Accurate approximation of the distributions of the 3D Poisson-Voronoi typical cell geometrical features

Martina Vittorietti, Geurt Jongbloed, Piet J.J. Kok, Jilt Sietsma

1Department of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
2Materials Innovation Institute (M2i), Mekelweg 4, 2628 CD Delft, The Netherlands
3Tata Steel, IJmuiden Technology Centre, Postbus 10.000, 1970 CA, IJmuiden, The Netherlands
4Department of Materials Science and Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

ABSTRACT. Although Poisson-Voronoi diagrams have interesting mathematical properties, there is still much to discover about the geometrical properties of its grains. Through simulations, many authors were able to obtain numerical approximations of the moments of the distributions of more or less all geometrical characteristics of the grain. Furthermore, many proposals on how to get close parametric approximations to the real distributions were put forward by several authors. In this paper we show that exploiting the scaling property of the underlying Poisson process, we are able to derive the distribution of the main geometrical features of the grain for every value of the intensity parameter. Moreover, we use a sophisticated simulation program to construct a close Monte Carlo based approximation for the distributions of interest. Using this, we also determine the closest approximating distributions within the mentioned frequently used parametric classes of distributions and conclude that these approximations can be quite accurate.

1. Introduction

In the past few years the use of Voronoi diagrams has rapidly increased. These diagrams represent an appealing structure, especially because they describe various natural processes quite well. In [1] an extensive list of fields in which Voronoi diagrams are adopted can be found. Among the many areas of applications of this model, the field of materials science stands out. In fact, they are now among the most used mathematical models for microstructure characterization and depending on the specific kind of materials, it is possible to use a proper category of Voronoi diagrams.

In this paper, we discuss the most basic instance of the model: Poisson-Voronoi diagrams. In this framework the nuclei or sites are generated by a homogeneous Poisson process with intensity parameter $\lambda$.

Although many interesting mathematical properties of Poisson-Voronoi diagrams are known, there is still much to discover about the distributions of the geometrical characteristics of its grains. Through simulations, many authors were able to obtain numerical approximations of the moments of the distribution of the volume, of the surface area, of the number of faces and many other geometrical characteristics of the grains. Nevertheless, analytic expressions of the distributions of many of these important features are not known, others are only known via complicated numerically intractable characterizations. Therefore, various proposals to obtain close approximations to the real distributions were put forward by several authors e.g.
Lognormal-, Generalized Gamma- and Rayleigh distributions. But as far as we know, there is no theoretical support for preferring one of these distributions.

The aim of this paper is twofold. After explaining that $\lambda$, the intensity parameter of the Poisson process, is the only parameter determining all distributional properties of the geometrical structure of the grain, we show that if we have the distribution of a given geometrical characteristic for $\lambda = 1$, the distribution of the same quantity for every value of $\lambda > 0$ can be obtained by rescaling. More precisely, we consider volume, surface area and number of faces of the grain, but the approach can be extended to other characteristics. Secondly, we find a close Monte Carlo based approximation for the previously mentioned geometrical characteristics of the grains and using it we determine the most closely approximating distribution within the mentioned frequently used parametric classes of distributions. As said before, several well known probability distributions were used for approximating the grain geometrical characteristics distributions, but in this study we determine the ‘best’ of these.

After briefly reviewing the basic concepts of Voronoi diagrams and the Poisson process in Section 2, in Section 3 we explain the scaling property of the Voronoi structure in terms of the intensity parameter and how it can be useful for studying distributional properties of the grain features. Since the intensity parameter $\lambda$ is the only parameter involved in generating a specific structure, it governs the distribution of all the geometrical characteristics of the Poisson-Voronoi typical cell. Later, we explain how the scaling acts on the different geometrical features and we show an empirical example of what happens changing the scale parameter. Section 4 describes our simulation approach and produces an accurate Monte Carlo approximation for the distribution of the grain volume and the grain surface area. In fact, we provide the approximate distributions of the volume and of the grain surface area for $\lambda = 1$ and we can adapt it for the other values of $\lambda$ using the aforementioned scaling properties. In Section 5, we study how well the true distributions of the geometrical characteristics can be approximated by some well-known and frequently used probability distributions in this context: the Gamma-, Generalized Gamma- and Lognormal distribution. Fitting these three distributions and comparing them through statistical measures such as the supremum distance between the Monte Carlo empirical distribution and its parametric approximations and Total Variation distance, we are not only able to identify the best approximation but also to give a measure of error if one of these parametric approximations is used. Finally, we discuss the possibility to extend our approach according to different Voronoi Diagrams cases, such as Multi-level Voronoi and/or Laguerre Voronoi Diagrams. We want to remark that for the 3D Voronoi diagrams generation we use TATA Steel software and for data analysis R.

2. Poisson-Voronoi Diagrams

We begin by reviewing the generic definition and the basic properties of the Poisson-Voronoi Diagram. Given a denumerable set of distinct points in $\mathbb{R}^d$, $X = \{x_i : i \geq 1\}$, the Voronoi diagram of $\mathbb{R}^d$ with nuclei $\{x_i\}$ (also called sites or generator points) is a partition of $\mathbb{R}^d$ consisting of cells

$$C_i = \{y \in \mathbb{R}^d : \|x_i - y\| \leq \|x_j - y\| \text{ for } j \neq i\}, \quad i = 1, 2, \ldots$$

where $\| \cdot \|$ is the usual Euclidean distance. This means that given a set of two or more but finitely many distinct points, we associate all locations in that space with the closest member(s) of the point set with respect to the Euclidean distance.

If we assume that $X = \Phi = \{x_i\}$ is the realization of a homogeneous Poisson point process, then we will refer to the resulting structure as the Poisson-Voronoi diagram, $\mathcal{V}_\Phi$. 
We find it useful to remind briefly what a Poisson process is and which are its basic properties. We follow Kingman’s approach in the mathematical definition of the process [cf. 2].

Let \( S \) be a measurable set in \( \mathbb{R}^d \), \( \mathcal{B}_S = \mathcal{B}(S) \) the \( \sigma \)-field of its Borel sets in \( S \) and \( \mu \) a boundedly finite Borel measure on \( \mathcal{B}_S \). Moreover, denote \( N(A) = \# \{ i : X_i \in A \} \). A Poisson process on \( S \) is then a random countable subset \( \Phi \) of \( S \), such that

- for every finite family of disjoint bounded Borel subsets \( A_1, A_2, \ldots, A_n \) of \( S \), the random variables \( N(A_1), N(A_2), \ldots, N(A_n) \) are independent
- \( N(A) \) has Poisson distribution \( \mathcal{P}(\lambda) \), where \( \lambda = \mu(A) \) lies in \( 0 \leq \lambda \leq \infty \).

From this it immediately follows that

\[
\mu(A) = \mathbb{E}\{N(A)\}.
\]

Therefore the measure \( \mu \) on \( S \) is often called the mean measure of the Poisson process \( \Phi \). When \( S = \mathbb{R}^d \), the mean measure is in most interesting cases given in terms of intensity. This is a positive measurable function \( \lambda \) on \( S \), in terms of which \( \mu \) is given by integrating \( \lambda \) with respect to \( d \)-dimensional Lebesgue measure:

\[
\mu(A) = \int_A \lambda(x) dx.
\]

If \( \lambda \) is continuous at \( x \), then eq. 2.1 implies that for small neighbourhoods \( A \) of \( x \),

\[
\mu(A) \approx \lambda(x)|A|,
\]

where \( |A| \) denotes the Lebesgue measure (length if \( d = 1 \), area if \( d = 2 \), volume if \( d = 3 \)) of \( A \). Thus \( \lambda(x)|A| \) is the approximate probability of a point of \( \Phi \) falling in the small set \( A \), and is larger in regions where \( \lambda \) is large than in those where \( \lambda \) is small. In the special case when \( \lambda \) is a constant, so that

\[
\mu(A) = \lambda|A|
\]

we speak of a uniform or homogeneous Poisson process.

In this paper, we assume that the sites of the Poisson-Voronoi diagrams are generated according to the particular case described by eq. 2.2.

As mentioned before, our aim is to find the distribution of the geometrical characteristics of the grains. In order to approximate these distributions, we generate a large sample of independent and identically distributed cells, more specifically typical cells. A typical Voronoi cell refers to a random polytope which loosely speaking has the same distribution as a randomly chosen cell from the diagram selected in such a way that every cell has the same chance of being sampled. Moreover, the distribution of the typical Poisson-Voronoi cell is by Slivnyak-Mecke formula [6] the same as the Voronoi cell obtained when the origin is added to the point process \( \Phi \). This corresponds formally to

\[
\mathcal{C} = \{ y \in \mathbb{R}^d : \| y \| \leq \| y - x \| \text{ for all } x \in \Phi \}
\]

Okabe et al. [1] synthesize previous research activity about the properties of Poisson Diagrams. Despite the fact that distributions of several geometrical characteristics are already known, the distributions of the main features, especially in 3D, are not. We describe a simulation approach to approximate these distributions in the next section.
3. DISTRIBUTION OF THE GEOMETRICAL PROPERTIES OF A TYPICAL CELL

Given the complexity in finding explicit formulae for the distributions of the Poisson-Voronoi tessellation geometrical characteristics, especially in 3D, many authors used Monte Carlo methods to approximate these. Among them Kiang [3], Kumar and Kurtz [4], Lorz and Hahn [5], Møller [6], Tanemura [7] obtained numerical results for the moments of the distribution of volume, surface area, and number of faces of the grain in 3D. They also give histogram estimates of these distributions and suggest approximations for them using various well known probability distributions. For instance, for the volume distribution, before 1990 most authors used the Lognormal distribution for approximating the grain size distribution in polycrystals. Nowadays, more flexible distributions such as Gamma or Generalized Gamma are commonly used (e.g. [4,7]). Although these models fit the observed data rather well (as we will see in the next section) our approach allows to find an accurate approximation of the true distribution and which parametric distribution optimally fits the data.

The main idea is that, given a Poisson-Voronoi diagram generated by a Poisson point process $\Phi$ with intensity parameter $\lambda$, this $\lambda$ is the only parameter determining the distributions of the geometrical features of the grains. Furthermore, the dependence of the distributions on the intensity parameter is via simple scaling of a ‘parent distribution’, due to the following important scaling property of the Poisson process.

**Lemma 3.1 (Scaling Property).** Let $\Phi = \{X_1, X_2, \ldots\}$ be a Poisson process on $\mathbb{R}^d$ with intensity $\lambda = 1$. Choose $\lambda > 0$ and define $\Phi_\lambda = \{X_1/\lambda^{1/d}, X_2/\lambda^{1/d}, \ldots\}$. Then $\Phi_\lambda$ is a Poisson process with intensity $\lambda$.

**Proof.** The fact that $\Phi_\lambda$ is a Poisson process is a special instance of the ‘Mapping theorem’ [see 2, Section 2.3], using states space $S = T = \mathbb{R}^d$ and $f(x_1, x_2, \ldots, x_d) = (x_1/\lambda^{1/d}, x_2/\lambda^{1/d}, \ldots, x_n/\lambda^{1/d})$. Denoting the mean measure of $\Phi$ (Lebesgue measure) by $\mu_1$, the induced mean measure $\mu_\lambda$ of $\Phi_\lambda$ is given by

$$\mu_\lambda(B) = \mu_1(f^{-1}(B)) = \int_{f^{-1}(B)} d\mu_1(x) = \lambda \int_B d\mu_1(x) = \lambda \mu_1(B)$$

□

In the following subsections Lemma 3.1 will be used to see how the distributions of volume, surface area and number of faces of the grains depend on the intensity parameter $\lambda$.

3.1. Grain Volume. We first focus our attention on the grain volume distribution because of the direct relationship of this Poisson-Voronoi geometrical characteristic and the grain size distribution in microstructure characterization of materials.

Exploiting the properties of the Poisson process, the distribution for the normalized length of the Voronoi cell in 1D or size measure in 1D, can be shown to have density [8]

$$f_{1D}(y) = 4y \exp(-2y) \ 1_{[0,\infty)}(y)$$

In dimension $d > 1$, it was conjectured that the area (2D) and the volume (3D) of the typical cell in a Poisson-Voronoi diagram may be distributed as the sum of two and three gamma variables with shape and scale parameters equal to 2 [3], but [9] and [10] showed the conjecture to be false. In 2D an analytic, though computationally challenging result is provided by Calka [11], which gives an expression for the distribution of the area of the typical cell in 2D given the number of vertices. In 3D, as we know so far, no analytic expression for the volume distribution exists.
Lemma 3.2. Denote by $F_\lambda$ the distribution function of the volume (length if $d = 1$, area if $d = 2$) of the typical cell of the Poisson-Voronoi diagram based on a homogeneous Poisson process on $\mathbb{R}^d$ with intensity parameter $\lambda > 0$. Then, for all $x \geq 0$,

\begin{equation}
F_\lambda(x) = F_1(\lambda x)
\end{equation}

Proof. Let $\Phi$ be a homogeneous Poisson process on $\mathbb{R}^d$ with intensity 1. Denote by $C$ the typical cell of the Voronoi diagram based on this process. Fix $\lambda > 0$ and consider the homogeneous Poisson process $\Phi_\lambda$ with intensity $\lambda$ as introduced in the statement of Lemma 3.1. Then the typical cell in the Voronoi diagram based on $\Phi_\lambda$ is a scaled version of the typical cell of the Voronoi diagram based on $\Phi$, in the sense that it is given by $C_\lambda = \{x/\lambda^{1/d} : x \in C\}$. This means that the volume $V_\lambda$ of $C_\lambda$ is exactly $\lambda^{-1}$ times the volume $V$ of $C$. Therefore, for $x \geq 0$,

$$F_\lambda(x) = P(V_\lambda \leq x) = P\left(\frac{V}{\lambda} \leq x\right) = P(V \leq \lambda x) = F_1(\lambda x)$$

\qed

3.2. Grain Surface area.

Lemma 3.3. Denote by $G_\lambda$ the distribution function of the surface area of the typical cell of the Poisson-Voronoi diagram based on a homogeneous Poisson process on $\mathbb{R}^3$ with intensity parameter $\lambda > 0$. Then, for all $x \geq 0$,

$$G_\lambda(x) = G_1\left(\lambda^{\frac{2}{3}}x\right)$$

Proof. The argument follows the proof of Lemma 3.2. Denote by $S_\lambda$ the surface area of $C_\lambda$ and note that scaling of $C_\lambda$ implies that $S_\lambda$ is $\lambda^{-\frac{2}{3}}$ times the surface area of $C$, $S$. Therefore

$$G_\lambda(x) = P(S_\lambda \leq x) = P\left(\frac{S}{\lambda^{\frac{2}{3}}} \leq x\right) = P\left(S \leq \lambda^{\frac{2}{3}}x\right) = G_1\left(\lambda^{\frac{2}{3}}x\right)$$

\qed

3.3. Number of grain faces. Finally, another (discrete) property of interest regards the number of grain faces of the typical cell. It is clear that using $\Phi$ or $\Phi_\lambda$ (from Lemma 3.1) as a basis for the Voronoi diagram, does not make any difference in the number of faces of the typical cell ($C$ or $C_\lambda$ respectively), leading to

Lemma 3.4. Denote by $N_\lambda$ the distribution function of the number of faces of the typical cell of the Poisson-Voronoi diagram based on a homogeneous Poisson process on $\mathbb{R}^d$ with intensity parameter $\lambda > 0$. Then, for all $x \geq 0$,

$$N_\lambda(x) = N_1(x)$$

The same lemma holds for number of corner points, $n_v$. In fact, exploiting the Euler-Poincaré relation, it is possible to determine $n_v$, when the number of faces is known.
4. Simulation Results

Now, we intend to approximate the real distribution function of the grain geometrical features, using the results obtained by a simulation based on 1 000 000 Voronoi diagrams. Our simulation approach also allows to determine which of the usual parametric models provides the best approximation to the true distribution and its deviation from it. We consider the volume, the surface area and the number of faces of a 3D Poisson Voronoi typical cell. There exist two possible approaches well described in [1]:

1. generate a large number of points inside a bounded region \( B \) according to \( \Phi \), construct \( \mathcal{V}_\Phi \) and measure the characteristics of all its cells.

2. generate a sequence of independent typical Poisson Voronoi cells, measure the characteristics of each and then aggregate them to obtain the required distributions.

We follow the second approach. The reason of this choice derives from the convenience of having a sample of independent and identically distributed Voronoi cells such that we can know how close we are to the real distribution. Moreover, we are able to control and eliminate the boundary effect that is present because the structure is actually only constructed on a bounded region. For our objective it is important that only the distributions of the geometrical properties of the typical cell are needed using \( \lambda = 1 \) in the simulations. By Lemma 3.2, 3.3 and 3.4, the distributions based on diagrams with different intensities can be obtained by scaling.

We conduct our simulation approach using the software provided by TATA Steel. The algorithm is based on the half plane intersection, which is closely related to the original definition of a Voronoi tessellation. Each Voronoi cell is constructed separately by intersecting \( n - 1 \) half spaces. A disadvantage is that this algorithm computes in \( O(n^2 \log n) \) time [12], while the most frequently used incremental algorithms can do it in \( O(n^2) \) time. To speed up the computations, the algorithm has been extended with a filter, which determines which neighboring points of a generator point (nucleus) are needed for the Voronoi cell construction of this nucleus point. This filter is built in such a way that it first sorts \( \sim 80\% \) of the points which are certainly needed for Voronoi cell construction. After that, the other \( \sim 20\% \) of the points are checked to see if they give half plane intersection with the Voronoi cell under construction. With this filter the computational speed is improved to be better than \( O(n \log n) \), which is the computational speed of the fastest algorithm by Fortune [13].

Then, we adopt the following Monte-Carlo procedure.
Repeat 1 000 000 times:

**Step 1:** Generate a 3D Poisson-Voronoi diagram with added generator point \((0,0,0)\) with \( \lambda = 1 \);

**Step 2:** Determine the geometrical characteristics of the realizations of the typical Voronoi cell, the cell that contains the point \((0,0,0), \mathcal{C}(0)\);

Then, aggregate the 1 000 000 values.

The main graphical results are shown in Figures 1, 2 and 3.

The values of the estimated densities of the previously mentioned geometrical characteristics are given at http://dutiosb.twi.tudelft.nl/ martina.vittorietti/. An R-package is under construction. In Tables 1 and 2, we report the estimated moments of the main geometrical characteristics and the estimated probabilities for the number of faces. They are coherent with both the theoretical and numerical results obtained by other authors [4,7].
Figure 1: (a) Kernel density estimate (Epanechnikov kernel, cross validation bandwidth $h = 0.05$) and (b) empirical cumulative distribution function of volume of 1,000,000 Poisson-Voronoi typical cells, $\lambda = 1$.

Figure 2: (a) Kernel density estimate (Epanechnikov kernel, cross validation bandwidth $h = 0.25$) and (b) empirical cumulative distribution function of surface area of 1,000,000 Poisson-Voronoi typical cells, $\lambda = 1$. 
Figure 3: (a) Relative frequencies and (b) empirical cumulative distribution function of number of faces of 1,000,000 Poisson-Voronoi typical cells, $\lambda = 1$

Table 1: Estimated moments of the geometrical features of 1,000,000 Poisson-Voronoi typical cells, $\lambda = 1$

|      | Volume       | Surface area | Number of faces |
|------|--------------|--------------|-----------------|
| (a)  | $\mu_1$ 1.00008 | $\mu_1$ 5.82670 | $\mu_1$ 15.53071 |
|      | $\sigma$ 0.41189 | $\sigma$ 1.43821 | $\sigma$ 3.33896 |
| (b)  | $\mu_2$ 1.16981 | $\mu_2$ 36.01888 | $\mu_2$ 252.35173 |
|      | $\mu_3$ 1.55900 | $\mu_3$ 234.69091 | $\mu_3$ 4277.80397 |
| (c)  | $\mu_4$ 2.32340 | $\mu_4$ 1603.48468 | $\mu_4$ 75464.60519 |
Table 2: Distribution of the number of faces $F$ of 1 000 000 Poisson-Voronoi typical cell, $\lambda = 1$

|   | $n_f$ | $p_f$   |
|---|---|---|
| 4 | 5  | 0.000005 |
| 5 | 35 | 0.000035 |
| 6 | 316| 0.000316 |
| 7 | 1822| 0.001822 |
| 8 | 6190| 0.006190 |
| 9 | 15051| 0.015051 |
| 10 | 30685| 0.030685 |
| 11 | 52528| 0.052528 |
| 12 | 77421| 0.077421 |
| 13 | 100094| 0.100094 |
| 14 | 114163| 0.114163 |
| 15 | 120015| 0.120015 |
| 16 | 115188| 0.115188 |
| 17 | 101151| 0.101151 |
| 18 | 82277| 0.082277 |
| 19 | 62408| 0.062408 |
| 20 | 44944| 0.044944 |
| 21 | 30477| 0.030477 |
| 22 | 19466| 0.019466 |
| 23 | 11682| 0.011682 |
| 24 | 6756| 0.006756 |
| 25 | 3631| 0.003631 |
| 26 | 1890| 0.001890 |
| 27 | 975| 0.000975 |
| 28 | 435| 0.000435 |
| 29 | 224| 0.000224 |
| 30 | 95| 0.000095 |
| 31 | 52| 0.000052 |
| 32 | 18| 0.000018 |
| 33 | 3| 0.000003 |
| 34 | 1| 0.000001 |
| 35 | 1| 0.000001 |
| 36 | 1| 0.000001 |
5. Parametric approximations to the distributions

Various proposals on how to estimate the distributions of the geometrical properties of the typical cell were put forward by several authors such as the Lognormal distribution \([14,15]\), Generalized Gamma distribution with 2 \([17]\) or 3 parameters \([7]\) and Rayleigh distribution \([16]\). Ferenc and Néda \([10]\) propose their own function for the volume distribution. Anyway, nobody was able to find an analytic expression for the grain geometrical characteristics distributions.

As noted in \([17]\) the use of the Lognormal distribution function for approximating grain size distribution lacks a solid physical basis and is not in general accurate. Nowadays, the debate regards mostly the Generalized Gamma Distribution with 2 or 3 parameters, but until now no physical explanation for using one preferential distribution exists. However, in view of the scaling properties described in the previous sections, it is natural to think that the distributions of the geometrical characteristics of the grain belong to a scale parametric family of distributions. Only then the distributions of the quantities for all \(\lambda\) can belong to the class. Let us look, for instance, at the Lognormal distribution. Its probability density is given by

\[
f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}.
\]

Let \(\hat{\mu}_1\) and \(\hat{\sigma}_1\) be the maximum likelihood estimates when \(\lambda = 1\) (based on the 1,000,000 simulated values). Define \(f_\lambda(x)\) as:

\[
f_\lambda(x) = \lambda f(\lambda x|\hat{\mu}_1, \hat{\sigma}_1) = \frac{1}{x\hat{\sigma}_1\sqrt{2\pi}}e^{-\frac{(\log(x)+\log(\lambda)-\hat{\mu}_1)^2}{2\hat{\sigma}_1^2}}
\]

which corresponds to a Lognormal distribution with parameter vector \((\hat{\mu}_1-\log(\lambda), \hat{\sigma}_1^2)\). Therefore, we have a log-addition scaling on the first parameter, which is not consistent with the \(\lambda\)-scaling that is found for real distributions. Now let us consider the Generalized Gamma distribution. Its density function, parameterized according to \([18]\), is given by

\[
f(x|a, b, k) = \frac{bx^{ak-1}}{\Gamma(k)} e^{-\left(\frac{x}{a}\right)^b} \quad \text{where } a \text{ and } b \text{ are the shape and the scale parameters.}
\]

Let \(\hat{a}_1, \hat{b}_1\) and \(\hat{k}_1\) be the maximum likelihood estimates for the parameters (based on the 1,000,000 simulated values) when \(\lambda = 1\). Define \(f_\lambda(x)\) as equal to:

\[
f_\lambda(x) = \lambda f(\lambda x|\hat{a}_1, \hat{b}_1, \hat{k}_1) = \frac{\hat{b}_1\hat{a}_1^{\frac{b_1}{k_1}}}{\Gamma(\hat{k}_1)} \left(\frac{\lambda}{\hat{a}_1}\right)^{\frac{b_1}{k_1}} \left(\frac{x}{\hat{a}_1}\right)^{\frac{b_1}{k_1}} e^{-\left(\frac{x}{\hat{a}_1}\right)^{\frac{b_1}{k_1}}}
\]

which corresponds to a Generalized Gamma with parameters \((\hat{a}_1, \hat{b}_1, \hat{k}_1)\). This suggests to look for a distribution that belongs to this scaling family. Special cases of this family are Gamma distribution with parameters \(a, k\) and \(b = 1\); Weibull distribution with parameters \(a, b\) and \(k = 1\). Beside the parametrization in eq.5.1, another one is provided by Prentice \([19]\). This is in general more stable in the estimation of the parameters but both parametrizations lead to the same estimates. In the next subsections, we report the estimated parameters of the best Generalized Gamma approximations and we statistically compare the fits based on the Gamma distribution, the Generalized Gamma distribution and the Lognormal distribution. This comparison is based on two criteria:
• Supremum distance between two distribution functions:

\[ D(F, G) = \sup_{x \in \mathbb{R}} |F(x) - G(x)| \]

This distance is computed between the empirical distribution function \( F_n \) of the sample and the maximum likelihood fit within the respective parametric families. The data are the 1,000,000 simulated values from the distribution of interest, with \( \lambda = 1 \).

• Total Variation distance between distributions with distribution functions \( F \) and \( G \) and densities \( f \) and \( g \) respectively on \( \mathbb{R} \):

\[ TV(F, G) := \sup_{A \in \mathcal{B}} \left| \int_A dF(x) - \int_A dG(x) \right| = \frac{1}{2} \int |f(x) - g(x)| \, dx \]

This distance is computed between the kernel estimate of the densities and maximum likelihood parametric fits based on the simulation results with \( \lambda = 1 \).

Note that information on these distances are based on the data obtained with \( \lambda = 1 \). If the estimates for more general values of \( \lambda \) are obtained via the rescaling, these distances do, however, not change under this scaling, so the distances also hold for the other values of \( \lambda \).

![Figure 4: Comparison of parametric approximations to the volume distribution of 1 000 000 Poisson-Voronoi typical cells, \( \lambda = 1 \)](image)

![Table 3: Estimated Generalized Gamma parameters for volume distribution approximation, \( \lambda=1 \)](table)

|       | \( \hat{a} \) | \( \hat{b} \) | \( \hat{k} \) |
|-------|------------|------------|------------|
| Estimate | 0.380 | 1.287 | 3.583 |
| Std. Error | 0.005 | 0.006 | 0.0322 |
Figure 5: (a) QQplot and (b) cumulative distribution function comparison of parametric approximations to the volume distribution of 1000000 Poisson-Voronoi typical cells, $\lambda = 1$

Table 4: Comparison of Gamma-, Generalized Gamma- and Lognormal approximations for volume distribution in terms of Supremum- and Total Variation distance

|                | Gamma | Generalized Gamma | Lognormal |
|----------------|-------|-------------------|-----------|
| Supremum distance | 0.013 | 0.005             | 0.041     |
| TV distance      | 0.018 | 0.005             | 0.089     |

Table 5: Estimated Generalized Gamma parameters for surface area distribution approximation

| $a$   | $\hat{b}$ | $\hat{k}$ |
|-------|-----------|-----------|
| Estimate | 3.174     | 2.102     | 3.839     |
| Std. Error | 0.025     | 0.011     | 0.036     |
Figure 6: Comparison of parametric approximations to the surface area distribution of 1,000,000 Poisson-Voronoi typical cells, $\lambda = 1$

Figure 7: (a) QQplot and (b) cumulative distribution function comparison of parametric approximations to the surface area of 1,000,000 Poisson-Voronoi typical cells, $\lambda = 1$
Table 6: Comparison of Gamma-, Generalized Gamma- and Lognormal approximations for surface area distribution in terms of Supremum- and Total Variation distance

|                | Gamma | Generalized Gamma | Lognormal |
|----------------|-------|-------------------|-----------|
| Supremum distance | 0.020 | 0.002             | 0.037     |
| TV distance     | 0.035 | 0.003             | 0.082     |

6. RESULTS AND DISCUSSION

This work provides a very accurate approximation for the distributions of all the main geometrical characteristics of Poisson-Voronoi typical cell. It is now possible to exploit our approximated distribution for generating observations of approximately every geometrical characteristics that define the grain size and for every possible value of $\lambda$. Moreover, we prove that the Generalized Gamma distribution, with parameter $\hat{a} = 0.380$, $\hat{b} = 1.287$, $\hat{k} = 3.583$ for the volume and $\hat{b} = 3.174$, $\hat{a} = 2.102$, $\hat{k} = 3.839$ for the surface area is the best approximation in the class of the commonly used parametric distributions for grain size distributions. Nevertheless, it is not the true underlying distribution. In fact, the interpretation of the total variation distance as measure of quality allow us to say that using Generalized Gamma approximation for the grain volume distribution we could commit an error of about 0.5% and about 0.3% for the grain surface area.

As noted in the introduction, Poisson-Voronoi diagrams are interesting and widely applied, but for modeling microstructures only constitute the most basic case. Their use in microstructure characterization is not fully evaluated yet. Therefore, we want to test their applicability and then extend our work to more general and less understood Voronoi structures, such as Multi-level Voronoi diagrams [20] or Laguerre-Voronoi tessellations [21] in which the convenient scaling properties present in the Poisson-Voronoi diagrams are less obvious.

ACKNOWLEDGEMENTS

This research was carried out under project number S41.5.14547b in the framework of the Partnership Program of the Materials innovation institute M2i (www.m2i.nl) and (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

[1] Okabe A., Boots B., Sugihara K., Chiu S.N. Spatial tessellations: concepts and applications of Voronoi diagrams. John Wiley & Sons; 2009; vol. 501.
[2] Kingman J.F.C. Poisson processes. Wiley Online Library; 1993.
[3] Kiang T. Random fragmentation in two and three dimensions. Zeitschrift fur Astrophysik; 1966; vol.64; p. 433.
[4] Kumar S., Kurtz S.K. Monte-Carlo study of angular and edge length distributions in a three-dimensional Poisson-Voronoi tesselation. Elsevier; Materials characterization; 1995; vol.34; p. 15–27.
[5] Lorz U., Hahn U. Geometric characteristics of random spatial Voronoi tessellations and planar sections. Dekan des Fachbereiches Mathematik; 1993.
[6] Møller J. Lectures on random Voronoi tessellations. Ney York: Notes Statist. Springer; 1994; vol. 87.
[7] Tanemura M. Statistical distributions of Poisson Voronoi cells in two and three dimensions. SCI-PRESS: FORMA-TOKYO; 2003; vol. 18; p. 221–247.
[8] Meijering J.L. Interface area, edge length, and number of vertices in crystal aggregates with random nucleation. Philips Res. Rep; 1953; vol. 8, p. 270–290.
[9] Zaninetti L. Poissonian and non-Poissonian Voronoi diagrams with application to the aggregation of molecules. Physics Letters A; 2009; vol. 373; p. 3223–3229.
[10] Ferenc J.S. Néda Z. On the size distribution of Poisson Voronoi cells. Elsevier; Physica A: Statistical Mechanics and its Applications; 2007; vol. 385; p. 518–526.
[11] Calka P. et al. Precise formulae for the distributions of the principal geometric characteristics of the typical cells of a two-dimensional Poisson-Voronoi tessellation and a Poisson line process. Advances in Applied Probability; 2003. p. 551–562
[12] O’Rourke J. Computational geometry in C. Cambridge university press; 1998.
[13] Fortune S. A sweepline algorithm for Voronoi diagrams. ACM: Proceedings of the second annual symposium on Computational geometry; 1986; p. 313–322.
[14] Smith J.E., Jordan M.L. Mathematical and graphical interpretation of the log-normal law for particle size distribution analysis. Elsevier; Journal of Colloid Science; 1964; vol. 19; p. 549–559.
[15] Söderlund J., Kiss L.B., Niklasson G.A., Granqvist C.G. Lognormal size distributions in particle growth processes without coagulation. APS: Physical review letters; 1998; vol. 80; p. 2386.
[16] Pande C.S. On a stochastic theory of grain growth. Elsevier; Acta Metallurgica; 1987; vol. 35; p. 2671–2678.
[17] Vaz M.F, Fortes M.A. Grain size distribution: The lognormal and the gamma distribution functions. Elsevier; Scripta metallurgica; 1988; vol. 22; p. 35–40.
[18] Stacy E.W., Milram G.A. Parameter estimation for a generalized gamma distribution. Technometrics; 1965; p. 349–358.
[19] Prentice R.L. A log gamma model and its maximum likelihood estimation. Biometrika; 1974; p. 539–544.
[20] Kok P.J.J., Spanjer W., Vegter H. A microstructure based model for the mechanical behaviour of multiphase steels. Key Engineering Materials; 2015; vols. 651-653; p. 975–980.
[21] Lautensack C., Zuyev S. Random Laguerre Tessellations. Advances in Applied Probability; 2008; vol. 40; no. 3; p. 630–650.