



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Statistical topology of cellular networks in two and three dimensions

J. K. Mason, E. A. Lazar, R. D. MacPherson, and D. J. Srolovitz

Phys. Rev. E **86**, 051128 — Published 26 November 2012

DOI: [10.1103/PhysRevE.86.051128](https://doi.org/10.1103/PhysRevE.86.051128)

Statistical topology of cellular networks in two and three dimensions

J. K. Mason*

Lawrence Livermore National Lab, Livermore, California 94550, USA.

E. A. Lazar[†] and R. D. MacPherson[‡]

School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540, USA.

D. J. Srolovitz[§]

Institute of High Performance Computing, 1 Fusionopolis Way, Singapore 138632.

Cellular networks may be found in a variety of natural contexts, from soap foams to biological tissues to grain boundaries in a polycrystal, and the characterization of these structures is therefore a subject of interest to a range of disciplines. An approach to describe the topology of a cellular network in two and three dimensions is presented. This allows for the quantification of a variety of features of the cellular network, including a quantification of topological disorder and a robust measure of the statistical similarity or difference of a set of structures. The results of this analysis are presented for numerous simulated systems including the Poisson–Voronoi and the steady-state grain growth structures in two and three dimensions.

PACS numbers: 02.40.Pc, 61.72.Mm

I. INTRODUCTION

Disordered structures that appear in a wide range of physical and biological systems often share characteristic geometric and structural features that allow them to be characterized as cellular networks. A cellular network in three dimensions consists of three-dimensional (3d) cells that meet at faces, two-dimensional (2d) faces that meet at edges, and one-dimensional (1d) edges that meet at vertices. This paper considers the common case where two cells meet at a face, three faces meet at an edge, and four edges meet at a vertex (for the analogous stable cellular network in two dimensions, two 2d cells meet at an edge and three 1d edges meet at a vertex). This type of cellular structure has been observed in the context of leaf areoles [1], epithelial cells [2, 3], crack patterns [4], patterned ground [5], Martian terrain [6], polycrystalline materials [7, 8] and foams [9, 10].

Although much work has been done to identify a unifying set of principles to explain the structure and behavior of cellular networks, many fundamental questions remain unresolved. Comparisons of different cellular networks are often practically restricted to scalar quantities such as the average cell size. While more detailed analyses of cellular networks have been performed (e.g., distribution of cell sizes or number of faces per cell), these do not fully characterize the cellular network—the remaining degrees of freedom still allow for considerable variations in network structure. This raises a question nearly as fundamental as the notion of a cellular network itself,

namely, “how can we rigorously identify the similarities or differences among a set of networks?”

The purpose of this study is to develop a statistical description of the topology of cellular networks. For specificity, we occasionally use language appropriate to the study of a grain boundary network in a single phase, polycrystalline material. Apart from the importance of grain boundary network structure for intergranular [11, 12] and transgranular [13, 14] processes, the grain boundary network topology is of special interest because of its sensitivity to variations in interface thermodynamics and to a plethora of deformation mechanisms. Improved characterization of grain boundary network topology would be useful in a variety of practical materials science contexts, from the quantification of microstructure variability for given processing conditions to measurements of the fraction of recrystallized material during an anneal.

The most widespread notion relating to grain boundary network topology in the literature is that of nearest neighbors, i.e., the set of grains that share a face with a central grain. The nearest neighbors are a natural subject of study since they have been observed to strongly influence the morphology and evolution of the central grain in 2d. This is evidenced by the Aboav–Weaire “law” [15, 16] that relates the number of edges of a grain to the average number of edges of its nearest neighbors, Lewis’ “law” [17] that relates the area of a grain to the number of its nearest neighbors, and the mathematically exact von Neumann–Mullins law [18, 19] that relates the rate of change of area of a grain to the number of its nearest neighbors. Concentrating on the nearest neighbors limits the available topological information about the microstructure though, since everything about the microstructure beyond the nearest neighbors is ignored. Additionally, these “laws” only hold strictly in two dimensions, while most physically and biologically interesting network structures occur in three dimensions.

* mason47@llnl.gov

† lazar@math.ias.edu

‡ rdm@math.ias.edu

§ srol@ihpc.a-star.edu.sg

Similar approaches have been used to study of jammed packings of hard spheres. For instance, the ‘granocentric’ model [20, 21] numerically predicts the distributions of the number of neighbors, the number of contacts, and the local packing fraction around a sphere in a poly-disperse sphere packing by considering the solid angle subtended by the neighboring particles. These results may be mapped to the study of cellular networks by constructing a tessellation of the space from the locations of the spheres, and offer the opportunity to study the effect of local disorder on the validity of Lewis’ “law” [22, 23]. Nevertheless, there does not appear to be a way to map an arbitrary cellular network to an equivalent sphere packing, meaning that the physical motivation for the granocentric model may not apply to the systems investigated here.

Closely related to these local approaches is the concept of a shell distance, where the distance between two cells is the minimum number of faces that must be crossed to go from the interior of one cell to the interior of the other. The set of cells at shell distance one is the set of nearest neighbors, and the sets of cells at further shell distances identify analogous portions of the cellular structure. Shell distance usually appears in the literature in the context of relating the number of edges of a central cell to the average number of edges of the cells at some shell distance [24–26], or of predicting the probability of a pair of cells with given numbers of sides occurring at some relative shell distance [27, 28]. The analysis of a cellular structure by shell distance is not suitable for our purposes though, since the shell distance does not distinguish between the different cells in a given shell.

The bond distance is complementary to the shell distance, and appeared nearly contemporaneously in the literature [29, 30]. Specifically, the bond distance between two cells is the minimum number of edges that must be followed on a continuous path from the boundary of one cell to the boundary of the other. This gives a more detailed measure of the cellular network topology than the shell distance. Nonetheless, the use of the bond distance has not been widely employed in the literature, appearing only in the context of counting the number of cells at a given bond distance [29] and of predicting the probability of a pair of cells with given numbers of sides occurring at some relative bond distance [30]. Moreover, since bond distances alone do not give sufficient information to reconstruct the cellular structure, analyses of the cellular network topology using bond distance are incomplete.

There is a considerable literature in the materials science community on the topological properties of a microstructure as a whole. Such approaches generally begin with a binary classification of grain boundaries based upon a property; e.g., as susceptible/resistant to corrosion. The analysis of the microstructure proceeds by considering clusters of susceptible boundaries, with the expectation that the material response changes dramatically when the maximum cluster size is comparable to the sample dimensions. More explicitly, this percolation

theory approach [31–34] asks a question like “is there a percolating cluster of contiguous boundaries that are susceptible to corrosion?” A recent and related development is the use of homology [35, 36] to analyze more subtle topological features of the grain boundary clusters.

In this paper, we develop a characterization of cellular networks that is complete from the standpoint of the network as a topological entity. That is, our description allows an exact reconstruction of the cellular network, up to a geometric deformation. We assume that the cellular structures are statistically homogeneous (see below) in order to be amenable to statistical analysis. Finally, for a characterization method to be practical, it must be efficiently computable, particularly if the intention is to apply this to large, statistically representative sections of cellular networks. To our knowledge, there are currently no characterization methods of this type, though the availability of one would significantly enhance our ability to rigorously and quantitatively describe cellular networks and the differences between networks formed via different processes.

II. TWO-DIMENSIONAL SWATCHES

We describe the cellular network topology by means of an object we call a *swatch*. The *World English Dictionary* defines a swatch as 1. a sample of cloth; 2. a number of such samples, usually fastened together in book form [37]. This paper defines a swatch as a collection of *strings*, or labels for the vertices, that describes a portion of the cellular network around a given vertex. Similarly, a *cloth* is a suitable collection of swatches that describes the topology of the overall network.

More formally, a swatch of a 2d cellular structure is a set of strings or labels for the vertices around some central vertex, and a set of relations indicating when more than one string is assigned to the same vertex. This description is complete in the sense that the topology of the cellular network may be reconstructed from the swatch by applying the relations to a reference structure, as described in Appendix A. A swatch of a 3d cellular structure is nearly the same, though the reference structure used for the reconstruction is more complicated, as described in Appendix B.

Similar to the analysis of a cellular structure by bond or shell distance, a swatch begins from a central location and gradually extends into the surrounding network. Specifically, a swatch is centered on a vertex known as the *root*. A string encodes one of the shortest paths from the root to a given vertex along the edges of the cellular network. This encoding is performed by using the individual *letters* (i.e., characters) of the string to indicate the direction followed at every vertex along the path. In two dimensions, our encoding scheme assumes that every vertex is connected to exactly three edges.

The construction of a swatch begins with the selection of a root and of the orientation of the normal to the plane

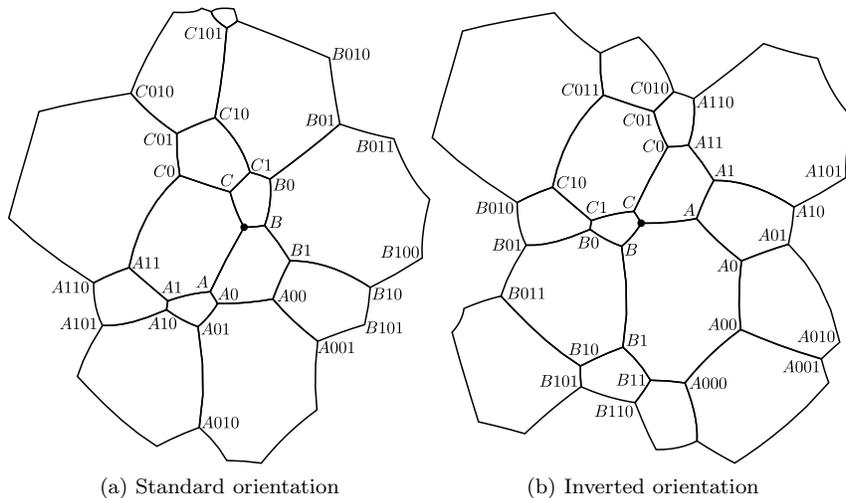


FIG. 1. Sections of the 2d cellular structure described by the swatches in (a) Table Ia and (b) Table Ib, respectively. The swatch assigns strings to the vertices around a root, indicated by the black dots in the center of the figures. The vertex labeled $A0$ in (a) is the root of the region in (b); i.e., the roots of the two swatches are separated by two edges. Additionally, the labeling convention is inverted in (b) relative to (a), corresponding to a switch of the normal vector for the 2d network from out-of-the-page to into-the-page.

0	1	2	3	4
	A	A0	A00 = B11	A001
	B	A1	A01 = A100	A010
	C	B0 = C11	A10 = A011	A101
		B1	A11 = C00	A110
		C0	B01	B010
		C1 = B00	B10	B011
			C01 = C100	B100
			C10 = C011	B101
				C010
				C101

(a) Swatch of Fig. 1a

0	1	2	3	4
	A	A0	A00	A000 = B111
	B	A1	A01 = A100	A001
	C	B0 = C11	A10 = A011	A010
		B1	A11 = C00	A101
		C0	B01	A110
		C1 = B00	B10	B010 = C101
			B11	B011 = B100
			C01	B101
			C10	B110
				C010
				C011 = C100

(b) Swatch of Fig. 1b

TABLE I. Swatches describing the sections of a 2d cellular structure shown in (a) Fig. 1a and (b) Fig. 1b, respectively. The construction is to radius four. The relations between strings, indicated by the equalities, completely specify the swatch and the local topology.

containing the 2d cellular network. For the moment, let the three edges connected to the root lead to three distinct vertices, and let the normal to the plane be oriented out of the page. We label the three vertices connected to the root with the strings A , B and C , arbitrarily selecting the vertex labeled A and assigning the remaining labels in a counter-clockwise manner consistent with the orientation of the normal.

Rather than simply a labeling scheme, a swatch is actually a table composed of the assigned strings, with separate columns for strings containing different numbers of letters. For instance, the swatch of the cellular network in Fig. 1a is given by Table Ia. The root is always labeled

by the null string; this means that there is one entry with no letters in the column to the left. The three strings A , B and C are arranged alphabetically in the column for strings of length one.

The *radius* of a swatch is the length of the longest currently assigned string, or equivalently the number of edges in one of the shortest paths from the root to the furthest labeled vertex (here, distance refers to the number of vertex to vertex steps rather than to a metric distance). We expand the area described by our swatch by continuing the construction to radius two. Notice that while there is a choice of three directions away from the root, a prohibition on retracing any edge means that there is a

choice of two directions at every subsequent vertex. This difference from the root is emphasized by the use of the letters 0 and 1 to indicate the left and right directions, respectively, relative to the most recently traversed edge. That is, the vertex reached by turning left at the end of the path to vertex A is labeled $A0$, and the vertex reached by turning right is labeled $A1$. The situation is similar for the paths emanating from vertices B and C , and the six distinct vertices of this type are labeled and entered in alphabetical order into the column of the table for strings of length two. This procedure may be repeated to continue expanding the swatch.

The most important components of a swatch are the relations between strings, which start appearing for the region in Fig. 1a when the swatch is of radius three. The procedure begins the same as above, with the vertex $A0$ leading to $A00$ and $A01$ and the vertex $A1$ leading to $A10$ and $A11$. Turning left from $B0$ would assign the string $B00$ to the vertex already labeled $C1$ though. Since the lengths of the strings $B00$ and $C1$ do not differ by more than one, the equivalence of these paths is indicated by equating $B00$ with $C1$ in Table Ia. $B00$ is not used as the basis for any longer strings since all vertices reached through $B00$ may be reached by shorter paths through $C1$. After recording this relation, the procedure continues as before with strings $B01$ and $B10$, until we observe that a right turn from $B1$ would assign the string $B11$ to the vertex already labeled $A00$. Since the lengths of the strings $B11$ and $A00$ do not differ by more than one, $B11$ is equated with $A00$ in Table Ia and $B11$ is not used as the basis for any longer strings. The remaining strings of length three are assigned in a similar fashion.

The restriction that relations occur between strings with lengths differing by at most one is due to the fact that this is already sufficient to identify the endpoints of the three edges around a given vertex. Consider Table Ia in isolation, without knowledge of Fig. 1a. From the construction given above, $A0$ must be connected to the three vertices A , $A00$ and $A01$ since these strings correspond to adding or subtracting an edge from the end of the path to $A0$. Similarly, $B0$ must be connected to the three vertices B , $C1$ and $B01$, where the relation $C1 = B00$ is used to obtain the string for the vertex that would have been labeled $B00$. The $A00$ vertex case is more subtle. This vertex is connected to $A0$ and $A001$ for the same reasons as before, but the string $A000$ does not appear anywhere in the table. A relation involving the string $A000$ is not necessary though, since the string $A00$ is equivalent to the string $B11$ and subtracting an edge from the end of the path to $B11$ leads to vertex $B1$. Therefore, the information in the swatch is already sufficient to determine that $A00$ is connected to $A0$, $B1$ and $A001$ without including a relation directly between the strings $A000$ and $B1$.

Occasionally, there may be three paths from the root to a given vertex, with the lengths of the paths differing by at most one. In this situation all three strings are written on the same line of the table describing the swatch, and

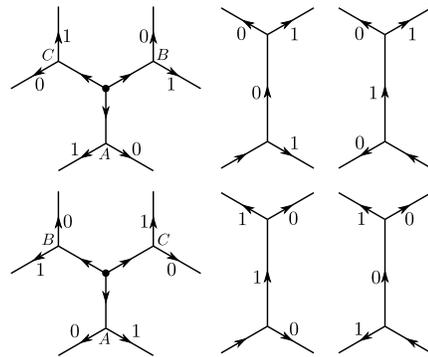


FIG. 2. Conventions for the labeling of vertices in 2d. The top and bottom rows correspond to the normal to the structure being oriented out of and into the page, respectively. Within a given row, the first figure indicates the labeling conventions around the root, while the second and third figures indicate the labeling conventions after a left or a right turn, respectively.

are indicative of a set of three transitive relations. This is the maximum number of strings that may be assigned to any vertex, since a vertex in a 2d cellular structure is connected to exactly three edges.

This completes the construction of a swatch associated with a particular root and choice of labeling of the neighboring vertices, apart from the effect of the orientation of the normal to the cellular network. Orienting the normal out of the page is considered to cause the strings A , B and C to be assigned in a counter-clockwise fashion and the letters 0 and 1 to correspond to the left and right directions, respectively. This set of conventions is reflected in the diagrams at the top of Fig. 2. Orienting the normal into the page reverses the conventions, causing the strings of length one to be assigned in a clockwise fashion and the letters 1 and 0 to correspond to the left and right directions, respectively. This is reflected in the set of diagrams at the bottom of Fig. 2; Fig. 1b and Table Ib show an example of a swatch constructed using these alternative settings. The two choices of normal orientation and three choices of the vertex labeled A allow six swatches to be constructed per root, all describing the same region of the cellular network and therefore regarded as equivalent. Practically speaking, we use a partial ordering of swatches to assign only the minimal swatch to a given root. This is not essential, but provides considerable computational benefits.

The reconstruction of the original 2d network from a swatch table is described, in detail, in Appendix A.

III. THREE-DIMENSIONAL SWATCHES

Although the construction of a swatch and the reconstruction of the cellular network topology in 3d is very similar to that in the 2d case, some complications arise. In particular, our encoding of a path by a string in 2d

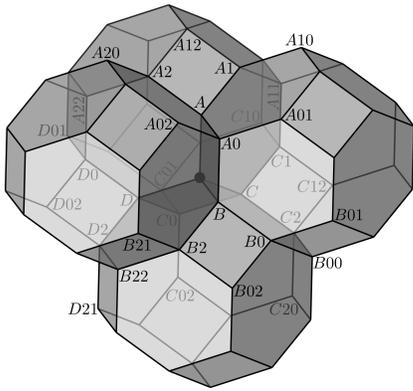


FIG. 3. The 3d region described by the swatch in Table II. The swatch assigns strings to the vertices around a root at the center of the four indicated cells.

relied on the sense of left and right that naturally occurs in the plane. A sensible definition of direction in 3d is more subtle, and is a central subject of this section.

As in the 2d case, a 3d swatch begins from a central vertex or root and gradually extends into the surrounding network. The strings assigned to the surrounding vertices encode the shortest paths from the root to the vertices along the edges of the cellular network. As before, the individual letters of the string indicate the direction followed at every vertex along the path. Our encoding scheme assumes that every vertex is connected to exactly four edges and every edge is connected to exactly three faces—conditions that are satisfied for the 3d cellular structures of interest.

The construction of a swatch begins with the selection of a root and an orientation for the space. For the moment, let the four edges connected to the root lead to four distinct vertices and the orientation of the space be positive. We label the four vertices connected to the root with the strings A , B , C and D , arbitrarily selecting the vertex labeled A . A positively oriented space corresponds to B , C and D being assigned to the remaining three vertices connected to the root in a clockwise fashion around the edge extending from the root to A , with the vertex B selected arbitrarily. The result is consistent with Fig. 3 for which a corresponding swatch is given by Table II. The root is always labeled by the null string and the four strings A , B , C and D are arranged alphabetically in the column for strings of length one.

When continuing the construction of the swatch to radius two, the prohibition on retracing an edge restricts us to a choice of three directions at every vertex apart from the root. We assign these three directions the letters 0, 1 and 2. A positively oriented space corresponds to these letters being assigned in a clockwise fashion to the three edges directed away from the end of the most recently traversed edge, from the perspective of an observer oriented along that edge. Still, this does not specify the setting of the letters; i.e., the choice of a direction to be labeled 0.

0	1	2	3
	A	A0 = B1	A01
	B	A1	A02
	C	A2	A10
	D	B0	A11 = C11
		B2	A12 = A21
		C0 = D1	A20
		C1	A22 = D00
		C2	B00 = C22
		D0	B01
		D2	B02 = B20
			B21
			B22 = D22
			C01
			C02
			C10
			C12 = C21
			C20
			D01
			D02 = D20
			D21

TABLE II. The swatch describing the 3d region in Fig. 3 constructed to radius three. The relations between strings, indicated by the equalities, completely specify the swatch and the local topology.

This ambiguity may be resolved by observing that the set of edges bounding a given face is independent of the orientation of the edges themselves. That is, the setting of the letters 0, 1 and 2 may be consistently assigned on the basis of the connectivity of faces to the edges in question. A set of conventions established along these lines is provided in Fig. 4. The three diagrams on the top right indicate that, for a general vertex, the single edge connected by a face to the two most recently traversed edges is to be labeled by the same letter as the most recently traversed edge. The advantage of this convention is that a set of consecutive and identical letters indicates a path around the same face. Since this convention may only be used after two edges have been traversed, the diagram on the top left is provided to establish the settings in the region around the root. Finally, the diagrams on the bottom row establish an equivalent set of conventions for a negatively oriented space.

With a set of conventions to establish a consistent labeling of the edges directed away from a vertex, the construction of the swatch to radius two proceeds as for the 2d case. The three vertices arrived at by passing through vertex A are labeled $A0$, $A1$ and $A2$, and the three strings are entered alphabetically into the column of Table II for strings of length two. Similarly, vertex B leads to vertex $B0$, but the vertex that would be labeled $B1$ is already assigned the string $A0$. Since the lengths of the strings $B1$ and $A0$ do not differ by more than one, the equiv-

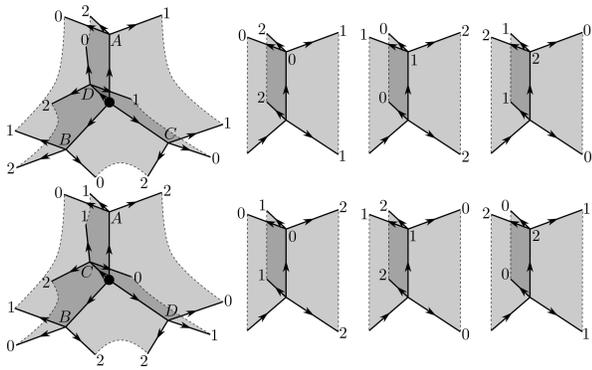


FIG. 4. Conventions for the labeling of vertices in 3d. The top and bottom rows correspond to positive and negative orientations of the space, respectively. Within a given row, the first figure indicates the labeling conventions around the root, while the remaining three figures indicate the labeling conventions after a turn of 0, 1 or 2, respectively.

absence of the associated paths is indicated by equating $B1$ with $A0$ in the third column of the table, and $B1$ is not used as the basis for any longer strings. The procedure continues as before with the strings $B2$, $C0$, $C1$, $C2$ and $D0$, until we observe that the vertex that would be labeled $D1$ is already assigned the string $C0$. This equivalence is indicated by equating $D1$ with $C0$ in the third column of the table, and $D1$ is not used as the basis for any longer strings. Finally, the string $D2$ is assigned and the construction to radius two is complete. This procedure may be repeated to continue expanding the swatch.

As in the 2d case, a 3d swatch that includes all relations between strings with lengths differing by at most one is sufficient to identify the endpoints of the four edges around a given vertex. The appearance of only relations between strings of the same length in the swatch for the region in Fig. 3 is therefore notable. This is because this portion of this cellular structure is composed of four identical truncated octahedra with all of their faces bounded by an even number of edges. This feature of the faces constrains the lengths of paths with the same endpoint to always differ by a multiple of two, limiting the types of relations that appear in this particular swatch. This is a special, rather than general, case.

Occasionally, there may be three or four paths from the root to a given vertex, with the lengths of each differing by at most one. In this situation, the three or four strings are written on the same line of the swatch, and are regarded as indicating sets of three or six transitive relations, respectively. The maximum number of strings that may be assigned to any vertex is four, since a vertex in a 4d cellular structure is connected to exactly four edges.

This completes the construction of a 3d swatch associated with a particular choice of a root and labeling of neighboring vertices to any desired radius (apart from the effect of the orientation of the space). Orienting the space negatively changes the handedness of the vertices

labeled A , B , C and D , and causes the letters 0, 1 and 2 to be assigned in a counterclockwise fashion from the perspective of an observer oriented along the most recently traversed edge. This set of conventions is indicated by the diagrams on the bottom row of Fig. 4. The two choices of space orientation and the twelve choices of vertices labeled A and B allow twenty-four swatches to be constructed per root, all describing the same region of the cellular network. All of the variants are considered equivalent, and the same partial ordering may be used in 2d and 3d to assign the minimal swatch to a given root.

Notice that despite the differences in the structure of 2d and 3d cellular networks, the resulting swatches (e.g., Tables I and II) show considerable similarities. This allows many procedures and calculations to be performed on a swatch independent of the dimensionality of the space in which the cellular network is embedded; this will be of considerable importance in Sections IV and V.

The reconstruction of the original 3d network from a swatch table is described, in detail, in Appendix B.

IV. ENTROPY

As indicated above, a swatch completely describes the topology of a cellular network in a region around the root, with the extent of the region described by the swatch radius. Nevertheless, a swatch is not entirely suitable for a direct description of the topology of the cellular network as a whole. Suppose that the cellular structure is statistically homogenous, i.e., any statistical feature measured within some bounded region converges to a definite limit as the area or volume of the region increases— independent of the location of the region in the cellular structure. This property of statistical homogeneity is not respected by the choice of a root at the swatch center.

We address this deficiency by constructing a swatch at every vertex of the cellular structure simultaneously. The resulting ensemble of swatches does not single out any particular vertex—respecting the assumption of statistical homogeneity. Intuitively, the procedure covers the entire cellular structure with small overlapping regions, with the topology of the cellular network in a given region described by the corresponding swatch. Provided that enough information is given to match the overlapping regions, this ensemble gives a complete description of the topology of the cellular structure.

We begin our statistical analysis by focusing on the relations between strings that arise in a swatch table. Let the length of a relation be the sum of the lengths of the equivalent strings; e.g., $C1 = B00$ is a relation of length five. Two swatches are k -equivalent when all of the relations up to length k are identical (i.e., the two swatches describe identical regions within a radius of roughly $k/2$). A k -equivalence class is the set of all swatches that are k -equivalent. Let $I(k)$ be the number of k -equivalence classes as a function of k (i.e., the number of distinct allowed configurations within a radius of roughly $k/2$).

Although the specific form of this function is not known, $I(k)$ is finite for all finite k .

Construct a swatch around a randomly chosen root of cellular structure X . The probability that this swatch belongs to the i th k -equivalence class is indicated by x_i^k , subject to the normalization condition $\sum_i^{I(k)} x_i^k = 1$ for every k . That is, x_i^k is the probability of observing the i th allowed configuration of cells in a region with a radius of roughly $k/2$ (where the radius is defined as for a swatch), centered on a vertex chosen at random from X . Hence, the set of probabilities x_i^k for all i and k give a description of the statistical topology of the cellular network; we refer to this as the *cloth* of X . Although there is no strict upper bound on the value of k , this is practically restricted either by the size of the cellular structure or by the available computational resources.

With the topology of the cellular network encoded by a discrete probability distribution, a variety of statistical techniques may be used to characterize the cellular structure. For example, the Shannon entropy [38] allows the k -entropy of X to be calculated as

$$H^k(x) = - \sum_i^{I(k)} x_i^k \ln x_i^k. \quad (1)$$

This quantity indicates the disorder of the structure at a length scale of about k edges.

We now examine the properties of the k -entropy by reference to five 2d structures, each containing N cells on the unit square with periodic boundary conditions:

1. The hexagonal structure is an array of hexagonal cells; i.e., the Voronoi tessellation of a triangular lattice.
2. The perturbed structure begins from the triangular lattice. Every point of the lattice is displaced in a random direction by a distance sampled from a normal distribution with a standard deviation of one-fifth the lattice spacing. A cellular structure is generated by taking the Voronoi tessellation of the result.
3. The evolved structure is the statistical steady-state resulting from normal grain growth (curvature flow) in an isotropic polycrystalline material following the procedure of Ref. [39].
4. The Poisson structure is the Voronoi tessellation of randomly (Poisson) distributed points.
5. The flip structure begins from the hexagonal structure. An edge is randomly selected and flipped, making the cells initially at the endpoints of the edge neighbors. This operation is not performed if it would result in a cell with fewer than three edges or a cell that shares an edge with itself. Attempting $100N$ edge flips is found to be more than sufficient to stabilize the statistical features of the result.

Representative areas of the reference structures appear in Fig. 5, ordered by increasing topological (Shannon) entropy. While not necessarily indicative of any actual physical systems, the reference structures are intended to be reproducible and to display a variety of cellular network topologies.

We first restrict our analysis of the k -entropy to a single type of cellular structure, namely, the evolved structure. Consider the dependence of the k -entropy on the number of cells N in the system. Fig. 6 shows the k -entropy curves for evolved cellular structures with $N = 10^2, 10^3, 10^4$, and 10^5 . The most striking features of this plot are that the curves are indistinguishable at small k and saturate for large k . The strong dependence of the saturation value on N clearly shows that this is a finite-size effect.

A 2d cellular structure with N cells contains $2N$ vertices. Since the number of available k -equivalence classes increases with k and the number of vertices in the system is a constant, the number of k -equivalence classes will eventually exceed the number of vertices. Suppose that there is a value of k above which this property holds and every swatch is in a distinct k -equivalence class. That is, $x_i^k = \frac{1}{2N}$ or $x_i^k = 0$ when the i th equivalence class is occupied or unoccupied, respectively, for all i . Referring to Eq. (1) indicates that the value of the k -entropy would be $\log(2N)$ in this situation. Inserting the values of N used in Fig. 6 gives 5.30, 7.60, 9.90 and 12.21 for the corresponding saturation values of the k -entropy, numerically identical to those observed. Hence, this is likely the source of the finite-size effect visible in the k -entropy for large values of k .

Meanwhile, the k -entropy goes to zero for small k in Fig. 6. This is a result of there being only a few k -equivalence classes available for small values of k , placing a sharp upper bound on the potential disorder of the system. For instance, only one k -equivalence class is available for $k = 0$ and $k = 1$ (i.e., the root vertex), and only two k -equivalence classes are available for $k = 2$ (i.e., three edges or one edge and a two-sided cell connected to the root). Since Eq. (1) indicates that the k -entropy must always be less than or equal to $\log(I(k))$, the k -entropy is always zero for $k = 0$ and $k = 1$ and is bounded above by $\log(2)$ for $k = 2$.

The most significant feature of Fig. 6 though is the convergence of the k -entropies to a limiting curve with increasing N . Consider measuring the x_i^k of an infinite cellular structure. Without any sampling error from a finite population of cells, the values of the x_i^k presumably reflect only the characteristics of the procedure employed to generate the network, rather than the details of any particular realization of the structure. With respect to Fig. 6 specifically, this would give a limiting k -entropy curve that is characteristic of the evolved structure. Restricting N to a finite value introduces sampling errors that cause the measured k -entropy curves to deviate from the limiting curve, most dramatically in the saturation of the k -entropy for large values of k . This cannot be

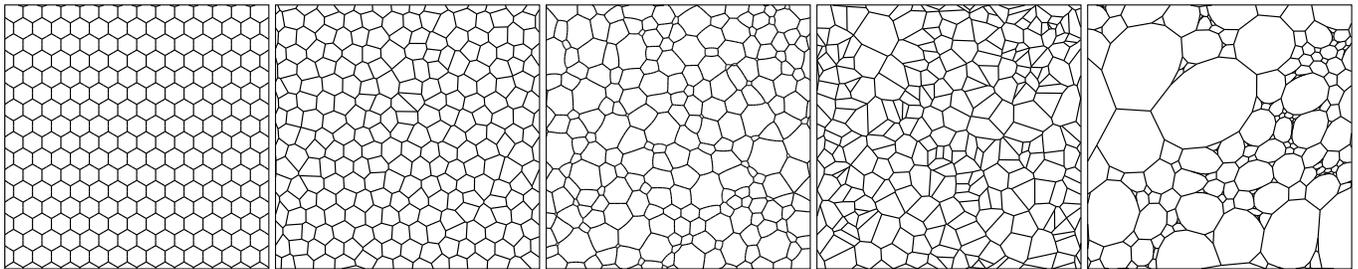


FIG. 5. From left to right, representative areas of the hexagonal, perturbed, evolved, Poisson and flip 2d reference structures, as described in the text. The corresponding k -entropies increase from left to right.

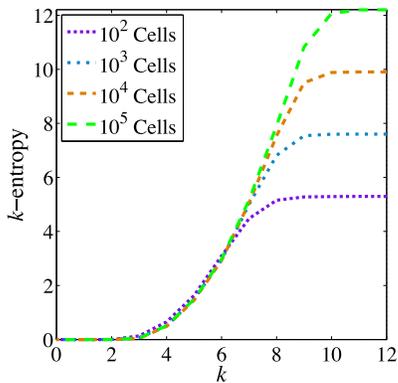


FIG. 6. (Color online) k -entropies for the 2d evolved structure with a varying number of cells in the system. As the system size increases, the k -entropies approach a limiting curve that is characteristic of the evolved structure. The saturation on the right occurs from finite-size effects.

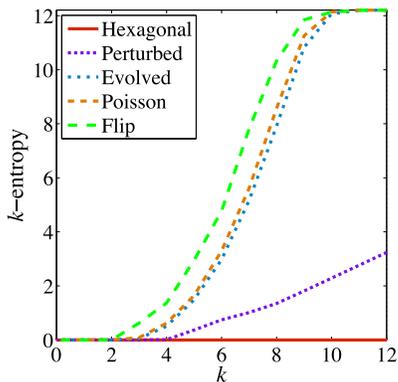


FIG. 7. (Color online) k -entropies of the five 2d reference structures with 10^5 cells.

avoided, given the restrictions imposed by finite sample size, but the convergence of the k -entropy curves with increasing N nevertheless gives an indication of the limiting curve.

Given the notion of a limiting k -entropy curve as a property of a particular generation procedure/network type, it is of interest to compare the limiting k -entropy curves for the different reference structures; this is done

in Fig. 7 for the five reference cellular structures with $N = 10^5$. Apart from the expected behavior at very small and large k , these curves do not intersect. Hence, they allow the structures in Fig. 5 to be sorted by topological entropy; the resultant ordering is consistent with visual determinations of the relative topological disorder of the structures.

The hexagonal structure clearly gives a lower bound to the k -entropy curves (i.e., the k -entropy is zero for all k), while we conjecture that the flip structure gives a corresponding upper bound for cellular structures (without two-sided cells). These bounds still allow for considerable variation in the k -entropy curves of the other cellular structures, particularly with increasing values of k . The similarity of the k -entropy curves of the evolved and Poisson structures is significant, particularly since the Poisson structure is often used as the initial condition for simulations of normal grain growth and foam coarsening. Specifically, the implied topological similarity of these cellular structures, relative to the variation that is conceivably allowed, suggests that there will be a short initial transient as the Poisson structure develops into the evolved one.

Properties of the k -entropy may also be investigated using 3d reference structures. We consider the following systems, each of which contain close to 9800 cells, set in the unit cube with periodic boundary conditions:

1. The body-centered cubic (BCC) structure is generated using the Voronoi tessellation of a BCC lattice.
2. The perturbed structure begins with a BCC lattice. Each point of the lattice is displaced in a random direction by a distance sampled from a normal distribution with a standard deviation of one-fifth the lattice spacing. A cellular structure is generated by taking the Voronoi tessellation of the perturbed lattice.
3. The evolved structure is the statistical steady-state resulting from normal grain growth (mean curvature flow) of an isotropic polycrystalline material, following the procedure of Ref. [40].
4. The Poisson structure is the Voronoi tessellation of randomly (Poisson) distributed points.

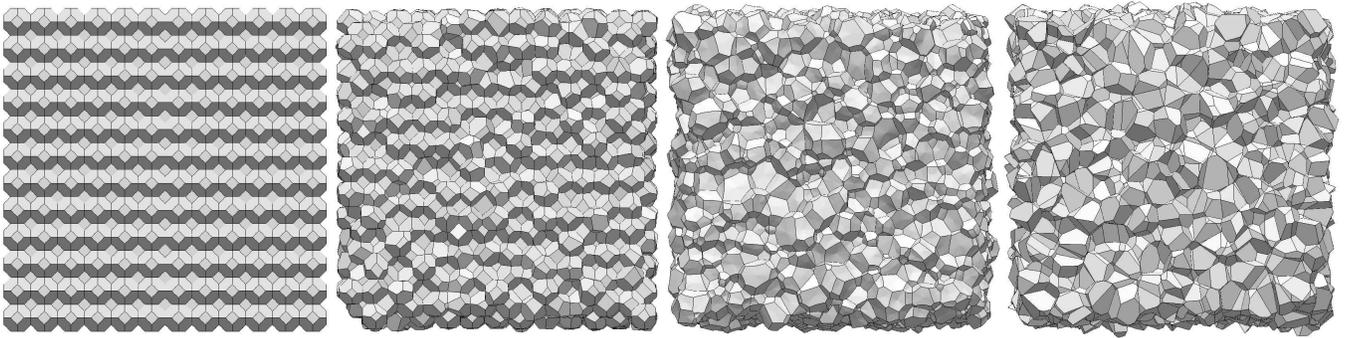


FIG. 8. From left to right, representative volumes of the BCC, perturbed, evolved and Poisson 3d reference structures; the flip structure is not illustrated for reasons described in the text.

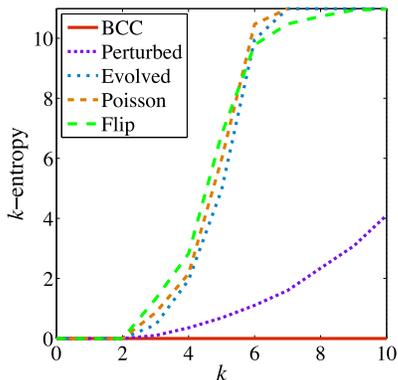


FIG. 9. (Color online) k -entropies of the five 3d reference structures with roughly 9800 cells.

- The flip structure begins from the Poisson structure. A randomly chosen edge is first selected and replaced with a triangular face; the five adjacent bodies are changed appropriately. Next, a triangular face is randomly selected and replaced with an edge. Neither of these operations are performed if they will create a face with fewer than three sides, a cell with fewer than four faces, or a cell that is adjacent to itself. Attempting this procedure $5000N$ times is found to be more than sufficient to stabilize the statistical features of the result.

Representative volumes of the reference structures appear in Fig. 8, ordered by increasing topological entropy. While the procedure for generating a flip structure specifies that certain topological changes should be made, this is not sufficient to specify the corresponding changes to the cell geometries. Indeed, there appears to be no straightforward way to adjust the geometry of this system to give something resembling a typical cellular structure. For this reason we do not illustrate the flip system, even though topological properties like the k -entropy remain well-defined.

Figure 9 compares the k -entropies for these five reference microstructures. Since our 3d system sizes are severely restricted, we do not consider cases with more

than 9800 cells. Many features of the 2d systems may still be observed for the 3d analogues though. For instance, the k -entropy is close to zero for small k in all systems. This results from the small number of available equivalence classes, just as in 2d, placing a sharp upper bound on the potential disorder of the system. Furthermore, all of the curves approach the same saturation value for large k , as determined by the number of vertices. Since the number of vertices is not uniquely specified by the number of cells in 3d, we allow the number of cells to vary and compare systems with similar numbers of vertices. Specifically, the five structures we consider have between 58,900 and 59,200 vertices, bounding the k -entropy for all structures at roughly $\log(59,000) \approx 10.98$. This is the limit observed in Fig. 9.

The 3d BCC structure clearly provides a lower bound to the k -entropy curves (i.e., the k -entropy is zero for all k). The flip structure appears to give a corresponding upper bound for $k \leq 5$, though the k -entropy dips below those Poisson and Evolved structures for $k \geq 6$. This illustrates that a system may have long-range order despite being relatively disordered at short distances. As in 2d, the similarity of the k -entropy curves for the Poisson and evolved structures suggests that the Poisson structure may be used as a reasonable initial condition to generate the evolved structure via normal grain growth. Perhaps most remarkably, the ordering of reference structures (with exception of the flip structure for $k \geq 6$) is identical to the ordering of the 2d analogs.

The similarity of k -entropy curves only gives an implicit measure of topological similarity of the corresponding structures though. This encourages the search for a more direct measure.

V. DISTANCE

In this section, we consider how to quantify the topological differences between distinct cellular structures. Let X and Y be two cellular structures and let the probability that a swatch of X or Y belongs to the i th k -equivalence class be x_i^k or y_i^k , respectively. (For any par-

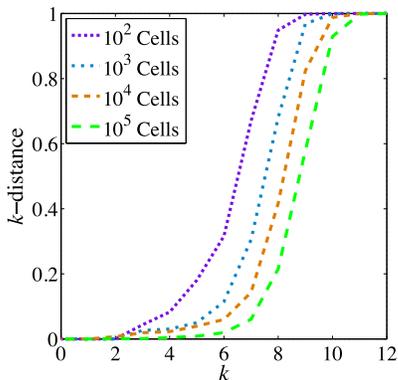


FIG. 10. (Color online) k -distances between distinct realizations of the 2d evolved structure. A curve shows the k -distances between two realizations with the same number of cells.

particular k , the x_i^k and y_i^k are discrete probability distributions on the space of k -equivalence classes.) Imagine drawing a swatch from X or Y with equal probability, without knowing from which cellular structure the swatch came. The probability that this swatch belongs to the i th k -equivalence class is $z_i^k = \frac{1}{2}x_i^k + \frac{1}{2}y_i^k$, the average of the corresponding probabilities for X and Y . The k -entropy of the corresponding random variable is $H^k(z)$. By comparison, an observer who knows the source of the swatch would calculate a k -entropy of $\frac{1}{2}H^k(y) + \frac{1}{2}H^k(x)$ [41]. The increase in the k -entropy from not knowing whether the swatch came from X or Y may therefore be expressed as

$$\begin{aligned} D_{JS}^k(x, y) &= H^k(z) - \frac{1}{2}H^k(y) - \frac{1}{2}H^k(x) \\ &= \frac{1}{2} \sum_i y_i^k \log \frac{y_i^k}{z_i^k} + \frac{1}{2} \sum_i x_i^k \log \frac{x_i^k}{z_i^k}. \end{aligned} \quad (2)$$

This quantity is frequently referred to as the Jensen–Shannon divergence [42] and is a standard measure of the similarity of discrete probability distributions. Although the Jensen–Shannon divergence does not satisfy the requirements of a metric, the square root of the Jensen–Shannon divergence does [41, 43]. That is, the function

$$d_{JS}^k(x, y) = \sqrt{\frac{1}{2 \log 2} \left(\sum_i y_i^k \log \frac{y_i^k}{z_i^k} + \sum_i x_i^k \log \frac{x_i^k}{z_i^k} \right)} \quad (3)$$

is a metric that measures the statistical difference between the cellular network topologies of X and Y . For the purposes of this paper, $d_{JS}^k(x, y)$ will be referred to simply as the k -distance. The k -distance has a minimum of zero when $x_i^k = y_i^k$ for every i , and has a maximum of one when at least one of x_i^k and y_i^k is zero for every i (i.e., when the probability distributions are disjoint).

As with the k -entropy, we initially restrict our attention to cellular structures with a single generation procedure and investigate the dependence of the k -distance on

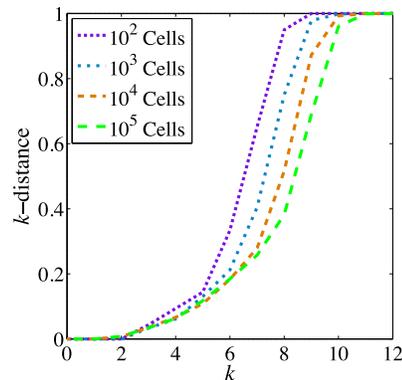


FIG. 11. (Color online) k -distances between a 2d evolved and 2d Poisson structure.

the number of cells N . Figure 10 shows four k -distance curves between two distinct realizations of the 2d evolved structure with the same but variable number of cells. Notably, as the total number of cells in the system goes to infinity, the k -distance appears to go to zero. Indeed, if the x_i^k of an infinite cellular structure reflects only the characteristics of the generation procedure, rather than the details of any particular realization of the structure, then the k -distances between two infinite structures will be exactly zero for all finite k . Note that the curves in Fig. 10 do indeed go to zero for small values of k and as N increases they tend to go to zero for larger and larger values of k . This is an important result, since this provides a test for whether two cellular structures are statistically equivalent.

Finite-size effects are strongly visible in Fig. 10 for large values of k ; this presumably occurs for the same reason as the saturation of the k -entropy in Fig. 6. That is, there may be some value of k above which every swatch of X and Y is in a distinct k -equivalence class. From Eq. 3, this would cause the k -distance to saturate at the maximum value of one at nearly the same value of k as for the saturation of the k -entropy in Fig. 6, as is visible from a comparison of the figures.

The k -distance between the 2d Poisson and the 2d evolved structures are shown in Fig. 11 for realizations with the same but variable number of cells. Similar to Fig. 10, the k -distance curves appear to be converging to a limiting curve as N increases. This limiting curve reflects the topological similarities and differences of infinite versions of the evolved structure and the Poisson structure. Deviations from the limiting curve occur due to sampling error from a finite number of cells, and is particularly visible in Fig. 11 for large values of k .

The existence of limiting k -distance curves that rigorously quantify the differences in cellular network topology raises the possibility that a cellular structure could be identified by the k -distances to a collection of reference structures. This is the subject of Fig. 12, which gives the k -distance between a 2d evolved structure and the four other reference 2d structures (plus a second real-

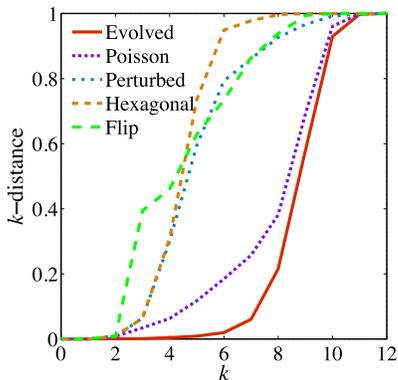


FIG. 12. (Color online) k -distances between a 2d evolved structure and the other four 2d reference structures. The evolved curve represents the distance between different realizations of the evolved structure and is shown for reference. All of the realizations contain 10^5 cells.

ization of another evolved structure) each with 10^5 cells. Not surprisingly, the distance between two realizations of evolved structures has the smallest k -distance for all k . Next, as expected from a visual inspection of Fig. 5 and from the analysis of the k -entropy curves in Fig. 7, the evolved structure is statistically quite similar to the Poisson structure (i.e., smaller k -distance for all k as compared with the other reference structures). The interpretation of the k -distance curves for the remaining reference structures is more subtle. The hexagonal and perturbed structures should be at about the same k -distance from the evolved structure for k small enough that the corresponding swatch covers only a few cells, as is observed. The increased number of k -equivalence classes inhabited by swatches of the perturbed structure makes this slightly more similar to the evolved structure with increasing k . Finally, the distance from the evolved to the flip structure is higher than for any other reference structure when k is equal to three, owing to the strong preponderance of three-sided cells in the flip structure. The curve for the flip reference structure then falls below that for the hexagonal structure; this is likely due to the increased variability of the flip structure relative to the hexagonal one.

The concept of the k -distance can also be used to study the difference between pairs of three-dimensional cellular structures. Figure 13 shows k -distances between a single 3d evolved structure and the four other reference 3d structures (plus a second realization of another 3d evolved structure). All structures contain roughly 9800 cells.

A number of observations are worthy of note. First, the k -distance between the pair of 3d evolved structures is lower than the distance between the evolved structure and any of the other 3d reference structures for all k . This is reasonable given that the two evolved structures are generated using the same process, implying that as the system size increases, this curve will approach zero

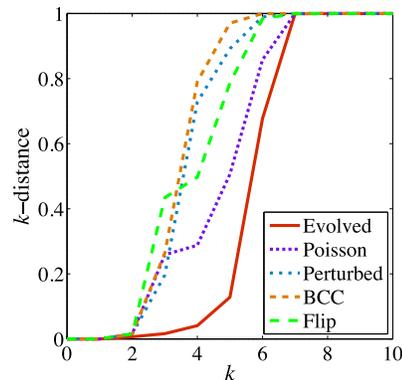


FIG. 13. (Color online) k -distances between a 3d evolved structure and the other four 3d reference structures. The evolved curve represents the distance between different realizations of the evolved structure and is shown for reference. All of the realizations contain roughly 9800 cells.

for all k .

Next, for most values of k , the 3d evolved structure is statistically most similar to the 3d Poisson structure and therefore has smaller k -distances as compared with the other reference structures. This should not be surprising given the similarity of the evolved and Poisson structures as illustrated in Fig. 8. For $k = 3$ though, the Poisson structure is more distant from the evolved structure than the perturbed one. This may be explained by the abundance of triangles in the Poisson structure (about 13.5% of all faces) and the scarcity of them in the evolved and perturbed structures (less than 1% of all faces).

As in 2d, the interpretation of the k -distance curves for the remaining 3d reference structures is more subtle. The BCC and perturbed structures are roughly the same k -distance from the evolved structure for k small enough that the corresponding swatch covers only a few cells. The increased number of k -equivalence classes inhabited by swatches of the perturbed structure makes this slightly more similar to the evolved structure than the BCC structure though, as is evident from the slightly smaller k -distance values.

Finally, the distance from the 3d evolved structure to the 3d flip structure is higher than for any other reference structure when k is equal to three, owing to the abundance of three-sided faces in the flip structure (over 30% of all faces). The curve for the flip reference structure then falls slightly below that for the BCC and perturbed structures, likely due to the increased variability of the flip structure relative to the others.

Remarkably, the ordering of the k -distance curves in Figures 12 and 13 is identical for most k . Aside from indicating the similarity of specific types of cell structures in different dimensions, this also helps illustrate the applicability of the concepts of k -equivalence class, k -entropy and k -distance in describing the topology of cellular structures in a manner that is independent of the dimension of the particular structure.

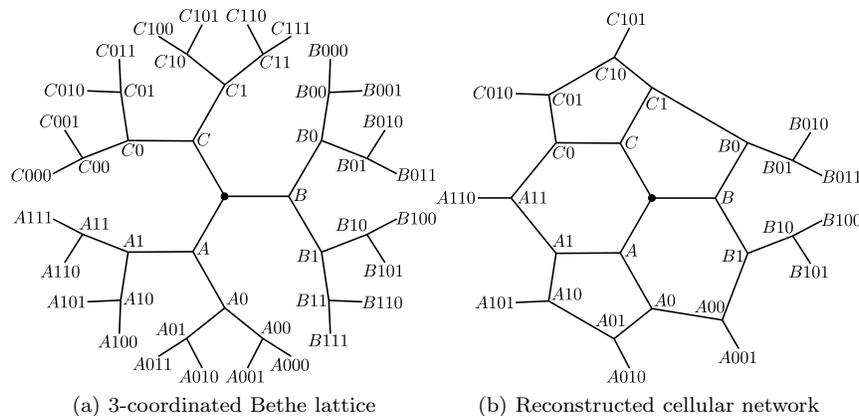


FIG. 14. (a) Our labeling scheme applied to the reference structure (3-coordinated Bethe lattice) for the 2d reconstruction, with a unique string for every vertex. (b) Reconstruction of the cellular structure in Fig. 1 from the the reference structure and the relations in Table Ia.

VI. CONCLUSIONS

We have developed the swatch as a mathematical object to describe the topology of a portion of a cellular network in two and three dimensions. By considering the probability of a swatch in a randomly chosen location being of a particular type, we are further able to characterize the statistical features of the topology of the cellular network as a whole. This allowed a meaningful definition of the entropy of a cellular structure to be developed, and the statistical difference of a pair of cellular structures to be rigorously measured by a distance function.

Historically, the materials science community has relied predominantly on the visual comparison of micrographs or on scalar measures such as the grain size to identify a cellular structure and reconstruct the processing history of a sample. While a full characterization of a material would include information relating to the presence of second phases particles, variations in solute content, grain morphology, etc., the distance measure presented here is an important step toward the complete and impartial quantification of a cellular structure. Built upon the topology of the grain boundary network, our distance function will enable a rigorous characterization of microstructure variability for a given processing procedure, and will allow the identification of the processing steps necessary to bring a material closer to the desired microstructure. From the standpoint of simulations, this distance function will enable measurements of the decay of the initial condition as the cellular network evolves toward a steady-state microstructure, and will enable the strict validation of simulation results using experimental microstructures.

More generally, our hope is that this paper stimulate interest in the broader scientific community in the study of cellular networks. The disorder that is frequently present in cellular networks has made the study of these structures resistant to reductionist techniques, and calls

out for the development of a language able to capture the properties of a complex structure as a whole.

ACKNOWLEDGMENTS

The authors would like to thank the Institute for Advanced Study where the original idea occurred, and Matthew Kahle for certain enlightening discussions in this direction. The first author was partially supported by grant HR0011-08-1-0093, and partially supported under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The second author was supported by grant HR0011-08-1-0093.

Appendix A: Reconstruction from a 2d swatch

One of the claims of Section II is that a swatch contains enough information to identify the endpoints of the three edges around any given vertex, provided that the swatch is constructed to a sufficient radius. This implies that the swatch labeling scheme described in Section II provides sufficient information for the reconstruction the cellular network topology in 2d. In fact, the cellular network topology is completely specified by the set of relations between strings with lengths that differ by at most one. Our intention is to substantiate this claim by explicitly reconstructing the topology of the region in Fig. 1a using only the relations that appear in Table Ia and the assumption that every vertex is 3-coordinated (connected to exactly three edges).

Let a *graph* be a collection of vertices connected by edges and a *tree* be a graph where any pair of vertices is connected by exactly one path. The n -coordinated Bethe lattice is defined as the tree where every vertex is directly connected to n other vertices. For instance, a portion of

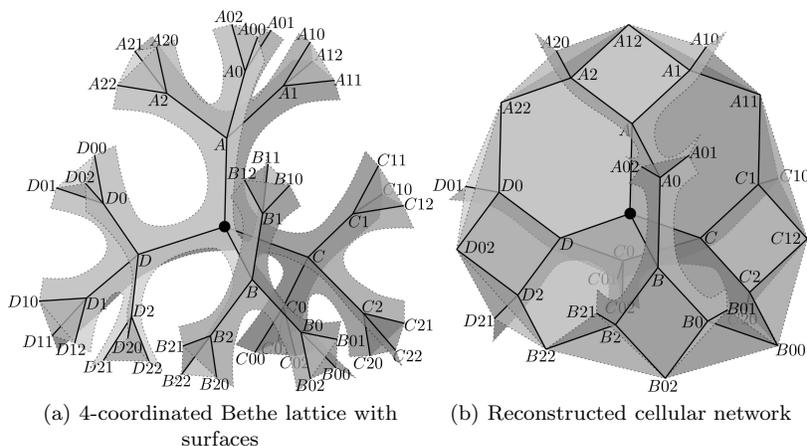


FIG. 15. (a) Our labeling scheme applied to the reference structure for the 3d reconstruction, with a unique string for every vertex. (b) Reconstruction of the cellular structure in Fig. 3 from the reference structure and the relations in Table II. Dashed lines indicate the continuation of the adjoining surface.

the 3-coordinated Bethe lattice around the root is given in Fig. 14a. This is used as the reference structure for our 2d reconstruction since the 3-coordinated Bethe lattice satisfies the minimal requirements to be a valid cellular network in 2d. Specifically, the property of a tree that any pair of vertices be connected by exactly one path precludes the existence of any closed cells in the reference structure, and more generally of any relations (equivalence of two labels for the same vertex, e.g., C1 and B00 in Fig. 1a) in a swatch of the reference structure. The absence of relations implies that every possible string occurs as the label of exactly one vertex.

The topology of a region described by a swatch may be reconstructed by identifying vertices and edges of the reference structure according to the relations of the swatch. This procedure induces the identification of certain of the surrounding edges and vertices as well, meaning that many vertices receive more than one label. When this occurs, only the lexicographically minimal string is retained for a given vertex. Our lexicographic ordering always places a shorter string before a longer one, and orders strings alphabetically otherwise.

There are two distinct types of relations in this context. Relations between strings with lengths differing by one always occur in pairs and indicate the identification of a pair of edges. Consider $B0 = C11$ and $C1 = B00$ in the third column of Table Ia. These relations are satisfied by identifying the edge from C1 to C11 with the edge from B0 to B00 in Fig. 14a. Continuity of the structure is maintained by overlaying the surrounding branches according to the relations $B = C111$, $B01 = C110$, $C = B000$, $C10 = B001$, etc., and retaining only the lexicographically minimal label for the vertices in every case. This gives the five-sided cell above and to the right of the root in Fig. 14b. Meanwhile, relations between strings of the same lengths occur individually and indicate the identification of a pair of vertices. The relation $A00 = B11$ in the fourth column of Table Ia

implies the identification of the vertices $A00$ and $B11$ and the overlaying of the surrounding branches according to the relations $A0 = B111$, $A001 = B110$, $B1 = A000$, etc. Retaining the lexicographically minimal string for every vertex gives the six-sided cell below and to the right of the root in Fig. 14b. Applying the remaining relations from Table Ia in a similar manner gives the graph in Fig. 14b; the reader is invited to confirm that this is topologically identical to the region around the root in Fig. 1a.

Since a swatch allows a network to be built that is topologically identical to the original 2d cellular network, the swatch gives a complete description of the cellular network topology. This appendix supports the stronger claim that the relations alone give a complete description as well. Furthermore, the relations provided by the procedure in Section II appear to be a minimal description in the sense that there is no redundant information. Since the relations are in correspondence with identifications of the vertices or edges of the 3-coordinated Bethe lattice, removing a relation from the set would result in a path not being closed, i.e., the reconstruction would no longer be topologically identical to the original network.

Appendix B: Reconstruction from a 3d swatch

The procedure for reconstructing the cellular network topology from a swatch in 3d closely follows the procedure of Appendix A for reconstruction in 2d. Specifically, relations between strings with lengths differing by at most one are applied to a reference structure. The reference structure must satisfy the basic requirements of a cellular network in 3d, namely, that every vertex be connected to exactly four edges and every edge be connected to exactly three surfaces. This set of properties is fulfilled by attaching surfaces to the 4-coordinated Bethe lattice to give the required reference structure, a portion of which appears in Fig. 15a. Analogous to the 2d case,

every possible string occurs as the label of a vertex somewhere in the structure.

The cellular network topology may be reconstructed by identifying the vertices and edges of the reference structure according to the relations of the swatch. As before, this procedure requires the identification of surrounding edges and vertices as well; i.e., many vertices receive more than one label. The same lexicographic ordering as in Appendix A is used to retain the lexicographically minimal string for a given vertex. That is, shorter strings always occur before longer ones and strings are ordered alphabetically otherwise.

Since a graph underlies our analysis of both 2d and 3d cellular networks, the same types of relations occur in both cases. Specifically, a pair of relations between strings with lengths differing by one indicates the identification of a pair of edges, while a relation between strings of the same lengths indicates the identification of a pair of vertices. Consider the relation $A0 = B1$ in the third column of Table II. Identifying this pair of vertices requires the application of the further relations $A = B11$ and $B = A00$ to maintain the continuity of the surface common to $A0$ and $B1$, and of $A02 = B12$ and $A01 = B10$ to satisfy the requirement that every vertex be connected

to four edges. The result is the four-sided face connected to the root and extending out of the page in Fig. 15b. Analogously, the relation $C0 = D1$ in the third column of Table II requires the identification of the surrounding vertices according to the relations $C = D11$, $D = C00$, $C01 = D10$ and $C02 = D12$ for the resulting structure to be a valid cellular network. This gives the four-sided face connected to the root and extending into the page in Fig. 15b. Although pairs of relations between strings with lengths differing by one do not occur for the swatch in Table II, the procedure in that case is essentially equivalent. Applying the remaining relations from Table II in the above manner gives the cellular network in Fig. 15b, which the reader is invited to confirm is topologically identical to the region around the root in Fig. 3.

The fact that a swatch allows the construction of an object that is topologically identical to the original 3d cellular network means that a swatch gives a complete description of the cellular network topology. Moreover, this demonstrates that the relations alone are sufficient in this respect. Since neglecting any one of the relations would result in the absence of one of the required vertex or edge identifications of the reference structure, our description appears to be minimal as well. This discussion is not a proof of the optimality of our description though.

-
- [1] B. Blonder, C. Violle, L. P. Bentley, and B. J. Enquist, *Ecology Letters* **14**, 91 (2011).
- [2] M. C. Gibson, A. B. Patel, R. Nagpal, and N. Perrimon, *Nature* **442**, 1038 (2006).
- [3] D. Staple, R. Farhadifar, J. Röper, B. Aigouy, S. Eaton, and F. Jülicher, *The European Physical Journal E: Soft Matter and Biological Physics* **33**, 117 (2010).
- [4] E. A. Jagla, *Physical Review E* **69**, 056212 (2004).
- [5] M. A. Kessler and B. T. Werner, *Science* **299**, 380 (2003).
- [6] P. Pina, J. Saraiva, L. Bandeira, and J. Antunes, *Planetary and Space Science* **56**, 1919 (2008).
- [7] C. Smith, *Metallurgical Reviews* **9**, 1 (1964).
- [8] F. Rhines, K. Craig, and R. DeHoff, *Metallurgical Transactions* **5**, 413 (1974).
- [9] S. Cox, M. Vaz, and D. Weaire, *European Physical Journal E* **11**, 29 (2003).
- [10] M. A. Fortes and P. I. C. Teixeira, *Journal of Physics A: Mathematical and General* **36**, 5161 (2003).
- [11] E. Lehockey, A. Brennenstuhl, and I. Thompson, *Corrosion Science* **46**, 2383 (2004).
- [12] S. Tsurekawa, S. Nakamichi, and T. Watanabe, *Acta Materialia* **54**, 3617 (2006).
- [13] A. Goyal, D. Norton, D. Kroeger, D. Christen, M. Paranthaman, E. Specht, J. Budai, Q. He, B. Saffian, F. List, D. Lee, E. Hatfield, P. Martin, C. Klabunde, J. Mathis, and C. Park, *Journal of Materials Research* **12**, 2924 (1997).
- [14] R. Haslinger and R. Joynt, *Physical Review B* **61**, 4206 (2000).
- [15] D. Weaire, *Metallography* **7**, 157 (1974).
- [16] D. A. Aboav, *Metallography* **13**, 43 (1980).
- [17] F. Lewis, *Anatomical Record* **38**, 341 (1928).
- [18] J. von Neumann, in *Metal Interfaces* (American Society for Metals, Cleveland, Ohio, 1952) pp. 108–110.
- [19] W. Mullins, *Journal of Applied Physics* **27**, 900 (1956).
- [20] M. Clusel, E. I. Corwin, A. O. N. Siemens, and J. Brujic, *Nature* **460**, 611 (2009).
- [21] E. I. Corwin, M. Clusel, A. O. N. Siemens, and J. Brujic, *Soft Matter* **6**, 2949 (2010).
- [22] M. P. Miklius and S. Hilgenfeldt, *Phys. Rev. Lett.* **108**, 015502 (2012).
- [23] K. A. Newhall, L. L. Pontani, I. Jorjadze, S. Hilgenfeldt, and J. Brujic, *Phys. Rev. Lett.* **108**, 268001 (2012).
- [24] K. Y. Szeto and W. Y. Tam, *Physical Review E* **53**, 4213 (1996).
- [25] T. Aste, K. Y. Szeto, and W. Y. Tam, *Physical Review E* **54**, 5482 (1996).
- [26] K. Y. Szeto, X. J. Fu, and W. Y. Tam, *Physical Review Letters* **88**, 138302 (2002).
- [27] B. Dubertret, N. Rivier, and M. Peshkin, *Journal of Physics A-Mathematical and General* **31**, 879 (1998).
- [28] K. Y. Szeto, T. Aste, and W. Y. Tam, *Physical Review E* **58**, 2656 (1998).
- [29] H. M. Ohlenbusch, T. Aste, B. Dubertret, and N. Rivier, *European Physical Journal B* **2**, 211 (1998).
- [30] B. Dubertret, K. Szeto, and W. Tam, *Europhysics Letters* **45**, 143 (1999).
- [31] V. Gertsman, M. Janecek, and K. Tangri, *Acta Materialia* **44**, 2869 (1996).
- [32] C. Schuh, R. Minich, and M. Kumar, *Philosophical Magazine* **83**, 711 (2003).
- [33] J. Basinger, E. Homer, D. Fullwood, and B. Adams, *Scripta Materialia* **53**, 959 (2005).
- [34] Y. Chen and C. A. Schuh, *Acta Materialia* **54**, 4709

- (2006).
- [35] G. S. Rohrer and H. M. Miller, *Acta Materialia* **58**, 3805 (2010).
- [36] T. Wanner, J. Fuller, Edwin R., and D. M. Saylor, *Acta Materialia* **58**, 102 (2010).
- [37] (2012).
- [38] C. E. Shannon, *Bell System Technical Journal* **27**, 379 (1948).
- [39] E. Lazar, R. MacPherson, and D. Srolovitz, *Acta Materialia* **58**, 364 (2010).
- [40] E. A. Lazar, J. Mason, R. D. MacPherson, and D. J. Srolovitz, *Acta Mater* **59**, 6837 (2011).
- [41] D. M. Endres and J. E. Schindelin, *Ieee Transactions On Information Theory* **49**, 1858 (2003).
- [42] J. H. Lin, *Ieee Transactions On Information Theory* **37**, 145 (1991).
- [43] F. Osterreich and I. Vajda, *Annals of the Institute of Statistical Mathematics* **55**, 639 (2003).