

# Fitting Laguerre tessellation approximations to tomographic image data

A. Spettl<sup>a\*</sup>, T. Brereton<sup>a</sup>, Q. Duan<sup>b</sup>, T. Werz<sup>c</sup>, C. E. Krill III<sup>c</sup>, D. P. Kroese<sup>b</sup> and V. Schmidt<sup>a</sup>

<sup>a</sup>*Institute of Stochastics, Ulm University, 89069 Ulm, Germany;* <sup>b</sup>*School of Mathematics and Physics, The University of Queensland, Brisbane 4072, Australia;* <sup>c</sup>*Institute of Micro and Nanomaterials, Ulm University, 89069 Ulm, Germany*

The analysis of polycrystalline materials benefits greatly from accurate quantitative descriptions of their grain structures. Laguerre tessellations approximate such grain structures very well. However, it is a quite challenging problem to fit a Laguerre tessellation to tomographic data, as a high-dimensional optimization problem with many local minima must be solved. In this paper, we formulate a version of this optimization problem that can be solved quickly using the cross-entropy method, a robust stochastic optimization technique that can avoid becoming trapped in local minima. We demonstrate the effectiveness of our approach by applying it to both artificially generated and experimentally produced tomographic data.

**Keywords:** image processing; microstructural characterization; grain boundary structure; polycrystalline; power diagram; inverse problem; cross-entropy method

## 1. Introduction

In recent years there have been significant advances in the tomographic characterization of materials. As a result, it is now possible to carry out detailed investigations of the 3D grain structures of polycrystalline materials; see, e.g., [1–7]. A fundamental ingredient in any such investigation is a suitable quantitative description of the grain morphology. Such a description contains the key features of the structure, ideally free from noise and imaging artifacts. A good description usually results in significant data compression, describing large 3D voxel data sets using only a small number of parameters. Data compression is necessary, for example, when carrying out analysis of sequences of tomographic data sets (e.g., the high time resolution in-situ synchrotron images considered in [8]). In addition, the description of tomographic data via tessellations provides a basis for the statistical analysis of grain structures and, in some cases, can be used to develop stochastic models of material microstructures; see, e.g., [8–16].

The most commonly used quantitative descriptions of space-filling grain ensembles are based on tessellations, which divide the space into disjoint regions called *cells*.

---

\*Corresponding author. Email: aaron.spettl@uni-ulm.de

The cells represent the individual grains. The most widely used tessellation model is the *Voronoi tessellation* (see, e.g., [17, 18]), which takes, as parameters, a set of generating points. The space is then divided into convex cells by assigning each point to its nearest generator. The *Laguerre tessellation* (see, e.g., [17, 18]) is a generalization of the Voronoi tessellation that also partitions the space into convex cells. Like the Voronoi tessellation, the Laguerre tessellation is generated by a set of points; however, unlike the Voronoi tessellation, these points are weighted, with the weights influencing the size of the cells. Consequently, the Laguerre tessellation is able to describe a wider range of structures than the Voronoi tessellation. For this reason, the Laguerre tessellation is a popular choice for modeling polycrystalline grain structures [8, 9, 19–22] and other materials, such as foams [12, 23, 24].

In order to describe a tessellation by a set of generating points, it is necessary to solve an inverse problem: that is, a set of generating points that produce the observed cells must be found. The *Voronoi inverse problem* (VIP) is well-studied; see, for example, [25–30]. Recently, Duan et al. [31] proposed an algorithm that finds solutions to the *Laguerre inverse problem* (LIP). Although the examples considered in [31] are restricted to 2D, the methodology is easily applied in higher dimensions.

The solutions to the VIP and the LIP assume that the empirical data constitute perfect descriptions of the observed cells. However, this is not true when working with tomographic data, which is distorted by noise and also contains imprecision arising from discretization during the imaging process. It is also worth noting that real-world materials are not perfectly described by Laguerre tessellations (though the descriptions can be quite good). These limitations mean that methods that attempt to invert a tessellation extracted from the tomographic data do not, in general, result in good fits. The LIP is solved by iteratively finding the generating points of the given tessellations. When applied to imperfect data, this iterative procedure propagates errors, resulting in tessellations that have little correspondence to the tomographic data. Thus, when dealing with empirical data, it is not appropriate to attempt to solve the LIP. Instead, the generating points of a Laguerre tessellation that is a good approximation of the material must be found. This is, at its core, an optimization problem. We call this problem the *Laguerre approximation problem* (LAP). The corresponding Voronoi approximation problem has been considered in the literature, beginning with [32].

A simple heuristic approach for solving the LAP was proposed in [6]. More sophisticated approaches, which formulate and solve an optimization problem, are described in [33–35]. Although these techniques provide good fits in certain settings, they are either limited to small sample sizes or require the considered tessellations to be sufficiently regular.

In this paper, we present a fast and robust algorithm for fitting Laguerre approximations to large data sets. More precisely, we formulate an optimization problem where we minimize the discrepancy between the grain boundaries observed in the

image data and the grain boundaries produced by our Laguerre approximation. The cost function is chosen so that it can be evaluated very efficiently and that all necessary information can be easily obtained from image data. We then solve the optimization problem using the cross-entropy (CE) method [36–39], a stochastic optimization algorithm that is able to escape local minima. We carry out experiments on both real and artificially-generated image data that show our approach is able to produce very good fits.

This paper is structured as follows. In Section 2, we review some key properties of Laguerre tessellations. In Section 3, we give a more complete description of the LAP and formulate our optimization problem. Then, in Section 4, we introduce the CE method as a robust tool for solving this optimization problem. Section 5 gives results for both artificial and experimental data that demonstrate the effectiveness of our approach. Finally, Section 6 summarizes our results and suggests directions for further research.

## 2. The Laguerre tessellation

In the following section, we define Voronoi and Laguerre tessellations and give a number of properties that we will use to solve the LAP. For notational convenience, we only consider tessellations in  $\mathbb{R}^3$ . However, our methods are easily applied in other dimensions.

The Voronoi tessellation is defined by a locally finite set of generating points,  $\{\mathbf{x}_i\}_{i \in \mathcal{I}}$ , where  $\mathcal{I} \subseteq \mathbb{N}$  denotes the index set consisting of natural numbers. The cell corresponding to the  $i$ th generating point,  $\mathbf{x}_i \in \mathbb{R}^3$ , is given by

$$C_i = \{\mathbf{y} \in \mathbb{R}^3 : \|\mathbf{y} - \mathbf{x}_i\| \leq \|\mathbf{y} - \mathbf{x}_j\| \text{ for all } j \in \mathcal{I}\}, \quad (1)$$

where  $\|\cdot\|$  is the Euclidean norm on  $\mathbb{R}^3$ . In the case of the Laguerre tessellation, the generating points,  $\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}$ , are weighted, where we assume that  $r_i > 0$ . The cells of the Laguerre tessellation,  $\{C_i\}_{i \in \mathcal{I}}$ , are then defined by

$$C_i = \{\mathbf{y} \in \mathbb{R}^3 : \text{pow}(\mathbf{y}, (\mathbf{x}_i, r_i)) \leq \text{pow}(\mathbf{y}, (\mathbf{x}_j, r_j)) \text{ for all } j \in \mathcal{I}\},$$

for all  $i \in \mathcal{I}$ , where the Euclidean norm used in (1) is replaced by the so-called power distance

$$\text{pow}(\mathbf{y}, (\mathbf{x}, r)) = \|\mathbf{y} - \mathbf{x}\|^2 - r^2.$$

Given the generator points, the faces of the cells can be computed efficiently; see, e.g., [40–42]. As we take the weights  $r_i$  to be positive real numbers, the generating points have a geometric interpretation as spheres. That is, we can represent a

generating point,  $(\mathbf{x}_i, r_i)$ , as a sphere with center  $\mathbf{x}_i$  and radius  $r_i$ .

The flexibility of the Laguerre tessellation comes at a cost. For example, while each cell of a Voronoi tessellation contains its generating point, the generating points of a Laguerre tessellation may not be contained in their corresponding cells. In some cases, a generating point of a Laguerre tessellation may not even produce a cell. That is, it is possible that there exists a point,  $(\mathbf{x}_i, r_i)$ , such that  $C_i = \emptyset$ ; see, for example, [40]. In addition, while the cells of a Voronoi tessellation uniquely determine its generating points, there are uncountably many sets of generating points that can generate a given Laguerre tessellation; see, [31, 43, 44]. Thus, while the VIP has a unique solution, the LIP has uncountably many solutions. Under mild conditions, however, it can be shown that the generating points of a Laguerre tessellation are uniquely determined given one generating point and the weight (or one coordinate) of a generating point in an adjacent cell; see [31].

We will often find it convenient to consider planes that are equidistant from two generating points (under their respective power distances). That is, given two generating points,  $(\mathbf{x}_i, r_i)$  and  $(\mathbf{x}_j, r_j)$ , we consider the separating plane given by

$$P_{i,j} = \{\mathbf{y} \in \mathbb{R}^3 : \mathbf{n}_{i,j}^\top \mathbf{y} + b_{i,j} = 0\},$$

where

$$\mathbf{n}_{i,j} = \frac{2 \cdot (\mathbf{x}_j - \mathbf{x}_i)}{\|2 \cdot (\mathbf{x}_j - \mathbf{x}_i)\|}$$

is the unit normal vector of  $P_{i,j}$  and

$$b_{i,j} = \frac{\|\mathbf{x}_i\|^2 - \|\mathbf{x}_j\|^2 + r_j^2 - r_i^2}{\|2 \cdot (\mathbf{x}_j - \mathbf{x}_i)\|}$$

is the shortest distance from the origin to the plane; see, e.g., [43, Section 2.1] for more details. The plane  $P_{i,j}$  defines a half-space

$$H_{i,j} = \{\mathbf{y} \in \mathbb{R}^3 : \mathbf{n}_{i,j}^\top \mathbf{y} + b_{i,j} \leq 0\},$$

which covers the cell  $C_i$ . Note that the intersection of all such half-spaces,  $\bigcap_{k \in \mathcal{I} \setminus \{i\}} H_{i,k}$ , defines the cell  $C_i$ . Because the normal vectors are taken to be unit vectors, the distance from an arbitrary point,  $\mathbf{x} \in \mathbb{R}^3$ , to the plane  $P_{i,j}$  is given by

$$d(\mathbf{x}, P_{i,j}) = |\mathbf{n}_{i,j}^\top \mathbf{x} + b_{i,j}|. \quad (2)$$

### 3. The Laguerre approximation problem (LAP)

Given an exact description of a Laguerre tessellation (e.g., in terms of the half-spaces defined above), it is not too difficult to solve the LIP. That is, it is straightforward to find a set of weighted points that is able to generate the given tessellation. Furthermore, these points can be chosen to satisfy certain constraints; see [31]. When dealing with tomographic data, however, the description of the tessellation is not exact. This is because, even if the material itself can be perfectly described by a Laguerre tessellation, the noise and discretization errors inherent in the imaging process mean that the cell boundaries extracted from the data are subject to error. The LIP can be solved as described in [31]. However, the iterative computation of generators is sensitive to imperfect data and errors propagate quickly. As a result, the ensuing tessellations often do not correspond well to the tomographic data. Therefore, solving the LIP for empirical data is generally ill advised. Instead, we wish to find generating points that produce a Laguerre tessellation that is as close as possible to the tomographic data (with respect to some metric). This is, at its core, an optimization problem. In order to properly formulate this optimization problem, a discrepancy measure needs to be defined. We then choose the generating points of the approximating tessellation in order to minimize this discrepancy. The choice of discrepancy measure depends on the nature of empirical data. In this paper, we work with tomographic data.

#### 3.1. *Tomographic image data*

We assume that the tomographic image data constitutes a collection of voxels in a convex window. Furthermore, we assume that the image has already been segmented — e.g., by the watershed transform; see [45, 46] — and that the voxels have been labeled by their corresponding grains. Thus, the data is of the form  $\{I(x, y, z) \in \{0, \dots, N\} : (x, y, z) \in W\}$ , where  $N \geq 1$  denotes the number of grains and  $W \subset \mathbb{N}^3$  is a grid of voxel coordinates. The  $i$ th grain region,  $R_i$ , is then given by  $R_i = \{(x, y, z) \in W : I(x, y, z) = i\}$ . Note that  $R_0$  does not correspond to an actual grain, but is either the empty set or a collection of one voxel thick layers that separate grains. Such thin layers often arise in segmentation procedures such as the watershed transformation. A grain region  $R_i$  may correspond to the grain itself or a region that contains the grain (if the space is not completely filled with grains), cf. Figure 1. We assume that the grain regions are roughly convex and that the segmentation is of a high quality.

#### 3.2. *Discrepancy measures*

The most direct way to measure the discrepancy between 3D image data and a Laguerre tessellation generated by the points  $\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}$  is by counting the number

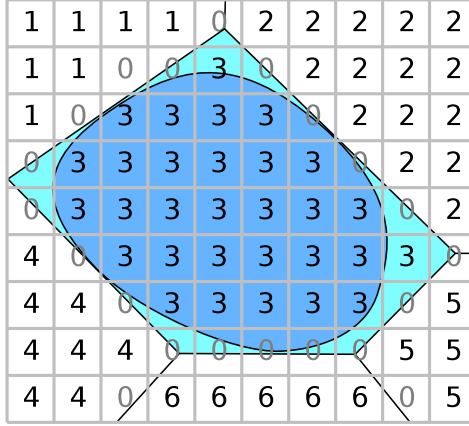


Figure 1. Illustration of voxels labeled by their corresponding grains. The grain regions are assumed to form a tessellation with convex cells. The grains themselves may not fill their corresponding regions completely, as depicted for grain number 3. Here, we show the grain region (i.e., voxels with label 3), a cell approximating the grain region (colored cyan) and the grain itself (colored blue).

of incorrectly assigned voxels. That is, we calculate

$$\mathcal{D}_{\text{vox}}(\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}) = \# \left\{ (x, y, z) \in W : I(x, y, z) > 0, I(x, y, z) \neq \tilde{I}(x, y, z) \right\},$$

where  $\tilde{I}$  is a discretized version of the tessellation generated by  $\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}$  and  $\#A$  denotes the number of elements in some set  $A$ . Many other discrepancy measures considered in the literature are of a similar form. For example, in [35], the difference between a cell in the approximating tessellation and its equivalent in the empirical data is measured by the intersection of the approximating cell with the corresponding adjacent cells in the empirical data. Minimizing the total overlap of the approximating cells is equivalent to minimizing the number of incorrectly assigned voxels. We call discrepancy measures that aim to minimize the volume of the difference between the empirical and approximating tessellations *volume-based* measures.

Although minimizing a volume-based measure gives a very good fit to the data, such measures are poorly suited to most numerical optimization methods. This is because evaluating such a discrepancy measure is computationally expensive. For example, when considering the number of incorrectly assigned voxels, a discretized version of the approximating tessellation needs to be generated. In practice, this limits the number of times such a discrepancy measure can be evaluated. As a result, approaches that aim to minimize volume-based discrepancies are restricted to small data sets with small numbers of grains (e.g., 109 grains in [34]) or are forced to use non-optimal optimization techniques, such as gradient-descent, which do not require too many evaluations of the discrepancy and, as such, limit the time taken by the fitting procedure (ideally to substantially less than 24 hours).

As we propose to use stochastic optimization techniques to solve this problem, we

need a discrepancy measure that is significantly faster to evaluate. We can achieve this by considering an *interface-based* discrepancy measure — a measure that considers only the boundaries between cells — instead of a volume-based one. If an approximating tessellation can accurately reproduce these interfaces, it will be a good approximation of the data. The primary advantage of interface-based discrepancy measures is that they can be calculated from the generating points of the approximating tessellation without the need to generate the tessellation itself.

In order to define such a discrepancy measure, we consider the sets of voxels that separate adjacent cells. We define the interface between two grains,  $i$  and  $j$ , in the empirical data by

$$N_{i,j} = \{(x, y, z) \in W : \mathcal{N}_{26}(x, y, z) \cap R_i \neq \emptyset \text{ and } \mathcal{N}_{26}(x, y, z) \cap R_j \neq \emptyset\},$$

where  $\mathcal{N}_{26}(x, y, z)$  denotes the 26-neighborhood of  $(x, y, z)$  — i.e., the voxels  $(x', y', z')$  that have a distance less than or equal to  $\sqrt{3}$  from  $(x, y, z)$ . This set contains all voxels that touch both grains. Note that  $N_{i,j} = \emptyset$  if the grains are not adjacent. If the tessellation generated by  $\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}$  is a good approximation to the empirical data, then the plane separating the generating points of adjacent cells  $i$  and  $j$  (which defines the boundary of the two cells) should be close to  $N_{i,j}$ . Thus, we measure the distance between  $N_{i,j}$  and  $P_{i,j}$ , the plane separating cells  $i$  and  $j$  in the approximating tessellation. The discrepancy between the approximating tessellation and the empirical data is then given by the sum of squares of these distances. That is, we define

$$\mathcal{D}(\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}) = \sum_{j=1}^N \sum_{k=j+1}^N \sum_{\mathbf{y} \in N_{j,k}} d(\mathbf{y}, P_{j,k})^2, \quad (3)$$

where  $d(\mathbf{y}, P_{j,k})$  is given in (2) and  $\{P_{j,k}\}_{j,k \in \mathcal{I}}$  is the set of separating planes determined by the generating points  $\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}$ . Note that this discrepancy measure can be calculated without generating the approximating tessellation. A separating plane can be computed for all pairs of generators (or generator candidates, when performing the optimization) — even if their cells are empty or not adjacent in the Laguerre tessellation. Therefore, such ‘degenerate’ cases are not a problem for our approach. By matching all separating planes to the corresponding test points at once, Laguerre cells computed based on ‘good’ configurations of generators will match the desired cell configuration well enough.

### 3.3. Minimizing the discrepancy

When aiming to minimize a volume-based discrepancy, the LAP reduces to the optimization problem of finding

$$\{(\mathbf{x}_i^*, r_i^*)\}_{i \in \mathcal{I}} = \arg \min_{\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}} \mathcal{D}_{\text{vox}}(\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}). \quad (4)$$

There are a number of significant difficulties which must be overcome in solving (4). In particular,

- (i) the optimization problem is high-dimensional,
- (ii) the optimization problem has many local minima,
- (iii) the discrepancy is expensive to evaluate.

The approach developed in [34] avoids problems (ii) and (iii) by finding the exact solution of a linear program (an optimization problem where a linear cost function is minimized subject to linear constraints). However, when the LAP is transformed into a linear program, the size of the resulting problem grows very quickly in both the number of voxels and the number of grains. This means that the linear programming approach cannot be applied to large 3D data sets containing many grains.

In [35], gradient-descent methods are used to obtain Laguerre approximations (with a relatively low number of evaluations of the discrepancy). Unlike the linear programming approach, the dimensionality of the optimization problem grows only linearly in the number of grains (and does not depend on the number of voxels). Thus, [35] avoids problem (iii) and, to a lesser extent, problem (i). However, the gradient-descent methods used there become stuck in local minima. Thus, the approach considered in [35] is reliant on the ability to find initial conditions that are close to global optima. Although it is often possible to find good initial conditions — e.g., when fitting tessellations to foams with regular structures — this is not always the case, as will be seen in Section 5. In addition, even when good initial conditions can be found, the quality of the approximating tessellation can usually be significantly improved when the optimization algorithm is able to escape local minima.

Stochastic optimization methods are widely used tools for solving high-dimensional optimization problems with many local minima; see [38] for an overview. Such methods have been used to solve problems related to the LAP. In [31], the CE method is used to find a solution to the LIP and, in [33], simulated annealing is used to fit 3D Laguerre tessellations to 2D image data. However, no stochastic approach has yet been proposed in order to solve the LAP for large, noisy 3D data sets.

In our approach to solving the LAP for large tomographic data sets, we use the CE method to minimize an interface-based discrepancy measure. The CE method has



been widely applied to high-dimensional multi-extremal optimization problems; see, [36–39]. It has a number of advantages over other stochastic optimization methods. For example, simulated annealing (described in [38, 47]) cannot be easily applied to the LAP, as it is very difficult to find an appropriate cooling schedule. In addition, the CE method can be easily parallelized.

By using the CE method to minimize an interface-based discrepancy measure, we have a method that can escape local minima and solves a problem whose dimensionality grows linearly in the number of grains. Although the discrepancy is expensive to evaluate, we reduce this cost significantly by minimizing an interface-based measure. In addition, we further reduce the cost of calculating the discrepancy by replacing  $\mathcal{D}$  defined in (3) with an approximation,  $\tilde{\mathcal{D}}$ .

### 3.4. Approximating the discrepancy

The interface-based discrepancy,  $\mathcal{D}$ , measures the distance between voxels on the boundaries between cells in the empirical data and the planes separating the generating points in the approximating tessellation. In order to calculate this discrepancy, every voxel in each grain boundary is considered. Although this is much faster to calculate than a volume-based discrepancy, it is still computationally expensive. However, very good approximations of  $\mathcal{D}$  can be obtained by instead considering sets of *test points* that describe the interfaces between the grains. We find that test points obtained by fitting approximating planes to the interfaces, using the orthogonal regression approach introduced in [48], work very well.

In order to calculate the test points, we consider only the points separating the two cells being considered (i.e., we ignore points of contact between three or more grains). Thus, instead of  $N_{i,j}$ , we consider

$$N_{i,j}^* = \{(x, y, z) \in N_{i,j} : (x, y, z) \notin N_{i,k} \text{ and } (x, y, z) \notin N_{k,j}, k \in \{1, \dots, N\}, k \notin \{i, j\}\}.$$

This avoids some numerical issues that could arise when using orthogonal regression.

We then approximate the boundary between the  $i$ th and  $j$ th grain using the plane that minimizes the total squared distance to all points in  $N_{i,j}^*$ . Determining this plane is a least-squares problem, which we solve using singular value decomposition. The approximating plane passes through the centroid,  $\mathbf{c} \in \mathbb{R}^3$ , of the set  $N_{i,j}^*$ . We obtain a normal vector,  $\mathbf{n} \in \mathbb{R}^3$ , to the plane by taking the right-singular vector corresponding to the smallest singular value of the matrix containing the voxel coordinates in  $N_{i,j}^*$  shifted by  $-\mathbf{c}$ . For more details on the singular value decomposition approach, see [49]. The test points,  $T_{i,j} \subset \mathbb{R}^3$  are chosen to lie on this approximating plane. More precisely, we put a circle with radius  $\sqrt{\#N_{i,j}^*}/4$  around the centroid  $\mathbf{c}$  with the same orientation as the plane. We then place 10 test points equidistantly on this circle. An illustration is given in Figure 2.

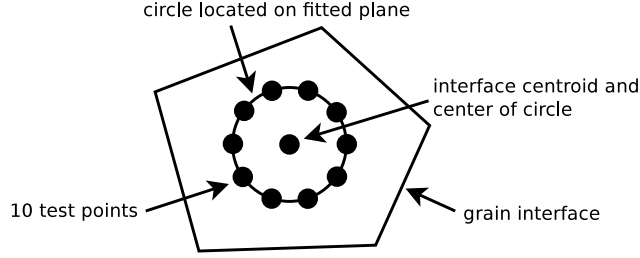


Figure 2. Illustration of test points construction. First, a circle is fitted to a grain interface. The center of the circle is given by the centroid of the voxels adjacent to both grains, its 3D orientation is obtained from the orthogonal regression plane. The radius is chosen with respect to the interface area. Finally, 10 test points are placed equidistantly on this circle.

In cases where the estimation of the plane is not ‘stable’ (i.e., the number of voxels in  $N_{i,j}^*$  is too low to determine the location and orientation), we simply use the centroid  $\mathbf{c}$  for all 10 test points. The criterion for ‘stable’ is simple: the smallest singular value must be smaller than half the second largest one. This way we can be confident that the third and shortest axis (which is perpendicular to the plane) is correctly identified.

Given the test points for the boundary between adjacent grains  $i$  and  $j$ , we define the approximate discrepancy at the boundary by

$$\tilde{\mathcal{D}}_{i,j}(\mathbf{x}_i, r_i, \mathbf{x}_j, r_j) = \sum_{\mathbf{y} \in T_{i,j}} d(\mathbf{y}, P_{i,j})^2,$$

where  $d(\mathbf{y}, P_{j,k})$  is given in (2) and  $P_{i,j}$  is the plane separating the generating points  $(\mathbf{x}_i, r_i)$  and  $(\mathbf{x}_j, r_j)$ . The total approximate discrepancy is then given by

$$\tilde{\mathcal{D}}(\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}) = \frac{1}{\sum_{j=1}^N \sum_{k=j+1}^N \#T_{j,k}} \sum_{j=1}^N \sum_{k=j+1}^N \tilde{\mathcal{D}}_{j,k}(\mathbf{x}_j, r_j, \mathbf{x}_k, r_k). \quad (5)$$

Note that, here, we normalize the discrepancy. This is done in order to make the cost function easier to interpret: it can be thought of as the average squared distance of test points from their corresponding separating plane.

## 4. Solving the LAP using the CE method

### 4.1. The CE method

The cross-entropy (CE) method is a stochastic optimization method that is able to solve many difficult optimization problems, including combinatorial optimization problems and continuous optimization problems with many local minima; see [36, 39]. The fundamental idea of the method is to describe the location of the global minimum of an  $m$ -dimensional cost function,  $S(\cdot)$ , in terms of a degener-

ate  $m$ -dimensional probability distribution. That is, a random variable with this distribution takes only a single value, namely  $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathbb{R}^m} S(\mathbf{x})$ . If there is more than one global minimum, each minimum will have a corresponding degenerate distribution. The CE method works by ‘learning’ one of these distributions. In the continuous setting, this is done as follows. A parametric density,  $f(\mathbf{x}; \boldsymbol{\theta})$ , is used to describe the possible locations of a global minimum. A sample of size  $M$  is drawn from this density and ordered by the corresponding values of  $S$ . The  $\lceil \rho M \rceil$  (where  $\lceil \cdot \rceil$  is the ceiling function) members of the sample with the lowest values of  $S$  are identified as the ‘elite’ sample, where  $\rho \in (0, 1)$ . The elite sample is then used to update the parameter vector,  $\boldsymbol{\theta}$ . The updating is done by choosing  $\boldsymbol{\theta}$  to minimize the cross-entropy distance (also called the Kullback-Leibler divergence) between  $f(\mathbf{x}; \boldsymbol{\theta})$  and the targeted degenerate probability distribution. The process is continued until a stopping condition is met, e.g., the probability distribution is nearly degenerate. Thus, the CE method uses information about good choices of arguments to find better choices. The CE method and its convergence properties are discussed extensively in [36–39].

Usually, the parametric density,  $f(\mathbf{x}; \boldsymbol{\theta})$ , is chosen to be a product of normal densities. That is, we choose

$$f(\mathbf{x}; \boldsymbol{\theta}) = \varphi(x_1; \mu_1, \sigma_1) \varphi(x_2; \mu_2, \sigma_2) \cdots \varphi(x_m; \mu_m, \sigma_m),$$

where  $\varphi(\cdot; \mu, \sigma)$  is the density of a normal distribution with mean  $\mu$  and standard deviation  $\sigma$  and  $\boldsymbol{\theta} = (\mu_1, \sigma_1, \dots, \mu_m, \sigma_m)$ . Thus, each component of the argument of  $S(\cdot)$  is drawn independently from a normal distribution. There are two main reasons for using normal distributions. First, as the standard deviation of a normal distribution goes to zero, the distribution converges to a degenerate distribution. The second reason is that the parameter vector,  $\boldsymbol{\theta}^*$ , which minimizes the cross-entropy distance, is simply given by the maximum-likelihood estimates obtained from the elite sample. In other words, we update the parameters by setting the means equal to the sample means of the elite sample and the standard deviations equal to the sample standard deviations of the elite sample.

Using this approach, we have the following general algorithm for estimating  $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathbb{R}^m} S(\mathbf{x})$ .

- 1) **Initialization.** Identify an initial parameter vector,  $\boldsymbol{\theta}^{(0)} = (\mu_1^{(0)}, \sigma_1^{(0)}, \dots, \mu_m^{(0)}, \sigma_m^{(0)})$  and set  $n = 0$ .
- 2) **Sampling.** Draw an independent sample  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(M)}$  from  $f(\mathbf{x}; \boldsymbol{\theta}^{(n)})$  and sort it so that  $S(\mathbf{X}^{(1)}) \leq \dots \leq S(\mathbf{X}^{(M)})$ . We denote the  $j$ th component of  $\mathbf{X}^{(i)}$  by  $X_j^{(i)}$ .
- 3) **Updating.** Calculate the sample means and standard deviations of the elite

sample. That is, calculate

$$\bar{X}_1 = \frac{1}{\lceil \rho M \rceil} \sum_{i=1}^{\lceil \rho M \rceil} X_1^{(i)}, \quad \dots, \quad \bar{X}_m = \frac{1}{\lceil \rho M \rceil} \sum_{i=1}^{\lceil \rho M \rceil} X_m^{(i)}$$

and

$$S_1 = \sqrt{\frac{1}{\lceil \rho M \rceil - 1} \sum_{i=1}^{\lceil \rho M \rceil} (X_1^{(i)} - \bar{X}_1)^2}, \quad \dots, \quad S_m = \sqrt{\frac{1}{\lceil \rho M \rceil - 1} \sum_{i=1}^{\lceil \rho M \rceil} (X_m^{(i)} - \bar{X}_m)^2}.$$

Set  $\mu_1^{(n+1)} = \bar{X}_1, \dots, \mu_m^{(n+1)} = \bar{X}_m$  and  $\sigma_1^{(n+1)} = S_1, \dots, \sigma_m^{(n+1)} = S_m$ .

Set  $\boldsymbol{\theta}^{(n+1)} = (\mu_1^{(n+1)}, \sigma_1^{(n+1)}, \dots, \mu_m^{(n+1)}, \sigma_m^{(n+1)})$ .

- 4) **Iteration.** If a predetermined stopping condition is met, terminate. Otherwise, set  $n = n + 1$  and repeat from step 2.

#### 4.1.1. Parameter choice and stopping conditions

The parameters of the CE algorithm help to improve the speed of convergence and quality of solutions. For example, a good choice of  $\boldsymbol{\theta}^{(0)}$  improves the convergence properties of the algorithm. The means,  $\mu_1^{(0)}, \dots, \mu_m^{(0)}$ , should be chosen as close as possible to the optimal solution. The initial standard deviations,  $\sigma_1^{(0)}, \dots, \sigma_m^{(0)}$ , should be chosen large enough that the algorithm is able to escape local minima but not so large that good configurations are quickly abandoned.

Ideally, the size of the sample in each step,  $M$ , should be quite large. However, there are often memory and performance constraints that limit this size. The parameter  $\rho$ , which controls the size of the elite sample, should be chosen large enough that a representative sample of good solutions are included in the elite sample. However, if  $\rho$  is chosen too large, the algorithm will take too long to converge. In general, the bigger  $M$  is, the smaller  $\rho$  should be.

A standard stopping condition for the algorithm is that it terminates when the cost function does not decrease significantly for a given number of steps.

#### 4.1.2. Variance injection and dynamic smoothing

When using the CE approach outlined above, the standard deviations of the normal densities may shrink too quickly. In this case, the CE algorithm can converge to a sub-optimal solution. In order to guard against this, we use variance injection. The basic idea is to occasionally increase the variances of the distributions so that the algorithm can easily escape local minima. Usually, variance is injected when the cost function does not decrease significantly enough over a given number of iterations. The magnitude of this increase can depend on the current value of the cost function. If variance injection is used a number of times without a significant decrease in the

cost function, then the algorithm is terminated.

An alternative to variance injection is to use smoothing when updating  $\boldsymbol{\theta}$ . The updating step for the means at the  $n$ th iteration is then of the form

$$\mu_1^{(n+1)} = \alpha \cdot \bar{X}_1 + (1 - \alpha) \cdot \mu_1^{(n)}, \quad \dots, \quad \mu_m^{(n+1)} = \bar{X}_m + (1 - \alpha) \cdot \mu_m^{(n)},$$

and the updating step for the standard deviations is given by

$$\sigma_1^{(n+1)} = \beta \cdot S_1 + (1 - \beta) \cdot \sigma_1^{(n)}, \quad \dots, \quad \sigma_m^{(n+1)} = \beta \cdot S_m + (1 - \beta) \cdot \sigma_m^{(n)},$$

where  $\alpha, \beta \in (0, 1]$  (typically  $[0.7, 1]$ ). It is also possible to carry out dynamic smoothing, with both  $\alpha$  and  $\beta$  taken to be functions of the number of steps. In our experience, variance injection is more effective than smoothing. For a detailed discussion of both variance injection and dynamic smoothing, see [37].

## 4.2. Solving the LAP

We solve the LAP by minimizing the approximate discrepancy described in Section 3.4. That is, we solve

$$\begin{aligned} \{(\mathbf{x}_i^*, r_i^*)\}_{i \in \mathcal{I}} &= \arg \min_{\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}} \tilde{\mathcal{D}}(\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}) \\ &= \arg \min_{\{(\mathbf{x}_i, r_i)\}_{i \in \mathcal{I}}} \sum_{j=1}^N \sum_{k=j+1}^N \tilde{\mathcal{D}}_{j,k}(\mathbf{x}_j, r_j, \mathbf{x}_k, r_k), \end{aligned} \quad (6)$$

where  $\tilde{\mathcal{D}}$  is given in (5). Putting this in the terminology of Section 4.1, we find the arguments that minimize the cost function  $\tilde{\mathcal{D}}$ , namely the  $N$  generating points, each of which is described by three coordinates and an associated radius. We associate a normal distribution with each of the  $m = 4N$  values that need to be determined. Thus, for  $i = 1, \dots, N$ , the coordinates and radius of the  $i$ th generating point,  $(\mathbf{x}_i, r_i) = (x_i, y_i, z_i, r_i)$ , are each described by a normal density. We denote the initial means and standard deviations by  $\mu_{x_i}^{(0)}, \mu_{y_i}^{(0)}, \mu_{z_i}^{(0)}, \mu_{r_i}^{(0)}$  and  $\sigma_{x_i}^{(0)}, \sigma_{y_i}^{(0)}, \sigma_{z_i}^{(0)}, \sigma_{r_i}^{(0)}$ . In order to ensure that the radii are positive, we truncate the corresponding normal densities to the positive real line (this is possible without changing the updating rules, see, e.g., [36, 37]).

### 4.2.1. Variance injection and stopping conditions

We apply variance injection when the cost function does not decrease significantly over a period of  $\tau > 0$  iterations. More precisely, at each iteration of the algorithm, we record  $\tilde{\mathcal{D}}_{\min}^{(n)}$ , the minimal value of the approximate discrepancies calculated from

the sample in that step. If, at the  $n$ th step,

$$\left| \frac{\tilde{\mathcal{D}}_{\min}^{(n)} - \max_{t \in \{n-\tau, \dots, n-1\}} \tilde{\mathcal{D}}_{\min}^{(t)}}{\tilde{\mathcal{D}}_{\min}^{(n)}} \right| < \delta_{\text{inject}},$$

we perform variance injection. Because many generating points may already be close to their optimal positions and sizes, variance injection is carried out locally with a magnitude controlled by a parameter  $\kappa > 0$ . This is done by calculating the local cost of each cell and increasing the variance of the associated densities accordingly. We calculate the average cost of the  $i$ th cell by

$$\bar{\mathcal{D}}_i = \frac{1}{\sum_{j=1}^N \#T_{i,j}} \sum_{j=1}^N \tilde{\mathcal{D}}_{i,j} \left( (\mu_{x_i}^{(n)}, \mu_{y_i}^{(n)}, \mu_{z_i}^{(n)}), \mu_{r_i}^{(n)}, (\mu_{x_j}^{(n)}, \mu_{y_j}^{(n)}, \mu_{z_j}^{(n)}), \mu_{r_j}^{(n)} \right).$$

The local cost of a cell,  $\mathcal{D}_i^*$ , is defined to be the maximum of its own average cost and the average cost of its adjacent cells. That is,

$$\mathcal{D}_i^* = \max \left\{ \bar{\mathcal{D}}_k : k = i \text{ or } k \text{ such that } N_{i,k}^* \neq \emptyset \right\}, \quad i = 1, \dots, N. \quad (7)$$

The variance injection is performed by setting

$$\begin{aligned} \sigma_{x_i}^{(n)} &= \sigma_{x_i}^{(n)} + \kappa \cdot \sqrt{\mathcal{D}_i^*}, & \sigma_{y_i}^{(n)} &= \sigma_{y_i}^{(n)} + \kappa \cdot \sqrt{\mathcal{D}_i^*}, \\ \sigma_{z_i}^{(n)} &= \sigma_{z_i}^{(n)} + \kappa \cdot \sqrt{\mathcal{D}_i^*}, & \sigma_{r_i}^{(n)} &= \sigma_{r_i}^{(n)} + \kappa \cdot \sqrt{\mathcal{D}_i^*}, \end{aligned}$$

for  $i = 1, \dots, N$ .

If, at any stage, the benefit of variance injection becomes negligible, we stop performing it. More precisely, if the current minimum cost,  $\mathcal{D}_{\min}^{(n)}$  divided by the minimum cost immediately prior to the last variance injection is larger than  $\gamma \in (0, 1)$ , we no longer carry out variance injection.

The algorithm is terminated when

$$\left| \frac{\tilde{\mathcal{D}}_{\min}^{(n)} - \max_{t \in \{n-\tau, \dots, n-1\}} \tilde{\mathcal{D}}_{\min}^{(t)}}{\tilde{\mathcal{D}}_{\min}^{(n)}} \right| < \delta_{\text{terminate}},$$

where  $\delta_{\text{terminate}} < \delta_{\text{inject}}$ .

#### 4.2.2. Initial configuration

Reasonably good initial tessellations can be obtained directly from the tomographic data [6]. The centroids and equivalent radii of the grains in the tomographic data are used for the coordinates and radii of the generating points. The centroids,

$\{(c_{x_i}, c_{y_i}, c_{z_i})\}_{i \in \mathcal{I}}$ , are given by

$$c_{x_i} = \sum_{(x,y,z) \in R_i} \frac{x}{\#R_i}, \quad c_{y_i} = \sum_{(x,y,z) \in R_i} \frac{y}{\#R_i}, \quad c_{z_i} = \sum_{(x,y,z) \in R_i} \frac{z}{\#R_i},$$

for  $i = 1, \dots, N$ . The equivalent radii are given by

$$\hat{r}_i = \sqrt[3]{\frac{3}{4\pi} \#R_i}$$

for  $i = 1, \dots, N$ . The initial means of the densities are then given by  $\mu_{x_i}^{(0)} = c_{x_i}$ ,  $\mu_{y_i}^{(0)} = c_{y_i}$ ,  $\mu_{z_i}^{(0)} = c_{z_i}$  and  $\mu_{r_i}^{(0)} = \hat{r}_i$  for  $i = 1, \dots, N$ .

The initial standard deviations are chosen proportional to the local cost of the cells in this approximation. The local costs are calculated as in Section 4.2.1. Thus,

$$\sigma_{x_i}^{(0)} = \sigma_{y_i}^{(0)} = \sigma_{z_i}^{(0)} = \sigma_{r_i}^{(0)} = \sqrt{\mathcal{D}_i^*},$$

for  $i = 1, \dots, N$ .

#### 4.2.3. Choice of control parameters for the CE method

Our investigations showed that a parameter choice of  $M = 4000$ ,  $\rho = 0.05$ ,  $\delta_{\text{inject}} = 0.05$ ,  $\delta_{\text{terminate}} = 0.01$ ,  $\tau = 10$ ,  $\gamma = 0.9$  and  $\kappa = 0.25$  works well for a large number of different data sets. The elite set then consists of  $\lceil \rho M \rceil = 200$  samples, which results in a sufficiently large sample of good solutions. Note that other parameter values may improve the speed of convergence (or, in the case of the parameters controlling variance injection and termination, improve the quality of the approximation). However, the basic performance of the CE algorithm is relatively robust to parameter choice. That is, the convergence behavior and quality of approximations are good for a wide range of parameters.

#### 4.2.4. Edge cells

The approximations obtained using the CE algorithm have unbounded cells at the edge of the observation window. This is because the corresponding grains in the tomographic data are not delimited by other grains. The unbounded cells are then intersected with the observation window,  $W$ . Note that, in most cases, edge grains are of little scientific interest and are not considered when analyzing the data. In practice, edge grains may be only partially observed or may not be representative (e.g., when studying the dynamics of grains undergoing grain coarsening). In this paper, we explicitly ignore grains and their approximating cells if they lie at the edge of the observation window.

## 5. Experimental results

In this section, we present the results of a number of numerical experiments which we carried out in order to demonstrate the effectiveness of the CE approach. We consider three distinct data sets. Two of these data sets are produced using stochastic models. These ‘artificial’ data sets allow us to investigate how well our method is able to reconstruct tessellations when we are certain the underlying tessellation is, indeed, Laguerre. The first artificial data set is produced using a stochastic model that describes a polycrystalline material undergoing grain coarsening. The second artificial data set is, by design, much more pathological. It exhibits large variation in the size and shape of its grains, as well as the number of neighbors each grain has. This data set allows us to explore the effectiveness of our approach in an extreme setting. In particular, it provides an example of a setting in which the standard choice of initial conditions is far from optimal. Finally, we consider an empirical data set: tomographic data obtained from a sample of Al-5 wt% Cu. We demonstrate that our method is able to produce an excellent approximation to this data set using a Laguerre tessellation.

### 5.1. *Artificial data*

The artificial data sets are produced using two distinct stochastic models. The polycrystalline model (PCM), developed in [8], describes a polycrystalline material undergoing grain coarsening. The second data set is produced using a randomly marked Poisson process; see [50]. The resulting tessellation is known as a Poisson-Laguerre tessellation (PLT); see [43]. In the first case, there is a strong correlation between the relative positions of the generating points and their weights. In the second example, the weights are independent of the positions of the generating points. As a result, the second data set is much less regular than the first data set. Figure 3 shows 2D cross sections of the data sets, together with their generating points.

#### 5.1.1. *Model descriptions*

In the PCM model, a random Laguerre tessellation is produced using a set of spheres that may overlap slightly. The centers of the spheres give the locations of the generating points and the radii define the weights. The spheres have hard-cores which may not overlap. This was shown to be a suitable model for polycrystalline materials in [8]. Both the density of the sphere packing and the radii of the spheres influence the sizes and shapes of the cells in the resulting tessellation. In particular, when highly dense packings are used, very “spherical” cells are generated with a narrow coordination number distribution. We generated a sample from this model in a bounded window,  $W = [0, 700]^3$  using parameters that were fitted to a sample



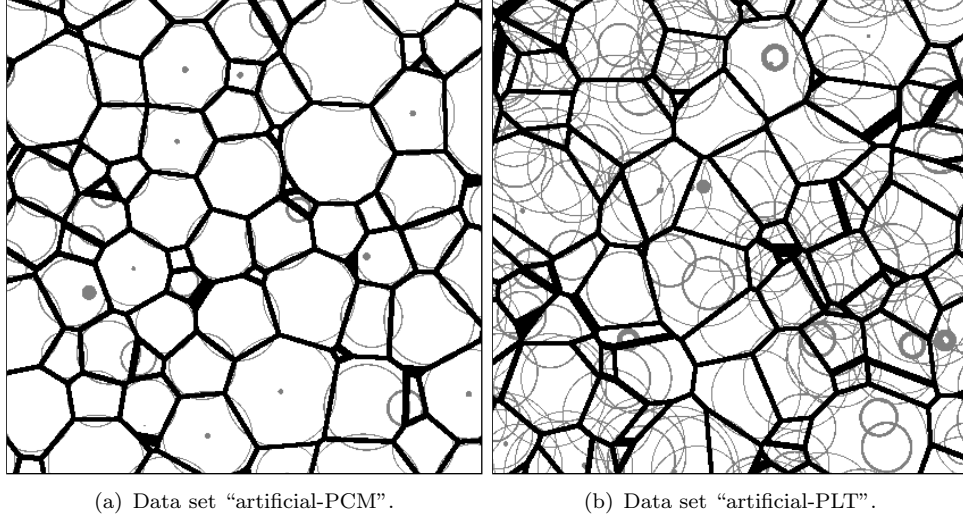


Figure 3. Partial cross-sections of artificial data sets. Spheres used as generators are drawn in gray (dots for sphere centers, gray circles for sphere surfaces). Radii are very large in the PLT case, therefore many surface voxels of spheres from other slices are visible.

of Al-5 wt% Cu annealed for 200 minutes. The fitting procedure is described in [8]. Note that the parameters we use were obtained by fitting the model to the empirical data described in Section 5.2. Having produced a sample, we removed cells that would correspond to a very small number of voxels after discretization (i.e., cells having a smaller volume than a ball with radius of 4 voxels). The final result is a tessellation consisting of roughly 2500 grains (approximately the same number as in the experimental data).

The PLT model is generated by simulating a homogeneous Poisson process with some intensity  $\lambda$  in a bounded window  $W$ . The points are independently marked by gamma-distributed random variables with some shape parameter  $\alpha$  and rate parameter  $\beta$ . Note that, unlike the PCM model, it is possible that some generating points will not produce cells in the PLT model. We generated a sample from this model in  $W = [0, 700]^3$  with intensity  $\lambda = 7.5 \times 10^{-6}$  (so that the expected number of points is roughly 2500), where we used parameters  $\alpha = 50$  and  $\beta = 1.125$  for the mark distribution.

### 5.1.2. Laguerre approximation

The realizations of the PCM and PLT models were both discretized on a voxel grid, resulting in two  $700 \times 700 \times 700$  images. These images are of the form described in Section 3.1. The CE method was then used to solve the LAP. The results are illustrated in Figure 4, where 2D cross-sections of the original tessellations are shown with the approximating tessellations superimposed.

Using multi-threading on a standard quad-core processor (Intel Core i5-3570K), the computing time was roughly 3 hours for the PCM data and 4 hours for the PLT

data. For smaller test sets, with 500 grains each, the time required was roughly 20 to 30 minutes. The memory requirements when fitting the full data sets were minimal (especially in comparison to approaches where it is necessary to store the complete voxelized data). Basically, it is necessary to store the test points and all the generators from one iteration of the CE method. The cost of storing additional variables, such as the means and standard deviations of the densities describing the generators and the cost values of the test points are negligible. For the PCM data, the number of test points used was approximately 140 000. Approximately 160 000 points were used for the PLT data. Together with  $N \cdot M \approx 2500 \cdot 4000$  generators in one iteration, this sums up to less than 200 MB of RAM when using 32-bit floating point coordinates.

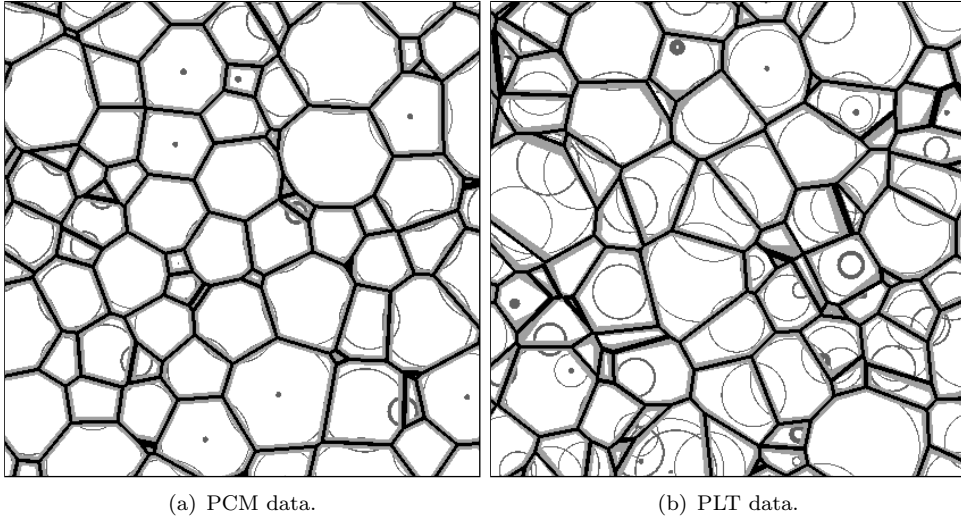


Figure 4. Partial cross-sections of artificial data: original cell boundaries (gray) superimposed with cell boundaries of our Laguerre approximation (black), with centers and radii of detected generators drawn in small gray dots and thin gray circles.

Figure 5 illustrates the convergence behavior of the CE algorithm. When approximating the PCM data, roughly 650 iterations of the algorithm were required. In the PLT case, about 900 iterations were required. It is not surprising that the CE algorithm took longer to terminate in the PLT case. As will be seen in Section 5.1.3 the initial conditions in this setting are much further from the optimal solution.

### 5.1.3. Results

In order to evaluate the quality of our approximations, we compare them to approximations obtained using the heuristic approach presented in [6] and the orthogonal regression method proposed in [48]. Note that, although the orthogonal regression approach is able to achieve quite good approximations of the cells, it does not result in a parametrized tessellation. Table 1 shows an evaluation of the approximations with respect to a volume-based discrepancy measure: the number of voxels that are

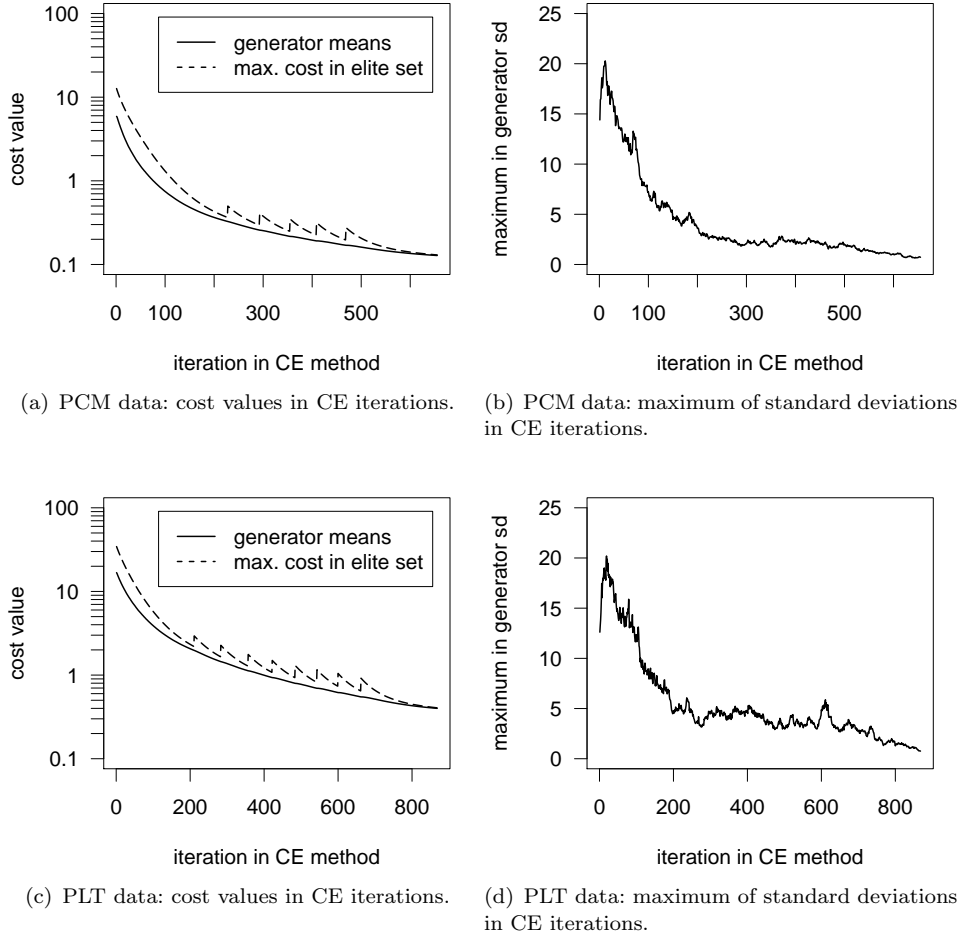


Figure 5. CE method for artificial data: convergence of the cost values  $\tilde{\mathcal{D}}$  (Eq. (5)) and standard deviations.

correctly labeled. Our method results in an almost perfect approximation of the PCM data. The heuristic approach considered in [6] also works very well. However, it is not able to reproduce the neighbor structure as successfully as our approach, as evidenced by rows 2 through 5 of Table 1. Our method is slightly less successful at approximating the PLT data but, in this case, considerably outperforms the heuristics of [6].

Table 1. Evaluation of the artificial data approximations: H denotes the heuristic approach of [6]; CE denotes the CE method considered in the present paper; OR denotes orthogonal regression proposed in [48].

	H	PCM CE	OR	H	PLT CE	OR
correctly labeled voxels [%]	97.3	99.1	99.8	87.8	96.2	99.8
cells with all neighbors correct [%]	75.6	88.5	78.9	18.6	45.1	61.3
cells with $\leq 1$ incorrect neighbors [%]	96.3	99.0	97.3	46.7	77.3	90.4
cells with $\leq 2$ incorrect neighbors [%]	99.6	99.8	99.7	72.2	92.1	98.2
mean number of erroneous neighbors/grain	0.28	0.13	0.24	1.83	0.90	0.51
average displacement of centroids [voxels]	0.56	0.24	0.09	2.56	1.33	0.06

The parameters used by the heuristic approach of [6] are very close to the initial conditions of the CE algorithm. Thus, from the difference in performance, it seems that the initial conditions are quite far from the optimal parameters. In order to find better configurations, the CE algorithm has to be able to escape local minima around these initial conditions. This is made clear by the fact that increasing the number of times variance injection is used (as well as the number of iterations), we are able to further improve the results for the CE algorithm. Namely, by manually increasing the number of variance injections to 20 (instead of the 8 used in the standard approach), we obtained an approximation which correctly labeled 97.3% of the voxels (instead of 96.2%). This significant improvement indicates that local minima are present in which the CE algorithm is becoming trapped (because variance injection works by helping the algorithm escape local minima). This, in turn, implies that there are local minima near to the initial conditions of the CE algorithm that are sufficiently deep to trap it. A direct implication of these results is that algorithms that converge to nearby local minima (such as gradient-descent) may not lead to good approximations when using a standard choice of initial conditions. Furthermore, it is not clear that it is always straightforward to find good initial conditions. For example, the initial generating points are usually chosen to lie inside their cells. But, using the methods presented in [31] and the exact description of the PLT tessellation, it was not possible to find a solution to the LIP subject to the restriction that generating points were contained in their cells. If a method such as gradient-descent starts with all of the points lying inside their cells, it is unclear that it will be able to obtain a solution where some generating points lie far outside their cells.

Figure 6 shows scatter plots of the original cell sizes vs. the cell sizes in the CE approximations. The quality of the CE approximation of the PCM data is immediately apparent. It is also clear that the difficulties in fitting the PLT data lie primarily in the small cells.

## 5.2. *Experimental data*

The experimental data we consider was used in [8] to develop the PCM model discussed in Section 5.1. An Al-5 wt% Cu sample (cylindrical with 8.5 mm length and 4 mm diameter) was heated to the semisolid state at a temperature of 592 °C, at which point coarsening processes were measured *in situ* using synchrotron X-ray tomographic imaging. For the present paper, we consider only data obtained after an annealing time of 200 minutes (which corresponds to the first annealing step in [8]). It consists of approximately 2500 grains, which is quite a large data set. A 2D cross-section of the data is shown in Figure 7. The experimental data set was segmented using the watershed transformation, resulting in a labeled image of grain regions, which are separated by a watershed layer of one-voxel thickness. Further

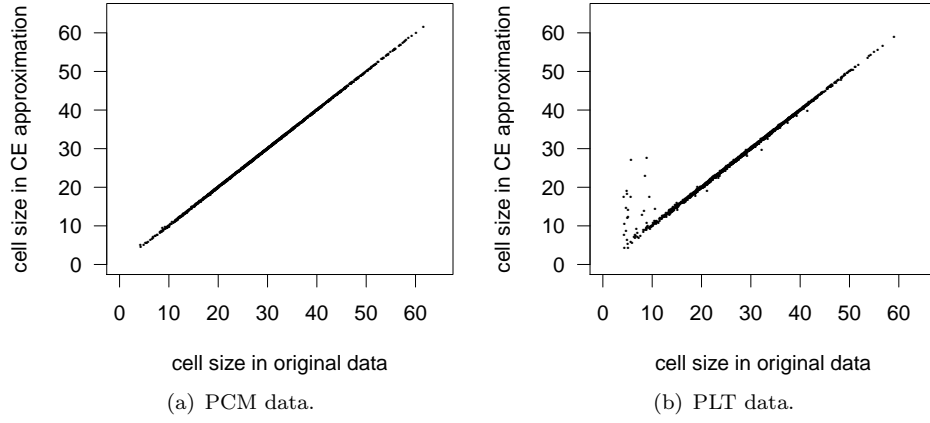


Figure 6. Scatter plot of cell sizes given as radii of volume-equivalent spheres obtained from the original data and cells extracted by the CE method.

details regarding sample preparation, imaging and segmentation can be found in [8].

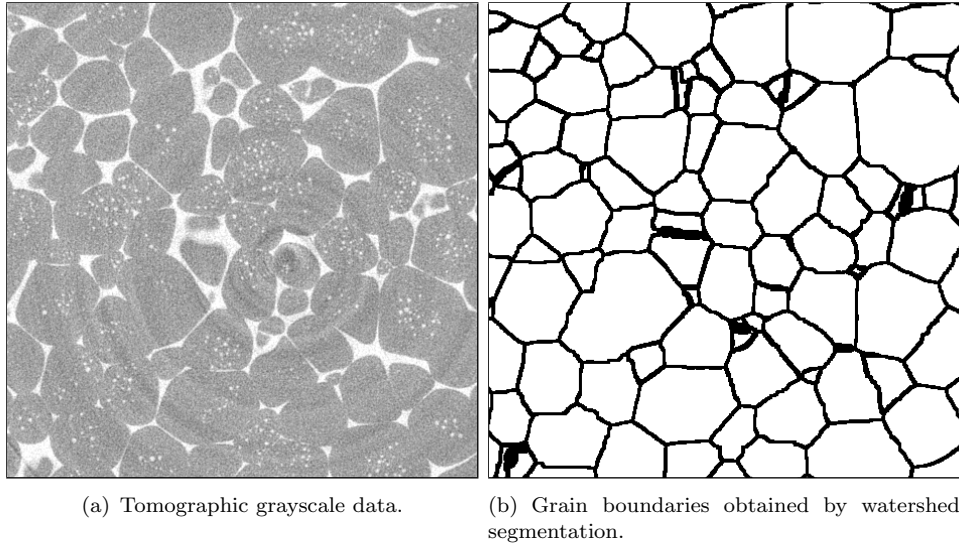


Figure 7. Partial cross-section of experimental data.

### 5.2.1. Laguerre approximation

The CE method was applied to solve the LAP for the experimental data set. A cross-section of the resulting Laguerre approximation is given in Figure 8.

Fitting the approximation took approximately 70 minutes on the same Intel Core i5-3570K quad-core processor used to fit the artificial data. The number of test points extracted from the image data was roughly 150 000, which is almost the same as for the artificial data sets. The convergence behavior of the CE method

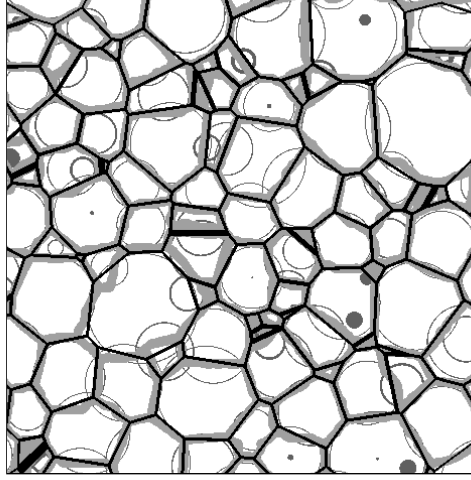


Figure 8. Partial cross-section of experimental data: original boundaries of grain regions (gray) superimposed with cell boundaries of our Laguerre approximation (black), with centers and radii of detected generators drawn in small gray dots and thin gray circles.

is illustrated in Figure 9. Using the same parameters, only 283 iterations were necessary, which is substantially fewer than the number required for the artificial data. This is mainly due to the fact that only one variance injection was required.

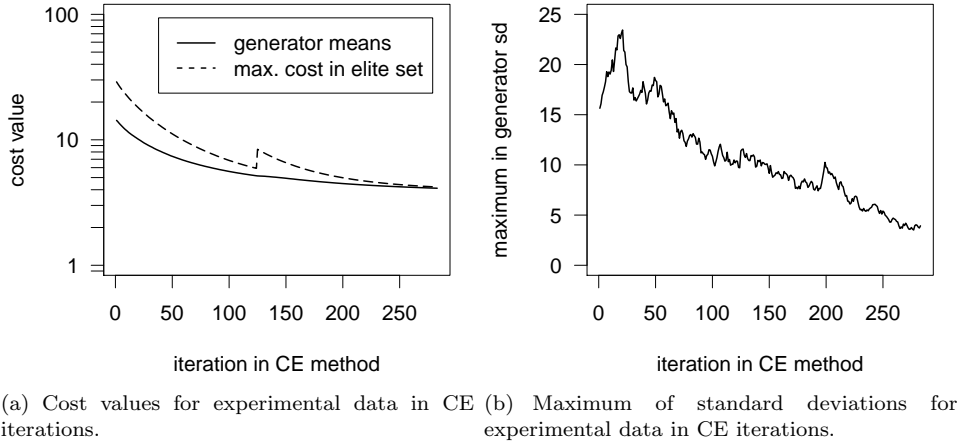


Figure 9. CE method for experimental data: convergence of the cost values  $\tilde{\mathcal{D}}$  (Eq. (5)) and standard deviations.

### 5.2.2. Results

As with the artificial data, we compare our results with the heuristic approach proposed in [6] and the orthogonal regression approach introduced in [48]. As mentioned above, the orthogonal regression approach reconstructs the individual cells quite well but does not result in a tessellation. The results are summarized in Table 2. The CE method correctly assigns 91.4 % of the voxels. In contrast, the heuristic approach

correctly assigns only 82.9 % of the voxels. The orthogonal regression yields 96.3 %.

Table 2. Evaluation of the experimental data approximations: H denotes the heuristic approach of [6]; CE denotes the CE method considered in the present paper; OR denotes orthogonal regression proposed in [48].

	H	CE	OR
correctly labeled voxels [%]	82.9	91.4	96.3
cells with all neighbors correct [%]	20.1	38.4	29.3
cells with $\leq 1$ incorrect neighbors [%]	48.4	73.1	50.2
cells with $\leq 2$ incorrect neighbors [%]	65.6	90.6	62.6
mean number of erroneous neighbors/grain	2.29	1.04	2.37
average displacement of centroids [voxels]	3.40	1.78	0.23

Note that the CE method significantly outperforms both the heuristic method and orthogonal regression when describing the neighborhood structure of the grains, although it does not seem possible for a tessellation containing only convex cells to accurately capture the full neighborhood structure. This is at least partially due to segmentation issues and contacts in the image data with very small areas. In contrast, the orthogonal regression approach performs surprisingly badly. It seems that the geometric properties of normal Laguerre tessellations (e.g., coinciding faces, edges and vertices) favor realistic reconstructions. We think this confirms that the Laguerre tessellation is a good choice for representing polycrystalline microstructures. It is possible, of course, that the quality of the fit could be improved using non-convex cells. For example, in [34], the linear programming method was used to obtain a generalized power diagram approximation of similar data (but with far fewer grains and voxels) resulting in a fit that correctly assigned 93.8 % of the voxels with non-convex cells.

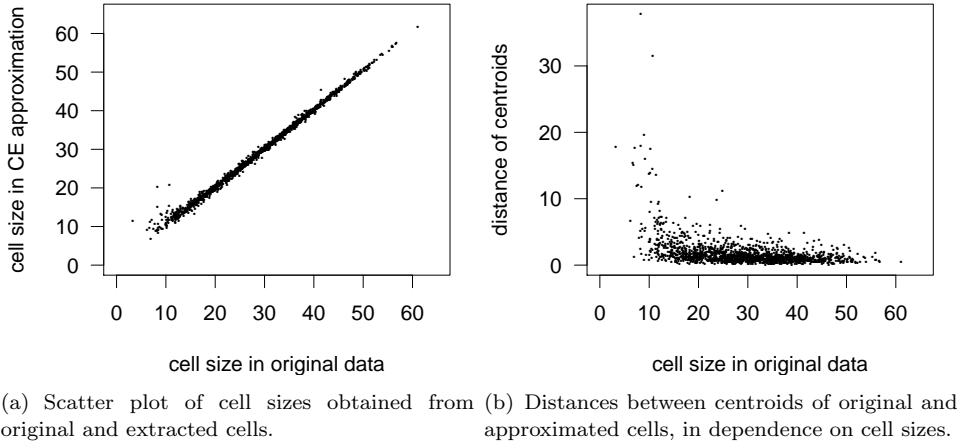


Figure 10. Experimental data: comparison of original data and cells extracted by the CE method. Cell sizes are given as the radii of their volume-equivalent spheres.

Figure 10(a) shows a scatter plot of the volumes of the original grains against the volumes of the Laguerre cells (with the volumes expressed as radii of volume-

equivalent spheres). The overall fit of the cell volumes is excellent. Figure 10(b) shows that the locations of the grains are also quite accurate. Note that the main issue with fitting seems to be small cells. This is not such a problem, however, as small cells (with equivalent radii of up to 10 voxels) are subject to image segmentation error and are, thus, not too reliable in the original data.

## 6. Conclusions and outlook

In this paper, we considered the problem of approximating tomographic data by a Laguerre tessellation. We expressed this problem as an optimization problem: the generating points of the approximating tessellation need to be chosen in order to minimize the discrepancy between the tessellation and the tomographic data. We considered an interface-based discrepancy measure, instead of the volume-based discrepancies more commonly considered in the literature. This allowed us to use the CE method, a stochastic optimization method that is able to escape local minima, to fit the approximating tessellation. We then carried out numerical experiments on both artificially generated and experimentally obtained tomographic data that demonstrated the broad effectiveness of our approach.

Our method is robust and easy to implement. Thus it can be applied to fit Laguerre tessellations to almost any material with a granular or cellular structure. An obvious next step is to extend our approach to tessellations that include non-convex cells. In particular, we believe the approach and philosophy outlined in this paper can be extended to fit generalized power diagrams (see [34, 51]) — generalizations of Laguerre tessellations that can include non-convex cells. This will be the subject of a forthcoming research paper.

## Funding

This work was partially supported by the Australian Research Council under Grant DP140101956 and by the Deutsche Forschungsgemeinschaft under Grant KR 1658/4-1.

## Supplemental material

A software package including Java code and data sets can be downloaded from <https://github.com/stochastics-ulm-university/laguerre-approximation>.

## References

- [1] K. M. Döbrich, C. Rau, and C. E. Krill III. Quantitative characterization of the three-



- dimensional microstructure of polycrystalline Al-Sn using X-ray microtomography. *Metallurgical and Materials Transactions A*, 35A:1953–1961, 2004.
- [2] D. J. Rowenhorst, J. P. Kuang, K. Thornton, and P. W. Voorhees. Three-dimensional analysis of particle coarsening in high volume fraction solid-liquid mixtures. *Acta Materialia*, 54:2027–2039, 2006.
  - [3] N. Limodin, L. Salvo, M. Suéry, and M. DiMichiel. In situ investigation by X-ray tomography of the overall and local microstructural changes occurring during partial remelting of an Al–15.8 wt.% Cu alloy. *Acta Materialia*, 55:3177–3191, 2007.
  - [4] M. Groeber, S. Ghosh, M. D. Uchic, and D. M. Dimiduk. A framework for automated analysis and simulation of 3D polycrystalline microstructures. Part 1: Statistical characterization. *Acta Materialia*, 56(6):1257–1273, 2008.
  - [5] W. Ludwig, A. King, P. Reischig, M. Herbig, E. M. Lauridsen, S. Schmidt, H. Proudhon, S. Forest, P. Cloetens, S. Rolland du Roscoat, J. Y. Buffière, T. J. Marrow, and H. F. Poulsen. New opportunities for 3D materials science of polycrystalline materials at the micrometre lengthscale by combined use of X-ray diffraction and X-ray imaging. *Materials Science and Engineering A*, 524(1–2):69–76, 2009.
  - [6] A. Lyckegaard, E. M. Lauridsen, W. Ludwig, R. W. Fonda, and H. F. Poulsen. On the use of Laguerre tessellations for representations of 3D grain structures. *Advanced Engineering Materials*, 13(3):165–170, 2011.
  - [7] T. Werz, M. Baumann, U. Wolfram, and C. E. Krill III. Particle tracking during Ostwald ripening using time-resolved laboratory X-ray microtomography. *Materials Characterization*, 90:185–195, 2014.
  - [8] A. Spetl, R. Wimmer, T. Werz, M. Heinze, S. Odenbach, C. E. Krill III, and V. Schmidt. Stochastic 3D modeling of Ostwald ripening at ultra-high volume fractions of the coarsening phase. *Modelling and Simulation in Materials Science and Engineering*, 23(6):065001, 2015.
  - [9] Z. Fan, Y. Wu, X. Zhao, and Y. Lu. Simulation of polycrystalline structure with Voronoi diagram in Laguerre geometry based on random closed packing of spheres. *Computational Materials Science*, 29:301–308, 2004.
  - [10] M. Beil, S. Eckel, F. Fleischer, H. Schmidt, V. Schmidt, and P. Walther. Fitting of random tessellation models to cytoskeleton network data. *Journal of Theoretical Biology*, 241:62–72, 2006.
  - [11] M. Groeber, S. Ghosh, M. D. Uchic, and D. M. Dimiduk. A framework for automated analysis and simulation of 3D polycrystalline microstructures. Part 2: Synthetic structure generation. *Acta Materialia*, 56(6):1274–1287, 2008.
  - [12] C. Lautensack. Fitting three-dimensional Laguerre tessellations to foam structures. *Journal of Applied Statistics*, 35(9):985–995, 2008.
  - [13] C. Lautensack, H. Ewe, P. Klein, and T. Sych. 3D characterization, modeling and effective thermal conductivity of open aluminium foams. In J. Hirsch, B. Skrotski, and G. Gottstein, editors, *Aluminium Alloys – Their Physical and Mechanical Properties*, pages 1368–1374, Weinheim, 2008. Wiley-VCH.
  - [14] V. Schmidt, G. Gaiselmann, and O. Stenzel. Stochastic 3D models for the microstructure of advanced functional materials. In V. Schmidt, editor, *Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms*, pages 95–141. Springer,

- Cham, 2015.
- [15] C. Redenbach and A. Liebscher. Random tessellations and their application to the modelling of cellular materials. In V. Schmidt, editor, *Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms*, pages 73–93. Springer, Cham, 2015.
  - [16] D. Westhoff, J. J. van Franeker, T. Brereton, D. P. Kroese, R. A. J. Janssen, and V. Schmidt. Stochastic modeling and predictive simulations for the microstructure of organic semiconductor films processed with different spin coating velocities. *Modelling and Simulation in Materials Science and Engineering*, 23(4):045003, 2015.
  - [17] S.N. Chiu, D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. J. Wiley & Sons, Chichester, 3rd edition, 2013.
  - [18] A. Okabe, B. Boots, K. Sugihara, and S.N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. J. Wiley & Sons, Chichester, 2nd edition, 2000.
  - [19] H. Telley, T.M. Liebling, A. Mocellin, and F. Righetti. Simulating and modelling grain growth as the motion of a weighted Voronoi diagram. *Materials Science Forum*, 94–96:301–306, 1992.
  - [20] H. Telley, T.M. Liebling, and A. Mocellin. The Laguerre model of grain growth in two dimensions: I. Cellular structures viewed as dynamical Laguerre tessellations. *Philosophical Magazine Part B*, 73(3):395–408, 1996.
  - [21] H. Telley, T.M. Liebling, and A. Mocellin. The Laguerre model of grain growth in two dimensions: II. Examples of coarsening simulations. *Philosophical Magazine Part B*, 73(3):409–427, 1996.
  - [22] X. Xue, F. Righetti, H. Telley, and T.M. Liebling. The Laguerre model for grain growth in three dimensions. *Philosophical Magazine Part B*, 75(4):567–585, 1997.
  - [23] C. Lautensack and T. Sych. 3D image analysis of open foams using random tessellations. *Image Analysis & Stereology*, 25:87–93, 2006.
  - [24] A. Liebscher. *Stochastic Modelling of Foams*. PhD thesis, TU Kaiserslautern, 2014.
  - [25] P.F. Ash and E.D. Bolker. Recognizing Dirichlet tessellations. *Geometriae Dedicata*, 19(2):175–206, 1985.
  - [26] D.G. Evans and S.M. Jones. Detecting Voronoi (area-of-influence) polygons. *Mathematical Geology*, 19(6):523–537, 1987.
  - [27] F. Aurenhammer. Recognising polytopical cell complexes and constructing projection polyhedra. *Journal of Symbolic Computation*, 3(3):249–255, 1987.
  - [28] D. Hartvigsen. Recognizing Voronoi diagrams with linear programming. *ORSA Journal on Computing*, 4(4):369–374, 1992.
  - [29] F.P. Schoenberg, T. Ferguson, and C. Li. Inverting Dirichlet tessellations. *The Computer Journal*, 46(1):76–83, 2003.
  - [30] A. Adamatzky. Massively parallel algorithm for inverting Voronoi diagram. *Neural Network World*, 5:385–92, 1993.
  - [31] Q. Duan, D.P. Kroese, T. Brereton, A. Spettl, and V. Schmidt. Inverting Laguerre tessellations. *The Computer Journal*, 57:1431–1440, 2014.
  - [32] A. Suzuki and M. Iri. Approximation of a tessellation of the plane by a Voronoi diagram. *Journal of the Operations Research Society of Japan*, 29(1):69–96, 1986.
  - [33] A. Liebscher, D. Jeulin, and C. Lantuéjoul. Stereological reconstruction of polycrys-

- talline materials. *Journal of Microscopy*, 258(3):190–199, 2015.
- [34] A. Alpers, A. Brieden, P. Gritzmann, A. Lyckegaard, and H. Poulsen. Generalized balanced power diagrams for 3D representations of polycrystals. *Philosophical Magazine*, 95(9):1016–1028, 2015.
  - [35] A. Liebscher. Laguerre approximation of random foams. *Philosophical Magazine*, 95(25):2777–2792, 2015.
  - [36] R. Y. Rubinstein and D. P. Kroese. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. Springer, New York, 2004.
  - [37] D. P. Kroese, S. Porotsky, and R. Y. Rubinstein. The cross-entropy method for continuous multi-extremal optimization. *Methodology and Computing in Applied Probability*, 8(3):383–407, 2006.
  - [38] D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo Methods*. J. Wiley & Sons, New Jersey, 2011.
  - [39] Z. I. Botev, D. P. Kroese, R. Y. Rubinstein, and P. L’Ecuyer. The cross-entropy method for optimization. In V. Govindaraju and C. R. Rao, editors, *Machine Learning: Theory and Applications*, pages 35–59. North-Holland, Oxford, 2013.
  - [40] F. Aurenhammer. Power diagrams: properties, algorithms and applications. *SIAM Journal on Computing*, 16(1):78–96, 1987.
  - [41] K. Sugihara. Three-dimensional convex hull as a fruitful source of diagrams. *Theoretical Computer Science*, 235:325–337, 2000.
  - [42] H. Imai, M. Iri, and K. Murota. Voronoi diagram in the Laguerre geometry and its applications. *SIAM Journal on Computing*, 14:93–105, 1985.
  - [43] C. Lautensack. *Random Laguerre Tessellations*. Verlag Lautensack, Weiler bei Bingen, 2007.
  - [44] F. Aurenhammer. A criterion for the affine equivalence of cell complexes in  $\mathbb{R}^d$  and convex polyhedra in  $\mathbb{R}^{d+1}$ . *Discrete and Computational Geometry*, 2:49–64, 1987.
  - [45] J. B. T. M. Roerdink and A. Meijster. The watershed transform: definitions, algorithms, and parallelization strategies. *Fundamenta Informaticae*, 41:187–228, 2001.
  - [46] O. Brunke, S. Odenbach, and F. Beckmann. Quantitative methods for the analysis of synchrotron- $\mu$ CT datasets of metallic foams. *The European Physical Journal Applied Physics*, 29(1):73–81, 2005.
  - [47] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.
  - [48] A. Spettl, T. Werz, C. E. Krill III, and V. Schmidt. Parametric representation of 3D grain ensembles in polycrystalline microstructures. *Journal of Statistical Physics*, 154:913–928, 2014.
  - [49] P. P. N. de Groen. An introduction to total least squares. *Nieuw Archief voor Wiskunde, 4th Series*, 14:237–253, 1996.
  - [50] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns*. J. Wiley & Sons, Chichester, 2008.
  - [51] H. Altendorf, F. Latourte, D. Jeulin, M. Faessel, and L. Saintoyant. 3D reconstruction of a multiscale microstructure by anisotropic tessellation models. *Image Analysis and Stereology*, 33(2):121–130, 2014.