Accurate modeling and evaluation of microstructures in complex materials

Pejman Tahmasebi*
Department of Petroleum Engineering, University of Wyoming, Laramie, Wyoming 82071, USA

(Received 24 November 2017; published 20 February 2018)

Accurate characterization of heterogeneous materials is of great importance for different fields of science and engineering. Such a goal can be achieved through imaging. Acquiring three- or two-dimensional images under different conditions is not, however, always plausible. On the other hand, accurate characterization of complex and multiphase materials requires various digital images (**I**) under different conditions. An ensemble method is presented that can take one single (or a set of) **I**(s) and stochastically produce several similar models of the given disordered material. The method is based on a successive calculating of a conditional probability by which the initial stochastic models are produced. Then, a graph formulation is utilized for removing unrealistic structures. A distance transform function for the **I**s with highly connected microstructure and long-range features is considered which results in a new **I** that is more informative. Reproduction of the **I** is also considered through a histogram matching approach in an iterative framework. Such an iterative algorithm avoids reproduction of unrealistic structures. Furthermore, a multiscale approach, based on pyramid representation of the large **I**s, is presented that can produce materials with millions of pixels in a matter of seconds. Finally, the nonstationary systems—those for which the distribution of data varies spatially—are studied using two different methods. The method is tested on several complex and large examples of microstructures. The produced results are all in excellent agreement with the utilized **I**s and the similarities are quantified using various correlation functions.

DOI: 10.1103/PhysRevE.97.023307

I. INTRODUCTION

Accurate knowledge of microstructure and heterogeneous materials is crucial for different aims such as performance evaluation, prognosis, and optimization of structures [1,2]. Furthermore, discovering new materials requires tremendous research on materials with complex structures and distributions. Mathematical modeling can help to characterize and evaluate microstructures with different properties. Such methods should also deal with a wide range of heterogeneous materials. For example, designing new materials with optimal structures requires considering several possibilities of morphologies. Thus, any new method in this field requires accurately quantifying the inherent complexities and heterogeneities. Then, the extracted properties are used to generate (i.e., reconstruct) new microstructures with the hope of statistically and physically representing the properties of the original samples.

Recent progress in nondestructive two- and three-dimensional (2D and 3D) imaging has enabled us to extract more information about the complexity and heterogeneity of materials through digital images (**I**) [3–6]. X-ray imaging [7,8] can eliminate destructive sectioning and provide high-resolution 3D images for various applications such as Sn-rich alloys [9], powder metallurgy steels [10], metamatrix composites [11–14], and light-weight alloys [15–19]. High-resolution focus ion-beam scanning electron microscopy (FIB-SEM) methods are an essential part of materials characterization as they reveal vital information regarding the spatial or morphological properties and connectivities between different phases. Providing 3D images, however, is costly and time consuming. Further, acquiring high-resolution images can cause one to miss different pore spaces and structures. Four-dimensional (4D; the fourth dimension being time) imaging has also provided the ability to investigate the structural evaluation over time under different stimuli (e.g., stress, temperature, and environmental conditions). 4D images can also help to better understand time-dependent processes such as deformation, phase transformations, magnetic domains, and compositional evaluation.

Improving our understanding of the existing nonlinear and complex deformational behaviors in materials science requires extensive research and characterization of such materials under different morphological and structural conditions. Besides, relying on a limited number of 3D images can put significant uncertainty in the characterization and evaluation of physical behaviors such as mechanical properties, structural deformation, fracture, and fatigue as well as contact and friction. Due to the aforementioned requirements and the existing limitations, namely, cost and time, providing several 2D or 3D images of materials is not feasible. Thus, estimating the range of variability and possible behaviors of the above crucial properties will not be likely. For instance, such a realistic evaluation can be very critical when the designed material is used in aircraft or biological devices.

Using a limited number of complex and heterogeneous images and producing a set of similar instances, but with different structural and morphological properties, has been recognized as a promising solution that can fill the existing gap. Therefore, producing comparable and, at the same time, accurate models of heterogeneous samples can quantify the discussed uncertainties. The provided variability and

*ptahmase@uwyo.edu

uncertainty can, indeed, help to revise the compositions and avoid over- or underestimation. To this end, several methods are available that can be used to describe and, consequently, produce various models. In one group, producing different models is achieved using an iterative process (e.g., genetic algorithm or simulated annealing) by which the difference between the generated model and the input image is decreased gradually [20]. Producing high-quality models using these methods requires considering various statistical quantifiers (e.g., lineal path, two-point cluster, and autocorrelation) in the objective function. In another group, called process-based methods, the medium is generated based on some physical rules such as sedimentation, erosion, and interaction between the phases [21]. These methods have been used for natural porous media (e.g., rocks) in which such physical rules can be defined.

Characterization of the multiphase random system requires extracting the correlation between several points, i.e., n-point correlation [2,22,23]. In practice, however, evaluation such a large number of points, usually $n \geqslant 5$, is impractical [24,25]. Several noncanonical, topological descriptors and interface information have also been developed for a wide spectrum of materials [26]. Although the low-order statistics, namely, when $n \leqslant 2$, do not provide accurate results, increasing the statistics to higher orders, such as four and five, even cannot be very useful for the constructing of complex materials [27]. Thus, one of the long-standing problems in the characterization of complex materials is the development of correlation functions that can accurately describe the complexity, only if such correlations can be measured. For this aim, several reconstruction methods have been developed in the past two decades. The Gaussian random field method was originally developed for modeling homogeneous materials using two-point correlation functions [28]. An example of such correlation functions is the field-field correlation function, which the result is used with level cut to produce a binary model of the target microstructure. This method has also been used for a variety of simple and isotropic microstructures, but its application is limited because it uses only the correlation between two points. Later, phase recovery was used to fully take into account the directional information extracted from two-point correlation functions, which can result in generating more heterogeneous microstructures [29,30]. This method is applied on an iterative framework in which the difference between the target and simulated model is reduced repeatedly. This method is, however, limited in the number of the used correlation functions that one can incorporate. Thus, process-based stochastic modeling [31] was later developed to address this shortcoming. In this method, an objective function, representing a set of correlation functions, is considered and the error level of the generated realization is decreased using the given objective function. This method, because it deals with many terms in the objective functions, is computationally expensive. Several implementations of this method have been developed so far that aim to address its inefficiencies [32–36]. Higher-order statistical methods, on the other hand, have also been proposed that can address the aforementioned issues via the direct use of the given complex microstructure in the initial image [37].

Aside from the considerable progress in the reconstruction of microstructures, the available methods still cannot provide an accurate representation of the disordered material. For example, the iterative methods are limited in terms of the number of utilized similarity quantifiers in the objective function. In other words, evaluating an extensive list of statistical terms in the objective function is computationally expensive. Similarly, process-based methods, because there are a lot of rules to consider, are extremely CPU demanding. All such methods are unable to translate the existing physics and complexity provided in the 2D or 3D images. As a result, they cannot reproduce the long-range and complex phases in the materials. Finally, these methods result in generating unrealistic and noisy or smooth models. It should be noted that such issues are more perceptible in 3D modeling as the current methods usually cannot handle very large and multiscale structures in a reasonable time. The importance of multiscale modeling in inhomogeneous microstructures has been discussed by Chen *et al.* [38] and the authors have used a single 2D image and successfully reconstructed polycrystalline microstructures.

In this paper, a method based on a direct use of 2D or 3D images of microstructure materials is proposed that can accurately model the complex materials. The complexity and, consequently, physics of such materials are transferred directly. The proposed method is then used in a multiscale framework by which very large, realistic, and exhaustive images can be modeled efficiently. Furthermore, various correction algorithms are also proposed by which one can remove unrealistic structures and artifacts.

II. METHODOLOGY

In this paper, an alternative concept of direct use of the available images of complex materials is proposed, in which not only are none of the current statistical descriptors are needed, but also different phases are used directly [37,39]. Thus, the first step in this algorithm is obtaining 2D or 3D images of the disordered material (i.e., **I**). It is worth noting that the **I** should be a representative volume element (RVE) of the microstructure. Thus, the **I**s must be checked in terms of being RVEs prior to this study and it will not be investigated in this paper [30]. Then, similar 2D or 3D images are generated using a conditional probability formulation. The proposed method is composed of several stages. First, the general framework is described and then the other components are discussed accordingly.

A. General framework

The algorithm uses the previously simulated regions to decide on the probability of the next cell in a modeling grid $\mathbf{G} = \{\mathcal{Z}_1, \mathcal{Z}_2, \ldots, \mathcal{Z}_N\}$, in which N indicates the number of points in the computational grid $\mathbf{G}$ on which the stochastic model is generated, and each point takes a single value $\mathcal{Z}_i$. The probability of having the Nth phase, given the previous $(N-1)$ phases, is given by

$$p(\mathbf{G}) = p(\mathcal{Z}_1)p(\mathcal{Z}_2|\mathcal{Z}_1)\cdots \underbrace{p(\mathcal{Z}_N|\mathcal{Z}_{N-1}, \mathcal{Z}_{N-2}, \ldots, \mathcal{Z}_2, \mathcal{Z}_1)}_{p(\mathcal{Z}_i|\mathbf{G}_{\Phi_i})}, \tag{1}$$

where $\mathbf{G}_{\Phi_i}$ indicates the conditional probability of $\mathcal{Z}_i$. Here, $\mathbf{G}_{\Phi_i}$ represents the configuration of the points in Φ_i. The

set Φ_i implies the sequential neighborhood of point N, i.e., $\{\Phi_i\}_{i=1,2,\ldots,N}$. One also may write the joint probability of Eq. (1) as

$$p(\mathbf{G}) = \prod_{i=1}^{N} p\left(\mathcal{Z}_i \middle| \mathbf{G}_{\Phi_i}\right). \tag{2}$$

The given probability in Eq. (2) should be calculated for each modeling point in $\mathbf{G}$. Nevertheless, the method would be computationally expensive. Thus, additional complementary strategies are introduced to alleviate the computational burden.

Based on the above conditional probability, instead of using all the previous points, one can consider only a small neighborhood around the visiting point. Thus, a small region around the visiting point can be used for the probability calculations. Such a region can be selected based the directional covariances in different x, y, and z directions. Furthermore, the conditional probability can be estimated using a product kernel such as a Gaussian radial based kernel, which is a measure of similarity. In other words, the conditional probability in Eq. (2) can be written as

$$p\left(\mathcal{Z}_i \middle| \mathbf{G}_{\Phi_i}\right) = \exp\left(-d\left\langle \mathcal{Z}_i, \mathbf{G}_{\Phi_i}\right\rangle\right) = \exp\left(-\sum_{i=1}^{n} \left|\mathcal{Z}_i - N_i^{\mathbf{G}_{\Phi_i}}\right|\right), \tag{3}$$

where N_i represents the neighborhood data in $\mathbf{G}_{\Phi_i}$ and $d\langle\rangle$ indicates Euclidean distance.

The above-discussed conditional probability allows one to consider a relatively small window around the visiting point. Thus, the Euclidean distance can be decomposed into three terms, among which only one of them requires being calculated for each visiting point. Such calculations can be performed using convolution that has been shown to be fast. First, the proposed algorithm selects a pattern (i.e., a set of points or a subset of $\mathbf{I}$) randomly from the $\mathbf{I}$ and inserts in a corner on $\mathbf{G}$. Afterwards, the next pattern is selected based on the similarity between the neighborhood points and the $\mathbf{I}$. Therefore, instead of considering all the previously simulated points, only the points in the current locations are used for convolution calculations [40]. The data event (the current points) at location (x,y) is denoted by $\mathbf{D_T}(x,y)$, where $\mathbf{T}$ represents the size of grid blocks or data event. Note that "data event" refers to data located around the visiting point and it changes for each point. Then, the similarity between the neighboring points and $\mathbf{I}$ is quantified based on the convolution:

$$\Omega(i,j;x,y) = \sum_{x=0}^{\ell_x-1} \sum_{y=0}^{\ell_y-1} \mathbf{I}(x+i, y+j)\mathbf{D_T}(x,y), \tag{4}$$

where $0 \leqslant i < L_x + \ell_x - 1$ and $0 \leqslant j < L_y + \ell_y - 1$. $\mathbf{I}(x,y)$ represents the location at point (x,y) of $\mathbf{I}$ of size $L_x \times L_y$, with $x \in \{0,\ldots,L_x-1\}$ and $y \in \{0,\ldots,L_y-1\}$. An overlap region of size $\ell_x \times \ell_y$ and a data event $\mathbf{D_T}$ are used to match the patterns in the $\mathbf{I}$. Note that the overlap region is the voxels we pick from the previously simulated points to be included in the similarity calculation for identifying the next pattern.

After calculating the similarity between the $\mathbf{I}$ and $\mathbf{D_T}$, one of the similar patterns—those for which the similarity is larger than a preset threshold—is selected and inserted in $\mathbf{G}$. The process is repeated until $\mathbf{G}$ is filled. As a rule of thumb, the neighboring regions might have an overlap of 1/5 to 1/6 of the size of the data event $\mathbf{T}$. Clearly, using large grid blocks will increase the computations, as the similarity function requires considering many points, while small regions may cause discontinuity. Thus, using the above-mentioned size can keep both the computation time and artifacts low.

B. The graph approach

For some complex and structural $\mathbf{I}$s, simple attaching of the heterogeneous patterns may create many discontinuities. This issue can be addressed by implementing the graph theory to remove the discontinuities optimally. Indeed, the graph theory can be used to identify the minimum error boundary by which two patterns are interlocked.

The differences between the pixel (or voxel in three dimensions) values in the neighboring regions are first calculated and represented by a graph $\mathbf{G}_\mathrm{p}$, where the value attributed to each node of the graph is the difference between the pixel values in the overlap regions. The nodes of $\mathbf{G}_\mathrm{p}$ are divided into two groups of terminals: (i) sources $\mathbf{s}$ from which "flow" begins and sinks $\mathbf{t}$ where the flow ends, and (ii) nonterminal nodes $\mathbf{P}$ that are located between the sinks and sources. The graph is given by the set $\mathbf{G}_p = \{V,E\}$, where $V = \mathbf{P}\mathrm{U}\{\boldsymbol{s},\boldsymbol{t}\}$ the edges. Similarly, the edges are also divided into two groups, namely, (i) the t edges $\{(\boldsymbol{s},\boldsymbol{t}),(\boldsymbol{u},\boldsymbol{t})\}$ that connect a nonterminal node in $\mathbf{P}$ with a terminal node, and (ii) the n edges that connect two nonterminal nodes and are denoted by $\mathcal{N}$. Two constraints are imposed on the path $f(u,v)$ between u and v: (i) mass conservation should be satisfied by $f(u,v)$, i.e., $\forall v \in V\backslash\{s,t\} : \sum_{\{u:(u,v)\in E\}} f(u,v) = \sum_{\{u:(v,u)\in E\}} f(v,u)$, and (ii) the antisymmetry relation, i.e., $\forall(u,v) \in V : f(u,v) = -f(v,u)$, must also be followed. The difference between the overlaps is expressed by two subsets $\mathcal{S}$ and $\mathcal{T}$, derived from a source s and a target sink t, respectively. Thus, the optimal path $\mathcal{C}$ is one that minimizes the overall cost function of flow, i.e., $\mathcal{C}(\mathcal{S},\mathcal{T}) = \sum_{u\in\mathcal{S}} \sum_{v\in\mathcal{T}} \mathcal{C}(u,v)$. This minimization problem is solved as a "fluid-flow" model [41] in which the optimal path is identified as one that allows minimum "fluid" passing from s to t with $|f| = \sum_{v\in V} f(s,v)$ and $f(\mathcal{S},\mathcal{T}) = \sum_{u\in\mathcal{S}} \sum_{v\in\mathcal{T}} f(u,v)$. Finally, one can use the graph method when the given $\mathbf{I}$ is too complex and represents very diverse patterns. The graph approach helps to remove the artifact due to interlocking the uncorrelated patterns.

C. Distance transform

Most of the available images for complex materials are mixed with unwanted noises that make their processing a challenging problem. Usually, one prefers to perform a series of postprocessing to remove the noises and artifacts and eventually summarize the pore and solid phase as a binary or categorical image [42]. The binary image does not, however, contain very rich information about the spatial distribution of the pores or target phase. One can acquire such information around each pixel using a distance transform $\mathcal{D}$. In other words, a more informative image can be constructed wherein each of its pixels represents extra information about the

surrounding phases [43,44]. This method, using a Euclidean or quasi-Euclidean distance, calculates the distance of each pixel from its closest objects (e.g., a solid phase). Mathematically speaking, given the closest set Ψ for any node n in a binary image within the $\mathbf{I}$, $\mathcal{D}$ is defined by

$$\mathcal{D}(n,\Psi) = \min_{q\in\Psi}\{d(n,t)\}, \tag{5}$$

where $d(n,q)$ is the Euclidean distance between nodes n and t.

The output of $\mathcal{D}$ represents an informative image that contains zero for the location of pores and nonzero, depending on the distance to the nearest solid phase. The distance transform is useful when the given $\mathbf{I}$ does represent a categorical and multiphase material. The binary or categorical images are not very informative and the distance transform can help producing new images by adding information about the surrounding pixels at each cell.

D. Complex distribution matching

Aside from the possible excessive smoothness and artifacts, complex structures and multimodal distributions make the modeling of the available $\mathbf{I}$s difficult. In some $\mathbf{I}$s, very high and low values are adjacent, in which the convolution ignores the low values. However, a meaningful structure is generated when all the existing phases are reproduced side by side. Thus, histogram matching is utilized for these circumstances to select the candidate patterns more precisely. For this aim, the histogram of the 2D or 3D model $\mathbf{M}$ is constructed per each candidate pattern (d_x,d_y,d_z) and the distance between the $\mathbf{I}$ and each model is quantified using the Jenson-Shannon divergence, $d_{js}(\mathbf{I},\mathbf{M})$. The distance is the average of two Kullback-Leibler divergences [45,46]:

$$d_{JS}(\mathbf{I}||\mathbf{M}) = \frac{1}{2}\sum_k \mathbf{I}_k \log_{10}\left(\frac{\mathbf{I}_k}{\mathbf{M}_k}\right) + \frac{1}{2}\sum_k M_k \log_{10}\left(\frac{\mathbf{M}_k}{\mathbf{I}_k}\right). \tag{6}$$

Then, the pattern with the minimum value of d_{js} is selected. In other words, instead of selecting the final pattern at random, the pattern that helps the current model to be closer to $\mathbf{I}$ is selected eventually. Therefore, the histogram matching indemnifies that the final realization represents a similar distribution as one observes in the initial $\mathbf{I}$. This approach is recommended to be used when one is interested to exactly reproduce the initial distribution given in $\mathbf{I}$. Although the proposed method can reproduce similar statistical distribution in $\mathbf{I}$, such a goal can be achieved in a more rigorous way through histogram matching.

E. Iterative rectification

For actual complex and multiphase materials, even the above methods might not result in a successful modeling. In this case, one may apply an iterative method to gradually improve pattern reproduction and remove the possible artifacts [47]. To this end, an objective function, containing important properties, can be constituted to generate models with closer statistical properties and texture to the original $\mathbf{I}$. One, however, should keep in mind that using several terms in the objective function may significantly increase the computational time. Therefore, there would be always a trade-off between the CPU time and accuracy. Therefore, Eq. (6) can be rewritten as

$$\mathrm{E}(\mathbf{M}) = \sum_z \sum_{p\in\{x,y,z\}} \sum_{u\in \mathrm{T}_p(z)} \|\mathbf{M}_{z,p,u} - d_{z,p,u}\|^2 + \frac{1}{2}\sum_k \mathbf{I}_k \times \log_{10}\left(\frac{\mathbf{I}_k}{\mathbf{M}_k}\right) + \frac{1}{2}\sum_k \mathbf{M}_k \log_{10}\left(\frac{\mathbf{M}_k}{\mathbf{I}_k}\right) + \cdots. \tag{9}$$

It should be noted that the $\mathbf{I}$ must be representative of the complexity and heterogeneity of the material. Thus, one can utilize extra images if a set of $\mathbf{I}$s can be provided. In that case, the presented algorithm in this paper remains the same, while the candidate patterns can be selected from images that are more diverse. Accordingly, the computational time increases since the similarity function is required to be calculated for several images. This approach can be used when the final realization, aside from using the above methods, still show an artifact. The iterative method, indeed, can remove the discontinuities iteratively and produce realizations that better match with the patterns shown in $\mathbf{I}$.

F. Multiscale algorithm

Although the above computational and rectification methods can produce high-quality models of complex materials, the proposed method up to this point can only be applied to small 2D and 3D images. In reality, with the recent progress in the field of imagining, current images of materials are mostly high-resolution 3D data sets. Thus, it is important to handle such complex and large images in a reasonable time. In this section, a multiscale method is proposed by which very large images can be modeled effectively.

The idea of the multiscale method is based on a pyramid representation of the images. In other words, resizing the large images, in most of the complex materials, does not impair the primary details. For instance, one can resize a large 1000×1000 image into two smaller images, namely, 500×500 and 250×250, all of which represent the same structures of the original large image. Thus, the initial large image of the complex material $\mathbf{I}_1$ is resized into two levels, $\mathbf{I}_2$ and $\mathbf{I}_3$. Then, one can start with the smallest image $\mathbf{I}_3$ and perform the described steps (i.e., calculating convolution) on this small image. After finding the location of the candidate pattern, its location is projected on the next finer image $\mathbf{I}_2$. The location of the selected candidate has already been recognized and, thus, its position can be projected on the next finer image $\mathbf{I}_2$. Clearly, one does not need to search the finer image thoroughly since the location of the matching pattern is roughly available from the previous coarse image. Nevertheless, only a small neighbor around the projected location is sufficient for finding the matching pattern. This searching window can save a considerable amount of time for large images. After finding the matching pattern on $\mathbf{I}_2$, its location is projected on the finest image $\mathbf{I}_1$ and only a small neighbor around the projected point is searched. Finally, the identified pattern is selected and inserted in the $\mathbf{G}$.

Using the multiscale method, one can reduce the computational time by identifying the location of the matching pattern

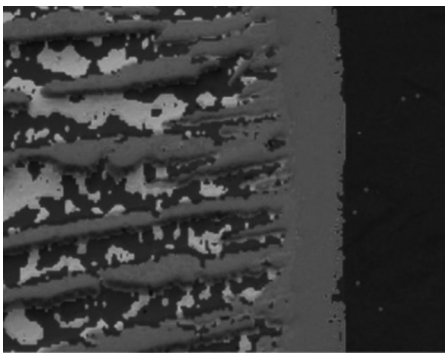

FIG. 1. SEM image of Ti alloy, which represents a nonstationary system [49].

rapidly. It should be noted that the high-resolution image does not require being always resized into three levels of images, but, depending on the initial resolution, one can perform the scaling in two or even four stages. The resizing should be continued until the initial structures of the **I** are preserved in the subsequent resized images. Clearly, this method can be used when the initial **I** is too large as it reduces the CPU time significantly. Thus, this method does not have any effect on the quality.

G. Nonstationary modeling

The described algorithm so far can only model those types of materials that are called stationary. In other words, in stationary systems, the statistical properties in various regions are practically similar. For this reason, the candidate patterns can be selected from any region in **I**. In nonstationary systems, however, the statistical properties in different regions are not necessarily similar [48]. An example is shown in Fig. 1. As can be seen, the phases in the Ti alloy are different in the lower right and left sides. Thus, patterns from any regions in **I** cannot be selected and interlocked with the previous ones in **G**. For instance, the patterns on the left side are not orientationally consistent with those on the right side. Here, two complementary algorithms are presented, by which the nonstationary systems can be modeled efficiently.

1. Using auxiliary data

In this method, the nonstationary factor is first recognized in the material. Nonstationarity can occur for various reasons. For example, it can be due to the rotation, thickness, proportion, size, and so on. Then, auxiliary data (AUX) for describing the nonstationarity is generated, by which the identified behavior is described. In this study, the nonstationarity is extracted manually, meaning that the user can decide on the nonstationary trend and behavior and produce a related descriptor. An example is shown in Fig. 2. The presented image in Fig. 2(a) represents a nonstationary system as the target material is distributed systematically (i.e., gradually) on the background phase. Proportion is an appropriate indicator that can describe the nonstationarity in this system. Since the target phase represents a gradually increase from left to right, similar AUX data, as shown in Fig. 2(b), can be considered for this image. The AUX implicitly quantifies the observed proportion in the **I** shown in Fig. 2(a). One can consider two or more AUX data for more complex nonstationary materials. An example is shown in Fig. 2(c) in which the orientation and thickness are both nonstationary. Thus, the thickness can be introduced using the AUX data in Fig. 2(d), while the orientation is shown using other AUX data in Fig. 2(e).

After generating the AUX data, the proposed algorithm should take the new information into account. In these systms, the selected patterns will not be only conditioned to the previously simulated points, but the AUX should also be considered. To this end, a similar AUX is generated for the **I.** Next, at each visiting point, a new data-even $\mathbf{D}_{\mathrm{AUX}}$ from the AUX data is extracted. Then, the similarity of these extracted data is calculated against the AUX data of **I**, which yield another similarity map. Afterward, the similarity maps derived from the neighboring points and AUX are added together, which results in a single similarity map. Finally, the patterns in **I** are sorted based on this map.

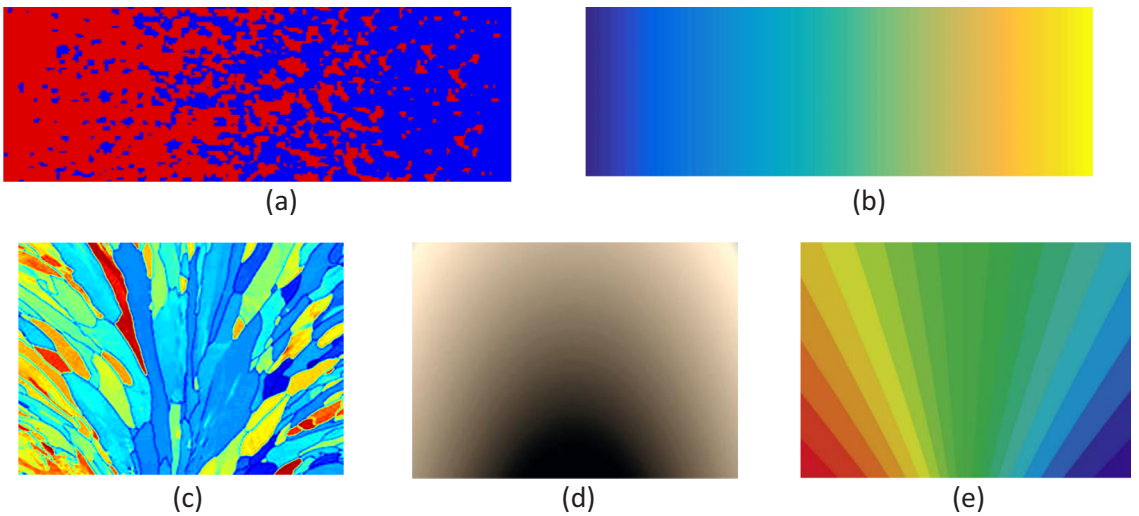

FIG. 2. (a), (c) Two examples of nonstationary systems and their corresponding **AUX** data in (b) and (d), (e), respectively.

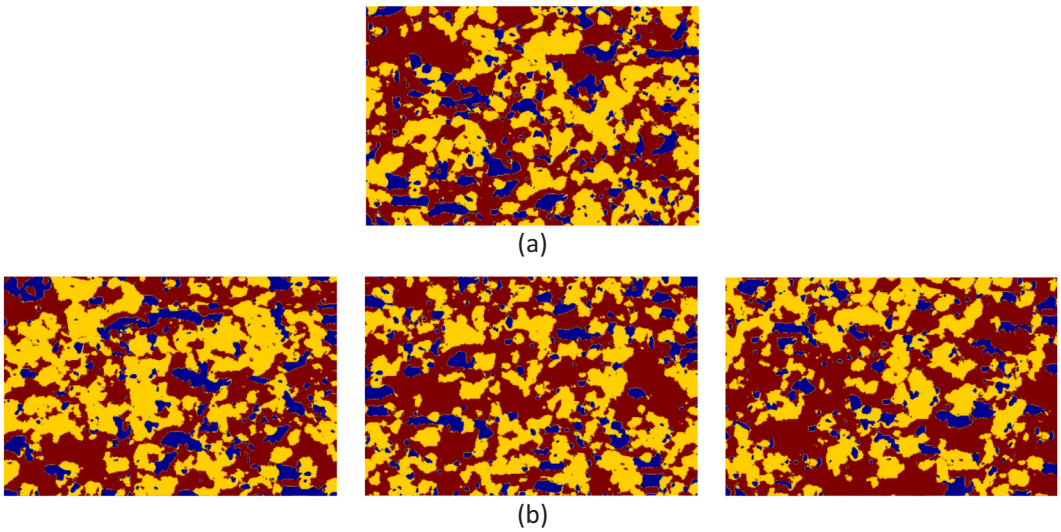

FIG. 3. SEM image of a porous Ni-YSZ sample [54]. Backscattered electron detector is used for this image (a). Yellow, red, and blue represent YSZ, Ni, and pore space, respectively. (b) Three realizations.

As mentioned, one can consider various (s) AUX data, depending on the number of nonstationary behaviors in the system. Thus, the final equation can be written as

$$\Omega_{\text{overall}} = \sum_{s=1}^{n} |\mathbf{I} - \mathbf{D}_{\text{I}}| + \beta_s \sum_{s=1}^{n} |\mathbf{I} - \mathbf{D}_{\text{AUX}}|, \qquad (10)$$

where β_s is a coefficient for each AUX data set. Intuitively, the nonstationary behaviors with higher impact will receive higher weight or coefficient.

2. *Partitioning*

In some **I**s, however, defining the AUX data is not straightforward. For example, the image shown in Fig. 1 is one such example. Partitioning the **I** into some stationary regions is another strategy that can be used for these complex materials. Each segment, in this case, will represent a stationary region that can be handled using the algorithm presented in this paper. However, the data in the comparted regions can only be used for simulating those specific parts and the patterns from other regions not allowed to be used elsewhere. This method is further discussed in Sec. III C.

III. RESULTS AND DISCUSSION

Several complex and heterogeneous examples of microstructure materials are selected for testing the accuracy of the proposed method. The samples are provided in both 2D and 3D images in the stationary and nonstationary distribution of different phases. Then, the generated realizations are compared visually and quantitatively based on the original **I**s. Furthermore, three statistical functions, namely, the two-point cluster correlation function (C_2), the autocorrelation function (ACF), and multiple-point connectivity (MPC), are used in this study to quantify the similarity between the initial **I** and produced realizations [2]. For example, C_2 calculates the

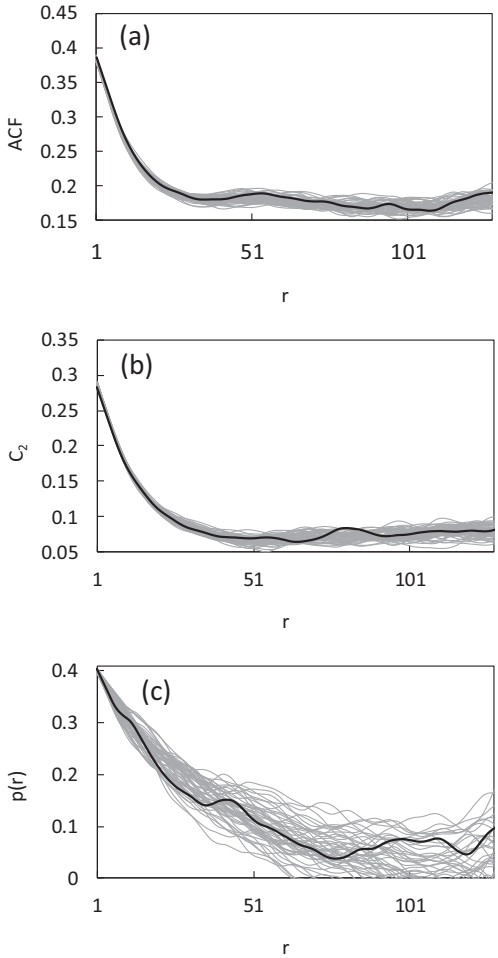

FIG. 4. Comparison between (a) ACF, (b) C_2, and (C) MPC between **I** (black curve) and generated realizations (gray curves) for the yellow phase.

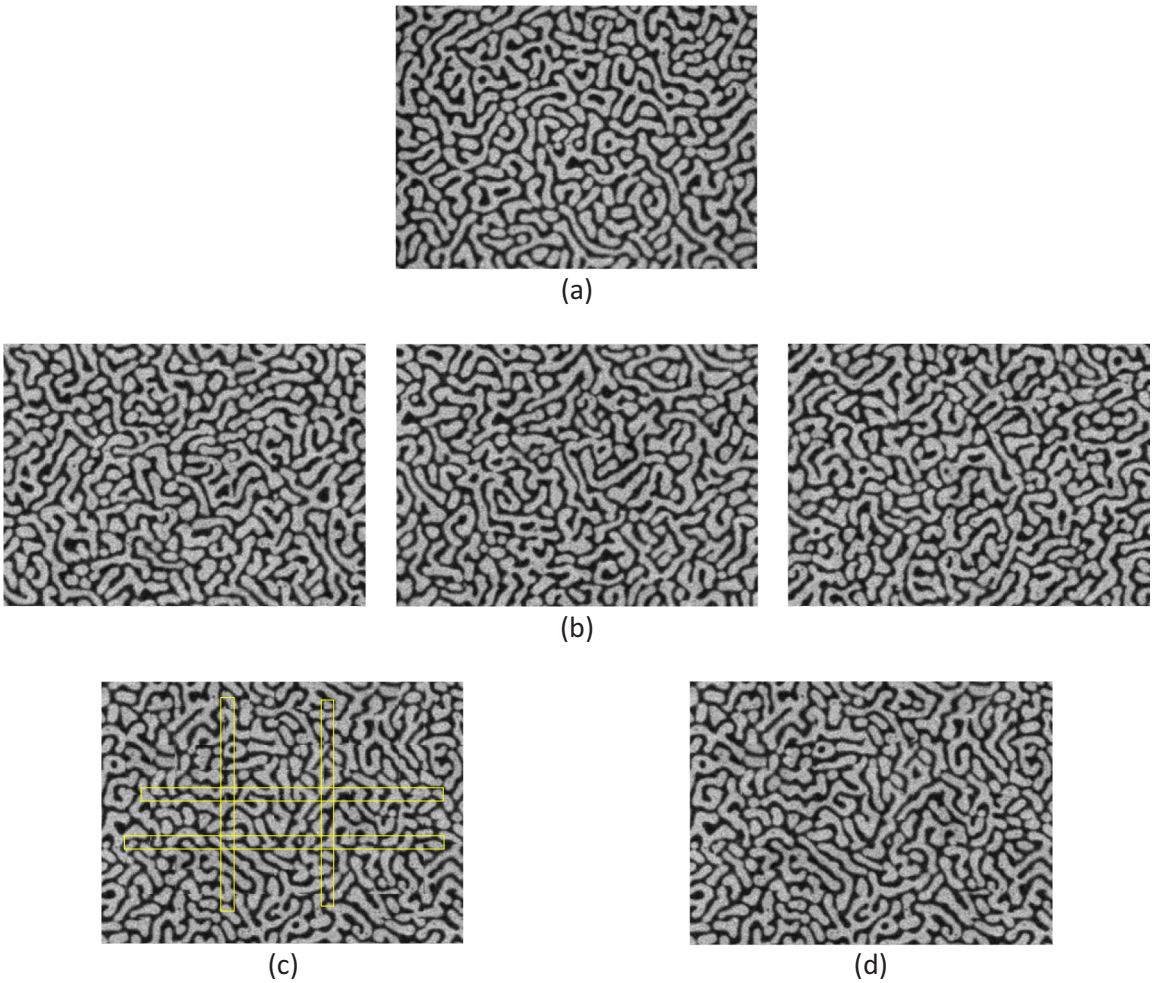

FIG. 5. (b) Generated realizations using (a) the stroboscopic optical reflectance image recorded during spin coating. (c) A raw realization without any correction and (d) the corrected realization in (c).

probability of finding both points r_1 and r_2 in the same cluster of phase i. In this function, the histogram of the distances between any two points in the same phase is calculated. This calculation is repeated for the other target phase in the $\mathbf{I}$ and their corresponding histograms are built. Finally, the C_2 is calculated based on the ratio of the histograms.

The MPC, as another accurate similarity function, represents the probability $p(\mathbf{r};m)$ of having a sequence of m points in a phase in a direction $\mathbf{r}$. Then, an indicator function $I^{(i)}(\mathbf{u})$ is defined by $I^{(i)}(\mathbf{u}) = 1$; if $\mathbf{u} \in$ phase i, and $I^{(i)}(\mathbf{u}) = 0$ otherwise, then

$$p(\mathbf{r};m) = \text{Prob}\{I^{(i)}(\mathbf{u}) = 1, I^{(i)}(\mathbf{u}+\mathbf{r}) = 1, \ldots, I^{(i)}(\mathbf{u}+m\mathbf{r}) = 1\}. \quad (11)$$

The MPC function extracts the multiple-point probability and represents a better tool for dissimilarity quantification.

A. Porous Ni-YSZ sample

Fuel cells are one of the resources that recently has emerged as a promising and efficient source of energy [50]. They, indeed, are in accordance with the recent policy for energy resources such as efficiency, cleanliness, and being economic. These materials convert chemical energy to electricity. A 2D microstructure of a porous Ni-YSZ (nickel yttria-stabilized zirconia) composite, as one of the recent fuel cells, is considered in this example, which provides excellent conductivity, high-temperature elastic properties, and high thermal conductivity [51–53]. Various factors in these samples, including volume fraction, average aggregate size, and the texture of phases, can significantly change their properties. One such sample using an image with the size of 425 × 686 pixels is shown in Fig. 3(b).

Fifty realizations are generated using the proposed method in this paper. Three of them are shown in Fig. 3(b). The CPU time for this modeling was 2 s. Visual inspection shows that the generated realizations represent an excellent agreement, in terms of structures and morphologies, with the provided $\mathbf{I}$ in Fig. 3(a). The discussed statistical tests are used to quantify such a similarity and the results of each of those functions are presented in Fig. 4. All the calculated similarity functions represent an excellent agreement with the original $\mathbf{I}$. As mentioned earlier, the MPC demonstrates a better representation of the existing variability between the generated models. Thus, one can ensure that the produced realizations have enough

diversity and they can be used to cover the uncertainty and possible structures.

B. Multiphase polymetric materials

Understanding of spin coating of multiphase polymeric materials is one of the long-standing problems in materials science. Such materials commonly contain various morphological patterns that represent self-organized subsets and regular micrometer features. Spin coating polymer of thin films, indeed, can produce multiphase layers of high uniformity on large surface areas. These materials can be used in a variety of fields such as photovoltaics and light-emitting diodes. A stroboscopically illuminated image during spin-coating of a PS:PI blend onto a silicon substrate at 1500 rpm is shown in Fig. 5(a) [55]. This image contains 466 × 642 pixels. This image is a challenging **I** for the reconstruction methods as it contains various curvilinear, complex, and continuous structures.

For the sake of comparison and demonstrating the ability of the proposed method for removing the artifacts and discontinuities, one realization with and without considering the discussed steps (Secs. II B to II F) is shown in Figs. 5(c) and 5(d). Those steps are inherently part of the proposed method. As can be seen, the main difference here is the removal of some small artifacts and decreasing the CPU time, while the realization largely remains unchanged. The artifacts are identified through the utilized similarity function in the algorithm. This function, indeed, can give the differences between each of the two patterns. Therefore, the artifacts can be found by their higher differences. Besides, the realization has not changed considerably, whereas the artifacts are not visible in the new realization shown in Fig. 5(d). The CPU time for the original realization is 60 s, while the computation time for the right one is around 3 s.

Several realizations are generated using the proposed method in this paper and three of them are shown in Fig. 5(b). The computational time for this modeling was 3 s. As can be seen, the generated realizations contain the same heterogeneity and complex structures as in the **I**. The variability between the realizations and **I** can be further verified using the results of similarity functions in Fig. 6. Once again, the generated realizations manifest an excellent agreement in accordance with the original **I**.

C. Nonstationary oxygen-free copper

Electron backscatter diffraction (EBSD) is one of the methods being used for identifying the deformation caused by corrosion in copper canisters. Such canisters contain different inhomogeneous structures and discontinuous welds that can define the places with a higher concentration of stress. These regions can significantly change the deformation and damage tolerance. In this section, a horizontal cross section of an electron beam weld is used [56], which demonstrates a nonstationary system [see Fig. 7(a)]. This image is composed of a grid with 312 × 437 pixels.

As discussed earlier, modeling of nonstationary systems requires a different strategy. In this example, the partitioning method is applied. Thus, the **I** is divided into three stationary

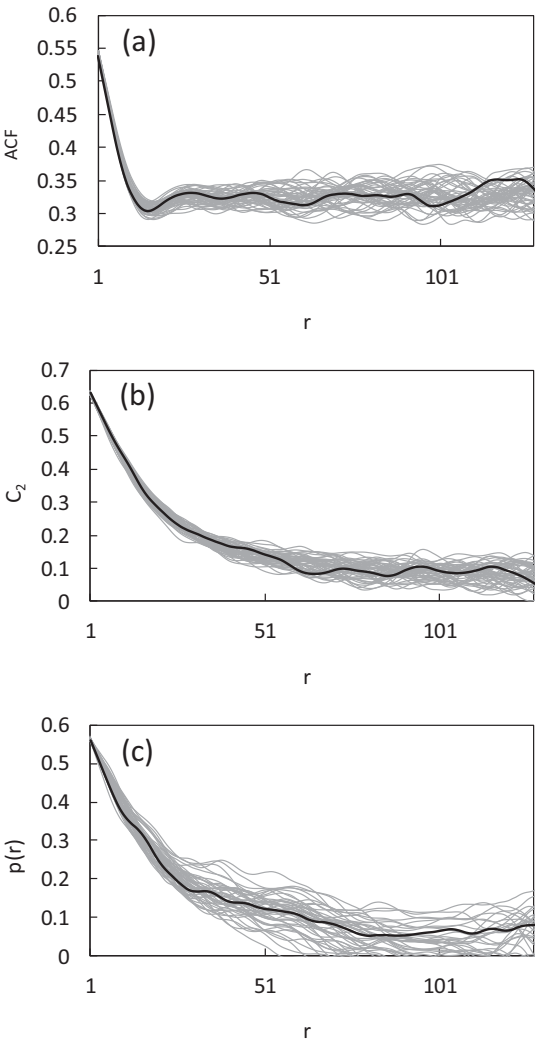

FIG. 6. Same as in Fig. 4, but for the multiphase polymetric material.

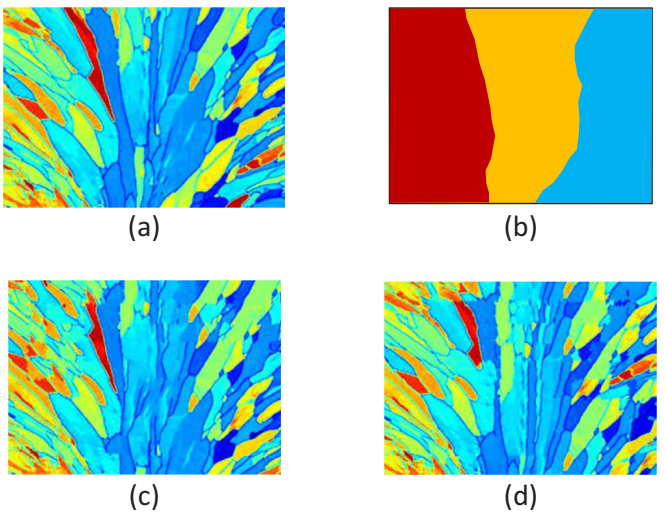

FIG. 7. (a) Modeling of the nonstationary oxygen-free copper sample. The original **I** in (a) is partitioned into three distinct regions (b) based on the orientations of phases. (c), (d) Two of the generated realizations are shown.

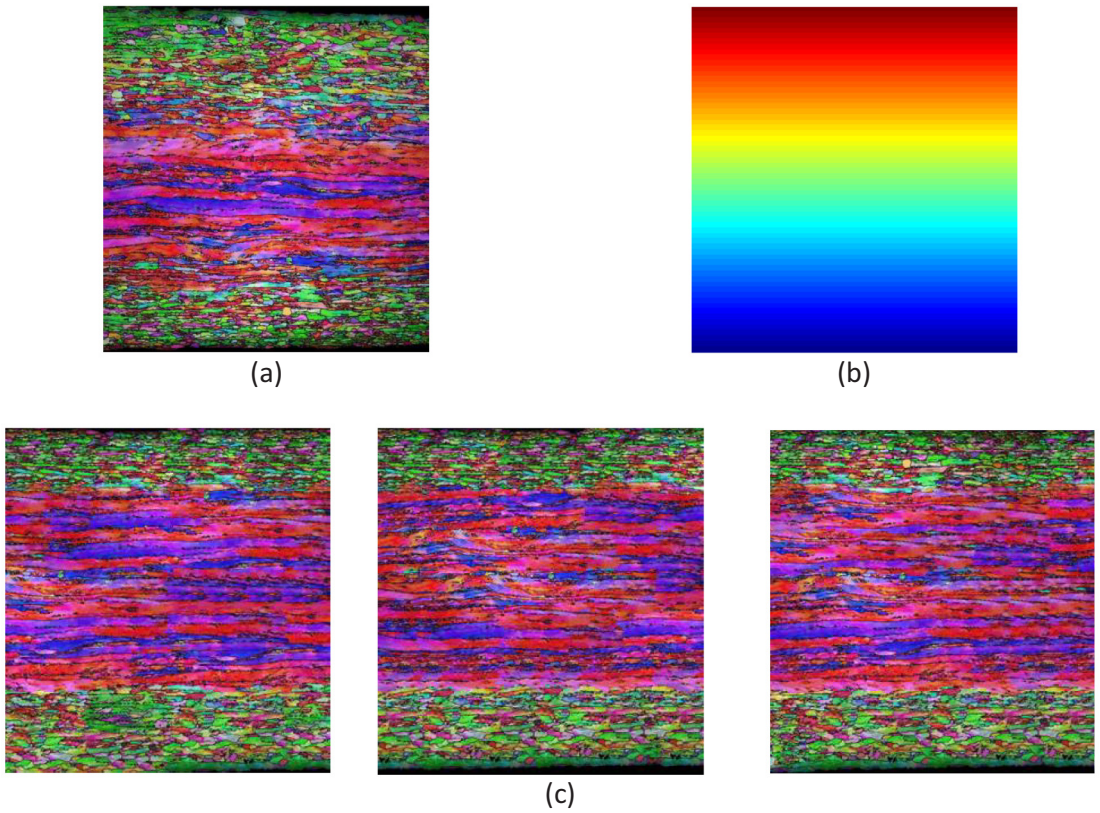

FIG. 8. (a) Modeling of nonstationary ferritic steel sample using the provided **AUX** data in (b). (c) Three of the generated realizations.

parts as shown in Fig. 7(b), which is only based on user interpretation of the observed structures in **I**. For example, one can notice that the structures represent three distinct regions; thus, assigning three polygons to separate such a behavior can help identify such regions. The modeling of each realization took 3 s. For each of these regions, their corresponding areas in **I** are used. For example, the structures in the right part not allowed to be used for the left side. Furthermore, as can be seen from **I** in Fig. 7(a), the orientations in each of these regions are different. Several realizations are generated based on this partitioning and two of them are shown in Figs. 7(c) and 7(d).

D. Nonstationary ferritic steel

This example represents another nonstationary system derived from the EBSD for ferritic steel [57]. This image reveals information about the crystallographic orientation and phase characteristics of the illuminated crystals [see Fig. 8(a)]. The EBSD manifests a clear nonstationary distribution of the middle phase around the sample, while the concentration in the central part is significant. This region contains more parallel layers, while the upper and lower margins represent a random distribution of other phases. Thus, AUX data, for representing such a behavior, as shown in Fig. 8(b), is utilized. The AUX data represent the vertical distribution of different phases. In this example, the best way to represent the nonstationarity is to use a vertical trend map, shown in Fig. 8(b), which can be created manually as the three main regions, depicted in the AUX data, are perceptible.

Various conditional realizations using the provided **I** and AUX data are generated and three of them are shown in Fig. 8(c). These realizations represent structure and texture similar to that of the **I**, while, for example, the parallel structures are only laid in the central part. The similarity between the realizations and original **I**, for the central structures, is quantified and shown in Fig. 9. It should be noted that the size of this modeling is 306 $\times$ 288 and the computations took 2 s.

E. Open-cell copper foam (3D)

Metallic foams manifest a high strength-to-density ratio and excellent properties of thermal and sound absorption. Such materials have wide applications in heat and mass transfer [58,59] and filtering for separation [60]. One way of producing efficient foam materials is through the methods in powder metallurgical processes in which space holder materials are generated by synthesizing open-cell metallic foams [61]. This process allows control of structures and physical properties.

Using different shapes of pores and repeating the time-demanding experiments for evaluating the mechanical properties is not, indeed, plausible. Thus, imaging techniques provide an alternative by which the costly step of physical experiments can be avoided. Providing various 3D images with different

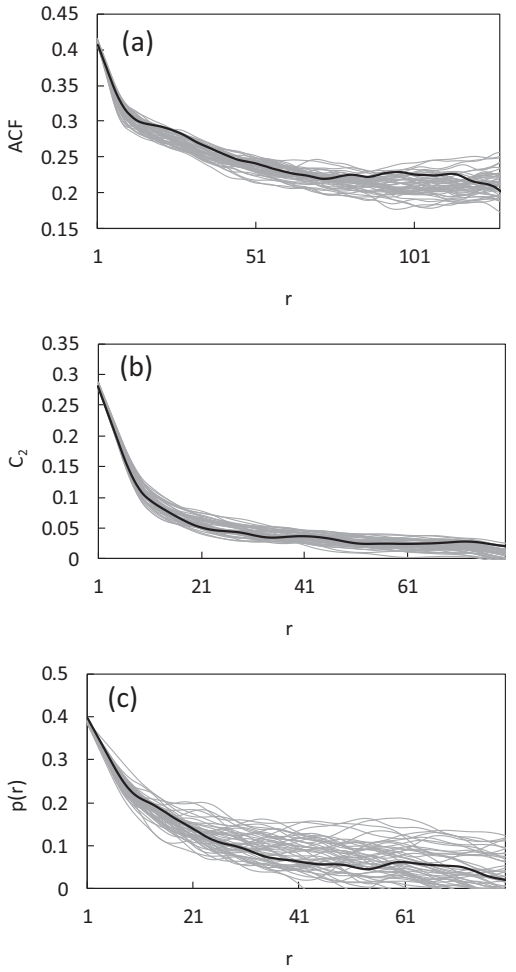

FIG. 9. Same as in Fig. 4, but for the nonstationary ferritic steel.

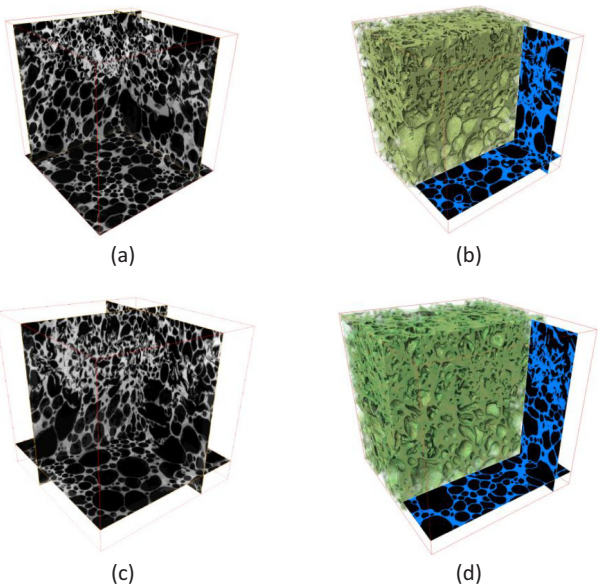

FIG. 10. (a), (b) Stochastic modeling of complex micro-CT image of copper foam sample. (c), (d) One of the generated realizations.

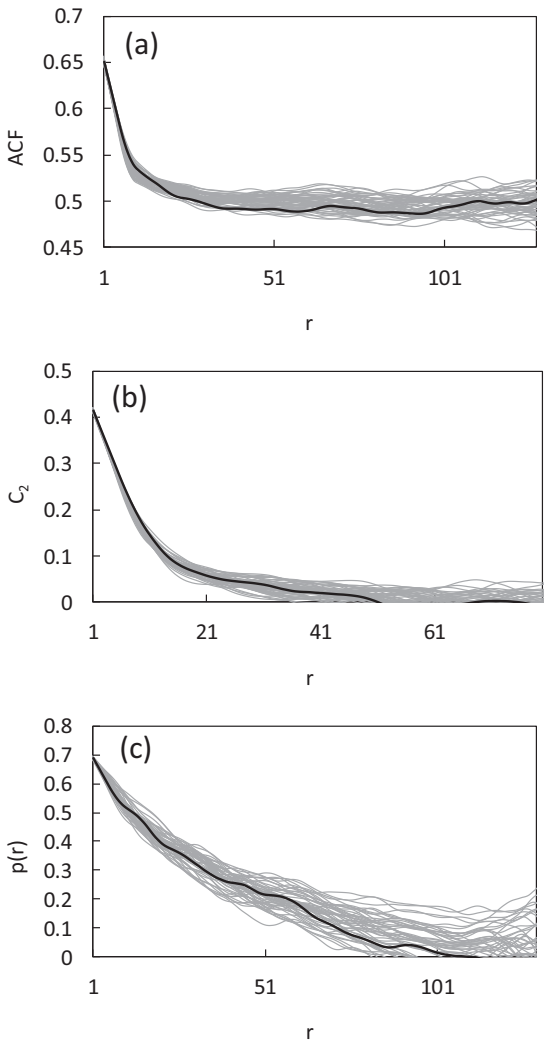

FIG. 11. Same as in Fig. 4, but for the open-cell copper foam.

possible pore-size distribution and required spatial variability is, however, very costly and tedious. Thus, an initial copper foam image produced by x-ray microtomography is considered as **I** [62,63] and several realizations, which result in different pore sizes, are generated using the proposed method in this paper [see Figs. 10(a) and 10(b)]. The size of this image is $350\times350\times225$ voxels. Accordingly, one does not need to take various 3D images under different conditions as the proposed method can produce most of such variabilities using a single image.

Several stochastic foam models are generated. One such realization is shown in Figs. 10(c) and 10(d) and took 500 s. Visually speaking, the produced realization exhibits the same pore structures as the **I** and the pore geometries are well reproduced. Then, the similarity functions are applied on **I** and the generated realizations to quantify the successful reproduction of topological features. The results are shown in Fig. 11 and they all are in excellent agreement with the original **I**. Thus, the generated 3D structure of foams can be used directly in computational mechanical methods such as the finite element method (FEM). The results of this example can provide an exhaustive understanding of the mechanical

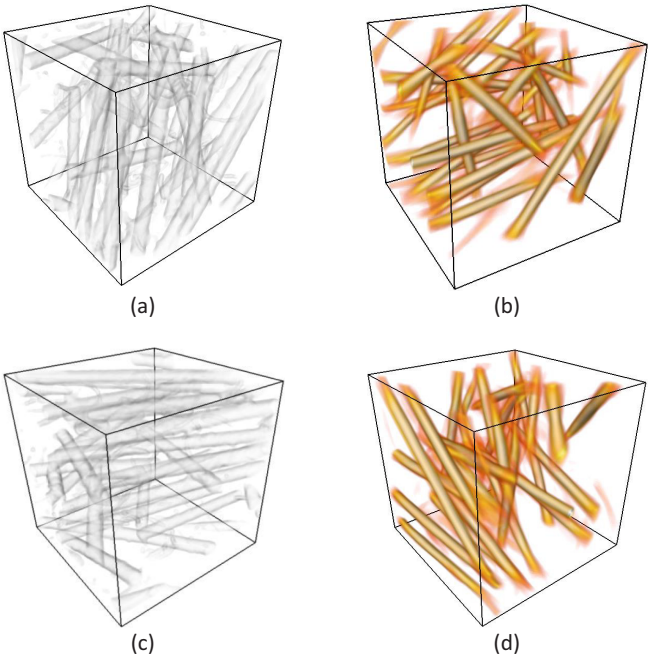

FIG. 12. Stochastic modeling of fiber concrete material. (a), (b) The original **I**, which contains long-range and connected fibers. (c), (d) One stochastic model is also provided.

properties of these complex samples when the size and shape of pore spaces vary.

F. Ultrahigh-performance fiber-reinforced materials

The structure of concrete is one of the most important factors that define its mechanical and physical properties, as well as transport phenomena, which all control the durability of these important materials. Thus, investigating such important structures using x-ray microtomography can provide invaluable insight into the measurable properties of different structures [64,65]. These images are provided using nondestructive synchrotron-based x-ray computed microtomography.

A sample with highly connected structures is selected in this section that can be thought of as a complicated fiber material [see Figs. 12(a) and 12(b)]. The size of this sample is $100\times100\times100$ voxels. Reproducing of long fibers for these materials is very important. Thus, the DT method, discussed in Sec. II C, was utilized extensively in this sample and the target fibers are stressed out, assisting the algorithm to reproduce them better. Various realizations for this 3D sample are generated and one of them is shown in Figs. 12(c) and 12(d). The computational time for this sample was 70 s. The produced realization represents very similar fibers in the concrete sample. The similarity is further quantified in Fig. 13.

Due to dealing with a very complex and relatively large sample size in this example, a sensitivity analysis is carried out for various **I** sizes. The results are shown in Fig. 14. As can be seen, the proposed method provides tremendous

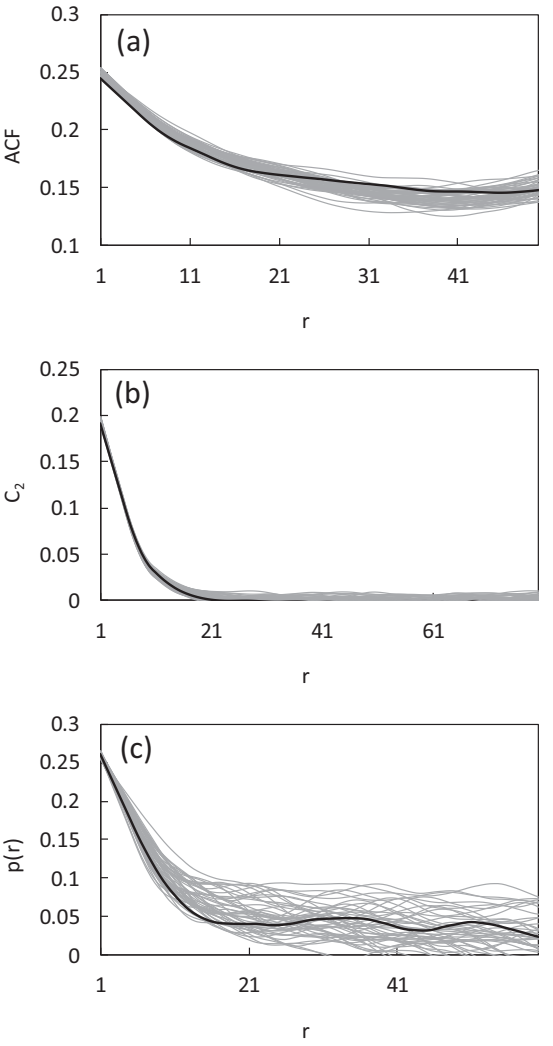

FIG. 13. Same as in Fig. 4, but for the fiber concrete material.

acceleration even when the initial input is very large. The reduction is, indeed, the result of the implemented multiscale method.

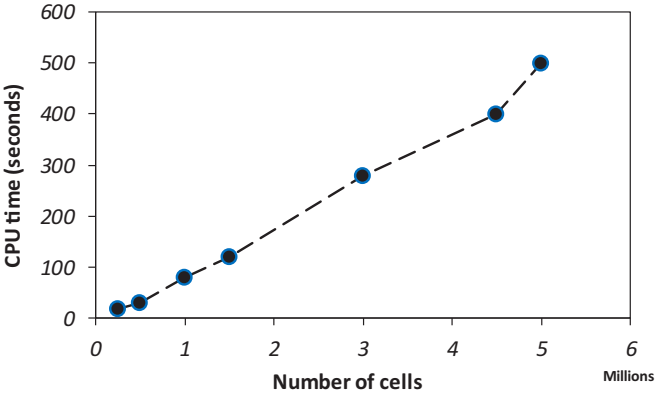

FIG. 14. CPU-time comparison for different **I** sizes.

IV. SUMMARY

A stochastic method based on a successive calculation of conditional probability was devised in this paper. For producing high-quality and realistic models of complex microstructures, several algorithms were integrated and employed. The method takes into account the complexity, histogram reproduction, and continuity among the structures, as well as the multiscale structures that exist in the microstructure materials. The algorithm is capable of producing models that represent both large- and small-scale features. The method was tested on several complicated systems of materials with the stationary and nonstationary distribution of phases. Two methods for dealing with nonstationary materials were also discussed. In summary, the aspects of the proposed method in this paper are listed as follows:

(1) The proposed method represents a multicomponent algorithm that can simulate materials with any level of complexity. Through using different strategies, such as graph theory, histogram matching, the iterative algorithm, and distance transform, one can produce high-quality realizations.

(2) Two alternative methods for dealing with nonstationary systems were presented in the paper that can handle most of the complex and nonstationary materials. None of the previous methods can simulate such complex nonstationary microstructures.

(3) The proposed multiscale method provides a very efficient framework for modeling large images of complex and multiphase materials. The method can produce microstructures with millions of pixels in a matter of seconds.

Thus, by producing various realistic realizations, the proposed method can be used readily for composites and alloys with complex structures. In other words, one is not requires to spend a significant amount of money and time to take 2D or 3D images under diverse situations as the proposed method can computationally produce such models.

ACKNOWLEDGMENTS

The author would like to thank the University of Wyoming for the financial support for this research. The significant help of Dr. Saadatfar from the Australian National University for donating the foam data is also appreciated. Finally, the constructive comments from anonymous referees, which helped to improve the initial submission, are also highly acknowledged.

[1] M. Sahimi, *Flow and Transport in Porous Media and Fractured Rock* (Wiley-VCH, Weinheim, 2011).
[2] S. Torquato, *Random Heterogeneous Materials* (Springer, New York, 2002).
[3] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001).
[4] G. T. Herman, *Fundamentals of Computerized Tomography* (Springer, London, 2009).
[5] D. G. Brandon and W. D. Kaplan, *Microstructural Characterization of Materials* (Wiley, New York, 2008).
[6] P. Tahmasebi, M. Sahimi, A. H. Kohanpur, and A. Valocchi, J. Pet. Sci. Eng. **155**, 21 (2016).
[7] J. Baruchel, P. Bleuet, A. Bravin, P. Coan, E. Lima, A. Madsen, W. Ludwig, P. Pernot, and J. Susini, C. R. Phys. **9**, 624 (2008).
[8] J. H. Kinney and M. C. Nichols, Annu. Rev. Mater. Sci. **22**, 121 (1992).
[9] E. Padilla, V. Jakkali, L. Jiang, and N. Chawla, Acta Mater. **60**, 4017 (2012).
[10] N. Chawla, J. J. Williams, X. Deng, C. McClimon, L. Hunter, and S. H. Lau, Int. J. Powder Metall. (Princeton, N. J.) **45**, 19 (2009).
[11] L. Babout, E. Maire, J. Y. Buffière, and R. Fougères, Acta Mater. **49**, 2055 (2001).
[12] A. Borbély, F. Csikor, S. Zabler, P. Cloetens, and H. Biermann, Mater. Sci. Eng. A **367**, 40 (2004).
[13] P. Kenesei, H. Biermann, and A. Borbély, Scr. Mater. **53**, 787 (2005).
[14] J. J. Williams, Z. Flom, A. A. Amell, N. Chawla, X. Xiao, and F. De Carlo, Acta Mater. **58**, 6194 (2010).
[15] A. Weck, D. S. Wilkinson, E. Maire, and H. Toda, Acta Mater. **56**, 2919 (2008).
[16] H. Toda, S. Yamamoto, M. Kobayashi, K. Uesugi, and H. Zhang, Acta Mater. **56**, 6027 (2008).
[17] M. Y. Wang, J. J. Williams, L. Jiang, F. De Carlo, T. Jing, and N. Chawla, Scr. Mater. **65**, 855 (2011).
[18] J. J. Williams, K. E. Yazzie, N. Connor Phillips, N. Chawla, X. Xiao, F. De Carlo, N. Iyyer, and M. Kittur, Metall. Mater. Trans. A **42**, 3845 (2011).
[19] M. Y. Wang, J. J. Williams, L. Jiang, F. De Carlo, T. Jing, and N. Chawla, Metallogr. Microstruct. Anal. **1**, 7 (2012).
[20] D. Westhoff, J. J. van Franeker, T. Brereton, D. P. Kroese, R. A. J. Janssen, and V. Schmidt, Model. Simul. Mater. Sci. Eng. **23**, 045003 (2015).
[21] P. Tahmasebi and M. Sahimi, Phys. Rev. E **85**, 066709 (2012).
[22] M. Sahimi, *Heterogeneous Materials I* (Springer-Verlag, New York, 2003).
[23] R. Bostanabad, A. T. Bui, W. Xie, D. W. Apley, and W. Chen, Acta Mater. **103**, 89 (2016).
[24] S. Torquato, Appl. Mech. Rev. **44**, 37 (1991).
[25] D. C. Pham and S. Torquato, J. Appl. Phys. **94**, 6591 (2003).
[26] B. Lu and S. Torquato, Phys. Rev. A **45**, 922 (1992).
[27] S. R. Kalidindi, S. R. Niezgoda, and A. A. Salem, JOM **63**, 34 (2011).
[28] A. Roberts, Phys. Rev. E **56**, 3203 (1997).
[29] D. T. Fullwood, S. R. Niezgoda, and S. R. Kalidindi, Acta Mater. **56**, 942 (2008).
[30] S. R. Niezgoda, D. M. Turner, D. T. Fullwood, and S. R. Kalidindi, Acta Mater. **58**, 4432 (2010).
[31] C. L. Y. Yeong and S. Torquato, Phys. Rev. E **57**, 495 (1998).
[32] M. G. Rozman and M. Utz, Phys. Rev. Lett. **89**, 135501 (2002).
[33] Y. Jiao, F. H. Stillinger, and S. Torquato, Proc. Natl. Acad. Sci. USA **106**, 17634 (2009).
[34] Y. Jiao, F. H. Stillinger, and S. Torquato, Phys. Rev. E **77**, 031135 (2008).
[35] Y. Jiao, E. Padilla, and N. Chawla, Acta Mater. **61**, 3370 (2013).

[36] S. S. Singh, J. J. Williams, Y. Jiao, and N. Chawla, Metall. Mater. Trans. A **43**, 4470 (2012).
[37] P. Tahmasebi and M. Sahimi, Phys. Rev. Lett. **110**, 078002 (2013).
[38] S. Chen, A. Kirubanandham, N. Chawla, and Y. Jiao, Metall. Mater. Trans. A **47**, 1440 (2016).
[39] P. Tahmasebi and M. Sahimi, Phys. Rev. E **91**, 032401 (2015).
[40] P. Tahmasebi, M. Sahimi, and J. E. Andrade, Geophys. Res. Lett. **44**, 4738 (2017).
[41] Y. Boykov, O. Veksler, and R. Zabih, IEEE Trans. Pattern Anal. Mach. Intell. **23**, 1222 (2001).
[42] P. Tahmasebi and M. Sahimi, Water Resour. Res. **52**, 2074 (2016).
[43] R. Fabbri, L. D. F. Costa, J. C. Torelli, and O. M. Bruno, ACM Comput. Surv. **40**, 1 (2008).
[44] A. Rosenfeld and J. Pfaltz, Pattern Recognit. **1**, 33 (1968).
[45] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 2006).
[46] D. M. Endres and J. E. Schindelin, IEEE Trans. Inf. Theory **49**, 1858 (2003).
[47] P. Tahmasebi and M. Sahimi, Water Resour. Res. **52**, 2099 (2016).
[48] P. Tahmasebi, Water Resour. Res. **53**, 5980 (2017).
[49] G. J. Mahon and J. M. Howe, Metall. Trans. A **21**, 1655 (1990).
[50] K. Kendall, Nat. Mater. **1**, 211 (2002).
[51] X. Zhang, S. H. Chan, G. Li, H. K. Ho, J. Li, and Z. Feng, J. Power Sources **195**, 685 (2010).
[52] W. Zhu and S. Deevi, Mater. Sci. Eng. A **362**, 228 (2003).
[53] S. A. Tabei, A. Sheidaei, M. Baniassadi, F. Pourboghrat, and H. Garmestani, J. Power Sources **235**, 74 (2013).
[54] M. Baniassadi, H. Garmestani, D. S. Li, S. Ahzi, M. Khaleel, and X. Sun, Acta Mater. **59**, 30 (2011).
[55] S. Ebbens, R. Hodgkinson, A. J. Parnell, A. Dunbar, S. J. Martin, P. D. Topham, N. Clarke, and J. R. Howse, ACS Nano **5**, 5124 (2011).
[56] K. Savolainen, T. Saukkonen, and H. Hänninen, World J. Nucl. Sci. Technol. **2**, 16 (2012).
[57] D. Raabe, D. Ponge, O. Dmitrieva, and B. Sander, Scr. Mater. **60**, 1141 (2009).
[58] W. Yuan, Y. Tang, X. Yang, and Z. Wan, Applied Energy **94**, 309 (2012).
[59] H. Tawfik, Y. Hung, and D. Mahajan, J. Power Sources **163**, 755 (2007).
[60] E. Vanhaecke, C. Pham-Huu, and D. Edouard, Catalysis Today **189**, 101 (2012).
[61] J. Banhart, Prog. Mater. Sci. **46**, 559 (2001).
[62] M. Saadatfar, F. Garcia-Moreno, S. Hutzler, A. P. Sheppard, M. A. Knackstedt, J. Banhart, and D. Weaire, Colloids Surf. A **344**, 107 (2009).
[63] A. M. Parvanian, M. Saadatfar, M. Panjepour, A. Kingston, and A. P. Sheppard, Mater. Des. **53**, 681 (2014).
[64] S. Lu, E. N. Landis, and D. T. Keane, Mater. Struct. **39**, 611 (2006).
[65] M. A. B. Promentilla, T. Sugiyama, T. Hitomi, and N. Takeda, Cem. Concr. Res. **39**, 548 (2009).