Tail Departure of Log-Normal Grain Size Distributions in Synthetic Three-Dimensional Microstructures

JOSEPH C. TUCKER, LISA H. CHAN, GREGORY S. ROHRER, MICHAEL A. GROEBER, and ANTHONY D. ROLLETT

Polycrystalline materials are generally thought to have grain size distributions that correspond to log-normal. Statistical volume elements can use such geometric shape distributions to simulate single-phase polycrystals. When the log-normal distribution is used for grain size, however, it can give rise to nonphysical large grains that cannot be practically accommodated in finite simulation volumes. The application of other distributions that afford better control of the upper tails, *e.g.*, truncated distributions, resolves the problem and allows more representative distributions to be generated. These points are illustrated with an example of generation of a synthetic three-dimensional (3-D) microstructure to represent the nickel-based superalloy Inconel 100, which exhibits significant upper tail departure from log-normal. Twin insertion, to represent annealing twins, will also be discussed.

DOI: 10.1007/s11661-011-0851-4

I. INTRODUCTION

COMPUTATIONAL modeling of materials has become a crucial and efficient field of materials science and engineering. Efforts are now under way to incorporate computational materials engineering into the development of materials and realize some savings in time and cost that have benefitted other areas of engineering such as aerospace.[1] Different strategies have been employed to generate three-dimensional (3-D) digital microstructures for use in the prediction of properties from microstructure.[2–8] Saylor *et al.*[2] and Brahme *et al.*[3] used experimental grain size, shape and orientation statistics in two dimensions to infer 3-D grain structures and to fit crystal orientations to the grains. Uchic *et al.*[4] reconstructed microstructures directly from serial section data. Groeber *et al.*[5,6] incorporated statistical nearest neighbor (NN) information into their generation procedure. Fan *et al.*[7] and St-Pierre *et al.*[8] used Voronoi tessellation to simulate polycrystalline structures.

It is commonly accepted that many single-phase fully dense polycrystals are described by a log-normal grain size distribution (GSD).[9–11] These studies confirm a log-normal fit to a high confidence because of limited data and the use of histograms that emphasize the region around the mean. No mention is made, however, concerning the consequences of deviation from the log-normal description. Of particular concern is the upper tail because a "fat tail" may result in large grains being inserted into the microstructure if the probability density is finite at grain sizes of the order of the simulation domain. If, however, the actual upper tail of the grain size distribution is less strong than log-normal, then nonphysical large grains may be present in the digital microstructure and the model will not be representative. Thus, subsequent property calculations using, *e.g.*, the finite-element method (FEM), will accordingly lose predictive power.[12,13]

The fundamental microstructure feature is the grain, which is described by shape, size, and orientation. The term representative volume element (RVE) is applied frequently to digital microstructures.[14] What must be addressed carefully is what part of the experimentally observed input statistics that the volume element represents. If a digital microstructure is an RVE, then the distribution of the entities of the microstructure must match the real material such that an applied stimulus evokes the real material response.

The difference between an RVE and a statistical volume element (SVE) is that an SVE does not represent all of the attributes of the observed distribution in one instantiation, but it will converge to it in multiple instantiations.[15] Accordingly, an SVE can be considered an RVE for only some material metrics (*e.g.*, GSD) but not others (*e.g.*, fatigue crack damage), which could be studied using multiple SVEs. SVEs that match lower order moments (mean and standard deviation) are suitable for material properties that probe the mean field (*e.g.*, elastic modulus). Matching higher order moments (skewness and kurtosis), however, that affect tails of the distribution would enable the analysis of responses that probes at the extreme of the GSD,[16,17] *e.g.*, fatigue crack initiation.

JOSEPH C. TUCKER, PhD Candidate, GREGORY S. ROHRER, Professor and Department Head, and ANTHONY D. ROLLETT, Professor, are with the Department of Materials Science and Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890. Contact e-mail: josephtu@andrew.cmu.edu LISA H. CHAN, Applications Engineer, is with EDAX Inc., 91 McKee Drive, Mahwah, NJ 07430. MICHAEL A. GROEBER, Engineer, is with the Air Force Research Laboratory, Materials and Manufacturing Directorate, Wright Patterson AFB, OH 45433.

Manuscript submitted April 4, 2011.
Article published online August 30, 2011

Based on various useful geometrical properties of ellipsoids, distributions of ellipsoids are used frequently as the geometric basis for 3-D digital microstructures.[2,3,5,6] The utility of ellipsoids is presented in Section II–B. The chosen distribution dictates the ellipsoids that are "seeded" into the structure and ultimately become grains. However, the various stages that are necessary for the generation of 3-D microstructures cause the distribution of grain sizes to depart from the initial ellipsoid size distribution. These steps all involve well-known mathematical techniques and are subsequently discussed briefly and in Section II. One important point is that the processes involved in packing a set of objects that approximate grains and then filling space to arrive at a fully dense microstructure mean that matching even just a measured size distribution requires a feedback loop. This is in contrast to matching orientation and misorientation distributions, which has been demonstrated to be feasible under a wide range of circumstances.[1]

A pseudorandom sampling of the chosen distribution is packed into a volume element of a designated size. The individual grain sizes are generally binned and ordered into an empirical cumulative distribution function (ecdf) to analyze GSDs. In dealing with probability distributions it is useful to think of bins as confidence intervals, indicating the likelihood of a particular range of ellipsoid sizes being chosen and placed into a structure. In a log-normal distribution, there is a well defined probability that an abnormally large ellipsoid is sampled and placed into the structure. "Abnormal" in this context means a grain whose size is greater than about three times the average size, which is known as an abnormal grain or "as large as" (ALA) grain. An investigation of the likelihood of ALA grain injection is detailed in Section II–C.

After the structure is "seeded" but before it is densified, a subset of the ellipsoid collection is chosen that optimizes space filling and minimizes overlap (between the ellipsoids). This is done by representing the ellipsoid centers as Monte Carlo (MC) integration points and performing a simulated anneal process.[1] Then, the resulting subset of ellipsoids is nucleated and grown to fill space with a cellular automaton (CA) based on ellipsoid size (and shape). At this point, the instantiation is complete, but several techniques are used to improve the quality of a structure that address small regions unrelated to the underlying set of ellipsoids and "relax" grain boundaries to provide more realism. These are discussed in Section II–G. All these steps impact the final GSD and are therefore pertinent to the feedback loop that matches the simulated GSD to the measured one.

The methods described in this study are linked to the nickel-based superalloy Inconel 100 (IN100). To develop physics-based models of IN100 such as FEM for material metrics such as damage and failure,[5,6,18] the generated models must sufficiently capture the pertinent statistics with high fidelity. Synthetic microstructures possessing this level of accuracy are sought after because a large number of experiments are needed to quantify variability meaningfully in fatigue life. In general, insufficient experimental data are available to support this testing.[18]

If theoretical distributions are successful in matching measured GSDs accurately, then an RVE for materials such as IN100 could be produced much more faithfully. Furthermore, more flexibility would be introduced with regard to structure size and scalability. The accuracy of physics-based microstructure simulations is affected directly by the geometry of the synthetic microstructure. Inefficient models on the front-end tax the efficiency of surface and volume meshing procedures and ultimately FEM simulation by creating more degrees of freedom.

II. REVIEW OF SYNTHETIC MICROSTRUCTURE GENERATION

A. *Log-Normal Distribution*

As mentioned in the Section I, many distributions of particles and single-phase polycrystals empirically fit a log-normal distribution.[9–11] Log-normally distributed variables are necessarily positive. The function $f(x)$ is log-normally distributed with mean (μ) and standard deviation (σ) in Eq. [1]. The cumulative distribution function (*cdf*) of a log-normal distribution is:

$$F(x;\mu,\sigma) = \frac{1}{2}\operatorname{erfc}\left[-\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right] = \Phi\left[\frac{\ln(x)-\mu}{\sigma}\right] \quad [1]$$

B. *Ellipsoids*

Ellipsoids are the chosen representative object because they are mathematically simple and are represented by a length:height:width aspect ratio, so they can embody anisotropy such as a rolling deformation (plane strain compression). The equation for an ellipsoid is written as

$$\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2+\left(\frac{z}{c}\right)^2 = 1 \quad [2]$$

where a, b, and c are the three semiaxes and x, y, and z are coordinates (relative to the ellipsoid center) of a point lying on the surface of the ellipsoid. Most convenient for defining the characteristics of polycrystalline microstructures is to have electron backscattered diffraction (EBSD) maps on three orthogonal material planes. Such maps are regular grids of measurement points, each of which contains the full crystal orientation. Such maps can be analyzed for grain orientation and shape. From the size distributions on the orthogonal planes, a set of ellipsoids can be generated whose cross-section statistics matches the measurements.[19] Extracting experimental data from a single two-dimensional (2-D) image produces ellipsoids limited to oblate and prolate. Spheres are used self-evidently for equiaxed grains. Provided that no spatial gradient in the grain size exists in the sample, a homogeneous distribution of ellipsoids is sufficient to represent the GSD.[3]

Different options exist for creating the ellipsoid distributions, including inputting a list or generating a set of ellipsoids that approximates a log-normal distribution. The list option is used for inputting ellipsoids

generated from other means (*e.g.*, EBSD or serial sectioning). If, for example, all three semi-axes are set equal to one another in Eq. [2], the ellipsoid geometry then reduces to spheres of constant radius. The result is grains that are not only equiaxed but also near monodisperse. Note that an exactly monodisperse size requires that the grain centers be placed on a simple lattice.

C. *Truncation in Preprocessing*

A log-normal distribution of ellipsoids can place too many small and large grains in a structure. The small grains can be "eroded" and "coarsened" as described in Section II–G; however, there is no simple provision for removing uncharacteristically large grains once they are placed into the structure. In short, the values of the standard deviation in a log-normal distribution that match observed values can easily produce large grains that are incompatible with the domain size.

To truncate the upper end of the distribution, the same log-normal distribution is used except a simple "if" statement is added which compares each sampled ellipsoid with an upper threshold. If the ellipsoid exceeds the threshold, then that object is discarded and the program continues with the sampling.

Table I shows a progression of successively more restrictive log-normal distribution truncations. Note that the log-normal mean and standard deviation inserted into Eq. [1] are those used to generate ellipsoids for the matched synthetic IN100 discussed in Section III ($\mu = 4.69 \times 10^{-2}, \sigma = 2.37 \times 10^{-2}$). To interpret Table I, the minimum simulation domain edge is scaled to 1, then the other edges are scaled accordingly. Because $\ln(0) = 1$, only the portion to the right of the decimal is retained. As the truncation threshold shrinks, more of the upper tail of the log-normal distribution is cutoff and the larger and less likely sampled semiaxis values are eliminated. Note that it is common to sample millions of ellipsoids in the packing process, meaning that a ~34,000 voxel (volume-pixel) ellipsoid is likely. Considering a $100 \times 100 \times 100$ domain, a grain this size occupies ~3 pct of the synthetic structure.

D. *Recursive Sampler*

With a description of the ellipsoids in hand, the next step is to seed them recursively into the simulation domain. The user can incorporate an ellipsoid overlap allowance (λ_{overlap}). The default value is 1.05, which allows for slight overlap as small degrees of overlap encouragement increases the fraction of space filled by the ellipsoids that remain after the packing optimization process. The criterion for recursive subdivision is

$$a_{\text{mid}}\lambda_{\text{overlap}} < bbox_{\text{max}} \quad [3]$$

with a_{mid} as the intermediate semiaxis of the ellipsoid and $bbox_{\text{max}}$ as the maximum dimension of the simulation domain. When the ellipsoids are placed, their centroid will lie within the subdivided cell, ensuring adequate dispersion.

E. *Packing Optimization (Simulated Annealing)*

An overfilling set of ellipsoids results from the recursive sampling procedure. A subset is desired that optimally fills space in the simulation domain.[3] Optimal filling means that each volume pixel (voxel) of the structure is contained in one and only one ellipsoid. Simulated annealing is a global optimization strategy. The simulated annealing algorithm chooses a subset of the initial set of ellipsoids based on minimizing the system energy.[20] The system energy (objective function) combines penalties for ellipsoid overlap and for gaps. First, a random subset is chosen roughly equal to the domain volume (Figure 1(a)) using an adjusted volume calculated with MC integration. This algorithm estimates the filled volume using MC integration.[2] MC integration computes an estimate of a multidimensional integral:

$$I = \int_v f(\bar{x})d\bar{x} \quad [4]$$

where $\bar{x} = \{x_1, \ldots x_n\}$ and V is a n-dimensional hypercube ($n = 3$ for 3-D). The estimate of Eq. [4] is

$$I = V\frac{1}{N}\sum_{i=1}^{N} f(\bar{x}_i) - V\langle f\rangle \quad [5]$$

where N the sample size and $\bar{x}_1, \ldots, \bar{x}_N$ are N random points in V. $\langle f\rangle$ is the sample mean of the integrand.[21] The more ellipsoid centers that fall within the ellipsoid of interest, the smaller its volume estimate. This allows a large enough subset to remain (with adequate dispersion) to attempt to pack the simulation volume with simulated annealing: an "initial screening" of ellipsoid fitness. The result is a more representative estimated filled volume of each ellipsoid.[3] The system energy is

Table I. A Progression of Truncation Thresholds of a Log-Normal cdf with Corresponding Percent of Distribution Truncated, Likelihood of Selection of the Threshold, and the Largest Ellipsoid that Could be Sampled Under the Enforced Truncation

Truncation ln(…)	Pct Log-Normal Distribution	1 in …	Largest Ellipsoid Sphere Equivalent Radius
1.20	5.50×10^{-9}	1.82×10^{8}	20
1.18	2.79×10^{-7}	3.59×10^{6}	18
1.16	9.17×10^{-6}	109,107	16
1.14	1.92×10^{-4}	5203	14
1.12	2.52×10^{-3}	396	12
1.10	2.04×10^{-2}	49	10

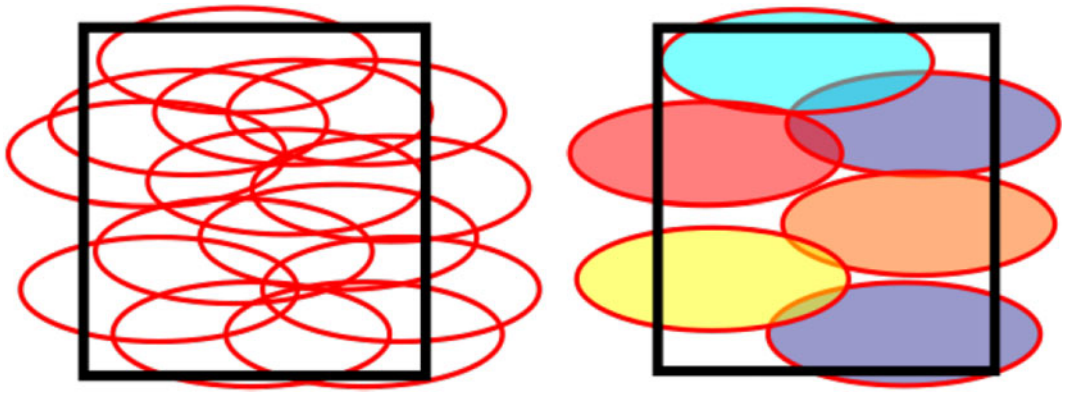


Fig. 1—(*a*) Overfilled simulation domain. (*b*) "Active" ellipsoids chosen through the simulated anneal algorithm.[19]

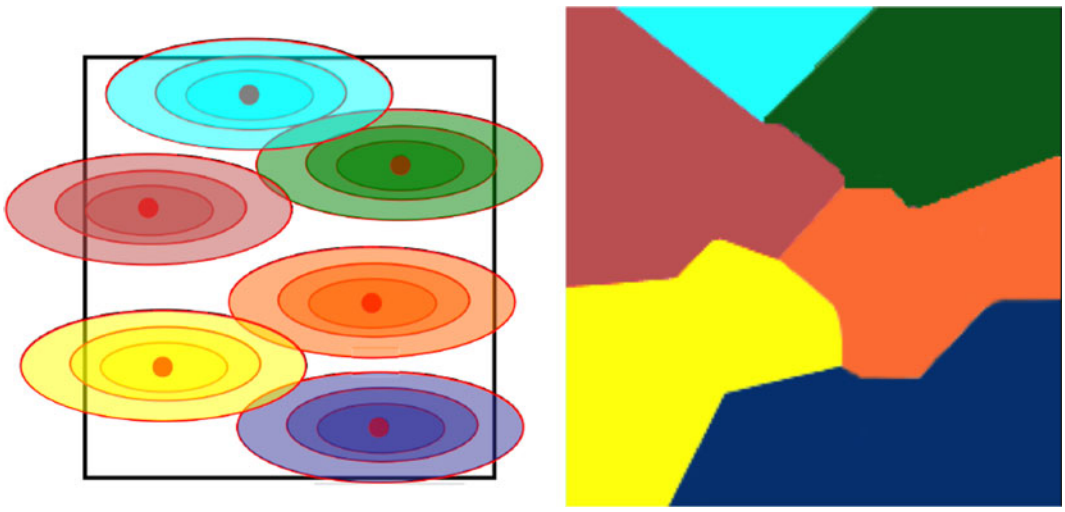


Fig. 2—(*a*) Ellipsoids growing through the cellular automaton (*b*) resulting dense polycrystalline structure.[19]

determined using the overlap and gaps between the ellipsoids

$$\alpha = \frac{-\omega}{1-\frac{1}{\zeta^2}}\left(1-\frac{E^2}{\zeta^2}\right) \qquad [6]$$

In Eq. [6], α is the overlap cost (energy), ω is an overlap encouragement factor, ζ is a zero overlap penalty, and E is an ellipsoid function that compares an ellipsoid to a given point (Eq. [7]). For all instantiations, $\omega = 1.0$ and $\zeta = 0.95$ (empirically chosen). This energy (Eq. [6]) is minimized through transformations (add, subtract, swap, and jog) on the collection of ellipsoids. If the transaction decreases the system energy, it is always accepted. If it increases the energy, then its acceptance is determined by how far along the annealing schedule the procedure is. As the anneal progresses, the criteria become more stringent[2] as the effective temperature is decreased. The result is illustrated in Figure 1(b).

$$E = \left(\frac{i-x}{a}\right)^2 + \left(\frac{j-y}{b}\right)^2 + \left(\frac{k-z}{c}\right)^2 \qquad [7]$$

F. *Cellular Automaton*

The simulated anneal algorithm generates an optimal set of ellipsoids, but regions remain that either contain multiple ellipsoids or are unoccupied. A CA densifies the structure.[2] The collection of "active" ellipsoids is placed on a discrete 3-D grid. The ellipsoid centroids are used as nucleation sites, with the larger ellipsoids being "nucleated" first. The ellipsoids are allowed to grow until they impinge and fill space (Figure 2).[2,3] The time increment is chosen so as to obtain an increase in radius of each grain of approximately one voxel.

G. *Postprocessing*

Three different postgeneration schemes are presented subsequently. The goal in postprocessing is to eliminate any unrealistic byproducts of the generation process and possibly to prepare the voxelized structure for surface meshing. Note that some aspects of the degree to which a synthetic microstructure resembles an experimentally observed one have yet to be quantified.

A minimum grain size threshold can be set in postprocessing. If a structure has millions of voxels, then it is reasonable to exclude grains consisting of only a few voxels (maybe larger, depending on circumstances). Setting a minimum size causes the algorithm to scan through the grains and merge any smaller grains with which it shares most of its grain boundary area. Setting a threshold is useful when trying to produce a synthetic microstructure from a measured structure because the minimum grain size is easily matched in this fashion.

Grains can also be merged that have fewer NN grains than a given value with their majority NN. This is effective for identifying and removing true "island" grains with only one NN. Note that apparent surface "island" grains frequently remain but have subsurface NN. This erosion process has also proved essential for preventing "dangling" voxels that lead to poorly formed surface meshes and that subsequently hinder volumetric meshing.

The grains have been constrained to grow as ellipsoids, which is what many of them resemble. Because each grain grows outward a constant rate, there is no opportunity for force balance to be established at triple lines and the dihedral angles to approach equilibrium as expected for annealed polycrystals. Applying (isotropic) MC grain growth "relaxes" the grain boundary networks that the CA produces. In the MC Potts model,[22] space is discretized on a simple cubic lattice (as also used in the CA) with each point belonging to a particular grain. Only points whose neighbors have different states increase the energy.[23,24] The structure is evolved by allowing each voxel to change or flip its grain ID if the change decreases the system energy or by comparison with a Boltzmann factor for an energy increase. A random number sequence governs which voxel attempts to flip. The algorithm results effectively in motion by curvature of the grain boundaries. This results in larger grains growing bigger at the expense of reduction in and consumption of smaller grains. MC coarsening is also effective in eliminating "island" grains. Anisotropy of grain shape is lost quickly with the MC Potts Model, however, so each voxel is typically given only 50 to 100 opportunities to flip. The effects of postprocessing are discussed in Section VII.

III. TWIN INSERTION

In face-centered-cubic metals in general and certainly the IN100 presented in this article, annealing twins are an important feature of the structure. Twinning events are one of the most important aspects of grain boundary engineering because they enhance the fraction of special boundaries. After the microstructure geometry has been defined by one of the microstructure generation methods described in Section II, the twin insertion algorithm takes the voxel structure and adds additional grains to simulate the annealing twins. The twin insertion algorithm assumes that twin grains always section the parent grain completely and that the parent grain resides on either side of the inserted twin or that the twin does not reside on the edge of the parent grain. The actual twin morphologies are admittedly more complex than this. Insertion of twins into the digital microstructures helps produce digital structures that are more representative of the experimentally observed microstructures that contain annealing twins.

Prior to inserting annealing twins into the specified microstructure, the frequency of $\Sigma 3$ clusters must be characterized. A $\Sigma 3$ cluster is a set of grains in which each shares a $\Sigma 3$ boundary with at least one other grain in the cluster. Because the $\Sigma 3$ misorientation is used to detect such clusters, no distinction is made between coherent and incoherent twins. Given the $\Sigma 3$ cluster distribution (referred to as the target or preexisting $\Sigma 3$ cluster distribution), the algorithm synthesizes annealing twins in the structure such that the final $\Sigma 3$ cluster distribution matches the target distribution. To specify the target $\Sigma 3$ distribution, the relative frequency by cluster size is converted to a list of $\Sigma 3$ cluster sizes, which are listed in order of decreasing size. At each iteration step, the algorithm attempts only to match one specified $\Sigma 3$ cluster size by inserting annealing twins into the grains and rotating the grain orientations to create incoherent $\Sigma 3$ grain boundaries. Therefore, the number of iterations is related directly to the desired number of target $\Sigma 3$ clusters.

After the target $\Sigma 3$ cluster size is specified, the algorithm initiates the $\Sigma 3$ clusters by using the preexisting $\Sigma 3$ clusters. If the target cluster size is smaller than the preexisting $\Sigma 3$ cluster, then the grains in the preexisting $\Sigma 3$ cluster are used as the seed grains for synthesizing twins and initiating the $\Sigma 3$ cluster growth. If no preexisting clusters satisfy the size criteria, then a grain is chosen at random to initiate the $\Sigma 3$ cluster.

Once the grain has been specified, twin insertion begins by identifying all the voxel positions that belong to the chosen grain and calculating the center of mass and spherical equivalent radius of the grain. If the radius of the grain is smaller than 5 voxels, the algorithm does not attempt to insert twins into the grain. One of the four variants of the $\langle 111 \rangle$ is chosen at random and the corresponding grain orientation for the twin is calculated by rotating the orientation of the parent grain 60 deg about the same $\langle 111 \rangle$ variant that was chosen. A fixed number of attempts is made (arbitrarily set to 40) to insert up to six twins into the chosen grain by identifying the orientation of the grain and calculating the $\langle 111 \rangle$ (the variant that was chosen) boundary plane for the grain in the sample reference frame. The calculation of the $\langle 111 \rangle$ boundary plane to insert into the grain is based on the parametric definition of a plane, where a plane is a set of points that fulfill the following conditions:

$$a(x - x'_o) + b(y - y'_o) + c(z - z'_o) < \text{tolerance} \quad [8a]$$

$$(x_o, y_o, z_o) = \text{centroid of the grain} \quad [8b]$$

$$(x'_o, y'_o, z'_o) = (x_o, y_o, z_o) + (D \cdot n) \quad [8c]$$

where D is the distance between the plane and the centroid of the grain, and the other definitions of the

variables are based on Figure 3.[24] The tolerance in Eq. [8a] determines the thickness of the grain and typically is set as the minimum tolerance required by the grain with an additional random value that ranges from 0.8 to 1.3. The distance from the centroid of the grain (D) in Eq. [8c] is also chosen at random to have a value that is in the range of 0 to 5 voxels away from the spherical equivalent radius of the grain. The choice of limiting the maximum value for D to be 5 voxels smaller than the grain radius is to avoid inserting twins on the edges of the chosen grain. The D value is tracked as each twin is inserted into the grain such that any subsequently added twins are not allowed to appear within approximately 3 voxels from any previously inserted twins. This ensures that, as more twins are added into the grain, the twins will not intersect or overlap with one another. The intersection of the twins is avoided to allow easier tracking of the Σ3 clusters.

Once twin insertion completes on the specified grain(s), all the neighbors of every grain included in the current Σ3 cluster, with priority given to the neighbors around the first and last grains in the cluster, are tested as potential candidates for connecting to the current cluster through an incoherent Σ3 grain boundary. To be accepted into the Σ3 cluster, the potential candidate must not cause the current Σ3 cluster to exceed the target cluster size. If the potential grain belongs to a preexisting Σ3 cluster or a Σ3 cluster that was constructed during a previous iteration step, the clusters are allowed to merge if the final cluster size does not exceed the target cluster size. When two Σ3 clusters need to be merged, the grain orientations of the smaller sized cluster will be rotated such that the grain boundary between the two clusters will be an incoherent Σ3 boundary. If the merged cluster was a preexisting cluster or if the single potential grain is accepted, then annealing twins are synthesized in the accepted grain(s). All the boundaries surrounding the current cluster are then analyzed to ensure that if any rotations performed on the grain orientations generated new Σ3 boundary relationships, the related grains are included in the current cluster count. The procedure of finding potential candidates to add to the cluster is then repeated until either the current cluster size is larger or equal to the current target cluster size or no grains can be added further to the current cluster. The pseudo code for the algorithm can be found in Appendix A.

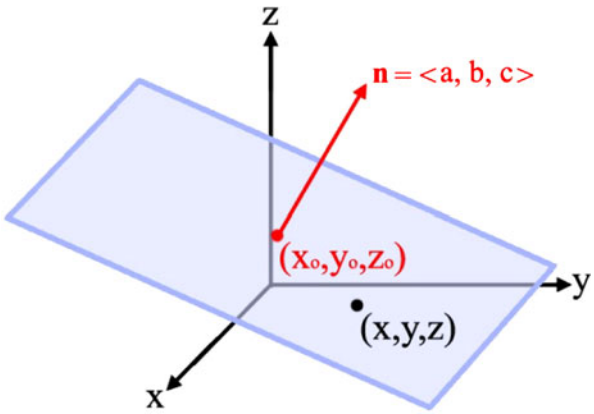


Fig. 3—Schematic showing the calculation of the equation of a plane.

An alternative to using Σ3 cluster sizes as the target for twin synthesis is to use only the number of twins as the target. When only a desired number of twins is declared (as in the case of this study), the algorithm chooses the grains randomly in the structure and performs the same twin insertion method outlined previously. Once twins are synthesized in a grain, the grain is removed from the list of potential grains for inserting the next set of twins.

A. *Controlling Twin Width and Placement of the Annealing Twin*

The variables of tolerance and D found in Eq. [8a] and Eq. [8c] allow the control of twin width and the location of the annealing twin, respectively. Figure 4 shows the increase in twin width or twin thickness as the tolerance value is increased from 0.5 to 3.0. Because the method of inserting the twins relies only on performing a dot product between the voxel positions and the <111> normal, no resolution concerns are associated with the

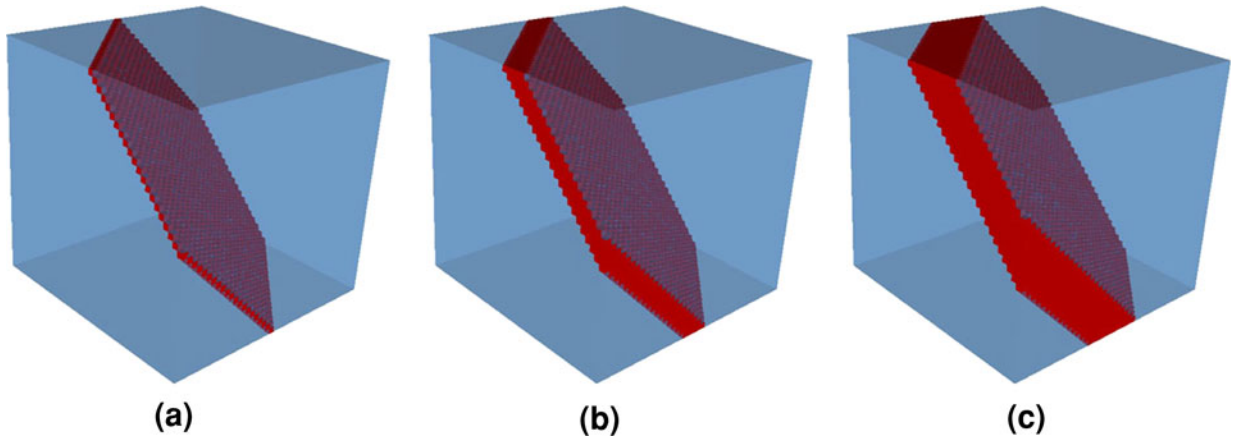


Fig. 4—Synthesized twin grain (red) with $D = 0$ and tolerance values of (*a*) 0.5, (*b*) 1.5, and (*c*) 3.0. (Color figure online).

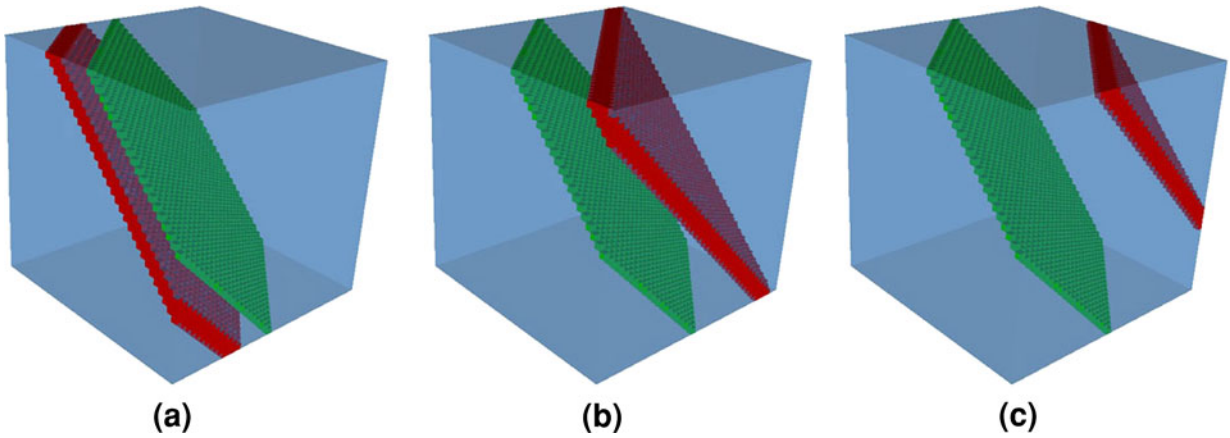


Fig. 5—Synthesized twin grain (red) with tolerance = 1.0 and D values of (*a*) –5, (*b*) 10, and (*c*) 20. The green twin grain marks the location of where the twin grain would be if D was set at 0.

twin width. The only concern involved with choosing an inappropriate tolerance value is that if the chosen parent grain is small, then a large tolerance value may convert the entire parent grain to a "twin" grain, whereas a small tolerance value may result in creating a twin that consists of a string of 5 voxels. Therefore, a reasonable tolerance value is in the range of 0.5 to 2.

Figure 5 shows the various placements of the synthesized twin (colored in red) as the distance from the center of the grain D is changed. The green colored twin in Figure 5 represents a twin that is inserted with $D = 0$. From Figure 5, it is clear that the variable D is a direct representation of the number of voxels to shift the center of the twin grain from the center of the parent grain. As can be observed in Figure 5(c), as the D value approaches the radius of the parent grain, the twin grain approaches the edges of the parent grain.

IV. NICKEL-BASED SUPERALLOY IN100

Nickel-based superalloys, like IN100, are used predominately in aircraft turbine engines because they have high strength and creep resistance at increased temperatures.[18] It is forged, subsolvus heat-treated, and produced by powder metallurgy processing. IN100 has a fine grain size (~3 μm), which presents an advantage for EBSD in allowing many grains to be counted on the mesoscale; the length scale in between microscopic and macroscopic. The grain size remains fine during processing because of the existence of carbide particles that are ignored for these purposes. The processing results in two microconstituents for each grain, ~80 pct γ-phase matrix with ~20 pct γ'-phase coherent precipitates. This investigation will consider only one type of grain,[5] and the multiphase nature of the microstructure is ignored. IN100 exhibits a log-normal distribution, particularly in the mean field. A random orientation is observed for IN100, which is assumed here.[25,26]

Variation in the high-cycle fatigue (HCF) life of IN100 specimens depends on the presence of hot spots or regions with increased local driving force for damage formation.[27,28] Fatigue damage formation in polycrystalline materials is driven mainly by microstructural slip, which is exacerbated by interacting microstructure attributes like grains, phases, and voids.

Microstructurally small fatigue crack (MSFC) formation in polycrystalline superalloys has been linked to voids and pores. However, as processing techniques have improved, cracks are observed increasingly at crystallographic planes along which slip has concentrated. In another nickel-based superalloy (René 88DT), cracks were observed to form predominately in grains much larger than the average size.[29,30] This finding suggests the importance of extreme values (*i.e.,* upper tails) in generating RVEs for MSFCs. The SVEs of IN100 have been generated using Voronoi tessellation optimized by a simulated annealing algorithm.[31,32] Przybyla and McDowell[18] investigated 200 SVEs of IN100 for MSFCs at 923 K (650 °C) using crystal plasticity constitutive relations. Higher order moments of grain size were not matched in this study; thus a large number of instantiations were required to converge to a significant result. The GSD of IN100 is shown and discussed in the Section VI.

V. GRAIN SIZE DISTRIBUTION ANALYSIS

The traditional presentation of grain size distributions has used histograms with continuous functions overlaid for comparison with analytical forms for theoretical probability densities. Such histograms have been the basis for the common conclusion that polycrystal grain sizes are log-normally distributed. Probability plots are useful, however, both for showing correlation of sample distributions to theoretical distributions on a linearized basis and for examination of tails. In probability plots, the x-axis displays sample quantiles and the y-axis scales percentiles of a theoretical distribution; a standard normal distribution ($N[0,1]$: $\mu = 0$, $\sigma = 1$) comparing the logarithm of grain size. A sample set is identical to $N[0,1]$ if the plot follows the $x = y$ line. If the data follow some other straight line, then it is a linear transformation of $N[0,1]$. If the probability plot is flatter than $x = y$, then the distribution on the horizontal axis

is more dispersed than the one on the vertical axis. The arcs and "*S*"-shapes indicate relative skewness in the compared distributions. To interpret a probability plot, *e.g.*, if a data point falls on the 40th percentile of *N*[0,1], then 40 pct of *N*[0,1] samplings will be smaller and 60 pct will be larger. If this same data point corresponds to a sample value at the 50th percentile, then the data point will be closer to the sample axis, meaning that the sample data are larger than that which *N*[0,1] predicts for the 50th percentile.[33]

VI. MATCHING GRAIN SIZE DISTRIBUTION STATISTICS

The GSD mean field is defined in this context as the grains that fall within ± 2 standard deviations from the mean. These grains are controlled most readily by the input log-normal mean and standard deviation which generate each ellipsoid. Note that all the synthetic microstructures are assumed equiaxed and thus reduce the ellipsoid description to that of a sphere unless otherwise indicated. All the synthetic structures have dimensions 100 × 100 × 100 and thus have 1 million voxels. The IN100 measured structure shown in Figure 6 has 5861 grains with dimensions 389 × 146 × 184 (~10 million voxels). The dimensions for the synthetic microstructures are among the largest a serial implementation can handle given the relatively high input statistical variability. Also, all the synthetic microstructures in this study have a random distribution of crystallographic orientations, which is what is observed for IN100.[25,26]

A. *Grain Size Distribution Mean Field*

For the case of IN100, the mean is calculated from the measured grain size statistics. The IN100 statistics are summarized in Table II. The mean sphere equivalent radius for the measured IN100 structure is 5.35 μm, which is the target used in the code set to match the synthetic mean sphere equivalent radius with the IN100 mean sphere equivalent radius. In this study, 1 μm = 1 voxel.

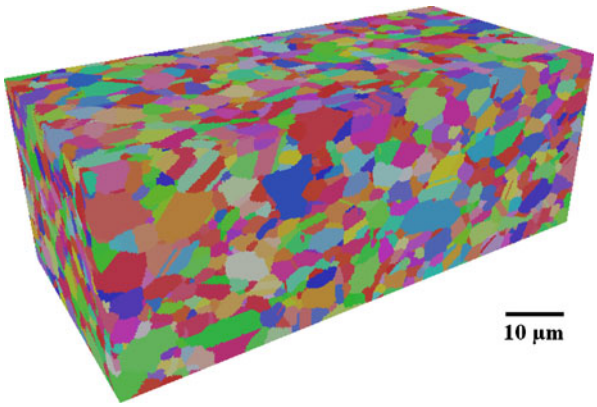


Fig. 6—Image of experimentally measured nickel-based superalloy IN100.[2,6]

1. *Instantiation 1*

The synthetic structure shown in Figure 7(a) has a mean sphere equivalent radius of 5.22 voxels. The resulting mean-matched synthetic microstructure probability plot is displayed in Figure 8(a). Notice that the mean-matched synthetic microstructure crosses the measured IN100 data at approximately zero as expected. Because Figure 8 displays the logarithm of the sphere equivalent radius over the mean sphere equivalent radius, *i.e.,* normalized by the mean, the new mean becomes one, equivalent to zero when transformed to logarithmic values. Because no attempt is made to match higher order statistics other than the mean in this instantiation, that is the only correspondence observed between the synthetic mean-matched and IN100 datasets.

2. *Instantiation 2*

To match the GSD standard deviation with the IN100 GSD standard deviation, the mean is enforced as in the first instantiation and for the second instantiation, the standard deviation is also enforced to generate the initial set of ellipsoids, again according to a log-normal distribution. As shown in Table II, the measured IN100 sphere equivalent radius standard deviation is 2.62. The IN100 mean and standard deviation are used as targets for the synthetic microstructure generator.

The second synthetic microstructure, with mean and standard deviation matched, is shown in Figure 7(b). Notice the variability in grain sizes that was introduced going from the mean-matched only first instantiation (Figure 7(a)). Also note the artificial grain boundaries created by matching the standard deviation and also the "island"-type grains appearance. These items are addressed in Section VII. The mean and standard deviation-matched probability plot is shown in Figure 8(a). The mean field grain data points now coincide with the IN100 measured mean field grain data points. However, the upper and lower tails do not correspond to the IN100 GSD because the latter does not follow a log-normal distribution. Of particular concern is the upper tail because of its fatigue implications, but both tails are addressed in the next section.

B. *Grain Size Distribution Tails*

The lower tail represents the limit of data collection. A log-normal distribution can predict too many small ellipsoids which results in too many small grains in the synthetic microstructure. By enforcing a minimum

Table II. Grain Size Statistics for Measured and Synthetic IN100

Sphere Equivalent Radius	Measured	Synthetic
Mean	5.35	5.17
Standard deviation	2.62	2.60
Minimum	1.79	0.52
Maximum	19.4	16.1
σ (Standard deviation of log (normalized size))	0.49	0.49

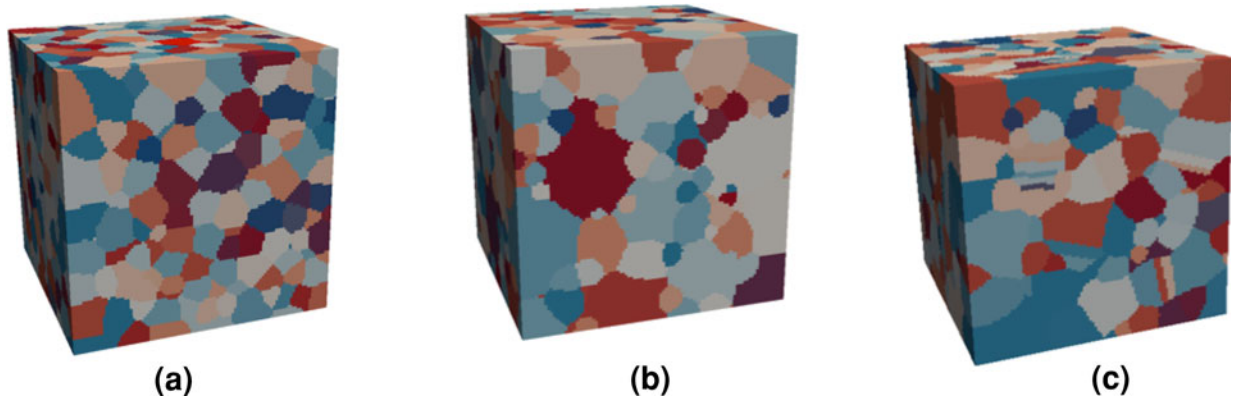


Fig. 7—Images of synthetic microstructures with grain size (*a*) mean matched; (*b*) mean and standard deviation matched; and (*c*) mean, standard deviation, minimum, and maximum matched with twins inserted and cleaned.

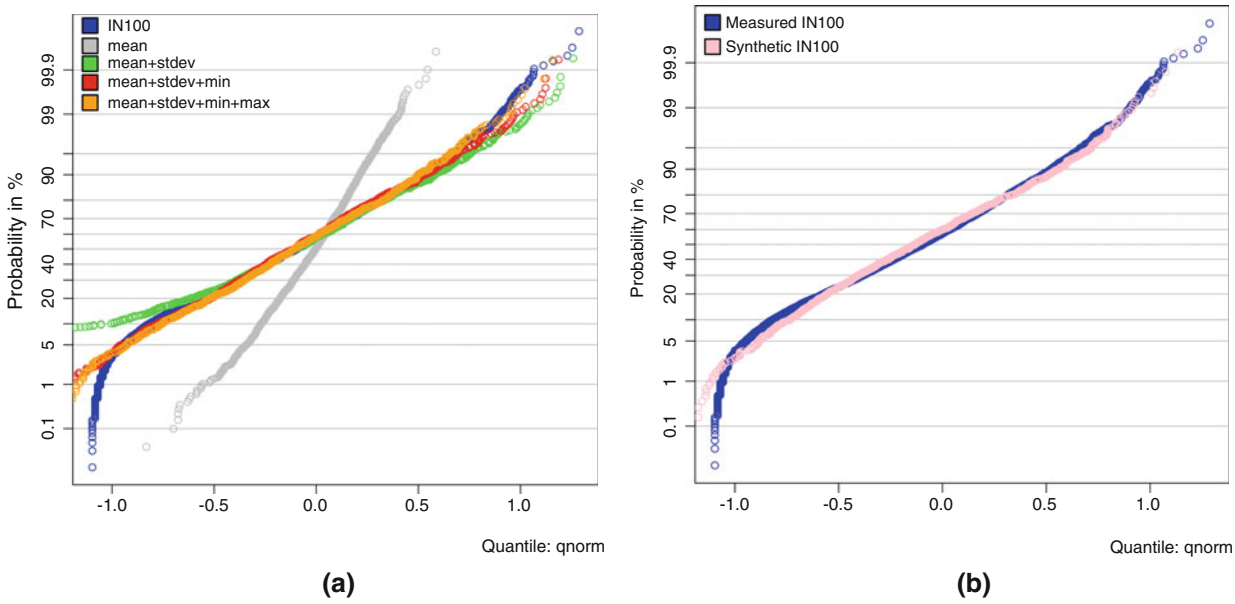


Fig. 8—Probability plot of (*a*) experimentally measured IN100 and synthetic microstructures with inclusively mean, standard deviation, minimum, and maximum matched synthetic microstructure; and (*b*) experimentally measured IN100 and mean, standard deviation, minimum, and maximum matched with twins inserted and cleaned synthetic microstructure.

threshold, the small grains are removed to match the limit of data collection. The algorithm proceeds by merging grains smaller than the input threshold with their majority nearest NN grain (Section II–G).

1. *Instantiation 3*

The probability plot of the third instantiation, with mean, standard deviation, and minimum-matched GSD, is displayed in Figure 8(a). Note that the lower tail now corresponds more closely to that of the IN100 GSD; a marked improvement was observed from the mean and standard deviation-matched microstructure.

It is apparent that grains smaller than the measured IN100 grains prevail in this third instantiation. Although it is important to reproduce grain sizes in the tails accurately, one would not expect the lower tail grains be "bad actors" in fatigue cracking. Additionally, the minimum threshold is effective in removing many island-type grains that were present in the standard deviation matched second instantiation.

2. *Instantiation 4*

Examining the probability plot of the inclusively matched mean, standard deviation, minimum, and maximum (size-truncated) synthetic structure in Figure 8(a) compared with the nonmaximum matched synthetic instantiation, the upper tail now more closely matches the upper tail of the measured IN100 microstructure. This is a significant result because these few data points in the upper tail represent the largest grains in the

volume element. These grains have been observed to be the "bad actors" in fatigue cracking.

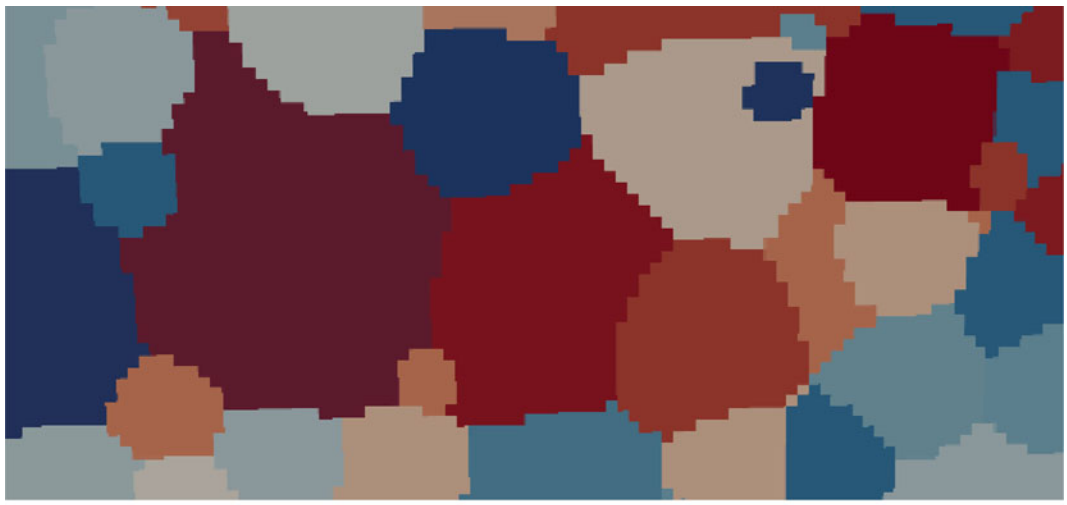
VII. MICROSTRUCTURE CLEANUP

The previous section described a series of progressive improvements that results in the size statistics of a synthetic microstructure being well matched to an IN100 measured microstructure. Now, a collection of additional processes is described that serve to clean up the size statistically matched synthetic microstructure by improving the overall grain shape while not significantly perturbing the size distribution.

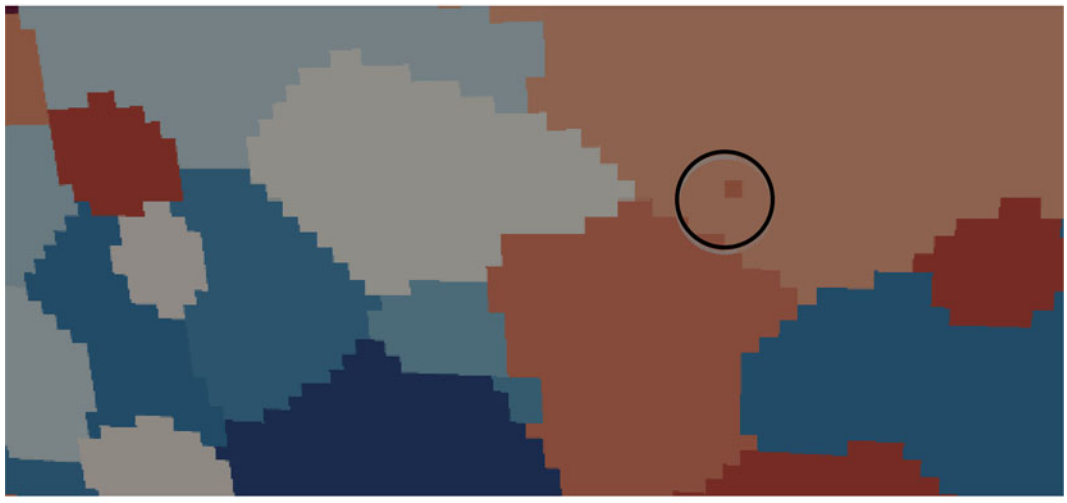
As shown by Figures 7(b) and 9(a), many grains bear a strong resemblance to their simple geometric origin, *i.e.,* spheres. Although the grains of IN100 are assumed equiaxed, by enforcing spheres, the final synthetic microstructure results in artificial boundaries. To provide more grain boundary "realism," the semiaxes are allowed to vary ± 20 pct randomly, resulting in a full scalene ellipsoid description. These ellipsoids can now overlap and impinge on one another with a greater range of shape.

When dealing with spheres, the axis orientation distribution function (ODF) was irrelevant. Since the grain geometry is now an ellipsoid, by applying a random axis ODF, the ellipsoids vary in their alignment with the edges of the simulation domain and the final structure exhibits more "realism." This claim of improved realism is acknowledged to be qualitative, and future publications will address its quantification. The Potts Model is employed in postprocessing to coarsen the grains and "relax" the boundaries. Five MC steps were performed on the volume element in Figure 7(c).

As mentioned previously, through the size statistic and packing matching processes, "island"-type grains are introduced (Figure 9(b)). Many of these grains seem



(a)



(b)

Fig. 9—Close-up images of Fig. 7(b) depicting (*a*) artificial grain boundary curvature and (*b*) an "island"-type grain.

to have only one NN; however, they may have subsurface NN. In any case, they are undesirable because they do not correspond to observed microstructures. Similar to the minimum threshold, these low-NN grains are "eroded" by merging them with their majority NN parent grain. A NN threshold of 7 was enforced for the synthetic microstructure in Figure 7(c).

The nickel-based superalloy IN100 has a high density of annealing twins. The final step consists of inserting twins into the synthetic microstructure, which is detailed in Section III.[34]

The final cleaned, mean field- and tail-matched GSD probability plot is shown in Figure 8(b) compared with the measured IN100. The matched statistics are displayed in Table II. The GSD from the uncleaned synthetic microstructure is preserved through the clean-up steps while vastly improving the grain shape, comparing Figures 7(b) and (c).

VIII. SUMMARY

The log-normal distribution fits the synthetic and IN100 sample grain size data very well close to the mean. However, neither the upper nor the lower tails fit well over multiple instantiations. Preprocessing and postprocessing tools were therefore introduced to tailor the tails of the GSD extracted from the synthetic microstructure and to match them to experimental data sets. Cleanup steps improve the grain shape and help pass the "eye test."

Matching just the mean sphere equivalent radius does not even result in a mean field SVE. Matching the mean and standard deviation sphere equivalent radius does result in a mean field SVE. This is the point where many modelers have stopped their statistic matching efforts and used hundreds of instantiations to converge to a significant result.[18] However, the importance of extreme values to crystallographic fatigue damage formation is well documented,[18] thus motivating the need for better control and quantification of the tails of synthetic GSDs. Matching the minimum grain size through a minimum threshold erosion technique represents matching the limit of data collection. This was shown to affect the entirety of the lower tail and permitted us to match the experimental microstructure of interest. Matching the upper tail through log-normal distribution truncation represents matching the largest grain size in a microstructure, which represents the common "bad actors" in fatigue cracking. Although this technique only enforces a maximum, its affect is felt throughout the whole upper tail because the resultant resampling of the log-normal distorts the log-normal distribution in such a way that results in it matching the upper tail of our microstructure of interest. At this point, a synthetic microstructure exists that matches the available size statistics. Shape statistics are also important in 3-D microstructures. The algorithm makes no attempt to match or track grain shape throughout the course of generation. Nevertheless certain tools used in post-processing also serve to improve the final grain shapes. The tools include a combination of MC Potts model, NN threshold, randomizing ellipsoid semi-axis alignments and length, and inserting twins correct for "island"-type grains, artificial grain boundary curvature, and microstructural features. These additional steps are compatible with the constraints that were applied to match the GSD over the available range of grain size, resulting in a synthetic 3-D microstructure that also passes the "eye test."

ACKNOWLEDGMENTS

This work was supported, in part, by the PETTT program from the High Performance Computing Modernization Office, the Air Force Office of Scientific Research, the General Electric Company, and the NSF-supported Carnegie Mellon MRSEC. Discussions with Mike Jackson and Jason Gruber are gratefully acknowledged.

APPENDIX A

The Method for Creating Annealing Twins in the Structure to Match Σ3 Cluster Distributions.

```
Input: Structure, list of target cluster size, and list of preexisting Σ3 clusters
Output: Structure with annealing twins
for number of target clusters, do
   Define target cluster size from input
   if current cluster size = 0
        if preexisting cluster exist and is smaller than target cluster size, then
           adopt preexisting cluster into current cluster list
           insert twins in the grains that reside in the preexisting cluster
        else
           randomly pick a grain and insert twins
        end
    else
        while current cluster size < target size do
           check all boundaries surrounding cluster to ensure all Σ3 have been
accounted for
           for neighbors of all grains in current cluster do
             if neighbor is part of preexisting cluster and adoption will not exceed
target size, do
                 adopt neighbor and preexisting cluster into current cluster list
                 insert twins in the newly adopted grains
             else
                 randomly pick a neighbor to adopt and insert twin
             end
           end for
       end while
   end
end for
```

REFERENCES

1. *Integrated Computational Materials Engineering: A Transformational Discipline for Improved Competitiveness and National Security*, The National Academies Press, Washington, DC, 2008.
2. D.M. Saylor, J. Fridy, B.S. El-Dasher, K. Jung, and A.D. Rollett: *Metall. Mater. Trans. A*, 2004, vol. 35A, pp. 1969–79.
3. A. Brahme, M.H. Alvi, D. Saylor, J. Fridy, and A.D. Rollett: *Scripta Mater.*, 2006, vol. 55, pp. 75–80.
4. M.D. Uchic, M.A. Groeber, D.M. Dimiduk, and J. Simmons: *Scripta Mater.*, 2006, vol. 55, pp. 23–28.
5. M. Groeber, S. Ghosh, M.D. Uchic, and D.M. Dimiduk: *Acta Mater.*, 2008, vol. 56, pp. 1257–73.
6. M. Groeber, S. Ghosh, M.D. Uchic, and D.M. Dimiduk: *Acta Mater.*, 2008, vol. 56, pp. 1274–87.
7. Z. Fan, Y. Wu, X. Zhao, and Y. Lu: *Comp. Mater. Sci.*, 2004, vol. 29, pp. 301–08.
8. L. St-Pierre, E. Heripre, M. Dexet, J. Crepin, G. Bertolino, and N. Bilger: *Int. J. Plast.*, 2008, vol. 24, pp. 1516–32.
9. M.F. Vaz and M.A. Fortes: *Scripta Metall.*, 1988, vol. 22, pp. 35–40.
10. B. Raeisinia and C.W. Sinclair: *Mater. Sci. Eng. A*, 2009, vol. 525, pp. 78–82.
11. T. Luther and C. Konke: *Eng. Frac. Mech.*, 2009, vol. 76, pp. 2332–43.
12. Y. Bandari, S. Sarkar, M. Groeber, M.D. Uchic, D.M. Dimiduk, and S. Ghosh: *Comp. Mater. Sci.*, 2007, vol. 41, pp. 222–35.
13. S. Ghosh, Y. Bhandari, and M. Groeber: *Comput. Aided Des.*, 2008, vol. 40, pp. 293–310.
14. T. Kanit, S. Forest, I. Galliet, V. Mounoury, and D. Jeulin: *Int. J. Solid Struct.*, 2003, vol. 40, pp. 3647–79.
15. M. Ostoja-Starzewski: *J. Appl. Mech.*, 2002, vol. 69, pp. 25–35.
16. S. Swaminathan, S. Ghosh, and S. Pagano: *J. Comp. Mater.*, 2006, vol. 40, pp. 583–604.
17. S. Swaminathan and S. Ghosh: *J. Comp. Mater.*, 2006, vol. 40, pp. 605–21.
18. C.P. Przybyla and D.L. McDowell: *Int. J. Plast.*, 2010, vol. 26, pp. 372–94.
19. S. Sintay: Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA, 2009.

20. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery: *Numerical Recipes. The Art of Scientific Computing*, 3rd ed., Cambridge University Press, New York, NY, 2007.
21. Z.T. Dimov: *Monte Carlo Methods for Applied Scientists*, World Scientific Publishing, Hackensack, NJ, 2008, pp. 13–14.
22. A.D. Rollett and P. Manohar: *Continuum Scale Simulation of Engineering Materials*, D. Raabe and F. Roters, eds., Wiley, New York, NY, 2004.
23. L.A. Barrales-Mora: *2D and 3D Grain Growth Modeling and Simulation*, Cuviller Verlag Gottinggen, Göttingen, Germany, 2008, pp. 24–25.
24. R. Fenn: *Geometry*, Springer-Verlag, London, UK, 2001.
25. J.K. Mackenzie: *Biometrika*, 1957, vol. 44, pp. 205–10.
26. J.K. Mackenzie: *Biometrika*, 1958, vol. 45, pp. 229–40.
27. D.L. McDowell: *Int. J. Fract.*, 1996, vol. 80, pp. 103–45.
28. S. Suresh: *Fatigue of Materials*, Cambridge University Press, New York, NY, 1998.
29. S.K. Jha, M.J. Larsen, J.M. Rosenberg, A.H. Li, and W.J. Porter: *Materials Damage Prognosis*, TMS, Warrendale, PA, 2005.
30. A. Shyman, C.J. Torbet, S.K. Jha, J.M. Larsen, M.J. Caton, C.J. Szczepanski, T. M. Pollock, and J.W. Jones: *Development of Ultrasonic Fatigue for Rapid, High Temperature Fatigue Studies in Turbine Engine Materials*, Champion, PA. 2004.
31. M. Zhang, J. Zhang, and D.L. McDowell: *Int. J. Plast.*, 2006, vol. 23, pp. 1328–48.
32. M. Shenoy, J. Zhang, and D.L. McDowell: *Fatig. Fract. Eng. Mater. Struct.*, 2007, vol. 30, pp. 889–904.
33. J.L. Devore: *Probability and Statistics for Engineering and the Sciences*, Brooks/Cole, Florence, KY, 2009, pp. 170–73.
34. L. Chan: Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA, 2010.