given 2*D* samples.

Here, the most important step lies in the choice of the set of descriptors for microstructure characterization. As discussed in Section 3.2, many descriptors can be considered but oftentimes a subset of them sufficiently characterizes a particular structure and so the decision depends on the material system and properties of interest. Prior experience or dimensionality reduction methods such as linear embedding or principal component analysis (see Section 3.2.3) can be used for this purpose. These descriptors (and their distributional characteristics) are either obtained from an available microstructure sample (via image analysis) or defined by the user (e.g., for designing a microstructure). In the latter case, it must be ensured that the chosen values for the set of descriptors correspond to a feasible microstructure.

The reconstruction procedure, though often involving optimization, is not costly and commonly takes minutes to hours depending on the problem. The parameters controlling the reconstruction step (e.g., those of the *SA* algorithm in the optimization step) have to be adjusted for optimal performance and cost but oftentimes this process is more straightforward than that discussed in Section 4.1.

Perhaps the most useful features of this method are on *sensible* and *parametric* characterization which make it straightforward to apply this method to material design. In particular, these features enable the integration of this framework with metamodeling techniques to build processing-structure or structure-property links (see Section 5.1 for an example).

4.3. SDF

SDF provides a systematic approach for characterizing the high-dimensional microstructure morphology in the frequency space. Though essentially equivalent to the two-point correlation function (see Section 3.3.1), microstructure quantification via the SDF better represents the prominent morphological features than the two-point correlation function (see, e.g., Fig. 15). This clear representation of microstructure, makes the parametrization of SDF (with potentially physics-aware parameters) easier.

As described in Section 3.3, with the SDF representation of a microstructure multiple methods (such as disk packing and Cahn's scheme) can be used to reconstruct statistically equivalent microstructures. These methods are fast (taking seconds to minutes depending on the problem), mostly analytical, and (as opposed to the optimization methods discussed above) don't entail much tuning. Such fast reconstruction techniques, along with a parameterized SDF representation, enable the use of SDF-based approaches for design and building processing-structure and structure-property links, esp., in optical applications (see Section 5.2 for an example). In addition, assuming an unknown 3*D* structure possesses the same SDF representation as a given 2*D* image, these reconstruction techniques can reconstruct 3*D* structures given a 2*D* sample.

Finally, we note that the equivalency between the SDF and $S_2(r)$ representations implies that SDF-based methods are applicable to isotropic binary materials that are completely (or to a good degree) characterized with their two-point correlation function.

4.4. Supervised learning

The supervised learning approach introduced in Section 3.4.3 can be applied to various material systems such as multiphase, isotropic, and anisotropic in 2*D* or 3*D*. However, it cannot recover (or reconstruct) a 3*D* structure given some 2*D* samples because the underlying joint distribution of pixels (voxels in 3*D*) cannot be learned from that in 2*D*. In addition, because the characterization step (without any user-defined predictors) does not result in necessarily sensible or physically meaningful parameters, the supervised learning approach is not directly applicable to material design or building PSP links.

Perhaps the most attractive features of this approach lie in its flexibility, simplicity, and speed. These features make the approach significantly efficient for building an ensemble of statistically equivalent samples for a wide range of material systems. The only adjustable parameter here is the neighborhood size which, as opposed to texture synthesis, does not noticeably affect the reconstruction accuracy as long as it is large enough. As for the cost, the characterization is done once and usually takes a few seconds (or minutes for 3*D* structures) while the reconstruction is almost instantaneous in 2*D* and takes seconds in 3*D*. Incorporating user-defined predictors into the characterization step (and using them accordingly in

reconstruction) can significantly improve the efficiency of this approach and more research needs to be conducted to investigate this.

Finally, we note that because the supervised-based learning approach learns (i.e., fits a model to) the underlying joint distribution of the morphology, it can be used to develop a methodology for diagnosing microstructural defects in manufactured products [47].

4.5. Texture synthesis

Although well established in computer graphics, (non-parametric) texture synthesis methods have only been recently applied to MCR. Nevertheless, they have been successfully adopted to a wide range of multi-phase, anisotropic, static, and dynamic (i.e., temporally evolving) material systems. In addition, reconstruction of a 3*D* structure given some 2*D* samples is also achievable with texture synthesis. Unlike other methods discussed in this section, the characterization step cannot be separated from the reconstruction step in texture synthesis as the original microstructure image (rather than some characterized information from it) is constantly used during the reconstruction.

Through multi-resolution reconstruction (see Section 3.5.1) and careful implementation, texture synthesis methods can reconstruct microstructures of moderate size relatively fast. Here, Perhaps the determination of the neighborhood size is the most important procedure. This parameter, as detailed in Section 3.5, greatly affects both the cost and accuracy and should be determined with caution. The texture synthesis methods applied to MCR do not characterize the structure with any physically meaningful parameters and so cannot be used for material design or building PSP links in a straightforward manner.

5. Applications of MCR in materials modeling and design

In this section, we provide two examples on the use of MCR in materials modeling and design. In the first example, physical descriptors are used to create a microstructure dataset which is then used for finding the constitutive law of a class of polymer composites. In the second example, the SDF is used to characterize the structure in nanodielectrics and map it to the system properties (i.e., permittivity and dielectric loss).

5.1. Constitutive law modeling of fibrous composites via physical descriptors

In order to accurately simulate the material properties of composites, it is necessary to perform multiscale FEA to account for the hierarchical material structure spanning multiple length scales. Because for most composites the general form of the constitutive law is unknown, it is necessary to rely on direct numerical simulation (DNS) techniques [278–280] in multiscale FEA. In a two-scale material (see Fig. 2), e.g., DNS requires solving nonlinear problems *simultaneously* at two scales: At each increment in the macroscopic simulation the strain state is prescribed to all the integration (aka Gauss or material) points. These strain states determine the boundary conditions for solving the corresponding nonlinear problems at the microscale where the solution gives the stress for the next iteration of the macroscopic FEM calculation.

While DNS is quite flexible and accurate for a wide range of materials, it is very expensive; rendering DNS-based multiscale design of materials insurmountable. Herein, we demonstrate how MCR via physical descriptors, coupled with machine learning techniques, can be applied to learn the constitutive law of a class of microstructures over a wide range of morphologies and strain states (i.e., boundary conditions). The practical significance would be to notably decrease the computational costs of a macroscale simulation by replacing the finite element simulations at the microscale with a supervised learner that models the material law. This dramatic speed-up, subsequently, would allow a designer to determine which microstructure morphology will result in the desired properties at the macroscale. For a more complete description of the case study, see [281].

The 2*D* material system consists of a matrix filled with identical over-lapping inclusions. The matrix and inclusions are governed by, respectively, Arruda-Boyce [282] and Neo Hookean [283] hyperelastic laws and the goal is to learn the constitutive law of the composite as a function of the morphology over a wide range of boundary conditions (i.e., the applied macroscopic strain field). In hyperelastic materials (see Appendix A.3 for more details), the work is independent of the load path and hence, such materials are completely characterized by the existence of a stored (or strain) energy function, $\bar{W}$, that is a potential for the stress, $\bar{\boldsymbol{S}}$. In particular:

$$\bar{\boldsymbol{S}} = \frac{\partial \bar{W}(\bar{\boldsymbol{\varepsilon}})}{\partial \bar{\boldsymbol{\varepsilon}}}, \tag{5-1}$$

where $\bar{\psi}$ is the stored energy potential, $\bar{\boldsymbol{\varepsilon}}$ is the applied macroscopic Green strain tensor, and the overhead bars of $\bar{W}$ and $\bar{\boldsymbol{S}}$ indicate that the values are homogenized over the microstructure domain.

The physical descriptors chosen to characterize the composite are: volume fraction (V_f), mean distance of nearest neighbors ($\bar{r}_d$), aspect ratio (δ_{asp}), and the number of inclusions (N). Given a feasible set of values for these descriptors, the procedure outlined in [30] (summarized in Section 3.2.2) is used to reconstruct a microstructure sample (see Fig. 27).

While no finite set of physical descriptors may completely characterize such a material system, studies on the uncertainty quantification illustrated [25] that these descriptors were sufficient and the reconstructed microstructures were large

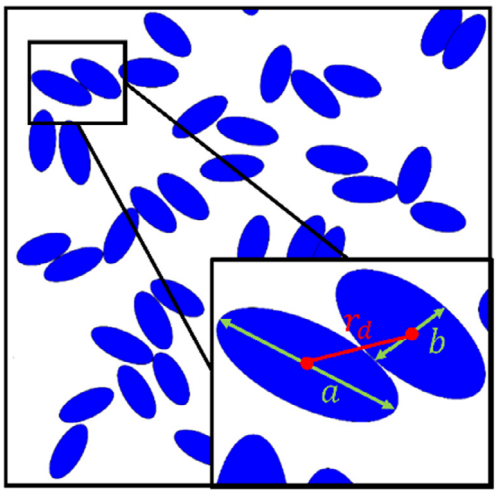

Fig. 27. A sample RVE: The microstructure is large enough to be an RVE and it is reconstructed via the descriptor-based approach discussed in Section 3.2. Here: $V_f = 0.267$, $\bar{r}_d = 33.286, N = 45$, and $\delta_{asp} = 2.249$.

enough to be RVE. The boundary conditions for the microstructure were set to be the macroscopic applied strain, $\bar{\boldsymbol{\varepsilon}} = (\bar{\varepsilon}_{11}\bar{\varepsilon}_{22}\bar{\varepsilon}_{12})^T$. This creates a design space with seven inputs, four physical descriptors and three boundary conditions (see Table 2 for the ranges), and a single response, the homogenized potential density, $\bar{W}$. Ultimately, it is desired to create a model that can quickly and accurately determine the gradient of the response (i.e., the homogenized 2^{nd} Piola Kirchhoff stress, $\bar{\boldsymbol{S}}$, in the macroscale) with respect to the applied boundary conditions as a function of the aforementioned seven descriptors.

To generate the data needed to learn the constitutive law and validate the results, a design of experiments (DoE) of size 10,000 was generated for the seven inputs via Sobol sequence [284,285]. The physical descriptor components of the DoE were next used to reconstruct the corresponding RVEs. The RVEs and the corresponding boundary conditions from the DoE then served as inputs to the DNS, from which the homogenized response was obtained.

Once the dataset was created, supervised learning techniques [286,287] were used to learn the constitutive law. As with most supervised learning models, once fitted, the model is extremely efficient at performing future predictions and can reduce the multiscale FEA costs by replacing the finite element simulations at the microscale. As for the supervised learner, subsets of the first 8000 data points were used to train single layer neural networks as well as polynomial regressors (where all the parameters were determined via CV) while the last 2000 data points were preserved for calculating the average prediction errors for both the potential (e_W) and the stress ($e_{||\boldsymbol{S}||}$):

$$e_W = \frac{100}{N_{samples}} \sum_{i=1}^{N_{samples}} \left(1 - \frac{\bar{W}_i^{predicted}}{\bar{W}_i^{target}}\right)^2 \%, \tag{5-1}$$

and

$$e_{||\boldsymbol{S}||} = \frac{100}{N_{samples}} \sum_{i=1}^{N_{samples}} \frac{(|\bar{\boldsymbol{S}}_i^{target} - \bar{\boldsymbol{S}}_i^{predicted}|_2)^2}{(|\bar{\boldsymbol{S}}_i^{target}|_2)^2} \%, \tag{5-2}$$

Table 2
Descriptors and their ranges: There are four physical descriptors and three strain components characterizing, respectively, the microstructure morphology and boundary conditions.

Descriptor	Range
V_f	$[2, 45]\%$
$\bar{r}_d$	$[0.3, 0.5]$mm
δ_{asp}	$[1, 5]$
N	$[40, 100]$
$\bar{\varepsilon}_{11}$	$[-10, 150]\%$
$\bar{\varepsilon}_{22}$	$[-10, 150]\%$
$\bar{\varepsilon}_{12}$	$[-40, 40]\%$

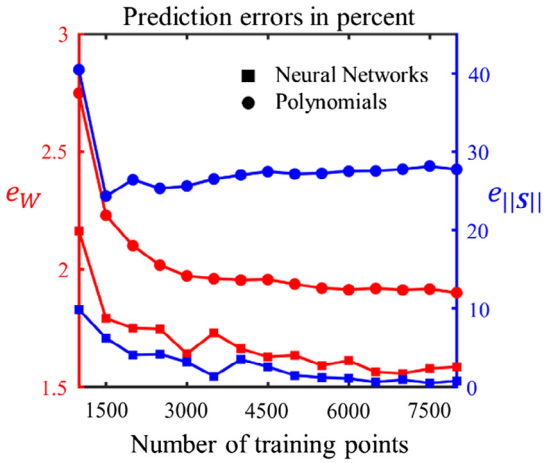

Fig. 28. Convergence studies for the fitted neural networks and the polynomial regressors: As the size of the training dataset increases, the prediction of the homogenized potential (left y-axis) improves, with neural networks outperforming polynomials. The prediction error of the magnitude of the homogenized 2^{nd} Piola Kirchhoff stress (right y-axis), however, decreases consistently only for the neural networks.

where $N_{samples} = 2000$, $\bar{W}_i^{predicted}$ and $\bar{\boldsymbol{S}}_i^{predicted}$ are the predictions from the supervised learner, and $\bar{W}_i^{target}$ and $\bar{\boldsymbol{S}}_i^{target}$ are obtained via FEA. Fig. 28 summarizes the results and illustrates that the fitted neural networks outperform the polynomials and can achieve negligible errors in predicting both the homogenized response (i.e., $\bar{W}$) as well as its gradient (i.e., $\bar{\boldsymbol{S}}$).

This case study demonstrates how MCR can be directly applied to the field of materials design by allowing the designer to quickly and efficiently explore a very large design space using response surface models of structure-property relations that would otherwise be too vast and expensive to explore. The approach we used can be readily generalized and applied to a wide range of materials whose microstructure can be well-defined by a set of physical descriptors and for which an accurate simulation model already exists. Additional applications of this work can include studying the impact of the choice of descriptors, as well as model uncertainty on the response of the microscale RVE, and more importantly, on the macrostructure (e.g., via Markov chain Monte Carlo).

5.2. Elucidation of structure-property links in nanodielectrics via SDF

The spatial dispersion of fillers in nanodielectrics is a critical structural feature that determines the dielectric properties such as permittivity and dielectric loss. The morphology of such materials possesses some special features including: (*i*) low volume fraction, (*ii*) uneven distribution of clusters (heterogeneity), and (*iii*) irregularly-shaped large agglomerates that cannot be modeled using simple geometries (e.g., a disk or ellipse). Although such structures are particulate, the above features affect the performance of the descriptor based approach (detailed in Section 3.2) because the associated reconstruction algorithm generally requires simple particle geometry and relatively even particle distribution. Here, we reconstruct realistic microstructures using the SDF based approach and illustrate the applications of SDF in establishing structure-property relations for microstructure design.

As illustrated in Fig. 29, a narrow range of frequencies in the SDF is a sign of having clustered and unevenly distributed particles in the real space. In contrast, with a broader SDF the particles are expected to be well dispersed. In this study, we choose the gamma distribution [288], governed by the two parameters κ and θ, to parametrize the SDF curves of nanodielectrics. By analyzing all the available samples (collected from several nanodielectric systems with similar polymer dielectric permittivity), it was found that the fitted κ is approximately a constant (1.1) while the fitted θ varies considerably;

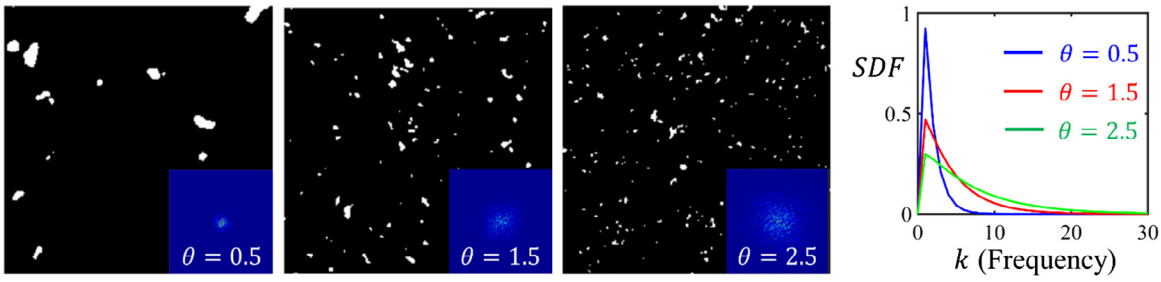

Fig. 29. Characterization of nanodielectrics with SDF: Each SDF (and hence the corresponding microstructure) can be characterized by a single parameter θ. As θ increases, the number of particles and their dispersion increases.

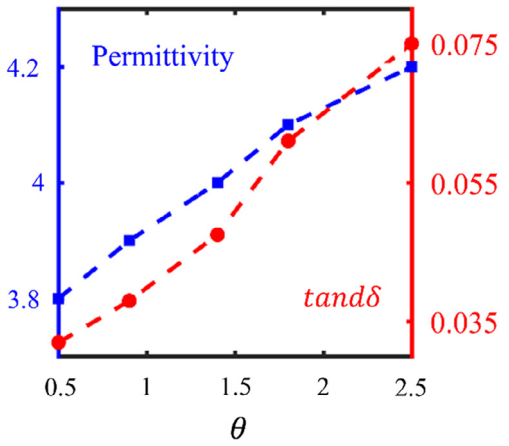

Fig. 30. Relationship between dielectric properties and the parameter of SDF: Linear correlation between permittivity (left y-axis) and dielectric loss ($tan\delta$, right y-axis) with θ.

with a large θ indicating more dispersion. This finding allows to characterize each microstructure by only one single parameter; θ.

To explore the relationship between θ and the dielectric properties, five structures were reconstructed using GRF (Section 3.3.2.1.1) with θ varying from 0.5 to 2.5, κ fixed (to 1.1), and a V_f of 2%. The dielectric permittivity and loss of these microstructures were then calculated via FEA for an epoxy-silica system. The results are summarized in Fig. 30 and indicate that both permittivity and dielectric loss ($tan\delta$) are positively correlated with the dispersion parameter θ. Such a correlation indicates that a trade-off needs to be made because higher permittivity is desired for better storage capabilities while a lower dielectric loss is preferred to reduce energy dissipation. The simple (almost) linear relationship also demonstrates the elegance of SDF: the parameter θ of SDF in the frequency space is a complex combination of various features in the real-space that represents the overall dispersion status of the structure and thus leads to a simple structure-property relationship. With the established relationships, one can optimize the parameter θ to achieve a desired balance between energy storage and dissipation.

6. Conclusion

Accelerated materials design calls for significant and concerted efforts on both experimental and computational aspects of materials science. On the computational side, one avenue of research has been devoted to building processing-structure-property relations for different material systems at various length-scales. We believe that microstructure characterization and reconstruction, coupled with machine learning and computer simulations, provides a systematic framework to generalize existing links or build new ones by statistically going beyond the limits where the data is available. To this end, herein we reviewed the recent progress on this topic in the hope to provide the scientific community with some guidance on how to apply the general MCR framework (see, e.g., Fig. 3) to materials analysis and (inverse) design.

The most common MCR methods (along with their algorithmic details, applicability, and associated challenges and costs) were reviewed in detail in Section 3 where, if possible, the characterization step was separated from the reconstruction step to help the reader better follow the overall procedure of each method. We briefly compared the various methods and summarized their advantages and disadvantages in Section 4. Two examples on the applications of MCR in materials design and modeling were also demonstrated in Section 5.

We believe the numerous works reviewed herein, though not perfectly comprehensive, clearly demonstrate the significance and the potentials of the role that MCR plays in materials science. More research and in-depth studies must be conducted to further develop the various MCR methods (esp. the recently developed ones based on machine learning and SDF) and seamlessly integrate them with computer simulations (such as FEA) and data-mining techniques to tackle the state-of-the-art materials design challenges including building sensible processing-structure links in additive manufacturing, constitutive modeling of composites, uncertainty quantification and propagation in multiscale materials design, Bayesian calibration and bias correction of phenomenological material models, and so on.

Acknowledgement

This work is supported by the U.S. Department of Commerce, National Institute of Standards and Technology as part of the Center for Hierarchical Materials Design (CHiMaD) award 70NANB14H012, National Science Foundation Award No. CMMI-1265709, and the Air Force Office of Scientific Research (AFOSR) Award No. FA9550-12-1-0458. The authors would also like to thank Stephen Lin for conducting the phase field simulations for Fig. 24.

Conflict of interest

The authors disclose no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations that could inappropriately influence, or be perceived to influence, their work.

Author Contributions

W.C. and R.B. initiated the review process. R.B. gathered the relevant literature and categorized the methods. All the authors determined the structure of the paper and contributed to the discussions. R.B. wrote the paper with help from Y.Z. (in Sections 3.3 and 5.2), X.L. (in Sections 3.2.3, 3.4.1, and 3.4.2), and T.K. (in Section 5.1). All the authors have read the manuscript and actively participated in the entire revision process.

Appendix A

A.1. Morphological operations: dilation and erosion

Morphological operations apply a structuring element object (SEO) to an input image (2D or 3D) to create an output image with the same size. In such an operation, the value of each pixel (voxel in 3D) in the output image is determined by comparing the corresponding pixel in the input image with an appropriately defined neighborhood of its surrounding pixels. The shape and size of the neighborhood is determined by the SEO and can make the morphological operation sensitive to specific shapes in the input image.

Dilation and erosion are the most basic morphological operations. They are available in almost all commercial software and take fractions of a second to execute. While dilation adds pixels (voxels in 3D) to the boundaries of the foreground (objects) in an image, erosion removes pixels from the object boundaries (see Fig. 31). More specifically, in dilation (erosion) the value of the output pixel is the *maximum* (*minimum*) value of all the pixels in the input pixel's neighborhood. In dilating (eroding) a *binary* image, if any of the pixels in the input image is 1 (0), the corresponding pixel in the output image will be 1 (0).

A.2. Gaussian pyramid for image compression

The Gaussian pyramid [244] (aka, hierarchical discrete correlation [143]) is commonly used in image analysis to decorrelate the neighboring pixels (voxel in 3*D*) and arrive at a compact representation of the image (see [289] for a review on image compression techniques). The compact representation of an image allows to, e.g., enhance the salient image features (such as edges and foreground objects) and reduce storage space or sampling rate.

Let $\boldsymbol{X}_0$ be the original image and $\boldsymbol{X}_1$ be the reduced version of it after low-pass filtering (which is done by convolving the image with a family of local and symmetric weighing functions). Define $\boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ similarly. The sequence of the images $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, as visualized for a 2*D* image in Fig. 26, resembles a pyramid where $\boldsymbol{X}_1$ and $\boldsymbol{X}_n$ are located, respectively, at the base and top of the pyramid. If the filtering function approximates a Gaussian distribution function, the resulting pyramid is called the Gaussian pyramid.

Although there are various image compression techniques available, the Gaussian pyramid has attracted much attention in MCR literature because, as described in Section 3.5, it reduces the overall computational costs through hierarchical reconstructions. Below, we demonstrate how a Gaussian pyramid can be built in either a bottom-up or top-down procedure for a 1*D* signal. The procedure is similar in 2*D* and 3*D*.

Let $w(u)$ be a discrete weighing function defined at integer values and non-zero only for $-m \leqslant u \leqslant m$. $w(u)$ is constrained by the following equations [143]:

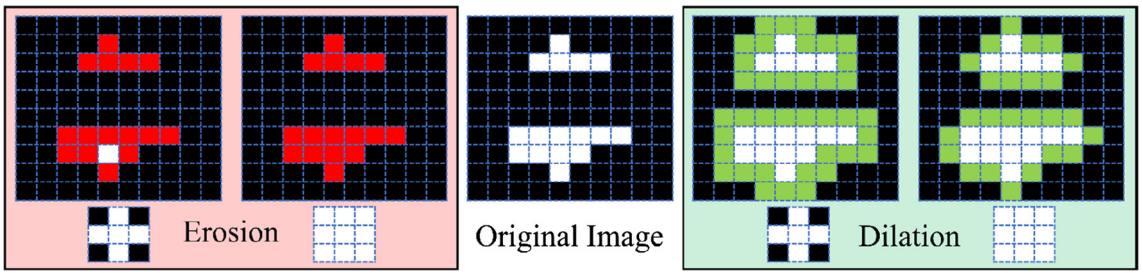

Fig. 31. Pedagogical illustration of dilation (right panel) and erosion (left panel) with two different binary SEOs on a binary image (black background and white foreground): Through dilation (erosion) with each SEO, the green (red) pixels will be added to (removed from) the foreground in the original image. The SEOs are shown below each image.

$$\text{Normalization}: \sum_{i=-m}^{m} w(i) = 1, \tag{A-1}$$

$$\text{Symmetry}: \; w(u) = w(-u) \text{ for all } u, \tag{A-2}$$

$$\text{Unimodal}: \; w(u_1) \geqslant w(u_2) \geqslant 0 \text{ for } 0 \leqslant u_1 \leqslant u_2, \tag{A-3}$$

$$\text{Equal contribution}: \sum_{i=-m}^{m} w(j+ir) = \frac{1}{r} \text{ for all } 0 \leqslant j < r, \tag{A-4}$$

where r is the sampling order (e.g., to reduce the number of nodes across the Gaussian levels to half or one-third, respectively, r would be 2 and 3). Eq. (7-4) stipulates that all the nodes at any particular level must contribute the same total weight to the nodes at the next higher level. If $m = 2$, e.g., the above constraints are satisfied when $w(0) = a$, $w(-1) = w(1) = \frac{1}{4}$, and $w(-2) = w(2) = \frac{1-2a}{4}$ with $\frac{1}{4} \leqslant a \leqslant \frac{1}{2}$. Having defined the weighing function, the value of node i at pyramid level n, $\boldsymbol{X}_{n,i}$, is then calculated as:

$$\boldsymbol{X}_{n,i} = \sum_{u=-m}^{m} w(u)\boldsymbol{X}_{n-1,ri+u}, \tag{A-5}$$

Eq. (7-5) is repeatedly used n times to build the Gaussian pyramid. This procedure of reducing (down-sampling or coarsening) a signal for one level is schematically illustrated in Fig. 32a. To up-sample or expand a signal (i.e., to do the opposite of the above procedure):

$$\boldsymbol{Y}_{n-1,i} = r^D \sum_{u=-m}^{m} w(u)\boldsymbol{Y}_{n,\frac{i-m}{r}}, \tag{A-6}$$

where D is the dimensionality of the signal (here $D = 1$) and only the terms for which $\frac{i-m}{r}$ is an integer are included in the summation. Expansion of a 1D signal for one level is illustrated in Fig. 32b.

Fig. 33 illustrates how the above procedure can be used in 3D to build a three-level Gaussian pyramid (i.e., by coarsening the original image twice). As it can be observed, each level of the pyramid contains less morphological details compared to its lower levels.

A.3. Hyperelastic materials

Hyperelastic materials provide a natural framework for the frame-invariant formulation of anisotropic material response by simply embodying the anisotropy in the potential, W. A consequence of the existence of a stored energy function is that the work done on a hyperelastic material is independent of the deformation path. This behavior is approximately observed in many rubber-like materials [283].

The energy density function of the Arruda-Boyce model, $\psi_{AB}(\boldsymbol{C})$, depends on three polymer properties (the initial bulk modulus K_0, the initial shear modulus μ_0, and the stretch at which the polymer chain network locks λ_m) and is given by:

$$\psi_{AB}(\boldsymbol{C}) = \mu \sum_{i=1}^{5} \alpha_i \beta^{i-1} (\hat{I}_1^i - 3^i) + \frac{K}{2}\left(\frac{J^2-1}{2} - \ln(J)\right), \tag{A-6}$$

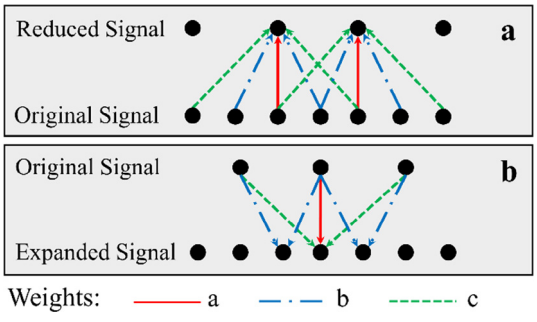

Fig. 32. Gaussian pyramid: 1D illustration of (a) reducing and (b) expanding a signal (with sampling order, r, of 2) to build a Gaussian pyramid. Only two levels of the Gaussian pyramids are shown where each row of nodes represents a level in the pyramid. In reduction (expansion), the value of each node at higher (lower) levels is obtained through a weighted sum of the node values at the lower (higher) level. Here, the node spacing doubles across the levels while the same filtering function would be used across the levels.

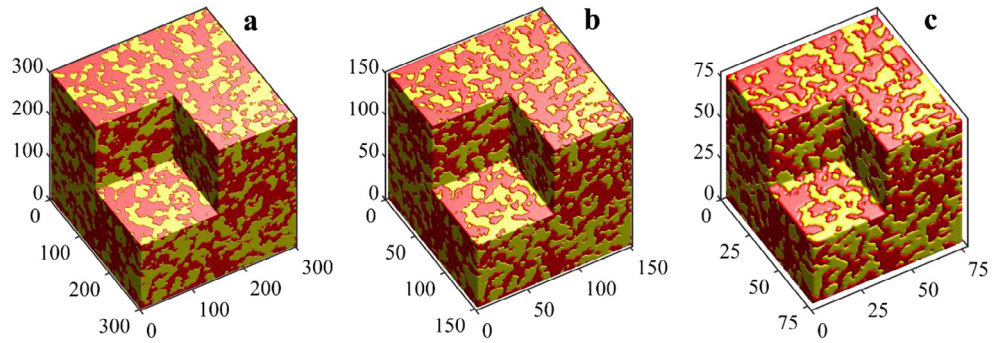

Fig. 33. Gaussian pyramid in 3D: By repeatedly applying Eq. (7-5) with $r = 2$ to the original image in (a), images with half (b) and one-fourth (c) resolution can be constructed. The numbers stand for pixel indices and the images in (b) and (c) are magnified, respectively, twice and four times to better illustrate the effect of coarsening.

where $\beta = \frac{1}{\lambda_m^2}$, $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{1}{20}$, $\alpha_3 = \frac{11}{1050}$, $\alpha_4 = \frac{19}{7000}$, and $\alpha_5 = \frac{519}{673{,}750}$ are all obtained via the first five terms of the inverse Langevin function, J is the determinant of the deformation gradient, $\hat{I}_1 = I_1 J^{-2/3}$ depends on the first invariant of the right Cauchy-Green deformation tensor (i.e., $I_1 = Tr(\boldsymbol{C})$), and μ is calculated from the material properties λ_m and μ_0:

$$\mu = \frac{\mu_0}{2}\left(\sum_{i=1}^{5} i\alpha_i \beta^{i-1} \hat{I}_1^{i-1}\right)^{-1}, \tag{A-6}$$

In the example of Section 5.1, the matrix material properties are chosen as follows:

$$K_0^{mat} = 800 \text{ MPa}, \quad \mu_0^{mat} = 180.5 \text{ MPa}, \quad \lambda_m^{mat} = 2.8$$

The energy density function of the Neo-Hookean model, $\psi_{NK}(\boldsymbol{C})$, is defined as [283]:

$$\psi_{NK}(\boldsymbol{C}) = \frac{\mu_0}{2}(\hat{I}_1 - 3) + \frac{K_0}{2}(J-1), \tag{A-7}$$

and the chosen values for the example in Section 5.1 read $K_0^{par} = 4.0$ GPa and $\mu_0^{par} = 1.9$ GPa.

References

[1] Ward C. Materials genome initiative for global competitiveness. In: 23rd Advanced aerospace materials and processes (AeroMat) conference and exposition. ASM; 2012.
[2] McDowell DL, Kalidindi SR. The materials innovation ecosystem: a key enabler for the materials genome initiative. MRS Bull 2016;41(4):326–35.
[3] Holdren JP. Materials genome initiative for global competitiveness. Washington, USA: National Science and Technology Council OSTP; 2011.
[4] Olson GB. Preface to the viewpoint set on: the materials genome. Scr Mater 2014;70(70):1–2.
[5] Olson GB. Computational design of hierarchically structured materials. Science 1997;277(5330):1237–42.
[6] Olson GB. Designing a new material world. Science 2000;288(5468):993.
[7] Fullwood DT et al. Microstructure sensitive design for performance optimization. Prog Mater Sci 2010;55(6):477–562.
[8] Committee on Integrated Computational Materials Engineering NRC. Integrated computational materials engineering: a transformational discipline for improved competitiveness and national security. National Academies Press; 2008.
[9] Torquato S. Random heterogeneous materials: microstructure and macroscopic properties, vol. 16. Springer; 2002.
[10] Kumar H, Briant CL, Curtin WA. Using microstructure reconstruction to model mechanical behavior in complex microstructures. Mech Mater 2006;38 (8–10):818–32.
[11] Agrawal A, Choudhary A. Perspective: materials informatics and big data: realization of the "fourth paradigm" of science in materials science. APL Mater 2016;4(5):053208.
[12] Kalidindi SR, Medford AJ, McDowell DL. Vision for data and informatics in the future materials innovation ecosystem. JOM 2016;68(8):2126–37.
[13] Curtarolo S et al. The high-throughput highway to computational materials design. Nat Mater 2013;12(3):191–201.
[14] Bhadeshia HH. Neural networks in materials science. ISIJ Int 1999;39(10):966–79.
[15] Sundararaghavan V, Zabaras N. Classification and reconstruction of three-dimensional microstructures using support vector machines. Comput Mater Sci 2005;32(2):223–39.
[16] Pilania G et al. Accelerating materials property predictions using machine learning. Sci Rep 2013;3:2810.
[17] Liu R et al. A predictive machine learning approach for microstructure optimization and materials design. Sci Rep 2015;5:11551.
[18] Xue D et al. Accelerated search for materials with targeted properties by adaptive design. Nat Commun 2016;7:11241.
[19] Matouš K et al. A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. J Comput Phys 2017;330:192–220.
[20] Qomi MJA et al. Combinatorial molecular optimization of cement hydrates. Nat Commun 2014;5.
[21] Masoumi S, Valipour H, Abdolhosseini Qomi MJ. Intermolecular forces between nanolayers of crystalline calcium-silicate-hydrates in aqueous medium. J Phys Chem C 2017;121(10):5565–72.
[22] Zhou Y et al. The contribution of propagons and diffusions in heat transport through calcium-silicate-hydrates. Appl Phys Lett 2017;110(4):043104.
[23] Greene MS et al. Computational uncertainty analysis in multiresolution materials via stochastic constitutive theory. Comput Methods Appl Mech Eng 2011;200(1–4):309–25.
[24] Kalinin SV, Sumpter BG, Archibald RK. Big-deep-smart data in imaging for guiding materials design. Nat Mater 2015;14(10):973–80.
[25] Bessa MA et al. A framework for data-driven analysis of materials under uncertainty: countering the curse of dimensionality. Comput Methods Appl Mech Eng 2017;320:633–67.

[26] Rahimi-Aghdam S, Bažant ZP, Abdolhosseini Qomi MJ. Cement hydration from hours to centuries controlled by diffusion through barrier shells of C-S-H. J Mech Phys Solids 2017;99:211–24.
[27] Salviato M, Esna Ashari S, Cusatis G. Spectral stiffness microplane model for damage and fracture of textile composites. Compos Struct 2016;137:170–84.
[28] Bažant Zdeněk P, Rahimi-Aghdam S. Diffusion-controlled and creep-mitigated ASR damage via microplane model. I: Mass concrete. J Eng Mech 2016;143(2):04016108.
[29] Breneman CM et al. Stalking the materials genome: a data-driven approach to the virtual design of nanostructured polymers. Adv Funct Mater 2013;23(46):5746–52.
[30] Xu HY et al. A descriptor-based design methodology for developing heterogeneous microstructural materials system. J Mech Des 2014;136 (5):051007.
[31] Wang C et al. Highly efficient light-trapping structure design inspired by natural evolution. Sci Rep 2013;3.
[32] Yu SC et al. Characterization and design of functional quasi-random nanostructured materials using spectral density function. In: Proceedings of the ASME international design engineering technical conferences and computers and information in engineering conference, 2016, vol. 2b; 2016. p. 135–45.
[33] Lee W-K, Yu S, Engel CJ, Reese T, Rhee D, Chen W, et al. Concurrent design of quasi-random photonic nanostructures. Proc Natl Acad Sci 2017;114 (33):8734–9. https://doi.org/10.1073/pnas.1704711114.
[34] Salvo L et al. X-ray micro-tomography an attractive characterisation technique in materials science. Nucl Instrum Methods Phys Res Sect B-Beam Interact Mater Atoms 2003;200:273–86.
[35] Requena G et al. 3D-Quantification of the distribution of continuous fibres in unidirectionally reinforced composites. Compos Part A-Appl Sci Manuf 2009;40(2):152–63.
[36] Li DS. Review of structure representation and reconstruction on mesoscale and microscale. JOM 2014;66(3):444–54.
[37] Jiao Y, Stillinger FH, Torquato S. Modeling heterogeneous materials via two-point correlation functions. II. Algorithmic details and applications. Phys Rev E Stat Nonlin Soft Matter Phys 2008;77(3 Pt 1):031135.
[38] Tang T et al. A pixel selection rule based on the number of different-phase neighbours for the simulated annealing reconstruction of sandstone microstructure. J Microsc 2009;234(3):262–8.
[39] Yeong C, Torquato S. Reconstructing random media. Phys Rev E 1998;57(1):495.
[40] Rozman MG, Utz M. Uniqueness of reconstruction of multiphase morphologies from two-point correlation functions. Phys Rev Lett 2002;89 (13):135501.
[41] Torquato S. Necessary conditions on realizable two-point correlation functions of random media. Ind Eng Chem Res 2006;45(21):6923–8.
[42] Jiao Y, Stillinger FH, Torquato S. Modeling heterogeneous materials via two-point correlation functions: basic principles. Phys Rev E Stat Nonlin Soft Matter Phys 2007;76(3 Pt 1):031110.
[43] Bogdanor MJ, Oskay C, Clay SB. Multiscale modeling of failure in composites under model parameter uncertainty. Comput Mech 2015;56(3):389–404.
[44] Greene MS et al. A generalized uncertainty propagation criterion from benchmark studies of microstructured material systems. Comput Methods Appl Mech Eng 2013;254:271–91.
[45] Kennedy MC, O'Hagan A. Bayesian calibration of computer models. J R Stat Soc: Ser B (Stat Methodol) 2001;63(3):425–64.
[46] Lee T-R et al. Quantifying uncertainties in the microvascular transport of nanoparticles. Biomech Model Mechanobiol 2014;13(3):515–26.
[47] Bui AT, Apley DW. A monitoring and diagnostic approach for stochastic textured surfaces. Technometrics 2017:1–13.
[48] Chau VT et al. The enigma of large-scale permeability of gas shale: pre-existing or frac-induced? J Appl Mech 2017;84(6). p. 061008-061008-11.
[49] Esna Ashari S, Mohammadi S. Fracture analysis of FRP-reinforced beams by orthotropic XFEM. J Compos Mater 2011. p. 0021998311418702.
[50] Saltelli A et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput Phys Commun 2010;181(2):259–70.
[51] Nagel JB, Sudret B. A unified framework for multilevel uncertainty quantification in Bayesian inverse problems. Probab Eng Mech 2016;43:68–84.
[52] Mehrez L, Fish J, Aitharaju V, Rodgers WR, Ghanem R. A PCE-based multiscale framework for the characterization of uncertainties in complex systems. Comput Mech 2017. https://doi.org/10.1007/s00466-017-1502-4.
[53] Debye P, Bueche AM. Scattering by an inhomogeneous solid. J Appl Phys 1949;20(6):518–25.
[54] Debye P, Anderson Jr H, Brumberger H. Scattering by an inhomogeneous solid. II. The correlation function and its application. J Appl Phys 1957;28 (6):679–83.
[55] Corson PB. Correlation functions for predicting properties of heterogeneous materials. II. Empirical construction of spatial correlation functions for two-phase solids. J Appl Phys 1974;45(7):3165–70.
[56] Corson PB. Correlation functions for predicting properties of heterogeneous materials. I. Experimental measurement of spatial correlation functions in multiphase solids. J Appl Phys 1974;45(7):3159–64.
[57] Corson PB. Correlation functions for predicting properties of heterogeneous materials. III. Effective elastic moduli of two-phase solids. J Appl Phys 1974;45(7):3171–9.
[58] Corson PB. Correlation functions for predicting properties of heterogeneous materials. IV. Effective thermal conductivity of two-phase solids. J Appl Phys 1974;45(7):3180–2.
[59] Joshi MY. A class of stochastic models for porous media. University of Kansas; 1974.
[60] Quiblier JA. A new three-dimensional modeling technique for studying porous media. J Colloid Interface Sci 1984;98(1):84–102.
[61] Hazlett RD. Statistical characterization and stochastic modeling of pore networks in relation to fluid flow. Math Geol 1997;29(6):801–22.
[62] Rintoul MD, Torquato S. Reconstruction of the structure of dispersions. J Colloid Interface Sci 1997;186(2):467–76.
[63] Yeong C, Torquato S. Reconstructing random media. II. Three-dimensional media from two-dimensional cuts. Phys Rev E 1998;58(1):224.
[64] Beran M. Statistical continuum theories. Trans Soc Rheol (1957–1977) 1965;9(1):339–55.
[65] Drugan WJ, Willis JR. A micromechanics-based nonlocal constitutive equation and estimates of representative volume element size for elastic composites. J Mech Phys Solids 1996;44(4):497–524.
[66] Šejnoha M, Zeman J. Micromechanical analysis of random composites. Citeseer; 2002.
[67] Niezgoda SR et al. Optimized structure based representative volume element sets reflecting the ensemble-averaged 2-point statistics. Acta Mater 2010;58(13):4432–45.
[68] Niezgoda SR, Yabansu YC, Kalidindi SR. Understanding and visualizing microstructure and microstructure variance as a stochastic process. Acta Mater 2011;59(16):6387–400.
[69] Gallager RG. Stochastic processes: theory for applications. Cambridge University Press; 2013.
[70] Quintanilla J. Microstructure and properties of random heterogeneous materials: a review of theoretical results. Polym Eng Sci 1999;39(3):559–85.
[71] Hill R. Elastic properties of reinforced solids – some theoretical principles. J Mech Phys Solids 1963;11(5):357–72.
[72] Rezakhani R, Cusatis G. Asymptotic expansion homogenization of discrete fine-scale models with rotational degrees of freedom for the simulation of quasi-brittle materials. J Mech Phys Solids 2016;88:320–45.
[73] Ostoja-Starzewski M. Microstructural randomness and scaling in mechanics of materials. CRC Press; 2007.
[74] Willis JR. Variational and related methods for the overall properties of composites. Adv Appl Mech 1981;21:1–78.
[75] Kanit T et al. Determination of the size of the representative volume element for random composites: statistical and numerical approach. Int J Solids Struct 2003;40(13–14):3647–79.
[76] Huet C. Coupled size and boundary-condition effects in viscoelastic heterogeneous and composite bodies. Mech Mater 1999;31(12):787–829.
[77] Hazanov S. On apparent properties of nonlinear heterogeneous bodies smaller than the representative volume. Acta Mech 1999;134(3–4):123–34.

[78] Rollett AD et al. Three-dimensional characterization of microstructure by electron back-scatter diffraction. Annu Rev Mater Res 2007;37:627–58.
[79] Rezakhani R, Zhou XW, Cusatis G. Adaptive multiscale homogenization of the lattice discrete particle model for the analysis of damage and fracture in concrete. Int J Solids Struct 2017;125:50–67.
[80] Li W et al. A multiscale framework for the simulation of the anisotropic mechanical behavior of shale. Int J Numer Anal Meth Geomech 2017.
[81] Rahimi-Aghdam S, Bažant ZP, Caner FC. Diffusion-controlled and creep-mitigated ASR damage via microplane model. II: material degradation, drying, and verification. J Eng Mech 2017;143(2).
[82] Salviato M et al. Experimental and numerical investigation of intra-laminar energy dissipation and size effect in two-dimensional textile composites. Compos Sci Technol 2016;135:67–75.
[83] Xu HY et al. Descriptor-based methodology for statistical characterization and 3D reconstruction of microstructural materials. Comput Mater Sci 2014;85:206–16.
[84] Otsu N. A threshold selection method from gray-level histograms. Automatica 1975;11(285–296):23–7.
[85] Torquato S, Beasley J, Chiew Y. Two-point cluster function for continuum percolation. J Chem Phys 1988;88(10):6540–7.
[86] Lu B, Torquato S. Lineal-path function for random heterogeneous materials. Phys Rev A 1992;45(2):922.
[87] Torquato S, Lu B. Chord-length distribution function for two-phase random media. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics 1993;47(4):2950–3.
[88] Jiao Y, Stillinger FH, Torquato S. A superior descriptor of random textures and its predictive capacity. Proc Natl Acad Sci USA 2009;106(42):17634–9.
[89] Pant LM, Mitra SK, Secanell M. Stochastic reconstruction using multiple correlation functions with different-phase-neighbor-based pixel selection. Phys Rev E 2014;90(2):023306.
[90] Bostanabad R et al. Stochastic microstructure characterization and reconstruction via supervised learning. Acta Mater 2016;103:89–102.
[91] Torquato S. Statistical description of microstructures. Annu Rev Mater Res 2002;32(1):77–111.
[92] Hanisch KH, Konig D, Stoyan D. The pair correlation function for point and fibre systems and its stereological determination by planar sections. J Microsc 1985;140(Pt 3):361–70.
[93] Yang S, Tewari A, Gokhale AM. Modeling of non-uniform spatial arrangement of fibers in a ceramic matrix composite. Acta Mater 1997;45(7):3059–69.
[94] Torquato S, Stell G. Microstructure of two-phase random media. I. The n-point probability functions. J Chem Phys 1982;77(4):2071–7.
[95] Berryman JG. Measurement of spatial correlation-functions using image-processing techniques. J Appl Phys 1985;57(7):2374–84.
[96] Fullwood DT et al. Gradient-based microstructure reconstructions from distributions using fast Fourier transforms. Mater Sci Eng A-Struct Mater Prop Microstruct Process 2008;494(1–2):68–72.
[97] Fullwood DT, Niezgoda SR, Kalidindi SR. Microstructure reconstructions from 2-point statistics using phase-recovery algorithms. Acta Mater 2008;56(5):942–8.
[98] Szapudi I. Introduction to higher order spatial statistics in cosmology. In: Data analysis in cosmology. Springer; 2009. p. 457–92.
[99] Marin FA et al. Modeling the galaxy three-point correlation function. Astrophys J 2008;672(2):849–60.
[100] Beran M. Use of the vibrational approach to determine bounds for the effective permittivity in random media. Il Nuovo Cimento 1965;38(2):771–82.
[101] Beran M, Molyneux J. Use of classical variational principles to determine bounds for the effective bulk modulus in heterogeneous media. Q Appl Math 1966:107–18.
[102] Drake JM, Klafter J, Levitz P. Chemical and biological microstructures as probed by dynamic processes. Science 1991;251(5001):1574–9.
[103] Robyr P, Bowtell R. Measuring Patterson functions of inhomogeneous liquids using the nuclear dipolar field. J Chem Phys 1997;107(3):702–6.
[104] Shepp L. On positive-definite functions associated with certain stochastic processes Technical report. Murray Hill: Bell Laboratories; 1963.
[105] Markov K. On the triangular inequality in the theory of two-phase random media Technical report. Universite de Sofia, Faculte de Mathematiques; 1995.
[106] Cule D, Torquato S. Generating random media from limited microstructural information via stochastic optimization. J Appl Phys 1999;86(6):3428–37.
[107] Liu XC, Shapiro V. Random heterogeneous materials via texture synthesis. Comput Mater Sci 2015;99:177–89.
[108] Roberts AP, Teubner M. Transport properties of heterogeneous materials derived from Gaussian random fields: bounds and simulation. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics 1995;51(5):4141–54.
[109] Quintanilla J, Torquato S. Lineal measures of clustering in overlapping particle systems. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics 1996;54(4):4027–36.
[110] Turner DM, Niezgoda SR, Kalidindi SR. Efficient computation of the angularly resolved chord length distributions and lineal path functions in large microstructure datasets. Modell Simul Mater Sci Eng 2016;24(7):075002.
[111] Singh H et al. Image based computations of lineal path probability distributions for microstructure representation. Mater Sci Eng A-Struct Mater Prop Microstruct Process 2008;474(1–2):104–11.
[112] Talukdar MS et al. Stochastic reconstruction of chalk from 2D images. Transp Porous Media 2002;48(1):101–23.
[113] Bostanabad R, Chen W, Apley DW. Characterization and reconstruction of 3D stochastic microstructures via supervised learning. J Microsc 2016;264(3):282–97.
[114] Manwart C, Torquato S, Hilfer R. Stochastic reconstruction of sandstones. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics 2000;62(1 Pt B):893–9.
[115] Aarts E, Korst J. Simulated annealing and Boltzmann machines; 1988.
[116] Collins BC, Matous K, Rypl D. Three-dimensional reconstruction of statistically optimal unit cells of multimodal particulate composites. Int J Multiscale Comput Eng 2010;8(5):489–507.
[117] Kumar NC, Matous K, Geubelle PH. Reconstruction of periodic unit cells of multimodal random particulate composites using genetic algorithms. Comput Mater Sci 2008;42(2):352–67.
[118] Deb K et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 2002;6(2):182–97.
[119] Capek P et al. Stochastic reconstruction of particulate media using simulated annealing: improving pore connectivity. Transp Porous Media 2009;76(2):179–98.
[120] Torquato S, Stillinger F. Controlling the short-range order and packing densities of many-particle systems. J Phys Chem B 2002;106(33):8354–9.
[121] Gerke KM, Karsanina MV, Skvortsova EB. Description and reconstruction of the soil pore space using correlation functions. Eurasian Soil Sci 2012;45(9):861–72.
[122] Latief FDE et al. Continuum reconstruction of the pore scale microstructure for Fontainebleau sandstone. Phys A-Stat Mech Appl 2010;389(8):1607–18.
[123] Zhao XC, Yao J, Yi YJ. A new stochastic method of reconstructing porous media. Transp Porous Media 2007;69(1):1–11.
[124] Piasecki R, Olchawa W. Speeding up of microstructure reconstruction: I. Application to labyrinth patterns. Modell Simul Mater Sci Eng 2012;20(5):055003.
[125] Guo EY et al. Accurate modeling and reconstruction of three-dimensional percolating filamentary microstructures from two-dimensional micrographs via dilation-erosion method. Mater Charact 2014;89:33–42.
[126] Zachary CE, Torquato S. Improved reconstructions of random media using dilation and erosion processes. Phys Rev E 2011;84(5):056102.
[127] Li DS et al. 3D reconstruction of carbon nanotube composite microstructure using correlation functions. J Comput Theor Nanosci 2010;7(8):1462–8.
[128] Baniassadi M et al. Three-phase solid oxide fuel cell anode microstructure realization using two-point correlation functions. Acta Mater 2011;59(1):30–43.
[129] Wu W, Jiang F. Simulated annealing reconstruction and characterization of the three-dimensional microstructure of a $LiCoO_2$ lithium-ion battery cathode. Mater Charact 2013;80:62–8.

[130] Chen D et al. Stable-phase method for hierarchical annealing in the reconstruction of porous media images. Phys Rev E Stat Nonlin Soft Matter Phys 2014;89(1):013305.
[131] Roberts AP. Statistical reconstruction of three-dimensional porous media from two-dimensional images. Phys Rev E 1997;56(3):3203–12.
[132] Rozman MG, Utz M. Efficient reconstruction of multiphase morphologies from correlation functions. Phys Rev E Stat Nonlin Soft Matter Phys 2001;63 (6 Pt 2):066701.
[133] Torquato S. Optimal design of heterogeneous materials. Annu Rev Mater Res 2010;40(40):101–29.
[134] Ballani F, Stoyan D. Reconstruction of random heterogeneous media. J Microsc 2015;258(3):173–8.
[135] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. Science 1983;220(4598):671–80.
[136] Talukdar MS et al. Stochastic reconstruction, 3D characterization and network modeling of chalk. J Petrol Sci Eng 2002;35(1–2):1–21.
[137] Liu Y et al. Computational microstructure characterization and reconstruction for stochastic multiscale material design. Comput Aided Des 2013;45 (1):65–76.
[138] Dueck G, Scheuer T. Threshold accepting – a general-purpose optimization algorithm appearing superior to simulated annealing. J Comput Phys 1990;90(1):161–75.
[139] Dueck G. New optimization heuristics – the great deluge algorithm and the record-to-record travel. J Comput Phys 1993;104(1):86–92.
[140] Jiang Z, Chen W, Burkhart C. Efficient 3D porous microstructure reconstruction via Gaussian random field and hybrid optimization. J Microsc 2013;252(2):135–48.
[141] Tang T, Teng Q-Z, He X-H. A hybrid reconstruction method of sandstone from 2D section image. In: Neural networks and signal processing, 2008 international conference on. IEEE; 2008.
[142] Alexander SK et al. Hierarchical annealing for synthesis of binary images. Math Geosci 2009;41(4):357–78.
[143] Burt PJ. Fast filter transforms for image-processing. Comput Graph Image Process 1981;16(1):20–51.
[144] Burt P, Adelson E. The Laplacian pyramid as a compact image code. IEEE Trans Commun 1983;31(4):532–40.
[145] Chiu SN et al. Stochastic geometry and its applications. John Wiley & Sons; 2013.
[146] Tewari A, Gokhale AM. Nearest-neighbor distances between particles of finite size in three-dimensional uniform random microstructures. Mater Sci Eng A-Struct Mater Prop Microstruct Process 2004;385(1–2):332–41.
[147] Al-Ostaz A, Diwakar A, Alzebdeh KI. Statistical model for characterizing random microstructure of inclusion-matrix composites. J Mater Sci 2007;42 (16):7016–30.
[148] Tewari A, Dighe M, Gokhale AM. Quantitative characterization of spatial arrangement of micropores in cast microstructures. Mater Charact 1998;40 (2):119–32.
[149] Fischer R, Miles R. The role of spatial pattern in the competition between crop plants and weeds. A theoretical analysis. Math Biosci 1973;18 (3):335–50.
[150] Zou G, Wu H. Nearest-neighbor distribution of interacting biological entities. J Theor Biol 1995;172(4):347–53.
[151] Xu HY et al. A machine learning-based design representation method for designing heterogeneous microstructures. J Mech Des 2015;137(5):051403.
[152] Zhang Y et al. Microstructure reconstruction and structural equation modeling for computational design of nanodielectrics. Integr Mater Manuf Innov 2015;4(1):1–26.
[153] Ganesh V, Chawla N. Effect of particle orientation anisotropy on the tensile behavior of metal matrix composites: experiments and microstructure-based simulation. Mater Sci Eng, A 2005;391(1):342–53.
[154] Kenney B et al. Computation of TPB length, surface area and pore size from numerical reconstruction of composite solid oxide fuel cell electrodes. J Power Sources 2009;189(2):1051–9.
[155] Pattan PC, Mytri V, Hiremath P. Classification of cast iron based on graphite grain morphology using neural network approach. In: Second international conference on digital image processing. International Society for Optics and Photonics; 2010.
[156] Klaysom C et al. The effects of aspect ratio of inorganic fillers on the structure and property of composite ion-exchange membranes. J Colloid Interface Sci 2011;363(2):431–9.
[157] Jean A et al. A multiscale microstructure model of carbon black distribution in rubber. J Microsc 2011;241(3):243–60.
[158] Pal NR, Pal SK. A review on image segmentation techniques. Pattern Recogn 1993;26(9):1277–94.
[159] Esna Ashari S, Buscarnera G, Cusatis G. A lattice discrete particle model for pressure-dependent inelasticity in granular rocks. Int J Rock Mech Min Sci 2017;91:49–58.
[160] Xu HY et al. Stochastic reassembly strategy for managing information complexity in heterogeneous materials analysis and design. J Mech Des 2013;135(10):101010.
[161] Sacks J et al. Design and analysis of computer experiments. Stat Sci 1989:409–23.
[162] Johnson ME, Moore LM, Ylvisaker D. Minimax and maximin distance designs. J Stat Plan Inference 1990;26(2):131–48.
[163] Simpson TW, Lin DK, Chen W. Sampling strategies for computer experiments: design and analysis. Int J Reliab Appl 2001;2(3):209–40.
[164] Jin R, Chen W, Sudjianto A. On sequential sampling for global metamodeling in engineering design. In: ASME 2002 international design engineering technical conferences and computers and information in engineering conference. American Society of Mechanical Engineers; 2002.
[165] Jin R, Chen W, Sudjianto A. An efficient algorithm for constructing optimal design of computer experiments. J Stat Plan Inference 2005;134 (1):268–87.
[166] Weibel ER. Stereological methods, vol. 2. London: Academic Press; 1980.
[167] Saltykov SA. Stereometric metallography. Moscow: Metallurgizdat; 1958. p. 267.
[168] Snyder VA, Alkemper J, Voorhees PW. The development of spatial correlations during Ostwald ripening: a test of theory. Acta Mater 2000;48 (10):2689–701.
[169] Dehoff RT. A geometrically general-theory of diffusion controlled coarsening. Acta Metall Mater 1991;39(10):2349–60.
[170] Reiss H, Frisch HL, Lebowitz JL. Statistical mechanics of rigid spheres. J Chem Phys 1959;31(2):369–80.
[171] Torquato S, Lu B, Rubinstein J. Nearest-neighbor distribution functions in many-body systems. Phys Rev A 1990;41(4):2059–75.
[172] Truskett TM et al. Structural precursor to freezing in the hard-disk and hard-sphere systems. Phys Rev E 1998;58(3):3083–8.
[173] Li M et al. Three dimensional characterization and modeling of particle reinforced metal matrix composites: part I – Quantitative description of microstructural morphology. Mater Sci Eng A-Struct Mater Prop Microstruct Process 1999;265(1–2):153–73.
[174] Nan CW, Clarke DR. The influence of particle size and particle fracture on the elastic/plastic deformation of metal matrix composites. Acta Mater 1996;44(9):3801–11.
[175] Karasek L, Sumita M. Characterization of dispersion state of filler and polymer-filler interactions in rubber carbon black composites. J Mater Sci 1996;31(2):281–9.
[176] Yuan MJ, Turng LS. Microstructure and mechanical properties of microcellular injection molded polyamide-6 nanocomposites. Polymer 2005;46 (18):7273–92.
[177] Baghgar M et al. Morphology-dependent electronic properties in cross-linked (P3HT-b-P3MT) block copolymer nanostructures. ACS Nano 2014;8 (8):8344–9.
[178] Jr B, Suresh S, Wienecke HA. Deformation of metal-matrix composites with continuous fibers – geometrical effects of fiber distribution and shape. Acta Metall Mater 1991;39(5):735–52.
[179] Christman T, Needleman A, Suresh S. An experimental and numerical study of deformation in metal ceramic composites. Acta Metall 1989;37 (11):3029–50.
[180] Moorthy S, Ghosh S. A Voronoi cell finite element model for particle cracking in elastic-plastic composite materials. Comput Methods Appl Mech Eng 1998;151(3–4):377–400.

[181] Ghosh S, Moorthy S. Particle fracture simulation in non-uniform microstructures of metal-matrix composites. Acta Mater 1998;46(3):965–82.
[182] Lewandowski JJ, Liu C, Hunt WH. Effects of matrix microstructure and particle distribution on fracture of an aluminum metal matrix composite. Mater Sci Eng A-Struct Mater Prop Microstruct Process 1989;107:241–55.
[183] Singh PM, Lewandowski JJ. The effects of reinforcement additions and heat-treatment on the evolution of the poisson ratio during straining of discontinuously reinforced aluminum-alloys. Metall Mater Trans A-Phys Metall Mater Sci 1995;26(11):2911–21.
[184] Mummery PM, Derby B, Scruby CB. Acoustic-emission from particulate-reinforced metal-matrix composites. Acta Metall Mater 1993;41(5):1431–45.
[185] Wilson S et al. Microstructural characterization and evolution in 3D. In: Proc 31st Risø int symp on materials science, Technical University of Denmark; 2010.
[186] Sintay SD, Rollett AD. Testing the accuracy of microstructure reconstruction in three dimensions using phantoms. Modell Simul Mater Sci Eng 2012;20(7):075005.
[187] Saylor DM et al. Statistically representative three-dimensional microstructures based on orthogonal observation sections. Metall Mater Trans A-Phys Metall Mater Sci 2004;35a(7):1969–79.
[188] Brahme A et al. 3D reconstruction of microstructure in a commercial purity aluminum. Scr Mater 2006;55(1):75–80.
[189] Gerard C et al. Comparison of experimental results and finite element simulation of strain localization scheme under cyclic loading. Comput Mater Sci 2009;46(3):755–60.
[190] Groeber M et al. A framework for automated analysis and simulation of 3D polycrystalline microstructures. Acta Mater 2008;56(6):1257–73.
[191] Groeber MA, Jackson MA. DREAM. 3D: a digital representation environment for the analysis of microstructure in 3D. Integr Mater Manuf Innov 2014;3(1):1.
[192] Mahin KW, Hanson K, Morris JW. Comparative-analysis of the cellular and Johnson-Mehl microstructures through computer-simulation. Acta Metall 1980;28(4):443–53.
[193] Avrami M. Kinetics of phase change I – general theory. J Chem Phys 1939;7(12):1103–12.
[194] Gilbert EN. Random subdivisions of space into crystals. Ann Math Stat 1962;33(3):958.
[195] Barbe F et al. Intergranular and intragranular behavior of polycrystalline aggregates. Part 1: F.E. model. Int J Plast 2001;17(4):513–36.
[196] Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3(Mar):1157–82.
[197] Robnik-Šikonja M, Kononenko I. An adaptation of relief for attribute estimation in regression; 1997.
[198] Hassinger I et al. Toward the development of a quantitative tool for predicting dispersion of nanocomposites under non-equilibrium processing conditions. J Mater Sci 2016;51(9):4238–49.
[199] Fabrigar LR et al. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 1999;4(3):272.
[200] Thompson B. Exploratory and confirmatory factor analysis: understanding concepts and applications. American Psychological Association; 2004.
[201] Jolliffe I. Principal component analysis. Wiley Online Library; 2002.
[202] Ganapathysubramanian B, Zabaras N. Modeling diffusion in random heterogeneous media: data-driven models, stochastic collocation and the variational multiscale method. J Comput Phys 2007;226(1):326–53.
[203] Choudhury A et al. Quantification and classification of microstructures in ternary eutectic alloys using 2-point spatial correlations and principal component analyses. Acta Mater 2016;110:131–41.
[204] Wen B, Zabaras N. A multiscale approach for model reduction of random microstructures. Comput Mater Sci 2012;63:269–85.
[205] Li Z, Wen B, Zabaras N. Computing mechanical response variability of polycrystalline microstructures through dimensionality reduction techniques. Comput Mater Sci 2010;49(3):568–81.
[206] Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. Science 2000;290(5500):2323–6.
[207] Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. In: NIPS; 2001.
[208] Donoho DL, Grimes C. Hessian eigenmaps: locally linear embedding techniques for high-dimensional data. Proc Natl Acad Sci 2003;100(10):5591–6.
[209] Tenenbaum JB, De Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. Science 2000;290(5500):2319–23.
[210] Saxena A, Gupta A, Mukerjee A. Non-linear dimensionality reduction by locally linear isomaps. In: Neural information processing. Springer; 2004.
[211] Abboud S, Bruderman I, Sadeh D. Frequency and time domain analysis of airflow breath patterns in patients with chronic obstructive airway disease. Comput Biomed Res 1986;19(3):266–73.
[212] Evans J, Morgan P, Renaud R. Simulation of electron spin resonance spectra by fast Fourier transform: a novel method of calculating spectra to include isotopic substitution, superhyperfine coupling, instrument time constant and modulation broadening in. Anal Chim Acta 1978;103(2):175–87.
[213] Berrut J-P. A Fredholm integral equation of the second kind for conformal mapping. J Comput Appl Math 1986;14(1–2):99–110.
[214] Rabiner LR, Gold B. Theory and application of digital signal processing, vol. 1. Englewood Cliffs, NJ: Prentice-Hall Inc; 1975. p. 777.
[215] Akiyama T. Pressure estimation from oscillatory signals obtained through BWR's instrument lines. J Dyn Syst Meas Contr 1986;108(1):80–5.
[216] Lim JS. Two-dimensional signal and image processing. Englewood Cliffs, NJ: Prentice Hall; 1990. p. 710.
[217] Salawu O. Detection of structural damage through changes in frequency: a review. Eng Struct 1997;19(9):718–23.
[218] Michel JC, Moulinec H, Suquet P. Effective properties of composite materials with periodic microstructure: a computational approach. Comput Methods Appl Mech Eng 1999;172(1):109–43.
[219] Moulinec H, Suquet P. A numerical method for computing the overall response of nonlinear composites with complex microstructure. Comput Methods Appl Mech Eng 1998;157(1–2):69–94.
[220] Michel J, Moulinec H, Suquet P. A computational scheme for linear and non-linear composites with arbitrary phase contrast. Int J Numer Meth Eng 2001;52(1–2):139–60.
[221] Brigham EO et al. The fast Fourier transform and its applications. Prentice Hall; 1988.
[222] Yu S et al. Design of non-deterministic quasi-random nanophotonic structures using Fourier space representations. Sci Rep 2017;7.
[223] Huang JS, Goldburg WI, Bjerkaas AW. Study of phase separation in a critical binary-liquid mixture – spinodal decomposition. Phys Rev Lett 1974;32(17):921–3.
[224] Huntington MD et al. Polymer nanowrinkles with continuously tunable wavelengths. ACS Appl Mater Interfaces 2013;5(13):6438–42.
[225] Chatfield C. The analysis of time series: an introduction. CRC Press; 2016.
[226] Teubner M. Level surfaces of Gaussian random-fields and microemulsions. Europhys Lett 1991;14(5):403–8.
[227] Grigoriu M. Random field models for two-phase microstructures. J Appl Phys 2003;94(6):3762–70.
[228] Rahman S. A random field model for generating synthetic microstructures of functionally graded materials. Int J Numer Meth Eng 2008;76(7):972–93.
[229] Levitz P. Off-lattice reconstruction of porous media: critical evaluation, geometrical confinement and molecular transport. Adv Colloid Interface Sci 1998;76:71–106.
[230] Berk NF. Scattering properties of a model bicontinuous structure with a well defined length scale. Phys Rev Lett 1987;58(25):2718–21.
[231] Cahn JW. Phase separation by spinodal decomposition in isotropic systems. J Chem Phys 1965;42(1):93.
[232] Zhang Y et al. Using cuttlefish ink as an additive to produce non-iridescent structural colors of high color visibility. Adv Mater 2015;27(32):4719–24.
[233] Conley GM et al. Light transport and localization in two-dimensional correlated disorder. Phys Rev Lett 2014;112(14):143901.
[234] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44.
[235] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;313(5786):504–7.
[236] Norouzi M. Convolutional restricted Boltzmann machines for feature learning. School of Computing Science-Simon Fraser University; 2009.
[237] Lee H et al. Unsupervised feature learning for audio classification using convolutional deep belief networks; 2009.
[238] Cang R et al. Microstructure representation and reconstruction of heterogeneous materials via deep belief network for computational material design. arXiv preprint arXiv:1612.07401; 2016.

[239] Cang RJ, Ren MY. Deep network-based feature extraction and reconstruction of complex material microstructures. Proceedings of the ASME international design engineering technical conferences and computers and information in engineering conference, Vol. 2b. p. 95–104.
[240] Lee H et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. ACM; 2009.
[241] Cang R, Ren MY. Deep network-based feature extraction and reconstruction of complex material microstructures. American Society of Mechanical Engineers; 2016.
[242] Aha DW, Kibler D, Albert MK. Instance-based learning algorithms. Mach Learn 1991;6(1):37–66.
[243] Casella G, George EI. Explaining the Gibbs sampler. Am Stat 1992;46(3):167–74.
[244] Burt PJ, Adelson EH. The Laplacian pyramid as a compact image code. IEEE Trans Commun 1983;31(4):532–40.
[245] Hastie T et al. The elements of statistical learning, vol. 2. Springer; 2009.
[246] Breiman L et al. Classification and regression trees. USA: CRC Press; 1984.
[247] Ripley BD. Pattern recognition and neural networks. Cambridge University Press; 1996.
[248] De Bonet JS. Multiresolution sampling procedure for analysis and synthesis of texture images. In: Proceedings of the 24th annual conference on computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co; 1997.
[249] Heeger DJ, Bergen JR. Pyramid-based texture analysis/synthesis. In: Proceedings of the 22nd annual conference on computer graphics and interactive techniques. ACM; 1995.
[250] Zhu SC, Wu YN, Mumford D. Filters, random fields and maximum entropy (FRAME): towards a unified theory for texture modeling. Int J Comput Vision 1998;27(2):107–26.
[251] Popat K, Picard RW. Novel cluster-based probability model for texture synthesis, classification, and compression. In: Visual communications' 93. International Society for Optics and Photonics; 1993.
[252] Efros AA, Leung TK. Texture synthesis by non-parametric sampling. In: Computer vision, 1999. The proceedings of the seventh IEEE international conference on. IEEE; 1999.
[253] Wei L-Y, Levoy M. Fast texture synthesis using tree-structured vector quantization. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co.; 2000. p. 479–88.
[254] Efros AA, Freeman WT. Image quilting for texture synthesis and transfer. In: Proceedings of the 28th annual conference on computer graphics and interactive techniques. ACM; 2001.
[255] Sundararaghavan V. Reconstruction of three-dimensional anisotropic microstructures from two-dimensional micrographs imaged on orthogonal planes. Integr Mater Manuf Innov 2014;3(1):1–11.
[256] Kumar A et al. A Markov random field approach for microstructure synthesis. Modell Simul Mater Sci Eng 2016;24(3):035015.
[257] Acar P, Sundararaghavan V. A Markov random field approach for modeling spatio-temporal evolution of microstructures. Modell Simul Mater Sci Eng 2016;24(7):075005.
[258] Krill Iii C, Chen L-Q. Computer simulation of 3-D grain growth using a phase-field model. Acta Mater 2002;50(12):3059–75.
[259] Gruber J et al. Sparse data structure and algorithm for the phase field method. Modell Simul Mater Sci Eng 2006;14(7):1189.
[260] Kopf J et al. Solid texture synthesis from 2d exemplars. In: ACM transactions on graphics (TOG). ACM; 2007.
[261] Tahmasebi P, Hezarkhani A, Sahimi M. Multiple-point geostatistical modeling based on the cross-correlation functions. Comput Geosci 2012;16(3):779–97.
[262] Tahmasebi P, Sahimi M. Cross-correlation function for accurate reconstruction of heterogeneous media. Phys Rev Lett 2013;110(7):078002.
[263] Mariethoz G, Lefebvre S. Bridges between multiple-point geostatistics and texture synthesis: review and guidelines for future research. Comput Geosci 2014;66:66–80.
[264] Wu KJ et al. An efficient Markov chain model for the simulation of heterogeneous soil structure. Soil Sci Soc Am J 2004;68(2):346–51.
[265] Elfeki A, Dekking M. A Markov chain model for subsurface characterization: theory and applications. Math Geol 2001;33(5):569–89.
[266] Hajizadeh A, Safekordi A, Farhadpour FA. A multiple-point statistics algorithm for 3D pore space reconstruction from 2D images. Adv Water Resour 2011;34(10):1256–67.
[267] Okabe H, Blunt MJ. Pore space reconstruction using multiple-point statistics. J Petrol Sci Eng 2005;46(1–2):121–37.
[268] Strebelle S. Conditional simulation of complex geological structures using multiple-point statistics. Math Geol 2002;34(1):1–21.
[269] Guardiano FB, Srivastava RM. Multivariate geostatistics: beyond bivariate moments. In: Geostatistics Troia'92. Springer; 1993. p. 133–44.
[270] Caers J, Journel AG. Stochastic reservoir simulation using neural networks trained on outcrop data. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers; 1998.
[271] Harrison P. A non-hierarchical procedure for re-synthesis of complex textures; 2001.
[272] Cressie N, Davidson JL. Image analysis with partially ordered markov models. Comput Stat Data Anal 1998;29(1):1–26.
[273] Cross GR, Jain AK. Markov random field texture models. IEEE Trans Pattern Anal Mach Intell 1983;1:25–39.
[274] Portilla J, Simoncelli EP. A parametric texture model based on joint statistics of complex wavelet coefficients. Int J Comput Vision 2000;40(1):49–70.
[275] Simoncelli EP, Portilla J. Texture characterization via joint statistics of wavelet coefficient magnitudes. In: Image processing, 1998. ICIP 98. Proceedings. 1998 international conference on. IEEE; 1998.
[276] Cressie N, Wikle CK. Statistics for spatio-temporal data. John Wiley & Sons; 2015.
[277] Bárdossy A, Li J. Geostatistical interpolation using copulas. Water Resour Res 2008;44(7).
[278] Le BA, Yvonnet J, He QC. Computational homogenization of nonlinear elastic materials using neural networks. Int J Numer Meth Eng 2015;104(12):1061–84.
[279] Feyel F. A multilevel finite element method (FE 2) to describe the response of highly non-linear structures using generalized continua. Comput Methods Appl Mech Eng 2003;192(28):3233–44.
[280] Feyel F, Chaboche J-L. FE 2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. Comput Methods Appl Mech Eng 2000;183(3):309–30.
[281] Bessa MA et al. A framework for data-driven analysis of materials under uncertainty: countering the curse of dimensionality. Comput Methods Appl Mech Eng 2017.
[282] Arruda EM, Boyce MC. A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials. J Mech Phys Solids 1993;41(2):389–412.
[283] Belytschko T et al. Nonlinear finite elements for continua and structures. John wiley & Sons; 2013.
[284] Sobol' IyM. On the distribution of points in a cube and the approximate evaluation of integrals. Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki 1967;7(4):784–802.
[285] Sobol IM. On quasi-Monte Carlo integrations. Math Comput Simul 1998;47(2–5):103–12.
[286] Hassaninia I et al. Characterization of the optical properties of turbid media by supervised learning of scattering patterns. Sci Rep 2017;7(1):15259.
[287] Bostanabad R et al. Leveraging the nugget parameter for efficient Gaussian process modeling. Int J Numer Methods Eng.
[288] Hoel PG. Introduction to mathematical statistics, 2nd ed.; 1954.
[289] Netravali AN, Limb JO. Picture coding: a review. Proc IEEE 1980;68(3):366–406.