Density of States for a Specified Correlation Function and the Energy Landscape

C. J. Gommes,[1,2,*] Y. Jiao,[2,†] and S. Torquato[2,3,‡]

[1]*Department of Chemical Engineering, University of Liège, B-4000 Liège, Belgium*
[2]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[3]*Program in Applied and Computational Mathematics and Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*

(Received 21 October 2011; published 22 February 2012)

The degeneracy of two-phase disordered microstructures consistent with a specified correlation function is analyzed by mapping it to a ground-state degeneracy. We determine for the first time the associated density of states via a Monte Carlo algorithm. Our results are explained in terms of the roughness of an energy landscape, defined on a hypercubic configuration space. The use of a Hamming distance in this space enables us to define a roughness metric, which is calculated from the correlation function alone and related quantitatively to the structural degeneracy. This relation is validated for a wide variety of disordered structures.

DOI: 10.1103/PhysRevLett.108.080601 PACS numbers: 05.20.−y, 61.43.−j

Spatial correlation functions are fundamental descriptors that arise in a variety of disciplines, including condensed matter physics [1], geostatistics [2], computer vision and image analysis [3], statistical physics [4], and materials science [5,6]. They notably provide a very general tool for characterizing the microstructure of materials and relating this information to their physical properties [5]. Moreover, most experimental techniques available for *in situ* studies with a nanometer resolution—notably scattering methods—yield information in the form of two-point correlation functions [7–10].

It is well known that standard two-point statistics are generally not sufficient to characterize fully a microstructure [11]. This is referred to as the phase problem in crystallography. The ambiguity of two-point information has been investigated theoretically from the perspective of crystallography [12,13], computer vision [14], materials science [15], and cosmology [15]. In some very special cases, distinct microstructures with identical correlation functions can be derived analytically [16,17]. It has also been shown that the structural ambiguity is considerably larger for a radial function without orientation information, which is the only data available from small-angle scattering experiments. However, if orientation information is employed, accurate microstructure reconstructions can often be obtained [18–21].

In this Letter, we determine for the first time a general means to calculate numerically the number of microstructures consistent with *any* specified correlation function for arbitrary systems in any Euclidean space dimension. For concreteness, the present analysis focuses on two-phase microstructures, which are suitable models for a host of natural and synthetic materials such as composites [5,6], colloidal suspensions and microemulsions [22], porous materials [23], etc. Moreover, the general method is applied to the radial two-point correlation function $S_2(r)$ defined as the probability that two random points at distance r from one another both belong to a given phase. Although the present work focuses on two-dimensional media, our procedure can be applied in any space dimensions.

The general methodology consists in mapping the determination of the microstructure ambiguity to the determination of a ground-state degeneracy. This is achieved in the general framework of the reconstruction methods, which aim at producing microstructures consistent with a target correlation function $\hat{S}_2(r)$. In that context, a microstructure having a correlation function $S_2(r)$ is associated with an ''energy'' defined as [24]

$$E = \sum_r [\hat{S}_2(r) - S_2(r)]^2, \tag{1}$$

which is equivalent to a norm-2 error. All the microstructures consistent with a specified $\hat{S}_2(r)$ are the ground states (with globally minimized energy) of the corresponding reconstruction problem. Examples of degenerate microstructures, obtained via simulated annealing [24,25], are given in Fig. 1 for correlation functions typical of hard-disk and polycrystal microstructures. In both cases the two displayed microstructures have the same two-point correlation function to within an error of 10^{-7}, but their higher-order correlation functions differ.

Using the language of solid-state physics, we refer to the number $\Omega(E)$ of microstructures having energy E as the density of states (DOS). An efficient method for estimating the DOS has been proposed by Wang and Landau [26,27], and applied to a wide variety of problems ranging from solid-state physics [28] to biophysics [29] and logic [30]. The algorithm is based on the observation that a Monte Carlo (MC) move from state i to j with acceptance probability

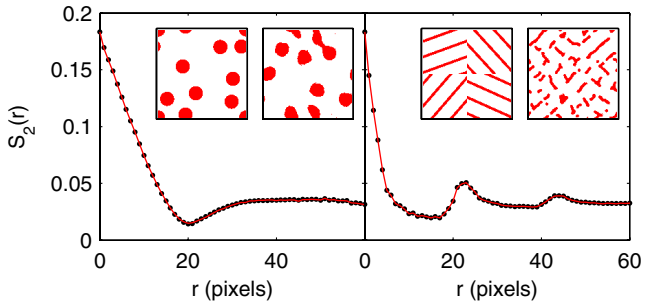

FIG. 1 (color online). Degenerate microstructures corresponding to a hard-disk (left) and a polycrystal (right) two-point correlation function: in each case, the correlation functions of the two degenerate microstructures (● and solid line) are indistinguishable on the scale of the figure.

$$p_{i\to j} = \min\{1, \Omega(E_i)/\Omega(E_j)\} \tag{2}$$

would lead the system to visit all energies with the same probability. Because $\Omega(E)$ is unknown, the DOS is initialized to $\Omega(E) = 1$ for all energies, and the MC algorithm updates this value until convergence is achieved.

We restrict the discussion to discrete two-phase microstructures, which can be thought of as binarized images composed of black and white pixels. Starting from an initial configuration with energy E_i, a black pixel is moved randomly to an unoccupied white pixel. The correlation function is updated, the new energy E_j is calculated through Eq. (1), and the move is accepted or rejected according to Eq. (2). Each time a given energy is visited, a histogram is updated, i.e., $H(E) \to H(E) + 1$, and the estimated density of states is updated according to $\Omega(E) \to F \times \Omega(E)$, where F is a numerical factor larger than 1. The evolution continues with the updated value of $\Omega(E)$ until the histogram $H(E)$ is flat. At this point $H(E)$ is reset to 0, F is reduced to $\sqrt{F}$, and the evolution starts over again. The entire procedure is repeated until F becomes lower than a prescribed accuracy.

Examples of DOS calculations are given in Fig. 2. The results are given in terms of the cumulative DOS,

$$N_\Omega(E) = \sum_{e\le E} \Omega(e), \tag{3}$$

which has been normalized for $N_\Omega(E \to \infty)$ to be equal to the total number of configurations $\Omega_{\text{tot}} = \binom{N}{N_1}$, with N the total number of pixels and N_1 the number of black pixels. The limit $N_\Omega(E \to 0)$ is the ground-state degeneracy Ω_0, i.e., the total number of microstructures consistent with $\hat{S}_2(r)$.

The "square" (a) and "sphere" (s) microstructures in Fig. 2 are uniquely specified by their correlation functions. The degeneracy Ω_0 is therefore only a trivial contribution resulting from the 64 possible translations on a 8×8 grid. The values found from the MC estimations are 66 ± 7 and 58 ± 8, for (a) and (s), respectively, with the error estimated from three independent runs. Configurations

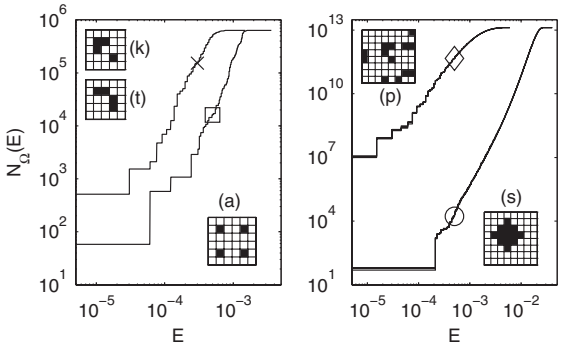

FIG. 2. Cumulative densities of states for various correlation functions: microstructures (a) and (s) are uniquely determined by their correlation functions (□ and ○); microstructures (k) and (t) have identical correlation functions (×); microstructure (p) is a realization of a Poisson process (◇). Each curve is the superposition of three independent Monte Carlo runs.

(k) and (t) are the "kite and trapezoid" configurations discussed in a previous paper [16]; they have identical correlation functions, which results in a nontrivial factor 2 to the ground-state degeneracy. Moreover, since the configurations lack rotational symmetry, the possible orientations contribute a trivial factor 4. The ground-state degeneracy is therefore $\Omega_0 = 512$, in excellent agreement with the MC estimate 500 ± 68. Finally, configuration (p) is a realization of a Poisson point process [3] used here as a model of a very disordered morphology. This particular microstructure is found to be extremely degenerate with Ω_0 as large as $(11 \pm 1) \times 10^6$.

It is generally acknowledged that large ground-state degeneracy is a characteristic of systems having a rough energy landscape [31]. Note that the configuration space $\mathcal{C}$

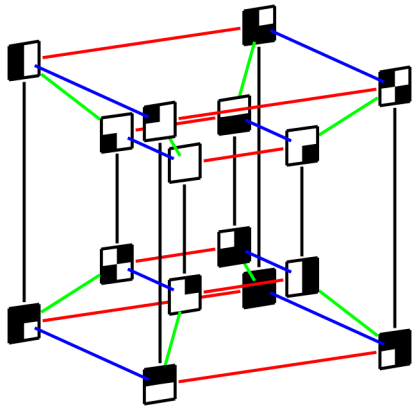

FIG. 3 (color online). The configuration space $\mathcal{C}$ of a two-phase microstructure is a N-dimensional hypercube, on which a Hamming distance is defined. Each dimension is associated with the phase of a particular pixel. For a 2×2 microstructure, $\mathcal{C}$ is a tesseract, represented here in three dimensions, where the fourth dimension is indicated by the edges joining the inner and outer cubes.

of discrete two-phase microstructures is the set of vertices of a N-dimensional hypercube (see Fig. 3). This results from the properties of the indicator vector $I(i)$, with components equal to 1 when point i is a black pixel and 0 otherwise, which can be seen as a coordinate in N-dimensional space. Moving the system along a given N-dimensional direction corresponds to swapping a particular pixel between black and white. Once a target correlation function $\hat{S}_2(r)$ is specified, each vertex is associated with an energy E according to Eq. (1).

The roughness of our energy landscape characterizes the spatial variability of E in configuration space. It is meaningful therefore to define a distance in this configuration space. A natural choice is the *Hamming* distance, which counts the number of edges between any two vertices. Interestingly, if the number of black pixels is kept constant, i.e., $\sum I(i) = N_1$, all the realizable microstructures lie on the intersection of the hypercube with a hyperplane. Since the Hamming distance within the hyperplane takes only even values, the distance $d[A, B]$ between two microstructures A and B is defined as half the Hamming distance, i.e.,

$$d[A, B] = \frac{1}{2} \sum_{i=1}^{N} (I_A(i) - I_B(i))^2.$$

In real space, the distance d is the smallest number of black-pixel displacements required to pass from A to B.

To address analytically the question of the roughness of the energy landscape, we use a *random walk in configuration space*, as illustrated in Fig. 4. Starting from a ground state, i.e., a microstructure with $E = 0$, the system makes successive random jumps of length $d = 1$ in configuration space $\mathcal{C}$. When the number of jumps n increases, the random walk explores the configuration space over increasingly larger distances d from the starting microstructure. The rate at which the energy increases with n characterizes the energy landscape.

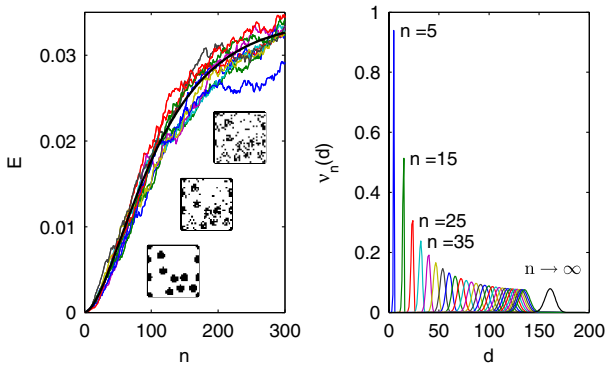

FIG. 4 (color online). Left: Energies visited during a random walk in configuration space, starting from a particular ground state of a hard-disk correlation function. Each irregular line is a particular realization and the thick black line is the average calculated analytically. Examples of microstructures visited for $n = 50$ and $n = 100$ are displayed. Right: Statistical distribution $\nu_n(d)$ of the Hamming distances d from the ground state of the microstructures visited after an increasing number of steps n of the random walk.

An analytical expression is derived in the Supplemental Material [32] for the average energy, and for the statistical distribution of distances to the ground state, both as a function of n. Combining these two pieces of information yields a characteristic energy profile for the basin of any ground state. A detailed analysis is provided elsewhere [33], and we focus here on two particular values on the average energy curve. The first is the average energy $\langle E \rangle$ in the limit $n \to \infty$ and the second is the average energy $E(1)$ reached for $n = 1$. Since random walk is ergodic, $\langle E \rangle$ is equal to the average energy of all Ω_{tot} possible states. The other value, $E(1)$, is a measure of the curvature of the energy landscape near the starting ground state. This results from the observation that $E(1)$ is the average energy of all states at distance $d = 1$ from the ground state, the latter having zero energy. The ratio $E(1)/\langle E \rangle$ therefore provides a metric for the local roughness of the energy landscape.

The complete mathematical expressions for $E(1)$ and $\langle E \rangle$ can be found in the Supplemental Material [32]. The quantity $\langle E \rangle$ is a global characteristic of the energy landscape, which accordingly depends only on $\hat{S}_2(r)$. In contrast, $E(1)$ depends also on the particular ground state used as the starting point for the random walk via a particular function $\sigma_C^2(r)$ with the following structural meaning. Imagine randomly choosing a black pixel in the ground state and drawing a circle of radius r centered on it (or a sphere in three dimensions). The fraction of the circle that overlaps a black pixel is a random variable φ_r that depends on the particular black pixel chosen as the center. The variance of φ_r is the function $\sigma_C^2(r)$, which can be thought of as a generalized "coarseness" [34].

The function $\sigma_C^2(r)$ contains three-point structural information in excess to $\hat{S}_2(r)$. Thus, $\sigma_C^2(r)$ could in principle differ significantly from one ground state to another. In practice, however, the asymptotic behavior of $\sigma_C^2(r)$ for both large and small r can be expressed in terms of $\hat{S}_2(r)$ alone, which enables us to derive a single approximation for $\sigma_C^2(r)$ common to all ground states [32]. Using that approximation yields a single value for $E(1)/\langle E \rangle$, which is a global metric of the roughness of the energy landscape.

Figure 5 shows the quantitative relation between $E(1)/\langle E \rangle$ and a normalized ground-state degeneracy $\Omega_0/\Omega_{\text{tot}}$, for a variety of microstructures defined on a 8×8 grid with the number of black pixels N_1 ranging from 2 to 32 (see Supplemental Material [32]). They comprise a series of structures uniquely characterized by their correlation functions, which have therefore only a trivial degeneracy, as well as a series of nontrivially degenerate realizations of Poisson point processes. In each case, the ground-state degeneracy was estimated via the MC algorithm, and the roughness metric was calculated from

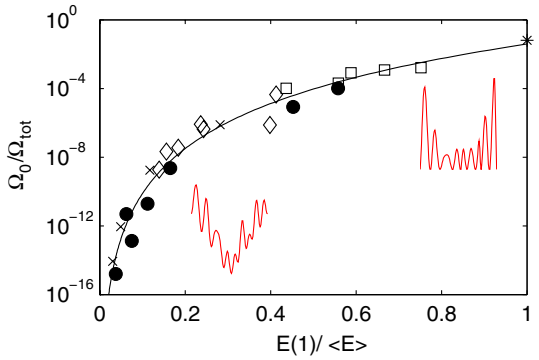

FIG. 5 (color online). Relation between ground-state degeneracy $\Omega_0/\Omega_{\text{tot}}$ and roughness of the energy landscape $E(1)/\langle E\rangle$. The various microstructures are disks of different sizes (●), realizations of Poisson processes (◇), hard-disk systems (×), as well as smaller systems with $N_1 = 4$ (□) and $N_1 = 2$ (*). The black line is a guide to the eye; the insets are sketches of the energy landscape for large and small values of $E(1)/\langle E\rangle$.

$\hat{S}_2(r)$ using the approximation for $\sigma_C^2(r)$. The ratio $\Omega_0/\Omega_{\text{tot}}$ is found to be *highly* correlated with the ratio $E(1)/\langle E\rangle$ over more than 15 orders of magnitude.

When passing from small to large values of $E(1)/\langle E\rangle$, the energy landscape changes qualitatively in the way suggested by the insets in Fig. 5. For low values of $E(1)/\langle E\rangle$, the energy landscape has an overall funnel structure, with low-energy barriers, which makes it well suited for optimization problems. In contrast, for large values of $E(1)/\langle E\rangle$, the landscape is very rough with a large number of ground states. However, it is interesting to note that the rightmost point in Fig. 5 is obtained for a system with $N_1 = 2$ having thus only a trivial degeneracy. The corresponding energy landscape is extremely rough because any possible energy can be found at a distance as short as $d = 1$ from the ground state, but the total number of configurations Ω_{tot} is also extremely small.

The data referred to as disks in Fig. 5 are a collection of nondegenerate microstructures with increasing values of N_1. As N_1 increases, the roughness $E(1)/\langle E\rangle$ decreases, but the degeneracy remains equal to its trivial translation contribution $\Omega_0 = N$. Interestingly, the values of $\Omega_0/\Omega_{\text{tot}}$ of these nondegenerate microstructures span the same curve as the realizations of Poisson processes, for which Ω_0 has a huge nontrivial contribution. The observation that the roughness-degeneracy relation does not discriminate trivial from nontrivial degeneracies suggests ways to determine analytically the degeneracy of large microstructures, which are out of the reach of MC methods [33].

To summarize, the long-standing problem of the degeneracy of microstructures compatible with a specified two-point correlation function can be tackled through the characterization of its associated energy landscape. The results of our MC calculations show that the roughness of the energy landscape is indeed highly correlated with the ground-state degeneracy. The MC algorithm converges only for very small systems [35], but the roughness metric we derived, $E(1)/\langle E\rangle$, can be calculated analytically from $\hat{S}_2(r)$ with no limit of size or dimensionality.

Our results have ramifications in the multitude of fields where correlation functions are useful. In the particular context of materials science, they may contribute to topics as diverse as the understanding of the sharpness of variational bounds for materials properties [5] and the quantitative analysis of scattering experiments [7]. In computer vision, they may help to assess the robustness of second-order texture classifiers [3]. They also have direct applications throughout many subfields of physics because Ω_0 is a lower bound for the physical ground-state degeneracy of any system with pairwise potential energy [33].

Our general methodology can be applied to correlation functions other than $S_2(r)$; this includes higher-order versions [36] as well as cluster correlation functions [15,37]. More generally, random walks in configuration space can be used to derive roughness metrics for the energy landscape of any physical problem, notably protein folding [38], complex chemical reactions [31], phase equilibria in nanopores [39], and glass transition [40]. For instance, the modeling of spin glasses with frustrated Ising models [41] yields an energy that is a linear functional of $S_2(r)$, which is comparatively simpler than the quadratic functional we considered here. We hope to investigate this in future work.

C. J. G. acknowledges support from the Fonds de la Recherche Scientifique (F.R.S.-FNRS, Belgium); S. T. and Y. J. were supported by the Office of Basic Energy Science, Department of Energy, Division of Materials Science and Engineering under Award No. DE-FG02-04-ER46108.

*cedric.gommes@ulg.ac.be
†yjiao@princeton.edu
‡torquato@electron.princeton.edu

[1] R. Zallen, *The Physics of Amorphous Solids* (Wiley, New York, 1983).
[2] J.-P. Chilès and P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty* (Wiley, New York, 1999).
[3] J. Serra, *Image Analysis and Mathematical Morphology* (Academic, London, 1982), Vol. 1.
[4] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
[5] S. Torquato, *Random Heterogeneous Materials* (Springer, New York, 2000).
[6] M. Sahimi, *Heterogeneous Materials I: Linear Transport and Optical Properties* (Springer, New York, 2003), Vol. 1.
[7] L. Feigin and D. Svergun, *Structure Analysis by Small-Angle X-Ray and Neutron Scattering* (Springer, Berlin, 1987).
[8] J. M. Drake, J. Klafter, and P. Levitz, Science **251**, 1574 (1991).

[9] G. A. Barral, L. Frydman, and G. C. Chingas, Science **255**, 714 (1992).
[10] A. Filipponi, A. DiCicco, and C. R. Natoli, Phys. Rev. B **52**, 15 122 (1995).
[11] G. Matheron, *Random Sets and Integral Geometry* (Wiley, New York, 1975).
[12] A. L. Patterson, Phys. Rev. **65**, 195 (1944).
[13] R. Hosemann and S. N. Bagchi, Acta Crystallogr. **7**, 237 (1954).
[14] A. Aubert and D. Jeulin, Pattern Recogn. **33**, 1083 (2000).
[15] Y. Jiao, F. H. Stillinger, and S. Torquato, Proc. Natl. Acad. Sci. U.S.A. **106**, 17 634 (2009).
[16] Y. Jiao, F. H. Stillinger, and S. Torquato, Phys. Rev. E **81**, 011105 (2010).
[17] Y. Jiao, F. H. Stillinger, and S. Torquato, Phys. Rev. E **82**, 011106 (2010).
[18] H. Hauptman, Science **233**, 178 (1986).
[19] M. G. Rozman and M. Utz, Phys. Rev. Lett. **89**, 135501 (2002).
[20] B. L. Hansen, B. L. Adams, M. E. Lyon, and A. Henrie, J. Comput.-Aided Mater. Des. **10**, 163 (2005).
[21] D. T. Fullwood, S. R. Niezgoda, and S. R. Kalidindi, Acta Mater. **56**, 942 (2008).
[22] P.-G. DeGennes, Rev. Mod. Phys. **64**, 645 (1992).
[23] T. J. Barton, L. M. Bull, W. G. Klemperer, D. A. Loy, B. McEnaney, M. Misono, P. A. Monson, G. Pez, G. W. Scherer, J. C. Vartuli, and O. M. Yaghi, Chem. Mater. **11**, 2633 (1999).
[24] C. L. Y. Yeong and S. Torquato, Phys. Rev. E **57**, 495 (1998); **58**, 224 (1998).
[25] Y. Jiao, F. H. Stillinger, and S. Torquato, Phys. Rev. E **77**, 031135 (2008).
[26] F. G. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[27] F. G. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[28] C. Yamaguchi and Y. Okabe, J. Phys. A **34**, 8781 (2001).
[29] N. Rathore and J. J. de Pablo, J. Chem. Phys. **116**, 7225 (2002).
[30] S. Ermon, C. Gomes, and B. Selman, in *Principles and Practice of Constraint Programming CP 2010*, Lecture Notes in Computer Science Vol. 6308, edited by D. Cohen (Springer, Berlin, 2010), pp. 38–52.
[31] D. J. Wales, M. A. Miller, and T. R. Walsh, Nature (London) **394**, 758 (1998).
[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.108.080601 for mathematical details and a list of microstructures used for Fig. 5.
[33] C. J. Gommes, Y. Jiao, and S. Torquato (to be published).
[34] B. L. Lu and S. Torquato, J. Chem. Phys. **93**, 3452 (1990).
[35] P. Dayal, S. Trebst, S. Wessel, D. Wurtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith, Phys. Rev. Lett. **92**, 097201 (2004).
[36] S. Torquato and G. Stell, J. Chem. Phys. **78**, 3262 (1983).
[37] C. Zachary and S. Torquato, Phys. Rev. E **84**, 056102 (2011).
[38] L. L. Chavez, J. N. Onuchic, and C. Clementi, J. Am. Chem. Soc. **126**, 8426 (2004).
[39] J. Puibasset, J. Chem. Phys. **133**, 104701 (2010).
[40] P. G. Debenedetti and F. H. Stillinger, Nature (London) **410**, 259 (2001).
[41] G. Parisi, Proc. Natl. Acad. Sci. U.S.A. **103**, 7948 (2006).