Stereological estimation for particle processes and random tessellations

van der Jagt, T.F.W.

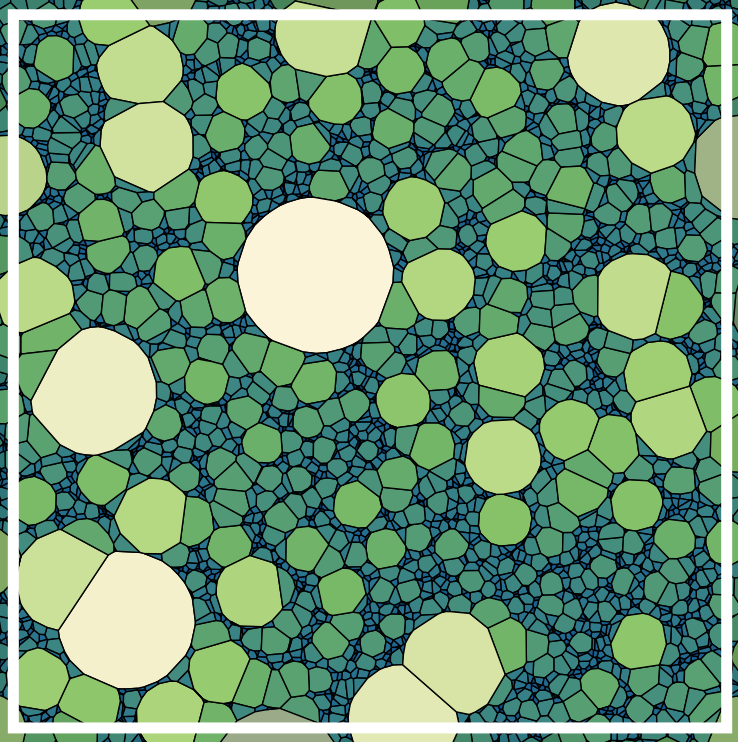
DOI
10.4233/uuid:ec3403bd-6129-42dc-8fa5-5adb2c46d648

Publication date
2026

Document Version
Final published version

Citation (APA)
van der Jagt, T. F. W. (2026). *Stereological estimation for particle processes and random tessellations*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:ec3403bd-6129-42dc-8fa5-5adb2c46d648

Stereological estimation for particle processes and random tessellations



Thomas van der Jagt

Stereological estimation for particle processes and random tessellations

Stereological estimation for particle processes and random tessellations

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on Tuesday the 3rd of March 2026 at 12:30

by

Thomas Folkert Willem VAN DER JAGT

Master of Science in Applied Mathematics,
Delft University of Technology, The Netherlands,
born in Rotterdam, The Netherlands.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. G. Jongbloed,	Delft University of Technology, *promotor*
Dr. M. Vittorietti,	Delft University of Technology, *copromotor*
Independent members:	
Prof. RNDr. V. Beneš,	Charles University
Prof. dr. A.J. Cabo,	Delft University of Technology
Prof. dr. M.N.M. van Lieshout,	Centrum Wiskunde & Informatica, University of Twente
Prof. dr. M.J. Santofimia Navarro,	Delft University of Technology
Prof. dr. A.W. van der Vaart,	Delft University of Technology

Keywords: Stereology, nonparametric estimation, consistency, point process, Laguerre tessellation, stochastic geometry, microstructure

Printed by: Proefschriftspecialist

Cover: A Poisson-Laguerre tessellation viewed through a square observation window.

ISBN 978-94-6384-895-4

An electronic copy of this dissertation is available at
https://repository.tudelft.nl/.

Contents

Chapter 1

Introduction

1.1 A gentle introduction to stereology

At the time of writing this thesis, it is precisely a century ago since the statistician Sven Dag Wicksell introduced the corpuscle problem in his paper [114], which is now commonly known as Wicksell's corpuscle problem. The problem he was dealing with originated from anatomy and may be described as follows. In post-mortem studies so-called follicles or corpuscles were observed in slices of organs. The follicles observed in the 2D slice were approximately circular, suggesting that the original follicles were approximately 3D balls. Anatomists at the time were interested in determining the statistical distribution of the radii of the (3D) follicles. This inspired Wicksell to define the following mathematical model. Suppose balls of varying sizes are randomly positioned in an opaque medium. These balls cannot be observed directly, instead the medium is intersected with a plane, and we observe the circle radii of the balls which happened to be hit by the plane. The problem is then to estimate the probability distribution of the 3D radii given the sample of the observed 2D radii. A visualization of the problem setting is shown in Figure 1.1.

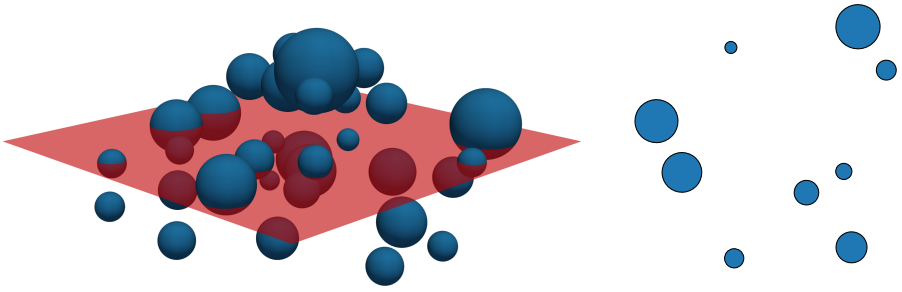
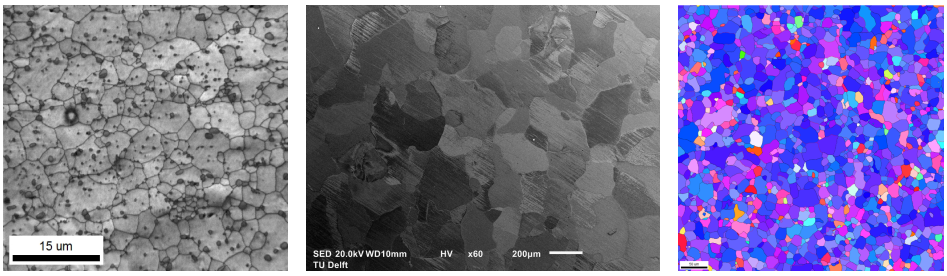


Figure 1.1: Left: Random spatial system of balls intersected with a plane. Right: Observed section profiles.

More formally, one may consider the centers of the balls to be a realization of a homogeneous Poisson point process, and their radii are drawn independently from a common probability distribution with distribution function F. As it turns out, it is possible to express this unknown distribution function F in terms of the distribution of the observed 2D radii. Despite of this inversion formula, estimation of F is a not a trivial matter, and various estimators for F have been proposed over the years. We refer to [18] for an overview of various estimators, and for some recent developments see for instance [33] and [34].

Wicksell's corpuscle problem belongs to a larger field of research which is known as stereology. Stereology deals with the estimation of higher dimensional information from lower dimensional observations. In many practical applications this often means that one is interested in 3D characteristics of some object, while only 2D observations are available. An important field of application for stereology is materials science, and throughout this thesis, applications to materials science will be the main motivation for our research. For steel, there are unfortunately no quick or cost-effective methods to perform a 3D scan of the steel sample. It is however possible to obtain 2D microscopic images of the surface, or cross sections of the steel sample. Examples of such images, obtained via different imaging techniques can be seen in Figure 1.2. In these images one can see that the so-called microstructure of a steel consists of some kind of cells, called grains. In the images the grains are the approximately polygonal areas delimited by lines called grain boundaries. The reason that accurately describing or characterizing the 3D microstructure of a material is of interest is because properties of the 3D microstructure, such as the grain size distribution, are closely related to mechanical properties of the material, such as its strength. If an accurate model of a materials microstructure is available, laborious experiments such as stress tests may be replaced by computer simulations, thereby saving time, energy and resources. Additionally, a better understanding of materials microstructures may eventually allow for designing new materials with desirable properties.

Throughout this thesis, we mainly take a model-based approach to stereology. This means that we mathematically define a 3D structure, which involves randomness, such as the random placement of balls of random sizes in the Wicksell model. We will then intersect this 3D structure with a plane, and then we aim to estimate the parameters of the



(a) (b) (c)

Figure 1.2: Images of metal microstructures, taken using various imaging techniques and different scales. Pictures taken by (a) Javier Hidalgo, (b) Wei Li, (c) Carola Celada-Casero.

underlying model based on the resulting 2D observations. We should also mention that design-based stereology is another approach to stereology. In design-based stereology the 3D structure at hand is considered deterministic as opposed to random, and randomness is introduced by sampling with random section planes. In this thesis, only in chapter 2 we take a design-based approach. For a more complete overview of design-based stereology we refer to [9] and [20]. For model-based stereology, see chapter 10 in [18].

1.2 Stochastic geometry

As described in the previous section, we study mathematical models for 3D structures, in particular materials microstructures. Because stochastic geometry is a branch of mathematics which deals with random geometric patterns and structures, it is an obvious starting point for our purposes. An essential component of many models in stochastic geometry is a so-called point process, which may also be described as a random point pattern. In the previous section we already mentioned that a homogeneous Poisson point process may be used to position the randomly sized balls in Wicksell's corpuscle problem. A realization of a homogeneous Poisson point process in 2D is shown in the left panel of Figure 1.3. Of course, the Wicksell model involves a homogeneous Poisson process in 3D, but this should provide some intuition on what these random point patterns may look like. The Poisson process is often considered to be the most important point process, and it will also play a significant role in this thesis. For more details on Poisson point processes and more general point processes we refer to [22], [23] and [53]. For a thorough treatment of stochastic geometry we refer to [86] and [18].

As is evident from the title of this thesis, we are interested in stereological estimation for particle processes and random tessellations. An example of a particle process is for example the random system of 3D balls in the Wicksell model. Mathematically, a particle process is a specific type of point process where each "point" is in fact a set, and may be

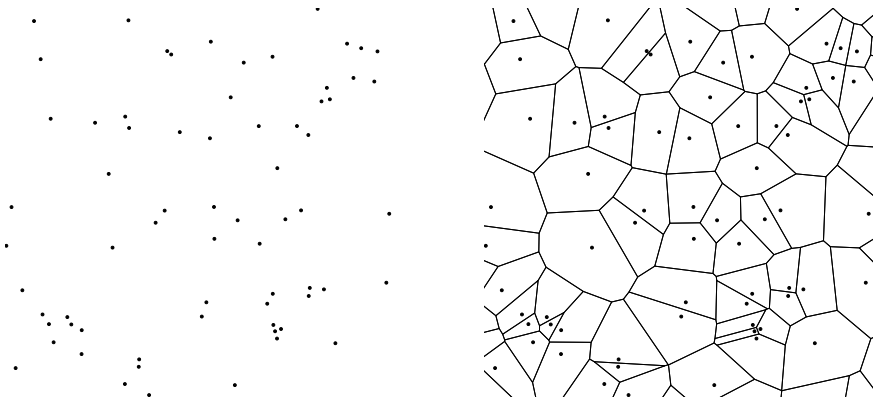


Figure 1.3: Left: A realization of a homogeneous Poisson point process. Right: The corresponding Poisson-Voronoi tessellation.

referred to as a particle or a grain. A particle in a particle process may for example be a (randomly sized) ball or a (random) polytope. If the particles of the particle process have a similar shape as the grains in a real materials microstructure, one can imagine that such a particle process may be a reasonable model for the microstructure at hand. That being said, even if the particles have similar shapes as the grains in a microstructure, this may not be enough to consider the model sufficiently realistic. For instance, if the particles are positioned via a Poisson point process, then the particles will not necessarily "fill space". Recall Figure 1.1, there is clearly empty space between the balls. Meanwhile, if one looks at the images of steel microstructures shown in Figure 1.2 then there is no empty space between grains. Evidently, replacing the balls in the Wicksell model with a shape that is similar to the shape of a grain in a real material is a step forward, compared to just using the Wicksell model. Moreover, an attractive feature of these kinds of models is that they are tractable for mathematical analysis. We should also note that there are materials which do not have a space-filling microstructure. Then, these models of randomly positioned particles are of interest as they may in fact accurately describe the real microstructure. However, in order to obtain more accurate models for steel microstructures, considering models which are space-filling seems to be crucial. Therefore, we also study random tessellations.

A tessellation or mosaic is a way to divide 3D or 2D space into non-overlapping pieces. Mathematically, a random tessellation may be described as a specific kind of particle process, which has non-overlapping particles and which does "fill space". An example of a random tessellation is shown on the cover of this thesis. The particles of a random tessellation are usually referred to as cells. A well-known example of a random tessellation is the so-called Poisson-Voronoi tessellation. A realization of a 2D Poisson-Voronoi tessellation is shown in the right panel of Figure 1.3. This tessellation is constructed as follows. First, one simulates a homogeneous Poisson process. Then, to each point (or seed) in the Poisson process a cell is assigned. In 2D, a cell in a Poisson-Voronoi tessellation is a polygon, and a point in 2D space is contained in this cell if it is closer to the seed of this cell than it is to any of the other seeds. Stereological estimation for Poisson-Voronoi tessellations was considered in [42]. Intersecting a 3D Poisson-Voronoi tessellation with a plane yields a tessellation which is no longer a Voronoi tessellation. Only recently a precise mathematical description of this random tessellation was given in [41]. For more details on random Voronoi tessellations we refer to [67]. As the Poisson-Voronoi tessellation only has one parameter (the intensity of the Poisson point process) it may not be flexible enough for some applications. One could then for instance consider the Poisson-Laguerre tessellation, which is a more advanced tessellation model. For an overview of random tessellation models see [78].

1.3 Overview of the thesis

In this section we present an overview of the contents of this thesis. In the previous sections we provided a motivation for studying stereological estimation and we informally introduced the main mathematical concepts considered in this thesis. In principle, the chapters in this thesis are self-contained, and all necessary definitions and notation are

introduced in each chapter. That being said, some of the described theory is needed in multiple chapters, and therefore we may occasionally refer the reader to sections in previous chapters. Readers who are not yet familiar with point processes, particle processes and random tessellations may consult Appendix A. In this appendix we provide a formal introduction of these concepts.

This thesis may be of interest to both mathematicians as well as materials scientists. We should however note that in most chapters the focus is on formal mathematical results. The chapters in this thesis are mainly based on published papers, as we outline below. A complete list of publications used for writing the chapters in this thesis is provided at the end of this thesis.

This thesis is organized in two parts. In part I, we consider stereological estimation for a specific class of particle processes. This may be seen as a generalization of Wicksell's corpuscle problem. In this generalization, 3D particles are randomly positioned and -oriented in an opaque body. These particles are all of the same convex shape, but not of the same size. Intersecting the opaque medium with a plane, some of the particles are also hit by this plane. This procedure yields a sample of observed 2D section profiles, and we wish to use the areas of these section profiles to estimate the size distribution of the 3D particles. In order to deal with this problem we first study random section areas of a fixed convex shape. This is the content of chapter 2, which is based on the paper [105]. The results in chapter 2 are then used in chapter 3 to define a nonparametric estimator for the 3D particle size distribution, given a sample of observed section profile areas. The statistical methodology developed in chapter 3 is based on the publication [106].

Next, we consider a practical application. We apply the estimation procedure described in chapter 3 to microscopic image data of a real steel sample in chapter 4. Additionally, we apply the estimation procedure to various simulated materials microstructures. We study the quality of the resulting estimates, and shed some light on how our statistical methodology may be of use for materials scientists. In this chapter we also briefly summarize the statistical methodology derived in chapter 3. The content of chapter 4 is based on the paper [108].

In part II of this thesis we focus on a class of random tessellations known as Poisson-Laguerre tessellations. This model is a more flexible generalization of the classical Poisson-Voronoi model, as it allows for more variation in cell sizes. We study these random tessellations not just from a stereological point of view (observations via a planar section) but also from the point of view of direct observations, as this already poses various challenges. A statistical methodology for Poisson-Laguerre tessellations is described in chapter 5, which is based on [107]. In Chapter 6 we study whether one can estimate the weighted generators corresponding to the observed cells of a Poisson-Laguerre tessellation. This is of interest because knowledge of these weighted generators is an essential ingredient for statistical inference for Poisson-Laguerre tessellations.

At the end of this thesis we provide some conclusions in chapter 7. Here, we also highlight several interesting future research directions. Finally, this thesis has two appendices. In Appendix A we provide a theoretical background for this thesis and in Appendix B we present auxiliary results for chapter 3.

Part I

Particle processes with a fixed grain shape

Chapter 2

The section volume distribution of a convex body

2.1 Introduction

In a typical stereological problem we are presented with observations which originate from a lower dimension than the dimension of interest. Recall from the introduction the classical example of the Wicksell corpuscle problem, which considers the following setting. Balls of varying size are randomly positioned in 3D space. This system of balls is intersected with a plane and the circular section profiles of the balls which happened to be cut by the section plane are observed. The problem is to determine the distribution of the radii of the 3D balls given the distribution of the radii of the observed 2D circular profiles.

An interesting generalization of this problem is to choose a convex shape other than the ball for the shape of the particles. Then, the distribution of observed section areas may be used to estimate the size distribution of the particles. The particles we consider are convex bodies, i.e. compact and convex sets with non-empty interior. In order to deal with such problems we study a class of distributions which is especially important in this setting. Suppose we take some convex body $K \subset \mathbb{R}^3$ of choice and intersect K with a random section plane. More generally, we may take a convex body $K \subset \mathbb{R}^d$, and intersect it with a random $(d-1)$-dimensional hyperplane. The random section planes we consider are Isotropic Uniformly Random (IUR) planes, which will be formally introduced in the next section. What can be said regarding the cumulative distribution function (CDF) G_K associated with the $(d-1)$-dimensional volume of such a random section of K? In this chapter we study this kind of distribution functions. In particular, we obtain results on absolute continuity. Whenever we refer to absolute continuity of a cumulative distribution function we mean absolute continuity with respect to Lebesgue measure. The existence and the accurate approximation of the density of G_K is an essential ingredient for defining likelihood-based estimators for particle size distributions as we will see in chapter 3.

Given a convex body $K \subset \mathbb{R}^2$ an IUR section of K is the intersection of K with a random line. The distribution function G_K is then also known as a chord-length distribution function. Some results regarding this function may be found in [32]. The author notes

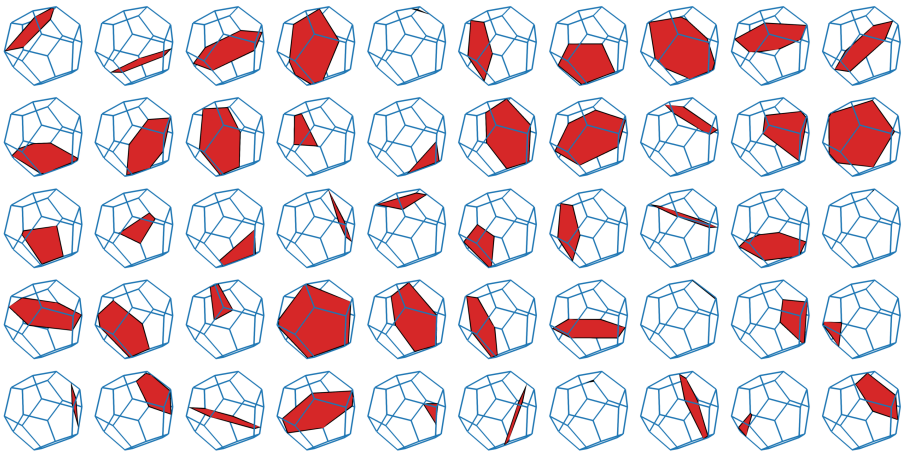


Figure 2.1: 50 IUR sections through a convex dodecahedron.

that it is typically assumed without proof that the CDF of a chord length distribution is absolutely continuous. Only for a limited set of convex polygons there are some results on absolute continuity. See for example [44] for the chord length distribution function of a regular polygon which is absolutely continuous.

For convex bodies $K \subset \mathbb{R}^3$ the distribution function G_K is sometimes called a cross section area distribution. In [84] it is noted that in a stereological setting it is of interest to obtain the density of G_K for some basic shapes such as the simplex or the cube. However, to the best of our knowledge there are no results on whether G_K has a density for a large class of convex bodies, especially in $\mathbb{R}^d$ with $d \geq 3$. To overcome the difficulty in obtaining an expression for G_K, simulations may be used to find an approximation. In [69] a description is given for approximating G_K (when K is a polytope in $\mathbb{R}^3$) and for how it may be used to estimate the size distribution of particles from a sample of observed section areas.

The outline of this chapter is as follows. First, we introduce necessary notation and definitions in section 2.2. In section 2.3 we present various results on the distribution function G_K. In particular, we show that for a large class of convex bodies, G_K is absolutely continuous. In section 2.4 we propose a Monte Carlo simulation scheme to approximate the corresponding probability density function g_K. Finally, we provide some conclusions in section 2.5.

2.2 Isotropic Uniformly Random planes

In this section we formally introduce Isotropic Uniformly Random (IUR) planes, along with some necessary notation and definitions. In particular, we introduce some terminology from convex geometry, a standard reference is [85]. In $\mathbb{R}^d$ a convex body is a convex and compact set with non-empty interior. Let $\mathcal{K}^d$ denote the class of convex bodies in $\mathbb{R}^d$.

Let $\nu_d(K)$ be the d-dimensional volume of K, its d-dimensional Lebesgue measure. K and L indicate convex bodies. Given a point $x \in \mathbb{R}^d$, the translation of K with x is given by: $K + x = \{k + x : k \in K\}$. The sum of two sets, also known as the Minkowski sum, is defined as: $K + L = \{k + l : k \in K, l \in L\}$. The dilatation or scaling of K with $\lambda > 0$ is given by: $\lambda K = \{\lambda k : k \in K\}$. ∂K denotes the boundary of K. Given $x \in \mathbb{R}^d$ and $r > 0$, we write $B(x, r) = \{y \in \mathbb{R}^d : \|x - y\| < r\}$ and $\bar{B}(x, r) = \{y \in \mathbb{R}^d : \|x - y\| \leq r\}$ for the open and closed ball respectively, with radius r centered at x. $\mathrm{SO}(d)$ denotes the rotation group of order d, containing all orthogonal $d \times d$-matrices of determinant one. Given $M \in \mathrm{SO}(d)$, the rotation of K with M is denoted by: $MK = \{Mk : k \in K\}$. We write $\operatorname{int} K$ and $\operatorname{relint} K$ to denote the interior and relative interior of K respectively. The set $\operatorname{relint} K$ contains all points $x \in K$ which are contained within the interior of the affine hull of K. That is, there exists a $r > 0$ such that $B(x, r) \cap \mathrm{aff(K)} \subset K$. Here, $\mathrm{aff}(K)$ is the affine hull of K, the smallest affine set containing K. A convex body $K \in \mathcal{K}^d$ is strictly convex if for all $x, y \in K$ and $\lambda \in (0, 1)$ we have $\lambda x + (1 - \lambda)y \in \operatorname{int} K$. A strictly convex body does not have any line segments in its boundary. The unit sphere in $\mathbb{R}^d$ is denoted by $\mathbb{S}^{d-1} = \{(x_1, \ldots, x_d) \in \mathbb{R}^d : x_1^2 + \cdots + x_d^2 = 1\}$. The upper hemisphere in $\mathbb{R}^d$ is given by: $\mathbb{S}_+^{d-1} = \{(x_1, \ldots, x_d) \in \mathbb{S}^{d-1} : x_d \geq 0\}$. Let σ_{d-1} denote the spherical measure on $\mathbb{S}^{d-1}$, also known as the spherical Lebesgue measure on $\mathbb{S}^{d-1}$. In integrals over (a subset of) $\mathbb{S}^{d-1}$ the notation $\mathrm{d}\theta$ should be interpreted as $\sigma_{d-1}(\mathrm{d}\theta)$. Recall that a hyperplane may be parameterized via a unit normal vector $\theta \in \mathbb{S}_+^{d-1}$ and its signed distance $s \in \mathbb{R}$ to the origin:

$$T_{\theta,s} = \{x \in \mathbb{R}^d : \langle x, \theta \rangle = s\}, \tag{2.1}$$

with $\langle \cdot, \cdot \rangle$ being the usual inner product in $\mathbb{R}^d$. The following definition gives a convenient parameterization of IUR planes, see [9] for IUR plane sections of convex bodies in $\mathbb{R}^3$ (the generalization to $\mathbb{R}^d$ is straightforward):

Definition 2.1 (IUR plane)**.** An IUR plane T hitting a fixed $K \in \mathcal{K}^d$, $d \geq 2$, is defined as $T = T_{\Theta,S}$ where (Θ, S) has joint probability density, $f_K : \mathbb{S}_+^{d-1} \times \mathbb{R} \to [0, \infty)$ given by:

$$f_K(\theta, s) = \begin{cases} \frac{1}{\mu([K])} & \text{if } K \cap T_{\theta,s} \neq \emptyset \\ 0 & \text{otherwise,} \end{cases}$$

with $T_{\theta,s}$ as in (2.1) and

$$\mu([K]) = \int_{\mathbb{S}_+^{d-1}} \int_{-\infty}^{\infty} \mathbb{1}\{K \cap T_{\theta,s} \neq \emptyset\} \mathrm{d}s \mathrm{d}\theta. \tag{2.2}$$

Loosely speaking this means that for an IUR plane through K, every plane which has a non-empty intersection with K has equal probability of occurring. The notion of IUR planes was originally introduced in [25]. It is important to highlight that there are other kinds of random planes which appear in stereological problems, hence care should be taken in considering the appropriate distribution. See [65] for more details. IUR planes may also be defined via invariant measures, see for instance section 8.4 in [86].

Note that the distribution in Definition 2.1 is a joint uniform distribution; the marginals are in general not uniform. We stress that the density f_K prescribes the probability associated with the possible locations and orientations of the section plane, not the volumes of

hyperplane sections. Fix $K \in \mathcal{K}^d$ and let f_K be as in Definition 2.1. Integrating out the variable s, we obtain the marginal density:

$$f_{K,\Theta}(\theta) = \frac{L(p_\theta(K))}{\mu([K])}, \quad \theta \in \mathbb{S}_+^{d-1}. \tag{2.3}$$

In (2.3), $p_\theta(K)$ represents the orthogonal projection of K on the line through the origin with direction θ. $L(p_\theta(K))$ is then the length of this orthogonal projection, hence $L(p_\theta(K))$ may also be called the width of K in direction θ. The constant $\mu([K])$ is related to the average width $\bar{b}(K)$, via:

$$\mu([K]) = \frac{\omega_d}{2}\bar{b}(K). \tag{2.4}$$

The average width is defined as:

$$\bar{b}(K) = \frac{2}{\omega_d}\int_{\mathbb{S}_+^{d-1}} L(p_\theta(K))\mathrm{d}\theta, \tag{2.5}$$

Here, $\omega_d = \sigma_{d-1}\left(\mathbb{S}^{d-1}\right)$ and $\omega_d/2 = \sigma_{d-1}\left(\mathbb{S}_+^{d-1}\right)$ represents the normalization constant. Conditioning an IUR plane on a fixed direction yields a so-called Fixed orientation Uniformly Random (FUR) plane. Fix $\theta \in \mathbb{S}^{d-1}$, let $a = a(\theta)$ be the smallest number such that $K \cap T_{\theta,a} \neq \emptyset$, similarly let $b = b(\theta)$ be the largest number such that $K \cap T_{\theta,b} \neq \emptyset$. Then, conditional on this direction $\Theta = \theta$, S is uniformly distributed on the interval $[a, b]$, and we denote this conditional density by:

$$f_{S|\Theta}(s|\theta) = \begin{cases} \frac{1}{b(\theta)-a(\theta)} & \text{if } s \in [a(\theta), b(\theta)] \\ 0 & \text{otherwise,} \end{cases} \tag{2.6}$$

We may also write: $S|\Theta = \theta \sim \mathrm{Unif}(a(\theta), b(\theta))$. We also need the following lemma, which appears as proposition 1 in [25]:

Lemma 2.1. *Suppose that $K, L \in \mathcal{K}^d$ with $K \subset L$. Let T be an IUR plane hitting L, then:*

1. *Hitting probability:*
$$\mathbb{P}(T \cap K \neq \emptyset) = \frac{\bar{b}(K)}{\bar{b}(L)}.$$
2. *Conditional property: Given that T hits K, i.e. $T \cap K \neq \emptyset$, T is an IUR plane hitting K.*

We are now ready to introduce the CDF of interest in this chapter.

Definition 2.2 (section volume CDF)**.** Fix $K \in \mathcal{K}^d$, $d \geq 2$, let f_K be as in Definition 2.1. Let $(\Theta, S) \sim f_K$, the random variable $Z = \nu_{d-1}(K \cap T_{\Theta,S})$ has cumulative distribution function G_K which is given by:

$$G_K(z) = \int_{S_+^{d-1}}\int_{\mathbb{R}} \mathbb{1}\{\nu_{d-1}(K \cap T_{\theta,s}) \leq z\} f_K(\theta, s)\mathrm{d}s\mathrm{d}\theta.$$

We refer to G_K as the section volume CDF of K.

We remark that the expression for the CDF G_K follows from the fact that $G_K(z) = P(Z \leq z) = \mathbb{E}(\mathbb{1}\{Z \leq z\})$ and the law of the unconscious statistician. In $\mathbb{R}^2$ we may still refer to G_K as chord length distribution function and in $\mathbb{R}^3$ we may call it cross section area distribution function.

2.3 Properties of the section volume CDF

In this section we derive various properties of the section volume CDF as described in Definition 2.2. The following lemma highlights some basic properties.

Lemma 2.2. *Fix $K, L \in \mathcal{K}^d$, let G_K, G_L be their section volume CDF respectively. Let $z \in \mathbb{R}$, then:*

1. *Translation invariance: $G_{K+x}(z) = G_K(z)$ for all $x \in \mathbb{R}^d$.*
2. *Rotation invariance: $G_{MK}(z) = G_K(z)$ for all $M \in \mathrm{SO}(d)$.*
3. *Scaling: $G_{\lambda K}(z) = G_K\left(z/\lambda^{d-1}\right)$ for all $\lambda > 0$.*
4. *Inclusion: If $K \subset L$ then:*

$$G_L(z) \leq G_K(z)\frac{\bar{b}(K)}{\bar{b}(L)} + \left(1 - \frac{\bar{b}(K)}{\bar{b}(L)}\right).$$

The translation and rotation invariance of IUR planes is a defining property of IUR planes, [25], and it may be used to prove property 1 and 2 in Lemma 2.2. The third property also appears in [84] for $d = 3$. All of these properties are well-known for chord length distributions and the generalization to $(d-1)$-dimensional sections of convex bodies in $\mathbb{R}^d$ is not difficult. For the sake of completeness, the proof of this lemma may be found in section 2.6. We need Brunn's theorem (see for example [52]) to prove our first result:

Theorem 2.3 (Brunn)**.** *Let $K \in \mathcal{K}^d$, $d \geq 2$. Fix $\theta \in \mathbb{S}^{d-1}$. The function $f_\theta : \mathbb{R} \to [0, \infty)$ given by:*

$$f_\theta(s) = \nu_{d-1}(K \cap T_{\theta,s})^{\frac{1}{d-1}}$$

is concave on its support.

Ignoring the exponent $1/(d-1)$ in the definition of f_θ, this function returns the volume of the intersection of K with $T_{\theta,s}$. Because we fix θ this means the function considers volumes of parallel slices of K, and it is a function of the (signed) distance of the section plane to the origin. The statement of Brunn's Theorem inspires us to study a distribution function which is closely related to G_K:

Definition 2.4 (Transformed section volume CDF)**.** Fix $K \in \mathcal{K}^d$, $d \geq 2$, let f_K be as in Definition 2.1. Let $(\Theta, S) \sim f_K$, the random variable $Z = \nu_{d-1}(K \cap T_{\Theta,S})^{1/(d-1)}$ has cumulative distribution function G_K^S which is given by:

$$G_K^S(z) = \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \mathbb{1}\left\{\nu_{d-1}(K \cap T_{\theta,s})^{\frac{1}{d-1}} \leq z\right\} f_K(\theta, s)\mathrm{d}s\mathrm{d}\theta.$$

We refer to G_K^S as the transformed section volume CDF of K.

This distribution function G_K^S turns out to be more natural to study compared to G_K. This will become clear in the proof of the upcoming theorem. Note that, for $K \in \mathcal{K}^d$, G_K and G_K^S are related as follows:

$$G_K^S(z) = G_K\left(z^{d-1}\right).$$

Remark 2.1. *G_K is absolutely continuous if and only if G_K^S is absolutely continuous. After all, suppose that G_K^S has probability density function g_K^S. Let $X \sim g_K^S$, then $X^{d-1} \sim G_K$ and via the well-known change of variables formula this random variable has the following probability density function:*

$$g_K(z) = g_K^S\left(z^{\frac{1}{d-1}}\right)\frac{z^{\frac{2-d}{d-1}}}{d-1}. \tag{2.7}$$

The converse case is analogous.

In the proof of the upcoming theorem we use some additional notation. Given a convex body $K \in \mathcal{K}^d$ its inner section function $m_K : \mathbb{S}^{d-1} \to [0, \infty)$ is defined by:

$$m_K(\theta) = \max_{s\in\mathbb{R}} \nu_{d-1}\left(K \cap T_{\theta,s}\right). \tag{2.8}$$

This function returns the maximal section volume for any given direction. We now present one of the main theorems in this chapter:

Theorem 2.5. *Let $K \in \mathcal{K}^d$, $d \geq 2$. Define the function $f_\theta : \mathbb{R} \to [0, \infty)$ by:*

$$f_\theta(s) = \nu_{d-1}(K \cap T_{\theta,s})^{\frac{1}{d-1}}.$$

If f_θ has a unique maximum and is continuous on $\mathbb{R}$ for almost all $\theta \in \mathbb{S}_+^{d-1}$, then G_K is absolutely continuous with respect to Lebesgue measure.

Proof. Given $K \in \mathcal{K}^d$, let G_K be its section volume CDF and let G_K^S be its transformed section volume CDF. We show that G_K^S is absolutely continuous, from this it follows that G_K is absolutely continuous by Remark 2.1. By conditioning the distribution function G_K^S on Θ having a particular value, G_K^S may be written as a mixture distribution:

$$G_K^S(z) = \mathbb{P}\left(f_\Theta(S) \leq z\right) = \int_{\mathbb{S}_+^{d-1}} \mathbb{P}\left(f_\theta(S) \leq z \middle| \Theta = \theta\right) f_{K,\Theta}(\theta)\mathrm{d}\theta,$$

with $f_{K,\Theta}(\theta)$ being the marginal density of Θ as in (2.3) and $f_\theta(\cdot)$ as in the statement of the theorem. For notation convenience, write:

$$G_K^S(z|\theta) := \mathbb{P}\left(f_\theta(S) \leq z \middle| \Theta = \theta\right). \tag{2.9}$$

Let $a = a(\theta)$ and $b = b(\theta)$ be as in (2.6) such that $S|\Theta = \theta \sim \mathrm{Unif}(a, b)$. Choose $\theta \in \mathbb{S}_+^{d-1}$ such that f_θ has a unique maximum and is continuous on $\mathbb{R}$. By definition of a we know that $T_{\theta,a}$ intersects K only through the boundary of K. By the assumed continuity of f_θ

we have $f_\theta(a) = 0$ and similarly: $f_\theta(b) = 0$. Note that f_θ has the following domain and codomain:

$$f_\theta : [a, b] \to D_\theta, \text{ with } D_\theta = \left[0, m_K(\theta)^{\frac{1}{d-1}}\right], \tag{2.10}$$

and $m_K(\cdot)$ as in (2.8). By Brunn's theorem f_θ is concave on its support and by assumption it attains its maximum in a single point c. As a result, f_θ is strictly increasing on (a, c) and strictly decreasing on (c, b). Therefore, f_θ restricted to (a, c) is invertible, and its inverse is convex and strictly increasing. Let:

$$f_\theta^+ : \left(0, m_K(\theta)^{\frac{1}{d-1}}\right) \to (a, c),$$

denote this inverse. Similarly, f_θ restricted to (c, b) has an inverse:

$$f_\theta^- : \left(0, m_K(\theta)^{\frac{1}{d-1}}\right) \to (c, b),$$

which is concave and strictly decreasing. Write:

$$p := \mathbb{P}(S \in (a, c) | \Theta = \theta).$$

By using the fact that $S|\Theta = \theta \sim \text{Unif}(a, b)$ we find $p = (c - a)/(b - a)$. Moreover, we obtain the following expression for $G_K^S(z|\theta)$:

$$\begin{aligned} G_K^S(z|\theta) &= \mathbb{P}\left(f_\theta(S) \le z \middle| \Theta = \theta, S \in (a, c)\right) p + \mathbb{P}\left(f_\theta(S) \le z \middle| \Theta = \theta, S \in (c, b)\right)(1 - p) \\ &= \mathbb{P}\left(S \le f_\theta^+(z) \middle| \Theta = \theta, S \in (a, c)\right) p + \mathbb{P}\left(S \ge f_\theta^-(z) \middle| \Theta = \theta, S \in (c, b)\right)(1 - p) \\ &= \frac{f_\theta^+(z) - a}{c - a} p + \left(1 - \frac{f_\theta^-(z) - c}{b - c}\right)(1 - p). \end{aligned}$$

Because f_θ^- is concave and strictly decreasing, $-f_\theta^-$ is convex and strictly increasing. Therefore, $G_K^S(\cdot|\theta)$ is a convex combination of two functions both of which are convex and strictly increasing on the interval D_θ (as in (2.10)). As a result, $G_K^S(\cdot|\theta)$ is convex, continuous, and strictly increasing on D_θ, which is the support of this distribution function. We conclude that for almost all $\theta \in \mathbb{S}_+^{d-1}$, $G_K^S(\cdot|\theta)$ is absolutely continuous because it is convex on its support and continuous on $\mathbb{R}$. Finally, this means that G_K^S as a mixture of absolutely continuous distribution functions is absolutely continuous by Fubini's theorem. □

Remark 2.2. *The arguments used in the proof of Theorem 2.5 do not hold for general convex bodies. For general convex bodies the function f_θ is concave by Brunn's theorem. Therefore, the set of points at which it attains its maximum may be an interval rather than a single point. When this is the case, $G_K^S(\cdot|\theta)$ is still convex on its support, but it is discontinuous in the point $m_K(\theta)^{1/(d-1)}$, which is the right boundary point of its support. As a result, for any convex body $K \in \mathcal{K}^d$, G_K^S is convex on the interval $(0, \tau_K)$ with: $\tau_K = \min_{\theta \in \mathbb{S}^{d-1}} m_K(\theta)^{1/(d-1)}$.*

2.3.1 Strictly convex bodies

Let us now consider a particular class of convex bodies known as strictly convex bodies. The class of strictly convex bodies is large in a precise sense. For one, the class of convex bodies which are not smooth or strictly convex form a set of first Baire category, see [117] for details. We have not yet mentioned smooth convex bodies, loosely speaking it means that their boundary is smooth. An important result we obtain in this section is that given that $K \in \mathcal{K}^d$ is strictly convex, then G_K is absolutely continuous. Therefore, we show that for a large class of convex bodies G_K is absolutely continuous. The main tools to obtain this result are the famous Brunn-Minkowski inequality and a variant of Brunn's theorem. In the field of convex geometry the importance of the Brunn-Minkowski inequality cannot be overstated, we refer to the review paper [31] for variants of the theorem and its applications.

Theorem 2.6 (Brunn-Minkowski)**.** *Given convex bodies $K, L \in \mathcal{K}^d$ and $0 < \lambda < 1$ the following inequality holds:*

$$v_d(\lambda K + (1-\lambda)L)^{\frac{1}{n}} \geq \lambda\, v_d(K)^{\frac{1}{n}} + (1-\lambda)\, v_d(L)^{\frac{1}{n}},$$

with equality if and only if K and L are equal up to translation and dilatation.

The equality condition in Theorem 2.6 means that there exist $\delta > 0$ and $x \in \mathbb{R}^d$ such that $K = \delta L + x$. In order to prove that G_K is absolutely continuous for strictly convex $K \in \mathcal{K}^d$ we show that the conditions in Theorem 2.5 are satisfied. First, we need the following Lemma:

Lemma 2.3. *Let $K, L \in \mathcal{K}^d$ with $K \subset \operatorname{int} L$, then $v_d(K) < v_d(L)$.*

Its proof is given in section 2.6. We show that the strict convexity of a convex body carries over to strict concavity of the function f_θ (as in Theorem 2.3).

Theorem 2.7. *Let $K \in \mathcal{K}^d$ be a strictly convex body, $d \geq 2$. Fix $\theta \in \mathbb{S}^{d-1}$. The function $f_\theta : \mathbb{R} \to [0, \infty)$ given by:*

$$f_\theta(s) = v_{d-1}(K \cap T_{\theta,s})^{\frac{1}{d-1}}$$

is continuous on $\mathbb{R}$ and strictly concave on its support.

Proof. The proof is a slight variation of a proof of Brunn's theorem using the Brunn-Minkowski inequality as found in [52] (pp 18, 19). Fix $\theta \in \mathbb{S}^{d-1}$. Choose r, t in the support of f_θ, such that $r < t$. Let $\lambda \in (0,1)$, set $s = \lambda r + (1-\lambda)t$ and consider the hyperplane sections $K_r := K \cap T_{\theta,r}$, $K_s := K \cap T_{\theta,s}$ and $K_t := K \cap T_{\theta,t}$. We show that:

$$\lambda K_r + (1-\lambda)K_t \subset (\operatorname{int} K) \cap T_{\theta,s}. \tag{2.11}$$

Let $z \in \lambda K_r + (1-\lambda)K_t$, then $z = \lambda x + (1-\lambda)y$ for some $x \in K_r$ and some $y \in K_t$. Because also $x, y \in K$ we have $z \in \operatorname{int} K$ due to the strict convexity of K. Also, note that $\langle z, \theta \rangle = \lambda \langle x, \theta \rangle + (1-\lambda)\langle y, \theta \rangle = \lambda r + (1-\lambda)t = s$. Hence, $z \in T_{\theta,s}$, which proves (2.11).

It can readily be verified that: $(\operatorname{int} K) \cap T_{\theta,s} = \operatorname{relint}(K_s)$. Combining this with (2.11) we find: $\lambda K_r + (1-\lambda)K_t \subset \operatorname{relint}(K_s)$. Let $\Pi(L)$ denote the orthogonal projection of L on the hyperplane $T_{\theta,0}$. Note that $\lambda K_r + (1-\lambda)K_t$ and K_s are subsets of $T_{\theta,s}$, projecting them on $T_{\theta,0}$ preserves the inclusion:

$$\Pi(\lambda K_r + (1-\lambda)K_t) \subset \operatorname{relint} \Pi(K_s).$$

Identifying $T_{\theta,0}$ with $\mathbb{R}^{d-1}$ we may regard $\Pi(\lambda K_r + (1-\lambda)K_t)$ and $\Pi(K_s)$ as convex bodies in $\mathbb{R}^{d-1}$. Under this identification $\Pi(\lambda K_r + (1-\lambda)K_t) \subset \operatorname{int} \Pi(K_s)$ and applying Lemma 2.3 yields:

$$\nu_{d-1}\left(\Pi(\lambda K_r + (1-\lambda)K_t)\right) < \nu_{d-1}\left(\Pi(K_s)\right).$$

Keep in mind that the sets which are projected on $T_{\theta,0}$ have the same $(d-1)$-dimensional volume as the original sets. Note that we may change the order of these projections and the Minkowski sum:

$$\lambda\Pi(K_r) + (1-\lambda)\Pi(K_t) = \Pi(\lambda K_r + (1-\lambda)K_t),$$

where the sum of sets is considered in the plane $T_{\theta,0}$. Hence,

$$\nu_{d-1}\left(\Pi(K_s)\right) > \nu_{d-1}(\Pi(\lambda K_r + (1-\lambda)K_t)) = \nu_{d-1}(\lambda\Pi(K_r) + (1-\lambda)\Pi(K_t)). \quad (2.12)$$

Once again, $\Pi(K_r)$ and $\Pi(K_t)$ may be identified as convex bodies in $\mathbb{R}^{d-1}$ and we may apply Brunn-Minkowski's (B.M.) inequality to obtain the desired result:

$$\begin{aligned}
f_\theta(s) &= \nu_{d-1}(K_s)^{\frac{1}{d-1}} \\
&= \nu_{d-1}(\Pi(K_s))^{\frac{1}{d-1}} \\
&> \nu_{d-1}(\lambda\Pi(K_r) + (1-\lambda)\Pi(K_t))^{\frac{1}{d-1}} && \text{(by equation (2.12))} \\
&\geq \lambda\, \nu_{d-1}(\Pi(K_r))^{\frac{1}{d-1}} + (1-\lambda)\, \nu_{d-1}(\Pi(K_t))^{\frac{1}{d-1}} && \text{(B.M.)} \\
&= \lambda\, \nu_{d-1}(K_r)^{\frac{1}{d-1}} + (1-\lambda)\, \nu_{d-1}(K_t)^{\frac{1}{d-1}} \\
&= \lambda f_\theta(r) + (1-\lambda) f_\theta(t).
\end{aligned}$$

Continuity of $f_\theta(\cdot)$ can be shown as follows. Let $a = a(\theta)$ and $b = b(\theta)$ be as in (2.6). By definition of a we know that $T_{\theta,a}$ intersects K only through the boundary of K. This intersection only contains a single point, if another point were in the intersection this would imply that the boundary of K contains a line segment which contradicts the strict convexity of K. As a result: $f_\theta(a) = 0$ and similarly: $f_\theta(b) = 0$. Because a and b are the only possible points of discontinuity, $f_\theta(\cdot)$ is continuous. □

Because a bounded concave function has a maximum, strict concavity then implies that the maximum is unique. We obtain as a direct consequence of Theorem 2.1 and Theorem 2.5:

Corollary 2.1. *Let $K \in \mathcal{K}^d$ strictly convex, and let G_K be its section volume CDF. Then, G_K is absolutely continuous.*

2.3.2 Polytopes

In this section we study polytopes, which are especially of interest for practical applications. Being examples of non-strictly convex bodies, they are not covered by Corollary 2.1. The main result we obtain in this section is that the section volume CDF of a full-dimensional convex polytope is absolutely continuous. In order to obtain this result for polytopes, we need to deal with the regions where the function f_θ, as in Brunn's theorem, is constant. The following lemma shows that this can only happen if the polytope has parallel edges.

Lemma 2.4. *Let $P \subset \mathbb{R}^d$ be a full-dimensional convex polytope, $d \geq 2$. Fix $\theta \in \mathbb{S}_+^{d-1}$ and define the function $f_\theta : \mathbb{R} \to [0, \infty)$ by:*

$$f_\theta(s) = \nu_{d-1}(P \cap T_{\theta,s})^{\frac{1}{d-1}}.$$

Suppose f_θ attains its maximum on the entire interval $[s_-, s_+]$, with $s_- < s_+$. Then, any plane $T_{\theta,s}$ with $s \in [s_-, s_+]$ intersects the same edges of P and these edges are parallel.

Proof. Let $a = a(\theta)$ and $b = b(\theta)$ as in (2.6). For $s \in (a, b)$ the intersection $P \cap T_{\theta,s}$ is a $(d-1)$-dimensional polytope, and its vertices are the intersections of $T_{\theta,s}$ with the edges of P. By Brunn's theorem we know that f_θ is concave on its support. The set of points at which a concave function attains its maximum is convex, hence it is a nondegenerate interval or a single point. By assumption it is the interval $[s_-, s_+]$. By Brunn-Minkowski's inequality, and in particular its equality condition, we know that all sections $\{P \cap T_{\theta,s} : s \in [s_-, s_+]\}$ are equal up to dilatation and translation. But, because all such sections have equal volume, these sections then have to be equal up to translations. Write $P_{s_-} = P \cap T_{\theta,s_-}$ and $P_{s_+} = P \cap T_{\theta,s_+}$. Because P_{s_+} is equal to P_{s_-} up to translation there exists a $x \in \mathbb{R}^d$ such that $P_{s_+} = P_{s_-} + x$. Let $s \in [s_-, s_+]$, we claim that:

$$P \cap T_{\theta,s} = P_{s_-} + \frac{s - s_-}{s_+ - s_-} x =: Q(s). \tag{2.13}$$

Let $z \in Q(s)$, then there exists a $y \in P_{s_-}$ such that:

$$z = y + \frac{s - s_-}{s_+ - s_-} x = \frac{s_+ - s}{s_+ - s_-} y + \left(1 - \frac{s_+ - s}{s_+ - s_-}\right)(x + y).$$

Since $y \in P_{s_-}$ and $(x + y) \in P_{s_-} + x = P_{s_+}$, it follows that z is the convex combination of two points in P, hence $z \in P$. Moreover, we have $\langle y, \theta \rangle = s_-$ and $\langle x + y, \theta \rangle = s_+$. A direct computation yields: $\langle z, \theta \rangle = s$. This means that $Q(s) \subset P \cap T_{\theta,s}$. Because $Q(s)$ is a translation of P_{s_-} and since $P \cap T_{\theta,s}$ is equal to P_{s_-} up to a translation we necessarily have that (2.13) holds. Therefore, for any vertex v of P_{s_-}, $v + ((s - s_-)/(s_+ - s_-))x$ is a vertex of $P \cap T_{\theta,s}$. It is evident that all vertices of the polytopes $\{P \cap T_{\theta,s} : s \in [s_-, s_+]\}$ lie on parallel line segments which are subsets of the edges of P, this finishes the proof. □

In the next theorem we combine some of the techniques used earlier in this chapter and Lemma 2.4 to show that the section volume CDF of any full-dimensional convex polytope is absolutely continuous.

Theorem 2.8. *Let $P \subset \mathbb{R}^d$ be a full-dimensional convex polytope, $d \geq 2$. Let G_P be its section volume CDF. Then, G_P is absolutely continuous.*

Proof. Given $\theta \in \mathbb{S}_+^{d-1}$ define the function $f_\theta : \mathbb{R} \to [0, \infty)$ by:

$$f_\theta(s) = \nu_{d-1}(P \cap T_{\theta,s})^{\frac{1}{d-1}}.$$

Let $B \subset \mathbb{R}$ be a Borel set of Lebesgue measure zero. Let f_P be as in Definition 2.2 and let $(\Theta, S) \sim f_P$. As in the proof of Theorem 2.5, we condition on $\Theta = \theta$ and write:

$$\mathbb{P}\Big(\nu_{d-1}(P \cap T_{\Theta,S})^{\frac{1}{d-1}} \in B \Big) = \mathbb{P}\left(f_\Theta(S) \in B\right) = \int_{S_+^{d-1}} \mathbb{P}\left(f_\theta(S) \in B \middle| \Theta = \theta\right) f_{P,\Theta}(\theta) \mathrm{d}\theta,$$

with $f_{P,\Theta}(\theta)$ being the marginal density of Θ as in (2.3). In order to show that G_P is absolutely continuous it is sufficient to show that $\mathbb{P}\left(f_\Theta(S) \in B\right) = 0$. Let $a = a(\theta)$ and $b = b(\theta)$ be as in (2.6) such that $S|\Theta = \theta \sim \mathrm{Unif}(a, b)$. Note that f_θ is continuous on $\mathbb{R}$ for almost all $\theta \in \mathbb{S}_+^{d-1}$. For almost all $\theta \in \mathbb{S}_+^{d-1}$ the section planes $T_{\theta,s}$ enter the polytope through a vertex as s runs from $a(\theta)$ to $b(\theta)$. For any such θ, $f_\theta(a) = 0$ and $f_\theta(b) = 0$, because a vertex has no $(d-1)$-dimensional volume. As a and b are the only possible points of discontinuity, f_θ is continuous on $\mathbb{R}$ for almost all $\theta \in \mathbb{S}_+^{d-1}$.

By Brunn's theorem we know that f_θ is concave on its support. The set of points at which a concave function attains its maximum is convex, hence it is a nondegenerate interval or a single point. Denote this set by: $[s_-(\theta), s_+(\theta)]$, in the case it consists of a single point, $s_-(\theta) = s_+(\theta)$. Write: $p_1 = \mathbb{P}(S \in (a, s_-)|\Theta = \theta)$, $p_2 = \mathbb{P}(S \in [s_-, s_+]|\Theta = \theta)$ and $p_3 = \mathbb{P}(S \in (s_+, b)|\Theta = \theta)$. We may write:

$$\begin{aligned}\mathbb{P}(f_\theta(S) \in B|\Theta = \theta) = \mathbb{P}(f_\theta(S) \in B|\Theta = \theta, S \in (a, s_-))p_1 + \\ + \mathbb{P}(f_\theta(S) \in B|\Theta = \theta, S \in [s_-, s_+])p_2 + \\ + \mathbb{P}(f_\theta(S) \in B|\Theta = \theta, S \in (s_+, b))p_3.\end{aligned}$$

Arguing as in the proof of Theorem 2.5 we obtain that for almost all $\theta \in \mathbb{S}_+^{d-1}$ the distribution functions $z \mapsto \mathbb{P}(f_\theta(S) \leq z|\Theta = \theta, S \in (a, s_-))$ and $z \mapsto \mathbb{P}(f_\theta(S) \leq z|\Theta = \theta, S \in (s_+, b))$ are continuous and convex on their support and therefore absolutely continuous with respect to Lebesgue measure. Hence, for any such θ we have $\mathbb{P}(f_\theta(S) \in B|\Theta = \theta, S \in (a, s_-)) = 0$ and $\mathbb{P}(f_\theta(S) \in B|\Theta = \theta, S \in (s_+, b)) = 0$. Clearly, for any $\theta \in \mathbb{S}_+^{d-1}$: $p_2(\theta) = (s_+(\theta) - s_-(\theta))/(b(\theta) - a(\theta))$. Further note that $\mathbb{P}(f_\theta(S) \in B|\Theta = \theta, S \in [s_-, s_+]) = \mathbb{1}\{m_P(\theta)^{1/(d-1)} \in B\}$, with $m_P(\cdot)$ is as in (2.8) and $\theta \in \mathbb{S}_+^{d-1}$. Combining these results we may therefore write:

$$\mathbb{P}(f_\Theta(S) \in B) = \int_{\mathbb{S}_+^{d-1}} \mathbb{1}\left\{m_P(\theta)^{\frac{1}{d-1}} \in B\right\} \frac{s_+(\theta) - s_-(\theta)}{b(\theta) - a(\theta)} f_{P,\Theta}(\theta) \mathrm{d}\theta. \tag{2.14}$$

In (2.14) we effectively only integrate over θ such that $s_+(\theta) > s_-(\theta)$. By Lemma 2.4 this strict inequality only holds if for all $s \in [s_-(\theta), s_+(\theta)]$ the same edges of P are intersected by $T_{\theta,s}$ and these edges are all parallel. Define:

$$D = \left\{\theta \in \mathbb{S}_+^{d-1} : s_+(\theta) > s_-(\theta)\right\}.$$

Hence, for any $\theta \in D$ and any $s \in [s_-(\theta), s_+(\theta)]$, we have $m_P(\theta) = \nu_{d-1}(P \cap T_{\theta,s})$ and the plane $T_{\theta,s}$ only intersects edges of P which are parallel. If $\sigma_{d-1}(D) = 0$, for example because P does not have any pair of parallel edges, then $\mathbb{P}(f_\Theta(S) \in B) = 0$, hence G_P is absolutely continuous.

Let us now consider the case $\sigma_{d-1}(D) > 0$. We may write D as a disjoint union $D = \cup_{i=1}^k D_i$ for some $k \in \mathbb{N}$. Here D_i is defined such that for all $\theta \in D_i$ all planes corresponding to $m_P(\theta)$ intersect the same parallel edges. Let $i \in \{1, \dots, k\}$ and let $e_1, \dots, e_m \subset P$ be the parallel edges of P corresponding to D_i. Consider the plane $T_{\phi_i,0}$, with $\phi_i \in \mathbb{S}_+^{d-1}$ such that this plane is orthogonal to the edges $e_1, \dots, e_m$. For any $L \subset \mathbb{R}^d$ let $\Pi_i(L)$ denote the orthogonal projection of L on the hyperplane $T_{\phi_i,0}$. Let $\theta \in D_i$, $s \in [s_-(\theta), s_+(\theta)]$ such that $m_P(\theta) = \nu_{d-1}(P \cap T_{\theta,s})$ and the plane $T_{\theta,s}$ intersects $e_1, \dots, e_m$. Note that for any $\theta \in D_i$ and $s \in [s_-(\theta), s_+(\theta)]$, $v_i := \nu_{d-1}(\Pi_i(P \cap T_{\theta,s}))$ attains the same value. After all, for any such plane, $\Pi_i(P \cap T_{\theta,s})$ is a polytope in $T_{\phi_i,0}$ and its vertices are given by the orthogonal projections of $e_1, \dots, e_m$ on $T_{\phi_i,0}$. Moreover, it is well known that the volume of $P \cap T_{\theta,s}$ and the volume of its projection on $T_{\phi_i,0}$ are related via:

$$v_i = \nu_{d-1}(\Pi_i(P \cap T_{\theta,s})) = |\langle \theta, \phi_i \rangle| \, \nu_{d-1}(P \cap T_{\theta,s}).$$

Hence,

$$m_P(\theta) = \nu_{d-1}(P \cap T_{\theta,s}) = \frac{v_i}{|\langle \theta, \phi_i \rangle|}, \quad \theta \in D_i. \tag{2.15}$$

If we were to draw $\xi \sim \mathrm{Unif}(\mathbb{S}^{d-1})$, then the Lebesgue density of the random variable $\langle \xi, \phi_i \rangle$ is given by:

$$t \in [-1, 1] \mapsto \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})}(1 - t^2)^{\frac{d-3}{2}}.$$

This density does not depend on ϕ_i due to symmetry. Because the probability measure corresponding to the uniform distribution on the sphere is the normalized spherical measure, we obtain $\sigma_{d-1}(\{\theta \in \mathbb{S}^{d-1} : \langle \theta, \phi_i \rangle \in B\}) = 0$. Via the change of variables formula it is easily verified that the random variable $(v_i/|\langle \xi, \phi_i \rangle|)^{1/(d-1)}$ also has a Lebesgue density. Therefore:

$$\sigma_{d-1}\left(\left\{\theta \in \mathbb{S}^{d-1} : (v_i/|\langle \theta, \phi_i \rangle|)^{\frac{1}{d-1}} \in B\right\}\right) = 0. \tag{2.16}$$

Finally, from its definition it is evident that the density $f_{P,\Theta}$ is bounded, see (2.3). Let $M > 0$ be an upper bound of this density. Using this fact and (2.15) and (2.16), the claim follows:

$$\begin{aligned}
\mathbb{P}(f_\Theta(S) \in B) &= \int_{\mathbb{S}_+^{d-1}} \mathbb{1}\left\{m_P(\theta)^{\frac{1}{d-1}} \in B\right\} \frac{s_+(\theta) - s_-(\theta)}{b(\theta) - a(\theta)} f_{P,\Theta}(\theta) \mathrm{d}\theta \\
&\leq \int_D \mathbb{1}\left\{m_P(\theta)^{\frac{1}{d-1}} \in B\right\} f_{P,\Theta}(\theta) \mathrm{d}\theta \\
&\leq M \sum_{i=1}^k \int_{D_i} \mathbb{1}\left\{m_P(\theta)^{\frac{1}{d-1}} \in B\right\} \mathrm{d}\theta \\
&= M \sum_{i=1}^k \int_{D_i} \mathbb{1}\left\{(v_i/|\langle \theta, \phi_i \rangle|)^{\frac{1}{d-1}} \in B\right\} \mathrm{d}\theta
\end{aligned}$$

$$\leq M \sum_{i=1}^{k} \sigma_{d-1}\left(\left\{\theta \in \mathbb{S}^{d-1} : (v_i/|\langle \theta, \phi_i \rangle|)^{\frac{1}{d-1}} \in B\right\}\right)$$
$$= 0.$$

□

2.4 Density approximation

In this section we consider convex bodies $K \in \mathcal{K}^d$ such that G_K is absolutely continuous. For most convex bodies $K \in \mathcal{K}^d$ there is no known explicit expression for G_K or its density g_K. In this section we focus on approximating the density g_K. This is achieved by obtaining a large sample from G_K along with a kernel density estimator (KDE). We use the following rejection sampling scheme, proposed in [65], to sample from the distribution G_K:

1. Enclose K inside a sphere: choose $R > 0$ such that $K \subset \bar{B}(0, R)$.
2. Choose an isotropic random direction $\Theta \sim \text{Unif}(\mathbb{S}^{d-1})$.
3. Sample $S \sim \text{Unif}(0, R)$.
4. The plane $T_{\Theta,S}$ hits $\bar{B}(0, R)$, if the plane also hits K we accept, and $Z = \nu_{d-1}(K \cap T_{\Theta,S})$ is a draw from G_K. If the plane does not hit K, we reject and go back to step 2.

In $\mathbb{R}^2$, step 2 may be achieved by sampling $\Phi \sim \text{Unif}(0, 2\pi)$ followed by setting $\Theta = (\cos(\Phi), \sin(\Phi))$. In $\mathbb{R}^3$ step 2 can be performed as follows. Sample $\Phi \sim \text{Unif}(0, 2\pi)$ and $X \sim \text{Unif}(-1, 1)$. Then, we may set $\Omega = \arccos(X)$ and $\Theta = (\sin(\Omega)\cos(\Phi), \sin(\Omega)\sin(\Phi), \cos(\Omega))$. In order to keep the rejection rate in the sampling scheme low, R should be as small as possible and K should be positioned at the origin, meaning that $0 \in \text{int}\, K$.

Of course, for $K \in \mathcal{K}^d$ with $d = 2$ we find: $G_K = G_K^S$. Note that G_K^S is initially convex on some initial interval (see Remark 2.2). As a result, its density g_K^S is non-decreasing on this interval. This means that g_K^S may even be constant initially. In addition, note that if $d = 2$ and K is a polygon then in [32] it has been shown that g_K^S is always constant on some initial interval. Because of the relation between g_K^S and g_K, if g_K^S is constant on an initial interval, then g_K behaves like $z^{(2-d)/(d-1)}$ on this interval. Hence, when $d = 3$ this means that g_K behaves like $1/\sqrt{z}$ for z close to zero. Clearly, this complicates the approximation of g_K near zero. Therefore, we choose to approximate the density g_K^S instead, and use (2.7) to obtain an approximation of g_K.

We will now introduce the Monte Carlo simulation scheme for approximating g_K^S. We choose a large $N \in \mathbb{N}$ and sample $Z_1, \ldots, Z_N \overset{\text{iid}}{\sim} G_K$ using the sampling scheme given above. Setting $X_i = Z_i^{1/(d-1)}$, we obtain that $X_1, \ldots, X_N \overset{\text{iid}}{\sim} G_K^S$. The following KDE is for example studied in [89], which we propose as an approximation for g_K^S:

$$\hat{g}_N^S(z) = \frac{1}{hN} \sum_{i=1}^{N} k\left(\frac{z - X_i}{h}\right) + k\left(\frac{z + X_i}{h}\right), \qquad z \geq 0, \tag{2.17}$$

with $h > 0$ the bandwidth parameter and k a symmetric kernel. The KDE in (2.17) is also known as the reflection method. A reason for using the reflection method over the classical (Parzen-Rosenblatt) KDE is that it ensures that no probability mass is assigned for $z < 0$. Recall that the classical KDE for the sample $X_1, \ldots, X_N$ is given by:

$$f(z) = \frac{1}{hN} \sum_{i=1}^{N} k\left(\frac{z - X_i}{h}\right), \qquad z \in \mathbb{R}. \tag{2.18}$$

Note that when computing the KDE in (2.18) for the following 'sample' of size $2N$:

$$X_1, X_2, \ldots, X_N, -X_1, -X_2, \ldots, -X_N,$$

we find: $f(z) = \hat{g}_N^S(z)/2$. This fact may be used to choose the bandwidth h, since most of the literature is devoted to bandwidth selection for the classical KDE. In the data examples in the next section we choose for k the Gaussian kernel and select the bandwidth with the popular Sheather-Jones method [97]. Whenever we want to approximate g_K instead of g_K^S, we simply follow the procedure given above to compute $\hat{g}_N^S$. Then, using (2.7) we set:

$$\hat{g}_N(z) = \hat{g}_N^S\left(z^{\frac{1}{d-1}}\right) \frac{z^{\frac{2-d}{d-1}}}{d-1}, \tag{2.19}$$

which is an approximation of g_K. A drawback of the KDE in (2.17) is that this density has (right)-derivative zero in $z = 0$. As mentioned before, when $d = 2$ and K is a convex polygon this is not an issue since the density g_K^S is then initially constant. In the data examples in the next section the approximations of G_P^S of some convex polytopes P in $\mathbb{R}^3$ appear initially (close to) linear. This suggests that the choice of boundary correction is reasonable. Should one consider a polytope P such that G_P^S is far from being initially linear then other boundary correction methods may be more appropriate.

2.4.1 Simulations

In this section we perform a few simulations to show that the Monte Carlo simulation scheme works well. For these simulations we focus on polytopes. Throughout this section, let $P \subset \mathbb{R}^d$ be a full-dimensional polytope. We have implemented the sampling scheme for drawing samples from G_P specifically for $d = 2$ and $d = 3$. The code used for the simulations may be found at https://github.com/thomasvdj/pysizeunfolder. The polytope can be entered into this program either by presenting a set of points, such that the polytope is given by the convex hull of these points, or by presenting a half-space representation of the polytope.

In the literature, similar simulations have been performed, e.g. for the cube and the dodecahedron. Therefore we also consider these shapes, such that we have a point of comparison. Besides approximating g_P and g_P^S we also approximate G_P^S. The distribution function G_P^S can be approximated arbitrarily closely by an empirical distribution function, given a large sample from G_P^S.

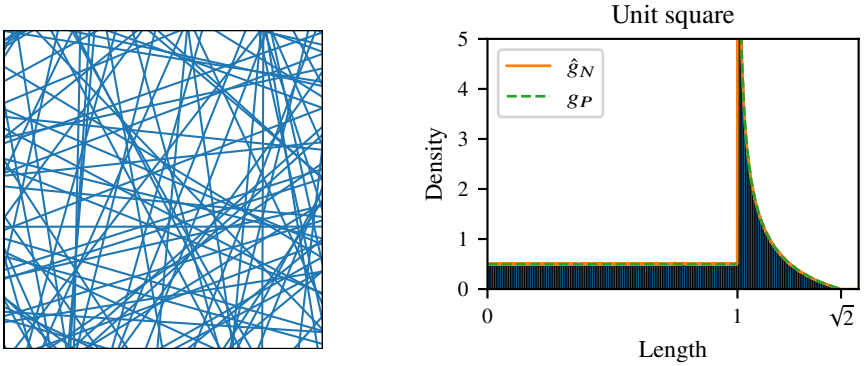


Figure 2.2: Left: 100 IUR sections through P the unit square in $\mathbb{R}^2$. Right: comparison of the density g_P to its approximation $\hat{g}_N$.

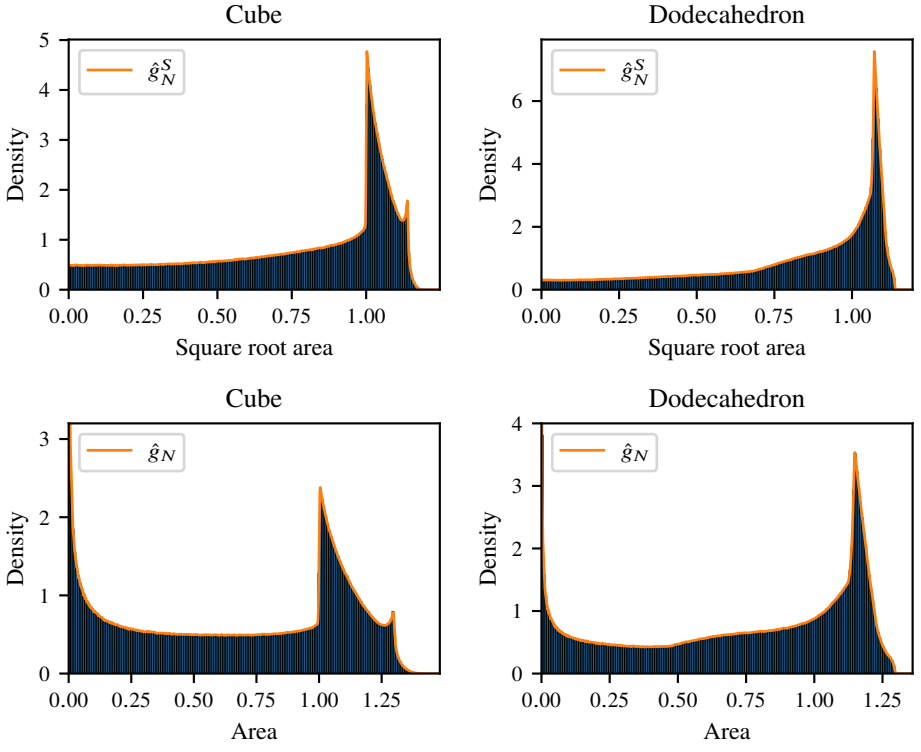


Figure 2.3: Approximations of g_P^S and g_P for P the unit cube (Left) and for P the dodecahedron with volume 1 (Right).

For all simulations, we set $N = 10^7$. For the first example, we choose the unit square in $\mathbb{R}^2$. The density of its chord length distribution may be found in [19], it is given by:

$$g_P(z) = \begin{cases} \frac{1}{2} & \text{if } 0 \leq z \leq 1 \\ \frac{1}{z^2\sqrt{z^2-1}} - \frac{1}{2} & \text{if } 1 < z \leq \sqrt{2} \end{cases}.$$

The approximation obtained via the proposed Monte Carlo scheme is shown in Figure 2.2. Figure 2.2 also contains a visualization of 100 IUR sections through the unit square. As can be seen in Figure 2.2, the approximation $\hat{g}_N$ is very close to the true probability density g_P.

We should stress that the proposed method is especially useful in the spatial setting $d = 3$. Naturally, whenever the analytical expression for g_P is available this is preferable. To the best of our knowledge, there are no known expressions for g_P of any polytope P in $\mathbb{R}^3$. In the planar case ($d = 2$) the density g_P is known for various polygons, for example for rectangles [19], and regular polygons [44]. In Figure 2.3 the approximations of g_P and g_P^S are shown for the cube and the dodecahedron, both shapes scaled to have volume 1.

Similar simulations were performed in [75] for the cube and dodecahedron, qualitatively the curves visualized there are close to the approximations of g_P shown in Figure 2.3. For the cube, one can easily see that for any direction $\theta \in \mathbb{S}_+^{d-1}$, there exists a section of area 1. By Remark 2.2, this means that g_P^S is non-decreasing on the interval $(0, 1)$, which can also be seen in Figure 2.3. Approximations of G_P^S for the cube and dodecahedron are shown in Figure 2.4. For these visualizations the same samples are used as in Figure 2.3. As mentioned before, these approximations of G_P^S appear initially (close to) linear, justifying the choice of boundary correction in the density approximation procedure.

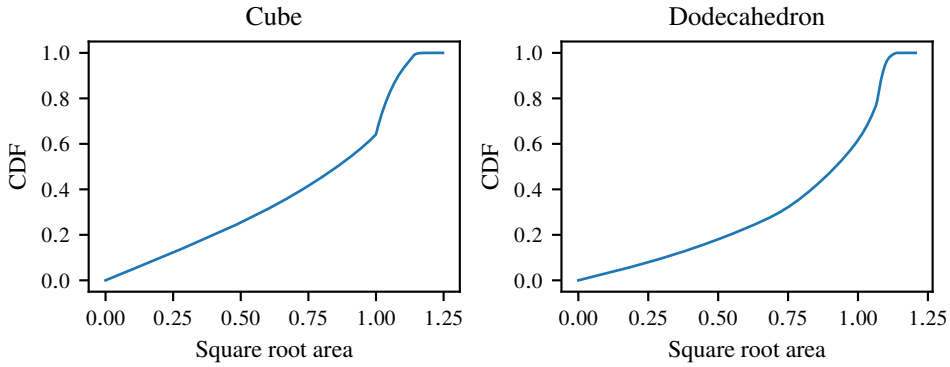


Figure 2.4: Approximations of G_P^S for P the unit cube (Left) and for P the dodecahedron with volume 1 (Right).

2.5 Concluding remarks

In this chapter we established absolute continuity of the (transformed) section volume CDF for various classes of convex bodies. As we will see in the next chapter, absolute continuity of these distribution functions is essential for stereological estimation of particle size distributions using likelihood-based inference. Whether these distribution functions are absolutely continuous for all convex bodies remains an open problem. From a theoretical perspective we cover a large class of convex bodies with the strictly convex bodies. With polytopes we cover a class of convex bodies which is especially important in practical applications. Moreover, for polytopes we provide a Monte Carlo simulation scheme for approximating the density corresponding to its (transformed) section volume CDF.

2.6 Additional proofs

Proof of Lemma 2.2. Let $x \in \mathbb{R}^d$, $\theta \in \mathbb{S}_+^{d-1}$ and $s \in \mathbb{R}$. It can be easily verified that the following holds:

$$(K + x) \cap T_{\theta,s} = \big(K \cap T_{\theta,s-\langle x,\theta\rangle}\big) + x. \tag{2.20}$$

Meaning that the intersection of a translated K with a plane is the same as the intersection of K with a translated plane and then translating the result. It follows that:

$$\begin{aligned}
G_{K+x}(z) &= \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}((K+x)\cap T_{\theta,s}) \le z\}\mathbb{1}\{(K+x)\cap T_{\theta,s} \neq \emptyset\}}{\mu([K+x])} \mathrm{d}s\mathrm{d}\theta \\
&\overset{(2.20)}{=} \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}((K\cap T_{\theta,s-\langle x,\theta\rangle})+x) \le z\}\mathbb{1}\{K\cap T_{\theta,s-\langle x,\theta\rangle} \neq \emptyset\}}{\mu([K+x])} \mathrm{d}s\mathrm{d}\theta \\
&= \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}(K\cap T_{\theta,s-\langle x,\theta\rangle}) \le z\}\mathbb{1}\{K\cap T_{\theta,s-\langle x,\theta\rangle} \neq \emptyset\}}{\mu([K+x])} \mathrm{d}s\mathrm{d}\theta \qquad (2.21) \\
&= \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}(K\cap T_{\theta,t}) \le z\}\mathbb{1}\{K\cap T_{\theta,t} \neq \emptyset\}}{\mu([K+x])} \mathrm{d}t\mathrm{d}\theta
\end{aligned}$$

In (2.21) we use the translation invariance of the Lebesgue measure. The final step is obtained by substituting $t = s - \langle x, \theta\rangle$. Via the same substitution it can be shown that $\mu([K + x]) = \mu([K])$. As a result we obtain: $G_{K+x}(z) = G_K(z)$. Moving on to the rotation invariance, let $M \in \mathrm{SO}(d)$, then the following can be shown:

$$MK \cap T_{\theta,s} = M(K \cap T_{M^T\theta,s}). \tag{2.22}$$

Using this, we find:

$$\begin{aligned}
G_{MK}(z) &= \frac{1}{2}\int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}(MK\cap T_{\theta,s}) \le z\}\mathbb{1}\{MK\cap T_{\theta,s} \neq \emptyset\}}{\mu([MK])} \mathrm{d}s\mathrm{d}\theta \qquad (2.23) \\
&\overset{(2.22)}{=} \frac{1}{2}\int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}(M(K\cap T_{M^T\theta,s})) \le z\}\mathbb{1}\{M(K\cap T_{M^T\theta,s}) \neq \emptyset\}}{\mu([MK])} \mathrm{d}s\mathrm{d}\theta \\
&= \frac{1}{2}\int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}(K\cap T_{M^T\theta,s}) \le z\}\mathbb{1}\{K\cap T_{M^T\theta,s} \neq \emptyset\}}{\mu([MK])} \mathrm{d}s\mathrm{d}\theta \qquad (2.24)
\end{aligned}$$

$$= \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\{\nu_{d-1}(K \cap T_{u,s}) \leq z\}\mathbb{1}\{K \cap T_{u,s} \neq \emptyset\}}{\mu([MK])} \mathrm{d}s\mathrm{d}u. \tag{2.25}$$

In (2.23) we use the fact that the inner integral does not change if we replace θ with $-\theta$, therefore we may integrate over $\mathbb{S}^{d-1}$ instead and divide the result by two. In (2.24) we use the rotation invariance of the Lebesgue measure. In (2.25) the substitution: $u = M^T\theta$ is applied. Because M is an orthogonal matrix of determinant one, the Jacobian corresponding to the transformation has determinant one. Since $M^T\mathbb{S}^{d-1} = \mathbb{S}^{d-1}$, the transformation does not affect the integration region. Via the same substitutions it can be shown that $\mu([MK]) = \mu([K])$ such that indeed $G_{MK}(z) = G_K(z)$. Next, we consider scaling of convex bodies. Let $\lambda > 0$, we remark that the following holds:

$$\lambda K \cap T_{\theta,s} = \lambda \left(K \cap T_{\theta,\frac{s}{\lambda}} \right). \tag{2.26}$$

Using this and the fact that $\nu_{d-1}(\lambda K) = \lambda^{d-1}\,\nu_{d-1}(K)$ for $K \in \mathcal{K}^{d-1}$, it is once again a matter of applying a substitution to obtain:

$$G_{\lambda K}(z) = \lambda \int_{\mathbb{S}_+^{d-1}} \int_{\mathbb{R}} \frac{\mathbb{1}\left\{\nu_{d-1}(K \cap T_{\theta,t}) \leq \frac{z}{\lambda^{d-1}}\right\}\mathbb{1}\{K \cap T_{\theta,t} \neq \emptyset\}}{\mu([\lambda K])} \mathrm{d}t\mathrm{d}\theta.$$

And similarly, via substitution we find: $\mu([\lambda K]) = \lambda\mu([K])$ such that indeed: $G_{\lambda K}(z) = G_K(z/\lambda^{d-1})$. We now consider the final statement of the lemma. Let T be an IUR plane hitting L. By Lemma 2.1, the probability that T hits K is given by $\bar{b}(K)/\bar{b}(L)$. Moreover, given that T hits K it is an IUR plane hitting K. It follows that:

$$\begin{aligned}
G_L(z) &= \mathbb{P}(\nu_{d-1}(L \cap T) \leq z) \\
&= \mathbb{P}(\nu_{d-1}(L \cap T) \leq z | K \cap T \neq \emptyset)\frac{\bar{b}(K)}{\bar{b}(L)} + \\
&\quad + \mathbb{P}(\nu_{d-1}(L \cap T) \leq z | K \cap T = \emptyset)\left(1 - \frac{\bar{b}(K)}{\bar{b}(L)}\right) \\
&\leq \mathbb{P}(\nu_{d-1}(K \cap T) \leq z | K \cap T \neq \emptyset)\frac{\bar{b}(K)}{\bar{b}(L)} + \left(1 - \frac{\bar{b}(K)}{\bar{b}(L)}\right) \\
&= G_K(z)\frac{\bar{b}(K)}{\bar{b}(L)} + \left(1 - \frac{\bar{b}(K)}{\bar{b}(L)}\right).
\end{aligned}$$

□

Proof of Lemma 2.3. Let $x \in \partial K$. Since $x \in \operatorname{int} L$ there exists a $R > 0$ such that $B(x, R) \subset \operatorname{int} L$. Because $x \in \partial K$ we know that $B(x, R) \cap (\mathbb{R}^d \setminus K) \neq \emptyset$. Choose $y \in B(x, R) \cap (\mathbb{R}^d \setminus K)$. Note that $y \in \operatorname{int} L$ and $\operatorname{int} L$ is open. Choose $r_1 > 0$ such that $B(y, r_1) \subset \operatorname{int} L$. Because K is closed, $\mathbb{R}^d \setminus K$ is open. Choose $r_2 > 0$ such that $B(y, r_2) \subset \mathbb{R}^d \setminus K$. Let $r = \min\{r_1, r_2\}$, then $B(y, r) \subset (\operatorname{int} L) \setminus K$ and this ball has a strictly positive volume. Hence, we find:

$$\nu_d(L) = \nu_d(L \setminus K) + \nu_d(K) \geq \nu_d(B(y, r)) + \nu_d(K) > \nu_d(K).$$

□

Chapter 3

Stereology for a particle process with a fixed grain shape

3.1 Introduction

Having studied random sections of a fixed particle in chapter 2, we are now ready to consider stereological estimation for a specific type of particle process. That is, in this chapter we study a generalization of Wicksell's corpuscle problem. This generalization may be described as follows. Consider 3D particles, convex bodies to be precise, which are randomly positioned in an opaque body and randomly oriented. A convex body is a compact and convex set with a non-empty interior. These particles all have the same known shape, but they do not have the same size. The particles cannot be observed directly, instead the medium is intersected with a plane, and we observe the 2D section profiles of the particles which happened to be hit by the plane. We address the statistical problem of estimating the size distribution of the particles, using a sample of observed areas of the section profiles. A visualization of the problem setting is given in Figure 3.1. In this particular example, each particle is a convex dodecahedron.

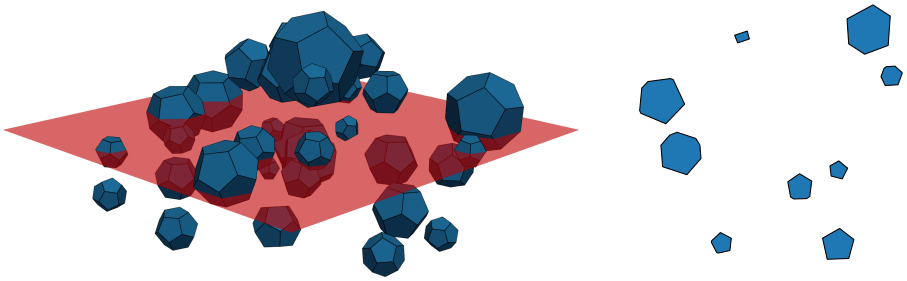


Figure 3.1: Left: Random spatial system of convex dodecahedra intersected with a plane. Right: Observed section profiles.

An overview of estimators for the size distribution in the spherical setting is presented in [18]. The problem has been studied for shapes other than spheres as well. In [70] the case of cubic particles is considered. In [68] a variation of the problem is studied: the particles are random polyhedra, and therefore not all particles have the same shape in this setting. A system of oriented cylinders is considered in [60].

More generally, the problem we consider in this chapter is a statistical inverse problem. For an overview of the key concepts in statistical inverse problems, illustrated with the deconvolution problem, we refer to [14]. Further aspects of statistical inverse problems are discussed in [17]. In this chapter we also make a connection with deconvolution problems, and the proposed estimator may also be applicable in more general settings. In particular, it can be used for what may be described as a multiplicative deconvolution problem. In the classical deconvolution setting the observations are contaminated with independent additive noise, in the multiplicative setting the (positive) observations are contaminated with independent (positive) multiplicative noise. A well-known instance of this problem is multiplicative censoring. In multiplicative censoring the multiplicative noise is uniformly distributed, this model was first introduced in [110].

The main contributions of this chapter are as follows. A key insight in our instance of the problem highlights that we can separate the shape of the particles from their sizes in the sense that an observed area may be interpreted as the product of two independent random variables, one related to the particle size and the other related to the known particle shape. The density function of the shape-related random variable is explicitly known only in exceptional cases, therefore we rely on the simulation procedure described in section 2.4 that can be used to approximate it arbitrarily well.

Using that shape-related distribution as an ingredient, we design a maximum likelihood procedure to estimate the size distribution of the particles, a procedure that can be used for a large class of possible shapes. Furthermore, we show the consistency of the resulting estimator and provide algorithms that can be used to compute it. Additionally, we assess the proposed estimator in a small simulation study in which we focus on convex polyhedra for the shape of the particles.

This chapter is organized as follows. In section 3.2 an integral equation is derived which describes the problem. Via this equation, we obtain an identifiability result in section 3.3 stating that the profile area distribution uniquely determines the 3D size distribution. We define an estimator for the so-called biased size distribution in section 3.4. In section 3.5 we prove the consistency of this estimator. Algorithms for computing the proposed estimator are discussed in section 3.6. In section 3.7 we describe how to estimate the particle size distribution via the biased size distribution. In section 3.8 some simulations are performed and at the end of the chapter we provide some conclusions in section 3.9.

We now briefly introduce some notation. Given a $\lambda > 0$ and a set $A \subset \mathbb{R}^3$ the scalar multiplication of λ with A is defined as: $\lambda A = \{\lambda x : x \in A\}$. Let $\mathrm{SO}(3)$ denote the rotation group of degree 3. It contains all 3×3 rotation matrices, which are orthogonal matrices of determinant 1. When a convex body $K \subset \mathbb{R}^3$ is hit by an Isotropic Uniformly Random (IUR) plane, we obtain a section with a random area (recall that IUR planes were formally defined in section 2.2). Let G_K denote the cumulative distribution function (CDF) associated with such a random area. It is sometimes referred to as cross section area

distribution. We studied the CDF G_K in chapter 2, and we recall the following properties:

Theorem 3.1. *Let $K \subset \mathbb{R}^3$ be a convex body and let T be an IUR plane hitting K. The random variable $Z := \text{area}(K \cap T)$ has distribution function G_K. Let G_K^S denote the distribution function of $\sqrt{Z}$. The following properties hold:*

1. *Motion invariance: G_K and G_K^S are invariant under translations and rotations of K.*
2. *Scaling of convex bodies: $G_{\lambda K}(z) = G_K\left(\frac{z}{\lambda^2}\right)$ for all $\lambda > 0$, $z \in \mathbb{R}$.*
3. *Absolute continuity: If K is strictly convex or if it is a polyhedron then G_K and G_K^S have a Lebesgue density.*
4. *Initial monotonicity: If G_K^S has Lebesgue density g_K^S, then g_K^S is non-decreasing on $(0, \tau_K)$ for some $\tau_K > 0$.*

Note in particular that for a large class of convex bodies, G_K has a Lebesgue density. In section 3.4 we define an estimator for the particle size distribution. It will then become clear why the square root transformation in Theorem 3.1 is relevant.

3.2 Derivation of the stereological integral equation

In this section we give a formal description of the model and derive a stereological integral equation. As mentioned in the introduction of this thesis, stereology deals with estimating higher dimensional information from lower dimensional samples. The stereological equation in this section directly relates the distribution of the 3D particle sizes to the distribution of observed 2D section profile areas. We derive an expression for the density f_A of the observed section areas. Another derivation of this density appears in chapter 16 of [84]. The derivation has two purposes, it provides an intuitive understanding of the problem and the equation is used for defining an estimator. Throughout this section we occasionally need to refer back to section 2.2 for properties of IUR planes.

Let $Q \subset \mathbb{R}^3$ be the opaque convex body containing the randomly positioned particles. The intersection of Q with a random plane yields a sample of observed section profile areas. For now, assume that Q contains just one particle, a convex body K_1. Assume that K_1 is similar to a known convex body $K \subset \mathbb{R}^3$, which we refer to as the reference particle. This means that there exists a rotation $M \in \text{SO}(3)$, a point $x \in \mathbb{R}^3$ and a scalar $\Lambda > 0$ such that $K_1 = \Lambda MK + x := \{\Lambda Mk + x : k \in K\}$. We refer to the scalar Λ as the size of K_1, which is distributed according to an unknown size distribution on $(0, \infty)$ with CDF H and PDF h. As such, the size is the scaling with respect to the reference particle, which has size 1. The mean size is denoted by:

$$\mathbb{E}(\Lambda) = \int_0^\infty \lambda h(\lambda) \mathrm{d}\lambda,$$

and we assume $0 < \mathbb{E}(\Lambda) < \infty$ throughout. Let T be an IUR plane hitting Q. Let $B := \{T \cap K_1 \neq \emptyset\}$ be the event that K_1 is hit by T. By Lemma 2.1, the probability that K_1 is

hit by T given that it has size λ is given by:

$$\mathbb{P}(B|\Lambda = \lambda) = \frac{\bar{b}(\lambda K)}{\bar{b}(Q)} = \lambda \frac{\bar{b}(K)}{\bar{b}(Q)}. \tag{3.1}$$

Recall that $\bar{b}(K)$, the mean width of K, was defined in (2.5). In (3.1) we use the fact that $\bar{b}(\lambda K) = \lambda \bar{b}(K)$ and $\bar{b}(K)$ is invariant under rotations and translations of K. While Λ is drawn from H, the size of a particle which appears in the plane section follows a different distribution from H. By this we mean that $\Lambda|B$ is not distributed according to H. Note that the probability in (3.1) is proportional to λ, via Bayes' rule the density of $\Lambda|B$, denoted by h^b is computed as:

$$h^b(\lambda) := f_{\Lambda|B}(\lambda) = \frac{\mathbb{P}(B|\Lambda = \lambda)h(\lambda)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|\Lambda = \lambda)h(\lambda)}{\int_0^\infty \mathbb{P}(B|\Lambda = \lambda)h(\lambda)\mathrm{d}\lambda} = \frac{\lambda h(\lambda)}{\mathbb{E}(\Lambda)}.$$

Throughout this chapter we refer to h^b as the density of the length-biased size distribution associated with h. Let H^b be the CDF corresponding to h^b and note that H and H^b are related via:

$$H^b(\lambda) = \frac{\int_0^\lambda x\mathrm{d}H(x)}{\int_0^\infty x\mathrm{d}H(x)} \quad \text{and} \quad H(\lambda) = \frac{\int_0^\lambda \frac{1}{x}\mathrm{d}H^b(x)}{\int_0^\infty \frac{1}{x}\mathrm{d}H^b(x)}, \quad \lambda \geq 0. \tag{3.2}$$

We refer to H^b as the length-biased size distribution function, or the length-biased version of H. For an elaborate overview of length-biased and more generally size-biased distributions we refer to [5]. The authors also prove the following general property of length-biased distributions: if $\Lambda_b \sim H^b$ and $\Lambda \sim H$, then: $\mathbb{P}(\Lambda_b \geq \lambda) \geq \mathbb{P}(\Lambda \geq \lambda)$. Hence, as H^b is the size distribution of the particles which appear in the plane section, this means that larger particles are more likely to appear in the cross section.

We can now derive the distribution of an observed section area, resulting from K_1 being hit by the section plane. Conditional on K_1 being hit let $A := \mathrm{area}(K_1 \cap T)$. By the conditional property of IUR planes in Lemma 2.1, given that T hits K_1 it is an IUR plane hitting K_1. Therefore, if K_1 with size λ appears in the section plane, its section area is distributed according to $G_{\lambda K}$. Recall that the definition and some properties of $G_{\lambda K}$ are given in Theorem 3.1. Using the rules of conditional probability we find:

$$\begin{aligned} F_A(a) := \mathbb{P}(A \leq a|B) &= \int_0^\infty \mathbb{P}(A \leq a|B, \Lambda = \lambda) f_{\Lambda|B}(\lambda)\mathrm{d}\lambda \\ &= \int_0^\infty G_{\lambda K}(a)\mathrm{d}H^b(\lambda). \end{aligned}$$

Using point 2 of Theorem 3.1, F_A may be written as:

$$F_A(a) = \int_0^\infty G_K\left(\frac{a}{\lambda^2}\right)\mathrm{d}H^b(\lambda) = \frac{1}{\mathbb{E}(\Lambda)}\int_0^\infty G_K\left(\frac{a}{\lambda^2}\right)\lambda\mathrm{d}H(\lambda). \tag{3.3}$$

Suppose now that we randomly position and orient non-overlapping particles $K_1, K_2, \ldots$ in Q, each similar to K. More specifically, the centers of the particles are distributed

according to a homogeneous Poisson point process. As for the orientations, all orientations of the particles are equally likely and independent. The sizes of the particles $\Lambda_1, \Lambda_2, \ldots$ are independent and identically distributed (iid) according to H. Intersecting Q with an IUR plane yields an iid sample $A_1, \ldots, A_n$ from F_A of observed section areas, for some random n. Let K be a convex body such that G_K has a density g_K, recall Theorem 3.1. Then, F_A has a density given by:

$$f_A(a) = \frac{1}{\mathbb{E}(\Lambda)} \int_0^\infty g_K\left(\frac{a}{\lambda^2}\right) \frac{1}{\lambda} \mathrm{d}H(\lambda). \tag{3.4}$$

Let $a_{\max}$ be the largest possible section area of K, such that g_K has support $(0, a_{\max})$. Hence, $g_K(a/\lambda^2) = 0$ whenever $a/\lambda^2 > a_{\max} \iff \lambda < \sqrt{a/a_{\max}}$. As a result, the lower bound of the integration region in (3.4) is effectively equal to $\sqrt{a/a_{\max}}$. The stereological equation (3.4) directly relates the sizes of the 3D particles to the areas of the observed 2D section profiles.

Example 1 (Wicksell's corpuscle problem)**.** Choose for the reference particle $K = \bar{B}(0,1) = \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$, the ball with radius 1. Then: $g_K(z) = 1/(2\pi\sqrt{1 - z/\pi})$, $0 < z < \pi$. We may interpret H as the distribution function of the radii of the 3D balls. Note that any plane section of a ball yields a circular disc. Given $A \sim f_A$ set $A = \pi R^2$, the density of the observed circle radii satisfies: $f_R(r) = f_A(\pi r^2) 2\pi r$, $r > 0$. Combining this with (3.4) yields:

$$f_R(r) = \frac{1}{\mathbb{E}(\Lambda)} \int_r^\infty \frac{1}{2\pi\sqrt{1 - \frac{r^2}{\lambda^2}}} \frac{2\pi r}{\lambda} \mathrm{d}H(\lambda) = \frac{r}{\mathbb{E}(\Lambda)} \int_r^\infty \frac{1}{\sqrt{\lambda^2 - r^2}} \mathrm{d}H(\lambda),$$

which corresponds to the well-known Wicksell's integral equation [114].

Remark 3.1. *By taking an appropriate choice for the reference particle, the size distribution may be directly related to a more convenient distribution. For example, if the reference particle has a diameter 1, then the size distribution corresponds to the distribution of the diameters of the particles. When choosing a reference particle with volume 1, then a particle with size λ has volume λ^3. The volume distribution function is then given by $F_V(x) = \mathbb{P}(\Lambda^3 \leq x) = H(x^{\frac{1}{3}})$.*

The derived stereological equation also holds under different assumptions. The random system of particles may be defined by choosing an isotropic typical particle, and then positioning the particles using a stationary point process on $\mathbb{R}^3$. A formal derivation of the stereological equation under these assumptions is presented in Appendix B. This model is also known as a germ-grain process. Relevant references for germ-grain processes are sections 6.5 and 10.5 in [18], as well as [69] and [70]. Hence, there is no need to restrict the particles to an opaque body or to position the particles via a Poisson point process.

In this setting, let N_V denote the expected number of 3D particles per unit volume, which corresponds to the intensity parameter of the stationary point process. Intersecting the system of particles with a plane, let N_A denote the expected number of observed 2D

section profiles per unit area. Here, F_A is the CDF associated with the typical section profile area. By combining the well known stereological equation (Theorem 10.1 in [18]):

$$N_A = N_V \bar{\bar{b}} \quad \text{with} \quad \bar{\bar{b}} := \bar{b}(K) \int_0^\infty \lambda \mathrm{d}H(\lambda),$$

and (3.3), yields:

$$N_A(1 - F_A(a)) = N_V \bar{b}(K) \int_0^\infty \lambda(1 - G_{\lambda K}(a))\mathrm{d}H(\lambda). \tag{3.5}$$

A derivation of a slightly more general version of (3.5) may be found in chapter 6 of [11]. We specifically mention (3.5) since it appears more frequently in the literature than (3.4).

In order to obtain a better understanding of the problem it helps to apply a transformation. We apply a square root transformation to (3.3). For $A \sim F_A$, set $S = \sqrt{A}$ such that $S \sim F_S$ and $F_S(s) = F_A(s^2)$ for $s \in \mathbb{R}$. As in Theorem 3.1, let $Z \sim G_K$ then $\sqrt{Z} \sim G_K^S$ and $G_K^S(z) = G_K(z^2)$ for $z \in \mathbb{R}$. Let $s \in \mathbb{R}$, then the following holds:

$$F_S(s) = \int_0^\infty G_K^S\left(\frac{s}{\lambda}\right) \mathrm{d}H^b(\lambda). \tag{3.6}$$

This expression may be recognized as the distribution function corresponding to a product of two independent random variables. This is a key insight which is made precise in the following lemma.

Lemma 3.1. *Consider a distribution function H with length-biased version H^b. Suppose $Z \sim G_K$ and $\Lambda_b \sim H^b$ with Z and Λ_b^2 independent. Set $A = Z\Lambda_b^2$. Then, $A \sim F_A$, and F_A, G_K and H^b are related via (3.3).*

Proof. Let X, Y, Z be non-negative random variables, with CDF F_X, F_Y and F_Z respectively. If $X = YZ$ with Y and Z independent, then their distribution functions are related via:

$$F_X(x) = \int_0^\infty F_Y\left(\frac{x}{z}\right) \mathrm{d}F_Z(z).$$

Comparing this with (3.6), the result is immediate. □

Let us provide some further intuition for Lemma 3.1. Note that point 2 of Theorem 3.1 means that for a given size $\lambda > 0$, if $Z \sim G_K$ then $Z\lambda^2 \sim G_{\lambda K}$. As the sizes of the particles in the section plane are distributed according to H^b, this hints towards the relationship given in Lemma 3.1.

Therefore, there are two main considerations in this problem. First, the size distribution of particles appearing in the cross section is a length-biased version of the actual size distribution. Second, we can separate the common shape of the particles and their sizes in some sense. Taking a random size from H^b, and independently taking an IUR section of the reference particle yields a sample from F_A via the relationship given in Lemma 3.1.

3.3 Identifiability of the particle size distribution

In this section, we present a general identifiability result for our model. This means that under appropriate conditions, given a known reference particle, there are no two size distributions which yield the same distribution of observed section areas. For this result we need the Mellin-Stieltjes transform, which we will also refer to as the Mellin transform. While characteristic functions appear naturally when studying sums of independent random variables, the Mellin transform is appropriate when studying products of independent random variables. We collect some properties of the Mellin transform, for details we refer to section 7.8 in [48] and to [118]. We note that the use of the Mellin transform for this problem was already considered in [51]. The authors obtain a slightly different expression due to the fact that an inversion formula for the density h was derived and because the density f_A in (3.4) was studied up to a normalization constant. The identifiability result in this section is new, a sufficient condition for identifiability in this context has not been derived before.

Definition 3.2 (Mellin-Stieltjes transform)**.** Given a non-negative random variable X, with CDF F, the Mellin-Stieltjes transform of X is defined as:

$$\mathcal{M}_X(s) = \mathbb{E}\left(X^{s-1}\right) = \int_0^\infty x^{s-1}\mathrm{d}F(x),$$

for $s \in \mathbb{C}$, whenever the integral is absolutely convergent.

Note in particular, that whenever $\int_0^\infty x^{c-1}\mathrm{d}F(x) < \infty$ for some $c \in \mathbb{R}$, then the Mellin transform exists for all $s = c + it$, $t \in \mathbb{R}$. Hence, existence of the Mellin transform corresponds to the existence of moments of a distribution. Let $\mathrm{St}(\alpha, \beta) := \{s \in \mathbb{C} : \alpha < \mathrm{Re}(s) < \beta\}$ denote the open strip parallel to the imaginary axis. Analogously, $\mathrm{St}[\alpha, \beta] := \{s \in \mathbb{C} : \alpha \leq \mathrm{Re}(s) \leq \beta\}$ denotes the closed strip. If we find $\alpha < \beta$ such that the Mellin transform of X converges absolutely on $\mathrm{St}[\alpha, \beta]$, then $\mathcal{M}_X$ is analytic on $\mathrm{St}(\alpha, \beta)$. Taking α as small as possible and β as large as possible, this open strip is referred to as the strip of analyticity of $\mathcal{M}_X$. A Mellin transform uniquely determines a distribution in the following sense:

Lemma 3.2 (Uniqueness of the Mellin transform)**.** *Let $X \sim F_1$ and $Y \sim F_2$. Assume the integrals in $\mathcal{M}_X$ and $\mathcal{M}_Y$ converge absolutely on $\mathrm{St}[\alpha, \beta]$, $0 \leq \alpha < \beta$. If $c \in (\alpha, \beta)$ and $\mathcal{M}_X(c + it) = \mathcal{M}_Y(c + it)$ for all $t \in \mathbb{R}$ then $F_1 = F_2$.*

The proof is given in section 3.10. A similar statement is proven in Theorem 8 in [16] for the case that the CDF has a Lebesgue density. Finally, we recall the Mellin convolution theorem. Let X, Y, Z be non-negative random variables, such that $X = YZ$ with Y and Z independent. For any $s \in \mathbb{C}$ such that $\mathcal{M}_Y(s)$ and $\mathcal{M}_Z(s)$ are finite:

$$\mathcal{M}_X(s) = \mathbb{E}\left(X^{s-1}\right) = \mathbb{E}\left((YZ)^{s-1}\right) = \mathbb{E}\left(Y^{s-1}\right)\mathbb{E}\left(Z^{s-1}\right) = \mathcal{M}_Y(s)\mathcal{M}_Z(s).$$

Having collected these properties we now state the identifiability result.

Theorem 3.3 (Identifiability)**.** *Suppose we are given densities f_A, g_K such that f_A can be expressed as in (3.4) for some CDF H.*

1. *If $\int_0^\infty z^{-\alpha} g_K(z)\mathrm{d}z < \infty$ for some $\alpha > 0$, then there is only one distribution function H on $(0,\infty)$ satisfying (3.4).*
2. *Assume $\int_0^\infty \lambda^{1+\delta}\mathrm{d}H(\lambda) < \infty$, for some $\delta > 0$. Then, there is only one such distribution function H on $(0,\infty)$ satisfying (3.4).*

Proof. We first consider statement 1 of the theorem. Let $\Lambda_b \sim H^b$, with H^b as in (3.2), and let $\Lambda \sim H$. Let $Z \sim g_K$ and $A \sim f_A$. We first determine on which strips the Mellin transforms of the random variables of interest are analytic. Since $\mathbb{E}(\Lambda_b^{-1}) = 1/\mathbb{E}(\Lambda)$ and $\mathbb{E}(\Lambda_b^0) = 1$ we obtain that $\mathcal{M}_{\Lambda_b}$ is analytic on $\mathrm{St}(0,1)$. Note that $1/2 < \mathrm{Re}(s) < 1 \iff 0 < \mathrm{Re}(2s-1) < 1$. As a result:

$$\mathcal{M}_{\Lambda_b^2}(s) = \mathbb{E}\left(\Lambda_b^{2s-2}\right) = \mathcal{M}_{\Lambda_b}(2s-1),$$

for all $s \in \mathrm{St}(1/2,1)$ and $\mathcal{M}_{\Lambda_b^2}$ is analytic on $\mathrm{St}(1/2,1)$. Choose $\alpha > 0$ such that $\mathbb{E}(Z^{-\alpha}) < \infty$. Because g_K has bounded support, all non-negative moments of Z exist and therefore $\mathcal{M}_Z$ is analytic on $\mathrm{St}(1-\alpha,\infty)$. By Lemma 3.1 and the Mellin convolution theorem we obtain:

$$\mathcal{M}_A(s) = \mathcal{M}_Z(s)\mathcal{M}_{\Lambda_b^2}(s),$$

for all $s \in \mathrm{St}(1-\alpha,\infty) \cap \mathrm{St}(1/2,1) = \mathrm{St}(\max\{1-\alpha,1/2\},1)$. Moreover, this also means that $\mathcal{M}_A$ is analytic on $\mathrm{St}(\max\{1-\alpha,1/2\},1)$. Let $c \in (\max\{1-\alpha,1/2\},1)$. Define:

$$L_Z := \{c+it : t \in \mathbb{R},\ \mathcal{M}_Z(c+it) \neq 0\}, \quad \text{and} \quad L := \{c+it : t \in \mathbb{R}\}.$$

For all $s \in L_Z$ we find: $\mathcal{M}_A(s)/\mathcal{M}_Z(s) = \mathcal{M}_{\Lambda_b^2}(s)$. Define $f : L_Z \to \mathbb{C}$ by $f(s) = \mathcal{M}_A(s)/\mathcal{M}_Z(s)$. Note that f is analytic on L_Z, because $s \mapsto \mathcal{M}_{\Lambda_b^2}(s)$ is analytic on the line L. As a result there is a unique analytic continuation of f to L. The uniqueness of this analytic continuation implies: $f(s) = \mathcal{M}_{\Lambda_b^2}(s)$, for all $s \in L$. Suppose $\bar{H}$ also satisfies (3.4), with $\bar{H}^b$ denoting its length-biased version and $\bar{\Lambda}_b \sim \bar{H}^b$. Then, following the same steps as before, we obtain: $f(s) = \mathcal{M}_{\bar{\Lambda}_b^2}(s)$ for all $s \in L$. By Lemma 3.2, Λ_b^2 and $\bar{\Lambda}_b^2$ have the same CDF. Therefore, for all $x \in \mathbb{R}$:

$$H^b(x) = \mathbb{P}\left(\Lambda_b^2 \leq x^2\right) = \mathbb{P}\left(\bar{\Lambda}_b^2 \leq x^2\right) = \bar{H}^b(x).$$

By (3.2) this also implies $H = \bar{H}$.

The proof of the second statement of the theorem is analogous, we simply highlight the differences. Let $\delta > 0$ be such that $\mathbb{E}(\Lambda^{1+\delta}) < \infty$. Note that $\mathbb{E}(\Lambda_b^{-1}) = 1/\mathbb{E}(\Lambda)$ and $\mathbb{E}(\Lambda_b^\delta) = \mathbb{E}(\Lambda^{\delta+1})/\mathbb{E}(\Lambda)$. It then follows that $\mathcal{M}_{\Lambda_b}$ is analytic on $\mathrm{St}(0,1+\delta)$ and $\mathcal{M}_{\Lambda_b^2}$ is analytic on $\mathrm{St}(1/2,1+\delta/2)$. Clearly, $\mathcal{M}_Z$ is analytic on $\mathrm{St}(1,\infty)$. Hence, $\mathcal{M}_A$ is analytic on $\mathrm{St}(1/2,1+\delta/2) \cap \mathrm{St}(1,\infty) = \mathrm{St}(1,1+\delta/2)$. In this case we take $c \in (1,1+\delta/2)$ and the remainder of the proof is as before. □

We obtain as a consequence:

Corollary 3.1. *In the following cases the distribution function H is identifiable:*

1. *The Wicksell corpuscle problem.*
2. *G_K^S has a bounded density g_K^S.*

Proof. Recall the expression for g_K in Example 1. Then:

$$\int_0^\pi z^{-\frac{1}{2}} g_K(z)\mathrm{d}z = \int_0^\pi \frac{1}{2\pi\sqrt{z}\sqrt{1-\frac{z}{\pi}}}\mathrm{d}z = \frac{\sqrt{\pi}}{2}.$$

This integral may be computed by substituting $t = z/\pi$ and recognizing the resulting integral as a constant multiple of the Beta function evaluated in $(1/2, 1/2)$. Hence, condition 1 of Theorem 3.3 is satisfied. If G_K^S has a density g_K^S with $g_K^S \leq m$, then:

$$\int_0^{a_{\max}} z^{-\frac{1}{4}} g_k(z)\mathrm{d}z = \int_0^{\sqrt{a_{\max}}} z^{-\frac{1}{2}} g_k^S(z)\mathrm{d}z \leq m \int_0^{\sqrt{a_{\max}}} z^{-\frac{1}{2}}\mathrm{d}z = 2m(a_{\max})^{\frac{1}{4}}.$$

Therefore, in this case condition 1 of Theorem 3.3 is also satisfied. □

Identifiability for the Wicksell problem is a classical result, in this case there is also a well-known explicit inverse relation.

Remark 3.2. *The condition in Theorem 3.3: $\int_0^\infty \lambda^{1+\delta}\mathrm{d}H(\lambda) < \infty$, for some $\delta > 0$, is also implied by the assumption $H(M) = 1$ for some $M > 0$. Recall the derivation of (3.4) in section 3.2, a maximum size of the particles is clearly enforced by that fact that they are contained within the body Q. This is a typical assumption in stereological problems.*

Note that the proof of Theorem 3.3 also presents (a rather implicit) inversion formula for H^b. Assume H^b is continuous. Let c be as in the proof of Theorem 3.3. Since analytic functions only have isolated zeros, $\mathcal{M}_Z(c + it) \neq 0$ for almost all $t \in \mathbb{R}$. By using the Mellin inversion formula as in the proof of Lemma 3.2:

$$H^b\left(\sqrt{x}\right) = \mathbb{P}\left(\Lambda_b^2 \leq x\right) = \lim_{T\to\infty} \frac{1}{2\pi i}\int_{c-iT}^{c+iT} -\frac{\mathcal{M}_A(s)}{\mathcal{M}_Z(s)}\frac{x^{-s+1}}{s}\mathrm{d}s, \quad x \geq 0. \tag{3.7}$$

H can then be retrieved via (3.2). We note that in [51] another expression for h was derived in terms of (inverse) Mellin transforms.

3.4 Estimator for the length-biased particle size distribution

In this section, we propose an estimator for the length-biased size distribution H^b. The proposed estimator is inspired by the approach taken in [46], for Wicksell's corpuscle problem. Given the random fraction interpretation of Lemma 3.1, first estimating H^b seems a natural intermediate step. We note that biased or weighted distributions frequently appear in stereology, see also section 7.5 in [69]. For the remainder of the chapter we

consider the number of observed section profiles n to be deterministic. This is a common convention for problems of this type. This means that we condition on a fixed number of particles being hit by the section plane. All further analysis is based on an iid sample of size n from f_A.

The reference particle K is considered to be known and we assume that it satisfies one of the conditions in Theorem 3.1 such that G_K^S has a density g_K^S. We stress that this also means that we consider g_K^S to be known. While there are very few shapes for which an explicit expression is known for g_K^S, we may use the Monte Carlo simulation scheme from section 2.4 to approximate such a density arbitrarily closely. To give some insight into what these densities look like, see Figure 3.2 for approximations of these densities for the cube, dodecahedron and tetrahedron. These approximations are obtained by computing a kernel density estimator with boundary correction, based on a sample of size $N = 10^7$.

Recall the square root transformation and the resulting expression (3.6). Because G_K^S has density g_K^S, F_S has a density f_S given by:

$$f_S(s) = \int_0^\infty g_K^S\left(\frac{s}{\lambda}\right)\frac{1}{\lambda}\mathrm{d}H^b(\lambda). \tag{3.8}$$

Keep in mind that g_K^S is supported on $(0, \sqrt{a_{\max}})$, such that the lower bound of the integration region is effectively $s/\sqrt{a_{\max}}$.

Suppose we have a sample of observed section areas: $A_1, \ldots, A_n \overset{\text{iid}}{\sim} f_A$. Let $S_i = \sqrt{A_i}$, then $S_1, \ldots, S_n \overset{\text{iid}}{\sim} f_S$, with f_S as in (3.8). Now, let $s_1 < s_2 < \cdots < s_n$ be a realization of the order statistics of $S_1, \ldots, S_n$. We use (3.8) to implicitly define an estimator for H^b via nonparametric maximum likelihood. This is achieved by considering a large class of distribution functions for H^b. Let $\mathcal{F}^+$ be the class of all distribution functions on $(0, \infty)$. Define:

$$\mathcal{F}_n^+ = \{F \in \mathcal{F}^+ : F \text{ is constant on } [s_{i-1}, s_i), i \in \{1, \ldots, n\}, \text{with } F(s_0) = 0\},$$

for some $0 < s_0 < s_1$. This means that $\mathcal{F}_n^+$ contains all piece-wise constant distribution functions with jump locations restricted to the set of observations, the s_i's. Note that as

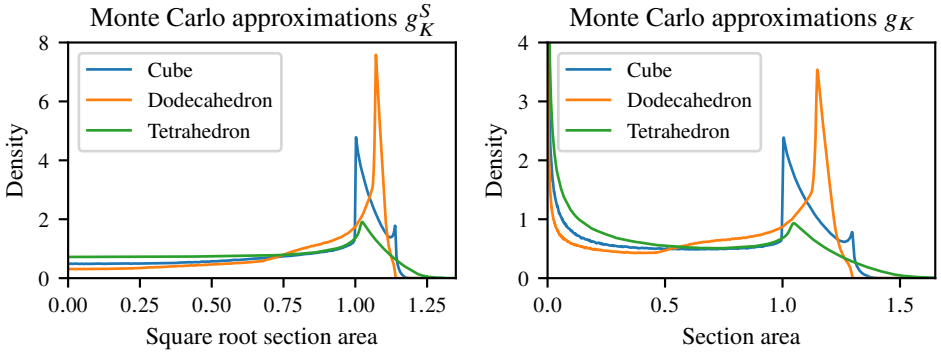


Figure 3.2: Left: Monte Carlo approximations of g_K^S, for various shapes K of unit volume. Right: Approximations of g_K, obtained via $g_K(z) = g_K^S(\sqrt{z})/(2\sqrt{z})$.

$n \to \infty$ the set of observed s_i's becomes dense in the support of f_S and the class $\mathcal{F}_n^+$ grows to the class of all distribution functions with the same support as f_S.

Remark 3.3. *If $H^b(M) = 1$ for some $M > 0$, then f_S is supported on the interval $(0, M\sqrt{a_{max}})$. When choosing the size of the reference particle K, it is important that $a_{max} \geq 1$. This is due to the choice of the sieve $\mathcal{F}_n^+$. Then, as n tends to infinity $\mathcal{F}_n^+$ grows to the class of distribution functions which also contains the true CDF H^b, since $M\sqrt{a_{max}} \geq M$. Taking a very large K means a_{max} is large, such that $M\sqrt{a_{max}}$ is much larger than M. Then, the s_i's will be quite sparse in $[0, M]$, which is also undesirable. For the sake of interpretability of H, recall Remark 3.1, we choose a K with volume 1. For the shapes considered in simulations we observed $a_{max} \geq 1$.*

For $H^b \in \mathcal{F}_n^+$ we define the (scaled by $\frac{1}{n}$) log-likelihood:

$$L(H^b) := \frac{1}{n}\sum_{i=1}^{n} \log(f_S(s_i)) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\int_0^\infty g_K^S\left(\frac{s_i}{\lambda}\right)\frac{1}{\lambda}\mathrm{d}H^b(\lambda)\right). \tag{3.9}$$

A maximum likelihood estimator (MLE) $\hat{H}_n^b$ for H^b is defined as a maximizer of the log-likelihood L, which may be written as:

$$\hat{H}_n^b \in \underset{H^b \in \mathcal{F}_n^+}{\arg\max} \frac{1}{n}\sum_{i=1}^{n} \log\left(\sum_{j=1}^{n} g_K^S\left(\frac{s_i}{s_j}\right)\frac{1}{s_j}(H^b(s_j) - H^b(s_{j-1}))\right). \tag{3.10}$$

The following theorem shows that this estimator is well-defined, and provides a sufficient condition for uniqueness:

Theorem 3.4 (Existence and uniqueness of $\hat{H}_n^b$)**.** *A maximizer of the log-likelihood L in $\mathcal{F}_n^+$ always exists. The maximizer is unique if the matrix $A = (\alpha_{i,j})$, with $\alpha_{i,j} = g_K^S(s_i/s_j)/s_j$, $i, j \in \{1, \ldots, n\}$, is full-rank.*

Proof. For $H^b \in \mathcal{F}_n^+$ define: $\beta_j = H^b(s_j)$ and write $\beta = (\beta_1, \beta_2, \ldots, \beta_n)^\mathsf{T}$. Consider the closed convex set:

$$C := \{\beta \in \mathbb{R}^n : 0 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_n \leq 1\}. \tag{3.11}$$

The maximization problem (3.10) is equivalent to maximizing $l : C \to \mathbb{R} \cup \{-\infty\}$ with l given by:

$$l(\beta) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\sum_{j=1}^{n} \alpha_{i,j}(\beta_j - \beta_{j-1})\right), \tag{3.12}$$

where $\alpha_{i,j} = g_K^S(s_i/s_j)/s_j$ and $\beta_0 = 0$. The set C is closed and bounded, and therefore compact. Because of the continuity of l on C, it has a maximum. We now show that l is strictly concave if and only if $A = (\alpha_{i,j})$ is full-rank. Strict concavity implies uniqueness of the maximum as well as the maximizer. Fix $\beta \in C$ such that $l(\beta) > -\infty$. Let $j, k \in$

$\{1, \dots, n\}$, computing the partial derivatives and Hessian of l yields:

$$\frac{\partial}{\partial \beta_j} l(\beta) = \frac{1}{n} \sum_{i=1}^{n} \frac{\alpha_{i,j} - \alpha_{i,j+1}}{\sum_{q=1}^{n} \alpha_{i,q}(\beta_q - \beta_{q-1})} \tag{3.13}$$

$$\frac{\partial^2}{\partial \beta_j \partial \beta_k} l(\beta) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\left(\alpha_{i,j} - \alpha_{i,j+1}\right)\left(\alpha_{i,k} - \alpha_{i,k+1}\right)}{\left(\sum_{q=1}^{n} \alpha_{i,q}(\beta_q - \beta_{q-1})\right)^2} =: H_{j,k}(\beta). \tag{3.14}$$

Here, we have set $\alpha_{i,n+1} = 0$ for all $i \in \{1, \dots, n\}$. Since $l(\beta) > -\infty$, there are no divisions by zero in (3.13) and (3.14). Note that the following holds for $k \in \{1, \dots, n\}$:

$$\sum_{j=1}^{n} \alpha_{k,j}(\beta_j - \beta_{j-1}) = \sum_{j=1}^{n} \alpha_{k,j}\beta_j - \sum_{j=0}^{n-1} \alpha_{k,j+1}\beta_j = \sum_{j=1}^{n} (\alpha_{k,j} - \alpha_{k,j+1})\beta_j.$$

Using this fact we show that the Hessian of l is negative definite if and only if A is full-rank. Let $\gamma \in \mathbb{R}^n$ and set $\gamma_0 = 0$, then:

$$\begin{aligned}
\gamma^\top H(\beta)\gamma &= \sum_{j=1}^{n}\sum_{k=1}^{n} H_{j,k}(\beta)\gamma_j\gamma_k \\
&= -\frac{1}{n} \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\left(\alpha_{i,j} - \alpha_{i,j+1}\right)\left(\alpha_{i,k} - \alpha_{i,k+1}\right)\gamma_j\gamma_k}{\left(\sum_{q=1}^{n} \alpha_{i,q}(\beta_q - \beta_{q-1})\right)^2} \\
&= -\frac{1}{n} \sum_{i=1}^{n} \frac{\left(\sum_{j=1}^{n} \alpha_{i,j}(\gamma_j - \gamma_{j-1})\right)^2}{\left(\sum_{q=1}^{n} \alpha_{i,q}(\beta_q - \beta_{q-1})\right)^2}
\end{aligned}$$

Clearly, $H(\beta)$ is negative semidefinite. Note that the following holds:

$$\gamma^\top H(\beta)\gamma = 0 \iff \sum_{j=1}^{n} \alpha_{i,j}(\gamma_j - \gamma_{j-1}) = 0, \quad \text{for all } i \in \{1, \dots, n\}. \tag{3.15}$$

Define $x \in \mathbb{R}^n$ via $x_j = \gamma_j - \gamma_{j-1}$, $j \in \{1, \dots, n\}$. Consider the matrix $A = (\alpha_{i,j})$, then the RHS of (3.15) may be written as $Ax = 0$. Since $\gamma_0 = 0$: $\gamma = 0 \iff x = 0$. Therefore, the Hessian is negative definite if and only if $Ax = 0 \iff x = 0$, which corresponds to A being full-rank. □

Recall that g_K^S is supported on $(0, \sqrt{a_{\max}})$. Suppose we choose the reference particle such that $\sqrt{a_{\max}} = 1 + \varepsilon$ for some small $\varepsilon > 0$. Then, $g_K^S(1) > 0$ ensuring that the diagonal of A contains positive entries. Whenever $s_i/s_j > 1 + \varepsilon$ for all $i > j$, A is an upper triangular matrix, because all entries below the diagonal are zero. It is well-known that such matrices are of full-rank. If $\varepsilon > 0$ is chosen sufficiently small, then with high probability $s_{j+1}/s_j > 1 + \varepsilon$ for all $j \in \{1, \dots, n\}$, such that the MLE is unique with high probability. For the sake of convenience, we will refer to $\hat{H}_n^b$ as the MLE, even though we

cannot always guarantee uniqueness. Note especially for the consistency result in the next section that consistency of the MLE should be interpreted as consistency of any sequence of MLE's.

From the proof of Theorem 3.4 it is clear that $\hat{H}_n^b$ may be computed by maximizing l. Because l is a concave function, computing $\hat{H}_n^b$ can be done efficiently as we will discuss in section 3.6.

3.5 Consistency of the maximum likelihood estimator

In this section, we show that the MLE $\hat{H}_n^b$ (3.10) for H^b is uniformly strongly consistent. In order to prove this, we transform the problem into a deconvolution problem. Deconvolution problems have been studied before quite extensively, see for example [37] and [38]. We use some results on estimators in deconvolution problems to show consistency of the MLE. In deconvolution problems, it is typical that assumptions are made on the so-called noise kernel to ensure consistency. For this problem, this translates into assumptions on the density g_K^S.

We start by rewriting the problem of estimating H^b into a deconvolution problem. Recall Lemma 3.1, for $S \sim f_S$, $\sqrt{Z} \sim g_K^S$ and $\Lambda_b \sim H^b$ we have: $S \overset{d}{=} \sqrt{Z}\Lambda_b$, with $\sqrt{Z}$ and Λ_b independent. Let us now perform a log-transformation, define: $Y = \log(S)$, $\epsilon = \log(\sqrt{Z})$ and $X = \log(\Lambda_b)$. The densities of Y and ϵ are related to those of S and $\sqrt{Z}$ by: $f_Y(y) = f_S(e^y)e^y$, $f_\epsilon(z) = g_K^S(e^z)e^z$. The distribution function of X is given by $F_X(x) = H^b(e^x)$. We then obtain:

$$Y \overset{d}{=} X + \epsilon,$$

with X and ϵ independent. Note that f_Y is the convolution of f_ε and F_X:

$$f_Y(y) = \int_{-\infty}^{\infty} f_\varepsilon(y - x)\mathrm{d}F_X(x) =: (f_\varepsilon * \mathrm{d}F_X)(y). \tag{3.16}$$

In this setting, F_X is the distribution function of interest. We do not have direct observations from F_X, there is additive noise from the known distribution of ϵ. Let $\mathcal{F}$ be the class of all distribution functions on $\mathbb{R}$. Define:

$$\mathcal{F}_n = \{F \in \mathcal{F} : F \text{ is constant on } [y_{i-1}, y_i), \text{ for } i \in \{1, \ldots, n\}, \text{with } F(y_0) = 0\}.$$

The observed order statistics $s_1, \ldots, s_n$ are transformed as well: $y_i = \log(s_i), i \in \{0, 1, \ldots, n\}$. We proceed similarly as before, the log-likelihood may be written as:

$$\tilde{L}(F) = \frac{1}{n}\sum_{i=1}^{n} \log\left(f_Y(y_i)\right) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\int_{-\infty}^{\infty} f_\epsilon\left(y_i - x\right)\mathrm{d}F_X(x)\right).$$

A maximum likelihood estimator $\hat{F}_n$ for F_X is defined as:

$$\hat{F}_n \in \underset{F_X \in \mathcal{F}_n}{\arg\max}\, \tilde{L}(F_X). \tag{3.17}$$

We now show that we may assume $\hat{H}_n^b(x) = \hat{F}_n(\log(x))$. The likelihoods of the two problems are related as follows. Let $F_X \in \mathcal{F}_n$, define $H^b(x) = F_X(\log(x))$ such that $H^b \in \mathcal{F}_n^+$. Then:

$$\begin{aligned}\tilde{L}(F_X) &= \frac{1}{n}\sum_{i=1}^{n}\log\left(\sum_{j=1}^{n} f_\epsilon\left(\log(s_i)-\log(s_j)\right)\left(F_X(\log(s_j))-F_X(\log(s_{j-1}))\right)\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\log\left(\sum_{j=1}^{n} g_K^S\left(\frac{s_i}{s_j}\right)\frac{s_i}{s_j}\left(H^b(s_j)-H^b(s_{j-1})\right)\right)\\ &= L(H^b)+\frac{1}{n}\sum_{i=1}^{n}\log(s_i).\end{aligned}$$

If we find a distribution function which provides a better likelihood in one of the problems, we immediately obtain a distribution function which provides the same improvement in likelihood in the other problem. So indeed, there exist MLE's which are related via: $\hat{H}_n^b(x) = \hat{F}_n(\log(x))$. The estimator $\hat{F}_n$ was studied in [37] and shown to be strongly uniformly consistent under some conditions on f_ε. This result may be used to show strong uniform consistency of $\hat{H}_n^b$, since:

$$\sup_{x>0}\left|\hat{H}_n^b(x)-H^b(x)\right| = \sup_{x>0}\left|\hat{F}_n(\log(x))-F_X(\log(x))\right| = \sup_{x\in\mathbb{R}}\left|\hat{F}_n(x)-F_X(x)\right|.$$

Let us now specify the assumptions we require for f_ϵ. We assume it belongs to the class $\mathcal{G}$ of upper semicontinuous functions that are of bounded variation on some compact interval and monotone outside this interval. Let $V_a^b(f)$ denote the total variation of the function f on the interval $[a,b]$, $a<b$. The class $\mathcal{G}$ may be written as:

$$\begin{aligned}\mathcal{G} = \Big\{ & g:\mathbb{R}\to[0,\infty) : g \text{ is an upper semicontinuous density such that } \exists M>0\\ & \text{with } V_{-M}^{M}(g)<\infty \text{ and } g \text{ is monotone on } (-\infty,-M] \text{ and } [M,\infty)\Big\}.\end{aligned}$$

This corresponds with the following assumptions on g_K^S:

Lemma 3.3. *Assume that g_K^S is upper semicontinuous and of bounded variation on its support. Then, the density $f_\epsilon:\mathbb{R}\to[0,\infty)$ given by $f_\epsilon(z) = g_K^S(e^z)e^z$ belongs to $\mathcal{G}$.*

The proof of this lemma can be found in section 3.10. We now collect some lemmas to obtain a consistency result for $\hat{F}_n$. The following result can be found in [37], as Corollary 1:

Lemma 3.4. *Let $\hat{F}_n$ be the MLE for F_X defined in (3.17). Assume $f_\epsilon \in \mathcal{G}$. Set $\hat{f}_n := f_\epsilon * \mathrm{d}\hat{F}_n$, $f_Y := f_\epsilon * \mathrm{d}F_X$, then almost surely:*

$$\lim_{n\to\infty}\|\hat{f}_n - f_Y\|_{L_1} = \lim_{n\to\infty}\int\left|\hat{f}_n(s)-f_Y(s)\right|\mathrm{d}s = 0.$$

The following lemma is a generalization of Lemma 3 in [37]. The proof is given in section 3.10.

Lemma 3.5. *Let f_ε be a Lebesgue density on $\mathbb{R}$. Let $(F_n)_{n\geq 1}$ be a sequence of distribution functions on $\mathbb{R}$, converging weakly to a distribution function F_X. Then, for $f_n := f_\epsilon * \mathrm{d}F_n$ and $f_Y := f_\epsilon * \mathrm{d}F_X$:*

$$\lim_{n\to\infty} \|f_n - f_Y\|_{L_1} = 0.$$

We now state the following theorem, which closely follows the proof of Theorem 3 in [37].

Theorem 3.5 (Consistency of $\hat{F}_n$)**.** *Let $f_\epsilon \in \mathcal{G}$. Assume that the deconvolution problem with this f_ϵ is identifiable. Then, with probability one, $\lim_{n\to\infty} \hat{F}_n(x) = F_X(x)$ for each x where F_X is continuous. If F_X is continuous then with probability one:*

$$\lim_{n\to\infty} \left\|\hat{F}_n - F_X\right\|_\infty = 0.$$

Proof. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting a sequence $Y_1, Y_2, \ldots$ of iid random variables, distributed according to f_Y as in (3.16). Set $\hat{f}_n := f_\epsilon * \mathrm{d}\hat{F}_n$. By Lemma 3.4 we know there exists a set $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ we have $\|\hat{f}_n(\cdot, \omega) - f_Y(\cdot)\|_{L_1} \to 0$ as $n \to \infty$. Fix $\omega \in \Omega_0$ and choose an arbitrary subsequence $(n_l)_{l\geq 1} \subset (n)_{n\geq 1}$. By Helly's selection principle, there exists a further subsequence $(n_k)_{k\geq 1} \subset (n_l)_{l\geq 1}$ such that $\hat{F}_{n_k}(\cdot, \omega)$ converges weakly to a distribution function F. By Lemma 3.5 this implies that $\hat{f}_{n_k}$ converges to $f_\epsilon * \mathrm{d}F$ in L_1. Because the whole sequence $\hat{f}_n$ converges to $f_\epsilon * \mathrm{d}F_X$ in L_1 this implies $F = F_X$ by identifiability of the deconvolution problem. Therefore, every subsequence of MLE's contains a further subsequence converging weakly to F_X. This implies weak convergence of the whole sequence to F_X. Finally, the uniform result follows from the monotonicity of all distribution functions in the sequence and F_X, and continuity of F_X. □

Turning to a consistency result for $\hat{H}_n^b$ as in (3.10), we need to make sure that g_K^S satisfies the conditions in Lemma 3.3. If g_k^S satisfies these conditions, its boundedness implies the problem is identifiable by Corollary 3.1. Note that identifiability in the original problem implies identifiability in the corresponding deconvolution problem.

Corollary 3.2 (Consistency of $\hat{H}_n^b$)**.** *Assume g_K^S is upper semicontinuous and of bounded variation on its support. Then, with probability one, $\lim_{n\to\infty} \hat{H}_n^b(\lambda) = H^b(\lambda)$ for each λ where H^b is continuous. If H is continuous, then so is H^b, and with probability one:*

$$\lim_{n\to\infty} \left\|\hat{H}_n^b - H^b\right\|_\infty = 0.$$

3.6 Algorithms

In this section we describe some algorithms for computing the maximum likelihood estimator $\hat{H}_n^b$ (3.10). Since a distribution function in $\mathcal{F}_n^+$ is discrete, it may be described by a

probability vector. Let $\mathcal{P}_n$ be the class of probability vectors in $\mathbb{R}^n$:

$$\mathcal{P}_n = \left\{ (p_1, \ldots, p_n) \in \mathbb{R}^n : \sum_{i=1}^{n} p_i = 1 \text{ and } p_i \geq 0 \text{ for all } i \in \{1, \ldots, n\} \right\}.$$

A distribution function $H^b \in \mathcal{F}_n^+$ may be associated with the probability vector $p \in \mathcal{P}_n$ defined as $p_j = H^b(s_j) - H^b(s_{j-1})$ (recall: $H^b(s_0) = 0$). We can switch between probability vectors and distribution functions via:

$$H^b(s_j) = \sum_{i=1}^{j} p_i, \quad \text{and} \quad p_j = H^b(s_j) - H^b(s_{j-1}). \tag{3.18}$$

3.6.1 Expectation Maximization (EM)

The EM algorithm was first thoroughly studied in [26]. While it is typically used in parametric settings, it may also be used for non-parametric estimation. It is especially appealing due to its ease of implementation and its interpretation of incomplete data models. For the application of EM to our problem, we follow the description of EM in [113]. The authors describe the EM algorithm for problems similar to the one we are facing. The class of problems they consider is the following.

Suppose we aim to estimate a distribution function F. We cannot directly observe a sample X from F. Instead, we observe $Y = T(X, C)$, with $X \sim F$ and C some random variable independent of X. Clearly, given Lemma 3.1 the problem of estimating H^b belongs to this class of problems with $T(x, c) = xc$ and $C \sim g_K^S$. Suppose we have an initial estimate $H_0^b \in \mathcal{F}_n^+$ of the CDF H^b. Let $p^{(0)}$ be the associated probability vector as in (3.18). Let $X_1, \ldots, X_n \sim H^b$. In [113] it is shown that in their general context the EM algorithm yields the following update rule:

$$p_j^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{P}_{p^{(k)}} \left(X_i = s_j | s_1, \ldots, s_n \right). \tag{3.19}$$

We use the notation $\mathbb{P}_p$ to indicate the probability measure associated with the probability vector p. For our 'random fraction' setting, we use Bayes' rule to obtain:

$$\mathbb{P}_{p^{(k)}} \left(X_i = s_j | s_1, \ldots, s_n \right) = \frac{g_K^S \left(\frac{s_i}{s_j} \right) \frac{1}{s_j} p_j^{(k)}}{\sum_{q=1}^{n} g_K^S \left(\frac{s_i}{s_q} \right) \frac{1}{s_q} p_q^{(k)}}.$$

Plugging this into (3.19) yields:

$$p_j^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} \frac{\alpha_{i,j}}{\sum_{q=1}^{n} \alpha_{i,q} p_q^{(k)}} p_j^{(k)}, \quad \text{with: } \alpha_{i,j} = g_K^S \left(\frac{s_i}{s_j} \right) \frac{1}{s_j}. \tag{3.20}$$

When terminating the EM algorithm after an appropriate number of iterations we obtain $\hat{H}_n^b$ from $p^{(k)}$ via (3.18). The EM algorithm may for example be terminated when successive iterations do not meaningfully change the log-likelihood anymore. We do not provide

a specific stopping criterion for EM, since we do not use it directly. We only use it in hybrid form with the Iterative Convex Minorant algorithm (ICM) which is described in the next section. For the ICM algorithm and the hybrid ICM-EM algorithm we do provide explicit termination conditions.

3.6.2 Iterative Convex Minorant (ICM)

The ICM algorithm was first introduced in [38]. The version of ICM we discuss is described in [45] and is sometimes called the modified ICM algorithm. This modification of ICM ensures convergence under fairly general conditions. The algorithm is designed to minimize a convex function ϕ over the closed convex cone:

$$C_+ := \{\beta \in \mathbb{R}^n : 0 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_n\}.$$

Recall equation (3.12), computing the estimator $\hat{H}_n^b$ is equivalent to solving the following optimization problem:

$$\hat{\beta} \in \underset{\beta \in C}{\arg\max}\, l(\beta) = \underset{\beta \in C}{\arg\max}\, \frac{1}{n} \sum_{i=1}^{n} \log\left(\sum_{j=1}^{n} \alpha_{i,j} (\beta_j - \beta_{j-1}) \right). \tag{3.21}$$

With $\beta_j = H^b(s_j)$ and $\hat{\beta} = (\hat{H}_n^b(s_1), \ldots, \hat{H}_n^b(s_n))$. From (3.21) it is clear that for any $\beta \in C_+$ with $\beta_n < 1$, the likelihood can be increased by setting $\beta_n = 1$, since $\alpha_{i,j} \geq 0$. Hence, we may incorporate the constraint $\beta_n = 1$ instead of $\beta_n \leq 1$. We achieve this via a Lagrange multiplier. Define the convex function $\phi : C_+ \to \mathbb{R} \cup \{\infty\}$ as:

$$\phi(\beta) = -l(\beta) + \beta_n = -\frac{1}{n} \sum_{i=1}^{n} \log\left(\sum_{j=1}^{n} \alpha_{i,j} (\beta_j - \beta_{j-1}) \right) + \beta_n.$$

Hereby we have incorporated the constraint, with a Lagrange multiplier equal to one. Also, the problem is now written as a convex minimization problem since $\hat{\beta} \in \arg\min_{\beta \in C_+} \phi(\beta)$. Therefore, the ICM algorithm may be used to compute the MLE. Suppose we have some initial estimate $\beta^{(0)}$. The idea of ICM is to locally approximate ϕ with the following quadratic form in iteration k:

$$\begin{aligned} \phi_{(k)}(\beta) = \left(\beta - \beta^{(k)} + W\left(\beta^{(k)}\right)^{-1} \nabla\phi\left(\beta^{(k)}\right) \right)^{\mathsf{T}} W\left(\beta^{(k)}\right) \cdot \\ \cdot \left(\beta - \beta^{(k)} + W\left(\beta^{(k)}\right)^{-1} \nabla\phi\left(\beta^{(k)}\right) \right). \end{aligned} \tag{3.22}$$

The notation $\nabla\phi$ is used for the gradient of ϕ, the vector of partial derivatives of ϕ. The matrix W is a diagonal matrix, its diagonal is often chosen equal to the diagonal of the Hessian matrix of ϕ:

$$W\left(\beta^{(k)}\right) = \operatorname{diag}\left(\frac{\partial^2}{\partial \beta_j^2} \phi\left(\beta^{(k)}\right) \right).$$

In the ICM algorithm, $\phi_{(k)}$ is minimized over C_+ instead of ϕ to obtain a candidate β for $\beta^{(k+1)}$. If this candidate β sufficiently decreases ϕ it is accepted, and we set $\beta^{(k+1)} = \beta$. Otherwise, a line-search is performed to obtain $\beta^{(k+1)}$, which is then given by a convex combination of β and $\beta^{(k)}$. A precise description of the algorithm is given in section 3.11. We remark that minimizing (3.22) is equivalent to computing the weighted least-squares estimator of a monotone regression function. This can be done efficiently, for more details see [45]. The partial derivatives of ϕ are related to those of l as in (3.13) and (3.14), via:

$$\frac{\partial}{\partial \beta_j}\phi(\beta) = -\frac{\partial}{\partial \beta_j} l(\beta) + \mathbb{1}\{j = n\} \quad \text{and} \quad \frac{\partial^2}{\partial \beta_j^2}\phi(\beta) = -\frac{\partial^2}{\partial \beta_j^2} l(\beta).$$

For ICM we use the following stopping criterion, stop whenever:

$$\max_{j \in \{1,\ldots,n\}} \left| \beta_j^{(k)} - \beta_j^{(k-1)} \right| < \varepsilon, \tag{3.23}$$

for 10 successive iterations. In simulations we set $\varepsilon = 10^{-4}$. The interpretation of this criterion is that we stop whenever the largest change in probability mass is below ε for 10 successive iterations. We note that this criterion could be inappropriate if ICM approaches the optimum very slowly. In simulations (section 3.8) this was not an issue.

3.6.3 Hybrid ICM-EM

In [113] it was proposed to combine ICM and EM into a hybrid algorithm. The idea is that a single iteration of this hybrid algorithm consists of first performing one iteration of the ICM algorithm followed by one iteration of the EM algorithm. The ICM algorithm appears somewhat slow initially, if it is started far from the MLE, whereas close to the optimal value it converges quickly. On the contrary, the EM algorithm seems quicker at the start but has trouble converging when close to the optimum. Moreover, when performing an EM step after an ICM step, ICM ensures that many of the p_j's are zero. From (3.20) we see that EM will never set such a p_j to a positive value, hence EM only needs to operate in a lower dimensional space. In practice, it seems that the hybrid ICM-EM algorithm inherits the strengths of both algorithms and is quicker than both ICM and EM. This was for example observed in simulations in [46] and [113]. As with ICM, the same termination condition (3.23) is used.

3.7 Regularization of the maximum likelihood estimator

In this section, we describe how the MLE $\hat{H}_n^b$ may be used to estimate H, the distribution function of interest. At first glance, it seems reasonable to plug in $\hat{H}_n^b$ for H^b in equation (3.2). Unfortunately, simulations indicate that this yields a poor estimate of H. In section 3.8 we describe in detail how simulations are performed. For now, Figure 3.3 shows the result of a single simulation run. This simulation corresponds to the case where each particle is a dodecahedron, $n = 1000$, and H corresponds to a standard exponential distribution. For this H, H^b corresponds to a gamma distribution. The left panel of Figure

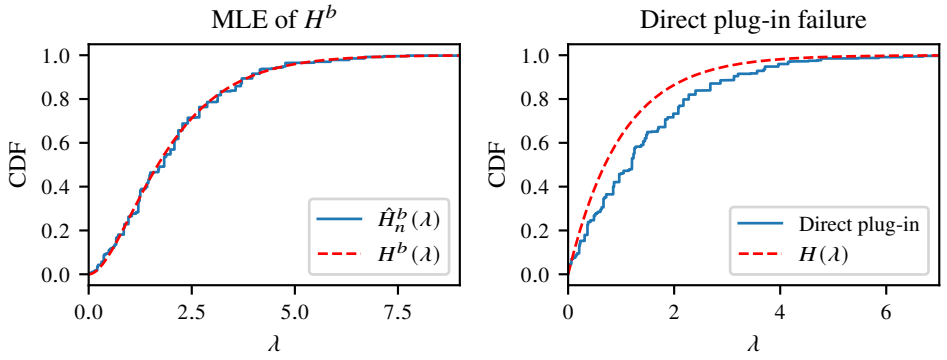


Figure 3.3: Left: MLE of H^b. Right: Direct plug-in estimate of H.

3.3 shows that $\hat{H}_n^b$ closely resembles H^b. Meanwhile, in the right panel of Figure 3.3 we observe that plugging in $\hat{H}_n^b$ for H^b in equation (3.2) yields a poor estimate of H. This is due to the influence of the behavior of $\hat{H}_n^b$ near zero.

We propose a regularization technique to resolve this issue. Let $t_n > 0$, truncating $\hat{H}_n^b$ at t_n yields:

$$\hat{H}_n^b(\lambda; t_n) := \begin{cases} \frac{\hat{H}_n^b(\lambda) - \hat{H}_n^b(t_n)}{1 - \hat{H}_n^b(t_n)} & \text{if } \lambda \geq t_n \\ 0 & \text{otherwise} \end{cases}.$$

Plugging this truncated version of $\hat{H}_n^b$ into (3.2) we obtain:

$$\hat{H}_n(\lambda; t_n) := \frac{\int_0^\lambda \frac{1}{x} \mathrm{d}\hat{H}_n^b(x; t_n)}{\int_0^\infty \frac{1}{x} \mathrm{d}\hat{H}_n^b(x; t_n)} = \begin{cases} \frac{\int_{t_n}^\lambda \frac{1}{x} \mathrm{d}\hat{H}_n^b(x)}{\int_{t_n}^\infty \frac{1}{x} \mathrm{d}\hat{H}_n^b(x)} & \text{if } \lambda \geq t_n \\ 0 & \text{if } 0 \leq \lambda < t_n \end{cases}. \tag{3.24}$$

Therefore, we introduce a new parameter t_n, which we refer to as the truncation parameter. In the following lemma, we show that for an appropriate choice of the truncation parameter t_n, a sequence of approximating CDFs converging to H^b may be de-biased to obtain a close approximation of H.

Lemma 3.6. *Let H be a continuous CDF on $(0, \infty)$, with finite first moment and length-biased version H^b. Let $(t_n)_{n\geq 1}$, $t_n > 0$ be a sequence such that $\lim_{n\to\infty} t_n = 0$. Let $(H_n^b)_{n\geq 1}$ be a sequence of CDFs. Assume H_n^b converges uniformly to H^b with rate at least t_n, that is: $\|H_n^b - H^b\|_\infty = o(t_n)$. Define:*

$$H_n(\lambda) = \begin{cases} \frac{\int_{t_n}^\lambda \frac{1}{x} \mathrm{d}H_n^b(x)}{\int_{t_n}^\infty \frac{1}{x} \mathrm{d}H_n^b(x)} & \text{if } \lambda \geq t_n \\ 0 & \text{if } 0 \leq \lambda < t_n \end{cases},$$

then: $\lim_{n\to\infty} \|H_n - H\|_\infty = 0$.

The proof is given in section 3.10. Lemma 3.6 shows that truncation is a viable approach for consistent estimation of H. Note that in Lemma 3.6, we may take $t_n = \sqrt{\|H_n^b - H^b\|_\infty}$. In practice the result cannot directly be applied to $\hat{H}_n^b$ since the quantity $\|\hat{H}_n^b - H^b\|_\infty$ is unknown. We propose a rule of thumb for t_n. Let $s \in \mathbb{R}$ and define:

$$\hat{F}_n^S(s;t) := \int_0^\infty G_K^S\left(\frac{s}{\lambda}\right) \mathrm{d}\hat{H}_n^b(\lambda;t) \tag{3.25}$$
$$\bar{F}_n^S(s) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{s_i \le s\}.$$

Note that $\hat{F}_n^S(\cdot\,;t)$ is the distribution function of observed square root section areas induced by the biased size distribution $\hat{H}_n^b(\cdot\,;t)$. That is, if $\hat{H}_n^b(\cdot\,;t)$ is the true biased size distribution, then $\hat{F}_n^S(\cdot\,;t)$ is the corresponding distribution of observed square root section areas. We propose the following choice for t_n:

$$\hat{t}_n := \operatorname*{arg\,min}_{t\in\{s_1,\dots,s_n\}} \int_0^\infty \left|\hat{F}_n^S(s;t) - \bar{F}_n^S(s)\right| \mathrm{d}s. \tag{3.26}$$

Hence, $\hat{t}_n$ minimizes the L^1-distance between the CDF of the observed square root section areas, induced by the estimated (biased) size distribution, and the empirical CDF of observed square root section areas. We minimize over $\{s_1, \dots, s_n\}$ for computational convenience. In practice, the integral in (3.26) can be computed via numerical integration.

3.8 Simulations

In the previous sections, we have introduced the MLE $\hat{H}_n^b$, and shown that under reasonable assumptions it is a consistent estimator of H^b. Also, a regularization technique was introduced to consistently estimate the size distribution function H using the MLE. In this section, some simulation results are presented to assess the performance of these estimators for H^b and H. The code used for the simulations may be found at https://github.com/thomasvdj/pysizeunfolder. Using this code the simulation and estimation procedure can be carried out in principle for any choice of convex polyhedron for the reference particle K.

Let us start by describing how to generate an iid sample of observed section areas, for a given H and a chosen reference particle K. Lemma 3.1 shows that it is sufficient to draw $Z \sim G_K$ and independently draw $\Lambda_b \sim H^b$, followed by setting $A := Z\Lambda_b^2$. A may be considered a random section area, and repeating these steps n times yields an iid sample $A_1, \dots, A_n$ distributed according to f_A. Taking the square root yields a sample of observed square root section areas. A sampling scheme for generating IUR planes through K is described in [25], see section 2.4 for sampling from G_K. Finally, we consider some well-known parametric distributions for H, for these choices H^b corresponds to some other well-known parametric distribution. Hence, drawing from H^b is straightforward. The following choices for H are considered, with the corresponding H^b:

1. *Exponential distribution*: For H we consider a standard exponential distribution, such that H^b corresponds to a gamma distribution.

$$H(\lambda) = 1 - e^{-\lambda}, \text{ and } H^b(\lambda) = 1 - (\lambda + 1)e^{-\lambda}, \ \lambda \geq 0.$$

2. *Lognormal distribution*: For H we consider a lognormal distribution with parameters μ and σ. For this H, H^b corresponds to a lognormal distribution with parameters $\mu + \sigma^2$ and σ. We set $\mu = 2$, $\sigma = 1/2$.

$$H(\lambda) = \Phi\left(\frac{\log(\lambda) - \mu}{\sigma}\right), \text{ and } H^b(\lambda) = \Phi\left(\frac{\log(\lambda) - \mu - \sigma^2}{\sigma}\right), \ \lambda > 0.$$

   Here, Φ denotes the CDF of a standard normal distribution.

For the simulations, we consider the following shapes for the particles: the dodecahedron, cube, and tetrahedron. As for the specific choice of the reference particle K, each of the shapes is scaled such that they have volume 1. Because these shapes are polyhedra,

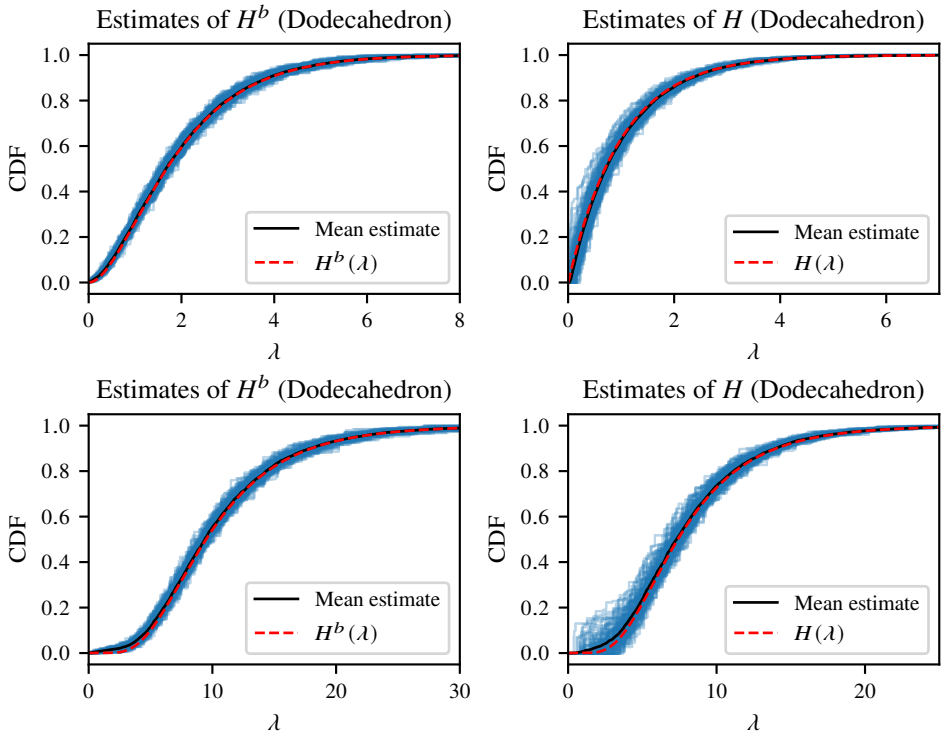


Figure 3.4: Simulation results for the dodecahedron, $n = 1000$. Top left: H^b is a gamma distribution. Top right: H is an exponential distribution. Bottom left: H is a lognormal distribution. Bottom right: H is a lognormal distribution.

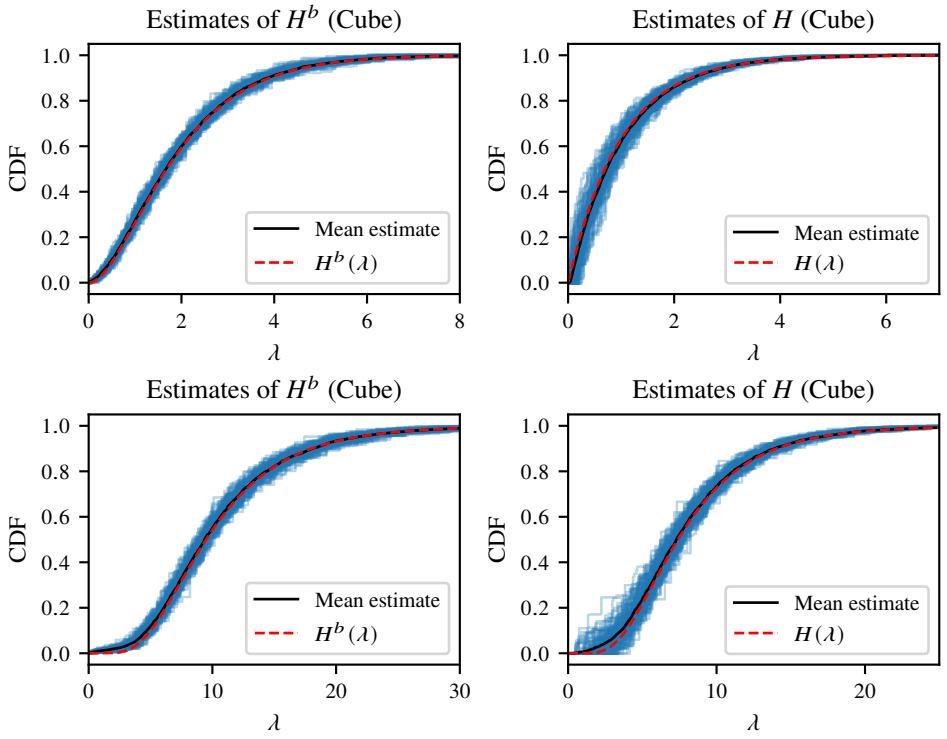


Figure 3.5: Simulation results for the cube, $n = 1000$. Top left: H^b is a gamma distribution. Top right: H is an exponential distribution. Bottom left: H is a lognormal distribution. Bottom right: H is a lognormal distribution.

each of the corresponding distribution functions G_K has a Lebesgue density by Theorem 3.1.

Now that we covered the simulation of iid samples we discuss the computation of estimators. For a given choice of n, H, and shape for the particles we generate a sample of n observed (square root) section areas. The MLE $\hat{H}_n^b$ is computed using the hybrid ICM-EM algorithm. The computation of the MLE requires that we can evaluate g_K^S in given points. As mentioned before, there is typically no explicit expression for g_K^S and we use the Monte Carlo simulation scheme described in section 2.4 for approximating g_K^S (recall Figure 3.2). For estimating H we compute $\hat{H}_n(\cdot, \hat{t}_n)$ as in (3.24), with $\hat{t}_n$ as in (3.26). Throughout this section, we refer to this estimator simply as $\hat{H}_n$. Note that for the computation of $\hat{t}_n$ we require G_K^S, which is also not explicitly known. Hence, similarly to g_K^S we use a Monte-Carlo approximation of G_K^S. In this case, we use an empirical distribution function based on the same sample used for approximating g_K^S.

We perform repeated simulations as follows. For various choices of n we generate a sample of n observed section areas. This is repeated 100 times for each choice of n, H, and shape for the particles. Simulation results for the dodecahedron and cube are shown

in figures 3.4 and 3.5 respectively. These results correspond to $n = 1000$. Each of the blue lines corresponds to one of the 100 estimates, each estimate based on a different

		$\|\hat{H}_n^b - H^b\|_\infty$		$\|\hat{H}_n - H\|_\infty$	
n	H	mean error	(2.5%, 97.5%)	mean error	(2.5%, 97.5%)
1000	Exponential	0.0577	(0.038, 0.076)	0.118	(0.064, 0.20)
1000	Lognormal	0.0657	(0.045, 0.10)	0.0924	(0.057, 0.18)
2000	Exponential	0.0452	(0.032, 0.063)	0.0972	(0.054, 0.17)
2000	Lognormal	0.0528	(0.034, 0.081)	0.0783	(0.044, 0.13)
5000	Exponential	0.0318	(0.024, 0.045)	0.0687	(0.036, 0.14)
5000	Lognormal	0.0380	(0.027, 0.054)	0.0586	(0.037, 0.097)
10000	Exponential	0.0260	(0.019, 0.035)	0.0578	(0.029, 0.12)
10000	Lognormal	0.0298	(0.023, 0.040)	0.0478	(0.028, 0.086)

Table 3.1: Simulation results for the dodecahedron.

		$\|\hat{H}_n^b - H^b\|_\infty$		$\|\hat{H}_n - H\|_\infty$	
n	H	mean error	(2.5%, 97.5%)	mean error	(2.5%, 97.5%)
1000	Exponential	0.0647	(0.045, 0.094)	0.134	(0.073, 0.23)
1000	Lognormal	0.0794	(0.055, 0.12)	0.107	(0.062, 0.17)
2000	Exponential	0.0509	(0.036, 0.070)	0.107	(0.059, 0.19)
2000	Lognormal	0.0630	(0.043, 0.087)	0.0911	(0.050, 0.17)
5000	Exponential	0.0394	(0.029, 0.053)	0.0787	(0.043, 0.13)
5000	Lognormal	0.0460	(0.033, 0.059)	0.0672	(0.040, 0.11)
10000	Exponential	0.0308	(0.022, 0.042)	0.0620	(0.036, 0.096)
10000	Lognormal	0.0368	(0.028, 0.047)	0.0544	(0.031, 0.091)

Table 3.2: Simulation results for the cube.

		$\|\hat{H}_n^b - H^b\|_\infty$		$\|\hat{H}_n - H\|_\infty$	
n	H	mean error	(2.5%, 97.5%)	mean error	(2.5%, 97.5%)
1000	Exponential	0.0948	(0.062, 0.15)	0.197	(0.090, 0.39)
1000	Lognormal	0.11	(0.078, 0.15)	0.163	(0.091, 0.30)
2000	Exponential	0.0792	(0.058, 0.10)	0.153	(0.085, 0.28)
2000	Lognormal	0.0930	(0.069, 0.13)	0.134	(0.082, 0.24)
5000	Exponential	0.0602	(0.046, 0.080)	0.120	(0.061, 0.25)
5000	Lognormal	0.0761	(0.058, 0.093)	0.0997	(0.068, 0.15)
10000	Exponential	0.0514	(0.038, 0.064)	0.101	(0.054, 0.20)
10000	Lognormal	0.0643	(0.048, 0.083)	0.0805	(0.059, 0.11)

Table 3.3: Simulation results for the tetrahedron.

	ICM		ICM-EM	
n	time (s)	# iterations	time (s)	# iterations
1000	3.99	314	0.511	26.4
2000	27.6	502	2.42	30.7
5000	415	784	25.8	62.5

Table 3.4: Algorithms mean run-times and mean number of iterations.

sample of size $n = 1000$. The black line is the point-wise average of all estimates. Further simulation results for the other shapes are summarized in Tables 3.1, 3.2 and 3.3. We quantify the error of the estimate as the supremum distance between the true H^b and $\hat{H}_n^b$, and similarly for the error of the estimates of H. The mean error is then the mean taken over the 100 resulting errors of the estimates. For these 100 resulting errors, the 2.5% and 97.5% quantiles are also shown.

Let us discuss the content of Tables 3.1, 3.2, and 3.3. As expected, as n increases the average error decreases, for all chosen shapes and size distributions, both for the estimates of H and H^b. Comparing the average supremum error for a fixed n, and a fixed size distribution, it is clear that the errors are smallest for the dodecahedron, followed by the cube and finally, the average error is largest for the tetrahedron. This is the case for both the average errors for estimating H as well as H^b. Note that estimating H instead of H^b increases the supremum error, and the corresponding mean supremum errors are also larger. These larger errors are also evident in figures 3.4 and 3.5. We note that for some practical applications, an estimate of H^b may be sufficient.

Finally, we briefly touch upon the computational efficiency of the algorithms for computing $\hat{H}_n^b$. We take for the shape of the particles the dodecahedron and for H the previously introduced lognormal distribution. In Table 3.4 the average run-times and iteration counts of the ICM and ICM-EM algorithms are shown, averaged over 10 simulation runs. The EM algorithm is not included in the table, in simulations it was several orders of magnitude slower than the other algorithms. Clearly, ICM-EM is considerably faster than ICM.

3.9 Concluding remarks

In this chapter we have studied a generalization of the classical Wicksell corpuscle problem, considering an arbitrary convex shape for the particles instead of spheres. In particular, for the problem of estimating the CDF H of the particle size distribution an identifiability result is derived. We also obtain an inversion formula via the Mellin transform. A nonparametric maximum likelihood estimator is proposed for the biased size distribution H^b and it is proven to be uniformly strongly consistent. Moreover, this estimator can be computed efficiently in practice. In a simulation study the proposed estimators for H^b and H perform well for various choices of particle shapes and particle size distributions.

3.10 Additional proofs

Proof of Lemma 3.2. The result follows almost immediately from theorem 7.8.2. in [48], which is a Mellin inversion theorem. Suppose $X \sim F_1$ and $Y \sim F_2$. By assumption, $\mathcal{M}_X$ and $\mathcal{M}_Y$ are analytic on $\mathrm{St}(\alpha, \beta)$, $0 \leq \alpha < \beta$. Let $c \in (\alpha, \beta)$, and assume $\mathcal{M}_X(c + it) = \mathcal{M}_Y(c + it)$ for all $t \in \mathbb{R}$. Let $x > 0$, by theorem 7.8.2. from [48] we obtain:

$$\bar{F}_1(x) := \frac{1}{2}(F_1(x+) + F_1(x-)) = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} -\mathcal{M}_X(s)\frac{x^{-s+1}}{s}\mathrm{d}s$$
$$\bar{F}_2(x) := \frac{1}{2}(F_2(x+) + F_2(x-)) = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} -\mathcal{M}_Y(s)\frac{x^{-s+1}}{s}\mathrm{d}s.$$

Here: $F(x+) := \lim_{h\downarrow 0} F(x+h)$ and $F(x-) := \lim_{h\uparrow 0} F(x+h)$. Note that for a continuity point x of F_1, $\bar{F}_1(x) = F_1(x)$. Because CDFs are right continuous we obtain: $F_1(x) = \bar{F}_1(x+)$ and $F_2(x) = \bar{F}_2(x+)$. Hence, it is sufficient to show $\bar{F}_1 = \bar{F}_2$. Because $\mathcal{M}_X(c+it) = \mathcal{M}_Y(c + it)$ for all $t \in \mathbb{R}$:

$$\bar{F}_1(x) - \bar{F}_2(x) = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} -(\mathcal{M}_X(s) - \mathcal{M}_Y(s))\frac{x^{-s+1}}{s}\mathrm{d}s = 0,$$

which finishes the proof. □

Proof of Lemma 3.3. f_ϵ is upper semicontinuous, as it is given by a product, and a composition of an upper semicontinuous function and a continuous function. By Theorem 3.1, g_K^S is non-decreasing on $(0, \tau_K)$ for some $0 < \tau_K \leq \sqrt{a_{\max}}$. Choose $M > \sqrt{a_{\max}}$ large enough such that $e^{-M} < \tau_K$. It now immediately follows that $f_\epsilon(z) = 0$ for $z \in [M, \infty)$ and f_ϵ is monotonically increasing on $(-\infty, -M]$. It remains to show that f_ϵ is of bounded variation on $[-M, M]$. Let $-M < z_0 < z_1 < \cdots < z_m < M$ be an arbitrary partition of $[-M, M]$. Then it follows:

$$\begin{aligned}
&\sum_{i=1}^{m} |f_\epsilon(z_i) - f_\epsilon(z_{i-1})| = \\
&\quad = \sum_{i=1}^{m} \left| g_K^S(e^{z_i})e^{z_i} - g_K^S(e^{z_i})e^{z_{i-1}} + g_K^S(e^{z_i})e^{z_{i-1}} - g_K^S(e^{z_{i-1}})e^{z_{i-1}} \right| \\
&\quad \leq \|g_K^S\|_\infty \sum_{i=1}^{m} |e^{z_i} - e^{z_{i-1}}| + e^M \sum_{i=1}^{m} \left| g_K^S(e^{z_i}) - g_K^S(e^{z_{i-1}}) \right| \qquad (3.27) \\
&\quad \leq \|g_K^S\|_\infty e^M + e^M V_0^{\sqrt{a_{\max}}}\left(g_K^S\right) < \infty.
\end{aligned}$$

Note that the first sum in (3.27) telescopes. In the final step we use the fact that g_K^S is bounded and is of bounded variation on its support. Because the above computation holds for arbitrary partitions of $[-M, M]$ we find: $V_{-M}^{M}(f_\epsilon) < \infty$, which finishes the proof. □

Proof of Lemma 3.5. Because $f_\epsilon \geq 0$ is a Lebesgue density, for every $m \in \mathbb{N}$ there exists a bounded continuous probability density function f_ϵ^m such that $\|f_\epsilon^m - f_\epsilon\|_{L^1} \leq 1/m$ (see

Lemma 3.7 in section 3.10). Let $m \in \mathbb{N}$, then:

$$\begin{aligned}\|f_n - f_Y\|_{L_1} &= \int \left| \int f_\epsilon(z-x) - f_\epsilon^m(z-x) + f_\epsilon^m(z-x)\mathrm{d}(F_n - F_X)(x) \right| \mathrm{d}z \\ &\leq \int \left| \int f_\epsilon(z-x) - f_\epsilon^m(z-x)\mathrm{d}(F_n - F_X)(x) \right| \mathrm{d}z \\ &\quad + \int \left| \int f_\epsilon^m(z-x)\mathrm{d}(F_n - F_X)(x) \right| \mathrm{d}z. \end{aligned} \tag{3.28}$$

Via the triangle inequality and Fubini, the first term in (3.28) is bounded by:

$$\begin{aligned} &\int \left| \int f_\epsilon(z-x) - f_\epsilon^m(z-x)\mathrm{d}F_n(x) \right| \mathrm{d}z \\ &\quad + \int \left| \int f_\epsilon(z-x) - f_\epsilon^m(z-x)\mathrm{d}F_X(x) \right| \mathrm{d}z \\ &\leq \int \int \left| f_\epsilon(z-x) - f_\epsilon^m(z-x) \right| \mathrm{d}z\mathrm{d}F_n(x) \\ &\quad + \int \int \left| f_\epsilon(z-x) - f_\epsilon^m(z-x) \right| \mathrm{d}z\mathrm{d}F_X(x) \leq 2\|f_\epsilon^m - f_\epsilon\|_{L^1} \leq \frac{2}{m}. \end{aligned}$$

The second term in (3.28) may be written as:

$$\int \left| \int f_\epsilon^m(z-x)\mathrm{d}(F_n - F_X)(x) \right| \mathrm{d}z = \|\varphi_{n,m} - \varphi_m\|_{L^1},$$

with $\varphi_{n,m}$ and φ_m defined as:

$$\varphi_{n,m}(z) = \int f_\epsilon^m(z-x)\mathrm{d}F_n(x), \quad \text{and} \quad \varphi_m(z) = \int f_\epsilon^m(z-x)\mathrm{d}F_X(x).$$

Because f_ϵ^m is a probability density, so are φ_m and $\varphi_{n,m}$ for all $n \in \mathbb{N}$. By the continuity of f_ϵ^m and the weak convergence of F_n to F_X we obtain that $\varphi_{n,m}$ converges pointwise to φ_m as $n \to \infty$. By Scheffé's Theorem pointwise convergence of probability densities to another probability density implies that these densities also converge in L^1. Combining all results yields:

$$\lim_{n\to\infty} \|f_n - f_Y\|_{L_1} \leq \lim_{n\to\infty} \frac{2}{m} + \|\varphi_{n,m} - \varphi_m\|_{L^1} = \frac{2}{m}.$$

Letting $m \to \infty$ we obtain the desired result. □

Lemma 3.7. *Let f be a Lebesgue density on $\mathbb{R}$, for every $\varepsilon > 0$ there exists a bounded continuous probability density function g such that $\|g - f\|_{L^1} < \varepsilon$.*

Proof. Recall that the space of compactly supported continuous functions is dense in L^1. For $n \in \mathbb{N}$ choose a continuous, compactly supported and non-negative function g_n such that $\|g_n - f\|_{L^1} \leq 1/(n+1)$. By the reverse triangle inequality:

$$|\|g_n\|_{L^1} - 1| = |\|g_n\|_{L^1} - \|f\|_{L^1}| \leq \|g_n - f\|_{L^1} \leq \frac{1}{n+1}. \tag{3.29}$$

Define: $\tilde{g}_n = g_n / \|g_n\|_{L^1}$. Note that by (3.29), $\|g_n\|_{L^1} > 0$. Hence, $\tilde{g}_n$ is a bounded and continuous probability density function. Combining all results:

$$\begin{aligned}\|\tilde{g}_n - f\|_{L^1} &= \frac{1}{\|g_n\|_{L^1}} \int |g_n(x) - f(x) + f(x) - \|g_n\|_{L^1} f(x)| \, \mathrm{d}x \\ &\leq \frac{1}{\|g_n\|_{L^1}} \left(\|g_n - f\|_{L^1} + |\|g_n\|_{L^1} - 1| \cdot \|f\|_{L^1}\right) \\ &\leq \frac{\frac{1}{n+1} + \frac{1}{n+1}}{1 - \frac{1}{n+1}} = \frac{2}{n}.\end{aligned}$$

Because this holds for all $n \in \mathbb{N}$ we obtain the desired result. □

Proof of Lemma 3.6. We first note the following:

$$\sup_{0 \leq \lambda < t_n} |H_n(\lambda) - H(\lambda)| = \sup_{0 \leq \lambda < t_n} H(\lambda) = H(t_n). \tag{3.30}$$

Let us now assume $\lambda \geq t_n$. By definition:

$$H(\lambda) - H_n(\lambda) = \frac{\int_0^\lambda \frac{1}{x} \mathrm{d}H^b(x) \int_{t_n}^\infty \frac{1}{x} \mathrm{d}H_n^b(x) - \int_{t_n}^\lambda \frac{1}{x} \mathrm{d}H_n^b(x) \int_0^\infty \frac{1}{x} \mathrm{d}H^b(x)}{\int_0^\infty \frac{1}{x} \mathrm{d}H^b(x) \int_{t_n}^\infty \frac{1}{x} \mathrm{d}H_n^b(x)}. \tag{3.31}$$

The numerator of (3.31) may be written as:

$$\begin{aligned}&\int_{t_n}^\infty \frac{1}{x} \mathrm{d}H_n^b(x) \left(\int_0^\lambda \frac{1}{x} \mathrm{d}H^b(x) - \int_{t_n}^\lambda \frac{1}{x} \mathrm{d}H_n^b(x) \right) \\ &\quad - \int_{t_n}^\lambda \frac{1}{x} \mathrm{d}H_n^b(x) \left(\int_0^\infty \frac{1}{x} \mathrm{d}H^b(x) - \int_{t_n}^\infty \frac{1}{x} \mathrm{d}H_n^b(x) \right) \\ &= \int_{t_n}^\infty \frac{1}{x} \mathrm{d}H_n^b(x) \left(\int_{t_n}^\lambda \frac{1}{x} \mathrm{d}(H^b - H_n^b)(x) + \int_0^{t_n} \frac{1}{x} \mathrm{d}H^b(x) \right) \\ &\quad - \int_{t_n}^\lambda \frac{1}{x} \mathrm{d}H_n^b(x) \left(\int_{t_n}^\infty \frac{1}{x} \mathrm{d}(H^b - H_n^b)(x) + \int_0^{t_n} \frac{1}{x} \mathrm{d}H^b(x) \right).\end{aligned} \tag{3.32}$$

Recall: $\mathbb{E}(\Lambda) = \int_0^\infty \lambda \mathrm{d}H(\lambda) = 1 / \int_0^\infty (1/x) \mathrm{d}H^b(x)$. Plugging (3.32) back into (3.31) yields:

$$\begin{aligned}H(\lambda) - H_n(\lambda) = \mathbb{E}(\Lambda) &\left(\int_{t_n}^\lambda \frac{1}{x} \mathrm{d}(H^b - H_n^b)(x) \right) + H(t_n) \\ &- \mathbb{E}(\Lambda) H_n(\lambda) \left(\int_{t_n}^\infty \frac{1}{x} \mathrm{d}(H^b - H_n^b)(x) \right) - H_n(\lambda) H(t_n).\end{aligned}$$

Therefore, we obtain the following bound:

$$\sup_{\lambda \geq t_n} |H(\lambda) - H_n(\lambda)| \leq 2\mathbb{E}(\Lambda) \sup_{\lambda \geq t_n} \left| \int_{t_n}^\lambda \frac{1}{x} \mathrm{d}(H^b - H_n^b)(x) \right| + H(t_n). \tag{3.33}$$

The integral in (3.33) may be computed via integration by parts:

$$\begin{aligned}
&\sup_{\lambda \geq t_n} \left| \int_{t_n}^{\lambda} \frac{1}{x} \mathrm{d}(H^b - H_n^b)(x) \right| = \\
&\quad = \sup_{\lambda \geq t_n} \left| \frac{H_n^b(\lambda) - H^b(\lambda)}{\lambda} - \frac{H_n^b(t_n) - H^b(t_n)}{t_n} - \int_{t_n}^{\lambda} \left(H_n^b(x) - H^b(x) \right) \mathrm{d}\frac{1}{x} \right| \\
&\quad \leq 2 \frac{\sup_{\lambda \geq t_n} \left| H_n^b(\lambda) - H^b(\lambda) \right|}{t_n} + \sup_{\lambda \geq t_n} \left| H_n^b(\lambda) - H^b(\lambda) \right| \cdot \left| \int_{t_n}^{\lambda} \mathrm{d}\frac{1}{x} \right| \\
&\quad \leq 3 \frac{\| H_n^b - H^b \|_\infty}{t_n}. \qquad (3.34)
\end{aligned}$$

Note that the bound in (3.33) is greater than $H(t_n)$, by (3.30) this means that the bound also holds when taking the supremum over $\lambda \geq 0$ instead. Combining (3.33) and (3.34) we finally obtain:

$$\| H_n - H \|_\infty \leq 6 \mathbb{E}(\Lambda) \frac{\| H_n^b - H^b \|_\infty}{t_n} + H(t_n). \qquad (3.35)$$

Letting n go to infinity, $H(t_n)$ converges to zero by the continuity of H. Using this and the fact that $(H_n^b)_{n \geq 1}$ converges uniformly to H^b with rate t_n (by assumption) the RHS of (3.35) converges to zero. □

3.11 Pseudo-code of algorithms

Algorithm 1 Expectation Maximization (EM)

Input: Observed order statistics: $s_1 < s_2 < \cdots < s_n$.
Output: The MLE $\hat{H}_n^b$.

$k := 0$
$p^{(0)} := (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) \in \mathcal{P}_n$
while Stopping criterion is not met **do**
    $p_j^{(k+1)} := \frac{1}{n} \sum_{i=1}^{n} \frac{\alpha_{i,j}}{\sum_{q=1}^{n} \alpha_{i,q} p_q^{(k)}} p_j^{(k)}, \quad$ with: $\alpha_{i,j} = g_K^S \left(\frac{s_i}{s_j} \right) \frac{1}{s_j}$
    $k := k + 1$
end while
$\hat{H}_n^b(s_j) := \sum_{i=1}^{j} p_i^{(k)}$ for $j \in \{1, \ldots, n\}$.
return $\hat{H}_n^b$

Algorithm 2 Iterative Convex Minorant (ICM)

Input: A convex function $\phi : C_+ \to \mathbb{R} \cup \{\infty\}$. $\beta^{(0)}$ with $\phi(\beta^{(0)}) < \infty$ and $\epsilon \in (0, 1/2)$.
Output: A minimizer of ϕ.

$k := 0$
$\beta^{(0)} := (\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}) \in C$
while Stopping criterion is not met **do**
  $\beta := \arg\min_{y \in C_+} \phi_{(k)}(y)$ ▷ With $\phi_{(k)}$ as in (3.22)
  if $\phi(\beta) < \phi(\beta^{(k)}) + \epsilon \nabla\phi(\beta^{(k)})^\intercal (\beta - \beta^{(k)})$ **then**
    $\beta^{(k+1)} := \beta$
  else
    $\lambda := 1$, $s := \frac{1}{2}$, $z := \beta$.
    while $\phi(z) < \phi(\beta^{(k)}) + (1 - \epsilon)\nabla\phi(\beta^{(k)})^\intercal (z - \beta^{(k)})$ (I) **or**
        $\phi(z) > \phi(\beta^{(k)}) + \epsilon \nabla\phi(\beta^{(k)})^\intercal (z - \beta^{(k)})$ (II) **do**
      if (I) **then** $\lambda := \lambda + s$
      if (II) **then** $\lambda := \lambda - s$
      $z := \beta^{(k)} + \lambda(\beta - \beta^{(k)})$
      $s := \frac{s}{2}$
    end while
    $\beta^{(k+1)} := z$
  $k := k + 1$
end while
return $\beta^{(k)}$

Chapter 4

Stereological estimation of grain size distributions in microstructures

4.1 Introduction

In this chapter, the statistical methodology proposed in chapter 3 is applied to polycrystalline materials, as well as mathematical models representing these materials. Polycrystalline materials, composed of multiple grains with distinct crystallographic orientations, exhibit intricate microstructures that significantly influence their mechanical properties [77]. Different types of microstructure-based simulations depend on proper estimation and characterization of the microstructure [80]. For example, a thorough physical understanding of the underlying mechanisms behind phenomena such as local stress fields [64], [49], fracture and damage initiation [24], [81], [102], shear banding [47], [93], [91], and recrystallization nucleation [50], [103], [96] depends on the initial microstructure. In addition to the microscale mechanical response, the macroscopic mechanical response, such as the stress-strain curve and yield surface, also depends on the microstructure [4], [10], [43], [90], [92], [98]. Therefore, accurate characterization of the microstructure of polycrystalline materials is essential and has become a crucial part of materials science research.

One of the critical microstructural features is the grain size distribution, which can affect many mechanical responses [4], [112]. For example, at a macro scale, it can affect the stress-strain curve. At a smaller scale, it can influence the stress and strain localization [1], [93], impacting for instance damage evolution. However, direct 3D measurement of the grain size distribution is costly and time-consuming [76]. Instead, more common two-dimensional (2D) characterization techniques, such as light microscopy, Scanning Electron Microscopy (SEM), and Electron Backscatter Diffraction (EBSD), are employed. These 2D techniques provide only surface information, leaving part of the critical information, like the 3D grain size distribution, unknown. This chapter aims to use 2D information obtained from section areas to estimate the 3D grain size distribution, offering a more efficient and accessible approach to microstructure characterization. Estimating the 3D grain size distribution from 2D observations is a well-known stereological problem, originally

addressed by [114], in the Wicksell corpuscle problem. In the Wicksell model, grains are represented by spheres of varying size, randomly positioned in 3D space. Intersecting this system of spheres with a plane results in a sample of observed circle radii, the distribution of which is uniquely related to the distribution of the 3D sphere radii. However, when considering real microstructures, the assumption of (approximately) spherical grains is often unrealistic. A more realistic approach is to generalize the Wicksell model, by replacing spheres with another convex shape, such as a polyhedron, as was done in chapter 3. Such a polyhedron then represents the typical grain shape. Instances of this polyhedron are randomly scaled, -oriented and -positioned in 3D space. Intersecting this system with a plane yields observed section profiles, the areas of which can be used to estimate the grain size distribution.

Various estimation procedures have been developed for these stereological problems. For spherical grains (Wicksell's problem), the Saltykov method [83] is widely used, estimating the underlying radius distribution as a discrete histogram, requiring bin size choices. Numerous variations of the Saltykov method exist, often differing in binning strategies, inversion procedures, or the choice of parametric distributions [58], [21], [12]. Other methods extend these principles to non-spherical grains, such as cubes or other specific polyhedra [70], [59], [61], [116]. For a more elaborate overview of estimators in the Wicksell problem we refer to [18].

The estimation procedure in chapter 3, offers a discrete distribution estimate for grain size via non-parametric maximum likelihood estimation. This method, when the shape of the 3D objects is known and fixed, is consistent, meaning that as the sample size increases, the estimate converges to the true distribution. Unlike the Saltykov method, which typically uses a fixed number of bins, this new method implicitly optimizes the number and size of bins through likelihood maximization, eliminating the need for manual binning. The estimation method from chapter 3 was developed (and theoretically studied) from the perspective of randomly placed, scaled and oriented particles in a 3D medium. In practice, this type of procedures is commonly applied to space-filling structures, as often seen in metal microstructures. As this estimation method works for arbitrary convex shapes, it enables to investigate which of these shapes actually works best in specific situations with space-filling structures.

In this chapter we estimate grain size distributions of various simulated- and real microstructures. We explore a range of grain shapes to assess how the chosen shape relates to the estimation accuracy of the actual grain size distributions. In the simulation setting, we consider Voronoi and the more general Laguerre-Voronoi diagrams as models for metal microstructures. These diagrams provide realistic approximations of grain shapes and distributions found in metals [72], making them ideal for studying the accuracy of 3D grain size distribution estimates derived from 2D cross-sectional images. In fact, though several theoretical properties of these models are known in the literature [62], [56], [40], there are no explicit relations or estimators of the 3D volume distributions from the 2D sections. Hence, applying the estimation procedure from chapter 3 in a simulation setting can provide insights into the behavior of the estimator for the underlying model. The contributions in this chapter are the following: i) we estimate the 3D grain size distribution using a novel statistical estimation procedure; ii) we investigate the influence of assumed grain shapes on the estimation of grain size distributions; iii) we conduct extensive sim-

ulations using random 3D (Laguerre)-Voronoi microstructures and single planar sections to estimate the underlying grain size distribution; and iv) we show that for the considered microstructures, a few grain shape choices yield accurate estimates, providing a practical procedure for selecting appropriate shapes for steel samples.

The outline of this chapter is as follows. In section 4.2 we describe the estimation procedure used for obtaining the 3D grain size distributions. Furthermore, we illustrate the methods used for the simulation of microstructures, as well as how cross sections are taken of these simulated microstructures. In section 4.3, the simulation results for 100 Laguerre Voronoi diagrams are presented. Estimates based on different shapes for the grains are compared and in section 4.4 a heuristic of the choice of the best shape is discussed. In section 4.5 the new estimation procedure is applied to real data. In this special case, the results of the volume distribution estimation based on 2D real data can be compared with the real 3D volume distribution obtained from 3D EBSD data. Final considerations and conclusions are discussed in section 4.6 and 4.7.

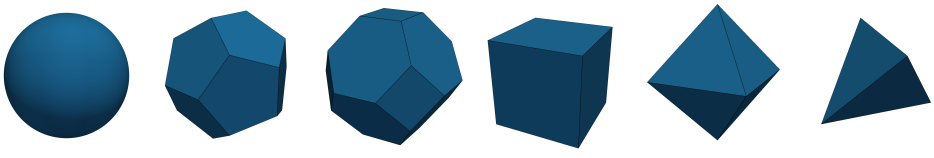
4.2 Methods

In this section, we describe how we estimate grain size distributions (subsection 4.2.1) using the methodology from chapter 3. Additionally, we describe the simulation procedure of the microstructures and the process of obtaining cross sections from these simulations (subsection 4.2.2).

4.2.1 Estimation of grain size distributions

Assume we have a sample of observed section areas: $a_1, a_2, \ldots, a_n$, with n being the sample size. This is assumed to be the sorted sample, meaning that: $a_1 < a_2 < \cdots < a_n$. We assume a particular 3D grain shape $K \subset \mathbb{R}^3$, representing the typical grain in the microstructure at hand. Consider taking a random section of the chosen shape K, with K scaled such that it has volume 1. The probability density function associated with the square-root of the area of such a random section is denoted by g_K^S. In principle g_K^S becomes known once the shape is chosen; in practice we use the simulation scheme from section 2.4 to obtain a very close approximation of this function. The estimation procedure can be described by two steps. First, we estimate the biased size distribution, denoted by H^b, both for mathematical as well as computational convenience. This can be interpreted as the distribution of the size of the typical grain which appears in the section plane. It is well known that larger grains are more likely to be hit by the section plane, meaning that larger grains are over represented in the plane section. Therefore, the actual grain size distribution is different, and it is estimated in a second step via a de-biasing procedure. Denote by H the size distribution function. Then, in step 1 we compute the estimator $\hat{H}_n^b$ for H^b, as defined in (3.10). In step 2 we use $\hat{H}_n^b$ to compute the estimator $\hat{H}_n$ for H, as defined via equations (3.24) and (3.26). For more details on the estimators $\hat{H}_n^b$ and $\hat{H}_n$ we refer to chapter 3.

A grain with size $\lambda > 0$ is up to a translation and rotation equal to λK. By λK we mean that K is scaled with a factor λ. As such, a grain with size λ has volume: $\text{Volume}(\lambda K) =$



(a) (b) (c) (d) (e) (f)

Figure 4.1: (a): Sphere. (b): Dodecahedron. (c): Kelvin cell/ Tetrakaidecahedron. (d) Cube. (e): Octahedron. (f): Tetrahedron.

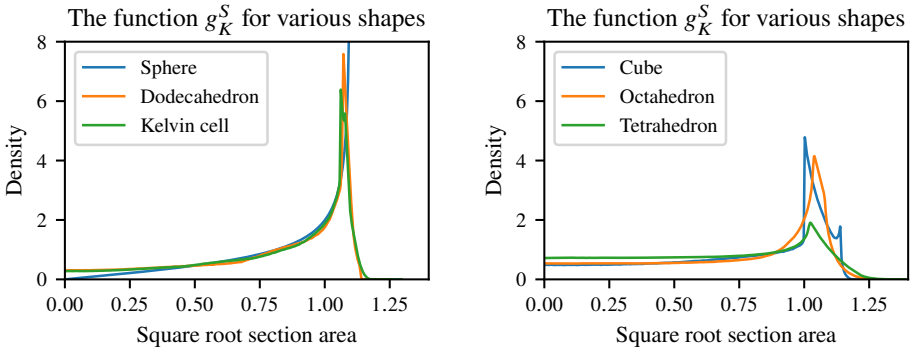


Figure 4.2: The function g_K^S for the shapes shown in Figure 4.1.

$\lambda^3 \text{Volume}(K) = \lambda^3$. As a result, the size distribution is related to the volume distribution function F_V via: $F_V(x) = H(x^{1/3})$. Hence, we estimate F_V as: $\hat{F}_V(x) = \hat{H}_n(x^{1/3})$. Additionally, the so-called biased volume distribution is given by $F_V^b(x) = H^b(x^{1/3})$ and may be estimated via $\hat{F}_V^b(x) = \hat{H}_n^b(x^{1/3})$. For simulations, we consider the shapes shown in Figure 4.1. For each shape K, we need the function g_K^S to carry out the estimation procedure. For all the considered shapes the function g_K^S is shown in Figure 4.2.

4.2.2 Simulation of random microstructures

We first describe the model chosen for our simulations. For studying the behavior of the estimator described in the previous section we run simulations using Voronoi diagrams as a mathematical model for microstructures. Voronoi diagrams and its generalizations are often referred as the state of the art for modelling microstructures [72]. Given some convex domain Ω in 3D space, a Voronoi diagram divides Ω into so-called cells, which are convex polyhedra. Given distinct points $x_1, \ldots, x_N \in \Omega$ and denoting by $\|\cdot\|$ the Euclidean norm, the Voronoi diagram generated by these points has cells $C_1, \ldots, C_N$ with:

$$C_i = \{x \in \Omega : \|x - x_i\| \leq \|x - x_j\|, \text{ for all } j \in \{1, \ldots, N\}\}.$$

In this chapter, we consider the Poisson-Voronoi diagram, meaning that the x_i's are a realization of a homogeneous Poisson process Φ on Ω. While Voronoi diagrams are attractive models for materials microstructures, the additional flexibility of its generalization,

the Laguerre-Voronoi diagram, also referred as Laguerre diagram, allows to more accurately model real microstructures. In [29] it was observed that geometric characteristics of Laguerre diagrams were closer to geometric characteristics of real polycrystalline microstructures in comparison to Voronoi diagrams. In [55] and [57] Laguerre diagrams were demonstrated to accurately model a foam microstructure. Moreover, [55] showed that Laguerre diagrams provided a better representation of foams compared to various types of Voronoi diagrams. The results in [28] indicate that Laguerre diagrams provide a superior representation for sintered alumina than Voronoi diagrams. Additionally, [115] demonstrated that Laguerre diagrams can accurately model two-phase composites.

Given some convex domain Ω in 3D space, a Laguerre diagram also divides Ω into cells, which are convex polyhedra. Given distinct points $x_1, \ldots, x_N \in \Omega$ and weights: $w_1, \ldots, w_N \in \mathbb{R}$. A Laguerre diagram generated by these weighted points has cells $L_1, \ldots, L_N$ with:

$$L_i = \{x \in \Omega : \|x - x_i\|^2 - w_i \leq \|x - x_j\|^2 - w_j, \text{ for all } j \in \{1, \ldots, N\}\}.$$

The Voronoi diagram is obtained if all weights are equal: $w_1 = w_2 = \cdots = w_N$. Hence, in order to describe how Laguerre diagrams are generated we need to specify how we choose the weights. The Laguerre diagrams considered in this chapter can be considered a sneak peek into part II of this thesis, where we will more thoroughly study a specific type of random Laguerre diagrams. A Laguerre diagram with periodic boundary conditions may be obtained by replacing the Euclidean distance $\|\cdot\|$ with a periodic distance.

We now describe the simulation setting. First, we define the domain as the unit cube $\Omega = [0,1] \times [0,1] \times [0,1]$. We fix a number of grains N, and generate a Poisson process conditioned on having N grains, that is equivalent to sample N uniformly distributed points: $x_1, \ldots, x_N$ in Ω. We choose a volume distribution function F_V and sample $v_1, \ldots, v_N \overset{\text{iid}}{\sim} F_V$. Then, we set:

$$r_i = \frac{v_i}{\sum_{i=1}^{N} v_i},$$

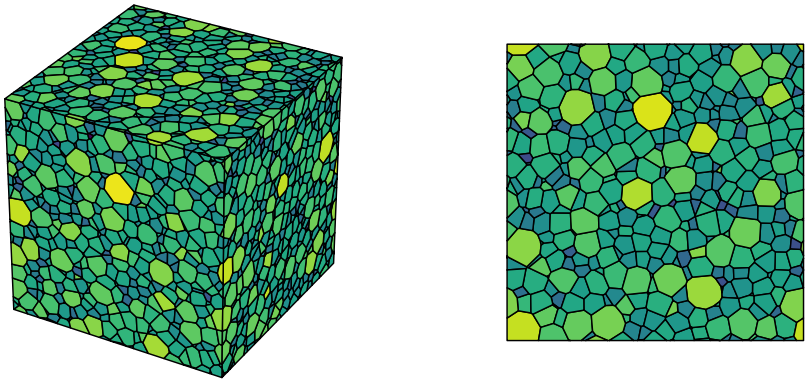


Figure 4.3: Example of a 3D Laguerre-Voronoi diagram with periodic boundary conditions, and a 2D planar section. Cells are colored according to their 3D volume.

such that the r_i's represent the volume fractions. Using the algorithm proposed in [15] (Algorithm 2), we generate a Laguerre diagram in Ω with n grains such that grain i has volume r_i. This algorithm is initialized with generator points $x_1, \ldots, x_N$. The final Laguerre diagram is approximately centroidal, meaning that the generator point of each cell is close to the center of mass of its cell.

Having generated a Laguerre diagram we take a random height z, sampled from the uniform distribution on $[0, 1]$, and intersect the diagram with a horizontal plane at height z. Throughout, the diagrams we consider have periodic boundary conditions. If parts of a cell appear in a section multiple times (due to periodic boundary conditions) the areas of the parts are added together, and this sum is considered as a single observed area.

4.3 Simulation results

4.3.1 Laguerre diagrams

In this section, we apply the estimation procedure to randomly generated Laguerre diagrams. We apply the procedure described in subsection 4.2.2 100 times. By this we mean that 100 times, a Laguerre diagram is generated, a planar section is taken, and estimates of the grain volume distribution are computed under various shape assumptions.

For the following simulations we generate Laguerre diagrams with $N = 50000$ grains. We choose for the volume distribution F_V a lognormal distribution with parameters $\sigma = 0.4$, $\mu = -\sigma^2/2$. Each of the 100 runs generated a random sample of observed section areas whose sample sizes are shown in the left panel of Figure 4.4.

The distribution of the observed section areas for one simulation is shown in the right panel of Figure 4.4. For each of the shapes we consider, the simulation results are given in Figures 4.5-4.6. Each blue line is an estimate corresponding to one of the 100 generated samples of section areas.

Looking at the estimates of the volume distribution, the estimates corresponding to the sphere and the tetrahedron are quite poor (Figures 4.5 (b) and 4.6 (f)). One may argue

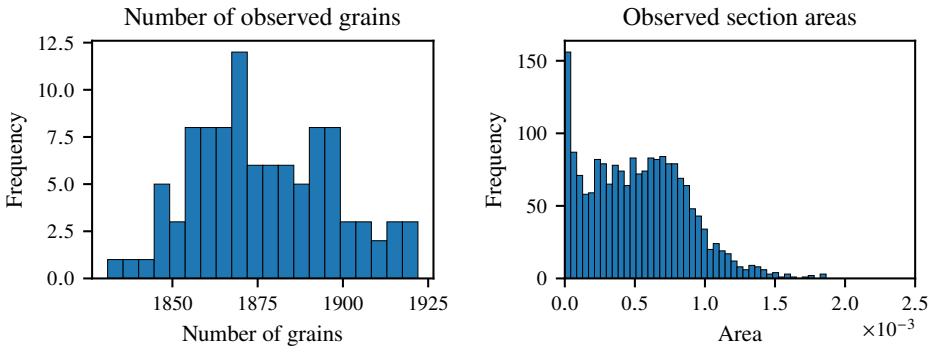
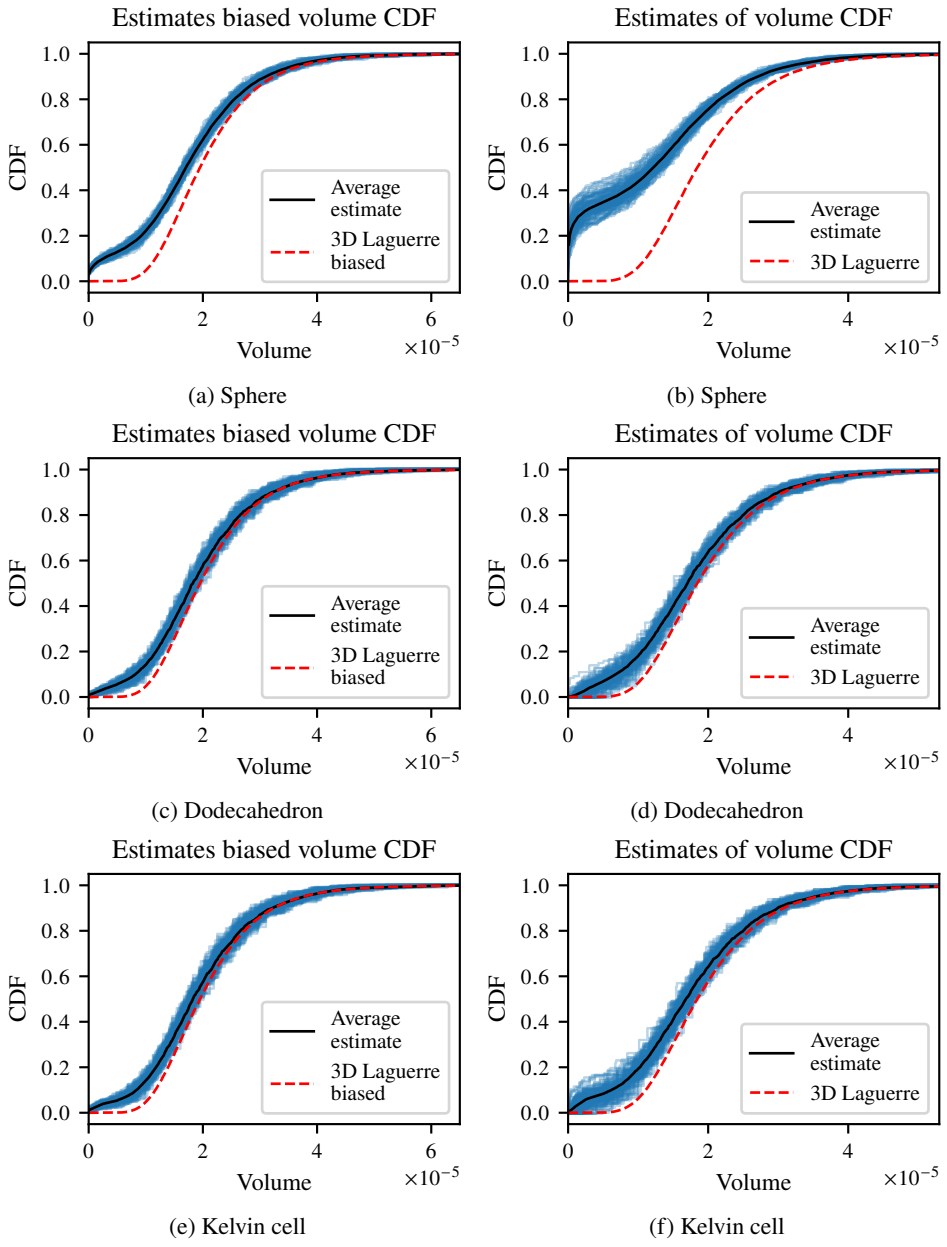


Figure 4.4: Left: Histogram of the number of observed grains for each of the 100 simulations. Right: Histogram of the observed section areas from one simulation.

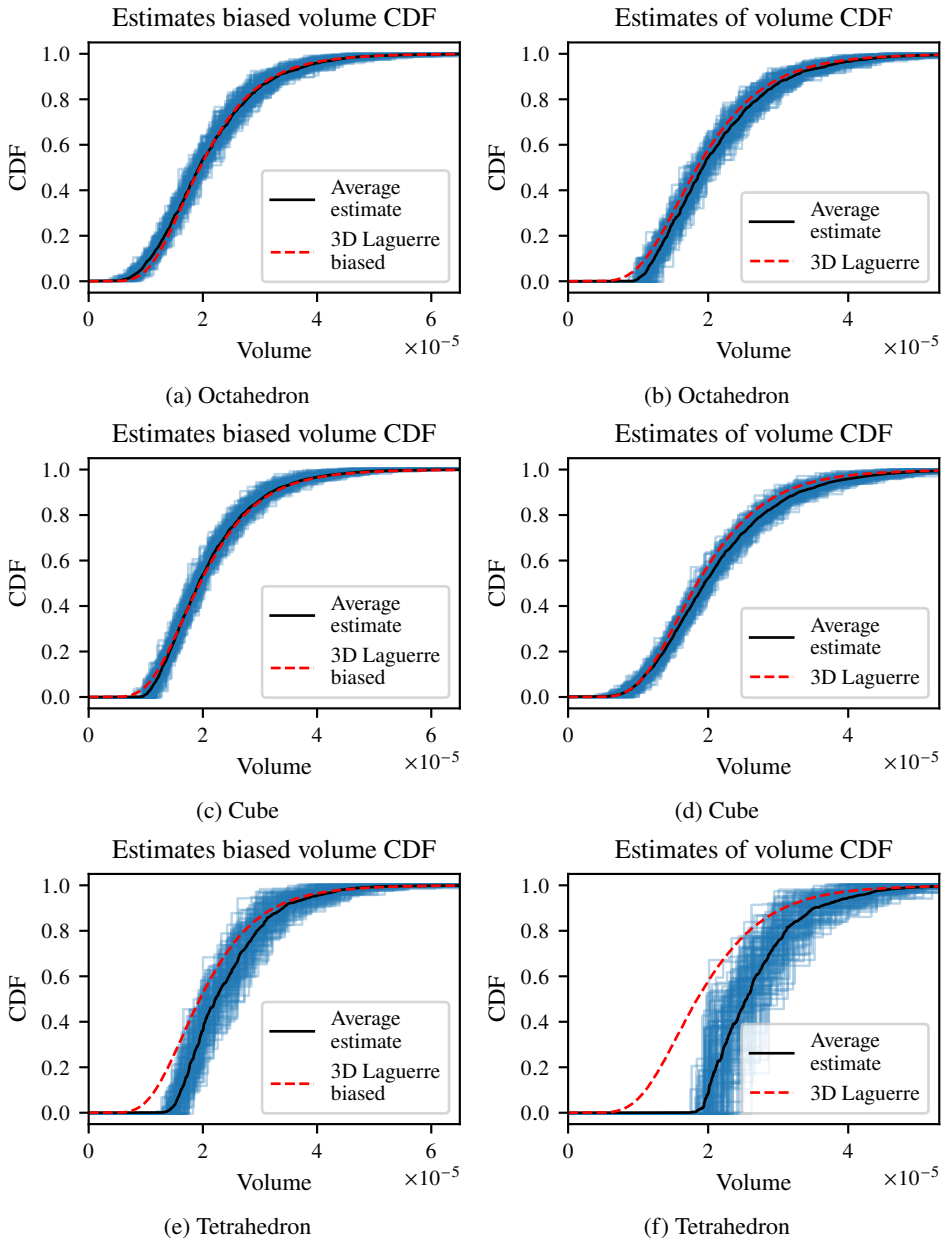


(a) Sphere (b) Sphere

(c) Dodecahedron (d) Dodecahedron

(e) Kelvin cell (f) Kelvin cell

Figure 4.5: Estimates of the biased volume distribution function (left) and estimates of the volume distribution function (right), based on the sphere, dodecahedron and Kelvin cell. The red dashed line represents F_V^b and F_V in the left and right panel respectively.



(a) Octahedron (b) Octahedron

(c) Cube (d) Cube

(e) Tetrahedron (f) Tetrahedron

Figure 4.6: Estimates of the biased volume distribution function (left) and estimates of the volume distribution function (right), based on the octahedron, cube and tetrahedron. The red dashed line represents F_V^b and F_V in the left and right panel respectively.

that a sphere is 'too round' to be a good representation of a typical grain shape and the tetrahedron has 'too sharp corners' to be representative of a real grain. All other shapes yield much better estimates, in particular the estimates corresponding to the octahedron and the cube (Figures 4.6 (b) and 4.6 (d)) appear quite close to the true volume distribution. The estimates corresponding to the biased volume distribution paint a similar, but slightly different picture. Note in particular that the variance of the estimates is rather different for each of the shapes. This variance is the smallest for the simulation results corresponding to the sphere, and largest for the results corresponding to the tetrahedron (Figures 4.5 (a) and 4.6 (e)).

In Tables 4.1-4.2 errors of the estimates in Figures 4.5 and 4.6 are shown. To be precise, the supremum error

$$\|\hat{F}_V - F_V\|_\infty := \sup_{x>0} \left|\hat{F}_V(x) - F_V(x)\right|.$$

and the L_1 error

$$\|\hat{F}_V - F_V\|_{L_1} := \int_0^\infty \left|\hat{F}_V(x) - F_V(x)\right| \mathrm{d}x.$$

are considered.

	$\|\hat{F}_V - F_V\|_\infty$		$\|\hat{F}_V^b - F_V^b\|_\infty$	
Shape	mean error	(2.5%, 97.5%)	mean error	(2.5%, 97.5%)
sphere	0.391	(0.34 , 0.46)	0.196	(0.17 , 0.22)
dodecahedron	0.167	(0.13 , 0.21)	0.143	(0.12 , 0.17)
Kelvin cell	0.182	(0.14 , 0.23)	0.146	(0.12 , 0.17)
octahedron	0.154	(0.10 , 0.22)	0.118	(0.089, 0.16)
cube	0.134	(0.098, 0.17)	0.114	(0.081, 0.16)
tetrahedron	0.619	(0.52 , 0.75)	0.345	(0.28 , 0.41)

Table 4.1: Mean supremum errors of estimates for the size- and biased size distribution function with 2.5% and 97.5% quantiles based on 100 simulations.

	$\|\hat{F}_V - F_V\|_{L_1}$ ($\times 10^{-6}$)		$\|\hat{F}_V^b - F_V^b\|_{L_1}$ ($\times 10^{-6}$)	
Shape	mean error	(2.5%, 97.5%)	mean error	(2.5%, 97.5%)
sphere	7.52	(6.37, 8.97)	3.51	(3.11 , 4.01)
dodecahedron	2.11	(1.42, 2.91)	1.91	(1.57 , 2.33)
Kelvin cell	2.41	(1.56, 3.74)	1.93	(1.61 , 2.34)
octahedron	1.46	(1.07, 1.97)	1.20	(1.00 , 1.45)
cube	1.45	(1.02, 1.99)	1.10	(0.850, 1.42)
tetrahedron	7.08	(5.29, 9.41)	3.54	(2.91 , 4.15)

Table 4.2: Mean L_1 errors of estimates for the size- and biased size distribution function with 2.5% and 97.5% quantiles based on 100 simulations.

The results in Tables 4.1-4.2 confirm what was found by the graphical inspection of Figures 4.5 and 4.6. For the volume distribution function, both the supremum and the L_1 error suggest that the cube and the octahedron are the best choice. What can be seen from both tables is that the sphere, the canonically used shape, leads to inferior approximations.

4.4 The choice of grain shape

In the previous section we have observed that some choices of grain shape yield much better results than other choices. We attempt to obtain a better understanding of why in particular the cube and the octahedron are often a good choice when the true data generating mechanism is a centroidal Laguerre diagram. To obtain preliminary results without making arbitrary choices on the distribution of weights, in this section Poisson-Voronoi diagrams that are Laguerre diagrams with all weights equal are used.

4.4.1 Characterizing shape via sphericity

Given a shape with volume V and surface area S, the isoperimetric inequality states: $S \geq \pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}$. Equality holds if and only if the shape is a sphere. It may also be stated as: among all shapes with a given surface area, a sphere has the maximum volume. In [111] this was used to define sphericity as:

$$\Psi = \frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{S}.$$

Then, $\Psi = 1$ for a sphere and for any other shape $0 < \Psi < 1$. In [111] the sphericity of quartz particles was studied. It is challenging to determine when two grains have approximately the same shape. One approach is to consider grains with approximately equal sphericity as being close in shape. The use of sphericity is in any case arbitrary. Other

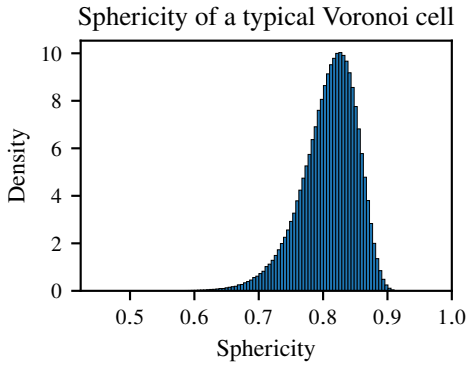


Figure 4.7: Simulated distribution of sphericity of a typical Poisson-Voronoi cell, based on 10^6 Voronoi cells.

Shape	Sphericity
Sphere	1
Dodecahedron	0.910
Kelvin cell	0.910
Octahedron	0.846
Mean typical Voronoi cell	0.808
Cube	0.806
Tetrahedron	0.671

Table 4.3: Approximate sphericity of various convex shapes.

shape parameters such as the number of facets or vertices are other reasonable choices. Because we apply the estimation procedure to space-filling microstructures, one may wonder what is the average sphericity of a grain in such a microstructure. As previously mentioned, as an example we consider a Poisson-Voronoi diagram, where all weights are equal. In [62] and [94] the sphericity has been used with other cell characteristics to describe the Poisson-Voronoi and the more general Laguerre diagram, respectively. Via simulations, we can generate Poisson-Voronoi cells and computing the sphericities of the individual cells yields the distribution in Figure 4.7.

In Table 4.3 the sphericity of various convex shapes is given, as well as the estimated mean sphericity of a typical Voronoi cell. Clearly, among all considered shapes the sphericity of the cube is closest to the mean sphericity of a typical Voronoi cell.

4.4.2 Verifying the choice of grain shape using the disector

A classical stereological technique for estimating the expected number of grains per unit volume N_V is called the disector [100]. It may also be used to estimate the mean grain volume since $\mathbb{E}(V) = 1/N_V$, for a space filling structure consisting of N_V cells in a unit volume body. The disector allows for unbiased estimation of N_V without assumptions on the grain shape. It requires two (close) parallel sections which are a known distance apart. Ideally, we would like to guarantee that no grains are lost between section planes, such that there are no grains between the two planes that we cannot observe. Let:

- Q^+: the number of grains which are observed in the top section but not in the lower section.
- Q^-: the number of grains which are observed in the lower section but not in the top section.
- A: the area of the observation window/ section plane.
- h: the distance between the section planes.

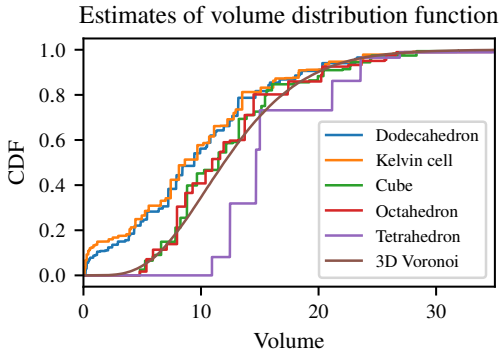


Figure 4.8: Estimates of the volume distribution function of the cells of a Poisson-Voronoi diagram based on different shapes.

When counting grains in a section plane it is important to deal with edge effects, as boundary grains are only partially observed. For obtaining an unbiased estimate, one may for instance use the Gundersen frame or the associated point rule (see section 3.3.5 in [9] and references therein) to determine which grains should be counted. Then, N_V may be estimated via:

$$\hat{N}_V = \frac{Q^+ + Q^-}{2Ah}.$$

Let us first proceed as before, we have various shapes we can consider to estimate the grain volume distribution using a single section. The result of the simulation of one Poisson-Voronoi diagram is shown in Figure 4.8. This particular realization of the Poisson-Voronoi diagram has 78862 cells.

For each of the estimates we can also compute the mean volume. We have estimates of the distribution function F_V denoted by $\hat{F}_V$. Then, the mean volume corresponding to $\hat{F}_V$ is given by:

$$\int_0^\infty x \mathrm{d}\hat{F}_V(x).$$

In Table 4.4 the estimates for all the shapes previously considered are shown. In Figure 4.9 mean grain volume estimates of the Poisson-Voronoi diagram are shown. These estimates

Shape	Estimated mean volume
Dodecahedron	9.959
Kelvin cell	9.621
Cube	12.35
Octahedron	12.35
Tetrahedron	16.20

Table 4.4: Estimated mean volume of the cells of a Poisson-Voronoi diagram corresponding to the estimates shown in Figure 4.8 based on different shapes.

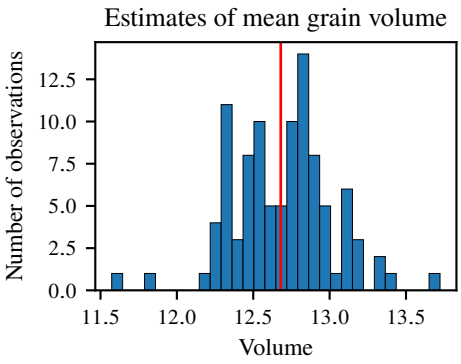


Figure 4.9: Histogram of mean grain volume estimates of a Poisson-Voronoi diagram obtained using the disector. The red line indicates the actual mean volume.

are obtained using the disector method using 100 pairs of parallel section planes. The actual mean volume (red line in Figure 4.9) is 12.68, which among the considered shapes, is closest to mean grain volume estimates obtained using the cube and the octahedron.

In the context of this chapter, we propose the disector method not as a tool for selecting an appropriate grain shape, but rather as a diagnostic approach to validate the grain shape assumptions made during the estimation process. In practice, one can take parallel sectional planes from a steel sample and then apply the disector technique to estimate the mean grain volume without any shape assumption. By comparing the estimated mean volumes from both a shape assumption and the disector method, which should ideally be close, one can validate whether a specific shape is a reasonable assumption for the given sample.

4.5 Application to experimentally measured EBSD data

In this section, we investigate an experimentally measured microstructure obtained using the Electron Backscatter Diffraction (EBSD) technique. The initial microstructure and crystallographic texture of the material were measured across the thickness (ND - normal direction) perpendicular to the rolling direction (RD). The EBSD scan area is 500 μm $\times$ 500 μm. Standard metallographic techniques were used to prepare the specimen for characterization. Analysis of the EBSD data was performed using TSL OIM software. The material used in this study is Interstitial-Free (IF) steel. For this example, we have the 3D EBSD information available using the serial sectioning technique [76].

Following a standard postprocessing procedure, and discarding the small grains located at grain boundaries, we obtained a sample of 1506 fully observed grains. A histogram of the observed section areas and a histogram of the observed grain diameters $((\frac{4\,\text{Area}}{\pi})^{1/2})$ is shown in Figure 4.11.

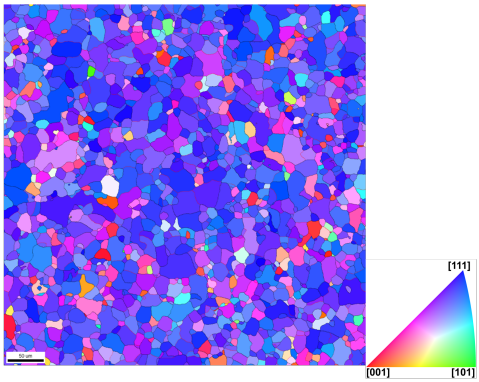


Figure 4.10: Electron backscatter diffraction (EBSD) measurements of an IF steel sample. The figure shows the IPF color map parallel to the normal direction for a section of 500μm $\times$ 500μm.

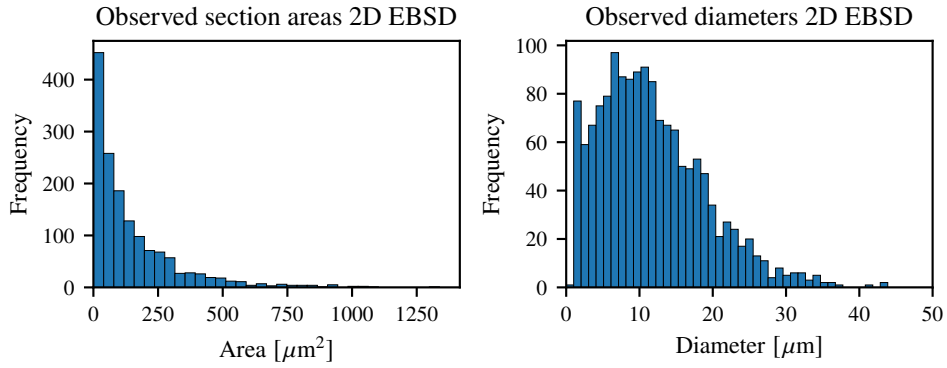


Figure 4.11: Left: Observed grain areas in 2D EBSD data set. Right: Observed grain diameters in 2D EBSD data set.

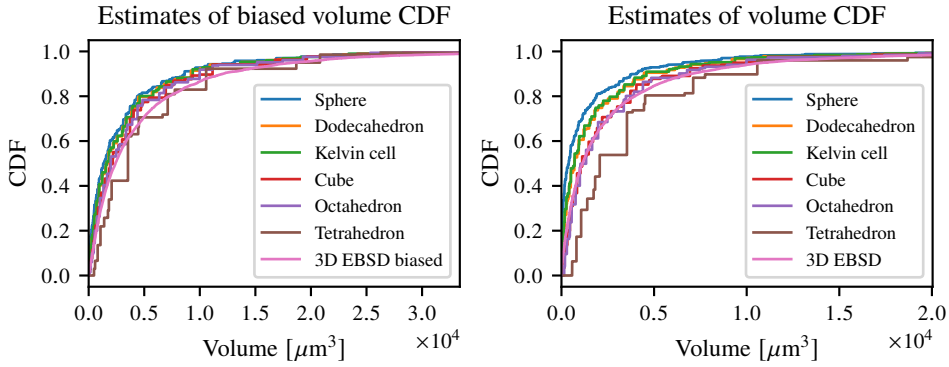


Figure 4.12: Estimates of the biased volume distribution function (left) and estimates of the volume distribution function (right) based on different shapes.

For all the shapes considered, we estimate the volume distribution function. These estimates can be directly compared to the experimental volume distribution in the 3D EBSD data set (Figure 4.12). From the comparison of the different shapes shown in Figure 4.12 as in the Laguerre Voronoi simulation the cube and octahedron appear to be the best shapes.

4.6 Discussion

This chapter critically addresses the problem of estimating 3D grain size distributions from 2D cross-sections, highlighting the importance of selecting an appropriate grain shape when applying the estimation procedure to space-filling microstructures. Our findings, supported by simulations using Laguerre Voronoi diagrams and a real-world data set, demonstrate that the choice of grain shape significantly affects the accuracy of the stere-

ological estimates. Notably, our results suggest that while traditional shapes like spheres and tetrahedrons often provide poor estimates probably due to their geometric simplicity or complexity, shapes such as cubes and octahedrons yield more reliable results.

The use of the disector method, not as a selection tool for grain shapes but as a diagnostic tool, has utility potential. It provides a baseline for validating the assumptions made during the estimation process, ensuring that the chosen grain shape assumptions are reasonable for the given sample.

While the study has provided valuable insights into estimating 3D grain size distributions from 2D sections, there are natural challenges to be addressed in the future. One is related to the fact that our estimation method is inspired by a model considering 'randomly sized and -oriented shapes randomly positioned in the 3D medium' and applied to space-filling structures. It is interesting to develop methods really based on models (like Voronoi, Laguerre Voronoi or newly developed models such as generalized balanced power diagrams [2]) leading to space-filling structures. In part II of this thesis we aim to make some progress in this research direction by studying so-called Poisson-Laguerre tessellations.

The assumption of isotropy in the grain structures is another issue. The models considered assume that the morphological properties of the grains are uniform in all directions, which is often not the case in real-world materials. Many materials exhibit anisotropic behavior due to directional cooling, applied stresses, or processing methods that align the grains in particular orientations. Ignoring anisotropy can lead to significant deviations between the estimated and actual grain size distributions.

As it comes to the simulated data, the use of periodic boundary conditions is a practical approach to manage computational boundaries. However, this assumption may not accurately reflect the true edge conditions of real materials. In natural or manufactured materials, the boundary effects can significantly influence the microstructural features near the edges, which are not captured by periodic boundary conditions. This can skew the estimation of grain size distributions, especially for materials where edge effects are pronounced.

Sphericity, as well as other shape parameters, may help in shedding some light on why some shapes work better than others. However, in the current status, the estimation procedure considered in this chapter does not incorporate grain shape information as an input, instead it is specific for the assumed grain shape. As a future development, the inclusion of shape parameters as an input of an estimation procedure must be explored.

Finally, both the sphericity and the disector method look promising but they need to be validated using real data. Having sectional data and data on the 3D grain surface areas, one can use the disector method and the sphericity measure as diagnostic tools to validate the grain shape assumption made during the estimation process.

4.7 Concluding remarks

The results of this chapter highlight the necessity for sensible grain shape selection in the estimation of 3D grain size distributions from 2D data. By using shapes that more accurately represent the microstructural characteristics of the material, such as cubes and octa-

hedrons, we achieved more accurate stereological estimates. Future work should focus on enhancing these methodologies by integrating models that incorporate space-filling characteristics, anisotropy and more realistic boundary conditions. Developing these advanced models will enable more accurate and generally applicable tools in materials science, contributing to the understanding and characterization of complex materials.

Part II

Poisson-Laguerre tessellations

Chapter 5

Nonparametric inference for Poisson-Laguerre tessellations

5.1 Introduction

In this part of the thesis we shift our focus towards tessellations, which may be used as space-filling models for microstructures. Tessellations have proven to be useful in a wide range of fields. For example, a Poisson-Voronoi tessellation may serve as a model for a wireless network [7]. In cosmology, Voronoi tessellations can be used to describe the distribution of galaxies [104]. There are several generalizations of the Voronoi tessellation, such as the Laguerre tessellation, which offers more flexibility compared to the Voronoi model. In the field of materials science, Laguerre tessellations have been fitted to various microstructures of materials. For instance, Laguerre tessellations were found to be accurate models for foams [55], [57], sintered alumina [28] and composites [115]. As we have established by now, a major challenge in this field is that in practice often only 2D microscopic images of cross sections of the 3D microstructure can be obtained. By studying a 3D object via a 2D slice there is evidently a loss of information.

In this chapter, we focus on statistical inference for a particular class of random tessellations known as Poisson-Laguerre tessellations. We do this both for the case where one directly observes a tessellation as well as for the case where the observed tessellation is obtained by intersecting a higher dimensional tessellation with a hyperplane. The latter type of tessellation is often referred to as a sectional tessellation. A Laguerre tessellation in $\mathbb{R}^d$ is defined via a set of weighted points $\eta = \{(x_1, h_1), (x_2, h_2), \dots\}$, called generators. Here, x_i is a point in $\mathbb{R}^d$ and $h_i > 0$ its weight. Each generator corresponds to a set, which is either a polytope or the empty set. This set is usually called a cell and we may also say that a generator generates this cell. The non-empty cells form a tessellation, meaning that these cells have disjoint interiors and the union of these cells equals $\mathbb{R}^d$. We refer to the subset $\eta^* \subset \eta$ of points which generate non-empty cells as the extreme points of η. We may write:

$$\eta^* := \{(x, h) \in \eta : C((x, h), \eta) \neq \emptyset\}, \tag{5.1}$$

where $C((x, h), \eta)$ denotes the cell associated with (x, h). A Poisson-Laguerre tessella-

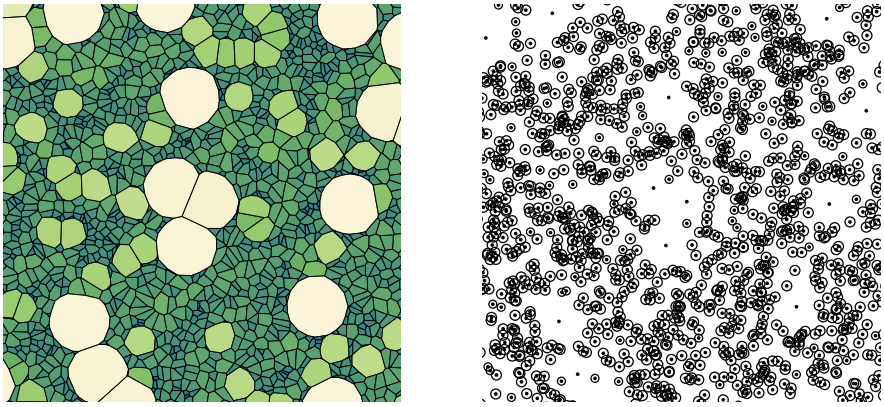


Figure 5.1: Left: A realization of a planar Poisson-Laguerre tessellation. Cells are colored according to their area. Right: The corresponding realization of extreme points. Around each point there is a circle with radius proportional to the weight of the point.

tion, which is a random tessellation, is obtained by taking η to be a Poisson (point) process on $\mathbb{R}^d \times (0, \infty)$. The intensity measure of η is assumed to be of the form $\nu_d \times \mathbb{F}$. Here, ν_d is Lebesgue measure on $\mathbb{R}^d$ and $\mathbb{F}$ is a non-zero locally finite measure concentrated on $(0, \infty)$. An example of a realization of a Poisson-Laguerre tessellation, and the corresponding realization of extreme points is shown in Figure 5.1. Random Laguerre tessellations generated by an independently marked Poisson process were first studied in [54] and [56]. We mostly follow the description of Poisson-Laguerre tessellations as given in [39]. Additionally, we will also rely on the result from [39] which states that the sectional Poisson-Laguerre tessellation is again a Poisson-Laguerre tessellation. The so-called β-Voronoi tessellation as introduced in [40] may be seen as a parametric model for a Poisson-Laguerre tessellation. In [41] it was shown that the sectional Poisson-Voronoi tessellation is in fact a β-Voronoi tessellation.

Because tessellations are usually not directly observed in nature, typically the first step towards statistical inference for tessellations is a reconstruction step. Such a reconstruction method is used to obtain a tessellation from an image, for details see section 9.10.1 in [18] and references therein. Therefore, when applying the methodology in this chapter to real data, it needs to be combined with such a reconstruction method. It is important to point out that the reconstruction methods used in [55], [57] and [95] reconstruct a Laguerre tessellation along with the extreme points simultaneously. Effectively, statistical inference for a Poisson-Laguerre tessellation is then reduced to statistical inference for the point process η^* as in (5.1). This appears to be the most common approach towards statistical inference for random Laguerre tessellations, and this is also the approach we take. For instance, in [94] a methodology is proposed for statistical inference for Laguerre tessellations, where parametric models are considered for the underlying point process. In [101], a Laguerre tessellation, along with the corresponding extreme points, is fitted to real data. Furthermore, a statistical analysis is performed on this point process of extreme

points.

Recall that the intensity measure of the underlying Poisson process η is assumed to be of the form $\nu_d \times \mathbb{F}$. For $z \geq 0$ we define $F(z) := \mathbb{F}((0, z])$, the distribution function of $\mathbb{F}$. Note that this distribution function is the only parameter in this model to be estimated. In this chapter, we define nonparametric estimators for F. These estimators for F depend on both the observed Laguerre cells in a bounded observation window as well as the points of η^* in the same window. Here, it is important to realize that η^* is not necessarily a Poisson process, as it is a dependent thinning of η. The proposed estimators are proven to be consistent as the observation window expands unboundedly to the whole of $\mathbb{R}^d$. Additionally, we consider the stereological setting where the observed Poisson-Laguerre tessellation in $\mathbb{R}^{d-1}$ is obtained by intersecting a Poisson-Laguerre tessellation in $\mathbb{R}^d$ with a hyperplane. Based on this observed sectional tessellation we introduce an estimator for the distribution function corresponding to the Poisson process of the higher dimensional tessellation.

This chapter is organized as follows. In section 5.2 we introduce necessary notation and definitions. Then, the main mathematical object of interest, the Poisson-Laguerre tessellation, is discussed in section 5.3. In section 5.4 we introduce our first estimator for F, which is based on a thinning of the extreme points. To the best of our knowledge, no estimators have been proposed in the context of Poisson-Laguerre tessellations as of yet. A second estimator for F is introduced in section 5.5, which depends on all observed extreme points, as well as the volumes of the corresponding Laguerre cells. In section 5.6 we consider statistical inference for Poisson-Laguerre tessellations in a stereological setting. In section 5.7 we perform a simulation study for the proposed estimators, to empirically verify their behavior. Based on the estimates observed in the simulation study we provide some intuition for their behavior in section 5.8. Finally, we conclude this chapter with a discussion in section 5.9.

5.2 Preliminaries

In this section we introduce notation and various definitions which we need throughout this chapter. Let ν_d denote Lebesgue measure on $\mathbb{R}^d$, and σ_{d-1} Lebesgue measure on the sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$, also known as the spherical measure. Given $x \in \mathbb{R}^d$ and $r > 0$, we write $B(x, r) = \{y \in \mathbb{R}^d : \|x-y\| < r\}$ and $\bar{B}(x, r) = \{y \in \mathbb{R}^d : \|x-y\| \leq r\}$ for the open and closed ball respectively, with radius r centered at x. We introduce the following constant:

$$\kappa_d := \nu_d\left(\bar{B}(0, 1)\right) = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(1 + \frac{d}{2}\right)}.$$

We may also use the following fact: $\sigma_{d-1}\left(\mathbb{S}^{d-1}\right) = d\kappa_d$. Let $A, B \subset \mathbb{R}^d$, then the sum of sets is defined as: $A + B = \{a + b : a \in A, b \in B\}$. If $x \in \mathbb{R}^d$, we also write: $A + x = \{a + x : a \in A\}$. Let $\mathcal{F}_+$ denote the space of all (not necessarily bounded) distribution functions on $(0, \infty)$.

We now introduce several definitions related to point processes. While these definitions are valid for point processes in much more general spaces, in this chapter we only

consider point processes on $\mathbb{R}^d \times (0,\infty)$. For more background on the theory of point processes we refer to Appendix A and references therein. Suppose $\mathbb{X} = \mathbb{R}^d \times (0,\infty)$, and let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A measure μ on $\mathbb{X}$ is locally finite if $\mu(B) < \infty$ for all bounded $B \in \mathcal{B}(\mathbb{X})$. Here, $\mathcal{B}(\mathbb{X})$ denotes the Borel σ-algebra of $\mathbb{X}$. Let $\mathbf{N}(\mathbb{X})$ denote the space of locally finite counting measures (integer-valued measures) on $\mathbb{X}$. We equip $\mathbf{N}(\mathbb{X})$ with the usual σ-algebra $\mathcal{N}(\mathbb{X})$, which is the smallest σ-algebra on $\mathbf{N}(\mathbb{X})$ such that the mappings $\mu \mapsto \mu(B)$ are measurable for all $B \in \mathcal{B}(\mathbb{X})$. A point process on $\mathbb{X}$ is a random element η of $(\mathbf{N}(\mathbb{X}), \mathcal{N}(\mathbb{X}))$, that is a measurable mapping $\eta : \Omega \to \mathbf{N}(\mathbb{X})$. The intensity measure of a point process η on $\mathbb{X}$ is the measure Λ defined by $\Lambda(B) := \mathbb{E}(\eta(B))$, $B \in \mathcal{B}(\mathbb{X})$.

Definition 5.1. Suppose Λ is a σ-finite measure on $\mathbb{X}$. A Poisson process with intensity measure Λ is a point process η on $\mathbb{X}$ with the following two properties:

1. For every $B \in \mathcal{B}(\mathbb{X})$, the random variable $\eta(B)$ is Poisson distributed with mean $\Lambda(B)$.
2. For every $m \in \mathbb{N}$ and pairwise disjoint sets $B_1, \dots, B_m \in \mathcal{B}(\mathbb{X})$, the random variables $\eta(B_1), \dots, \eta(B_m)$ are independent.

Let δ denote the Dirac measure, hence for $x \in \mathbb{X}$ and $B \in \mathcal{B}(\mathbb{X})$: $\delta_x(B) = \mathbb{1}\{x \in B\}$. A counting measure μ on $\mathbb{X}$ is called simple if $\mu(\{x\}) \leq 1$ for all $x \in \mathbb{X}$. As such, a simple counting measure has no multiplicities. Similarly, a point process η on $\mathbb{X}$ is called simple if $\mathbb{P}(\eta(\{x\}) \leq 1,\ \forall x \in \mathbb{X}) = 1$. Let $\mathbf{N}_s(\mathbb{X})$ be the subset of $\mathbf{N}(\mathbb{X})$ containing all simple measures. Define: $\mathcal{N}_s(\mathbb{X}) := \{A \cap \mathbf{N}_s(\mathbb{X}) : A \in \mathcal{N}(\mathbb{X})\}$. Then, a simple point process on $\mathbb{X}$ may be seen as a random element η of $(\mathbf{N}_s(\mathbb{X}), \mathcal{N}_s(\mathbb{X}))$. If a point process is simple it is common to identify the point process with its support, and view the point process as a random set of discrete points in $\mathbb{X}$. We may for example write $x \in \eta$ instead of $x \in \operatorname{supp}(\eta)$. It is common practice to switch between the interpretations of a simple point process as a random counting measure or as a random set of points, depending on whichever interpretation is more convenient. We will also do this throughout this chapter. Enumerating the points of a simple point process in a measurable way we may write:

$$\eta = \{x_1, x_2, \dots\}, \quad \text{and} \quad \eta = \sum_{i=1}^{\eta(\mathbb{X})} \delta_{x_i}.$$

For $v \in \mathbb{R}^d$ let S_v denote the shift operator. Suppose $\eta = \{(x_1, h_1), (x_2, h_2), \dots\}$ is a point process with $x_i \in \mathbb{R}^d$ and $h_i > 0$. Then, we define $S_v\eta := \{(x_1 - v, h_1), (x_2 - v, h_2), \dots\}$. Additionally, for a deterministic set $B \subset \mathbb{R}^d \times (0,\infty)$ we define $S_vB := \{(x + v, h) : (x, h) \in B\}$. Note that in the random counting measure interpretation of a point process, the definition is as follows: $S_v\eta(B) := \eta(S_vB)$, for $B \in \mathcal{B}(\mathbb{R}^d \times (0,\infty))$. This is indeed consistent with the previous definition since $S_v\eta(B) = \sum_i \delta_{(x_i,h_i)}(S_vB) = \sum_i \delta_{(x_i - v, h_i)}(B)$. We call η stationary if $S_v\eta$ and η are equal in distribution for all $v \in \mathbb{R}^d$. Throughout this chapter, $(W_n)_{n\geq 1}$ is a fixed convex averaging sequence. That is, each $W_n \subset \mathbb{R}^d$ is convex and compact, and the sequence is increasing: $W_n \subset W_{n+1}$. Finally, the sequence $(W_n)_{n\geq 1}$ expands unboundedly: $\sup\{r \geq 0 : B(x,r) \subset W_n \text{ for some } x \in W_n\} \to \infty$ as $n \to \infty$.

5.3 Poisson-Laguerre tessellations

In this section we describe the main mathematical object of interest in this chapter, the Poisson-Laguerre tessellation. This random tessellation is a generalization of the well-known Poisson-Voronoi tessellation, and was first studied in [54] and [56]. We will mostly follow the description of the Poisson-Laguerre tessellation as given in [39], which is subtly different. Let us start with the definition of a tessellation:

Definition 5.2. A tessellation of $\mathbb{R}^d$ is a countable collection $T = \{C_i : i \in \mathbb{N}\}$, of sets $C_i \subset \mathbb{R}^d$ (the cells of the tessellation) such that:

- $\text{int}(C_i) \cap \text{int}(C_j) = \emptyset$, if $i \neq j$.
- $\cup_{i\in\mathbb{N}} C_i = \mathbb{R}^d$.
- T is locally finite: $\#\{i \in \mathbb{N} : C_i \cap B \neq \emptyset\} < \infty$ for all bounded $B \in \mathcal{B}(\mathbb{R}^d)$.
- Each C_i is a compact and convex set with interior points.

Now, we introduce the Laguerre diagram. Let $\varphi = \{(x_i, h_i)\}_{i\in\mathbb{N}}$, with $x_i \in \mathbb{R}^d$ and $h_i > 0$. Assume moreover that $x_i \neq x_j$ for $i \neq j$. The Laguerre cell associated with $(x, h) \in \varphi$ is defined as:

$$C((x,h),\varphi) = \left\{y \in \mathbb{R}^d : \|y - x\|^2 + h \leq \|y - x'\|^2 + h' \text{ for all } (x', h') \in \varphi\right\}. \tag{5.2}$$

For $i \in \mathbb{N}$, the cell $C((x_i, h_i), \varphi)$ may be written as the intersection of half spaces:

$$C((x_i,h_i),\varphi) = \bigcap_{j\in\mathbb{N}} \left\{y \in \mathbb{R}^d : 2\langle y, x_j - x_i\rangle \leq \|x_i\|^2 - \|x_j\|^2 + h_j - h_i\right\}. \tag{5.3}$$

The Laguerre diagram generated by φ is the set of non-empty Laguerre cells, and is denoted by $L(\varphi)$:

$$L(\varphi) := \{C((x,h),\varphi) : (x,h) \in \varphi \text{ and } C((x,h),\varphi) \neq \emptyset\}.$$

A Laguerre diagram is not necessarily a tessellation, conditions on φ are needed to ensure that $L(\varphi)$ is locally finite and that all cells are bounded. Note that we consider a different parameterization of Laguerre diagrams in this chapter, compared to the Laguerre diagrams considered in chapter 4. Additionally, here we consider Laguerre diagrams in $\mathbb{R}^d$ instead of Laguerre diagrams in bounded domains. As we will discuss in a moment, the random Laguerre diagrams we consider are in fact tessellations. A Laguerre diagram has an interesting interpretation as a crystallization process. From the definition of a Laguerre cell it follows that:

$$x \in C((x_i,h_i),\varphi) \iff \exists t \geq h_i : x \in \bar{B}\left(x_i, \sqrt{t - h_i}\right) \text{ and } x \notin \bigcup_{j\neq i} B\left(x_j, \sqrt{(t-h_j)_+}\right),$$

with $(x)_+ = \max\{x, 0\}$. Hence, we may consider the ball $B_i(t) := \bar{B}(x_i, \sqrt{(t - h_i)_+})$ which starts growing at time $t = h_i$. The ball initially grows fast, and then its growth slows down

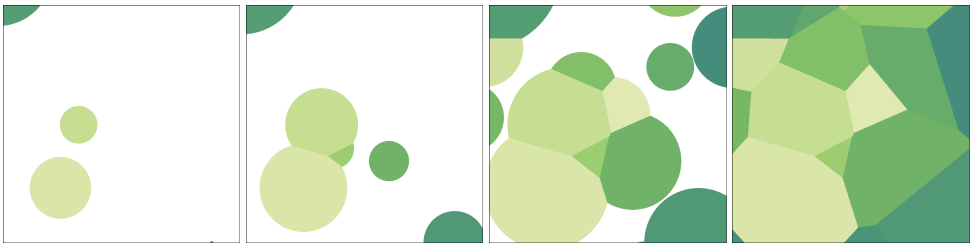


Figure 5.2: Visualization of the crystallization process. From left to right, the crystallization process is shown at times $t = 60$, $t = 80$, $t = 120$ and $t = 280$.

because the rate of increase of the function $t \mapsto \sqrt{t - h_i}$ slows down ($t \geq h_i$). If B_i is the first ball to hit a given point $x \in \mathbb{R}^d$, then $x \in C((x_i, h_i), \varphi)$. It is possible that x_i lies in another cell $C((x_j, h_j), \varphi)$, $i \neq j$ and yet $C((x_i, h_i), \varphi)$ may be non-empty. It is also possible that a pair (x_i, h_i) does not generate a cell, essentially because its ball starts growing too late. A visualization of the crystallization process is given in Figure 5.2.

In the literature one can also find other parameterizations of Laguerre diagrams. For instance, the parameterization in [54] can be obtained as follows. Let $\psi = \{(x_i, r_i)\}_{i \in \mathbb{N}} \subset \mathbb{R}^d \times (0, \infty)$, set $h_i = -r_i^2$ for all $i \in \mathbb{N}$, and then consider the Laguerre diagram generated by $\varphi = \{(x_i, h_i)\}_{i \in \mathbb{N}}$. The pair (x_i, r_i) is then often associated with the sphere centered at x_i with radius r_i. Note that this choice leads to negative weights in our choice of parameterization. We would like to note that the weights are allowed to be negative, we consider positive weights for mathematical convenience.

Throughout this chapter we assume that η is a Poisson process on $\mathbb{R}^d \times (0, \infty)$ with intensity measure $\nu_d \times \mathbb{F}$. Here, $\mathbb{F}$ is a locally finite measure concentrated on $(0, \infty)$. Because the measure $\nu_d \times \mathbb{F}$ has no atoms, η is a simple point process. From proposition 3.6. in [39] it follows that $L(\eta)$, the Laguerre diagram generated by the Poison process η, is with probability one a tessellation. We refer to $L(\eta)$ as the Poisson-Laguerre tessellation generated by η. We do note that in the aforementioned paper it is additionally assumed that $\mathbb{F}$ is absolutely continuous with respect to Lebesgue measure. However, this assumption is not needed for $L(\eta)$ to be a tessellation with probability one, as this is a straightforward modification of the proofs given in [39].

For $z \geq 0$ we define:

$$F(z) := \mathbb{F}((0, z]). \tag{5.4}$$

Thereby, this monotone function F is the only parameter in this model to be estimated. Note that F is not necessarily bounded, it is bounded if and only if $\mathbb{F}$ is a finite measure. If $\mathbb{F}$ is a finite measure one may define the constant $\lambda = \lim_{z \to \infty} F(z)$ and the probability measure $Q(\cdot) = \mathbb{F}(\cdot)/\lambda$. The intensity measure of η is then given by $\lambda \nu_d \times Q$, and η may be seen as an independently marked homogeneous Poisson process on $\mathbb{R}^d$ with mark space $(0, \infty)$. Its intensity is given by λ and Q represents its mark distribution. In view of the crystallization process interpretation of a Laguerre diagram, we may also say that Q, and thereby F, describes the distribution of the arrival times of the generator points.

Remark 5.1. *While we consider a Poisson process η on $\mathbb{R}^d \times (0, \infty)$, in the context of*

Poisson-Laguerre tessellations, also Poisson processes on $\mathbb{R}^d \times E$ have been considered for other choices of $E \subset \mathbb{R}$. For the purposes of this chapter we could also have considered $E = (a, \infty)$ for any choice of $a \in \mathbb{R}$, and have obtained analogous results. For the sake of convenience we have chosen to take $a = 0$. However, the particular value of a does not matter in practice. Note from the definition of the Laguerre cell that adding a fixed constant c to all weights does not affect the cell. Analogously, replacing F by $\tilde{F}$, where $\tilde{F}(z) = F(z - c)$ does not affect the distribution of the resulting Poisson-Laguerre tessellation, as it shifts the distribution of the arrival times by a constant. Hence, as a consequence, when dealing with real data, one may always add a sufficiently large constant c to all observed weights to ensure positivity of the weights. After all, the resulting weighted points may be considered to be a realization from the same data generating mechanism, but now with $\tilde{F}$ as the underlying parameter.

In the introduction of this chapter we explained that we are interested in estimators for F which depend on the observed Laguerre cells and the extreme points of η, which we denote by η^*, as defined in (5.1). To be precise, the estimators we propose for F depend on the points of η^* in the observation window W_n, as well as the Laguerre cells corresponding to these points of η^* in W_n. Recall, $(W_n)_{n\geq 1}$ is some fixed convex averaging sequence. The reader may for example keep $W_n = [-n, n]^d$ in mind as an explicit example. Note that the point process η^* may be seen as a (dependent) thinning of η, and is not necessarily a Poisson process. We conclude this section with a simulation example, with the purpose of providing an intuitive understanding of Poisson-Laguerre tessellations.

Example 2. In Figure 5.1 a realization is shown of a planar Poisson-Laguerre tessellation along with its realization of extreme points. The side length of the square observation window is equal to 40. For this example we have taken $\mathbb{F}$ to be a discrete probability measure on $\{1, 8, 10\}$. Specifically, $\mathbb{F}$ is defined as: $\mathbb{F}(\{1\}) = 0.01$, $\mathbb{F}(\{8\}) = 0.04$ and $\mathbb{F}(\{10\}) = 0.95$. Hence, η may be seen as an independently marked homogeneous Poisson process, with points in $\mathbb{R}^2$ and marks in $\{1, 8, 10\}$. The homogeneous Poisson process has intensity 1 and the marks are distributed according to $\mathbb{F}$. Let us briefly discuss the image in Figure 5.1 in view of the crystallization process interpretation. Given the choice of $\mathbb{F}$, we expect a small number of balls corresponding to points with weight $h = 1$, these balls start growing early, and result in large cells. A larger number of points with weight $h = 8$ have balls associated with them which start growing later, yielding cells which are a bit smaller. Finally, a very large number of points with weight $h = 10$ will generate even smaller cells.

5.4 Inference via a dependent thinning

5.4.1 Definition of an estimator

In this section, we define our first estimator for F. This estimator only depends on points (x, h) of η^* with $x \in W_n$ and for which x is located in its own Laguerre cell. The estimator is easy to compute, and the techniques used in this section will be important when we define an estimator for F based on all points of η^* in $W_n \times (0, \infty)$. Recall from the previous section that η is a Poisson process on $\mathbb{R}^d \times (0, \infty)$, $d \geq 2$, with intensity measure $\nu_d \times \mathbb{F}$.

We may also write: $\eta = \{(x_1, h_1), (x_2, h_2), \dots\}$, with $x_i \in \mathbb{R}^d$, $h_i > 0$. We start as follows, let $y \in \mathbb{R}^d$, and consider the following thinning of η:

$$\eta^y := \{(x, h) \in \eta : x + y \in C((x, h), \eta)\}. \tag{5.5}$$

In (5.2) we defined $C((x, h), \eta)$, which denotes the Laguerre cell associated with the weighted point $(x, h) \in \eta$. Evidently, for every $y \in \mathbb{R}^d$, η^y only contains a subset of points of η^*. Hence, we have: $\eta^y \subset \eta^* \subset \eta$. In particular, for $y = 0$ we obtain the set of points of η^* which are contained within their own Laguerre cell. In the following lemma we compute the intensity measure of η^y.

Lemma 5.1. *Let $B \in \mathcal{B}(\mathbb{R}^d)$, $y \in \mathbb{R}^d$ and $z \geq 0$, the intensity measure Λ^y of η^y satisfies:*

$$\Lambda^y (B \times (0, z]) = \nu_d(B) \int_0^z \exp\left(-\kappa_d \int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) \mathrm{d}F(h).$$

This intensity measure can be computed via the Mecke equation, which may for example be found in Theorem 4.1 in [53]. The statement is as follows:

Theorem 5.3 (Mecke equation)**.** *Let Λ be a σ-finite measure on a measurable space $(\mathbb{X}, \mathcal{X})$ and let η be a point process on $\mathbb{X}$. Then η is a Poisson process with intensity measure Λ if and only if:*

$$\mathbb{E}\left(\sum_{x \in \eta} f(x, \eta)\right) = \int \mathbb{E}\left(f(x, \eta + \delta_x)\right) \Lambda(\mathrm{d}x),$$

for all non-negative measurable functions $f : \mathbb{X} \times \mathbf{N}(\mathbb{X}) \to [0, \infty]$.

Proof of Lemma 5.1. By definition, the intensity measure of η^y is given by:

$$\begin{aligned}\Lambda^y (B \times (0, z]) &:= \mathbb{E}\left(\eta^y (B \times (0, z])\right) \\ &= \mathbb{E}\left(\sum_{(x,h) \in \eta} \mathbb{1}_B(x) \mathbb{1}_{(0,z]}(h) \mathbb{1}\left\{x + y \in C((x, h), \eta)\right\}\right). \end{aligned} \tag{5.6}$$

We rewrite the final indicator function in (5.6) into a more convenient form. By the definition of a Laguerre cell, we obtain:

$$\begin{aligned} x + y \in C((x, h), \eta) &\iff \|y\|^2 + h - h' \leq \|x + y - x'\|^2, \text{ for all } (x', h') \in \eta \\ &\iff \eta\left(A_{x,h,y}\right) = 0, \end{aligned}$$

where we define the set $A_{x,h,y}$ as:

$$A_{x,h,y} = \left\{(x', h') \in \mathbb{R}^d \times (0, \infty) : \|y\|^2 + h - h' > \|x + y - x'\|^2\right\}.$$

Since η is a Poisson process, the random variable $\eta\left(A_{x,h,y}\right)$ is Poisson distributed with parameter $\mathbb{E}(\eta\left(A_{x,h,y}\right))$. As a consequence, the probability that $\eta\left(A_{x,h,y}\right) = 0$ is given

by:

$$\begin{aligned}\mathbb{P}\left(\eta\left(A_{x,h,y}\right)=0\right) &= \exp\left(-\mathbb{E}\left(\eta\left(A_{x,h,y}\right)\right)\right)\\ &= \exp\left(-\int_{\mathbb{R}^d}\int_0^{\|y\|^2+h} \mathbb{1}\left\{\|x+y-x'\|<\sqrt{\|y\|^2+h-t}\right\}\mathrm{d}F(t)\mathrm{d}x'\right)\\ &= \exp\left(-\int_0^{\|y\|^2+h}\int_{\mathbb{R}^d} \mathbb{1}\left\{\|x'\|<\sqrt{\|y\|^2+h-t}\right\}\mathrm{d}x'\mathrm{d}F(t)\right)\\ &= \exp\left(-\kappa_d\int_0^{\|y\|^2+h}\left(\|y\|^2+h-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right). \end{aligned} \tag{5.7}$$

Note that (5.7) does not depend on x. Using (5.7) and the Mecke equation, the expectation in (5.6) can be computed as follows:

$$\begin{aligned}&\mathbb{E}\left(\sum_{(x,h)\in\eta}\mathbb{1}_B(x)\mathbb{1}_{(0,z]}(h)\mathbb{1}\{\eta\left(A_{x,h,y}\right)=0\}\right)=\\ &=\int_0^\infty\int_{\mathbb{R}^d}\mathbb{1}_B(x)\mathbb{1}_{(0,z]}(h)\mathbb{P}\left(\eta\left(A_{x,h,y}\right)=0\right)\mathrm{d}x\mathrm{d}F(h) \qquad (5.8)\\ &=\int_0^\infty\int_{\mathbb{R}^d}\mathbb{1}_B(x)\mathbb{1}_{(0,z]}(h)\exp\left(-\kappa_d\int_0^{\|y\|^2+h}\left(\|y\|^2+h-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right)\mathrm{d}x\mathrm{d}F(h)\\ &=\nu_d(B)\int_0^z\exp\left(-\kappa_d\int_0^{\|y\|^2+h}\left(\|y\|^2+h-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right)\mathrm{d}F(h).\end{aligned}$$

In (5.8) we used the fact that $(x,h)\notin A_{x,h,y}$ such that $\eta\left(A_{x,h,y}\right)=\left(\eta+\delta_{(x,h)}\right)\left(A_{x,h,y}\right)$. □

Recall that $\mathcal{F}_+$ denotes the space of all (not necessarily bounded) distribution functions on $(0,\infty)$. Given the statement of Lemma 5.1 we focus on the case $y=0$ and define for $F\in\mathcal{F}_+$ the function $G_F:[0,\infty)\to[0,\infty)$ via:

$$G_F(z):=\int_0^z\exp\left(-\kappa_d\int_0^h(h-t)^{\frac{d}{2}}\mathrm{d}F(t)\right)\mathrm{d}F(h). \tag{5.9}$$

For functions G_F with $F\in\mathcal{F}_+$ as in (5.9) we obtain the following important identifiability result:

Theorem 5.4. *Let $F_1,F_2\in\mathcal{F}_+$, $R>0$. If $G_{F_1}(z)=G_{F_2}(z)$ for all $z\in[0,R)$ then $F_1(z)=F_2(z)$ for all $z\in[0,R)$. In particular, if $G_{F_1}=G_{F_2}$ then $F_1=F_2$.*

The key ingredient for the proof of this theorem is a variant of the Grönwall inequality. This inequality is in particular known for its applications in integral- and differential equations. We refer to [74] for more variants of this inequality and their applications.

Theorem 5.5 (Theorem 1.3.3. in [74]). *Suppose u, α and β are measurable non-negative functions on $[0,\infty)$. Assume that α is non-decreasing. Assume for all $z \geq 0$: $u, \alpha, \beta \in L^1([0,z])$. If for all $z \geq 0$ the following holds:*

$$u(z) \leq \alpha(z) + \beta(z) \int_0^z u(s)\mathrm{d}s.$$

Then, for all $z \geq 0$:

$$u(z) \leq \alpha(z) \left(1 + \beta(z) \int_0^z \exp\left(\int_s^z \beta(r)\mathrm{d}r\right)\mathrm{d}s\right).$$

Note that if u, α and β satisfy the conditions in Theorem 5.5 and β is non-decreasing, then:

$$u(z) \leq \alpha(z)\left(1 + \beta(z) z \exp\left(\beta(z) z\right)\right). \tag{5.10}$$

We need to point out that in [74] this theorem also includes the assumption that u, α and β are continuous. However, as noted on p. 14 in the same reference, this assumption is not needed.

Proof of Theorem 5.4. Let $z \geq 0$. For $i \in \{1,2\}$ note that the (Lebesgue-Stieltjes) measures associated with G_{F_i} and F_i are mutually absolutely continuous. The corresponding Radon-Nikodym derivative is given by:

$$\frac{\mathrm{d}G_{F_i}}{\mathrm{d}F_i}(z) = \exp\left(-\kappa_d \int_0^z (z-t)^{\frac{d}{2}}\,\mathrm{d}F_i(t)\right).$$

Hence, we may also write:

$$F_i(z) = \int_0^z \frac{\mathrm{d}F_i}{\mathrm{d}G_{F_i}}(h)\mathrm{d}G_{F_i}(h) = \int_0^z \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F_i(t)\right)\mathrm{d}G_{F_i}(h).$$

Via integration by parts we may write:

$$\begin{aligned}\int_0^z (z-t)^{\frac{d}{2}}\mathrm{d}F_i(t) &= 0 \cdot F_i(z) - z^{\frac{d}{2}}F_i(0) - \int_0^z F_i(t)\mathrm{d}\left((z-t)^{\frac{d}{2}}\right)(t) \\ &= \frac{d}{2}\int_0^z F_i(t)(z-t)^{\frac{d}{2}-1}\mathrm{d}t. \end{aligned} \tag{5.11}$$

Moreover, the expression in (5.11) is a non-decreasing function of z. We now derive a general upper bound for $|F_1(z) - F_2(z)|$:

$$\begin{aligned}&|F_1(z) - F_2(z)| = \\ &= \left|\int_0^z \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F_1(t)\right)\mathrm{d}G_{F_1}(h) + \right. \\ &\quad \left. - \int_0^z \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F_2(t)\right)\mathrm{d}G_{F_2}(h)\right|\end{aligned}$$

$$\leq \left| \int_0^z \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_1(t)\right) - \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \mathrm{d}G_{F_1}(h) \right| +$$
$$+ \left| \int_0^z \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \mathrm{d}(G_{F_1} - G_{F_2})(h) \right|. \tag{5.12}$$

Let us now consider the first term of (5.12). For $h \geq 0$ define:

$$C(h) := \max\left\{ \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_1(t)\right), \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \right\}.$$

Note that C is increasing. Since $|e^x - e^y| \leq \max\{e^x, e^y\} \, |x-y|$ for $x, y \geq 0$ the first term in (5.12) is bounded by:

$$\begin{aligned}
&\int_0^z \left| \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_1(t)\right) - \exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \right| \mathrm{d}G_{F_1}(h) \\
&\leq \int_0^z C(h)\kappa_d \left| \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_1(t) - \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t) \right| \mathrm{d}G_{F_1}(h) \\
&= \int_0^z C(h)\kappa_d \left| \frac{d}{2} \int_0^h (F_1(t) - F_2(t)) \, (h-t)^{\frac{d}{2}-1} \mathrm{d}t \right| \mathrm{d}G_{F_1}(h) \\
&\leq \frac{d\kappa_d}{2} C(z) \int_0^z \int_0^h |F_1(t) - F_2(t)| \, (h-t)^{\frac{d}{2}-1} \mathrm{d}t \mathrm{d}G_{F_1}(h) \\
&\leq \frac{d\kappa_d}{2} C(z) z^{\frac{d}{2}-1} \int_0^z \int_0^z |F_1(t) - F_2(t)| \, \mathrm{d}t \mathrm{d}G_{F_1}(h) \\
&= \frac{d\kappa_d}{2} C(z) z^{\frac{d}{2}-1} G_{F_1}(z) \int_0^z |F_1(t) - F_2(t)| \, \mathrm{d}t.
\end{aligned}$$

Via integration by parts, the second term of (5.12) is bounded by:

$$\begin{aligned}
&\left| \exp\left(\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \left(G_{F_1}(z) - G_{F_2}(z)\right) \right| + \\
&+ \left| \int_0^z \left(G_{F_1}(h) - G_{F_2}(h)\right) \mathrm{d}\left(\exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \right)(h) \right| \\
&\leq \left| G_{F_1}(z) - G_{F_2}(z) \right| \exp\left(\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) + \\
&+ \sup_{h \in [0,z]} \left| G_{F_1}(h) - G_{F_2}(h) \right| \int_0^z \mathrm{d}\left(\exp\left(\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right) \right)(h) \\
&\leq \sup_{h \in [0,z]} \left| G_{F_1}(h) - G_{F_2}(h) \right| 2 \exp\left(\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t)\right).
\end{aligned}$$

Combining all results, we obtain:

$$|F_1(z) - F_2(z)| \leq \frac{d\kappa_d}{2} C(z) z^{\frac{d}{2}-1} G_{F_1}(z) \int_0^z |F_1(t) - F_2(t)| \, \mathrm{d}t + \\ + \sup_{h \in [0,z]} \left| G_{F_1}(h) - G_{F_2}(h) \right| 2 \exp\left(\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t) \right).$$

Applying Theorem 5.5 and (5.10) with $u(z) = |F_1(z) - F_2(z)|$ yields:

$$|F_1(z) - F_2(z)| \leq K(z) \sup_{h \in [0,z]} \left| G_{F_1}(h) - G_{F_2}(h) \right|. \tag{5.13}$$

Here, $K(z)$ is given by:

$$K(z) := \left(1 + \frac{d\kappa_d}{2} C(z) z^{\frac{d}{2}} G_{F_1}(z) \exp\left(\frac{d\kappa_d}{2} C(z) z^{\frac{d}{2}} G_{F_1}(z) \right) \right) \cdot \\ \cdot 2 \exp\left(\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \, \mathrm{d}F_2(t) \right).$$

The statement of the theorem immediately follows from (5.13). □

Suppose we wish to estimate G_F, and we observe the extreme points of η within the bounded observation window W_n, as well as their Laguerre cells. We define the following unbiased estimator for G_F:

$$\hat{G}_n(z) := \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) \mathbb{1}_{(0,z]}(h) \mathbb{1}\{x \in C((x,h), \eta)\} \\ = \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta^0} \mathbb{1}_{W_n}(x) \mathbb{1}_{(0,z]}(h). \tag{5.14}$$

In (5.14), η^0 represents the point process η^y as in (5.5) with $y = 0$. Hence, G_F is a function which we can estimate and which uniquely determines F, this motivates the following definition:

Definition 5.6 (First inverse estimator of F)**.** Define $\hat{F}_n^0$ to be the unique function $\hat{F}_n^0 \in \mathcal{F}_+$ which satisfies: $G_{\hat{F}_n^0}(z) = \hat{G}_n(z)$ for all $z \geq 0$, with $\hat{G}_n$ as in (5.14).

Let us now discuss why $\hat{F}_n^0$ is well-defined. Clearly, if there exists a function $\hat{F}_n^0 \in \mathcal{F}_+$ which satisfies $G_{\hat{F}_n^0}(z) = \hat{G}_n(z)$ for all $z \geq 0$ then it is unique by Theorem 5.4. Suppose $(x_1, h_1), (x_2, h_2), \ldots, (x_k, h_k)$ is the sorted realization of the points of η^0 with $x_1, \ldots, x_k \in W_n$ and $h_1 \leq h_2 \leq \cdots \leq h_k$. We may write:

$$\hat{G}_n(z) = \frac{1}{\nu_d(W_n)} \sum_{i=1}^{k} \mathbb{1}\{h_i \leq z\}.$$

Set $h_0 = 0$ such that $\hat{F}_n^0(h_0) = 0$. Clearly, $\hat{G}_n$ is piecewise constant, with jump locations at $h_1, \ldots, h_k$. Recall from the proof of Theorem 5.4 that the Lebesgue-Stieltjes measures

associated with $\hat{F}_n^0$ and $G_{\hat{F}_n^0}$ are mutually absolutely continuous. As a consequence, if $\hat{F}_n^0$ exists, it is necessarily also piecewise constant with the same jump locations as $G_{\hat{F}_n^0} = \hat{G}_n$. Therefore, if we can uniquely specify the value of $\hat{F}_n^0$ at $h_1, \ldots, h_k$, existence and uniqueness of $\hat{F}_n^0$ is established. Let $i \in \{1, \ldots, k\}$ then, for the $\hat{F}_n^0$ we are looking for:

$$\begin{aligned}
&\hat{G}_n(h_i) = \\
&= \hat{G}_n(h_{i-1}) + \int_{h_{i-1}}^{h_i} \exp\left(-\kappa_d \int_0^h (h-t)^{\frac{d}{2}} \, \mathrm{d}\hat{F}_n^0(t)\right) \mathrm{d}\hat{F}_n^0(h) \\
&= \hat{G}_n(h_{i-1}) + \exp\left(-\kappa_d \sum_{j=1}^{i} \left(h_i - h_j\right)^{\frac{d}{2}} \left(\hat{F}_n^0(h_j) - \hat{F}_n^0(h_{j-1})\right)\right)\left(\hat{F}_n^0(h_i) - \hat{F}_n^0(h_{i-1})\right) \\
&= \hat{G}_n(h_{i-1}) + \exp\left(-\kappa_d \sum_{j=1}^{i-1} \left(h_i - h_j\right)^{\frac{d}{2}} \left(\hat{F}_n^0(h_j) - \hat{F}_n^0(h_{j-1})\right)\right)\left(\hat{F}_n^0(h_i) - \hat{F}_n^0(h_{i-1})\right).
\end{aligned}$$

Since $\hat{F}_n^0(h_0) = 0$, $\hat{F}_n^0$ is recursively defined via:

$$\begin{aligned}
\hat{F}_n^0(h_i) = \hat{F}_n^0(h_{i-1}) + \left(\hat{G}_n(h_i) - \hat{G}_n(h_{i-1})\right) \cdot \\
\cdot \exp\left(\kappa_d \sum_{j=1}^{i-1} \left(h_i - h_j\right)^{\frac{d}{2}} \left(\hat{F}_n^0(h_j) - \hat{F}_n^0(h_{j-1})\right)\right).
\end{aligned} \tag{5.15}$$

Note that the RHS of (5.15) only depends on the values $\hat{F}_n^0(h_j)$ with $j < i$. So indeed, (5.15) completely defines $\hat{F}_n^0$. Moreover, this expression is also a convenient formula for computing $\hat{F}_n^0$ in practice.

5.4.2 Consistency

In this section we show that $\hat{F}_n^0$, as in Definition 5.6, is a strongly consistent estimator for F. A single realization of $\hat{G}_n$ and $\hat{F}_n^0$ are shown in Figure 5.3. We present additional simulation results in section 5.7. The first step to proving consistency is to show that the estimator $\hat{G}_n$ as in (5.14) for G_F is strongly consistent. For empirical estimators such as $\hat{G}_n$, their consistency follows from a spatial ergodic theorem. From Proposition 13.4.I. in [23], and the ergodicity of the Poisson process under consideration, we obtain:

Theorem 5.7 (Spatial ergodic theorem)**.** *Let η be a Poisson process on $\mathbb{X} = \mathbb{R}^d \times (0, \infty)$ with intensity measure $\nu_d \times \mathbb{F}$. Here, $\mathbb{F}$ is a locally finite measure concentrated on $(0, \infty)$. Let $g(\psi, h)$ be a measurable non-negative function on $\mathbf{N}(\mathbb{X}) \times (0, \infty)$. Then, for any convex averaging sequence $(W_n)_{n \geq 1}$:*

$$\lim_{n \to \infty} \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) g(S_x \eta, h) \overset{a.s.}{=} \int_0^\infty \mathbb{E}\left(g\left(\eta + \delta_{(0,h)}, h\right)\right) \mathbb{F}(\mathrm{d}h).$$

We do note that Proposition 13.4.I in [23] is phrased in the context that $\mathbb{F}$ is a finite measure. However, like Theorem 12.2.IV in the same reference (another spatial ergodic

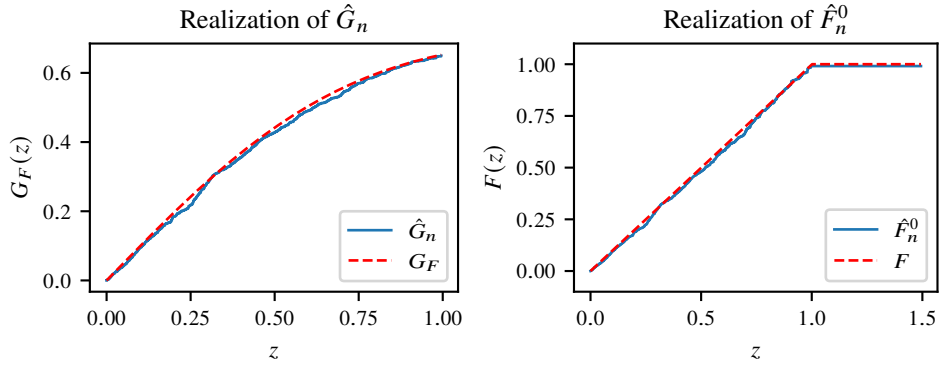


Figure 5.3: Left: A realization of $\hat{G}_n$. Right: The corresponding realization of $\hat{F}_n^0$. The actual underlying F is equal to the CDF of a uniform distribution on $(0, 1)$.

theorem), which is stated under the assumption that $\mathbb{F}$ is locally finite, the result remains valid if $\mathbb{F}$ is locally finite. Besides the spatial ergodic theorem we also need the following useful lemma for estimators of monotone functions:

Lemma 5.2. *Let $(F_n)_{n\geq 1}$ be a random sequence of monotone functions on $\mathbb{R}$, and let F be a deterministic monotone function on $\mathbb{R}$. If for all $z \in \mathbb{R}$: $\mathbb{P}(\lim_{n\to\infty} F_n(z) = F(z)) = 1$, then: $\mathbb{P}(\lim_{n\to\infty} F_n(z) = F(z),\ \forall z \in \mathbb{R}) = 1$.*

The proof of Lemma 5.2 is given in section 5.10. We obtain the following result:

Corollary 5.1. *With probability one: $\lim_{n\to\infty} \hat{G}_n(z) = G_F(z)$ for all $z \geq 0$.*

Proof. Let $z \geq 0$, by Lemma 5.2 it is sufficient to show that $\lim_{n\to\infty} \hat{G}_n(z) = G_F(z)$ almost surely. Using the same notation as in the proof of Lemma 5.1, note that $\eta(A_{x,h,0}) = S_x\eta(A_{0,h,0})$ for all $(x, h) \in \eta$ almost surely. As a consequence, $\hat{G}_n(z)$ may be written as follows:

$$\hat{G}_n(z) = \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)\mathbb{1}\{S_x\eta(A_{0,h,0}) = 0\}.$$

Following the computation in the proof of Lemma 5.1, it is readily verified that applying the spatial ergodic theorem with $g(\psi, h) = \mathbb{1}_{(0,z]}(h)\mathbb{1}\{\psi(A_{0,h,0}) = 0\}$ yields the result with the desired limit. □

Finally, we need the following continuity result:

Lemma 5.3. *Let $(F_n)_{n\geq 1}$ be a sequence of functions in $\mathcal{F}_+$ and let $F \in \mathcal{F}_+$. Let $R > 0$. If $\lim_{n\to\infty} F_n(z) = F(z)$ for all $z \in [0, R)$, then $\lim_{n\to\infty} G_{F_n}(z) = G_F(z)$ for all $z \in [0, R)$. In particular, if $\lim_{n\to\infty} F_n(z) = F(z)$ for all $z \geq 0$, then $\lim_{n\to\infty} G_{F_n}(z) = G_F(z)$ for all $z \geq 0$.*

The proof of Lemma 5.3 is given in section 5.10. Combining the previous results with Theorem 5.4 we prove the following consistency result.

Theorem 5.8 (Consistency of $\hat{F}_n^0$)**.** *With probability one,* $\lim_{n\to\infty} \hat{F}_n^0(z) = F(z)$ *for all* $z \geq 0$.

Proof. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting a Poisson process η, with intensity measure $\nu_d \times \mathbb{F}$. By Corollary 5.1 there exists a set $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ and $z \geq 0$ we have $\lim_{n\to\infty} \hat{G}_n(z;\omega) = G_F(z)$. Let $\omega \in \Omega_0$ and $z \geq 0$, we show that $\lim_{n\to\infty} \hat{F}_n^0(z;\omega) = F(z)$.

Pick $M > 0$ such that $F(z) < M$. For $n \in \mathbb{N}$ and $h \geq 0$, define: $\bar{F}_n(h) = \min\{\hat{F}_n^0(h;\omega), M\}$. Then, $(\bar{F}_n)_{n\geq 1}$ is a uniformly bounded sequence of monotone functions. Let $(n_l)_{l\geq 1} \subset (n)_{n\geq 1}$ be an arbitrary subsequence. By Helly's selection principle there exists a further subsequence $(n_k)_{k\geq 1} \subset (n_l)_{l\geq 1}$ such that $\bar{F}_{n_k}$ converges pointwise to some monotone function $\bar{F}$ as $k \to \infty$. This implies that $\lim_{k\to\infty} \hat{F}_{n_k}^0(h;\omega) = \lim_{k\to\infty} \bar{F}_{n_k}(h) = \bar{F}(h)$ for all $h \in [0, R)$ with $R := \sup\{h \geq 0 : \bar{F}(h) < M\}$. By Lemma 5.3 we obtain:

$$\lim_{k\to\infty} \hat{G}_{n_k}(h;\omega) := \lim_{k\to\infty} G_{\hat{F}_{n_k}^0(\,\cdot\,;\omega)}(h) = G_{\bar{F}}(h) \quad \text{for all } h \in [0, R).$$

Because the whole sequence $\hat{G}_n(h;\omega)$ converges to $G_F(h)$ as $n \to \infty$, for $h \geq 0$, we obtain $G_F(h) = G_{\bar{F}}(h)$ for all $h \in [0, R)$. Theorem 5.4 now yields $F(h) = \bar{F}(h)$ for all $h \in [0, R)$, and since $z \in [0, R)$ we have in particular $F(z) = \bar{F}(z)$. As a consequence: $\lim_{k\to\infty} \hat{F}_{n_k}^0(z;\omega) = F(z)$. Because the initial subsequence was chosen arbitrarily, the whole sequence converges: $\lim_{n\to\infty} \hat{F}_n^0(z;\omega) = F(z)$. □

5.5 Inference via the volume-biased weight distribution

5.5.1 Definition of an estimator

In this section we define a second estimator for F, which depends on all points of η^* in $W_n \times (0, \infty)$ as well as the volumes of the Laguerre cells corresponding to these points. As such, this estimator depends on more data compared to the estimator in the previous section. First, we present a result for Poisson-Laguerre tessellations in $\mathbb{R}^d$, the estimator itself is defined specifically for the planar case ($d = 2$). Suppose for now that $\mathbb{F}$ is a finite measure, such that η may be interpreted as an independently marked homogeneous Poisson process. Because $\mathbb{F}$ then determines the distribution of the weights (h-coordinates) of the points of η, a natural question is to ask how the distribution of the weights of the points of η^* is related to $\mathbb{F}$. As it turns out, it is more tractable to study a biased or weighted version of this distribution. We introduce the so-called volume-biased weight distribution in the following definition, which is also well-defined if $\mathbb{F}$ is not a finite measure:

Definition 5.9 (volume-biased weight distribution)**.** Let η be a Poisson process on $\mathbb{R}^d \times (0, \infty)$, $d \geq 2$, with intensity measure $\nu_d \times \mathbb{F}$. Here, $\mathbb{F}$ is a locally finite measure concentrated on $(0, \infty)$. Let $A \in \mathcal{B}(\mathbb{R})$, define the following probability measure:

$$\mathbb{F}^V(A) := \mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{[0,1]^d}(x) \mathbb{1}_A(h) \nu_d\left(C((x,h),\eta)\right) \right). \tag{5.16}$$

Consider the Poisson-Laguerre tessellation generated by η, then the interpretation of $\mathbb{F}^V$ is as follows. $\mathbb{F}^V$ describes the distribution of the random weight associated with a randomly chosen Laguerre cell, the probability of picking any given cell being proportional to its volume. Because there is an infinite number of Laguerre cells in the tessellation, care needs to be taken in making this statement precise. This can be done via Palm calculus for marked point processes, see for instance chapter 3 in [86]. Note that the sum in (5.16) effectively only sums over points $(x, h) \in \eta$ with a Laguerre cell $C((x, h), \eta)$ of positive volume. Hence, it can also be seen as a sum over points of η^*. From its definition it is not immediately obvious that $\mathbb{F}^V$ is a well-defined probability measure. Specifically, it is not immediately evident that $\mathbb{F}^V(\mathbb{R}) = 1$. We address this in the proof of the following theorem, where we derive the CDF (Cumulative Distribution Function) associated with $\mathbb{F}^V$.

Theorem 5.10. *Let η be a Poisson process as in Definition 5.9, and let $z \geq 0$. Define $F(z) := \mathbb{F}((0, z])$ and $F^V(z) := \mathbb{F}^V((0, z])$, the distribution functions corresponding to $\mathbb{F}$ and $\mathbb{F}^V$ respectively. The measure $\mathbb{F}^V$ is a probability measure and F^V is given by:*

$$F^V(z) = 1 - \exp\left(-\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \mathrm{d}F(t)\right) + \\ + \frac{d\kappa_d}{2} \int_z^\infty \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}}\, \mathrm{d}F(t)\right) \int_0^z (u-h)^{\frac{d}{2}-1} \mathrm{d}F(h)\mathrm{d}u.$$

Proof. By the translation invariance of Lebesgue measure and Fubini's theorem, we may write:

$$\begin{aligned} F^V(z) &= \mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{[0,1]^d}(x)\mathbb{1}_{(0,z]}(h)\nu_d\left(C((x,h),\eta) - x\right)\right) \\ &= \mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{[0,1]^d}(x)\mathbb{1}_{(0,z]}(h) \int_{\mathbb{R}^d} \mathbb{1}\{y \in C((x,h),\eta) - x\}\mathrm{d}y\right) \\ &= \int_{\mathbb{R}^d} \mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{[0,1]^d}(x)\mathbb{1}_{(0,z]}(h)\mathbb{1}\{x + y \in C((x,h),\eta)\}\right)\mathrm{d}y \\ &= \int_{\mathbb{R}^d} \mathbb{E}\left(\eta^y([0,1]^d \times (0,z])\right)\mathrm{d}y. \end{aligned} \tag{5.17}$$

With η^y as in (5.5). In Lemma 5.1 we computed the expectation in (5.17). Plugging in this expression, and passing to polar coordinates by substituting $y = r\theta$, with $r \geq 0$ and $\theta \in \mathbb{S}^{d-1}$, we obtain:

$$\begin{aligned} F^V(z) &= \int_{\mathbb{R}^d} \int_0^z \exp\left(-\kappa_d \int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right)\mathrm{d}F(h)\mathrm{d}y \\ &= d\kappa_d \int_0^z \int_0^\infty \exp\left(-\kappa_d \int_0^{r^2+h} \left(r^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) r^{d-1}\mathrm{d}r\mathrm{d}F(h) \end{aligned} \tag{5.18}$$

$$= \frac{d\kappa_d}{2} \int_0^z \int_h^\infty \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right) (u-h)^{\frac{d}{2}-1} \mathrm{d}u \mathrm{d}F(h) \tag{5.19}$$

$$= \frac{d\kappa_d}{2} \int_0^\infty \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right) \int_0^{\min\{u,z\}} (u-h)^{\frac{d}{2}-1} \mathrm{d}F(h) \mathrm{d}u. \tag{5.20}$$

In (5.18) and (5.20) we apply Fubini's theorem, and in (5.19) we substitute $u = r^2 + h$. We can now write $F^V(z)$ as a sum of two integrals:

$$\begin{aligned} F^V(z) = &\frac{d\kappa_d}{2} \int_0^z \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right) \int_0^u (u-h)^{\frac{d}{2}-1} \mathrm{d}F(h) \mathrm{d}u + \\ &+ \frac{d\kappa_d}{2} \int_z^\infty \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right) \int_0^z (u-h)^{\frac{d}{2}-1} \mathrm{d}F(h) \mathrm{d}u. \end{aligned} \tag{5.21}$$

The first integral of (5.21) can be calculated explicitly since the integrand has an explicit primitive. The first term of (5.21) is given by:

$$\left[-\exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right)\right]_0^z = 1 - \exp\left(-\kappa_d \int_0^z (z-t)^{\frac{d}{2}} \mathrm{d}F(t)\right).$$

Plugging this back into (5.21) yields the expression for F^V as stated in the theorem. Finally, via (5.20) we can show that $\lim_{z\to\infty} F^V(z) = 1$. After all, the integrand in (5.20) (considering the integral w.r.t. u) can be bounded from above using the inequality $\min\{u, z\} \leq u$. Via the dominated convergence theorem it follows that:

$$\begin{aligned} \lim_{z\to\infty} F^V(z) &= \frac{d\kappa_d}{2} \int_0^\infty \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right) \lim_{z\to\infty} \int_0^{\min\{u,z\}} (u-h)^{\frac{d}{2}-1} \mathrm{d}F(h) \mathrm{d}u \\ &= \int_0^\infty \exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right) \frac{d\kappa_d}{2} \int_0^u (u-h)^{\frac{d}{2}-1} \mathrm{d}F(h) \mathrm{d}u \\ &= \left[-\exp\left(-\kappa_d \int_0^u (u-t)^{\frac{d}{2}} \, \mathrm{d}F(t)\right)\right]_0^\infty = 1. \end{aligned}$$

□

The Stieltjes integrals in the expression for F^V may be written as Lebesgue integrals using integration by parts. For instance:

$$\begin{aligned} \int_0^z (z-t)^{\frac{d}{2}} \mathrm{d}F(t) &= 0 \cdot F(z) - z^{\frac{d}{2}} F(0) - \int_0^z F(t) \mathrm{d}\left((z-t)^{\frac{d}{2}}\right)(t) \\ &= \frac{d}{2} \int_0^z F(t)(z-t)^{\frac{d}{2}-1} \mathrm{d}t. \end{aligned} \tag{5.22}$$

As announced in the beginning of this section, we will now focus on the case $d = 2$, which is important for practical applications. In that case, Theorem 5.10 and (5.22) yield the following expression for F^V.

Corollary 5.2. *Let $z \geq 0$, if $d = 2$ the CDF F^V is given by:*

$$F^V(z) = 1 - \exp\left(-\pi \int_0^z F(t)\mathrm{d}t\right) + \pi F(z) \int_z^\infty \exp\left(-\pi \int_0^u F(t)\mathrm{d}t\right) \mathrm{d}u. \tag{5.23}$$

Let us now introduce some convenient notation which will be used throughout this section. For $z \geq 0$, $F \in \mathcal{F}_+$ and $m \geq 0$ we define:

$$V(z; F, m) := 1 - \exp\left(-\pi \int_0^z F(t)\mathrm{d}t\right) + \pi F(z) \left(m - \int_0^z \exp\left(-\pi \int_0^u F(t)\mathrm{d}t\right) \mathrm{d}u\right).$$

$$m_F := \int_0^\infty \exp\left(-\pi \int_0^u F(t)\mathrm{d}t\right) \mathrm{d}u.$$

Note that if $m = m_F$, then $V(\cdot; F, m) = F^V$, with F^V as in (5.23). In other words, $V(\cdot; F, m)$ is then the volume-biased weight distribution induced by F. We obtain the following identifiability result:

Theorem 5.11. *Let $F_1, F_2 \in \mathcal{F}_+$, let $R > 0$. If $m_{F_1} = m_{F_2}$ and $V(z; F_1, m_{F_1}) = V(z; F_2, m_{F_2})$ for all $z \in [0, R)$, then $F_1(z) = F_2(z)$ for all $z \in [0, R)$. Consequently, if $m_{F_1} = m_{F_2}$ and $V(\cdot; F_1, m_{F_1}) = V(\cdot; F_2, m_{F_2})$, then $F_1 = F_2$.*

The proof of Theorem 5.11 as well as the proofs of most of the remaining lemmas in this section are postponed to section 5.10. The techniques used for proving these results are similar to the techniques used in section 5.4. We now define the following natural estimator for the distribution function F^V:

$$\tilde{F}_n^V(z) := \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)\nu_d\left(C((x,h),\eta)\right). \tag{5.24}$$

Alternatively, the following estimator for F^V may be defined:

$$\hat{F}_n^V(z) := \frac{\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)\nu_d\left(C((x,h),\eta)\right)}{\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\nu_d\left(C((x,h),\eta)\right)}. \tag{5.25}$$

Remark 5.2. *Note that the estimators $\hat{F}_n^V$ and $\tilde{F}_n^V$ for F^V do not incorporate edge effects. For instance, a Laguerre cell may be partially observed through the observation window W_n, such that computation of the estimators requires information outside of the window. In practice one could artificially shrink the observation window such that the estimators can be computed based on this smaller window.*

Similarly to $\hat{F}_n^0$, we can define an inverse estimator for F using an estimator for F^V. We choose to use $\hat{F}_n^V$ for this purpose, since it satisfies $\lim_{z\to\infty} \hat{F}_n^V(z) = 1$, in general this is not the case for $\tilde{F}_n^V$. In view of Theorem 5.11 we need to keep in mind that the constant m_F is unknown. We can resolve this by first using $\hat{F}_n^0$ to estimate m_F. That is, we define:

$$\hat{m}_n := m_{\hat{F}_n^0} = \int_0^\infty \exp\left(-\pi \int_0^u \hat{F}_n^0(t)\mathrm{d}t\right) \mathrm{d}u. \tag{5.26}$$

Finally, we define our second estimator for F as follows:

Definition 5.12 (Second inverse estimator of F)**.** Define $\hat{F}_n$ to be the unique function $\hat{F}_n \in \mathcal{F}_+$ which satisfies for all $z \geq 0$:

$$\hat{F}_n^V(z) = 1 - \exp\left(-\pi \int_0^z \hat{F}_n(t)\mathrm{d}t\right) + \pi \hat{F}_n(z) \left(\hat{m}_n - \int_0^z \exp\left(-\pi \int_0^u \hat{F}_n(t)\mathrm{d}t\right)\mathrm{d}u\right), \tag{5.27}$$

with $\hat{F}_n^V$ as in (5.25) and $\hat{m}_n$ as in (5.26). That is, $\hat{F}_n$ is the unique function $\hat{F}_n \in \mathcal{F}_+$ which satisfies for all $z \geq 0$: $V(z; \hat{F}_n, \hat{m}_n) = \hat{F}_n^V(z)$.

We again discuss why $\hat{F}_n$ is well-defined. If there exists a function $\hat{F}_n \in \mathcal{F}_+$ which satisfies $V(z; \hat{F}_n, \hat{m}_n) = \hat{F}_n^V(z)$ for all $z \geq 0$ then it is unique by Theorem 5.11. From (5.27) we see that $\hat{F}_n$ cannot be the zero function. Moreover, we see that $\hat{F}_n$ should satisfy the following:

$$\begin{aligned} m_{\hat{F}_n} &= \lim_{z\to\infty} \int_0^z \exp\left(-\pi \int_0^u \hat{F}_n(t)\mathrm{d}t\right)\mathrm{d}u \\ &= \hat{m}_n - \lim_{z\to\infty} \frac{\hat{F}_n^V(z) - 1 + \exp\left(-\pi \int_0^z \hat{F}_n(t)\mathrm{d}t\right)}{\pi \hat{F}_n(z)} = \hat{m}_n. \end{aligned}$$

The final equality follows from the fact that $\lim_{z\to\infty} \hat{F}_n^V(z) = 1$ and $\lim_{z\to\infty} \hat{F}_n(z) > 0$, since $\hat{F}_n$ is non-zero. Therefore, $\hat{F}_n$ necessarily satisfies:

$$\hat{F}_n^V(z) = 1 - \exp\left(-\pi \int_0^z \hat{F}_n(t)\mathrm{d}t\right) + \pi \hat{F}_n(z) \int_z^\infty \exp\left(-\pi \int_0^u \hat{F}_n(t)\mathrm{d}t\right)\mathrm{d}u. \tag{5.28}$$

Recall from (5.23) that this means that $\hat{F}_n^V$ is the volume-biased weight distribution induced by $\hat{F}_n$. Suppose $(x_1, h_1), (x_2, h_2), \ldots, (x_m, h_k)$ is the sorted realization of the points of η^* with $x_1, \ldots, x_k \in W_n$ and $h_1 \leq h_2 \leq \cdots \leq h_k$. Clearly, $\hat{F}_n^V(z)$ is piecewise constant, with jump locations at $h_1, \ldots, h_k$. In the proof of Theorem 5.11 we observe that the Lebesgue-Stieltjes measures associated with $\hat{F}_n$ and $V(\cdot; \hat{F}_n, \hat{m}_n) = V(\cdot; \hat{F}_n, m_{\hat{F}_n}) = \hat{F}_n^V$ are mutually absolutely continuous. As a consequence, $\hat{F}_n$ is necessarily also piecewise constant with the same jump locations as $\hat{F}_n^V$. Therefore, we simply need to specify the value of $\hat{F}_n$ at $h_1, \ldots, h_k$. Taking $z = h_1$ in (5.27), and using the fact that $\int_0^{h_1} \hat{F}_n(t)\mathrm{d}t = 0$ we can solve for $\hat{F}_n(h_1)$:

$$\hat{F}_n(h_1) = \frac{\hat{F}_n^V(h_1)}{\pi\,(\hat{m}_n - h_1)}. \tag{5.29}$$

In section 5.11 an explicit formula for $\hat{m}_n$ is given, which also shows that $\hat{m}_n > h_1$. Let $i \in \{2, \ldots, k\}$ then, via equation (5.19) from the proof of Theorem 5.10, it follows that for the $\hat{F}_n$ we are looking for:

$$\begin{aligned} \hat{F}_n^V(h_i) &= \hat{F}_n^V(h_{i-1}) + \pi \int_{h_{i-1}}^{h_i} \int_h^\infty \exp\left(-\pi \int_0^u \hat{F}_n(t)\mathrm{d}t\right)\mathrm{d}u\mathrm{d}\hat{F}_n(h) \\ &= \hat{F}_n^V(h_{i-1}) + \pi \int_{h_i}^\infty \exp\left(-\pi \int_0^u \hat{F}_n(t)\mathrm{d}t\right)\mathrm{d}u \left(\hat{F}_n(h_i) - \hat{F}_n(h_{i-1})\right). \end{aligned}$$

Hence,

$$\pi \int_{h_i}^{\infty} \exp\left(-\pi \int_0^u \hat{F}_n(t)\mathrm{d}t\right) \mathrm{d}u = \frac{\hat{F}_n^V(h_i) - \hat{F}_n^V(h_{i-1})}{\hat{F}_n(h_i) - \hat{F}_n(h_{i-1})}. \tag{5.30}$$

Equation (5.28) may be used to obtain an expression for $\hat{F}_n^V(h_i)$, plugging (5.30) into this expression and solving for $\hat{F}_n(h_i)$ yields:

$$\begin{aligned}\hat{F}_n(h_i) &= \hat{F}_n(h_{i-1}) \left(\frac{\hat{F}_n^V(h_i) - 1 + \exp\left(-\pi \int_0^{h_i} \hat{F}_n(t)\mathrm{d}t\right)}{\hat{F}_n^V(h_{i-1}) - 1 + \exp\left(-\pi \int_0^{h_i} \hat{F}_n(t)\mathrm{d}t\right)} \right) \\ &= \hat{F}_n(h_{i-1}) \left(\frac{\hat{F}_n^V(h_i) - 1 + \exp\left(-\pi \sum_{j=1}^{i-1} (h_i - h_j) \left(\hat{F}_n(h_j) - \hat{F}_n(h_{j-1})\right)\right)}{\hat{F}_n^V(h_{i-1}) - 1 + \exp\left(-\pi \sum_{j=1}^{i-1} (h_i - h_j) \left(\hat{F}_n(h_j) - \hat{F}_n(h_{j-1})\right)\right)} \right).\end{aligned} \tag{5.31}$$

Note that the RHS of (5.31) only depends on the values $\hat{F}_n(h_j)$ with $j < i$. Hence, (5.29) along with (5.31) completely defines $\hat{F}_n$. From (5.31) it is evident that $\hat{F}_n \in \mathcal{F}_+$, and this expression may be used to compute $\hat{F}_n$ in practice.

5.5.2 Consistency

In this section we show that $\hat{F}_n$, as in Definition 5.12, is a strongly consistent estimator for F. We start with a Lemma which implies that $\hat{m}_n$ is a strongly consistent estimator for m_F.

Lemma 5.4. *Let $(F_n)_{n\geq 1}$ be a sequence in $\mathcal{F}_+$, and let $F \in \mathcal{F}_+$ be non-zero. If $\lim_{n\to\infty} F_n(z) = F(z)$ for all $z \geq 0$, then $\lim_{n\to\infty} m_{F_n} = m_F$.*

Next, we show that $\tilde{F}_n^V$ and $\hat{F}_n^V$ are strongly consistent and uniformly strongly consistent estimators of F^V respectively.

Lemma 5.5. *With probability one, $\lim_{n\to\infty} \tilde{F}_n^V(z) = F^V(z)$ for all $z \geq 0$. Additionally, with probability one we have $\lim_{n\to\infty} \|\hat{F}_n^V - F^V\|_\infty = 0$. Here, $\tilde{F}_n^V$ and $\hat{F}_n^V$ are given by (5.24) and (5.25) respectively.*

Proof. We first show that with probability one, $\lim_{n\to\infty} \tilde{F}_n^V(z) = F^V(z)$ for all $z \geq 0$. Let $z \geq 0$, by Lemma 5.2 it is sufficient to show that $\lim_{n\to\infty} \tilde{F}_n^V(z) = F^V(z)$ almost surely. Again, we apply the spatial ergodic theorem (Theorem 5.7). This can be done since for all $(x, h) \in \eta$ we have: $C((x, h), \eta) - x = C((0, h), S_x\eta)$. Hence, by the translation invariance of Lebesgue measure, $\tilde{F}_n^V(z)$ may be written as:

$$\tilde{F}_n^V(z) = \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)\nu_d\left(C((0, h), S_x\eta)\right).$$

So indeed, the spatial ergodic theorem yields $\lim_{n\to\infty} \tilde{F}_n^V(z) = F^V(z)$ almost surely. Similarly, we may argue that $\lim_{n\to\infty} \tilde{F}_n^V(\infty) = 1$ almost surely. Since $\hat{F}_n^V(z) = \tilde{F}_n^V(z)/\tilde{F}_n^V(\infty)$,

we obtain via the continuous mapping theorem that $\lim_{n\to\infty} \hat{F}_n^V(z) = F^V(z)$ almost surely. The uniform strong consistency follows from repeating the steps in the proof of the Glivenko-Cantelli theorem. □

We need one more lemma before we prove the consistency result for $\hat{F}_n$.

Lemma 5.6. *Let $(F_n)_{n\geq 1}$ be a sequence in $\mathcal{F}_+$, and let $F \in \mathcal{F}_+$. Let $(m_n)_{n\geq 1}$ be a sequence in $(0,\infty)$ and let $m > 0$. If $\lim_{n\to\infty} F_n(z) = F(z)$ for all $z \geq 0$ and $\lim_{n\to\infty} m_n = m$, then $\lim_{n\to\infty} V(z; F_n, m_n) = V(z; F, m)$ for all $z \geq 0$.*

Theorem 5.13 (Consistency of $\hat{F}_n$)**.** *With probability one, $\lim_{n\to\infty} \hat{F}_n(z) = F(z)$ for all $z \geq 0$.*

Proof. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting a Poisson process η, with intensity measure $\nu_2 \times \mathbb{F}$. By Lemma 5.4 and Lemma 5.5 there exists a set $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ and $z \geq 0$ we have $\lim_{n\to\infty} \hat{F}_n^V(z;\omega) = V(z; F, m_F)$ and $\lim_{n\to\infty} \hat{m}_n(\omega) = m_F$. Let $\omega \in \Omega_0$ and $z \geq 0$, we show that $\lim_{n\to\infty} \hat{F}_n(z;\omega) = F(z)$.

Pick $M > 0$ such that $F(z) < M$. For $n \in \mathbb{N}$ and $h \geq 0$, define: $\bar{F}_n(h) = \min\{\hat{F}_n(h;\omega), M\}$. Then, $(\bar{F}_n)_{n\geq 1}$ is a uniformly bounded sequence of monotone functions. Let $(n_l)_{l\geq 1} \subset (n)_{n\geq 1}$ be an arbitrary subsequence. By Helly's selection principle there exists a further subsequence $(n_k)_{k\geq 1} \subset (n_l)_{l\geq 1}$ such that $\bar{F}_{n_k}$ converges pointwise to some monotone function $\bar{F}$ as $k \to \infty$. This implies that $\lim_{k\to\infty} \hat{F}_{n_k}(h;\omega) = \lim_{k\to\infty} \bar{F}_{n_k}(h) = \bar{F}(h)$ for all $h \in [0, R)$ with $R := \sup\{h \geq 0 : \bar{F}(h) < M\}$. By Lemma 5.6 we obtain along this subsequence:

$$\lim_{k\to\infty} \hat{F}_{n_k}^V(h) := \lim_{k\to\infty} V(h; \hat{F}_{n_k}(\,\cdot\,,\omega), \hat{m}_{n_k}(\omega)) = V(h; \bar{F}, m_F) \text{ for all } h \in [0, R).$$

Because the whole sequence $\hat{F}_n^V(h;\omega)$ converges to $V(h; F, m_F)$ as $n \to \infty$, for $h \geq 0$, we obtain $V(h; F, m_F) = V(h; \bar{F}, m_F)$ for all $h \in [0, R)$. Theorem 5.11 now yields $F(h) = \bar{F}(h)$ for all $h \in [0, R)$, since $z \in [0, R)$, we have in particular $F(z) = \bar{F}(z)$. As a consequence: $\lim_{k\to\infty} \hat{F}_{n_k}(z;\omega) = F(z)$. Because the initial subsequence was chosen arbitrarily, the whole sequence converges: $\lim_{n\to\infty} \hat{F}_n(z;\omega) = F(z)$. □

Remark 5.3 (Density estimation)**.** *For many practical applications it is reasonable to assume $\mathbb{F}$ is absolutely continuous with respect to Lebesgue measure and has a density f. One could then define estimators for f via kernel smoothing. Let k be a symmetric kernel, $\tau > 0$ and $z \in \mathbb{R}$, then we may define:*

$$\hat{f}_n^0(z) = \int_0^\infty \frac{1}{\tau} k\left(\frac{z-t}{\tau}\right) \mathrm{d}\hat{F}_n^0(t) \quad \text{and} \quad \hat{f}_n(z) = \int_0^\infty \frac{1}{\tau} k\left(\frac{z-t}{\tau}\right) \mathrm{d}\hat{F}_n(t). \tag{5.32}$$

In the classical context of kernel density estimation it is well-known which choices of the bandwidth parameter τ lead to consistent and/ or optimal rates of convergence, see for instance chapter 24 in [109]. We also refer to [36] for various examples of density estimators obtained via smoothing of estimators of distribution functions. In our setting it is not yet clear which choices of τ lead to consistent estimators, because much of the behavior of the estimators $\hat{F}_n$ and $\hat{F}_n^0$ is unknown. In practice one can still apply the density estimators in (5.32) by manually choosing a value for τ, being careful to take a value which does not lead to under- or oversmoothing.

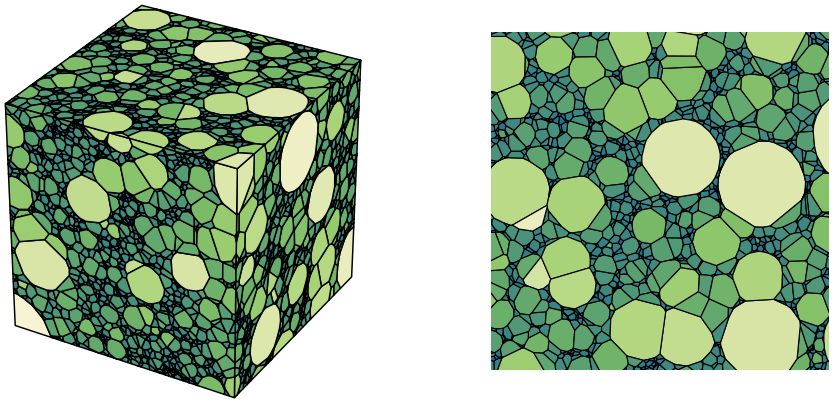


Figure 5.4: A realization of a 3D Poisson-Laguerre tessellation and a corresponding 2D sectional Poisson-Laguerre tessellation. Cells are colored according to their 3D volume.

5.6 Stereology

In this section we study a special type of Poisson-Laguerre tessellations, namely sectional Poisson-Laguerre tessellations. By this we mean that we intersect a Poisson-Laguerre tessellation with a hyperplane, and we consider the resulting tessellation in this hyperplane. In Theorem 4.1. in [39] it was shown that intersecting a Poisson-Laguerre tessellation in $\mathbb{R}^d$ with a hyperplane, yields a tessellation in this hyperplane which is again a Poisson-Laguerre tessellation. Because our parameterization is subtly different to the setting in [39] we derive the intensity measure of the Poisson process corresponding to the sectional Poisson-Laguerre tessellation, for which we also use a different argument.

Suppose we observe the extreme points and the corresponding cells, of a Poisson-Laguerre tessellation $L(\eta)$ in $\mathbb{R}^{d-1}$, through the observation window W_n. The underlying Poisson process η has intensity measure $\nu_{d-1} \times \mathbb{F}$, where $\mathbb{F}$ is a locally finite measure concentrated on $(0, \infty)$. Hence, we may use any of the estimators in the previous two sections to estimate $F(z) = \mathbb{F}((0, z])$. Throughout this section, we assume that $\bar{F}_n$ is a piecewise constant, strongly consistent estimator for F. Now, this Poisson-Laguerre tessellation in $\mathbb{R}^{d-1}$ is the sectional tessellation corresponding to a Poisson-Laguerre tessellation $L(\Psi)$ in $\mathbb{R}^d$. The Poisson process Ψ of this higher-dimensional tessellation has intensity measure $\nu_d \times \mathbb{H}$, where $\mathbb{H}$ is a locally finite measure concentrated on $(0, \infty)$. For $z \geq 0$ define: $H(z) := \mathbb{H}((0, z])$. In this section, we show how F is related to H, and how a consistent estimator for F can be used to obtain a (locally) consistent estimator for H. Thereby, we have a solution to the stereological problem. First, we need the following lemma for obtaining an expression for F in this stereological setting:

Lemma 5.7. *Let $\varphi \subset \mathbb{R}^d \times (0, \infty)$ be an at most countable set. Let $\theta \in \mathbb{S}^{d-1}$ and $s \in \mathbb{R}$. Define the hyperplane $T := \{y \in \mathbb{R}^d : \langle \theta, y \rangle = s\}$. For $(x, h) \in \varphi$, with $x \in \mathbb{R}^d$ and $h > 0$*

let:

$$x' := x - (\langle \theta, x\rangle - s)\theta$$
$$h' := h + \|x' - x\|^2 = h + (\langle \theta, x\rangle - s)^2.$$

Note that $x' \in T$ *and define* $\varphi' := \{(x', h') : (x, h) \in \varphi\}$. *Then, for all* $(x, h) \in \varphi$: $C((x, h), \varphi) \cap T = C'((x', h'), \varphi')$ *with:*

$$C'((x', h'), \varphi') = \{y \in T : \|y - x'\|^2 + h' \leq \|y - \bar{x}\|^2 + \bar{h} \text{ for all } (\bar{x}, \bar{h}) \in \varphi'\}.$$

Proof. Let $y \in T$ and $(x, h) \in \varphi$, then a direct computation yields:

$$\|x' - y\|^2 = \|x - y\|^2 - 2(\langle \theta, x\rangle - s)\langle x - y, \theta\rangle + (\langle \theta, x\rangle - s)^2 = \|x - y\|^2 - h' + h.$$

Since $\|x' - y\|^2 + h' = \|x - y\|^2 + h$, the claim follows. □

This lemma describes the set of weighted points which generates a sectional Laguerre diagram. We now apply this to the Poisson-Laguerre tessellation generated by the Poisson process Ψ. Because a Poisson-Laguerre tessellation is stationary and isotropic the choice of hyperplane does not affect the distribution of the sectional tessellation. For $x \in \mathbb{R}^d$ write: $x = (\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_d)$. We choose the hyperplane $\mathrm{x}_d = 0$ which corresponds to taking $\theta = (0, \ldots, 0, 1) \in \mathbb{S}^{d-1}$ and $s = 0$ in Lemma 5.7. In view of Lemma 5.7 consider the function which maps a pair $(x, h) \in \Psi$ to the corresponding (x', h'). Hence, this function is given by $(\mathrm{x}_1, \ldots, \mathrm{x}_d, h) \mapsto (\mathrm{x}_1, \ldots, \mathrm{x}_{d-1}, 0, h + \mathrm{x}_d^2)$. Naturally, the d-th component of the resulting vector is always zero. We identify the hyperplane $\mathrm{x}_d = 0$ with $\mathbb{R}^{d-1}$ and therefore we consider the function $\tau : \mathbb{R}^d \times (0, \infty) \to \mathbb{R}^{d-1} \times (0, \infty)$ which is defined via: $\tau(x, h) = (\mathrm{x}_1, \ldots, \mathrm{x}_{d-1}, h + \mathrm{x}_d^2)$. Hence, the point process $\eta := \tau(\Psi)$ generates the sectional tessellation. By the mapping theorem (see Theorem 5.1 in [53]) η is again a Poisson process on $\mathbb{R}^{d-1} \times (0, \infty)$ with intensity measure: $\mathbb{E}(\Psi(\tau^{-1}(\cdot)))$. Let $B \subset \mathbb{R}^{d-1}$ be a Borel set and let $z \geq 0$. Note that $h + \mathrm{x}_d^2 \leq z$ if and only if $h \leq z$ and $\mathrm{x}_d \in [-\sqrt{z - h}, \sqrt{z - h}]$. As a result:

$$\tau(x, h) \in B \times (0, z] \iff x \in B \times \left[-\sqrt{z - h}, \sqrt{z - h}\right] \text{ and } h \leq z. \tag{5.33}$$

Via the Campbell formula and (5.33) we find:

$$\begin{aligned}
\mathbb{E}\left(\Psi(\tau^{-1}(B \times (0, z]))\right) &= \int_{\mathbb{R}^d} \int_0^\infty \mathbb{1}\{\tau(x, h) \in B \times (0, z]\} \mathrm{d}H(h)\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \int_0^z \mathbb{1}\left\{x \in B \times \left[-\sqrt{z - h}, \sqrt{z - h}\right]\right\} \mathrm{d}H(h)\mathrm{d}x \\
&= \nu_{d-1}(B) 2 \int_0^z \sqrt{z - h} \mathrm{d}H(h).
\end{aligned}$$

Hence, we obtain:

$$F(z) = 2 \int_0^z \sqrt{z - h} \mathrm{d}H(h).$$

Let us discuss some properties of this function F. First of all, F is not a bounded function. Indeed, choose $z_0 > 0$ such that $H(z_0) > 0$, and let $z > z_0$, via integration by parts we observe:

$$F(z) = \int_0^z H(h) \frac{1}{\sqrt{z-h}} \mathrm{d}h \geq \int_{z_0}^z H(h) \frac{1}{\sqrt{z}} \mathrm{d}h \geq H(z_0) \frac{z - z_0}{\sqrt{z}}. \tag{5.34}$$

It immediately follows that $\lim_{z\to\infty} F(z) = \infty$. Another property of F is that it is absolutely continuous, and has a Lebesgue density f given by:

$$f(z) = \int_0^z \frac{1}{\sqrt{z-t}} \mathrm{d}H(t).$$

Indeed, via Fubini's theorem we can verify that f is a density of F:

$$\begin{aligned}
\int_0^z f(s)\mathrm{d}s &= \int_0^z \int_0^s \frac{1}{\sqrt{s-t}} \mathrm{d}H(t)\mathrm{d}s \\
&= \int_0^z \int_t^z \frac{1}{\sqrt{s-t}} \mathrm{d}s\mathrm{d}H(t) \\
&= 2\int_0^z \sqrt{z-t}\mathrm{d}H(t) = F(z).
\end{aligned}$$

It is possible to express H in terms of F, because this is an Abel integral equation. For a direct derivation of the inversion formula see for example [99]. Here, we simply show that the following expression is indeed an inversion formula for $H(z)$:

$$\begin{aligned}
\frac{1}{\pi}\int_0^z \frac{1}{\sqrt{z-t}} \mathrm{d}F(t) &= \frac{1}{\pi}\int_0^z \frac{1}{\sqrt{z-t}} \int_0^t \frac{1}{\sqrt{t-s}} \mathrm{d}H(s)\mathrm{d}t \\
&= \int_0^z \int_s^z \frac{1}{\pi} \frac{1}{\sqrt{z-t}\sqrt{t-s}} \mathrm{d}t\mathrm{d}H(s) \\
&= \int_0^z \int_0^1 \frac{1}{\pi}(1-u)^{-\frac{1}{2}} u^{-\frac{1}{2}} \mathrm{d}u\mathrm{d}H(s) \qquad (5.35) \\
&= H(z). \qquad (5.36)
\end{aligned}$$

In (5.35) we substituted $u = (t-s)/(z-s)$. Finally, (5.36) follows from the fact that the inner integral in (5.35) is equal to one, since this integral represents the Beta function evaluated in $(1/2, 1/2)$. A plugin estimator for $H(z)$ is therefore given by:

$$H_n(z) := \frac{1}{\pi}\int_0^z \frac{1}{\sqrt{z-t}} \mathrm{d}\bar{F}_n(t),$$

where $\bar{F}_n$ is a piecewise constant, strongly consistent estimator for F. This estimator is however rather ill-behaved. While H is a monotone function, H_n is not. Because $\bar{F}_n$ is piecewise constant, H_n is decreasing between jump locations of $\bar{F}_n$. Moreover, if z_0 is a jump location of $\bar{F}_n$, then $\lim_{z\downarrow z_0} H_n(z) = \infty$. Therefore, we use isotonization to obtain an estimator for H which is monotone, and show that it is consistent. We note that our

estimator is similarly defined as the isotonic estimator in [35]. For the remainder of this section, let $k = k(n)$ be the number of jump locations of $\bar{F}_n$. Let $h_1, h_2, \ldots, h_k$ with $0 < h_1 < h_2 < \cdots < h_k < \infty$ be the jump locations of $\bar{F}_n$. In order to introduce the isotonic estimator we define for $z \geq 0$:

$$U_n(z) := \int_0^z H_n(t)\mathrm{d}t = \frac{2}{\pi}\int_0^z \sqrt{z-t}\mathrm{d}\bar{F}_n(t). \tag{5.37}$$

Choose (a large) $M > 0$ and write $z_M := \min\{h_k, M\}$. Let U_n^M be the greatest convex minorant of U_n on $[0, z_M]$. That is, U_n^M is the greatest convex function on $[0, z_M]$ which lies below U_n. Then, define:

$$\hat{H}_n^M(z) := \begin{cases} U_n^{M,r}(z) & \text{if } z \in \big[0, z_M\big) \\ U_n^{M,l}(z_M) & \text{if } z \geq z_M, \end{cases} \tag{5.38}$$

where $U_n^{M,l}$, $U_n^{M,r}$ denote the left- and right-derivative of U_n^M respectively. The reason we cannot simply extend the definition of U_n^M to the whole of $[0, \infty)$ is due to the fact that U_n is concave on $[h_k, \infty)$. As a result, the greatest convex minorant of U_n on $[0, \infty)$ is the zero function. Because of the convexity of U_n^M on $[0, z_M]$, $\hat{H}_n^M$ is guaranteed to be non-decreasing, and is referred to as an isotonic estimator. Analogously to (5.37) we define for $z \geq 0$:

$$U(z) := \int_0^z H(t)\mathrm{d}t = \frac{2}{\pi}\int_0^z \sqrt{z-t}\mathrm{d}F(t). \tag{5.39}$$

Note that $U^r(z) = H(z)$, so indeed, the right-derivative of U_n^M is a natural choice for an estimator of H. In the next theorem we prove consistency of $\hat{H}_n^M$. Currently, it is not known whether $\hat{H}_n := \hat{H}_n^\infty$ is a globally consistent estimator.

Theorem 5.14 (Consistency of $\hat{H}_n^M$)**.** *Let $M > 0$ and let $\hat{H}_n^M$ be as in (5.38). Let $z \in [0, M)$, then with probability one:*

$$H(z-) \leq \liminf_{n\to\infty} \hat{H}_n^M(z) \leq \limsup_{n\to\infty} \hat{H}_n^M(z) \leq H(z).$$

In particular, if z is a continuity point of H: $\lim_{n\to\infty} \hat{H}_n^M(z) = H(z)$ almost surely.

Proof. Let $z \in [0, M)$. Because $\bar{F}_n$ is piecewise constant and a consistent estimator of the unbounded function F (recall equation (5.34)), it follows that $\lim_{n\to\infty} h_{k(n)} = \infty$ almost surely. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting a Poisson process η, with intensity measure $\nu_{d-1} \times \mathbb{F}$. Choose $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ we have $\lim_{n\to\infty} h_{k(n)}(\omega) = \infty$ and $\lim_{n\to\infty} \bar{F}_n(h;\omega) = F(h)$ for all $h \geq 0$. For the remainder of the proof, let $\omega \in \Omega_0$ and take n sufficiently large such that $h_{k(n)}(\omega) > M$. Note how U_n and U depend on $\bar{F}_n$ and F respectively, see (5.37) and (5.39). As a consequence, the pointwise convergence of $\bar{F}_n(\,\cdot\,;\omega)$ to F implies: $\lim_{n\to\infty} U_n(x;\omega) = U(x)$ for all $x \geq 0$. Note that U is non-decreasing and continuous, therefore the convergence is also uniform on $[0, M]$. That is, $\lim_{n\to\infty} \sup_{x\in[0,M]} |U_n(x;\omega) - U(x)| = 0$. Because U is defined as

the integral of a non-decreasing function, it is convex. A variant of Marshall's lemma (the convex analogue of 7.2.3. on p. 329 in [79]) directly yields:

$$\sup_{x\in[0,M]} \left|U_n^M(x;\omega) - U(x)\right| \leq \sup_{x\in[0,M]} |U_n(x;\omega) - U(x)|\,.$$

Therefore, we also have $\lim_{n\to\infty} \sup_{x\in[0,M]} \left|U_n^M(x;\omega) - U(x)\right| = 0$. Take $\delta > 0$ such that $z + \delta < M$. Then, for each $0 < h < \delta$ we have by the convexity of U_n^M:

$$\frac{U_n^M(z;\omega) - U_n^M(z-h;\omega)}{h} \leq U_n^{M,l}(z;\omega) \leq U_n^{M,r}(z;\omega) \leq \frac{U_n^M(z;\omega) - U_n^M(z+h;\omega)}{h}.$$

By using $\lim_{n\to\infty} \sup_{x\in[0,M]} \left|U_n^M(x;\omega) - U(x)\right| = 0$, the following holds:

$$\frac{U(z) - U(z-h)}{h} \leq \liminf_{n\to\infty} U_n^{M,r}(z;\omega) \leq \limsup_{n\to\infty} U_n^{M,r}(z;\omega) \leq \frac{U(z) - U(z+h)}{h}.$$

The result follows from letting $h \downarrow 0$ and by recognizing that $U^l(z) = H(z-)$ and $U^r(z) = H(z)$. □

Remark 5.4. *By choosing $M > 0$ very large, the estimators $\hat{H}_n := \hat{H}_n^\infty$ and $\hat{H}_n^M$ will in practice often coincide, since we will typically observe $h_k < M$. Therefore, in the remainder of this chapter we will only consider computational aspects and simulation performance of the estimator $\hat{H}_n$.*

We now show that computing the isotonic estimator $\hat{H}_n$ is equivalent to solving an isotonic regression problem. This is achieved via the following lemma, which is a straightforward modification of Lemma 2 in [35].

Lemma 5.8. *Let $M > 0$, and let φ be an a.e. continuous non-negative function on $[0, M]$. Define the function Φ, for $z \geq 0$ as:*

$$\Phi(z) = \int_0^z \varphi(x)\mathrm{d}x.$$

Let Φ^* be the greatest convex minorant of Φ on $[0, M]$. Let $\Phi^{*,r}$ be the right-derivative of Φ^*, then:

$$\int_0^M (\varphi(x) - \psi(x))^2\mathrm{d}x \geq \int_0^M (\varphi(x) - \Phi^{*,r}(x))^2\mathrm{d}x + \int_0^M (\Phi^{*,r} - \psi(x))^2\mathrm{d}x,$$

for all functions ψ in the set:

$$\mathcal{F}_M := \{\psi : [0, M] \to [0, \infty) : \psi \text{ is non-decreasing and right-continuous}\}\,.$$

We use Lemma 5.8 to show that $\hat{H}_n$ may be interpreted as the L^2-projection of H_n on the space of monotone functions. Recall that $h_1, h_2, \ldots, h_k$ are the unique jump locations

of $\bar{F}_n$. Additionally, let $h_0 = 0$. Define $\tilde{H}_n$ to be the piece-wise constant function on $[0, h_k)$ which is given by:

$$\tilde{H}_n(z) = \frac{U_n(h_{i+1}) - U_n(h_i)}{h_{i+1} - h_i}, \qquad z \in [h_i, h_{i+1}), \;\; i \in \{0, 1, \ldots, k-1\}.$$

For $z \in [0, h_k]$, let: $\tilde{U}_n(z) = \int_0^z \tilde{H}_n(t)\mathrm{d}t$. Then, $U_n(h_i) = \tilde{U}_n(h_i)$ for all $i \in \{0, 1, \ldots, k\}$. While U_n is concave between successive jump locations (due to the square root), $\tilde{U}_n$ is linear between successive jump locations. As a consequence, U_n and $\tilde{U}_n$ have the same greatest convex minorant. Hence, $\hat{H}_n(z) = U_n^{*,r}(z) = \tilde{U}_n^{*,r}(z)$, for $z \in [0, h_k)$. Finally, by taking $\varphi = \tilde{H}_n$ (and $\varphi = H_n$) and $M = h_k$ in Lemma 5.8 we see that:

$$U_n^{*,r} = \operatorname*{arg\,min}_{H \in \mathcal{F}_{h_k}} \int_0^{h_k} (H(x) - H_n(x))^2 \mathrm{d}x = \operatorname*{arg\,min}_{H \in \mathcal{F}_{h_k}} \int_0^{h_k} (H(x) - \tilde{H}_n(x))^2 \mathrm{d}x. \tag{5.40}$$

Because $\tilde{H}_n$ is piece-wise constant, in (5.40) we may even minimize over all functions in $\mathcal{F}_{h_k}$ which are also piece-wise constant with jump locations at $h_1, h_2, \ldots, h_k$. Hence, $\hat{H}_n$ is piece-wise constant and when solving the minimization problem in (5.40) we only seek to determine the values $\hat{H}_n$ attains at these jump locations. Let $y_i = \tilde{H}_n(h_i)$, and $w_i = h_{i+1} - h_i$. Then, by setting $\hat{\beta} = (\hat{H}_n(h_1), \hat{H}_n(h_2), \ldots, \hat{H}_n(h_{k-1}))$, (5.40) may be written as:

$$\hat{\beta} = \operatorname*{arg\,min}_{\beta \in C_+} \sum_{i=1}^{k-1} (\beta_i - y_i)^2 w_i, \tag{5.41}$$

where the closed convex cone C_+ is given by: $C_+ := \{\beta \in \mathbb{R}^{k-1} : 0 \le \beta_1 \le \beta_2 \le \cdots \le \beta_{k-1}\}$. Finally, observe that $\hat{H}_n(h_k) = \hat{H}_n(h_{k-1})$. The optimization problem in (5.41) is indeed an isotonic regression problem. We note that implementations for solving this problem are widely available. In Figure 5.5 a realization of H_n and the corresponding realization of $\hat{H}_n$ is shown.

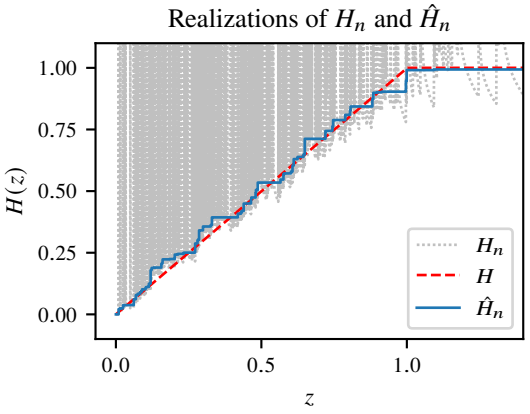


Figure 5.5: A comparison of the plugin estimator H_n and the isotonic estimator $\hat{H}_n$. The actual underlying H is equal to the CDF of a uniform distribution on $(0, 1)$.

5.7 Simulations

In previous sections we have derived consistent estimators for the distribution function corresponding to the underlying Poisson process η, both in the direct setting ($\mathbb{R}^d$) and in the stereological setting ($\mathbb{R}^{d-1}$). Additionally, we have shown how to compute these estimators. In this section we perform some simulations such that we can assess their performance. We note that edge effects which occur in practice, were not taken into account for the simulations in this section. We discuss practical issues arising from edge effects in section 5.9. For the simulations we compute Laguerre tessellations using the Voro++ software [82]. For the estimators $\hat{F}_n^0$ and $\hat{F}_n$ we focus on the case $d = 2$. Let $M > 0$ and $z \geq 0$, we consider the following choices for the underlying F.

$$F_1(z; M) = z \cdot \mathbb{1}\{z < M\} + M \cdot \mathbb{1}\{z \geq M\} \tag{5.42}$$

$$F_2(z) = 0.01 \cdot \mathbb{1}\{z \geq 1\} + 0.04 \cdot \mathbb{1}\{z \geq 8\} + 0.95 \cdot \mathbb{1}\{z \geq 10\}. \tag{5.43}$$

Note that F_2 corresponds to the F in Example 2. For both choices of F it is simple to simulate a corresponding Poisson process, because these Poisson processes can be recognized as independently marked homogeneous Poisson processes. Throughout this section

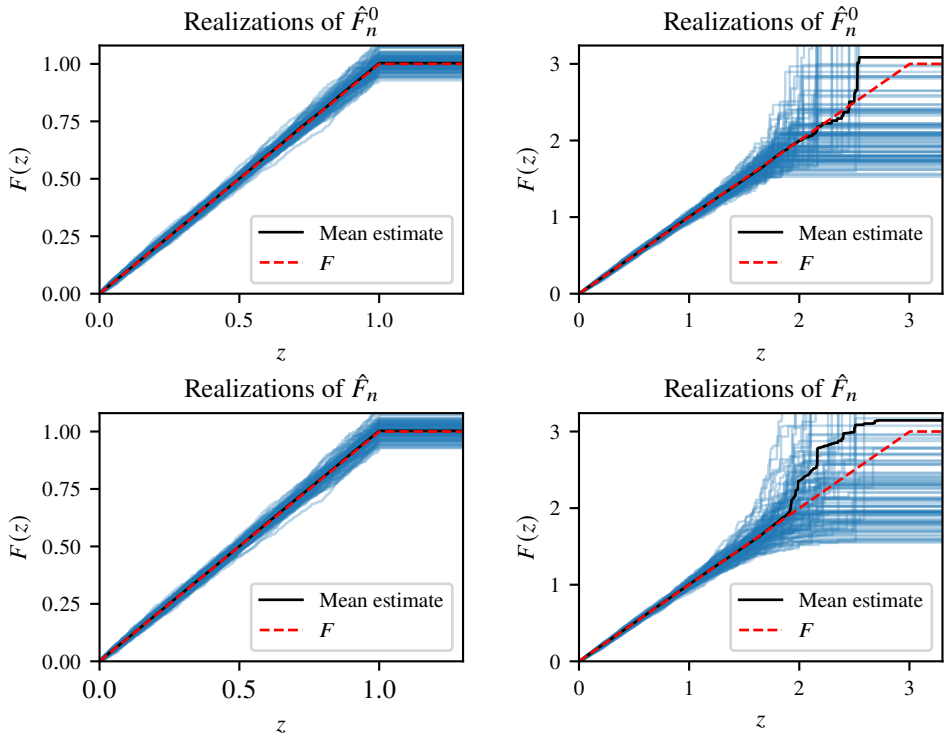


Figure 5.6: Simulation results for $\hat{F}_n^0$ and $\hat{F}_n$, where F is given by (5.42), with $M = 1$ (Left) and $M = 3$ (Right).

we write $P_n := \mathbb{E}(\eta^0(W_n \times (0,\infty)))$. We choose a square observation window W_n such that $P_n = 1000$. In words, we choose a square W_n with an area such that the expected number of observed points of η^0 in W_n is equal to 1000. First, we consider F_1 as the underlying truth. For $M = 1$ and $M = 3$ we repeat the simulation procedure with this F 100 times, such that we obtain 100 realizations of $\hat{F}_n^0$ and $\hat{F}_n$ for each value of M. For each estimator, and each choice of M, we also compute the pointwise average of all realizations. The results are shown in Figure 5.6. A blue line is a realization of an estimator, a black line is a pointwise average. We can clearly see that estimates of $F(z)$ for z close to zero are much more accurate than estimates of $F(z)$ for large values of z. This is especially evident for the results corresponding to $M = 3$. This is not too surprising in view of the crystallization interpretation of a Laguerre tessellation as described in section 5.3. We expect that points with large weights are less likely to generate non-empty cells. As a result we sample points with large weights less often, which makes estimation of $F(z)$ for large values of z more difficult. This also means that we expect that the accuracy of an estimate of F near zero is much more important if we wish to use this estimate to simulate a Poisson-Laguerre tessellation which is similar to the observed tessellation. From Figure 5.6 it is not very clear whether there are significant differences between $\hat{F}_n^0$ and $\hat{F}_n$, though it does seem that $\hat{F}_n^0$ performs slightly better on average when z is large.

Now, we consider F_2 as the underlying truth. For this choice of F we only observe points with weights in the set $\{1, 8, 10\}$. As a result, realizations of $\hat{F}_n^0$ and $\hat{F}_n$ will only have jumps at these values. We can therefore easily quantify the error of $\hat{F}_n^0$ by computing $F(z) - \hat{F}_n^0(z)$ for $z \in \{1, 8, 10\}$. Of course, we can do the same for $\hat{F}_n$. Again, we repeat the simulation procedure 100 times. This time however, we also repeat this for multiple choices of observation windows. We choose W_n such that P_n is equal to 500, 1000, 2000 or 5000. The simulation results are shown in Tables 5.1 and 5.2. This table contains the mean over all 100 absolute errors for each choice of W_n and for each choice of z. We also

	$F(1) - \hat{F}_n^0(1)$		$F(8) - \hat{F}_n^0(8)$		$F(10) - \hat{F}_n^0(10)$	
P_n	mean $(\lvert\cdot\rvert)$	(2.5%, 97.5%)	mean $(\lvert\cdot\rvert)$	(2.5%, 97.5%)	mean $(\lvert\cdot\rvert)$	(2.5%, 97.5%)
500	0.002 84	(−0.0057, 0.0052)	0.007 34	(−0.018 , 0.018)	0.0430	(−0.11 , 0.10)
1000	0.001 97	(−0.0042, 0.0046)	0.004 24	(−0.010 , 0.0099)	0.0327	(−0.065, 0.091)
2000	0.001 45	(−0.0034, 0.0029)	0.003 49	(−0.0079, 0.0081)	0.0192	(−0.042, 0.046)
5000	0.000 845	(−0.0019, 0.0020)	0.002 06	(−0.0046, 0.0052)	0.0146	(−0.030, 0.031)

Table 5.1: Simulation results for $\hat{F}_n^0$, where F is given by (5.43).

	$F(1) - \hat{F}_n(1)$		$F(8) - \hat{F}_n(8)$		$F(10) - \hat{F}_n(10)$	
P_n	mean $(\lvert\cdot\rvert)$	(2.5%, 97.5%)	mean $(\lvert\cdot\rvert)$	(2.5%, 97.5%)	mean $(\lvert\cdot\rvert)$	(2.5%, 97.5%)
500	0.002 94	(−0.0069, 0.0057)	0.007 62	(−0.019 , 0.017)	0.386	(−1.9 , 0.47)
1000	0.001 98	(−0.0039, 0.0050)	0.004 42	(−0.010 , 0.010)	0.267	(−1.1 , 0.33)
2000	0.001 51	(−0.0037, 0.0031)	0.003 54	(−0.0074, 0.0083)	0.170	(−0.61, 0.25)
5000	0.000 885	(−0.0020, 0.0020)	0.002 13	(−0.0048, 0.0045)	0.107	(−0.30, 0.21)

Table 5.2: Simulation results for $\hat{F}_n$, where F is given by (5.43).

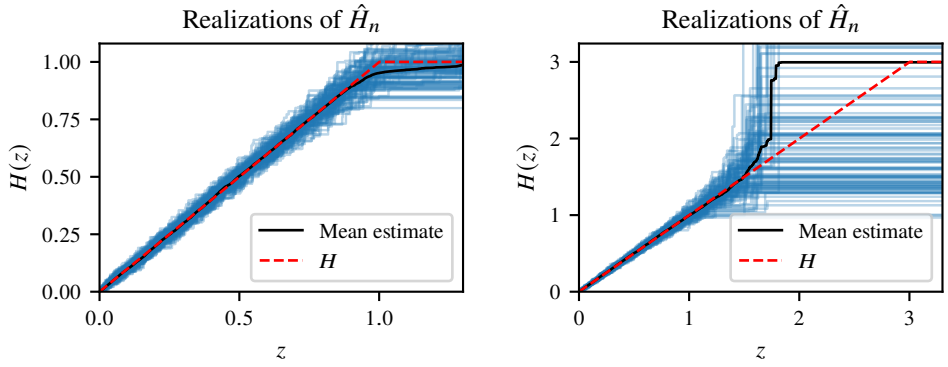


Figure 5.7: Simulation results for $\hat{H}_n$ where H is given by (5.42), with $M = 1$ (Left) and $M = 3$ (Right).

include the 2.5% and 97.5% quantiles of these 100 errors. We can see that at $z = 1$ and $z = 8$ the performance of the estimators $\hat{F}_n^0$ and $\hat{F}_n$ is quite similar. However, at $z = 10$ it is clear that $\hat{F}_n^0$ performs much better. This is somewhat surprising, after all, $\hat{F}_n$ takes into account more information than $\hat{F}_n^0$. We do not yet know whether the difference in performance is due to differences in numerical stability of the inversion procedures or due to different rates of convergence of the estimators. We also should point out that for a single realization of $\hat{F}_n$ corresponding to $P_n = 500$ we observed a numerical overflow. That is, we observed: $\hat{F}_n(10) < \hat{F}_n(8)$. It may therefore be of future interest to study whether there are more numerically stable ways to compute $\hat{F}_n$.

Finally, we show some simulation results for $\hat{H}_n$. In the previous simulations we observed that $\hat{F}_n^0$ performs better than $\hat{F}_n$. Therefore, we compute $\hat{H}_n$ via $\bar{F}_n = \hat{F}_n^0$. We consider $d = 3$ such that we observe a 2D sectional tessellation. We take the underlying H equal to F_1 as in (5.42). Again, we choose W_n such that $P_n = 1000$ and perform 100 repeated simulations for both $M = 1$ and $M = 3$. The results are shown in Figure 5.7. As expected, in the stereological setting we observe a bigger variance in realizations of $\hat{H}_n$ compared to the realizations shown in Figure 5.6. Overall, all estimators seem to perform satisfactorily.

5.8 Connection to the Boolean model

In the previous section we observed that estimates of $F(z)$ for large values of z can be quite inaccurate. We attempt to gain some additional insight in this phenomenon, and try to understand how the behavior of $F(z)$ for large values of z influences the resulting Poisson-Laguerre tessellation. This insight is achieved via a connection between Poisson-Laguerre tessellations and the Boolean model, a classical model in stochastic geometry, which is well-studied. A Boolean model may be described as follows. Let ψ be an independently marked homogeneous Poisson process on $\mathbb{R}^d$ with mark space $(0, \infty)$. Writing $\psi = \{(x_1, r_1), (x_2, r_2), \dots\}$ with $x_i \in \mathbb{R}^d$ and $r_i > 0$, the union set $\cup_{(x,r)\in\psi} \bar{B}(x, r)$ is a

Boolean model with spherical grains. For more details on the Boolean model we refer to [86] and [18].

Recall the interpretation of a Laguerre tessellation as a crystallic growth process. Then, we may interpret F as the distribution of the arrival times of generator points. Intuitively, we can understand that the behavior of $F(z)$ for large values of z has less influence compared to the behavior of $F(z)$ for small values of z. After all, we associate generator points with late arrival times with large values of z. In order to better understand the behavior of the crystallic growth process, we may look at the so-called coverage process and its coverage function, as defined below:

Definition 5.15. For $t \in \mathbb{R}$, define the random closed set $\Xi(t)$ as:

$$\Xi(t) = \bigcup_{(x,h)\in\eta:h\leq t} \bar{B}\left(x, \sqrt{t-h}\right).$$

We refer to $\Xi(t)$ as the coverage process. For a bounded $W \in \mathcal{B}(\mathbb{R}^d)$, with $\nu_d(W) > 0$, define the coverage function $K_F : \mathbb{R} \to [0,1]$ via:

$$K_F(t) = \frac{1}{\nu_d(W)} \mathbb{E}\left(\nu_d(\Xi(t) \cap W)\right).$$

Note that $\Xi(t)$ represents the union of all balls in the crystallic growth process at time t. Also observe that $\Xi(t)$ only depends on the restriction of η to $\mathbb{R}^d \times (0,t]$, and this restriction is an independently marked homogeneous Poisson process. The key insight is that $\Xi(t)$ is an instance of a Boolean model. If $F(t) = 0$ then $\Xi(t) = \emptyset$ almost surely. Now consider $F(t) > 0$. Suppose $Z \sim \mathbb{F}((0,t] \cap \cdot)/F(t)$, then the distribution function of $\sqrt{t-Z}$ is given by $G(z) = 1 - (F(t-z^2)/F(t))$ for $z \in [0, \sqrt{t}]$. Hence, letting $\psi = \{(x_1, r_1), (x_2, r_2), \dots\}$ be an independently marked homogeneous Poisson process on $\mathbb{R}^d$ with intensity $F(t)$, and mark distribution function G on the mark space $[0, \sqrt{t}]$, the random closed set $\cup_{(x,r)\in\psi} \bar{B}(x,r)$ is a Boolean model which is equal in distribution to $\Xi(t)$. The coverage function K_F describes the expected fraction of a set covered by the coverage process at a given time. As a consequence, known results of the Boolean model yield the following, (see section 9.3 in [86], or equations 3.4 and 3.15 in [18]):

Theorem 5.16. *The definition of K_F does not depend on the choice of W, and for $t \in \mathbb{R}$, $K_F(t)$ is given by:*

$$K_F(t) = 1 - \exp\left(-\kappa_d \int_0^t (t-h)^{\frac{d}{2}} \mathrm{d}F(h)\right). \tag{5.44}$$

Moreover, for any convex averaging sequence $(W_n)_{n\geq 1}$ we have with probability one:

$$\lim_{n\to\infty} \frac{\nu_d(\Xi(t) \cap W_n)}{\nu_d(W_n)} = K_F(t).$$

Proof. As noted previously, the result follows from known results of the Boolean model. We derive the expression for $K_F(t)$ for the sake of completeness. Let $y \in \mathbb{R}^d$ and note that:

$$y \notin \Xi(t) \iff \|x - y\| > \sqrt{t-h} \text{ for all } (x,h) \in \eta \cap (\mathbb{R}^d \times (0,t]) \iff \eta(A_y(t)) = 0,$$

with:

$$A_y(t) = \left\{(x,h) \in \mathbb{R}^d \times (0,t] : \|x-y\| \leq \sqrt{t-h}\right\}.$$

Because η is a Poisson process, we may compute:

$$\begin{aligned}\mathbb{P}(y \notin \Xi(t)) &= \mathbb{P}(\eta(A_y(t)) = 0) \\ &= \exp\left(-\int_{\mathbb{R}^d}\int_0^t \mathbb{1}\{\|x-y\| \leq \sqrt{t-h}\}\mathrm{d}F(h)\mathrm{d}x\right) \\ &= \exp\left(-\kappa_d \int_0^t (t-h)^{\frac{d}{2}}\mathrm{d}F(h)\right).\end{aligned}$$

Note that this probability does not depend on y. Via Fubini we obtain the result:

$$\begin{aligned}K_F(t) &= \frac{1}{\nu_d(W)}\mathbb{E}\left(\int_W \mathbb{1}\{y \in \Xi(t)\}\mathrm{d}y\right) \\ &= \frac{1}{\nu_d(W)}\int_W \mathbb{E}\left(\mathbb{1}\{y \in \Xi(t)\}\right)\mathrm{d}y \\ &= \frac{1}{\nu_d(W)}\int_W \mathbb{P}(0 \in \Xi(t))\mathrm{d}y \\ &= 1 - \mathbb{P}(0 \notin \Xi(t)).\end{aligned}$$

□

Remark 5.5. *If $\mathbb{F}$ is instead a locally finite measure concentrated on (t_0, ∞) for some $t_0 \in \mathbb{R}$, then in the expression for $K_F(t)$ in Theorem 5.16 one may replace the 0 as the lower integration bound by t_0.*

In view of the limit result in Theorem 5.16, we may in some sense interpret $K_F(t)$ as the actual fraction of $\mathbb{R}^d$ covered by the coverage process, not just the expected fraction in some region. Suppose that $\tilde{F}$ is an approximation of the function F. This function $\tilde{F}$ satisfies $\tilde{F}(t) = F(t)$ for all $t \in [0,T]$ for some $T > 0$. Moreover, suppose that on the interval (T,∞), $\tilde{F}$ is a poor approximation of F. If $K_F(T)$ is large, say $K_F(T) = 0.9$, then that means that 90% of space is already covered by the coverage process at time T. Weighted generators with weights larger than T, can only contribute to Laguerre cells which cover at most 10% of the whole space. Of course, the balls associated with weights smaller than T will continue growing and therefore it is likely that if the weighted generators with weights larger than T produce a Laguerre cell, those cells will cover far less than 10% of space. As a consequence, if one simulates a Poisson-Laguerre tessellation with either F or with $\tilde{F}$ as the underlying distribution function, it is likely that the effect on the resulting tessellation is rather small.

Based on the previous observations, the coverage function K_F may be useful to determine in which part of the domain of F estimation is hard. Let us now discuss the simulation results of section 5.7 in this context. The simulation results corresponding to the F in (5.42) with $M = 3$ yielded estimates of $F(z)$ which were rather inaccurate for large values of z. For this F we have for $t \in [0,M]$: $K_F(t) = 1 - \exp\left(-\pi t^2/2\right)$. Note that

$K_F(1) \approx 0.79$, which explains that the estimates corresponding to $M = 1$ were quite accurate. Meanwhile, if $M = 3$, then $K_F(1.22) \approx 0.9$, $K_F(1.39) \approx 0.95$ and $K_F(1.72) \approx 0.99$. This does explain why estimation of $F(z)$ is very hard for large values of z if $M = 3$. All things considered, the estimates of F seem rather well-behaved for $z \in [0, 1.72]$.

In future research it may be of interest to study whether this connection between Poisson-Laguerre tessellations and the Boolean model can provide additional insights.

5.9 Discussion

In this chapter we have defined two estimators for the distribution function F, which describes the distribution of the arrival times of the generators. For these estimators we have established their consistency, and studied their performance in simulations. We have also considered statistical inference in a stereological setting. When computing the proposed estimators in practice, one has to deal with edge effects, and we now briefly discuss some of the challenges caused by this phenomenon. Throughout this chapter, all points of η^* in $W_n \times (0, \infty)$ are considered to be known, as well as the Laguerre cells corresponding to these points. In practice it is often the case that some of the cells near the boundary of W_n are only partially observed. Then, one could follow the suggestion in Remark 5.2 and compute the estimator based on a smaller window $W_n' \subset W_n$. Here, one should choose $W_n' \subset W_n$ as large as possible such that $C((x, h), \eta) \subset W_n$ for all $(x, h) \in \eta^* \cap (W_n' \times (0, \infty))$. The estimator of interest can then be computed, replacing W_n by W_n' in the definition of the estimator.

In a practical setting, a Laguerre tessellation may have been fitted to some image data using a reconstruction algorithm. It could be the case that a generator point in the observation window cannot be reconstructed because its cell is located outside of the observation window. Because there are various ways of reconstructing Laguerre tessellations and their extreme points, detailed simulations which consider the constraints induced by those reconstruction methods are out of the scope of this chapter. We believe it will be useful to perform those kinds of simulations in future research, to better understand how well the estimators perform when all practical considerations are taken into account. For now, it should also be possible to mitigate these edge effects caused by the reconstruction approach by computing the estimators based on a smaller observation window $W_n' \subset W_n$. We address some of the challenges associated with the weighted generators in the next chapter.

Besides challenges arising from edge effects we also would like to discuss a few possible directions which may be pursued in future research. There are various important properties of the estimators $\hat{F}_n^0$ and $\hat{F}_n$ which are still unknown. For instance, at present we do not know which of the estimators $\hat{F}_n^0$ or $\hat{F}_n$ should in general be preferred in practice. Hence, it may be of interest to study the rates of convergence of these estimators. Another important challenge for future research is the derivation of the asymptotic distributions of $\hat{F}_n^0$ and $\hat{F}_n$. Knowledge of these asymptotic distributions is essential for deriving (asymptotic) confidence intervals for $\hat{F}_n^0$ and $\hat{F}_n$. This information may then also be used to determine guidelines for required observation window sizes.

5.10 Additional proofs

Proof of Lemma 5.2. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting the sequence $(F_n)_{n\geq 1}$. For $z \in \mathbb{R}$, there exists by assumption a set $\Omega_z \in \mathcal{A}$ be such that $\lim_{n\to\infty} F_n(z;\omega) = F(z)$ for all $\omega \in \Omega_z$ and $\mathbb{P}(\Omega_z) = 1$. Let $D := \{z \in \mathbb{R} : z \in \mathbb{Q} \text{ or } z \text{ is a discontinuity point of } F\}$. Because monotone functions have at most countably many discontinuity points and because the rationals are countable it follows that D is countable. Letting $\Omega' := \cap_{z\in D}\Omega_z$ we obtain $\mathbb{P}(\Omega') = 1$. Let $z \in \mathbb{R}$ and $\omega \in \Omega'$, we show that $\lim_{n\to\infty} F_n(z;\omega) = F(z)$. If $z \in \mathbb{Q}$ or if z is a discontinuity point of F, then the result is immediate. Suppose that $z \in \mathbb{R}\setminus\mathbb{Q}$ is a continuity point of F. For $m \in \mathbb{N}$ choose $\delta_m > 0$ such that $|F(z) - F(x)| < 1/m$ whenever $|z - x| < \delta_m$. Choose $r_m, s_m \in \mathbb{Q}$ such that $r_m \leq z \leq s_m$ and $|r_m - z| < \delta_m$ and $|s_m - z| < \delta_m$. By the monotonicity of each F_n, and since $\omega \in \Omega'$ we have for all $m \in \mathbb{N}$:

$$F(r_m) = \lim_{n\to\infty} F_n(r_m;\omega) \leq \lim_{n\to\infty} F_n(z;\omega) \leq \lim_{n\to\infty} F_n(s_m;\omega) = F(s_m).$$

Due to the choice of r_m and s_m we obtain:

$$-\frac{1}{m} < F(r_m) - F(z) \leq \lim_{n\to\infty} F_n(z;\omega) - F(z) \leq F(s_m) - F(z) < \frac{1}{m}.$$

The result now follows since $|\lim_{n\to\infty} F_n(z;\omega) - F(z)| < 1/m$ for all $m \in \mathbb{N}$. □

Proof of Lemma 5.3. Let $R > 0$ and assume $\lim_{n\to\infty} F_n(z) = F(z)$ for all $z \in [0, R)$. Fix $z \in [0, R)$. We introduce the following shorthand notation, for $h \in [0, z]$:

$$\phi_n(h) := \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F_n(t), \text{ and } \phi(h) := \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F(t).$$

The triangle inequality yields the following bound:

$$\begin{aligned}
\left|G_{F_n}(z) - G_F(z)\right| \leq \\
&\leq \left|\int_0^z \exp\left(-\kappa_d \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F_n(t)\right) - \exp\left(-\kappa_d \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F(t)\right)\mathrm{d}F_n(h)\right| + \\
&\quad + \left|\int_0^z \exp\left(-\kappa_d \int_0^h (h-t)^{\frac{d}{2}}\,\mathrm{d}F(t)\right)\mathrm{d}(F_n - F)(h)\right| \\
&\leq \sup_{h\in[0,z]} \left|\exp\left(-\kappa_d\phi_n(h)\right) - \exp\left(-\kappa_d\phi(h)\right)\right| F_n(z) + \\
&\quad + \left|\int_0^z \exp\left(-\kappa_d\phi(h)\right)\mathrm{d}(F_n - F)(h)\right|.
\end{aligned} \tag{5.45}$$

Let us consider the first term of (5.45). Fix $h \in [0, z]$. Since F_n converges pointwise to F on $[0, z]$ and $t \mapsto (h-t)^{\frac{d}{2}}\mathbb{1}\{h \geq t\}$ is continuous and bounded on $[0, z]$ it follows that $\lim_{n\to\infty}\phi_n(h) = \phi(h)$. Hence, the sequence of monotone functions $\exp(-\kappa_d\phi_n(\cdot))$ converges pointwise to the monotone function $\exp(-\kappa_d\phi(\cdot))$ on $[0, z]$. Because ϕ is (absolutely) continuous, the limit function $\exp(-\kappa_d\phi(\cdot))$ is continuous. The convergence is therefore uniform on $[0, z]$, and we obtain:

$$\lim_{n\to\infty} \sup_{h\in[0,z]} \left|\exp\left(-\kappa_d\phi_n(h)\right) - \exp\left(-\kappa_d\phi(h)\right)\right| F_n(z) = 0 \cdot F(z) = 0.$$

Let us now consider the second term of (5.45). Because $\exp(-\kappa_d \phi(\cdot))$ is continuous and bounded, it immediately follows from the pointwise convergence of F_n to F on $[0, z]$ that this second term vanishes as $n \to \infty$. This proves that $\lim_{n\to\infty} G_{F_n}(z) = G_F(z)$. The proof remains valid when $R = \infty$. □

Proof of Theorem 5.11. Let $z \geq 0$. For $i \in \{1, 2\}$ write $F_i^V := V(\cdot; F_i, m_{F_i})$. From equation (5.19) it can be seen that the (Lebesgue-Stieltjes) measures associated with F_i^V and F_i are mutually absolutely continuous. The corresponding Radon-Nikodym derivative is given by:

$$\frac{\mathrm{d}F_i^V}{\mathrm{d}F_i}(z) = \pi \int_z^\infty \exp\left(-\pi \int_0^u F_i(t)\mathrm{d}t\right)\mathrm{d}u =: p_i(z).$$

Hence, we may also write:

$$F_i(z) = \int_0^z \frac{\mathrm{d}F_i}{\mathrm{d}F_i^V}(h)\mathrm{d}F_i^V(h) = \int_0^z \frac{1}{p_i(h)}\mathrm{d}F_i^V(h).$$

Since $m_{F_i} < \infty$ (this is shown in the proof of Lemma 5.4) we have $p_i(0) < \infty$ and from its definition it is clear that p_i is a decreasing function. Because $x \mapsto 1/x$ is Lipschitz on (c, ∞) for $c > 0$ with Lipschitz constant $1/c^2$ we have for $h \in [0, z]$:

$$\left|\frac{1}{p_1(h)} - \frac{1}{p_2(h)}\right| \leq \max\left\{\frac{1}{p_1(h)^2}, \frac{1}{p_2(h)^2}\right\} |p_1(h) - p_2(h)| \leq C(h)\,|p_1(h) - p_2(h)|\,.$$

Here we have defined $C(h) := \max\{1/p_1(h)^2, 1/p_2(h)^2\}$, which is increasing. As a consequence, we obtain the following upper bound for $|F_1(z) - F_2(z)|$:

$$\begin{aligned}|F_1(z) - F_2(z)| &= \left|\int_0^z \frac{1}{p_1(h)} - \frac{1}{p_2(h)}\mathrm{d}F_1^V(h) + \int_0^z \frac{1}{p_2(h)}\mathrm{d}(F_1^V - F_2^V)(h)\right| \\ &\leq C(z)\int_0^z |p_1(h) - p_2(h)|\,\mathrm{d}F_1^V(h) + \left|\int_0^z \frac{1}{p_2(h)}\mathrm{d}(F_1^V - F_2^V)(h)\right|.\end{aligned} \tag{5.46}$$

We consider the two terms in (5.46) separately. The first term of (5.46) is bounded by:

$$\begin{aligned}&\pi C(z)\int_0^z \left|\int_0^\infty \exp\left(-\pi\int_0^u F_1(t)\mathrm{d}t\right) - \exp\left(-\pi\int_0^u F_2(t)\mathrm{d}t\right)\mathrm{d}u\right|\mathrm{d}F_1^V(h) + \\ &+ \pi C(z)\int_0^z \left|\int_0^h \exp\left(-\pi\int_0^u F_1(t)\mathrm{d}t\right) - \exp\left(-\pi\int_0^u F_2(t)\mathrm{d}t\right)\mathrm{d}u\right|\mathrm{d}F_1^V(h).\end{aligned} \tag{5.47}$$

The first term of (5.47) is equal to $\pi C(z) F_1^V(z)\left|m_{F_1} - m_{F_2}\right|$ and the second term of (5.47) is bounded by:

$$\begin{aligned}&\pi C(z)\int_0^z\int_0^h \left|\exp\left(-\pi\int_0^u F_1(t)\mathrm{d}t\right) - \exp\left(-\pi\int_0^u F_2(t)\mathrm{d}t\right)\right|\mathrm{d}u\mathrm{d}F_1^V(h) \\ &\leq \pi^2 C(z)\int_0^z\int_0^h\int_0^u |F_1(t) - F_2(t)|\,\mathrm{d}t\mathrm{d}u\mathrm{d}F_1^V(h) \\ &\leq \pi^2 C(z) F_1^V(z) z\int_0^z |F_1(t) - F_2(t)|\,\mathrm{d}t.\end{aligned} \tag{5.48}$$

In (5.48) we used the fact $|e^{-x} - e^{-y}| \leq |x - y|$ for $x, y \geq 0$. Via the integration by parts formula, the second term of (5.46) is bounded by:

$$\begin{aligned}
&\left| \left(F_1^V(z) - F_2^V(z)\right) \frac{1}{p_2(z)} - \int_0^z F_1^V(h) - F_2^V(h) \mathrm{d}\left(\frac{1}{p_2(h)}\right)(h) \right| \\
&\leq \frac{1}{p_2(z)} \left|F_1^V(z) - F_2^V(z)\right| + \sup_{h\in[0,z]} \left|F_1^V(h) - F_2^V(h)\right| \left| \int_0^z \mathrm{d}\left(\frac{1}{p_2(h)}\right)(h) \right| \\
&\leq \frac{2}{p_2(z)} \sup_{h\in[0,z]} \left|F_1^V(h) - F_2^V(h)\right|.
\end{aligned}$$

Collecting all results, we obtain:

$$\begin{aligned}
|F_1(z) - F_2(z)| \leq \pi C(z) F_1^V(z) \left|m_{F_1} - m_{F_2}\right| + \pi^2 C(z) F_1^V(z) z \int_0^z |F_1(t) - F_2(t)| \,\mathrm{d}t + \\
+ \frac{2}{p_2(z)} \sup_{h\in[0,z]} \left|F_1^V(h) - F_2^V(h)\right|.
\end{aligned}$$

Applying Theorem 5.5 and (5.10) yields:

$$|F_1(z) - F_2(z)| \leq K(z) \left(\pi C(z) F_1^V(z) \left|m_{F_1} - m_{F_2}\right| + \frac{2}{p_2(z)} \sup_{h\in[0,z]} \left|F_1^V(h) - F_2^V(h)\right| \right). \tag{5.49}$$

Here, $K(z)$ is given by:

$$K(z) := \left(1 + \pi^2 C(z) F_1^V(z) z^2 \exp\left(\pi^2 C(z) F_1^V(z) z^2\right)\right)$$

The statement of the theorem immediately follows from (5.49). □

Proof of Lemma 5.4. We first note that we may assume without loss of generality that $(F_n)_{n\geq 1}$ is a sequence of functions not containing the zero function. Indeed, we could take an arbitrary subsequence $(n_l)_{l\geq 1} \subset (n)_{n\geq 1}$, and then use the pointwise convergence of F_n to F to choose a further subsequence $(n_k)_{k\geq 1} \subset (n_l)_{l\geq 1}$ such that $(F_{n_k})_{k\geq 1}$ is a sequence which does not contain the zero function. If we then show $\lim_{k\to\infty} m_{F_{n_k}} = m_F$ then the whole sequence also converges: $\lim_{n\to\infty} m_{F_n} = m_F$.

We introduce the following notation, for $u \geq 0$ let:

$$p_n(u) := \exp\left(-\pi \int_0^u F_n(t)\mathrm{d}t\right), \quad p(u) := \exp\left(-\pi \int_0^u F(t)\mathrm{d}t\right).$$

Via the inequality $|e^{-x} - e^{-y}| \leq |x - y|$ for $x, y \geq 0$ and (5.22) we obtain the following upper bound for $|p_n(u) - p(u)|$:

$$|p_n(u) - p(u)| \leq \pi \left| \int_0^u F(t) - F_n(t)\mathrm{d}t \right| = \pi \left| \int_0^u (u - t)\mathrm{d}(F - F_n)(t) \right|. \tag{5.50}$$

Due to the pointwise convergence of F_n to F we obtain that p_n converges pointwise to p as $n \to \infty$. The triangle inequality yields:

$$\left|m_{F_n} - m_F\right| \leq \int_0^z |p_n(u) - p(u)| \, \mathrm{d}u + \int_z^\infty |p_n(u) - p(u)| \, \mathrm{d}u. \tag{5.51}$$

The first term of (5.51) vanishes as $n \to \infty$. Indeed, p_n converges pointwise to p as $n \to \infty$, and since $|p_n(u) - p(u)| \leq 1$ the dominated convergence theorem may be applied. The dominated convergence theorem can also be used to show that the second term of (5.51) vanishes as $n \to \infty$. We now show which dominating function g may be used. Choose $z \geq 0$ large enough such that $F(z) > 0$ and set $c := F(z)$. We show that $m_F < \infty$:

$$\begin{aligned} m_F &= \int_0^z p(u)\mathrm{d}u + \int_z^\infty p(u)\mathrm{d}u \\ &\leq z + \exp\left(-\pi \int_0^z F(t)\mathrm{d}t\right) \int_z^\infty \exp\left(-\pi \int_z^u F(t)\mathrm{d}t\right) \mathrm{d}u \\ &\leq z + \int_z^\infty \exp\left(-\pi c \int_z^u \mathrm{d}t\right) \mathrm{d}u = z + \frac{1}{\pi c}. \end{aligned} \tag{5.52}$$

Choose $N \in \mathbb{N}$ large enough such that $F_n(z) \geq c/2$ for all $n \geq N$. This can be done since F_n converges pointwise to F. Applying the same bound as in (5.52) yields $|p_n(u)| \leq \exp\left(-\pi c(u - z)/2\right)$ for all $u \geq z$ and all $n \geq N$. Hence, we may define the dominating function $g : [z, \infty) \to [0, \infty)$ as:

$$g(u) := p(u) + \max\left\{ \max_{k \in \{1,\ldots,N\}} p_k(u), \exp\left(-\pi \frac{c}{2}(u - z)\right) \right\}.$$

Note that $m_F < \infty$ and $m_{F_k} < \infty$ for all $k \in \{1, \ldots, N\}$ by (5.52), applied to F and F_k respectively. As a consequence, g is integrable on $[z, \infty)$. Because $|p_n(u) - p(u)| \leq g(u)$ for all $u \geq z$ the proof is finished. □

Proof of Lemma 5.6. Let $z \geq 0$, we readily obtain the following bound:

$$\begin{aligned} &\left|V(z; F_n, m_n) - V(z; F, m)\right| \leq \\ &\quad \left|\exp\left(-\pi \int_0^z F(t)\mathrm{d}t\right) - \exp\left(-\pi \int_0^z F_n(t)\mathrm{d}t\right)\right| + \\ &\quad + \pi \left|F(z) - F_n(z)\right| \left|m - \int_0^z \exp\left(-\pi \int_0^u F(t)\mathrm{d}t\right) \mathrm{d}u\right| + \\ &\quad + \pi F_n(z) \left(|m_n - m| + \left|\int_0^z \exp\left(-\pi \int_0^u F_n(t)\mathrm{d}t\right) - \exp\left(-\pi \int_0^u F(t)\mathrm{d}t\right) \mathrm{d}u\right|\right) \end{aligned} \tag{5.53}$$

Each of the three terms of (5.53) vanishes as $n \to \infty$, and each of the terms appearing here also appear in the proof of Lemma 5.4. The fact that the first term vanishes follows from (5.50). The second term vanishes due to the pointwise convergence of F_n to F. The third term vanishes since $\lim_{n\to\infty} F_n(z) = F(z)$, $\lim_{n\to\infty} m_n = m$, and by using the same argument as for the first term in (5.51). □

5.11 Computational formula

First of all, note that:

$$\hat{m}_n = \int_0^{h_1} \exp\left(-\pi \int_0^u \hat{F}_n^0(t)\mathrm{d}t\right)\mathrm{d}u + \int_{h_1}^{\infty} \exp\left(-\pi \int_0^u \hat{F}_n^0(t)\mathrm{d}t\right)\mathrm{d}u \tag{5.54}$$

The first integral of (5.54) is equal to h_1, since $\hat{F}_n^0$ is zero on $[0, h_1)$. Let $h_{k+1} > h_k$, via a direct computation we obtain:

$$\begin{aligned}
&\int_{h_1}^{h_{k+1}} \exp\left(-\pi \int_0^u \hat{F}_n^0(t)\mathrm{d}t\right)\mathrm{d}u = \\
&= \sum_{i=2}^{k+1} \int_{h_{i-1}}^{h_i} \exp\left(-\pi \int_0^u \hat{F}_n^0(t)\mathrm{d}t\right)\mathrm{d}u \\
&= \sum_{i=2}^{k+1} \exp\left(-\pi \int_0^{h_{i-1}} \hat{F}_n^0(t)\mathrm{d}t\right) \int_{h_{i-1}}^{h_i} \exp\left(-\pi \int_{h_{i-1}}^u \hat{F}_n^0(t)\mathrm{d}t\right)\mathrm{d}u \\
&= \sum_{i=2}^{k+1} \exp\left(-\pi \int_0^{h_{i-1}} \hat{F}_n^0(t)\mathrm{d}t\right) \int_{h_{i-1}}^{h_i} \exp\left(-\pi(u - h_{i-1})\hat{F}_n^0(h_{i-1})\right)\mathrm{d}u \\
&= \sum_{i=2}^{k+1} \exp\left(-\pi \sum_{j=1}^{i-1} \hat{F}_n^0(h_j)(h_j - h_{j-1})\right) \frac{1}{\pi \hat{F}_n^0(h_{i-1})} \left(1 - \exp\left(-\pi \hat{F}_n^0(h_{i-1})(h_i - h_{i-1})\right)\right).
\end{aligned}$$

Letting $h_{k+1} \to \infty$ we obtain:

$$\begin{aligned}
\hat{m}_n = h_1 &+ \exp\left(-\pi \sum_{j=1}^{k} \hat{F}_n^0(h_j)(h_j - h_{j-1})\right) \frac{1}{\pi \hat{F}_n^0(h_k)} + \\
&+ \sum_{i=2}^{k} \exp\left(-\pi \sum_{j=1}^{i-1} \hat{F}_n^0(h_j)(h_j - h_{j-1})\right) \frac{1}{\pi \hat{F}_n^0(h_{i-1})} \left(1 - \exp\left(-\pi \hat{F}_n^0(h_{i-1})(h_i - h_{i-1})\right)\right).
\end{aligned}$$

Chapter 6

Inverting Poisson-Laguerre tessellations

6.1 Introduction

In this chapter we continue studying Poisson-Laguerre tessellations. We address an issue with random Laguerre tessellations that as far as we know has not yet been addressed in the literature. Resolving this issue is an essential step for being able to perform statistical inference on Poisson-Laguerre tessellations in practice. This issue is related to the weighted generator points used for defining a Laguerre tessellation.

Recall, that like the Laguerre tessellation, a Voronoi tessellation is defined via a set of points, referred to as the generators, and each generator corresponds to a cell in the tessellation. For Laguerre tessellations, the generators also carry a weight. As opposed to generators in the Voronoi model, not all generators will necessarily generate a cell in the Laguerre setting. That is, the cell corresponding to some weighted generator may be the empty set. For Voronoi- and Laguerre tessellations, algorithms have been developed for computing the cells in the tessellation for a given set of (weighted) generator points, see for instance [71] and [6]. Additionally, in the case of the Voronoi tessellation it is known how to perform the inverse process. That is, given some description of the cells of a Voronoi tessellation it is possible to retrieve the generator points used to obtain this Voronoi tessellation. We refer to [87] for an algorithm to perform this procedure. In the case of Laguerre tessellations it is not possible to uniquely determine the set of weighted generator points used to obtain a given Laguerre tessellation. More specifically, in [27] it was shown via simulations that multiple configurations of weighted points may result in the same Laguerre tessellation, meaning that Laguerre tessellations are overparameterized. In Figure 6.1 two configurations of weighted generators are shown which yield the same Laguerre tessellation. The circle around each generator point has radius equal to the weight of this generator. To the best of our knowledge there is no known result which presents a characterization of all configurations of weighted points which yield the same Laguerre tessellation. The process of retrieving (weighted) generators from a given Voronoi- or Laguerre tessellation is referred to as inverting the Voronoi- or Laguerre tessellation, which explains the title of this chapter.

In the literature, various instances of random Laguerre tessellations have been studied,

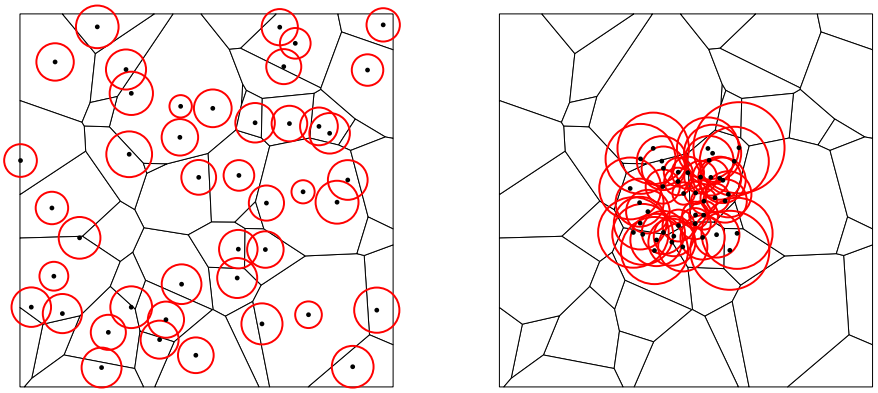


Figure 6.1: Example highlighting the overparameterization of a Laguerre tessellation

see [54], [39] and [95]. These random Laguerre tessellations are obtained by taking the weighted generator points to be a realization of a (marked) point process. The fact that it is not possible to uniquely recover the weighted generator points of a given Laguerre tessellation is somewhat disappointing in the context of statistical inference for random Laguerre tessellations. This being the case because the usual approach to performing statistical inference for random Laguerre tessellations is to consider the point process of weighted generators corresponding to observed cells to be known, and to then use these points to estimate the parameter(s) of the underlying point process model. This approach is for instance taken in [95] as well as in chapter 5.

In this chapter we first study general Laguerre tessellations, and then we shift our focus towards Poisson-Laguerre tessellations. For general Laguerre tessellations, we show that under a set of commonly satisfied regularity conditions it is possible to fully characterize all configurations of weighted generator points which yield the same Laguerre tessellation. A Poisson-Laguerre tessellation in $\mathbb{R}^d$ is obtained by taking the weighted generators to be a realization of a Poisson process on $\mathbb{R}^d \times E$ for some set $E \subset \mathbb{R}$. As in chapter 5, in this chapter we also consider the choice $E = (0, \infty)$, and we assume that the intensity measure of the underlying Poisson process η has intensity measure $\nu_d \times \mathbb{F}$. Here, ν_d is Lebesgue measure on $\mathbb{R}^d$ and $\mathbb{F}$ is a locally finite measure on $(0, \infty)$. While a Poisson-Laguerre tessellation is a tessellation of $\mathbb{R}^d$, in many practical settings it is common to take a so-called observation window $W \subset \mathbb{R}^d$, and to only observe a part of the tessellation through this window. We provide sufficient conditions for consistently inverting the observed Poisson-Laguerre tessellation, as the observation window expands unboundedly to the whole space. This essentially means that while it is in general not possible to uniquely determine the weighted generator points of a Laguerre tessellation, if one observes a Poisson-Laguerre tessellation one can get a closer and closer approximation of the original weighted generator points corresponding to observed cells as one observes the Poisson-Laguerre tessellation through observation windows of increasing sizes.

Taking these results into account, we are interested in whether the estimators for $F(z) =$

$\mathbb{F}((0,z])$, $z \geq 0$, as proposed in Chapter 5 still perform well if these estimators are computed based on our proposed approximation of the generator points instead of the true generator points. That is, we want to know if it is possible to estimate F, when we only rely on the cells of the tessellation which are (partially) observed through the observation window, without prior knowledge of the generators. We address this question via a simulation study.

In a practical context, it is often the case that one does not actually have Laguerre tessellation data. Instead, one may have some image data, for example a microscopic image of a materials microstructure. Then, one may fit a Laguerre tessellation to this image data. We refer to [3] and references therein for an overview of methods which can fit a unique Laguerre tessellation to image data. Typically, these procedures also provide a configuration of weighted generators which generate the fitted Laguerre tessellation. If one has a Laguerre tessellation without a corresponding configuration of weighted generators, then algorithm 1 in [27] may be used to obtain such a configuration. Once this fitting procedure is performed one may apply the methodology proposed in this chapter, to obtain a specific configuration of the weighted generators which is suitable for statistical analysis. This approach will especially be sensible if the image data at hand may approximately be considered a realization of a Poisson-Laguerre tessellation.

This chapter is organized as follows. In section 6.2 we introduce necessary notation and definitions. We present a characterization of the overparameterization of Laguerre tessellations in section 6.3. Inspired by this result we propose a method for inverting a Poisson-Laguerre tessellation in section 6.4. Sufficient conditions for consistent inversion of Poisson-Laguerre tessellations are given in section 6.5. Then, we perform some simulations in section 6.6. Here, we apply the proposed inversion procedure and use the resulting approximation of the weighted generators to compute estimates of F. Finally, we provide some conclusions in section 6.7.

6.2 Preliminaries

In this section we introduce necessary notation and definitions. As in the previous chapter we focus on Poisson-Laguerre tessellations, and therefore we also need the notation and definitions as used in chapter 5. We refer the reader to sections 5.2 and 5.3 for a reminder of those concepts. In this chapter we do need some additional definitions, which are introduced below.

Let $P \subset \mathbb{R}^d$ be a full-dimensional polytope. A subset F of P is called a face of P if either $F = \emptyset$, $F = P$ or if there exists a supporting hyperplane H of P such that $F = P \cap H$. H is a supporting hyperplane of P if $H \cap P \neq \emptyset$ and if P is contained in only one of the closed half spaces bounded by H. Note that each face of a convex polytope is again a convex polytope. A k-dimensional face of P is called a k-face. Let $\mathscr{F}^k(P)$ denote the set of all k-faces of P. Usually, 0-faces are called vertices, 1-faces edges and $(d-1)$-faces facets. A tessellation $T = \{C_i : i \in \mathbb{N}\}$ is called face-to-face if the intersection of any two cells C_i and C_j is either empty or a k-face ($k \leq d-1$) of both cells. This condition may

be written as:

$$C_i \cap C_j \in \bigcup_{k=0}^{d-1} \left(\mathscr{F}^k(C_i) \cap \mathscr{F}^k(C_i) \right) \cup \{\emptyset\} \text{ for all } i \neq j.$$

If moreover every k-face is contained in the boundary of exactly $d - k + 1$ cells ($k = 0, \ldots, d-1$), then the tessellation is called normal. Let $\varphi = \{(x_i, h_i)\}_{i \in \mathbb{N}} \subset \mathbb{R}^d \times \mathbb{R}$. We say that φ satisfies the regularity conditions if:

1. $\text{conv}\{x_i : (x_i, h_i) \in \varphi\} = \mathbb{R}^d$.
2. Only finitely many $(x_i, h_i) \in \varphi$ satisfy $\|x_i - y\|^2 + h_i \leq t$ for any $y \in \mathbb{R}^d$ and any $t \in \mathbb{R}$.

We say that the points of φ are in general position if:

1. A $(k-1)$-dimensional affine subspace contains at most k x_i's ($k = 2, \ldots, d$).
2. At most $d + 1$ points $(x, h) \in \varphi$ satisfy $\|x - y\|^2 + h = t$ for any $y \in \mathbb{R}^d$, $t \in \mathbb{R}$.

If φ satisfies the regularity conditions, then the Laguerre diagram generated by φ is a face-to-face tessellation. If the points of φ are also in general position, then all cells of this Laguerre tessellation have dimension d and this tessellation is normal. The aforementioned results can be found in [54].

Throughout, η is a Poisson process on $\mathbb{R}^d \times (0, \infty)$ with intensity measure $\nu_d \times \mathbb{F}$. Here, $\mathbb{F}$ is a locally finite measure concentrated on $(0, \infty)$. In [39] it was shown that with probability one η satisfies the regularity conditions and its points are in general position. As a consequence, $L(\eta)$ is with probability one a tessellation, which is known as the Poisson-Laguerre tessellation. One may define the distribution function $F(z) = \mathbb{F}((0, z])$ for $z \geq 0$. Consistent estimators for F were introduced in chapter 5.

In this chapter we will also use the following notation. As shown in Lemma 5.1, for $x, y \in \mathbb{R}^d$ and $h > 0$ we have $y \in C((x, h), \eta) - x \iff \eta(A_{x,h,y}) = 0$, with:

$$A_{x,h,y} = \left\{ (x', h') \in \mathbb{R}^d \times (0, \infty) : \|y\|^2 + h - h' > \|x + y - x'\|^2 \right\}. \tag{6.1}$$

Moreover,

$$\mathbb{P}\left(y \in C((x, h), \eta) - x\right) = \exp\left(-\kappa_d \int_0^{\|y\|^2 + h} \left(\|y\|^2 + h - t \right)^{\frac{d}{2}} \mathrm{d}F(t) \right). \tag{6.2}$$

For deterministic $(x, h) \in \mathbb{R}^d \times \mathbb{R}$, the set $C((x, h), \eta)$ is a so-called random closed set. Let $\mathcal{F}(\mathbb{R}^d)$ denote the system of closed subsets of $\mathbb{R}^d$. $\mathfrak{F}(\mathbb{R}^d)$ denotes the σ-algebra on $\mathcal{F}(\mathbb{R}^d)$ which is generated by all families $\mathcal{F}^K = \{F \in \mathcal{F}(\mathbb{R}^d) : F \cap K = \emptyset\}$, $K \in \mathcal{K}^d$. Here, $\mathcal{K}^d$ denotes the space of convex bodies in $\mathbb{R}^d$, and a convex body is a convex and compact set with non-empty interior. Then, a random closed set is a random element of $(\mathcal{F}(\mathbb{R}^d), \mathfrak{F}(\mathbb{R}^d))$. Robbins' theorem states that for any random closed set $X \subset \mathbb{R}^d$ and $p \in \mathbb{N}$ we have:

$$\mathbb{E}\left(\nu_d(X)^p\right) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \mathbb{P}(y_1, \ldots, y_p \in X) \mathrm{d}y_1 \ldots \mathrm{d}y_p. \tag{6.3}$$

Robbin's theorem is essentially Fubini's theorem in the context of random closed sets. We refer to [66] for more details on random closed sets and Robbins' theorem. From the proof of Theorem 5.10 in chapter 5 we recall that for any $x \in \mathbb{R}^d$:

$$\begin{aligned}
&\int_0^\infty \mathbb{E}\left(\nu_d(C((x,h),\eta))\right) \mathrm{d}F(h) = \\
&= \int_0^\infty \int_{\mathbb{R}^d} \mathbb{P}\left(y \in C((x,h),\eta) - x\right) \mathrm{d}y \mathrm{d}F(h) \\
&= \int_0^\infty \int_{\mathbb{R}^d} \exp\left(-\kappa_d \int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) \mathrm{d}y \mathrm{d}F(h) \\
&= 1. \qquad (6.4)
\end{aligned}$$

6.3 The overparameterization of Laguerre tessellations

For a Laguerre tessellation it can be seen from the definition of a Laguerre cell that transforming all weights h_i via $h_i \mapsto h_i + z$ for a fixed $z \in \mathbb{R}$ does not change the tessellation. In [27] it was shown via simulations that there also exist various other choices of generator points and weights which generate the same Laguerre tessellation. However, to the best of our knowledge there is no explicit characterization of the class of all configurations of points and weights which yield the same Laguerre tessellation. In this section, we obtain such a characterization for a class of Laguerre tessellations generated by weighted points satisfying commonly used regularity conditions. The following lemma from [63] highlights a large class of generator points and weights which result in the same Laguerre tessellation.

Lemma 6.1 (Proposition 6 in [63])**.** *Let $\mathcal{I} \subset \mathbb{N}$ and $\varphi = \{(x_i, h_i)\}_{i\in\mathcal{I}} \subset \mathbb{R}^d \times \mathbb{R}$. Let $\lambda > 0, c \in \mathbb{R}^d$ and $z \in \mathbb{R}$. For $i \in \mathcal{I}$ define:*

$$\begin{aligned}
x_i' &:= \lambda x_i + c \\
h_i' &:= \lambda h_i - \lambda(\lambda - 1)\|x_i\|^2 - 2\lambda\langle x_i, c\rangle + z.
\end{aligned}$$

Set $\psi = \{(x_i', h_i')\}_{i\in\mathcal{I}}$. Then, $L(\varphi) = L(\psi)$. In fact, $C((x_i, h_i), \varphi) = C((x_i', h_i'), \psi)$ for all $i \in \mathcal{I}$.

Remark 6.1. *If we exclude the case $\lambda = 1$ in Lemma 6.1, then x_i' and h_i' may be written in the following form:*

$$\begin{aligned}
x_i' &:= \lambda(x_i - c') + c' \\
h_i' &:= \lambda h_i - \lambda(\lambda - 1)\|x_i - c'\|^2 + z',
\end{aligned}$$

for some $c' \in \mathbb{R}^d$ and $z' \in \mathbb{R}$. This form highlights in particular how each weight h_i' depends on the distance of x_i to c'.

Remark 6.2. *Lemma 6.1 was used to construct the example in Figure 6.1.*

In order to prove a converse of Lemma 6.1 we need some additional assumptions. For instance, from the crystallic growth interpretation of a Laguerre tessellation as described in section 5.3, we see that one can always add a weighted point to a configuration of weighted points with a sufficiently large weight such that this weighted point will generate an empty cell. After all, the cell corresponding to this weighted point will essentially "start growing too late". As such, we can at best characterize the class of weighted points which generate the same Laguerre tessellation, when restricting ourselves to the weighted points corresponding to the non-empty cells. In the theorem below we present such a characterization.

Theorem 6.1. *Let $\varphi = \{(x_i, h_i)\}_{i\in\mathbb{N}} \subset \mathbb{R}^d \times \mathbb{R}$, $d \geq 2$. Assume that φ satisfies the regularity conditions and the points of φ are in general position. Let $\psi \subset \mathbb{R}^d \times \mathbb{R}$ be a countable set of distinct points. Assume that $C((x,h),\varphi) \neq \emptyset$ for each $(x,h) \in \varphi$ and $C((x,h),\psi) \neq \emptyset$ for each $(x,h) \in \psi$. Then, $L(\varphi) = L(\psi)$ if and only if $\psi = \{(x_i', h_i')\}_{i\in\mathbb{N}}$ with:*

$$\begin{aligned} x_i' &:= \lambda x_i + c \\ h_i' &:= \lambda h_i + \lambda(\lambda - 1)\|x_i\|^2 + 2\lambda\langle x_i, c\rangle + z, \end{aligned}$$

for some $\lambda > 0$, $c \in \mathbb{R}^d$ and $z \in \mathbb{R}$.

Proof. Let $i \in \mathbb{N}$. If x_i' and h_i' are as in the statement of the theorem then $L(\varphi) = L(\psi)$ by Lemma 6.1. Now it remains to show the converse. Hence, we assume $L(\varphi) = L(\psi)$. Note that the points of ψ can always be relabeled in such a way that we obtain $C((x_i,h_i),\varphi) = C((x_i',h_i'),\psi)$ for all $i \in \mathbb{N}$. Throughout this proof we write $C_i = C((x_i,h_i),\varphi)$ and $C_i' = C((x_i',h_i'),\psi)$ for $i \in \mathbb{N}$. Hence it remains to show that if $C_i = C_i'$ for all $i \in \mathbb{N}$, then x_i' and h_i' are as in the statement of the theorem. Choose $i, j \in \mathbb{N}$ with $i \neq j$ such that $C_i \cap C_j \neq \emptyset$. Since the Laguerre tessellation $L(\varphi)$ is a normal face-to-face tessellation, C_i and C_j share a facet. By (5.3) this facet is contained within the supporting hyperplane:

$$H_{ij} = \{x \in \mathbb{R}^d : 2\langle x, x_i - x_j\rangle = \|x_i\|^2 - \|x_j\|^2 - h_j + h_i\}.$$

We now seek conditions such that ψ generates the same Laguerre cells. If so, we necessarily have $H_{ij} = H_{ij}'$ with:

$$H_{ij}' = \{x \in \mathbb{R}^d : 2\langle x, x_i' - x_j'\rangle = \|x_i'\|^2 - \|x_j'\|^2 - h_j' + h_i'\},$$

this being the case because H_{ij}' still needs to be a supporting hyperplane of C_i' and C_j'. Note that the normal vectors of H_{ij} and H_{ij}' are given by $x_i - x_j$ and $x_i' - x_j'$ respectively. As a result there exists a $\lambda_{ij} \neq 0$ such that: $x_i' - x_j' = \lambda_{ij}(x_i - x_j)$. Now take another $k \in \mathbb{N}$, with $k \neq i$ and $k \neq j$ such that C_i, C_j and C_k share a vertex. Due to normality of the Laguerre tessellation such a k exists. As a result, C_j and C_k share a facet and C_i and C_k share a facet. Arguing as before, there exists $\lambda_{ik} \neq 0$ and $\lambda_{jk} \neq 0$ such that:

$$x_i' - x_j' = \lambda_{ij}(x_i - x_j) \tag{6.5}$$

$$x_i' - x_k' = \lambda_{ik}(x_i - x_k) \tag{6.6}$$

$$x_j' - x_k' = \lambda_{jk}(x_j - x_k). \tag{6.7}$$

Suppose that $\lambda_{ik} \neq \lambda_{ij}$, we show that this leads to a contradiction. Equation (6.5) may be written as: $\lambda_{ij}x_i = \lambda_{ij}x_j + x'_i - x'_j$. Subtracting (6.7) from (6.6) yields: $x'_i - x'_j = \lambda_{ik}(x_i - x_k) - \lambda_{jk}(x_j - x_k)$. Combining these two expressions yields:

$$\lambda_{ij}x_i = \lambda_{ij}x_j + \lambda_{ik}(x_i - x_k) - \lambda_{jk}(x_j - x_k).$$

By solving for x_i we obtain:

$$x_i = \left(1 - \frac{\lambda_{jk} - \lambda_{ik}}{\lambda_{ij} - \lambda_{ik}}\right)x_j + \left(\frac{\lambda_{jk} - \lambda_{ik}}{\lambda_{ij} - \lambda_{ik}}\right)x_k.$$

This means that x_i is a linear combination of x_j and x_k, and therefore these three points are collinear. This is in contradiction with the assumption that the points of φ are in general position. Hence, $\lambda_{ik} = \lambda_{ij}$. By symmetry: $\lambda \equiv \lambda_{ij} = \lambda_{ik} = \lambda_{jk}$. Because the Laguerre tessellation is a face-to-face tessellation we can iteratively consider neighboring cells such that we eventually find:

$$x'_i - x'_j = \lambda(x_i - x_j) \text{ for all } i, j \in \mathbb{N}. \tag{6.8}$$

Fix $i \in \mathbb{N}$, and let $j \in \mathbb{N}$. Choose $c \in \mathbb{R}^d$ such that $x'_i = \lambda x_i + c$. From (6.8) we now obtain that also $x'_j = \lambda x_j + c$. Hence, $x'_i = \lambda x_i + c$ for all $i \in \mathbb{N}$. We will argue that $\lambda > 0$ at the end of this proof. Now, we need to determine weights $h'_1, h'_2, \cdots \in \mathbb{R}$ such that the weighted points $\psi := \{(x'_i, h'_i)\}_{i\in\mathbb{N}}$. generate the same Laguerre cells. Consider once again $j \neq i$ such that $C_i \cap C_j \neq \emptyset$. Following the notation for H_{ij} and H'_{ij} as before, plugging in $x'_i = \lambda x_i + c$ we obtain:

$$H'_{ij} = \{x \in \mathbb{R}^d : 2\langle x, \lambda(x_i - x_j)\rangle = \lambda^2\|x_i\|^2 + 2\lambda\langle x_i, c\rangle - \lambda^2\|x_j\|^2 - 2\lambda\langle x_j, c\rangle - h'_j + h'_i\}.$$

It is now evident that $H_{ij} = H'_{ij}$ if and only if:

$$\begin{aligned}
&\lambda^2\|x_i\|^2 + 2\lambda\langle x_i, c\rangle - \lambda^2\|x_j\|^2 - 2\lambda\langle x_j, c\rangle - h'_j + h'_i = \lambda\left(\|x_i\|^2 - \|x_j\|^2 - h_j + h_i\right)\\
\iff\ & h'_i - h'_j = \lambda\left(h_i - (\lambda-1)\|x_i\|^2 - 2\langle x_i, c\rangle\right) - \lambda\left(h_j - (\lambda-1)\|x_j\|^2 - 2\langle x_j, c\rangle\right)\\
\iff\ & h'_i - h'_j = \lambda(f_i - f_j),
\end{aligned} \tag{6.9}$$

with $f_i := h_i - (\lambda-1)\|x_i\|^2 - 2\langle x_i, c\rangle$. Analogous to (6.8) we may argue that (6.9) holds for all $i, j \in \mathbb{N}$. Fix $i \in \mathbb{N}$ and let $j \in \mathbb{N}$, then choose $z \in \mathbb{R}$ such that $h'_i = \lambda f_i + z$. From (6.9) we now obtain that also $h'_j = \lambda f_j + z$. Hence, for all $i \in \mathbb{N}$ we have:

$$h'_i = \lambda f_i + z = \lambda h_i - \lambda(\lambda-1)\|x_i\|^2 - 2\lambda\langle x_i, c\rangle + z.$$

This is precisely the form of h'_i as in the statement of the theorem. Finally, we conclude that we must have $\lambda > 0$. This is the case because for $i, j \in \mathbb{N}$ with $C_i \cap C_j \neq \emptyset$ the supporting hyperplane H_{ij} also defines a half space containing C_i, recall (5.3). Similarly, the half space induced by the supporting hyperplane H_{ji} contains C_j. In order to have $C_i = C'_i$ for all $i \in \mathbb{N}$ we need not only to preserve these supporting hyperplanes but also the orientation of the corresponding half spaces, which requires $\lambda > 0$. After all, choosing $\lambda < 0$ will flip the orientation of each half space. □

6.4 Inverting Poisson-Laguerre tessellations via weighted least-squares

In view of Theorem 6.1 it is evident that whenever one observes a (random) Laguerre tessellation through a bounded window it is unfortunately not possible to exactly reconstruct the original set of of weighted generators corresponding to the observed cells. However, we can already intuitively understand that in many cases some configurations of the weighted points are more likely than others. Consider for example Figure 6.1. If one is given the information that the original generator points are uniformly distributed within the observation window, then the configuration in the left panel seems far more likely than the configuration in the right panel. This suggests that it may still be possible to get very close to the original configuration, even though exact reconstruction is not possible.

We now turn our attention towards Poisson-Laguerre tessellations. Recall that η is a Poisson process on $\mathbb{R}^d \times (0,\infty)$ with intensity measure $\nu_d \times \mathbb{F}$. Suppose we observe the Poisson-Laguerre tessellation $L(\eta)$ through a bounded observation window W_n, where $(W_n)_{n\geq 1}$ is a convex averaging sequence. In view of algorithm 1 in [27] we may assume that we have a configuration of weighted generators corresponding to the (partially) observed cells of the tessellation. Because we are in some sense interested in the "best" configuration of the weighted generators, and because all configurations are related to the original configuration via the form presented in Theorem 6.1, it does not matter which configuration we start with. As such we consider it to be of no loss of generality that we observe weighted generator points of η^* up to a deterministic transformation of the form presented in Theorem 6.1. As in chapter 5, η^* denotes the extreme points of η, and is given by:

$$\eta^* = \{(x,h) \in \eta : C((x,h),\eta) \neq \emptyset\}.$$

Hence, we consider the following. For all points $(x,h) \in \eta^*$ with $x \in W_n$ we assume that the Laguerre cells corresponding to these points are observed. Additionally, the original points (elements of η^*) themselves are considered to be known up to a transformation of the form in Theorem 6.1. Specifically, suppose that $\lambda_0 > 0$, $c_0 \in \mathbb{R}^d$, and define the function $f_0 : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d \times \mathbb{R}$ via:

$$f_0(x,h) = \left(\frac{x}{\lambda_0} - \frac{c_0}{\lambda_0}, \frac{1}{\lambda_0} h + \frac{1}{\lambda_0}\left(\frac{1}{\lambda_0} - 1\right)\|x\|^2 - \frac{2}{\lambda_0}\left\langle x, \frac{c_0}{\lambda_0}\right\rangle \right).$$

We do not consider the additive constant z in Theorem 6.1. We do this because the distribution of a Poisson-Laguerre tessellation is invariant under shifts of the distribution function $F(z) = \mathbb{F}((0,z])$. As such, we cannot do better than estimating F up to a shift.

For any point $(x,h) \in \eta^*$ with $x \in W_n$ we observe $f_0(x,h)$ instead of (x,h) but λ_0 and c_0 are considered to be unknown. Define: $W_n^0 = \frac{1}{\lambda_0} W_n - \frac{c_0}{\lambda_0}$. We wish to estimate λ_0 and c_0. For this purpose we define the following criterion function $T_n : \mathbb{R} \times \mathbb{R}^d \to [0,\infty)$ via:

$$\begin{aligned} T_n(\lambda, c) :=& \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) \int_{C((x,h),f_0(\eta))} \|\lambda x + c - y\|^2 \mathrm{d}y \\ =& \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) \int_{C((x,h),\eta)} \left\| \lambda\left(\frac{x}{\lambda_0} - \frac{c_0}{\lambda_0}\right) + c - y \right\|^2 \mathrm{d}y \end{aligned} \tag{6.10}$$

Note that while T_n contains a sum over elements of η (or $f_0(\eta)$), only elements from η^* (or $f_0(\eta^*)$) contribute to this sum. Intuitively, we believe that this criterion function is asymptotically minimized when taking λ and c such that $\lambda(x/\lambda_0 - c_0/\lambda_0) + c = x$ which corresponds to $\lambda = \lambda_0$ and $c = c_0$. That is, given the transformation f_0 we are interested in finding the inverse transformation. First, we compute the expected value of $T_n(\lambda, c)$. To this end, we need some additional notation. Let $y \in \mathbb{R}^d$ and define the probability density function $p_F : \mathbb{R}^d \to [0, \infty]$ via:

$$p_F(y) = \int_0^\infty \exp\left(-\kappa_d \int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) \mathrm{d}F(h).$$

The fact that p_F is a probability density function follows from (6.4).

Lemma 6.2. *For all $y \in \mathbb{R}^d \setminus \{0\}$: $p_F(y) \leq 1/(\kappa_d \|y\|^d)$. Additionally, for all $q \in \mathbb{N}$: $\int_{\mathbb{R}^d} \|y\|^q p_F(y)\mathrm{d}y < \infty$.*

The proof of Lemma 6.2 is given in section 6.8. Lemma 6.2 ensures that the expected value of $T_n(\lambda, c)$, as given in the next lemma, is finite.

Lemma 6.3. *Let $\lambda > 0$ and $c \in \mathbb{R}^d$, then:*

$$\mathbb{E}\left(T_n(\lambda, c)\right) = \frac{1}{\nu_d(W_n)} \int_{W_n} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x + c - \frac{\lambda c_0}{\lambda_0} \right\|^2 \mathrm{d}x + \int_{\mathbb{R}^d} \|y\|^2 p_F(y)\mathrm{d}y.$$

And in particular:

$$\mathbb{E}\left(T_n(\lambda_0, c)\right) = \|c - c_0\|^2 + \int_{\mathbb{R}^d} \|y\|^2 p_F(y)\mathrm{d}y.$$

Proof. Substituting $\tilde{y} = y - x$, and then writing y instead of $\tilde{y}$ in (6.10) we obtain:

$$\begin{aligned}
&T_n(\lambda, c) = \\
&= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x) \int_{C((x,h),\eta)-x} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - y \right\|^2 \mathrm{d}y \\
&= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x) \int_{\mathbb{R}^d} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - y \right\|^2 \mathbb{1}\{y \in C((x,h),\eta) - x\}\mathrm{d}y
\end{aligned}$$

Note that $C((x,h),\eta) = C((x,h),\eta + \delta_{x,h})$. Hence, via the Mecke equation (Theorem 5.3) and Fubini we obtain:

$$\begin{aligned}
&\mathbb{E}\left(T_n(\lambda, c)\right) = \\
&= \frac{1}{\nu_d(W_n)} \int_{W_n} \int_{\mathbb{R}^d} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - y \right\|^2 \int_0^\infty \mathbb{P}\left(y \in C((x,h),\eta) - x\right) \mathrm{d}F(h)\mathrm{d}y\mathrm{d}x \\
&= \frac{1}{\nu_d(W_n)} \int_{W_n} \int_{\mathbb{R}^d} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - y \right\|^2 p_F(y)\mathrm{d}y\mathrm{d}x
\end{aligned}$$

$$= \frac{1}{\nu_d(W_n)} \int_{W_n} \left\| \left(\frac{\lambda}{\lambda_0} - 1 \right) x - \frac{\lambda c_0}{\lambda_0} + c \right\|^2 \left(\int_{\mathbb{R}^d} p_F(y) \mathrm{d}y \right) \mathrm{d}x +$$
$$- \frac{2}{\nu_d(W_n)} \left\langle \int_{W_n} \left(\frac{\lambda}{\lambda_0} - 1 \right) x - \frac{\lambda c_0}{\lambda_0} + c \mathrm{d}x, \int_{\mathbb{R}^d} y p_F(y) \mathrm{d}y \right\rangle +$$
$$+ \frac{1}{\nu_d(W_n)} \int_{W_n} \mathrm{d}x \int_{\mathbb{R}^d} \|y\|^2 p_F(y) \mathrm{d}y.$$

Because the probability density function p_F is symmetric around 0 and because $\int_{\mathbb{R}^d} \|y\| p_F(y) \mathrm{d}y < \infty$ by Lemma 6.2, it follows that $\int_{\mathbb{R}^d} y p_F(y) \mathrm{d}y = 0$. As a consequence:

$$\mathbb{E}\left(T_n(\lambda, c)\right) = \frac{1}{\nu_d(W_n)} \int_{W_n} \left\| \left(\frac{\lambda}{\lambda_0} - 1 \right) x + c - \frac{\lambda c_0}{\lambda_0} \right\|^2 \mathrm{d}x + \int_{\mathbb{R}^d} \|y\|^2 p_F(y) \mathrm{d}y.$$

□

Remark 6.3. *By considering any non-empty cell $C((x,h),\eta)$ with $(x,h) \in \eta$ and $x \in W_n$ to be fully observed we do not incorporate edge effects. We discuss this issue in section 6.6.*

6.4.1 Definition and computation of an estimator

From its expression, it is evident that the function $(\lambda, c) \mapsto \mathbb{E}\left(T_n(\lambda, c)\right)$ attains its global minimum in (λ_0, c_0). This inspires the definition of the following estimators for λ_0 and c_0:

$$(\hat{\lambda}_n, \hat{c}_n) = \underset{(\lambda,c) \in \mathbb{R} \times \mathbb{R}^d}{\arg\min} \; T_n(\lambda, c).$$

Applying this inversion procedure in practice means the following. For $\lambda \in \mathbb{R}, c \in \mathbb{R}^d$ define $f(\cdot; \lambda, c) : \mathbb{R}^{d+1} \to \mathbb{R}^d \times \mathbb{R}$ via:

$$f((x,h); \lambda, c) = (\lambda x + c, \lambda h - \lambda(\lambda - 1)\|x\|^2 - 2\lambda \langle x, c \rangle).$$

Then, the inversion procedure boils down to computing the following point process:

$$\hat{\eta}_n^* := f(f_0(\eta); \hat{\lambda}_n, \hat{c}_n) \cap (W_n \times (0, \infty)), \tag{6.11}$$

which may be considered an approximation of $\eta^* \cap (W_n \times (0, \infty))$. Let us provide some further motivation for estimating (λ_0, c_0) by minimizing T_n. We intuitively expect that for a point $(x,h) \in \eta^*$, x should be quite close to the center of its cell $C((x,h),\eta)$. Suppose that we consider for the center of a cell its centroid, also known as center of mass. We will now show that minimizing T_n corresponds to minimizing the sum of the volume weighted squared distances of each generator to the center of mass of its cell. For a Borel set $K \subset \mathbb{R}^d$ of positive volume its centroid $c(K)$ is defined as:

$$c(K) = \frac{1}{\nu_d(K)} \int_K x \mathrm{d}x.$$

The following equation highlights a convenient property of $c(K)$, for $a \in \mathbb{R}^d$ we have:

$$\int_K \|x-a\|^2 \mathrm{d}x = \int_K \|x-c(K)\|^2 \mathrm{d}x + \nu_d(K)\|a-c(K)\|^2. \tag{6.12}$$

Note that the LHS of (6.12) is minimized by $a = c(K)$. In words, the centroid minimizes the integrated squared distance to a given set. Let us introduce the following notation, for the volume and the centroid corresponding to the Laguerre cell associated with (x,h):

$$v_{x,h} := \nu_d(C((x,h),\eta)), \quad c_{x,h} := c(C((x,h),\eta))$$

Suppose that $(x',h') = f_0(x,h)$ for $(x,h) \in \eta^*$, because the function f_0 does not affect the resulting Laguerre tessellation we have:

$$v_{x',h'} := \nu_d(C((x',h'),f_0(\eta))) = \nu_d(C((x,h),\eta)) =: v_{x,h}.$$

Similarly we also have $c_{x',h'} = c_{x,h}$. We will frequently use this when we switch from summing over elements from η to summing over elements from $f_0(\eta)$, or vice-versa. Throughout, we will set $c_{x,h} = 0$ whenever $v_{x,h} = 0$. Applying (6.12) to the definition of T_n, we obtain:

$$\begin{aligned}
&T_n(\lambda,c) = \\
&= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) \left(\int_{C((x,h),f_0(\eta))} \|y-c_{x,h}\|^2 \mathrm{d}y + v_{x,h} \left\|\lambda x + c - c_{x,h}\right\|^2 \right) \\
&= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) \left(\int_{C((x,h),\eta)} \|y-c_{x,h}\|^2 \mathrm{d}y + v_{x,h} \left\|\lambda \left(\frac{x}{\lambda_0} - \frac{c_0}{\lambda_0}\right) + c - c_{x,h}\right\|^2 \right).
\end{aligned}$$

As such, we have written T_n as the sum of two terms, and only the second term depends on λ and c. Therefore, we also have:

$$(\hat{\lambda}_n, \hat{c}_n) = \operatorname*{arg\,min}_{(\lambda,c)\in\mathbb{R}\times\mathbb{R}^d} \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \left\|\lambda x + c - c_{x,h}\right\|^2. \tag{6.13}$$

So indeed, $(\hat{\lambda}_n, \hat{c}_n)$ minimizes the sum of volume weighted squared distances of the generators to the centers of mass of their cells. From its expression, we can see that T_n is a convex function. As a consequence, we can obtain $(\hat{\lambda}_n, \hat{c}_n)$ by computing the critical point of T_n (indeed, there is only one critical point). Taking the partial derivatives of T_n w.r.t. λ and c yields:

$$\begin{aligned}
\frac{\partial T_n(\lambda,c)}{\partial \lambda} &= \frac{2}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \left(\lambda\|x\|^2 + \left\langle c - c_{x,h}, x\right\rangle\right) \\
\frac{\partial T_n(\lambda,c)}{\partial c} &= \frac{2}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \left(\lambda x + c - c_{x,h}\right).
\end{aligned}$$

Setting these partial derivatives equal to zero and solving for λ and c yields a unique solution, which is given by:

$$\begin{aligned}\hat{\lambda}_n &= \frac{\frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \langle c_{x,h}, x\rangle}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} - \left\langle \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} c_{x,h}}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}}, \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} x}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} \right\rangle}{\frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \|x\|^2}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} - \left\| \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} x}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} \right\|^2} \\ \hat{c}_n &= \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} c_{x,h}}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} - \hat{\lambda}_n \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} x}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}}.\end{aligned} \tag{6.14}$$

Or, if we write $f_0(\eta) \cap (W_n^0 \times \mathbb{R}) = \{(x_1, h_1), \ldots, (x_m, h_m)\}$ and $c_i = c_{x_i,h_i}$, $v_i = v_{x_i,h_i}$, then $\hat{\lambda}_n$ may be written more compactly as:

$$\begin{aligned}\hat{\lambda}_n &= \frac{\left(\frac{\sum_{i=1}^m v_i \langle c_i, x_i\rangle}{\sum_{i=1}^m v_i}\right) - \left\langle \frac{\sum_{i=1}^m v_i c_i}{\sum_{i=1}^m v_i}, \frac{\sum_{i=1}^m v_i x_i}{\sum_{i=1}^m v_i} \right\rangle}{\left(\frac{\sum_{i=1}^m v_i \|x_i\|^2}{\sum_{i=1}^m v_i}\right) - \left\| \frac{\sum_{i=1}^m v_i x_i}{\sum_{i=1}^m v_i} \right\|^2} \\ &= \frac{\left(\sum_{i=1}^m v_i\right)\left(\sum_{i=1}^m v_i \langle c_i, x_i\rangle\right) - \left\langle \sum_{i=1}^m v_i c_i, \sum_{i=1}^m v_i x_i \right\rangle}{\left(\sum_{i=1}^m v_i\right)\left(\sum_{i=1}^m v_i \|x_i\|^2\right) - \left\|\sum_{i=1}^m v_i x_i\right\|^2}.\end{aligned}$$

Similarly, $\hat{c}_n$ may be written as:

$$\hat{c}_n = \frac{\sum_{i=1}^m v_i c_i}{\sum_{i=1}^m v_i} - \hat{\lambda}_n \frac{\sum_{i=1}^m v_i x_i}{\sum_{i=1}^m v_i}.$$

In order for $\hat{\lambda}_n$ to be well-defined we need to verify that the denominator in its definition is not equal to zero. Whenever $x_1, \ldots, x_m$ are distinct points, we have by Jensen's inequality:

$$\left\| \frac{\sum_{i=1}^m v_i x_i}{\sum_{i=1}^m v_i} \right\|^2 < \frac{\sum_{i=1}^m v_i \|x_i\|^2}{\sum_{i=1}^m v_i}.$$

Because $x_1, \ldots, x_m$ are distinct points with probability one, $\hat{\lambda}_n$ is almost surely well-defined. As is frequently the case with least-squares estimators, these can often be computed via an explicit formula such as (6.14) or by solving a system of linear equations. In this case, note that we may write:

$$\frac{\sum_{i=1}^m v_i \|\lambda x_i + c - c_i\|^2}{\sum_{i=1}^m v_i} = \beta^T X \beta - 2u^T \beta + K.$$

Here, X is a $(d+1) \times (d+1)$ matrix which is given by the following block matrix:

$$X = \begin{pmatrix} I_d & \left(\frac{\sum_{i=1}^m v_i x_i}{\sum_{i=1}^m v_i}\right) \\ \left(\frac{\sum_{i=1}^m v_i x_i}{\sum_{i=1}^m v_i}\right)^T & \frac{\sum_{i=1}^m v_i \|x_i\|^2}{\sum_{i=1}^m v_i} \end{pmatrix}.$$

Each vector is to be interpreted as a column-vector and β^T denotes the transpose of β. The matrix I_d is the unit matrix in $\mathbb{R}^{d\times d}$. Finally, u and β are vectors in $\mathbb{R}^{d+1}$ and K is a constant, these are given by:

$$u = \begin{pmatrix} \frac{\sum_{i=1}^m v_i c_i}{\sum_{i=1}^m v_i} \\ \frac{\sum_{i=1}^m v_i \langle c_i, x_i\rangle}{\sum_{i=1}^m v_i} \end{pmatrix},\ \beta = \begin{pmatrix} c \\ \lambda \end{pmatrix},\ K = \frac{\sum_{i=1}^m v_i \|c_i\|^2}{\sum_{i=1}^m v_i}.$$

Then, $(\hat{\lambda}_n, \hat{c}_n)$ may be computed by solving the linear system $X\beta = u$ for β. In simulations this was observed to be more numerically stable compared to using (6.14).

6.4.2 Limiting behavior of the criterion function

As a first important step towards understanding how $(\hat{\lambda}_n, \hat{c}_n)$ may behave as $n \to \infty$, we study how the criterion function T_n behaves as $n \to \infty$. Before we can determine the limiting behavior of T_n, we need two lemmas which allow us to control the sizes of the Laguerre cells that are summed over in the definition of T_n.

Lemma 6.4. *Let $r > 0$ and $\varphi = \{(x_i, h_i)\}_{i\in\mathbb{N}} \subset \mathbb{R}^d \times \mathbb{R}$. Let $(x, h) \in \varphi$. Define the following set:*

$$D_{x,h} = \left\{(x', h') \in \mathbb{R}^d \times \mathbb{R} : h' \le r^2 + h,\ \|x - x'\| < r + \sqrt{r^2 + h - h'}\right\}.$$

Then, if $(\varphi \setminus \{(x, h)\}) \cap D_{x,h} = \emptyset$ we have $\bar{B}(x, r) \subset C((x, h), \varphi)$.

Proof. Suppose $(\varphi \setminus \{(x, h)\}) \cap D_{x,h} = \emptyset$. Let $y \in \bar{B}(x, r)$, it remains to show that $\|x - y\|^2 + h \le \|x' - y\|^2 + h'$ for all $(x', h') \in \varphi \setminus \{(x, h)\}$. Let $(x', h') \in \varphi \setminus \{(x, h)\}$. If $h' \le r^2 + h$ then by assumption $\|x - x'\| \ge r + \sqrt{r^2 + h - h'}$. As a consequence of this fact and $\|x - y\| \le r$ we obtain:

$$\|x-y\|^2+h-h' \le r^2+h-h' \le (\|x - x'\| - r)^2 \le (\|y - x'\| + \|x - y\| - r)^2 \le \|x' - y\|^2.$$

So indeed, $\|x - y\|^2 + h \le \|x' - y\|^2 + h'$. It remains to consider the case $h' > r^2 + h$. In that case we also obtain the desired result:

$$\|x - y\|^2 + h \le r^2 + h < h' \le \|x' - y\|^2 + h'.$$

□

The lemma below is essentially an intermediate result obtained in the proof of Proposition 3.1 in [30]. For the sake of completeness, the proof is presented in section 6.8.

Lemma 6.5. *Let $z > 0$, $R > 2\sqrt{z}$ and $\varphi = \{(x_i, h_i)\}_{i\in\mathbb{N}} \subset \mathbb{R}^d \times (0, z]$. Let $(x, h) \in \varphi$. Suppose that $J \in \mathbb{N}$, and $C_1, \ldots, C_J$ are convex cones with non-empty interiors with $\cup_{j=1}^J C_j = \mathbb{R}^d$. Additionally, assume these convex cones have disjoint interiors and satisfy $\langle u, v\rangle \ge \frac{3}{4}\|u\| \cdot \|v\|$ whenever $u, v \in C_j$. Define the following sets for $j \in \{1, \ldots, J\}$:*

$$B_{x,j} = \left(\left(B(x, R) \setminus \bar{B}(x, 2\sqrt{z})\right) \cap (C_j + x)\right) \times (0, z].$$

If $\varphi \cap B_{x,j} \neq \emptyset$ for all $j \in \{1, \ldots, J\}$ then $C((x, h), \varphi) \subset B(x, R)$.

In the theorem below we show that T_n almost surely converges pointwise to a function T. Note that the integral in the definition of T is finite by Lemma 6.2.

Theorem 6.2. *Let $\lambda > 0$ and $c \in \mathbb{R}^d$, then with probability one we have :*

$$\lim_{n\to\infty} T_n(\lambda, c) = T(\lambda, c) := \begin{cases} \|c - c_0\|^2 + \int_{\mathbb{R}^d} \|y\|^2 p_F(y)\mathrm{d}y & \text{if } \lambda = \lambda_0 \\ \infty & \text{otherwise.} \end{cases}$$

The proof of Theorem 6.2 is given in section 6.8. In the remainder of this section we derive a few additional limit results, which we need later.

Lemma 6.6. *For $n \in \mathbb{N}$ define:*

$$D_n = \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x) v_{x,h} \tag{6.15}$$

$$B_n = \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x) v_{x,h} \left(c_{x,h} - x\right). \tag{6.16}$$

Then, with probability one: $\lim_{n\to\infty} D_n = 1$, *and* $\lim_{n\to\infty} B_n = 0$

Proof. The fact that $\lim_{n\to\infty} D_n = 1$ almost surely, was shown in the proof of Lemma 5.5 as we have $D_n = \tilde{F}_n^V(\infty)$. Let $c \in \mathbb{R}^d$, we may write:

$$\begin{aligned} T_n(\lambda_0, c) &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x) \int_{C((x,h),\eta)} \|x - c_0 + c - y\|^2 \mathrm{d}y \\ &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x) \Bigg(v_{x,h}\|c - c_0\|^2 + 2\left\langle c - c_0, \int_{C((x,h),\eta)} x - y \mathrm{d}y \right\rangle + \\ &\quad + \int_{C((x,h),\eta)} \|x - y\|^2 \mathrm{d}y \Bigg) \\ &= D_n\|c - c_0\|^2 - 2\langle c - c_0, B_n\rangle + T_n(\lambda_0, c_0). \end{aligned}$$

Indeed, note that for any $(x, h) \in \eta^*$:

$$\int_{C((x,h),\eta)} y - x \mathrm{d}y = \int_{C((x,h),\eta)-x} y \mathrm{d}y = v_{x,h} c(C((x,h),\eta) - x) = v_{x,h}(c_{x,h} - x).$$

As a consequence:

$$\langle c - c_0, B_n\rangle = \frac{1}{2}\left(D_n\|c - c_0\|^2 + T_n(\lambda_0, c_0) - T_n(\lambda_0, c)\right). \tag{6.17}$$

By Theorem 6.2, and since $\lim_{n\to\infty} D_n = 1$ almost surely, we obtain via the continuous mapping theorem the following almost sure limit:

$$\begin{aligned} \lim_{n\to\infty} \langle c - c_0, B_n\rangle &= \frac{1}{2}\left(\|c - c_0\|^2 + \int_{\mathbb{R}^d} \|y\|^2 p_F(y)\mathrm{d}y - \|c - c_0\|^2 - \int_{\mathbb{R}^d} \|y\|^2 p_F(y)\mathrm{d}y\right) \\ &= 0. \end{aligned}$$

Keeping in mind that this holds for any $c \in \mathbb{R}^d$, we may for instance take $c = c_0 + e_j$ where e_j with $j \in \{1, \dots, d\}$ is the j-th standard unit basis vector of $\mathbb{R}^d$. Taking this c we observe that the j-th component of B_n converges to 0 almost surely. Because this holds for any $j \in \{1, \dots, d\}$, $\lim_{n\to\infty} B_n = 0$ almost surely. □

For any $c \in \mathbb{R}^d$, let $\bar{\lambda}_n(c) = \arg\min_\lambda T_n(\lambda, c)$. Note that $\lambda \mapsto T_n(\lambda, c)$ is convex. Recalling that the partial derivatives of T_n were derived in section 6.4.1, a straightforward computation yields:

$$\bar{\lambda}_n(c) = \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \left\langle c_{x,h} - c, x \right\rangle}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \|x\|^2}. \tag{6.18}$$

Computing the partial second derivative of T_n w.r.t. λ yields:

$$\frac{\partial^2 T_n(\lambda, c)}{\partial \lambda^2} = \frac{2}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} \|x\|^2 =: 2M_n. \tag{6.19}$$

From this, it immediately follows that for any $c \in \mathbb{R}^d$, $T_n(\cdot, c)$ is strongly convex with parameter $2M_n$. Recall that a function $f : \mathbb{R}^d \to \mathbb{R}$ is strongly convex with parameter $m > 0$ if for all $x, y \in \mathbb{R}^d$:

$$f(y) \geq f(x) + \nabla \langle f(x), y - x \rangle + \frac{m}{2} \|y - x\|^2.$$

Observe that if x^* is the global minimizer of f, then:

$$\|y - x^*\|^2 \leq \frac{2}{m} \left(f(y) - f(x^*) \right). \tag{6.20}$$

A sufficient condition for $f : \mathbb{R} \to \mathbb{R}$ to be strongly convex with parameter $m > 0$ is that f is twice differentiable and $f''(x) \geq m$ for all $x \in \mathbb{R}$. We now show that the strong convexity parameter of $T_n(\cdot, c)$ diverges to infinity as $n \to \infty$:

Lemma 6.7. *With probability one,* $\lim_{n\to\infty} M_n = \infty$. *Here,* M_n *is as in (6.19).*

Proof. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting the Poisson process η. As in the proof of Theorem 6.2 it is sufficient to show that for all $M > 0$ there exists a $\Omega_M \in \mathcal{A}$ with $\mathbb{P}(\Omega_M) = 1$ such that for all $\omega \in \Omega_M$: $\liminf_{n\to\infty} M_n(\omega) > M$. Let $M > 0$. Choose $\Omega_M \in \mathcal{A}$ with $\mathbb{P}(\Omega_M) = 1$ such that for all $\omega \in \Omega_M$: $L(\eta(\omega))$ is a tessellation and $\lim_{n\to\infty} D_n(\omega) = 1$. Such a Ω_M exists by Lemma 6.6 and the fact that Poisson-Laguerre tessellations are well-defined. Let $\omega \in \Omega_M$, then we may write $M_n(\omega)$ as:

$$\begin{aligned} M_n(\omega) &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta(\cdot;\omega)} \mathbb{1}_{W_n}(x) v_{x,h}(\omega) \left\| \frac{x}{\lambda_0} - \frac{c_0}{\lambda_0} \right\|^2 \\ &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta(\cdot;\omega)} \mathbb{1}_{W_n}(x) v_{x,h}(\omega) \left\| \frac{x}{\lambda_0} - \frac{c_0}{\lambda_0} \right\|^2 \mathbb{1}\left\{ \left\| \frac{x}{\lambda_0} - \frac{c_0}{\lambda_0} \right\| \leq \sqrt{M} \right\} + \end{aligned} \tag{6.21}$$

$$+\frac{1}{\nu_d(W_n)}\sum_{(x,h)\in\eta(\,\cdot\,;\omega)}\mathbb{1}_{W_n}(x)v_{x,h}(\omega)\left\|\frac{x}{\lambda_0}-\frac{c_0}{\lambda_0}\right\|^2\mathbb{1}\left\{\left\|\frac{x}{\lambda_0}-\frac{c_0}{\lambda_0}\right\|>\sqrt{M}\right\} \tag{6.22}$$

We consider (6.21) and (6.22) separately. Note that the following expression is an upper bound for (6.21):

$$\frac{M}{\nu_d(W_n)}\left(\sum_{(x,h)\in\eta(\,\cdot\,;\omega)}\mathbb{1}_{W_n\cap\bar{B}(c_0,\lambda_0 M)}(x)v_{x,h}(\omega)\right). \tag{6.23}$$

For sufficiently large n, $W_n\cap\bar{B}(c_0,\lambda_0 M)=\bar{B}(c_0,\lambda_0 M)$. Hence, for such n, the term in brackets in (6.23) does not depend on n. Because the tessellation $L(\eta(\omega))$ is locally finite and has cells of finite volume, this term is finite. Because $\nu_d(W_n)\to\infty$ as $n\to\infty$ we see that (6.21) vanishes. For (6.22) we obtain the following lower bound:

$$\begin{aligned}&\frac{1}{\nu_d(W_n)}\sum_{(x,h)\in\eta(\,\cdot\,;\omega)}\mathbb{1}_{W_n}(x)v_{x,h}(\omega)\left\|\frac{x}{\lambda_0}-\frac{c_0}{\lambda_0}\right\|^2\mathbb{1}\left\{\left\|\frac{x}{\lambda_0}-\frac{c_0}{\lambda_0}\right\|>\sqrt{M}\right\}\\ &\quad>D_n(\omega)M-\frac{1}{\nu_d(W_n)}\sum_{(x,h)\in\eta(\,\cdot\,;\omega)}\mathbb{1}_{W_n}(x)v_{x,h}(\omega)M\mathbb{1}\left\{\left\|\frac{x}{\lambda_0}-\frac{c_0}{\lambda_0}\right\|\le\sqrt{M}\right\}.\end{aligned} \tag{6.24}$$

By the choice of Ω_M, the first term of (6.24) converges to M as $n\to\infty$, while the second term of (6.24) vanishes, as $n\to\infty$. The fact that the second term of (6.24) vanishes can be shown via the same argument as used for (6.21). Hence: $\liminf_{n\to\infty}M_n(\omega)>M$. □

As a consequence we find that a strongly consistent estimator for λ_0 can be obtained even if we do not optimize for c.

Corollary 6.1. *Let $c\in\mathbb{R}^d$, then with probability one:* $\lim_{n\to\infty}\bar{\lambda}_n(c)=\lambda_0$.

Proof. Applying (6.20) to the function $T_n(\,\cdot\,,c)$, and taking into account that $T_n(\,\cdot\,,c)$ is non-negative, we obtain:

$$\left\|\lambda_0-\bar{\lambda}_n(c)\right\|^2\le\frac{1}{M_n}\left(T_n(\lambda_0,c)-T_n(\bar{\lambda}_n(c),c)\right)\le\frac{T_n(\lambda_0,c)}{M_n}. \tag{6.25}$$

By Lemma (6.7), Theorem 6.2 and the continuous mapping theorem we obtain that the RHS of (6.25) vanishes almost surely, which yields the result. □

6.5 Consistency of the inversion procedure

In the previous section we obtained various limit results, which we will need in this section for proving consistency of the inversion procedure. That is, we will show that under

reasonable conditions the proposed estimator $(\hat{\lambda}_n, \hat{c}_n)$ for (λ_0, c_0) is consistent. For $n \in \mathbb{N}$ define the following random vector:

$$A_n = \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\nu_d(C(x,h),\eta)x. \tag{6.26}$$

That is, A_n may be seen as a weighted average of the generators corresponding to non-empty cells, the weights being the volumes of the cells. One might argue that for a proper weighted average we should divide A_n by D_n, with D_n as in (6.15). However, by Lemma 6.6 we know that D_n converges to 1 almost surely as $n \to \infty$. Hence, when studying the behavior of A_n as $n \to \infty$, it is not needed to divide A_n by D_n because this normalization will not lead to a different limit of A_n, if it exists in some sense of stochastic convergence. As a first observation, the expected value of A_n is the centroid of W_n:

$$\mathbb{E}(A_n) = \frac{1}{\nu_d(W_n)} \int_{W_n} \int_0^\infty \mathbb{E}(\nu_d(C(x,h),\eta))\mathrm{d}F(h)x\mathrm{d}x = \frac{1}{\nu_d(W_n)} \int_{W_n} x\mathrm{d}x = c(W_n).$$

This follows from the Mecke equation and (6.4). As it turns out, the behavior of A_n is essential for obtaining a consistency result for $(\hat{\lambda}_n, \hat{c}_n)$, as is highlighted by the theorem below.

Theorem 6.3 (Consistent inversion)**.**

1. *If the sequence $(A_n)_{n\geq 1}$ is uniformly tight, then:*

$$\lim_{n\to\infty} (\hat{\lambda}_n, \hat{c}_n) \overset{\mathbb{P}}{=} (\lambda_0, c_0).$$

2. *If the random variable $\sup_{n\geq 1} \|A_n\|$ is almost surely finite, then:*

$$\lim_{n\to\infty} (\hat{\lambda}_n, \hat{c}_n) \overset{a.s.}{=} (\lambda_0, c_0).$$

Proof. Let A_n, B_n and D_n be as in (6.26), (6.16) and (6.15) respectively. For $n \in \mathbb{N}$ observe that:

$$\frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)v_{x,h}c_{x,h} = B_n + A_n.$$

Let $\bar{\lambda}_n(0)$ and M_n be as in (6.18) and (6.19) respectively. We may write $\hat{\lambda}_n$ (recall (6.14)) as follows:

$$\begin{aligned}
\hat{\lambda}_n &= \frac{D_n \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x)v_{x,h}\langle c_{x,h}, x\rangle - \langle B_n + A_n, A_n\rangle}{D_n\left(\frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x)v_{x,h}\|x\|^2\right) - \|A_n\|^2} \\
&= \frac{D_n \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x)v_{x,h}\langle c_{x,h}, x\rangle - \|A_n\|^2 - \langle B_n, A_n\rangle}{D_n\left(\frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x)v_{x,h}\|x\|^2\right)} \cdot \\
&\quad \cdot \frac{D_n\left(\frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x)v_{x,h}\|x\|^2\right)}{D_n\left(\frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x)v_{x,h}\|x\|^2\right) - \|A_n\|^2}
\end{aligned}$$

$$= \left(\bar{\lambda}_n(0) - \frac{\|A_n\|^2 + \langle B_n, A_n \rangle}{D_n M_n} \right) \frac{1}{1 - \frac{\|A_n\|^2}{D_n M_n}}$$

By Corollary 6.1 we have $\lim_{n\to\infty} \bar{\lambda}_n(0) = \lambda_0$ almost surely. Suppose condition 1 of the theorem is satisfied. By the uniform tightness of A_n, and since $\lim_{n\to\infty} B_n = 0$ almost surely by Lemma 6.6, we obtain that $\|A_n\|^2 + \langle B_n, A_n \rangle$ is uniformly tight. Because $\lim_{n\to\infty} M_n = \infty$ almost surely this yields:

$$\lim_{n\to\infty} \frac{\|A_n\|^2 + \langle B_n, A_n \rangle}{D_n M_n} \overset{\mathbb{P}}{=} 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{1 - \frac{\|A_n\|^2}{D_n M_n}} \overset{\mathbb{P}}{=} 1. \tag{6.27}$$

Here we also used the rule which is often written as $o_P(1)O_P(1) = o_P(1)$ in stochastic o notation. Hence, via the continuous mapping theorem we obtain $\lim_{n\to\infty} \hat{\lambda}_n = \lambda_0$ in probability. If condition 2 of the theorem is satisfied then the limits in (6.27) become almost sure limits, yielding $\lim_{n\to\infty} \hat{\lambda}_n = \lambda_0$ almost surely. We now consider $\hat{c}_n$ (recall (6.14)), which may be written as follows:

$$\begin{aligned}
\hat{c}_n &= \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} c_{x,h}}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} - \hat{\lambda}_n \frac{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h} x}{\sum_{(x,h)\in f_0(\eta)} \mathbb{1}_{W_n^0}(x) v_{x,h}} \\
&= \frac{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h} c_{x,h}}{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h}} - \hat{\lambda}_n \frac{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h} \left(\frac{x}{\lambda_0} - \frac{c_0}{\lambda_0} \right)}{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h}} \\
&= \frac{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h} \left(c_{x,h} - x \right)}{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h}} + \frac{1}{\lambda_0}(\lambda_0 - \hat{\lambda}_n) \frac{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h} x}{\sum_{(x,h)\in \eta} \mathbb{1}_{W_n}(x) v_{x,h}} + \frac{\hat{\lambda}_n}{\lambda_0} c_0. \\
&= \frac{B_n}{D_n} + \frac{1}{\lambda_0}(\lambda_0 - \hat{\lambda}_n) \frac{A_n}{D_n} + \frac{\hat{\lambda}_n}{\lambda_0} c_0.
\end{aligned} \tag{6.28}$$

By Lemma 6.6, $\lim_{n\to\infty} B_n/D_n = 0$ almost surely. If condition 1 of the theorem is satisfied, then $\lim_{n\to\infty} \hat{\lambda}_n = \lambda_0$ in probability. Hence,

$$\lim_{n\to\infty} \frac{(\lambda_0 - \hat{\lambda}_n)}{\lambda_0 D_n} \overset{\mathbb{P}}{=} 0,$$

here was also used that $\lim_{n\to\infty} D_n = 1$ almost surely by Lemma 6.6. As a consequence, the second term of (6.28) vanishes when taking the limit in probability ($o_P(1)O_P(1) = o_P(1)$), since A_n is uniformly tight. Via the continuous mapping theorem we obtain $\lim_{n\to\infty} \hat{c}_n = c_0$ in probability. If condition 2 of the theorem is satisfied, then $\lim_{n\to\infty} \hat{\lambda}_n = \lambda_0$ almost surely. Via similar arguments we obtain $\lim_{n\to\infty} \hat{c}_n = c_0$ almost surely. □

It seems necessary that some conditions on $(c(W_n))_{n\geq 1}$ need to be imposed for any of the two conditions in Theorem 6.3 to hold. As is commonly done we may for instance choose a convex averaging sequence $(W_n)_{n\geq 1}$ which is centered at the origin. That is, we may take W_n such that $c(W_n) = 0$ for all $n \in \mathbb{N}$. Obviously, if A_n then converges almost surely to 0, this would imply condition 2 in Theorem 6.3. While we do not know if A_n converges in general, we will now highlight via an example why the almost sure convergence of A_n to 0 should not necessarily be expected, especially if $d = 2$.

Example 3. The vector A_n may essentially be seen as a weighted average of Poisson points, the weights being the volumes of the Laguerre cells. Let us now consider what happens if we consider the arithmetic mean of Poisson points instead. Suppose ψ is a homogeneous Poisson process on $\mathbb{R}^d$ with intensity 1. For convenience, we take $W_n = [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$. That is, W_n is the centered cube of volume $\nu_d(W_n) = n$. Then, we define:

$$\bar{A}_n = \frac{1}{\nu_d(W_n)} \sum_{x \in \psi} \mathbb{1}_{W_n}(x) x.$$

Of course, for an actual arithmetic mean we should normalize with $\psi(W_n)$ instead of $\nu_d(W_n)$. However, by the spatial ergodic theorem $\lim_{n\to\infty} \psi(W_n)/\nu_d(W_n) = 1$ with probability one. So this normalization will not lead to a different limit of $\bar{A}_n$, if it exists in some sense of stochastic convergence. Let $t \in \mathbb{R}^d$, it can be shown (the proof is given in section 6.8) that the characteristic function of $\bar{A}_n$ satisfies:

$$\lim_{n\to\infty} \varphi_{\bar{A}_n}(t) = \begin{cases} \exp\left(-\frac{\|t\|^2}{24}\right) & \text{if } d = 2 \\ 1 & \text{if } d \geq 3. \end{cases}$$

By Levy's continuity theorem this implies the following limit in distribution:

$$\lim_{n\to\infty} \bar{A}_n = \begin{cases} \mathcal{N}\left(0, \frac{1}{12} I_2\right) & \text{if } d = 2 \\ 0 & \text{if } d \geq 3. \end{cases}$$

Here, $\mathcal{N}(0, \frac{1}{12} I_2)$ denotes the multivariate normal distribution on $\mathbb{R}^2$ with mean 0 and covariance matrix $\frac{1}{12} I_2$, I_2 being the 2×2 identity matrix. Recall that convergence in distribution to a constant implies convergence in probability to the same constant. So if $d \geq 3$ the convergence also holds in probability. Additionally, the L^2-norm reveals that $\bar{A}_n$ also converges to 0 in L^2 if $d \geq 3$:

$$\mathbb{E}\left(\|\bar{A}_n\|^2\right) = \frac{d}{12 n^{1-\frac{2}{d}}}.$$

Again, the proof of this statement is presented in section 6.8. The L^2-norm highlights that the rate of convergence of $\bar{A}_n$ to 0 is faster in higher dimensions.

While the characteristic function of $\bar{A}_n$ as in Example 3 is tractable, this does not appear to be the case for A_n. Still, for $\bar{A}_n$, knowledge of its L^2-norm is sufficient for deriving uniform tightness of $\bar{A}_n$ via Markov's inequality. Therefore, we derive an upper bound for the L^2-norm of A_n. In order to compute this bound we need several results which are stated below. The proofs of these statements are given in section 6.8.

Proposition 6.1. *Let $x \in \mathbb{R}^d$, $h \geq 0$ and $p \in \mathbb{N}$. We have the following inequalities:*

$$\begin{aligned} \mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) &\leq p d \kappa_d^p \int_0^\infty \exp\left(-\kappa_d \int_0^{r^2+h} \left(r^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) r^{pd-1} \mathrm{d}r \\ &\leq \frac{p!}{F(h)^p}. \end{aligned}$$

Here, we set $p!/F(h)^p = \infty$ if $F(h) = 0$.

For proving Proposition 6.1 we use a similar technique as was used in Proposition 2.3 in [73] to derive a bound on the second moment of the volume of a typical Poisson-Voronoi cell. While the second inequality in Proposition 6.1 provides a finite bound for large values of h it is not yet clear whether the p-th moment is finite for all $h \geq 0$. We show this in the lemma below.

Lemma 6.8. *Let $p \in \mathbb{N}$. There exists a constant $0 < \alpha_p < \infty$ such that for all $x \in \mathbb{R}^d$ and $h \geq 0$, we have $\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) < \alpha_p$.*

Theorem 6.4. *Let $p \in \mathbb{N}$. There exists a constant $0 < \beta_p < \infty$ such that for all $x \in \mathbb{R}^d$:*

$$\int_0^\infty \mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) \mathrm{d}F(h) \leq \beta_p.$$

Lemma 6.9. *Let $x_1, x_2 \in \mathbb{R}^d$, then:*

$$\int_0^\infty \int_0^\infty \mathbb{E}\left(\nu_d(C((x_1,h_1),\eta+\delta_{(x_2,h_2)}))\nu_d(C((x_2,h_2),\eta+\delta_{(x_1,h_1)}))\right) \mathrm{d}F(h_1)\mathrm{d}F(h_2) \leq 4.$$

We set $W_n = n^{\frac{1}{d}}W$ for $n \in \mathbb{N}$, where $W \subset \mathbb{R}^d$ is a convex body satisfying $\nu_d(W) = 1$ and $c(W) = 0$. Via a direct computation we may write:

$$\begin{aligned}\mathbb{E}\left(\|A_n\|^2\right) &= \frac{1}{\nu_d(W_n)^2}\mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\nu_d(C(x,h),\eta)^2\|x\|^2\right) + \\ &+ \frac{1}{\nu_d(W_n)^2}\mathbb{E}\left(\sum_{(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}} \mathbb{1}_{W_n}(x_1,x_2)\nu_d(C(x_1,h_1),\eta)\nu_d(C(x_2,h_2),\eta)\langle x_1,x_2\rangle\right).\end{aligned} \tag{6.29}$$

Here, $(\eta)^2_{\neq}$ denotes the set of all distinct pairs of points of η. That is, if $(x_1,h_1),(x_2,h_2) \in (\eta)^2_{\neq}$, then $(x_1,h_1) \neq (x_2,h_2)$. Via the Mecke equation and Theorem 6.4 the first term of (6.29) is bounded from above by:

$$\begin{aligned}&\frac{1}{\nu_d(W_n)^2}\mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\nu_d(C(x,h),\eta)^2\|x\|^2\right) = \\ &\quad= \frac{1}{n^2}\int_{W_n}\int_0^\infty \mathbb{E}\left(\nu_d(C(x,h),\eta)^2\right)\mathrm{d}F(h)\|x\|^2\mathrm{d}x \\ &\quad\leq \frac{\beta_2}{n^2}\int_{W_n}\|x\|^2\mathrm{d}x \\ &\quad\leq \frac{\beta_2}{n^2}\int_{W_n}\mathrm{diam}(W_n)^2\mathrm{d}x \\ &\quad= \frac{\beta_2\mathrm{diam}(W)^2}{n^{1-\frac{2}{d}}}.\end{aligned} \tag{6.30}$$

Here we also used: $\mathrm{diam}(W_n) = n^{1/d}\mathrm{diam}(W)$. Notice that this upper bound is similar to the L^2-norm of $\bar{A}_n$. In the case of $\bar{A}_n$, the cross-terms vanish, due to independence. If this is also the case for A_n, this would mean the second expectation in (6.29) is equal to zero. Define the function $\phi : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$ via:

$$\phi(x_1, x_2) = \int_0^\infty \int_0^\infty \mathbb{E}\left(\nu_d(C((x_1, h_1), \eta + \delta_{(x_2,h_2)}))\nu_d(C((x_2, h_2), \eta + \delta_{(x_1,h_1)})\right) \mathrm{d}F(h_1)\mathrm{d}F(h_2).$$

Then, via the multivariate Mecke equation (Theorem 4.4 in [53]) the second term in (6.29) is given by:

$$\frac{1}{n^2}\int_{W_n}\int_{W_n} \phi(x_1, x_2)\langle x_1, x_2\rangle \mathrm{d}x_1\mathrm{d}x_2. \tag{6.31}$$

By Lemma 6.9 we know that $|\phi(x_1, x_2)| \leq 4$. Taking this into account and by applying the Cauchy-Schwarz inequality, we obtain the following upper bound for (6.31):

$$\left|\frac{1}{n^2}\int_{W_n}\int_{W_n} \phi(x_1, x_2)\langle x_1, x_2\rangle \mathrm{d}x_1\mathrm{d}x_2\right| \leq \frac{4}{n^2}\int_{W_n}\int_{W_n} |\langle x_1, x_2\rangle| \,\mathrm{d}x_1\mathrm{d}x_2 \leq \mathrm{diam(W)}^2 4n^{\frac{2}{d}}.$$

Combining this expression with (6.30) yields:

$$\mathbb{E}\left(\|A_n\|^2\right) \leq \frac{\beta_2 \mathrm{diam}(W)^2}{n^{1-\frac{2}{d}}} + \mathrm{diam(W)}^2 4n^{\frac{2}{d}}. \tag{6.32}$$

Unfortunately, the RHS of (6.32) is not uniformly bounded in n. It does yield $\mathbb{E}(\|A_n\|^2) < \infty$ for each $n \in \mathbb{N}$ which we will need later. Currently, it seems that deriving a tight upper bound for $\mathbb{E}\left(\|A_n\|^2\right)$ is hard in general. However, under certain assumptions on $\mathbb{F}$ it becomes possible to obtain a bound of order $O(n^{-1+2/d})$. To prove this, we require the notion of stabilization, which is often used in stochastic geometry for deriving laws of large numbers and central limit theorems, we refer to [88] for an overview. In the context of Poisson-Laguerre tessellations stabilization techniques were used in [30] for proving asymptotic normality of estimators of geometric characteristics of cells.

Theorem 6.5. *[Propositions 3.1 and 3.2 in [30]] Suppose $\mathbb{F}$ is concentrated on $(0, M)$ for some $M > 0$, and let $(x, h) \in \mathbb{R}^d \times (0, M)$. There exists a random variable $R_x(\eta) > 0$ (the radius of stabilization) such that:*

1. *With probability one, for all $(x', h') \in \mathbb{R}^d \times (0, M)$, with $\|x - x'\| > R_x(\eta)$:*

$$C((x, h), \eta) = C\left((x, h), \eta + \delta_{(x',h')}\right).$$

2. *With probability one:*

$$C((x, h), \eta) \subset \bar{B}(x, R_x(\eta)).$$

3. *The random variable $R_x(\eta)$ has exponentially decaying tails. That is, there exists constants $c_1, c_2 > 0$ such that for all $r \geq 0$:*

$$\mathbb{P}(R_x(\eta) > r) < c_1 e^{-c_2 r}.$$

4. *The distribution of $R_x(\eta)$ does not depend on x. Additionally, $R_x(\eta + \delta_x) = R_x(\eta)$ and $R_x(\eta + \delta_{x'}) \leq R_x(\eta)$ for all $x' \in \mathbb{R}^d$.*

Additionally, we recall the Poincaré inequality, which is useful for deriving bounds of the variance of so-called Poisson functionals if one can control the influence of inserting an additional point into the point process.

Lemma 6.10 (Poincaré inequality, Theorem 18.7 in [53])**.** *Suppose η is a Poisson process on the measurable space $(\mathbb{X}, \mathcal{X})$ with σ-finite intensity measure Λ. Suppose $f : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ satisfies $\mathbb{E}(f(\eta)^2) < \infty$, then:*

$$\mathrm{Var}(f(\eta)) \leq \int_{\mathbb{X}} \mathbb{E}\left((D_x f(\eta))^2\right) \Lambda(\mathrm{d}x),$$

with $D_x f(\eta) = f(\eta + \delta_x) - f(\eta)$, for $x \in \mathbb{X}$.

Proposition 6.10 and Theorem 6.5 may be used to prove the theorem below. The proof of Theorem 6.6 is given in section 6.8.

Theorem 6.6. *Let A_n be as in (6.26) with $W_n = n^{\frac{1}{d}} W$ for $n \in \mathbb{N}$, where $W \subset \mathbb{R}^d$ is a convex body satisfying $\nu_d(W) = 1$ and $c(W) = 0$. If $\mathbb{F}$ is concentrated on $(0, M)$ for some $M > 0$, then there exists a constant $0 < \gamma < \infty$ such that for all $n \in \mathbb{N}$:*

$$\mathbb{E}\left(\|A_n\|^2\right) \leq \frac{\gamma}{n^{1-\frac{2}{d}}}.$$

As a consequence, $\lim_{n\to\infty} A_n = 0$ in L^2 for $d \geq 3$. If $d \geq 2$, the sequence $(A_n)_{n\geq 1}$ is uniformly tight.

Theorem 6.6 guarantees consistency of the inversion procedure for a large class of choices for $\mathbb{F}$. Whether the inversion procedure is consistent for all locally finite measures $\mathbb{F}$ on $(0, \infty)$ is an open problem.

Remark 6.4. *In general, if one does not wish to assume a uniform upper bound for the weights, it is not clear whether there is still a way to apply stabilization techniques. In the recent paper [13] a different technique, called region-stabilization, is used in the context of Poisson-Laguerre tessellations for deriving central limit theorems. Here, various parametric models are considered with unbounded weights. Perhaps similar techniques can be used to derive a bound for $\mathbb{E}(\|A_n\|^2)$ without assuming an upper bound for the weights.*

6.6 Simulations

In this section we perform various simulations to empirically study the behavior of the estimator $(\hat{\lambda}_n, \hat{c}_n)$ proposed in this chapter. We also consider a variant of $(\hat{\lambda}_n, \hat{c}_n)$, which takes into account edge effects. By computing these estimators, we effectively invert the observed Poisson-Laguerre tessellations. Keeping in mind the main motivation for inverting Poisson-Laguerre tessellations, we then study how well the function F can be

estimated if the original weighted generators are unknown. That is, first we observe a Poisson-Laguerre tessellation through a bounded window and apply the inversion procedure. Then, we compute a variant of the estimator $\hat{F}_n^0$ (recall Definition 5.6 in chapter 5) for F, which is based on the weighted generator points obtained via the inversion procedure instead of the actual weighted generators. For all simulations we simulate planar ($d = 2$) Poisson-Laguerre tessellations.

6.6.1 Estimation of λ_0 and c_0

Recall the definition of $(\hat{\lambda}_n, \hat{c}_n)$ as a solution to the weighted least squares problem in (6.13). Then, $(\hat{\lambda}_n, \hat{c}_n)$ minimizes a function which cannot actually be computed in practice. After all, when a Poisson-Laguerre tessellation is observed through a window W_n, the cells at the boundary of the window are only partially observed and therefore their volumes and centroids cannot be computed. Therefore, we also consider the following estimator for (λ_0, c_0):

$$(\bar{\lambda}_n, \bar{c}_n) = \underset{(\lambda,c)\in\mathbb{R}\times\mathbb{R}^d}{\arg\min} \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in f_0(\eta)} \mathbb{1}\{C((x,h), f_0(\eta)) \subset W_n\} v_{x,h} \left\|\lambda x + c - c_{x,h}\right\|^2.$$

Because $(\bar{\lambda}_n, \bar{c}_n)$ depends on less information compared to $(\hat{\lambda}_n, \hat{c}_n)$, it is to be expected that its performance will be slightly worse. Loosely speaking, $(\bar{\lambda}_n, \bar{c}_n)$ is based on information contained in W_n and we expect $(\bar{\lambda}_n, \bar{c}_n)$ to behave like $(\hat{\lambda}_n, \hat{c}_n)$ if $(\hat{\lambda}_n, \hat{c}_n)$ is computed based on a smaller observation window $W_n' \subset W_n$ instead of W_n. While $(\hat{\lambda}_n, \hat{c}_n)$ cannot be computed in practice, this is not an issue in a simulation setting, where we can effectively also observe cells outside of W_n.

For the simulations in this section we simulate Poisson-Laguerre tessellations in $\mathbb{R}^2$ with the following choices for the underlying function F, with $z \geq 0$:

$$F_1(z) = z \cdot \mathbb{1}\{z < 1\} + \mathbb{1}\{z \geq 1\} \tag{6.33}$$

$$F_2(z) = 0.01 \cdot \mathbb{1}\{z \geq 1\} + 0.04 \cdot \mathbb{1}\{z \geq 8\} + 0.95 \cdot \mathbb{1}\{z \geq 10\}. \tag{6.34}$$

Note that we also considered these choices for F in section 5.7. For both choices of F it is simple to simulate a corresponding Poisson process, because these Poisson processes can be recognized as independently marked homogeneous Poisson processes. As before, we write $P_n := \mathbb{E}(\eta^0(W_n \times (0,\infty))$, and we choose a square observation window W_n such that $P_n = 1000$. In words, we choose a square W_n with an area such that the expected number of observed points of η^0 in W_n is equal to 1000. Recall that $\eta^0 = \{(x,h) \in \eta : x \in C((x,h),\eta)\}$. For each choice of F and P_n, 100 Laguerre tessellations are simulated, yielding 100 realizations of $(\bar{\lambda}_n, \bar{c}_n)$ and $(\hat{\lambda}_n, \hat{c}_n)$. For the sake of convenience we take $\lambda_0 = 1$ and $c_0 = 0$. The results of these simulations are summarized in Tables 6.1-6.4. In these tables the average absolute errors are shown, as well as the 2.5% and 97.5% quantiles of these absolute errors.

Let us now discuss the contents of these tables. For one, it can be seen that the estimates of λ_0 and c_0 corresponding to F_1 are more accurate compared to estimates corresponding to F_2. A Poisson-Laguerre tessellation with F_1 as the underlying distribution function

	$\lvert\bar{\lambda}_n - \lambda_0\rvert$		$\lVert\bar{c}_n - c_0\rVert$	
P_n	mean	(2.5%, 97.5%)	mean	(2.5%, 97.5%)
500	0.000 766	(0.000 066 , 0.0015)	0.021	(0.0036, 0.042)
1000	0.000 334	(0.000 017 , 0.000 73)	0.0120	(0.0020, 0.024)
2000	0.000 143	(0.000 009 6, 0.000 29)	0.006 75	(0.0010, 0.013)
5000	0.000 061 7	(0.000 025 , 0.000 11)	0.004 33	(0.0013, 0.0077)

Table 6.1: Estimates of λ_0 and c_0 obtained with edge correction, the underlying F is given by (6.33).

	$\lvert\hat{\lambda}_n - \lambda_0\rvert$		$\lVert\hat{c}_n - c_0\rVert$	
P_n	mean	(2.5%, 97.5%)	mean	(2.5%, 97.5%)
500	0.000 386	(0.000 026 , 0.0012)	0.0113	(0.000 91, 0.035)
1000	0.000 138	(0.000 012 , 0.000 37)	0.005 53	(0.000 80, 0.013)
2000	0.000 071 8	(0.000 002 1, 0.000 18)	0.003 48	(0.000 65, 0.0082)
5000	0.000 017 9	(0.000 001 6, 0.000 047)	0.001 46	(0.000 23, 0.0037)

Table 6.2: Estimates of λ_0 and c_0 obtained without edge correction, the underlying F is given by (6.33).

	$\lvert\bar{\lambda}_n - \lambda_0\rvert$		$\lVert\bar{c}_n - c_0\rVert$	
P_n	mean	(2.5%, 97.5%)	mean	(2.5%, 97.5%)
500	0.001 26	(0.000 077 , 0.0039)	0.0450	(0.0058 , 0.14)
1000	0.000 489	(0.000 030 , 0.0014)	0.0222	(0.0037 , 0.062)
2000	0.000 239	(0.000 007 5, 0.000 74)	0.0137	(0.0023 , 0.043)
5000	0.000 086 9	(0.000 002 9, 0.000 27)	0.007 35	(0.000 91, 0.022)

Table 6.3: Estimates of λ_0 and c_0 obtained with edge correction, the underlying F is given by (6.34).

	$\lvert\hat{\lambda}_n - \lambda_0\rvert$		$\lVert\hat{c}_n - c_0\rVert$	
P_n	mean	(2.5%, 97.5%)	mean	(2.5%, 97.5%)
500	0.000 937	(0.000 047 , 0.0029)	0.0339	(0.0042 , 0.11)
1000	0.000 414	(0.000 013 , 0.0010)	0.0194	(0.0019 , 0.045)
2000	0.000 158	(0.000 013 , 0.000 46)	0.0104	(0.0011 , 0.028)
5000	0.000 048 7	(0.000 002 4, 0.000 15)	0.004 39	(0.000 55, 0.011)

Table 6.4: Estimates of λ_0 and c_0 obtained without edge correction, the underlying F is given by (6.34).

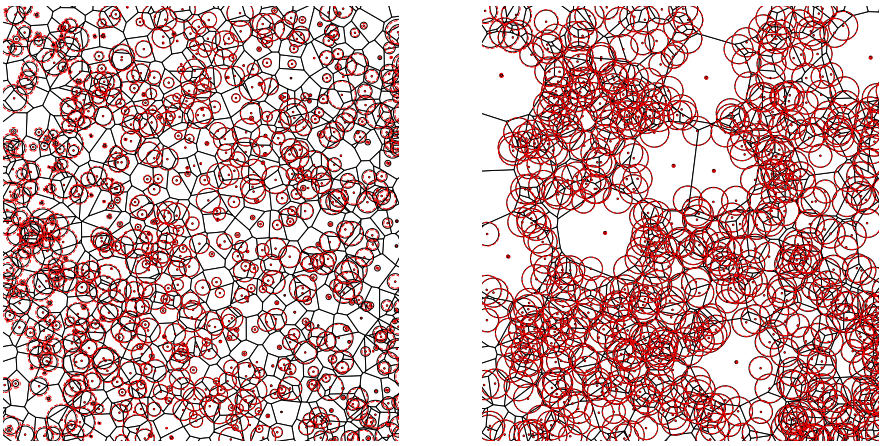


Figure 6.2: Realizations of Poisson-Laguerre tessellations with F_1 (Left) and F_2 (Right) as the underlying distribution function. Black points and black circles represent the original weighted generators (η^*), red points and red dashed circles represent weighted generators ($\hat{\eta}_n^*$) obtained via the estimator $(\hat{\lambda}_n, \hat{c}_n)$, see (6.11).

will have mostly cells of similar sizes. If F_2 is the underlying distribution function then a corresponding Poisson-Laguerre tessellation will have a few large cells, a larger number of medium sized cells and a lot of small cells. Perhaps estimation of (λ_0, c_0) is more difficult when there is more variation in the cell sizes. As anticipated we can also see that the estimator $(\bar{\lambda}_n, \bar{c}_n)$ yields larger mean absolute errors compared to $(\hat{\lambda}_n, \hat{c}_n)$. Finally, it should be noted that the errors corresponding to estimates of λ_0 are far smaller compared to the errors corresponding to estimates of c_0. This is perhaps caused by the rather extreme behavior of the criterion function T_n as $n \to \infty$ if $\lambda \neq \lambda_0$ (Theorem 6.2). Overall, both estimation procedures appear to yield rather close estimates of λ_0 and c_0. Indeed, the errors are sufficiently small that visualizing both the original configuration of generators as well as the generators obtained via any of the two inversion procedures ($(\bar{\lambda}_n, \bar{c}_n)$ and $(\hat{\lambda}_n, \hat{c}_n)$) we see that these points appear to overlap. This is shown in Figure 6.2. Here, the circle radii represent the weights corresponding to the generators. For visualization purposes the circle radii for the realization corresponding to F_2 were normalized (divided by 10). When looking very closely, one can see that the weights of the original generators are slightly different than the weights of the weighted generators obtained via the inversion procedure.

6.6.2 Estimation of F

In this section we essentially investigate whether it is possible to estimate F, if the only available information is a region of a Poisson-Laguerre tessellation observed through a bounded observation window W_n. That is, the underlying weighted generators are considered to be unknown. This also means that cells at the boundary of the window are only

partially observed. The simulations in this section may be seen as a continuation of the previous section, as the same choices for F are considered, and the estimates of F are computed based on weighted generators obtained via the inversion procedure defined via $(\bar{\lambda}_n, \bar{c}_n)$.

Let $\hat{\eta}_n^*$ denote the configuration of weighted generators obtained via the inversion procedure defined via $(\bar{\lambda}_n, \bar{c}_n)$, recall (6.11). Then, $\hat{\eta}_n^*$ may be seen as an approximation of $\eta^* \cap (W_n \times \mathbb{R})$. Write $\hat{\eta}_n^* = \{(\hat{x}_1, \hat{h}_1), \ldots, (\hat{x}_m, \hat{h}_m)\}$, with $\hat{h}_1 \leq \hat{h}_2 \leq \cdots \leq \hat{h}_m$. Then, analogously to the estimator $\hat{F}_n^0$ for F as defined in chapter 5, we now define the estimator $\bar{F}_n^0$ for F as follows. For $z \geq 0$ define:

$$\bar{G}_n(z) := \frac{1}{\nu_d(W_n)} \sum_{(\hat{x},\hat{h}) \in \hat{\eta}_n^*} \mathbb{1}_{W_n}(\hat{x}) \mathbb{1}_{(0,z]}(\hat{h}) \mathbb{1}\{\hat{x} \in C((\hat{x}, \hat{h}), \hat{\eta}_n^*)\}. \tag{6.35}$$

This is essentially an adaptation of the estimator $\hat{G}_n$ for G_F as defined in (5.14) in chapter 5. We define the estimator $\bar{F}_n^0$ as a piece-wise constant distribution function with jump locations at $\hat{h}_1, \ldots, \hat{h}_m$. Hence, we only need to specify the values $\bar{F}_n^0(\hat{h}_i)$ for $i \in \{1, \ldots, m\}$. Let $\hat{h}_0 < \hat{h}_1$ and set $\bar{F}_n^0(\hat{h}_0) = 0$. For $i \in \{1, \ldots, m\}$, the function $\bar{F}_n^0$ is recursively defined via:

$$\begin{aligned} \bar{F}_n^0(\hat{h}_i) = \bar{F}_n^0(\hat{h}_{i-1}) + \left(\bar{G}_n(\hat{h}_i) - \bar{G}_n(\hat{h}_{i-1})\right) \cdot \\ \cdot \exp\left(\kappa_d \sum_{j=1}^{i-1} \left(\hat{h}_i - \hat{h}_j\right)^{\frac{d}{2}} \left(\bar{F}_n^0(\hat{h}_j) - \bar{F}_n^0(\hat{h}_{j-1})\right)\right). \end{aligned} \tag{6.36}$$

As such, we now have an estimator $\bar{F}_n^0$ for F which can be computed if the weighted generators of the (partially) observed cells are a priori unknown. We compare the obtained estimates to realizations of $\hat{F}_n^0$, which can be computed if the weighted generators of the (partially) observed cells are considered known. Then, we can see whether not knowing the generators has a noticeable effect on the resulting estimates. We should mention that for the obtained realizations of $\bar{F}_n^0$ it was needed to apply a shift. Recall that the distribution of a Poisson-Laguerre tessellation is not affected by shifts of the underlying distribution function. As such, if two estimates of F are equal op to a shift, these estimates are considered equally accurate. Hence, to each realization of $\bar{F}_n^0$ a shift was applied which essentially minimizes the average distance to the true function F. The results of these simulations are shown in Figure 6.3.

First, consider the estimates corresponding to F_1. Visually, it appears as if $\bar{F}_n^0$ yields estimates which are smoothed versions of $\hat{F}_n^0$. Overall, both estimators appear to yield accurate estimates. The estimator $\bar{F}_n^0$ appears to be less accurate in the left and right tail of F, and it is very accurate in the middle of its support (z close to 0.5). Meanwhile, realizations of $\hat{F}_n^0$ are very accurate for z close to 0 and slowly become more inaccurate as z becomes larger. The estimates corresponding to F_2 paint a rather different picture. While both estimators appear to provide accurate estimates of F for say $z < 9.5$, the estimates obtained via $\bar{F}_n^0$ become very inaccurate for larger values of z. Some of these realizations of $\bar{F}_n^0$ have an upper bound which is much too large to fit in the frame of the plot.

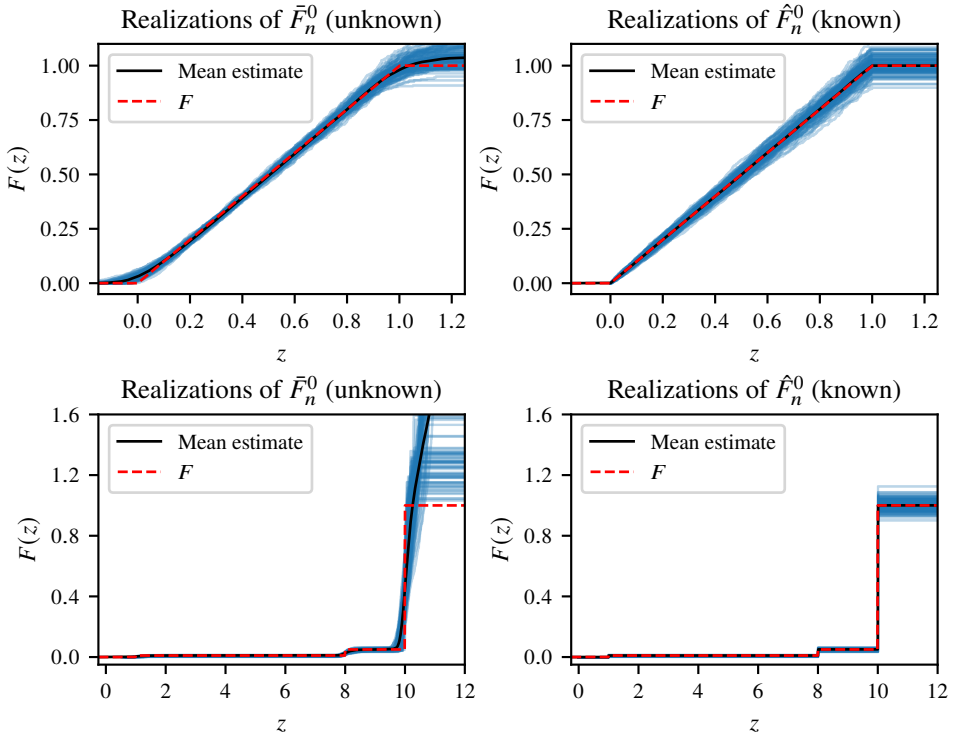


Figure 6.3: Realizations of the estimators $\bar{F}_n^0$ (left panel, weighted generators are unknown) and $\hat{F}_n^0$ (right panel, weighted generators are considered known). In the upper panel F is given by (6.33), in the lower panel F is given by (6.34). In these simulations, $P_n = 1000$.

Recall the coverage function as defined in section 5.8. This function essentially describes how the crystallic growth process behaves on average over time. Besides being useful for determining in which regions estimation of F is difficult, it is also an interesting summary statistic. Therefore, instead of directly comparing estimates of F, we compare the coverage functions induced by the estimates to the coverage function induced by the underlying F. Recall that for a given distribution function F, the coverage function K_F is given by (5.44). The coverage functions corresponding to the estimates of F in Figure 6.3 are shown in Figure 6.4. We can clearly see that if F_1 is the underlying distribution function, then the coverage functions of the obtained estimates of F_1 are very close to the actual coverage function. Meanwhile, if F_2 is the underlying distribution function, the coverage functions induced by the obtained estimates show a significantly larger variance. This holds for coverage functions corresponding to both $\bar{F}_n^0$ and $\hat{F}_n^0$. Curiously, the coverage functions corresponding to $\bar{F}_n^0$ appear to match the true coverage function rather closely for $t \geq 10$. That is, while the realizations of $\bar{F}_n^0$ provide a rather poor estimate of F_2 for large values of z, this barely affects the coverage function in the same regime.

Because $K_{F_2}(10) \approx 0.41$, we would not expect estimation of $F_2(z)$ to be difficult for

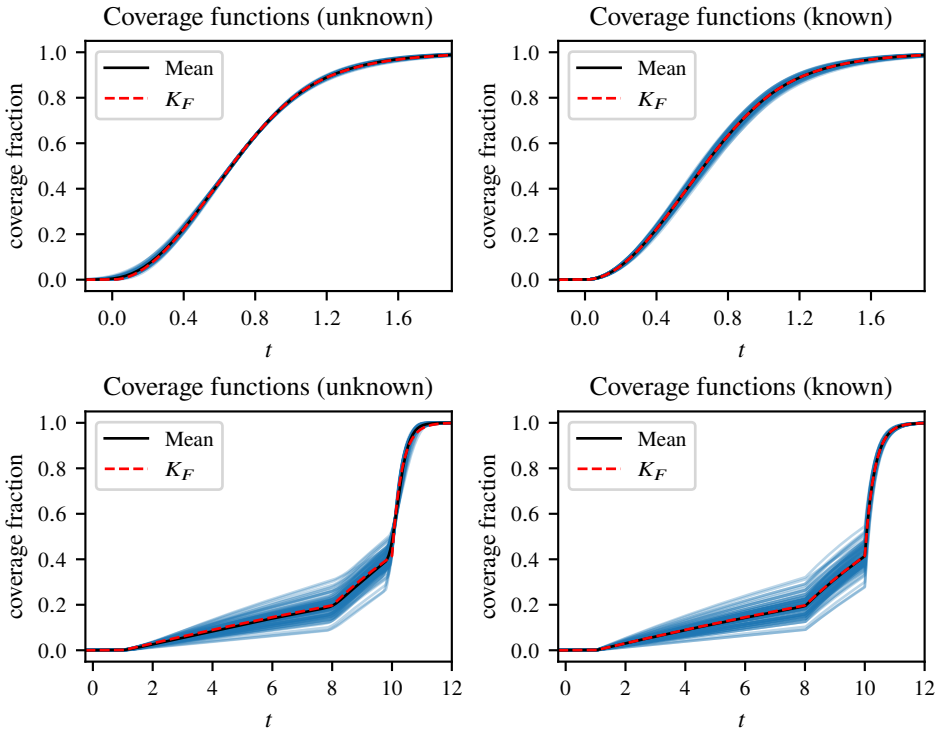


Figure 6.4: Realizations of the coverage functions corresponding to $\bar{F}_n^0$ (left panel, weighted generators are unknown) and $\hat{F}_n^0$ (right panel, weighted generators are considered known). In the upper panel F is given by (6.33), in the lower panel F is given by (6.34). In these simulations, $P_n = 1000$.

z close to 10. Indeed, this is also reflected by the fact that $\hat{F}_n^0$ performs very well in this region. Therefore, we may wonder whether combining the inversion procedure proposed in this chapter and the estimation procedure from chapter 5, leads to a consistent estimator. Perhaps consistency can only be obtained under certain conditions. If the estimator $\bar{F}_n^0$ does turn out to be consistent, then it would be very useful to know how important the behavior of $\bar{F}_n^0(z)$ is for large values of z. The simulations in this section suggest that its behavior barely affects the corresponding coverage function. However, we do not know how the distribution of a Poisson-Laguerre tessellation is affected by taking either F_2, or a realization of $\bar{F}_n^0$ as the underlying distribution function.

6.7 Concluding remarks

In this chapter we derived an inversion procedure for Poisson-Laguerre tessellations. The main motivation for wanting to retrieve the weighted generators of the observed cells is for

their use in statistical inference, as in chapter 5. We obtained various theoretical results and provided conditions for the inversion procedure to be consistent. In simulations we observed that the inversion procedure yields a very close approximation of the original weighted generators. Additionally, we have shown how these weighted generators may be used to estimate the distribution function F. It is however evident that if there is no prior knowledge of the original weighted generators, the resulting estimates of F are less accurate compared to when estimates of F can be computed based on the original weighted generators. Whether the estimators for F based on weighted generators obtained via the inversion procedure are in general consistent is an important open problem.

6.8 Additional proofs

Proof of Lemma 6.2. Let $y \in \mathbb{R}^d \setminus \{0\}$, by (6.48) we have:

$$\int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t) \geq F(h)\|y\|^d.$$

Via this inequality we obtain:

$$\begin{aligned}
p_F(y) &= \int_0^{\infty} \exp\left(-\kappa_d \int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) \mathrm{d}F(h) \\
&\leq \int_0^{\infty} \exp\left(-\kappa_d F(h)\|y\|^d\right) \mathrm{d}F(h) \\
&= \int_{F(0)}^{F(\infty)} \exp\left(-\kappa_d u\|y\|^d\right) \mathrm{d}u \\
&= \frac{1}{\kappa_d\|y\|^d}\left(\exp\left(-\kappa_d \cdot 0 \cdot \|y\|^d\right) - \exp\left(-\kappa_d \cdot F(\infty) \cdot \|y\|^d\right)\right) \\
&\leq \frac{1}{\kappa_d\|y\|^d}.
\end{aligned}$$

Here, we substituted $u = F(h)$. Let $q \in \mathbb{N}$, then substituting $y = r\theta$, with $r \geq 0$ and $\theta \in \mathbb{S}^{d-1}$ we obtain:

$$\begin{aligned}
\int_{\mathbb{R}^d} \|y\|^q p_F(y)\mathrm{d}y &= \int_{\mathbb{R}^d} \|y\|^q \int_0^{\infty} \exp\left(-\kappa_d \int_0^{\|y\|^2+h} \left(\|y\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) \mathrm{d}F(h)\mathrm{d}y \\
&= d\kappa_d \int_0^{\infty}\int_0^{\infty} \exp\left(-\kappa_d \int_0^{r^2+h} \left(r^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) r^{q+d-1}\mathrm{d}r\mathrm{d}F(h).
\end{aligned}$$

This expression is very similar to an expression which appears in the proof of Theorem 6.4. Indeed, if we were to set $q = (p-1)d$, $p \in \mathbb{N}$, then the proof of Theorem 6.4 yields $\int_{\mathbb{R}^d} \|y\|^q p_F(y)\mathrm{d}y < \infty$. Because Theorem 6.4 holds for any $d, p \in \mathbb{N}$ we may simply choose p sufficiently large such that $q \leq (p-1)d$ and then we obtain the result via

Hölder's inequality:

$$\int_{\mathbb{R}^d} \|y\|^q p_F(y)\mathrm{d}y \leq \left(\int_{\mathbb{R}^d} \|y\|^{(p-1)d} p_F(y)\mathrm{d}y\right)^{\frac{(p-1)d}{q}} < \infty.$$

□

Proof of Lemma 6.5. Assume $\varphi \cap B_{x,k} \neq \emptyset$ for all $k \in \{1,\ldots,J\}$. Let $y \in C((x,h),\varphi) - x$, then also $y \in C_j$ for some $j \in \{1,\ldots,J\}$. Choose $x_j \in \varphi \cap B_{x,j}$. Then, $2\sqrt{z} < \|x_j - x\| < R$ and $x_j - x \in C_j$. Because $y \in C((x,h),\varphi) - x$ and $x_j \in \varphi$ we have:

$$\|y\|^2 + h \leq \|x_j - x - y\|^2 + h_j = \|x_j - x\|^2 + \|y\|^2 - 2\langle x_j - x, y\rangle + h_j$$

Using $2\sqrt{z} < \|x_j - x\|$ we now obtain:

$$2\langle x_j - x, y\rangle \leq \|x_j - x\|^2 + h_j - h \leq \|x_j - x\|^2 + z < \frac{5}{4}\|x_j - x\|^2. \tag{6.37}$$

Because $y \in C_j$, $x_j - x \in C_j$ we have by the definition of C_j:

$$\langle x_j - x, y\rangle \leq \frac{3}{4}\|x_j - x\| \cdot \|y\| \iff \|y\| \leq \frac{4}{3}\frac{\langle x_j - x, y\rangle}{\|x_j - x\|}$$

Combining this with (6.37) we obtain the desired result:

$$\|y\| \leq \frac{4}{3}\frac{\langle x_j - x, y\rangle}{\|x_j - x\|} < \frac{5}{6}\frac{\|x_j - x\|^2}{\|x_j - x\|} < \|x_j - x\| < R.$$

Indeed, $C((x,h),\varphi) - x \subset B(0,R) \iff C((x,h),\varphi) \subset B(x,R)$. □

Proof of Theorem 6.2. First, consider the case $\lambda = \lambda_0$. Equation (6.10) may be written as:

$$\begin{aligned} T_n(\lambda_0, c) &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \int_{C((x,h),\eta)} \|x - c_0 + c - y\|^2 \mathrm{d}y \\ &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \int_{C((x,h),\eta)-x} \|c - c_0 - y\|^2 \mathrm{d}y \\ &= \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \int_{C((0,h),S_x\eta)} \|c - c_0 - y\|^2 \mathrm{d}y \end{aligned}$$

Applying the spatial ergodic theorem (Theorem 5.7), $T_n(\lambda_0, c)$ converges almost surely to its expected value (see Lemma 6.3). Now, consider the case $\lambda \neq \lambda_0$. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting the Poisson process η. In order to show that $\lim_{n\to\infty} T_n(\lambda, c) = \infty$ almost surely, it is sufficient to show that for all $M > 0$ there exists a $\Omega_M \in \mathcal{A}$ with $\mathbb{P}(\Omega_M) = 1$ such that for all $\omega \in \Omega_M$: $\liminf_{n\to\infty} T_n(\lambda, c; \omega) > M$. Indeed, setting $\Omega_0 = \cap_{M\in\mathbb{N}}\Omega_M$, we have $\mathbb{P}(\Omega_0) = 1$ and $\lim_{n\to\infty} T_n(\lambda, c; \omega) = \infty$ for all $\omega \in \Omega_0$. Let $M > 0$. We need some additional notation before we choose Ω_M. Pick $z > 0$ such that

$F(z) > 0$. Take $R > 0$ large enough such that $R > 2\sqrt{z}$. Take $r > 0$ small enough such that $r + \sqrt{r^2 + z} < 2\sqrt{z}$. One may for instance take $r = \sqrt{z/3}$. For any $(x, h) \in \mathbb{R}^d \times (0, \infty)$ let the set $D_{x,h}$ be as in Lemma 6.4 and the sets $B_{x,j}$, $j \in \{1, \dots, J\}$ as in Lemma 6.5. For $(x, h) \in \mathbb{R}^d \times (0, \infty)$ define the following event:

$$O_{x,h} = \{\eta(D_{x,h} \setminus \{(x, h)\}) = 0, \eta(B_{x,j}) > 0 \ \ \forall j \in \{1, \dots, J\}\}.$$

Note that $D_{x,h} \cap B_{x,j} = \emptyset$ for all $j \in \{1, \dots, J\}$. Additionally, the $B_{x,j}$'s have disjoint interiors and therefore we have:

$$\mathbb{P}(O_{x,h}) = \mathbb{P}(\eta(D_{x,h} \setminus \{(x, h)\}) = 0) \prod_{j=1}^{J} \mathbb{P}\left(\eta(B_{x,j}) > 0\right).$$

We will now argue that $\mathbb{P}(O_{x,h}) = \mathbb{P}(O_{0,h}) \geq \alpha > 0$, where α is a constant which does not depend on h. Note that proving this implies the following:

$$\int_0^z \mathbb{P}(O_{0,h}) \mathrm{d}F(h) \geq \alpha F(z) > 0. \tag{6.38}$$

Since $(\nu_d \times \mathbb{F})(\{(x, h)\}) = 0$ we have $\mathbb{P}(\eta(D_{x,h} \setminus \{(x, h)\}) = 0) = \mathbb{P}(\eta(D_{x,h}) = 0)$, which may be computed as:

$$\begin{aligned}
\mathbb{P}(\eta(D_{x,h}) = 0) &= \exp\left(-\int_{\mathbb{R}^d} \int_0^{r^2+h} \mathbb{1}\{\|x - x'\| < r + \sqrt{r^2 + h - h'}\} \mathrm{d}F(h') \mathrm{d}x'\right) \\
&= \exp\left(-\kappa_d \int_0^{r^2+h} \left(r + \sqrt{r^2 + h - h'}\right)^d \mathrm{d}F(h')\right) \\
&\geq \exp\left(-\kappa_d \int_0^{r^2+z} \left(z + \sqrt{r^2 + z}\right)^d \mathrm{d}F(h')\right) \\
&= \exp\left(-\kappa_d F(r^2 + z) \left(z + \sqrt{r^2 + z}\right)^d\right) > 0.
\end{aligned}$$

Let $j \in \{1, \dots, J\}$, then:

$$\begin{aligned}
\mathbb{P}\left(\eta(B_{x,j}) > 0\right) &= 1 - \mathbb{P}\left(\eta(B_{x,j}) = 0\right) \\
&= 1 - \exp\left(-F(z)\nu_d\left(\left(B(x, R) \setminus \bar{B}(x, 2\sqrt{z})\right) \cap (C_j + x)\right)\right) \\
&= 1 - \exp\left(-F(z)\nu_d\left(\left(B(0, R) \setminus \bar{B}(0, 2\sqrt{z})\right) \cap C_j\right)\right) > 0.
\end{aligned}$$

So indeed, there exists some $\alpha > 0$ such that for all $h \in [0, z]$ we have $\mathbb{P}(O_{x,h}) =$

$\mathbb{P}(O_{0,h}) \geq \alpha > 0$. We now use the event $O_{x,h}$ to obtain a lower bound for $T_n(\lambda, c)$:

$$
\begin{aligned}
T_n(\lambda, c) \geq & \\
&\geq \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)v_{x,h} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x) \right\|^2 \\
&\geq \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)v_{x,h} \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x) \right\|^2 \mathbb{1}\{O_{x,h}\} \\
&\geq \frac{\kappa_d r^d}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h) \left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x) \right\|^2 \mathbb{1}\{O_{x,h}\}.
\end{aligned}
$$

The final inequality follows from the fact that if the event $O_{x,h}$ occurs then $B(x, r) \subset C((x, h), \eta)$, hence $v_{x,h} \geq \kappa_d r^d$. Let $N > 0$ be large enough such that:

$$
\kappa_d r^d \left(\frac{N - \|c_0\|}{\lambda_0} |\lambda - \lambda_0| - \|c - c_0\| - R \right)^2 \int_0^z \mathbb{P}(O_{0,h}) \mathrm{d}F(h) > M.
$$

Choose $\Omega_M \in \mathcal{A}$ with $\mathbb{P}(\Omega_M) = 1$ such that for all $\omega \in \Omega_M$ the following two properties hold:

$$
\lim_{n\to\infty} \frac{1}{\nu_d(W_n)} \sum_{(x,h)\in\eta(\cdot;\omega)} \mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)\mathbb{1}\{O_{x,h}(\omega)\} = \int_0^z \mathbb{P}(O_{0,h}) \mathrm{d}F(h)
$$

$$
\eta(\bar{B}(0, N) \times (0, z]; \omega) < \infty
$$

Such a set exists by the spatial ergodic theorem (Theorem 5.7) and by the fact that a Poisson random variable with a finite rate parameter is almost surely finite. Suppose $(x, h) \in \eta$, $\|x\| > N$ and the event $O_{x,h}$ occurs such that $C((x, h), \eta) \subset B(x, R)$ and therefore $\|c_{x,h} - x\| < R$. Then, we claim that the following holds:

$$
\left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x) \right\| \geq \frac{N - \|c_0\|}{\lambda_0} |\lambda - \lambda_0| - \|c - c_0\| - R. \tag{6.39}
$$

We will now verify this claim.

$$
\left\| \left(\frac{\lambda}{\lambda_0} - 1\right) x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x) \right\| = \frac{1}{\lambda_0} |\lambda - \lambda_0| \cdot \left\| x - \left(\frac{\lambda c_0 - \lambda_0 c + \lambda_0 (c_{x,h} - x)}{\lambda - \lambda_0} \right) \right\|. \tag{6.40}
$$

Consider the term in brackets in the RHS of (6.40), we obtain the following upper bound

on its norm:

$$\begin{aligned}\left\|\frac{\lambda c_0 - \lambda_0 c + \lambda_0(c_{x,h} - x)}{\lambda - \lambda_0}\right\| &\leq \frac{1}{|\lambda - \lambda_0|}\left(\|\lambda c_0 - \lambda_0 c\| + \lambda_0 \left\|c_{x,h} - x\right\|\right)\\ &\leq \frac{1}{|\lambda - \lambda_0|}\left(\|\lambda c_0 - \lambda_0 c_0\| + \|\lambda_0 c_0 - \lambda_0 c\| + \lambda_0 R\right)\\ &\leq \frac{1}{|\lambda - \lambda_0|}\left(|\lambda - \lambda_0|\|c_0\| + \lambda_0\|c - c_0\| + \lambda_0 R\right)\\ &\leq \|c_0\| + \frac{\lambda_0}{|\lambda - \lambda_0|}\left(\|c - c_0\| + R\right).\end{aligned}$$

Note that for $x, y \in \mathbb{R}^d$ with $\|x\| > N > K > \|y\|$ we have by the reverse triangle inequality: $\|x - y\| \geq |\|x\| - \|y\|| = \|x\| - \|y\| > N - K$. Applying this to the RHS of (6.40) we obtain the inequality in (6.39):

$$\begin{aligned}&\left\|\left(\frac{\lambda}{\lambda_0} - 1\right)x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x)\right\| \geq\\ &\quad\geq \frac{1}{\lambda_0}|\lambda - \lambda_0|\left(N - \left(\|c_0\| + \frac{\lambda_0}{|\lambda - \lambda_0|}\left(\|c - c_0\| + R\right)\right)\right)\\ &\quad= \frac{N - \|c_0\|}{\lambda_0}|\lambda - \lambda_0| - \|c - c_0\| - R.\end{aligned}$$

As a consequence, we may write (where $\bar{B}(0,N)^c = \mathbb{R}^d \setminus \bar{B}(0,N)$):

$$\begin{aligned}&T_n(\lambda, c) \geq\\ &\geq \frac{\kappa_d r^d}{\nu_d(W_n)}\sum_{(x,h)\in\eta}\mathbb{1}_{W_n\cap\bar{B}(0,N)^c}(x)\mathbb{1}_{(0,z]}(h)\left\|\left(\frac{\lambda}{\lambda_0} - 1\right)x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x)\right\|^2\mathbb{1}\{O_{x,h}\}\\ &\quad+ \frac{\kappa_d r^d}{\nu_d(W_n)}\sum_{(x,h)\in\eta}\mathbb{1}_{W_n\cap\bar{B}(0,N)}(x)\mathbb{1}_{(0,z]}(h)\left\|\left(\frac{\lambda}{\lambda_0} - 1\right)x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x)\right\|^2\mathbb{1}\{O_{x,h}\}\\ &\geq \frac{\kappa_d r^d}{\nu_d(W_n)}\sum_{(x,h)\in\eta}\mathbb{1}_{W_n\cap\bar{B}(0,N)^c}(x)\mathbb{1}_{(0,z]}(h)\left(\frac{N - \|c_0\|}{\lambda_0}|\lambda - \lambda_0| - \|c - c_0\| - R\right)^2\mathbb{1}\{O_{x,h}\}\\ &\quad+ \frac{\kappa_d r^d}{\nu_d(W_n)}\sum_{(x,h)\in\eta}\mathbb{1}_{W_n\cap\bar{B}(0,N)}(x)\mathbb{1}_{(0,z]}(h)\left\|\left(\frac{\lambda}{\lambda_0} - 1\right)x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x)\right\|^2\mathbb{1}\{O_{x,h}\}\end{aligned} \tag{6.41}$$

We now consider the two terms in (6.41) separately. The first term may be written as:

$$\frac{\kappa_d r^d}{\nu_d(W_n)}\left(\frac{N - \|c_0\|}{\lambda_0}|\lambda - \lambda_0| - \|c - c_0\| - R\right)^2\sum_{(x,h)\in\eta}\mathbb{1}_{W_n}(x)\mathbb{1}_{(0,z]}(h)\mathbb{1}\{O_{x,h}\}+ \tag{6.42}$$

$$- \frac{\kappa_d r^d}{\nu_d(W_n)}\left(\frac{N - \|c_0\|}{\lambda_0}|\lambda - \lambda_0| - \|c - c_0\| - R\right)^2\sum_{(x,h)\in\eta}\mathbb{1}_{W_n\cap\bar{B}(0,N)}(x)\mathbb{1}_{(0,z]}(h)\mathbb{1}\{O_{x,h}\}. \tag{6.43}$$

By the choice of Ω_M, the expression in (6.42) converges for all $\omega \in \Omega_M$ to:

$$\kappa_d r^d \left(\frac{N - \|c_0\|}{\lambda_0} |\lambda - \lambda_0| - \|c - c_0\| - R \right)^2 \int_0^z \mathbb{P}(O_{0,h}) \mathrm{d}F(h).$$

The expression in (6.43) converges for all $\omega \in \Omega_M$ to 0, since we have:

$$\begin{aligned} 0 &\leq \frac{\kappa_d r^d}{\nu_d(W_n)} \left(\frac{N - \|c_0\|}{\lambda_0} |\lambda - \lambda_0| - \|c - c_0\| - R \right)^2 \sum_{(x,h)\in\eta} \mathbb{1}_{W_n \cap \bar{B}(0,N)}(x) \mathbb{1}_{(0,z]}(h) \mathbb{1}\{O_{x,h}\} \\ &\leq \kappa_d r^d \left(\frac{N - \|c_0\|}{\lambda_0} |\lambda - \lambda_0| - \|c - c_0\| - R \right)^2 \frac{\eta((W_n \cap \bar{B}(0,N)) \times (0,z])}{\nu_d(W_n)}. \end{aligned}$$

Indeed, for sufficiently large n we have $W_n \cap \bar{B}(0,N) = \bar{B}(0,N)$ and since $\eta(\bar{B}(0,N) \times (0,z];\omega) < \infty$ for all $\omega \in \Omega_M$, the fact that $\nu_d(W_n) \to \infty$ as $n \to \infty$ yields the convergence to zero. A similar argument may be used to show that the second term of (6.41) converges to zero, this follows from the following inequalities:

$$\begin{aligned} 0 &\leq \frac{\kappa_d r^d}{\nu_d(W_n)} \sum_{(x,h)\in\eta} \mathbb{1}_{W_n \cap \bar{B}(0,N)}(x) \mathbb{1}_{(0,z]}(h) \left\| \left(\frac{\lambda}{\lambda_0} - 1 \right) x - \frac{\lambda c_0}{\lambda_0} + c - (c_{x,h} - x) \right\|^2 \mathbb{1}\{O_{x,h}\} \\ &\leq \kappa_d r^d \left(\left| \frac{\lambda}{\lambda_0} - 1 \right| N + \left\| c - \frac{\lambda c_0}{\lambda_0} \right\| + R \right)^2 \frac{\eta((W_n \cap \bar{B}(0,N)) \times (0,z])}{\nu_d(W_n)}. \end{aligned}$$

Hence, combining all results, we have for all $\omega \in \Omega_M$:

$$\liminf_{n\to\infty} T_n(\lambda, c; \omega) \geq \kappa_d r^d \left(\frac{N - \|c_0\|}{\lambda_0} |\lambda - \lambda_0| - \|c - c_0\| - R \right)^2 \int_0^z \mathbb{P}(O_{0,h}) \mathrm{d}F(h) > M.$$

□

Proof of Proposition 6.1. We start with proving the first inequality in the proposition for $p \geq 2$. Note that: $\nu_d(C((x,h),\eta)) = \nu_d(C((x,h),\eta) - x)$ by the translation invariance of Lebesgue measure. For $y \in \mathbb{R}^d$ recall that $y \in C((x,h),\eta) - x \iff \eta(A_{x,h,y}) = 0$ with $A_{x,h,y}$ as in (6.1). By (6.3) we now obtain:

$$\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \mathbb{P}\left(\eta(A_{x,h,y_1}) = 0, \ldots, \eta(A_{x,h,y_p}) = 0\right) \mathrm{d}y_1 \ldots \mathrm{d}y_p. \tag{6.44}$$

Instead of integrating over $(\mathbb{R}^d)^p$ we may also integrate over the union of all sets of the form:

$$C_I = \left\{ (y_1, \ldots, y_p) \in (\mathbb{R}^d)^p : \|y_{i_1}\| \leq \|y_{i_2}\| \leq \cdots \leq \|y_{i_p}\| \right\},$$

where $I = (i_1, i_2, \ldots, i_p)$ is a permutation of $(1, 2, \ldots, p)$. Note that the integrand in (6.44) is symmetric in $y_1, \ldots, y_p$. Hence, when integrating over any set C_I the result is

the same for every choice of I. Because there are $p!$ permutations of $(1, 2, \dots, p)$, we may write:

$$\begin{aligned}
&\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) = \\
&\quad = p! \int \cdots \int_C \mathbb{P}\left(\eta(A_{x,h,y_1}) = 0, \dots, \eta(A_{x,h,y_p}) = 0\right) \mathrm{d}y_1 \dots \mathrm{d}y_p \\
&\quad \leq p! \int \cdots \int_C \mathbb{P}\left(\eta(A_{x,h,y_p}) = 0\right) \mathrm{d}y_1 \dots \mathrm{d}y_p \\
&\quad = p! \int \cdots \int_C \exp\left(-\kappa_d \int_0^{\|y_p\|^2+h} \left(\|y_p\|^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) \mathrm{d}y_1 \dots \mathrm{d}y_p. \qquad (6.45)
\end{aligned}$$

The final equality follows from (6.2) and C is given by:

$$C = \left\{(y_1, \dots, y_p) \in (\mathbb{R}^d)^p : \|y_1\| \leq \|y_2\| \leq \cdots \leq \|y_p\|\right\}.$$

Next, we substitute $y_i = r_i\theta_i$ with $r_i \geq 0$ and $\theta_i \in \mathbb{S}^{d-1}$ for $i \in \{1, \dots, p\}$. Then (6.45) is given by:

$$\begin{aligned}
&p!(d\kappa_d)^p \int_0^\infty \int_0^{r_p} \cdots \int_0^{r_3} \int_0^{r_2} \varphi(r_p) r_1^{d-1} \cdots r_p^{d-1} \mathrm{d}r_1 \dots \mathrm{d}r_p \\
&= p!(d\kappa_d)^p \int_0^\infty \varphi(r_p) \left(\int_0^{r_p} \cdots \int_0^{r_3} \int_0^{r_2} r_1^{d-1} \cdots r_p^{d-1} \mathrm{d}r_1 \dots \mathrm{d}r_{p-1}\right) \mathrm{d}r_p, \qquad (6.46)
\end{aligned}$$

with:

$$\varphi(r_p) = \exp\left(-\kappa_d \int_0^{r_p^2+h} \left(r_p^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right).$$

The integral in brackets in (6.46) may be computed as:

$$\int_0^{r_p} \cdots \int_0^{r_3} \int_0^{r_2} r_1^{d-1} \cdots r_p^{d-1} \mathrm{d}r_1 \dots \mathrm{d}r_{p-1} = \frac{1}{(p-1)!d^{p-1}} r_p^{dp-1}.$$

Plugging this back into (6.46) we obtain the first inequality of the proposition:

$$\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) \leq pd\kappa_d^p \int_0^\infty \exp\left(-\kappa_d \int_0^{r_p^2+h} \left(r_p^2 + h - t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) r_p^{pd-1} \mathrm{d}r_p. \qquad (6.47)$$

In the case $p = 1$ (6.3) may be used to show that (6.47) becomes an equality as none of the techniques used to obtain an upper bound for the case $p \geq 2$ need to be used. Consider

the second inequality for general $p \in \mathbb{N}$. Via integration by parts, we obtain:

$$\begin{aligned}\int_0^{r_p^2+h} \left(r_p^2+h-t\right)^{\frac{d}{2}} \mathrm{d}F(t) &= \frac{d}{2}\int_0^{r_p^2+h} \left(r_p^2+h-t\right)^{\frac{d}{2}-1} F(t)\mathrm{d}t \\ &\geq \frac{d}{2}\int_h^{r_p^2+h} \left(r_p^2+h-t\right)^{\frac{d}{2}-1} F(t)\mathrm{d}t \\ &\geq F(h)\frac{d}{2}\int_h^{r_p^2+h} \left(r_p^2+h-t\right)^{\frac{d}{2}-1} \mathrm{d}t \\ &= F(h) r_p^d. \end{aligned} \tag{6.48}$$

We now plug this lower bound into (6.47) and assume $F(h) > 0$:

$$\begin{aligned}\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) &\leq p d \kappa_d^p \int_0^\infty \exp\left(-\kappa_d F(h) r_p^d\right) r_p^{pd-1}\mathrm{d}r_p \\ &= \frac{p}{F(h)^p}\int_0^\infty e^{-u}u^{p-1}\mathrm{d}u \\ &= \frac{p!}{F(h)^p}.\end{aligned}$$

Here, we substituted $u = \kappa_d F(h) r_p^d$. The resulting integral may be recognized as the Gamma function evaluated in p. □

Proof of Lemma 6.8. We derive the result by showing that the first upper bound in the statement of Proposition 6.1 is finite. Choose $z > 1$ sufficiently large such that $z - \sqrt{z} > 1$ and $F(z) > 0$.

$$\begin{aligned}&\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) \\ &\leq p d \kappa_d^p \int_0^z r^{pd-1}\mathrm{d}r + p d \kappa_d^p \int_z^\infty \exp\left(-\kappa_d \int_0^{r^2+h} \left(r^2+h-t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) r^{pd-1}\mathrm{d}r \\ &= \kappa_d^p z^{pd} + p d \kappa_d^p \int_z^\infty \exp\left(-\kappa_d \int_0^{r^2+h} \left(r^2+h-t\right)^{\frac{d}{2}} \mathrm{d}F(t)\right) r^{pd-1}\mathrm{d}r. \end{aligned} \tag{6.49}$$

Via integration by parts, we obtain:

$$\begin{aligned}\int_0^{r^2+h} \left(r^2+h-t\right)^{\frac{d}{2}} \mathrm{d}F(t) &= \frac{d}{2}\int_0^{r^2+h} \left(r^2+h-t\right)^{\frac{d}{2}-1} F(t)\mathrm{d}t \\ &\geq \frac{d}{2}\int_z^{r^2+h} \left(r^2+h-t\right)^{\frac{d}{2}-1} F(z)\mathrm{d}t \\ &\geq F(z)\frac{d}{2}\int_z^{r^2+h} \left(r^2+h-t\right)^{\frac{d}{2}-1} \mathrm{d}t \\ &= F(z)(r^2+h-z)^{\frac{d}{2}}\end{aligned}$$

$$\begin{aligned}
&\geq F(z)(r^2 - z)^{\frac{d}{2}} \\
&= F(z)(r - \sqrt{z})^{\frac{d}{2}}(r + \sqrt{z})^{\frac{d}{2}} \\
&\geq F(z)r^{\frac{d}{2}}.
\end{aligned}$$

Plugging this back into (6.49) and substituting $u = \kappa_d r^{\frac{d}{2}} F(z)$ as in the proof of Proposition 6.1 yields:

$$\begin{aligned}
\mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) &\leq \kappa_d^p z^{pd} + pd\kappa_d^p \int_z^\infty \exp\left(-\kappa_d F(z) r^{\frac{d}{2}}\right) r^{pd-1} \mathrm{d}r \\
&\leq \kappa_d^p z^{pd} + \kappa_d^p z^{pd} + pd\kappa_d^p \int_0^\infty \exp\left(-\kappa_d F(z) r^{\frac{d}{2}}\right) r^{pd-1} \mathrm{d}r \\
&= \kappa_d^p z^{pd} + \frac{p!}{F(z)^p} =: \alpha_p.
\end{aligned}$$

□

Proof of Theorem 6.4. For $p = 1$ the result follows from (6.4). For the remainder of the proof consider $p \geq 2$. Choose $z > 0$ large enough such that $F(z) > 0$. Let $0 < \alpha_p < \infty$ be as Lemma 6.8. Using the second statement of Proposition 6.1 we may write:

$$\begin{aligned}
\int_0^\infty \mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) \mathrm{d}F(h) &\leq \int_0^z \alpha_p \mathrm{d}F(h) + \int_z^\infty \mathbb{E}\left(\nu_d(C((x,h),\eta))^p\right) \mathrm{d}F(h) \\
&\leq \alpha_p F(z) + \int_z^\infty \frac{p!}{F(h)^p} \mathrm{d}F(h) \\
&= \alpha_p F(z) + \int_{F(z)}^{F(\infty)} \frac{p!}{u^p} \mathrm{d}u \qquad (6.50) \\
&= \alpha_p F(z) + \frac{p!}{p-1}\left(\frac{1}{F(z)^{p-1}} - \frac{1}{F(\infty)^{p-1}}\right) \\
&\leq \alpha_p F(z) + \frac{p!}{(p-1)F(z)^{p-1}} =: \beta_p.
\end{aligned}$$

In (6.50) we substituted $u = F(h)$. □

Proof of Lemma 6.9. Let $h_1, h_2 > 0$ and define the following two half spaces:

$$\begin{aligned}
H_1 &= \left\{y \in \mathbb{R}^d : \|x_1 - y\|^2 + h_1 \leq \|x_2 - y\|^2 + h_2\right\} \\
H_2 &= \left\{y \in \mathbb{R}^d : \|x_2 - y\|^2 + h_2 \leq \|x_1 - y\|^2 + h_1\right\}.
\end{aligned}$$

Additionally, let $\hat{H}_1 = H_1 \times (0, \infty)$ and $\hat{H}_2 = H_2 \times (0, \infty)$. The sets $\hat{H}_1$ and $\hat{H}_2$ have disjoint interiors, they only intersect at their boundaries. Because $\mathbb{E}(\eta(\hat{H}_1 \cap \hat{H}_2)) = 0$, we have $\eta(\hat{H}_1 \cap \hat{H}_2) = 0$ almost surely. As a consequence, the Poisson processes $\eta_{\hat{H}_1}$ and $\eta_{\hat{H}_2}$ are independent. Here, $\eta_{\hat{H}_1}$ and $\eta_{\hat{H}_2}$ denote the restrictions of η to $\hat{H}_1$ and $\hat{H}_2$ respectively. Note that $C((x_1,h_1), \eta + \delta_{(x_2,h_2)}) \subset C((x_1,h_1), \eta_{\hat{H}_1} + \delta_{(x_2,h_2)})$. Because

$\eta_{\hat{H}_1}$ and $\eta_{\hat{H}_2}$ are independent, the random variables $\nu_d(C((x_1,h_1),\eta_{\hat{H}_1}+\delta_{(x_2,h_2)}))$ and $\nu_d(C((x_2,h_2),\eta_{\hat{H}_2}+\delta_{(x_1,h_1)}))$ are also independent. As a consequence:

$$\begin{aligned}
&\mathbb{E}\left(\nu_d(C((x_1,h_1),\eta+\delta_{(x_2,h_2)}))\nu_d(C((x_2,h_2),\eta+\delta_{(x_1,h_1)}))\right) \leq \\
&\quad \leq \mathbb{E}\left(\nu_d(C((x_1,h_1),\eta_{\hat{H}_1}+\delta_{(x_2,h_2)}))\nu_d(C((x_2,h_2),\eta_{\hat{H}_2}+\delta_{(x_1,h_1)}))\right) \\
&\quad = \mathbb{E}\left(\nu_d(C((x_1,h_1),\eta_{\hat{H}_1}+\delta_{(x_2,h_2)}))\right)\mathbb{E}\left(\nu_d(C((x_2,h_2),\eta_{\hat{H}_2}+\delta_{(x_1,h_1)}))\right). \qquad (6.51)
\end{aligned}$$

We will now derive a bound for the first expectation in (6.51). By definition of the Laguerre cell, note that $C((x_1,h_1),\eta_{\hat{H}_1}+\delta_{(x_2,h_2)}) = C((x_1,h_1),\eta_{\hat{H}_1})\cap H_1$. Let $\bar{x}_1$ be the orthogonal projection of x_1 onto H_1. Then, for all $y\in H_1$: $\|\bar{x}_1-y\|\leq\|x_1-y\|$. Therefore, we obtain the following inclusion:

$$\begin{aligned}
C\left((x_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1 &= \left\{y\in H_1 : \|x_1-y\|^2+h_1 \leq \|x'-y\|^2+h' \ \forall (x',h')\in\eta_{\hat{H}_1}\right\} \\
&\subseteq \left\{y\in H_1 : \|\bar{x}_1-y\|^2+h_1 \leq \|x'-y\|^2+h' \ \forall (x',h')\in\eta_{\hat{H}_1}\right\} \\
&= C\left((\bar{x}_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}\left(\nu_d\left(C\left((x_1,h_1),\eta_{\hat{H}_1}+\delta_{(x_2,h_2)}\right)\right)\right) &= \mathbb{E}\left(\nu_d\left(C\left((x_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1\right)\right) \\
&\leq \mathbb{E}\left(\nu_d\left(C\left((\bar{x}_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1\right)\right). \qquad (6.52)
\end{aligned}$$

Write:

$$B_{x,h,y} = \left\{(x',h')\in\mathbb{R}^d\times(0,\infty) : \|x-y\|^2+h > \|x-x'\|^2+h'\right\}$$

Then, $y\in C((x,h),\eta) \iff \eta(B_{x,h,y})=0$. Via (6.3) we now compute the expectation in (6.52) as follows:

$$\begin{aligned}
\mathbb{E}\left(\nu_d\left(C\left((\bar{x}_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1\right)\right) &= \int_{\mathbb{R}^d}\mathbb{P}\left(y\in C\left((\bar{x}_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1\right)\mathrm{d}y \\
&= \int_{H_1}\mathbb{P}\left(y\in C\left((\bar{x}_1,h_1),\eta_{\hat{H}_1}\right)\right)\mathrm{d}y \\
&= \int_{H_1}\mathbb{P}\left(\eta_{\hat{H}_1}\left(B_{\bar{x}_1,h_1,y}\right)=0\right)\mathrm{d}y \qquad (6.53)
\end{aligned}$$

Because $\eta_{\hat{H}_1}\left(B_{\bar{x}_1,h_1,y}\right)$ is Poisson distributed we can compute the probability of it being zero as:

$$\begin{aligned}
&\mathbb{P}\left(\eta_{\hat{H}_1}\left(B_{\bar{x}_1,h_1,y}\right)=0\right) = \\
&\quad = \exp\left(-\int_{H_1}\int_0^\infty \mathbb{1}\left\{\|\bar{x}_1-x'\|^2+t<\|\bar{x}_1-y\|^2+h_1\right\}\mathrm{d}F(t)\mathrm{d}x'\right) \\
&\quad = \exp\left(-\int_{H_1}\int_0^{\|\bar{x}_1-y\|^2+h_1} \mathbb{1}\left\{\|\bar{x}_1-x'\|<\sqrt{\|\bar{x}_1-y\|^2+h_1-t}\right\}\mathrm{d}F(t)\mathrm{d}x'\right)
\end{aligned}$$

$$= \exp\left(-\int_0^{\|\bar{x}_1-y\|^2+h_1} \nu_d\left(\bar{B}\left(\bar{x}_1, \sqrt{\|\bar{x}_1-y\|^2+h_1-t}\right)\cap H_1\right) \mathrm{d}F(t)\right) \tag{6.54}$$

Because $\bar{x}_1 \in H_1$, if we consider the volume of the intersection of the half space H_1 and the ball centered at $\bar{x}_1$ as in (6.54), it follows that the volume of that intersection is at least half the volume of said ball. Hence:

$$\mathbb{P}\left(\eta_{\hat{H}_1}\left(B_{\bar{x}_1,h_1,y}\right)=0\right) \leq \exp\left(-\frac{\kappa_d}{2}\int_0^{\|\bar{x}_1-y\|^2+h_1}\left(\|\bar{x}_1-y\|^2+h_1-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right).$$

Plugging this bound back into (6.53) we obtain:

$$\begin{aligned}
&\mathbb{E}\left(\nu_d\left(C\left((\bar{x}_1,h_1),\eta_{\hat{H}_1}\right)\cap H_1\right)\right) \leq \\
&\quad\leq \int_{\mathbb{R}^d}\exp\left(-\frac{\kappa_d}{2}\int_0^{\|\bar{x}_1-y\|^2+h_1}\left(\|\bar{x}_1-y\|^2+h_1-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right)\mathrm{d}y \\
&\quad= \int_{\mathbb{R}^d}\exp\left(-\frac{\kappa_d}{2}\int_0^{\|u\|^2+h_1}\left(\|u\|^2+h_1-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right)\mathrm{d}u
\end{aligned} \tag{6.55}$$

The final equality follows from substituting $u = y - \bar{x}_1$. Note that the bound in (6.55) does not depend on x_1, x_2 and h_2. By symmetry, replacing h_1 with h_2 in (6.55) serves as an upper bound for the second expectation in (6.51). We now conclude:

$$\begin{aligned}
&\int_0^\infty\int_0^\infty \mathbb{E}\left(\nu_d(C((x_1,h_1),\eta+\delta_{(x_2,h_2)}))\nu_d(C((x_2,h_2),\eta+\delta_{(x_1,h_1)}))\right)\mathrm{d}F(h_1)\mathrm{d}F(h_2) \\
&\quad\leq \left(\int_0^\infty\int_{\mathbb{R}^d}\exp\left(-\frac{\kappa_d}{2}\int_0^{\|u\|^2+h}\left(\|u\|^2+h-t\right)^{\frac{d}{2}}\mathrm{d}F(t)\right)\mathrm{d}u\mathrm{d}F(h)\right)^2 \\
&\quad= 2^2.
\end{aligned}$$

The fact that the final integral equals 2 follows from (6.4) by substituting $\tilde{F} = F/2$. □

Proof of Example 3. Let $t \in \mathbb{R}^d$, and let $\bar{A}_{n,j}$ denote the j-th component of $\bar{A}_n$, $j \in \{1,\ldots,d\}$. Because the marginals of the points of a homogeneous Poisson process are independent, the characteristic function of $\bar{A}_n$ satisfies:

$$\varphi_{\bar{A}_n}(t) = \mathbb{E}\left(\exp\left(i\langle t,\bar{A}_n\rangle\right)\right) = \prod_{j=1}^d \mathbb{E}\left(\exp\left(it_j\bar{A}_{n,j}\right)\right)$$

Fix $m \in \mathbb{N}$, and let $U_1,\ldots,U_m \overset{\text{iid}}{\sim} \text{Unif}(W_n)$. Then, conditional on $\psi(W_n) = m$, the restriction of ψ to W_n is equal in distribution to $\sum_{i=1}^m \delta_{U_i}$. Moreover, each U_i is equal in distribution to the vector $(V_{i,1},\ldots,V_{i,d})^T$ with $V_{i,1},\ldots,V_{i,d} \overset{\text{iid}}{\sim} \text{Unif}(-\frac{1}{2}n^{1/d},\frac{1}{2}n^{1/d})$. Via

the law of total expectation we may now compute:

$$\begin{aligned}
\mathbb{E}\left(\exp\left(it_j\bar{A}_{n,j}\right)\right) &= \sum_{m=0}^{\infty}\mathbb{E}\left(\exp\left(it_j\bar{A}_{n,j}\right)\big|\psi(W_n)=m\right)\mathbb{P}(\psi(W_n)=m)\\
&= \sum_{m=0}^{\infty}\mathbb{E}\left(\exp\left(it_j\frac{1}{n}\sum_{k=1}^{m}V_{k,j}\right)\right)\frac{n^m e^{-n}}{m!}\\
&= \sum_{m=0}^{\infty}\varphi_{V_{1,1}}\left(\frac{t_j}{n}\right)^m\frac{n^m e^{-n}}{m!}\\
&= \exp\left(n\left(\varphi_{V_{1,1}}\left(\frac{t_j}{n}\right)-1\right)\right).
\end{aligned}$$

Here, we used the fact that the characteristic function of a sum of independent random variables is simply the product of their individual characteristic functions. Plugging in the characteristic function of $V_{1,1}$ and expressing the sine in terms of its Taylor series centered at zero we obtain for $t_j \neq 0$:

$$\begin{aligned}
\mathbb{E}\left(\exp\left(it_j\bar{A}_{n,j}\right)\right) &= \exp\left(\frac{\sin\left(\frac{t_j}{2n^{1-\frac{1}{d}}}\right)-\frac{t_j}{2n^{1-\frac{1}{d}}}}{\frac{t_j}{2n^{2-\frac{1}{d}}}}\right)\\
&= \exp\left(-\frac{t_j^2}{24n^{1-\frac{2}{d}}}+o\left(\frac{n^{2-\frac{1}{d}}}{n^{5(1-\frac{1}{d})}}\right)\right)\\
&= \exp\left(-\frac{t_j^2}{24n^{1-\frac{2}{d}}}+o\left(n^{\frac{4}{d}-3}\right)\right)
\end{aligned}$$

Taking the product of the marginal characteristic functions and letting $n\to\infty$, we obtain the desired result:

$$\lim_{n\to\infty}\varphi_{\bar{A}_n}(t)=\begin{cases}\exp\left(-\frac{\|t\|^2}{24}\right) & \text{if } d=2\\ 1 & \text{if } d\geq 3.\end{cases}$$

Via similar techniques, the L^2-norm of $\bar{A}_n$ may be computed as:

$$\begin{aligned}
\mathbb{E}\left(\|\bar{A}_n\|^2\right) &= \frac{1}{n^2}\sum_{m=0}^{\infty}\mathbb{E}\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\langle U_i,U_j\rangle\right)\mathbb{P}\left(\psi(W_n)=m\right)\\
&= \frac{1}{n^2}\sum_{m=0}^{\infty}\mathbb{E}\left(\sum_{i=1}^{m}\|U_i\|^2+\sum_{j=1}^{m}\sum_{k=1,k\neq j}^{m}\langle U_i,U_j\rangle\right)\frac{n^m e^{-n}}{m!}\\
&= \frac{1}{n^2}\sum_{m=0}^{\infty}\left(\frac{mdn^{\frac{2}{d}}}{12}+\sum_{j=1}^{m}\sum_{k=1,k\neq j}^{m}\langle\mathbb{E}(U_i),\mathbb{E}(U_j)\rangle\right)\frac{n^m e^{-n}}{m!}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n^2} \sum_{m=0}^{\infty} \left(\frac{m d n^{\frac{2}{d}}}{12} \right) \frac{n^m e^{-n}}{m!} \\
&= \frac{1}{n^2} \frac{d n^{\frac{2}{d}}}{12} e^{-n} \sum_{m=0}^{\infty} m \frac{n^m}{m!} \\
&= \frac{d}{12 n^{1-\frac{2}{d}}}.
\end{aligned}
$$

□

Proof of Theorem 6.6. For $j \in \{1, \ldots, d\}$, let $A_{n,j}$ denote the j-th component of the random vector A_n. Because $\mathbb{E}(A_n) = 0$ we may write:

$$
\mathbb{E}\left(\|A_n\|^2 \right) = \mathbb{E}\left(A_{n,1}^2 + A_{n,2}^2 + \cdots + A_{n,d}^2 \right) = \sum_{j=1}^{d} \mathrm{Var}(A_{n,j}). \tag{6.56}
$$

We will apply Lemma 6.10 to $\mathrm{Var}(A_{n,j})$ for each j. The fact that $\mathbb{E}(A_{n,j}^2) < \infty$ follows from (6.32). Fix $j \in \{1, \ldots, d\}$ and set $f(\eta) = A_{n,j}$. Let $(x', h') \in \mathbb{R}^d \times (0, M)$. For $(x, h) \in \eta$ let $R_x(\eta)$ be as in Theorem 6.5, via property 1 in Theorem 6.5 we may write:

$$
\begin{aligned}
&f(\eta + \delta_{(x',h')}) = \\
&\quad = \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta + \delta_{(x',h')}} \mathbb{1}_{W_n}(x) \nu_d(C((x,h), \eta + \delta_{(x',h')}))x \\
&\quad = \frac{1}{\nu_d(W_n)} \mathbb{1}_{W_n}(x') \nu_d(C((x',h'), \eta))x' + \\
&\qquad + \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) \nu_d(C((x,h), \eta + \delta_{(x',h')}))x \\
&\quad = \frac{1}{\nu_d(W_n)} \mathbb{1}_{W_n}(x') \nu_d(C((x',h'), \eta))x' + \\
&\qquad + \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) \mathbb{1}\{\|x - x'\| \le R_x(\eta)\} \nu_d((C(x,h), \eta + \delta_{(x',h')}))x \\
&\qquad + \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) \mathbb{1}\{\|x - x'\| > R_x(\eta)\} \nu_d(C((x,h), \eta))x.
\end{aligned}
$$

Similarly, we may write:

$$
\begin{aligned}
f(\eta) = &\frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) \mathbb{1}\{\|x - x'\| > R_x(\eta)\} \nu_d(C((x,h), \eta))x + \\
&+ \frac{1}{\nu_d(W_n)} \sum_{(x,h) \in \eta} \mathbb{1}_{W_n}(x) \mathbb{1}\{\|x - x'\| \le R_x(\eta)\} \nu_d(C((x,h), \eta))x.
\end{aligned}
$$

Using these expressions for $f(\eta + \delta_{(x',h')})$ and $f(\eta)$ we derive an upper bound for $|D_{(x',h')}f(\eta)|$. Taking into account that $\|x\| \leq \operatorname{diam}(W)n^{1/d}$, we obtain:

$$
\begin{aligned}
&|D_{(x',h')}f(\eta)| \leq \\
&\leq \frac{1}{n}\mathbb{1}_{W_n}(x')\nu_d(C((x',h'),\eta))\|x'\| + \\
&\quad + \frac{1}{n}\sum_{(x,h)\in\eta}\mathbb{1}_{W_n\cap\bar{B}(x',R_x(\eta))}(x)\left|\nu_d(C((x,h),\eta)) - \nu_d(C((x,h),\eta+\delta_{(x',h')}))\right|\|x\| \\
&\leq \frac{\operatorname{diam}(W)}{n^{1-\frac{1}{d}}}\mathbb{1}_{W_n}(x')\nu_d(C(x',h'),\eta) + \\
&\quad + \frac{\operatorname{diam}(W)}{n^{1-\frac{1}{d}}}\sum_{(x,h)\in\eta}\mathbb{1}_{W_n}(x)\mathbb{1}\{\|x-x'\|\leq R_x(\eta)\}\nu_d(C((x,h),\eta)).
\end{aligned}
$$

Here, we also used the fact: $C((x,h),\eta+\delta_{(x',h')}) \subset C((x,h),\eta)$. Note that this upper bound does not depend on j. Via Lemma 6.10 and (6.56) we obtain:

$$
\begin{aligned}
\mathbb{E}\left(\|A_n\|^2\right) \leq \frac{\operatorname{diam}(W)^2 d}{n^{2-\frac{2}{d}}}\Bigg(&\int_{\mathbb{R}^d}\int_0^M \mathbb{1}_{W_n}(x')\mathbb{E}\left(\nu_d(C((x',h'),\eta))^2\right)\mathrm{d}F(h')\mathrm{d}x' \\
&+ \int_{\mathbb{R}^d}\int_0^M \mathbb{E}\left(\left(\sum_{(x,h)\in\eta}\mathbb{1}_{W_n}(x)\mathbb{1}\{\|x-x'\|\leq R_x(\eta)\}\nu_d(C((x,h),\eta))\right)^2\right)\mathrm{d}F(h')\mathrm{d}x'\Bigg).
\end{aligned}
\tag{6.57}
$$

We now separately determine upper bounds for both of the integrals appearing in (6.57). For the first integral we obtain the following:

$$
\int_{\mathbb{R}^d}\int_0^M \mathbb{1}_{W_n}(x')\mathbb{E}\left(\nu_d(C((x',h'),\eta))^2\right)\mathrm{d}F(h')\mathrm{d}x' \leq \int_{\mathbb{R}^d}\mathbb{1}_{W_n}(x')\beta_2\mathrm{d}x' = \beta_2 n. \tag{6.58}
$$

Here, β_2 is as in Theorem 6.4. Let $(\eta)^2_{\neq}$ denotes the set of all distinct pairs of points of η. That is, if $(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}$, then $(x_1,h_1)\neq(x_2,h_2)$. Now, consider the second integral in (6.57). Expanding the square and applying Fubini, we obtain:

$$
\begin{aligned}
&\int_{\mathbb{R}^d}\int_0^M \mathbb{E}\left(\left(\sum_{(x,h)\in\eta}\mathbb{1}_{W_n}(x)\mathbb{1}\{\|x-x'\|\leq R_x(\eta)\}\nu_d(C((x,h),\eta))\right)^2\right)\mathrm{d}F(h')\mathrm{d}x' \\
&= \int_{\mathbb{R}^d}\int_0^M \mathbb{E}\left(\sum_{(x,h)\in\eta}\mathbb{1}_{W_n}(x)\mathbb{1}\{\|x-x'\|\leq R_x(\eta)\}\nu_d(C((x,h),\eta))^2\right)\mathrm{d}F(h')\mathrm{d}x' + \\
&\quad + \int_{\mathbb{R}^d}\int_0^M \mathbb{E}\Bigg(\sum_{(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}}\mathbb{1}_{W_n}(x_1)\mathbb{1}_{W_n}(x_2)\mathbb{1}\{\|x_1-x'\|\leq R_{x_1}(\eta)\}\cdot \\
&\quad \cdot\mathbb{1}\{\|x_2-x'\|\leq R_{x_2}(\eta)\}\nu_d(C((x_1,h_1),\eta))\nu_d(C((x_2,h_2),\eta))\Bigg)\mathrm{d}F(h')\mathrm{d}x'
\end{aligned}
$$

$$
\begin{aligned}
&= F(M)\mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)\int_{\mathbb{R}^d} \mathbb{1}\{\|x-x'\|\le R_x(\eta)\}\mathrm{d}x'\nu_d(C((x,h),\eta))^2\right)\\
&\quad + F(M)\mathbb{E}\Bigg(\sum_{(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}} \mathbb{1}_{W_n}(x_1)\mathbb{1}_{W_n}(x_2)\int_{\mathbb{R}^d} \mathbb{1}\{\|x_1-x'\|\le R_{x_1}(\eta)\}\cdot\\
&\quad \cdot \mathbb{1}\{\|x_2-x'\|\le R_{x_2}(\eta)\}\mathrm{d}x'\nu_d(C((x_1,h_1),\eta))\nu_d(C((x_2,h_2),\eta))\Bigg)\\
&= F(M)\kappa_d\mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)R_x(\eta)^d\nu_d(C((x,h),\eta))^2\right) \qquad (6.59)\\
&\quad + F(M)\mathbb{E}\Bigg(\sum_{(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}} \mathbb{1}_{W_n}(x_1)\mathbb{1}_{W_n}(x_2)\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))\cdot\\
&\quad \cdot \nu_d(C((x_1,h_1),\eta))\nu_d(C((x_2,h_2),\eta))\Bigg). \qquad (6.60)
\end{aligned}
$$

An upper bound for the term in (6.59) may be computed via the Mecke equation as follows:

$$
\begin{aligned}
&F(M)\kappa_d\mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)R_x(\eta)^d\nu_d(\bar{B}(x,R_x(\eta)))^2\right)\le\\
&\qquad \le F(M)\kappa_d^3\mathbb{E}\left(\sum_{(x,h)\in\eta} \mathbb{1}_{W_n}(x)R_x(\eta)^{3d}\right)\\
&\qquad = F(M)\kappa_d^3\int_{\mathbb{R}^d}\int_0^M \mathbb{1}_{W_n}(x)\mathbb{E}(R_x(\eta+\delta_x)^{3d})\mathrm{d}F(h)\mathrm{d}x\\
&\qquad \le F(M)^2\kappa_d^3\mathbb{E}\left(R_0(\eta)^{3d}\right)n. \qquad (6.61)
\end{aligned}
$$

Here, we also used point 4 of Theorem 6.5. Because $R_0(\eta)$ has exponentially decaying tails by point 3 of Theorem 6.5, all of its moments are finite, hence $\mathbb{E}(R_0(\eta)^{3d}) < \infty$. Next, we consider the term in (6.60). This term may be computed via the multivariate Mecke equation (Theorem 4.4 in [53]). By also using points 2 and 4 of Theorem 6.5 we obtain:

$$
\begin{aligned}
&F(M)\mathbb{E}\Bigg(\sum_{(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}} \mathbb{1}_{W_n}(x_1)\mathbb{1}_{W_n}(x_2)\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))\cdot\\
&\cdot \nu_d(C((x_1,h_1),\eta))\nu_d(C((x_2,h_2),\eta))\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&\leq F(M)\kappa_d^2\mathbb{E}\Bigg(\sum_{(x_1,h_1),(x_2,h_2)\in(\eta)^2_{\neq}} \mathbb{1}_{W_n}(x_1)\mathbb{1}_{W_n}(x_2)\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))\cdot\\
&\cdot R_{x_1}(\eta)^d R_{x_2}(\eta)^d\Bigg)\\
&= F(M)\kappa_d^2\int_{W_n}\int_0^M\int_{W_n}\int_0^M \mathbb{E}\Bigg(\nu_d(\bar{B}(x_1,R_{x_1}(\eta+\delta_{x_2}))\cap\bar{B}(x_2,R_{x_2}(\eta+\delta_{x_1})))\cdot\\
&\cdot R_{x_1}(\eta+\delta_{x_2})^d R_{x_2}(\eta+\delta_{x_1})^d\Bigg)\mathrm{d}F(h_1)\mathrm{d}x_1\mathrm{d}F(h_2)\mathrm{d}x_2\\
&\leq F(M)^3\kappa_d^2\int_{W_n}\int_{W_n}\mathbb{E}\Bigg(\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))R_{x_1}(\eta)^d R_{x_2}(\eta)^d\Bigg)\mathrm{d}x_1\mathrm{d}x_2
\end{aligned}
\tag{6.62}
$$

Note that for any $x_1, x_2 \in \mathbb{R}^d$:

$$
\begin{aligned}
\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta))) = \nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))\cdot\\
\cdot\,\mathbb{1}\{\|x_1-x_2\|\leq R_{x_1}(\eta)+R_{x_2}(\eta)\}
\end{aligned}
$$

As a consequence, (6.62) may be written as:

$$
\begin{aligned}
&F(M)^3\kappa_d^2\int_{W_n}\int_{W_n}\mathbb{E}\Bigg(\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))R_{x_1}(\eta)^d R_{x_2}(\eta)^d\cdot\\
&\cdot\,\mathbb{1}\{\|x_1-x_2\|\leq R_{x_1}(\eta)+R_{x_2}(\eta)\}\Bigg)\mathrm{d}x_1\mathrm{d}x_2\\
&= F(M)^3\kappa_d^2\int_{W_n}\int_{W_n}\mathbb{E}\Bigg(\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))R_{x_1}(\eta)^d R_{x_2}(\eta)^d\cdot\\
&\cdot\,\mathbb{1}\{\|x_1-x_2\|\leq R_{x_1}(\eta)+R_{x_2}(\eta)\}\mathbb{1}\{R_{x_1}(\eta)\leq R_{x_2}(\eta)\}\Bigg)\mathrm{d}x_1\mathrm{d}x_2+
\end{aligned}
\tag{6.63}
$$

$$
\begin{aligned}
&+F(M)^3\kappa_d^2\int_{W_n}\int_{W_n}\mathbb{E}\Bigg(\nu_d(\bar{B}(x_1,R_{x_1}(\eta))\cap\bar{B}(x_2,R_{x_2}(\eta)))R_{x_1}(\eta)^d R_{x_2}(\eta)^d\cdot\\
&\cdot\,\mathbb{1}\{\|x_1-x_2\|\leq R_{x_1}(\eta)+R_{x_2}(\eta)\}\mathbb{1}\{R_{x_1}(\eta)> R_{x_2}(\eta)\}\Bigg)\mathrm{d}x_1\mathrm{d}x_2
\end{aligned}
\tag{6.64}
$$

We now derive an upper bound for the expression in (6.63), the obtained upper bound will

also hold for (6.64) by symmetry in x_1 and x_2. The following is an upper bound for (6.63):

$$\begin{aligned}
&F(M)^3\kappa_d^2\int_{W_n}\int_{W_n}\mathbb{E}\Bigg(\nu_d(\bar{B}(x_2,R_{x_2}(\eta)))R_{x_2}(\eta)^{2d}\mathbb{1}\{\|x_1-x_2\|\le 2R_{x_2}(\eta)\}\cdot\\
&\cdot\mathbb{1}\{R_{x_1}(\eta)\le R_{x_2}(\eta)\}\Bigg)\mathrm{d}x_1\mathrm{d}x_2\\
&\le F(M)^3\kappa_d^3\int_{W_n}\mathbb{E}\Bigg(R_{x_2}(\eta)^{3d}\int_{\mathbb{R}^d}\mathbb{1}\{\|x_1-x_2\|\le 2R_{x_2}(\eta)\}\mathrm{d}x_1\Bigg)\mathrm{d}x_2\\
&= F(M)^3\kappa_d^4 2^d\mathbb{E}(R_0(\eta)^{4d})n. \qquad (6.65)
\end{aligned}$$

Plugging (6.58), (6.61) and (6.65) back into (6.57) yields the desired result.

$$\mathbb{E}\left(\|A_n\|^2\right)\le\frac{\mathrm{diam}(W)^2 d}{n^{1-\frac{2}{d}}}\left(\beta_2+F(M)^2\kappa_d^3\mathbb{E}\left(R_0(\eta)^{3d}\right)+F(M)^3\kappa_d^4 2^{d+1}\mathbb{E}\left(R_0(\eta)^{4d}\right)\right).$$

□

Chapter 7

Conclusion

In this thesis we have studied mathematical models which may describe materials microstructures, and we have developed statistical methodology for estimating the parameter(s) of these models in a stereological setting. In particular, we considered a class of particle processes with particles of the same known shape in part I of this thesis, and we studied Poisson-Laguerre tessellations in part II of this thesis. In both cases the underlying parameter of interest is a distribution function. For these distribution functions we derived nonparametric estimators, proved strong consistency, addressed computational aspects and studied their performance in simulations. As such, our methodology contributes novel estimation procedures, as consistent estimators did not yet exist within some of the models we studied. We would like to note that the particle processes in part I are more flexible compared to the commonly considered Wicksell model, and Laguerre tessellations provide additional flexibility over the more commonly considered Voronoi tessellations. As such, we provide materials scientists with statistical tools for models which may describe real materials microstructures more accurately compared to previously studied models.

There is still a wide range of interesting future research directions which can be pursued. From the statistical point of view, the asymptotic distributions of our estimators are especially of interest. In the case of the nonparametric maximum likelihood estimator proposed in chapter 3, studying its asymptotic behavior is difficult due to its implicit definition. Meanwhile, the asymptotic behavior of the estimators in chapter 5 is challenging to study due to dependence between observations. Results on these asymptotic distributions may then be used to construct asymptotic confidence intervals for these estimators. These confidence intervals may then for instance be used to determine guidelines for required sample sizes. Another research direction is to attempt proving (strong) consistency of the inversion method proposed in chapter 6 without assuming an upper bound for the weights. A related important problem is to study whether applying the estimators for F from chapter 5 to weighted generators obtained via the inversion procedure from chapter 6, yields consistent estimators for F. Finally, one may be interested in adapting the methodology in this thesis to incorporate edge effects. For the models studied in this thesis we only considered edge effects in observations to a limited extent, perhaps more sophisticated edge

correction techniques can be developed, to ensure minimal loss of information.

From the application point of view there are also a number of interesting research questions. First of all, it would be of interest to apply the estimation procedure we introduced for Poisson-Laguerre tessellations to real materials data. Then, one can study whether the obtained estimates of the distribution function F (or H) lead to simulated Poisson-Laguerre tessellations which are similar to the real 2D (or 3D) materials data. It is evident that the models considered in this thesis are more flexible compared to for instance Wicksell's model or the Poisson-Voronoi tessellation. However, there are still various properties of real materials which cannot be accurately described by the models studied in this thesis. More advanced models are needed to describe for instance materials microstructures with non-convex grains or anisotropy. A possible approach that could be taken here is to study a further generalization of the Laguerre tessellation, known as the generalized balanced power diagram. Here, it is again needed to choose an appropriate underlying point process model, a natural starting point would be to consider a Poisson process. Alternatively, one may consider studying random Laguerre tessellations generated by point processes other than the Poisson process.

Appendix A

Mathematical background

In this appendix we introduce definitions from the theory of point processes and stochastic geometry, these definitions provide a formal background for the models studied in this thesis. The presentation in this appendix is far from complete and we refer the interested reader to [22], [23], [86] and [53] for further background reading. Throughout, let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space supporting all considered random elements.

A.1 Point processes

While one may think of a point process as a random set of points, a point process is often defined as a random counting measure. Let $\mathbb{X}$ be a complete separable metric space, e.g. $\mathbb{X} = \mathbb{R}^d$. A measure μ on $\mathbb{X}$ is locally finite if $\mu(B) < \infty$ for all bounded $B \in \mathcal{B}(\mathbb{X})$. Here, $\mathcal{B}(\mathbb{X})$ denotes the Borel σ-algebra of $\mathbb{X}$. Let $\mathbf{N}(\mathbb{X})$ denote the space of locally finite counting measures (integer-valued measures) on $\mathbb{X}$. We equip $\mathbf{N}(\mathbb{X})$ with the usual σ-algebra $\mathcal{N}(\mathbb{X})$, which is the smallest σ-algebra on $\mathbf{N}(\mathbb{X})$ such that the mappings $\mu \mapsto \mu(B)$ are measurable for all $B \in \mathcal{B}(\mathbb{X})$.

Definition A.1 (Point process)**.** A point process on $\mathbb{X}$ is a random element η of $(\mathbf{N}(\mathbb{X}), \mathcal{N}(\mathbb{X}))$, that is a measurable mapping $\eta : \Omega \to \mathbf{N}(\mathbb{X})$.

For any $B \in \mathcal{B}(\mathbb{X})$, the random variable $\eta(B)$ represents the number of points of η in B. The analogue of the first moment of a random variable for a point process is the so-called intensity measure. This measure describes the expected number of points of the point process in a given region.

Definition A.2 (Intensity measure)**.** Let η be a point process on $\mathbb{X}$. The intensity measure Λ of η is the measure defined via $\Lambda(B) = \mathbb{E}(\eta(B))$ for $B \in \mathcal{B}(\mathbb{X})$.

The intensity measure is sometimes called the first order moment measure, there are also higher order moment measures, which we will not discuss here. If one wishes to study a point process as a random set of points, care needs to be taken. A point process η as defined above may have multiplicities, that is, there may be duplicates of points. Formally this means that with a positive probability one may observe a realization of η satisfying

$\eta(\{x\}) > 1$ for some $x \in \mathbb{X}$. It becomes possible to switch between the interpretations of a point process as a random counting measure or as a random set of points if one restricts oneself to point processes without multiplicities, which are called simple point processes. Let δ denote the Dirac measure, hence for $x \in \mathbb{X}$ and $B \in \mathcal{B}(\mathbb{X})$: $\delta_x(B) = \mathbb{1}\{x \in B\}$. A counting measure μ on $\mathbb{X}$ is called simple if $\mu(\{x\}) \leq 1$ for all $x \in \mathbb{X}$. As such, a simple counting measure has no multiplicities. Similarly, a point process η on $\mathbb{X}$ is called simple if $\mathbb{P}(\eta(\{x\}) \leq 1,\ \forall x \in \mathbb{X}) = 1$. Let $\mathbf{N}_s(\mathbb{X})$ be the subset of $\mathbf{N}(\mathbb{X})$ containing all simple measures. Define: $\mathcal{N}_s(\mathbb{X}) := \{A \cap \mathbf{N}_s(\mathbb{X}) : A \in \mathcal{N}(\mathbb{X})\}$. Then, a simple point process on $\mathbb{X}$ may be seen as a random element η of $(\mathbf{N}_s(\mathbb{X}), \mathcal{N}_s(\mathbb{X}))$. Let us now formalize what is meant by a random set, or more specifically, a random closed set. Let $\mathcal{F}(\mathbb{R}^d)$ denote the system of closed subsets of $\mathbb{R}^d$. $\mathfrak{F}(\mathbb{R}^d)$ denotes the σ-algebra on $\mathcal{F}(\mathbb{R}^d)$ which is generated by all families $\mathcal{F}^K = \{F \in \mathcal{F}(\mathbb{R}^d) : F \cap K = \emptyset\}$, $K \in \mathcal{K}^d$. Recall that $\mathcal{K}^d$ denotes the space of convex bodies in $\mathbb{R}^d$, and a convex body is a convex and compact set with non-empty interior. The definition of a random closed set in $\mathbb{R}^d$ is as follows:

Definition A.3 (Random closed set)**.** A random closed set in $\mathbb{R}^d$ is a random element X of $(\mathcal{F}(\mathbb{R}^d), \mathfrak{F}(\mathbb{R}^d))$, that is a measurable mapping $X : \Omega \to \mathcal{F}(\mathbb{R}^d)$.

For $x \in \mathbb{R}^d, r > 0$ let $\bar{B}(x, r) = \{y \in \mathbb{R}^d : \|x - y\| \leq r\}$ denote the closed ball with radius r, centered at x. As an example of a random closed set, let Q be a probability measure on $(0, \infty)$ and draw $R \sim Q$, then the random ball $\bar{B}(0, R)$ is a random closed set. The following theorem highlights how one can switch between the two interpretations of a point process:

Theorem A.4 (Theorem 2.5. in [11])**.** *If η is a point process on $\mathbb{X}$, then* supp η *is a random closed set. On the other hand, if X is a locally finite random closed set (that is* $\mathrm{card}(X \cap B) < \infty$ *almost surely for bounded $B \in \mathcal{B}(\mathbb{X})$) then $\eta(\cdot) = \mathrm{card}(X \cap \cdot)$ is a simple point process on $\mathbb{X}$.*

Throughout this thesis we only consider simple point processes, and therefore we may switch between the two interpretations of a point process as needed. By corollary 1.6.12. from [8] we may even enumerate the points of a simple point process η on $\mathbb{X}$, and write:

$$\eta = \{x_1, x_2, \dots\} \quad \text{and} \quad \eta = \sum_{i=1}^{\eta(\mathbb{X})} \delta_{x_i}.$$

As an example of the flexibility obtained by considering a simple point process as a random counting measure or a random closed set, consider the following. Let $f : \mathbb{X} \to \mathbb{R}$ be a measurable function, then the following notations are equivalent:

$$\sum_{x \in \eta} f(x) = \int_{\mathbb{X}} f(x)\eta(\mathrm{d}x).$$

The Campbell formula allows for computing expectations of these kinds of random variables obtained by summing over points of the point process:

Theorem A.5 (Campbell formula)**.** *Let η be a simple point process on $\mathbb{X}$ with intensity measure Λ. Let $f : \mathbb{X} \to \mathbb{R}$ be a measurable function satisfying $f \geq 0$ or $\int_{\mathbb{X}} |f(x)|\Lambda(\mathrm{d}x) < \infty$, then:*

$$\mathbb{E}\left(\sum_{x\in\eta} f(x)\right) = \int_{\mathbb{X}} f(x)\Lambda(\mathrm{d}x).$$

Arguably the most important point process is the Poisson process, which is defined as follows:

Definition A.6 (Poisson process)**.** Suppose Λ is a σ-finite measure on $\mathbb{X}$. A Poisson process with intensity measure Λ is a point process η on $\mathbb{X}$ with the following two properties:

1. For every $B \in \mathcal{B}(\mathbb{X})$, the random variable $\eta(B)$ is Poisson distributed with mean $\Lambda(B)$.
2. For every $m \in \mathbb{N}$ and pairwise disjoint sets $B_1, \ldots, B_m \in \mathcal{B}(\mathbb{X})$, the random variables $\eta(B_1), \ldots, \eta(B_m)$ are independent.

If η is a Poisson process on $\mathbb{X} = \mathbb{R}^d$ and $\Lambda = \lambda\nu_d$ with $\lambda > 0$ and ν_d Lebesgue measure on $\mathbb{R}^d$, then η is referred to as the homogeneous Poisson process with intensity λ. A Poisson process is parameterized via its intensity measure, as two Poisson processes with the same intensity measure are equal in distribution. A Poisson process is a simple point process if and only if its intensity measure Λ has no atoms. That is, $\Lambda(\{x\}) = 0$ for all $x \in \mathbb{X}$. The following Theorem highlights how one can simulate a Poisson process:

Theorem A.7 (Proposition 3.8 in [53])**.** *Let η be a Poisson process on $\mathbb{X}$ with intensity measure Λ. Suppose $0 < \Lambda(\mathbb{X}) < \infty$, then η is equal in distribution to the following point process:*

$$\sum_{i=1}^{\kappa} \delta_{X_i},$$

where $X_1, X_2, \ldots \overset{iid}{\sim} \Lambda(\cdot)/\Lambda(\mathbb{X})$ and independently of the X_i's: $\kappa \sim \mathrm{Poisson}(\Lambda(\mathbb{X}))$.

If $\Lambda(\mathbb{X}) = \infty$ one can still simulate the restriction of this Poisson process to a subset $B \subset \mathbb{X}$ with $B \in \mathcal{B}(\mathbb{X})$ and $0 < \Lambda(B) < \infty$. This is the case, because this restriction is a Poisson process on B, with intensity measure $\Lambda(B \cap \cdot)$.

A point process on $\mathbb{R}^d$ is stationary if its behavior is the same in every region. More formally, for $v \in \mathbb{R}^d$ let S_v denote the shift operator. Suppose $\eta = \{x_1, x_2, \ldots\}$ is a simple point process on $\mathbb{R}^d$. Then, we define $S_v\eta := \{x_1 - v, x_2 - v, \ldots\}$. Additionally, for a deterministic set $B \subset \mathbb{R}^d$ we define $S_v B := \{x + v : x \in B\}$. Note that in the random counting measure interpretation of a point process, the definition is as follows: $S_v\eta(B) := \eta(S_v B)$, for $B \in \mathcal{B}(\mathbb{R}^d)$. This is indeed consistent with the previous definition since $S_v\eta(B) = \sum_i \delta_{x_i}(S_v B) = \sum_i \delta_{x_i - v}(B)$. We call η stationary if $S_v\eta$ and η are equal in distribution for all $v \in \mathbb{R}^d$. The homogeneous Poisson process on $\mathbb{R}^d$ is an example of a stationary point process. The intensity measure of a stationary point process on $\mathbb{R}^d$ is given by $\lambda\nu_d$ for some $\lambda > 0$. We then refer to λ as the intensity of this stationary point process.

The Poisson process is frequently used as a building block for defining more complicated models. Those models are often tractable thanks to the attractive independence properties of the Poisson process. However, many interesting models arise when one does introduce dependence into the model. We will illustrate this via the Mecke equation:

Theorem A.8 (Mecke equation)**.** *Let Λ be a σ-finite measure on $\mathbb{X}$ and let η be a point process on $\mathbb{X}$. Then η is a Poisson process with intensity measure Λ if and only if:*

$$\mathbb{E}\left(\sum_{x\in\eta} f\left(x,\eta\right)\right) = \int \mathbb{E}\left(f\left(x,\eta+\delta_x\right)\right)\Lambda(\mathrm{d}x),$$

for all non-negative measurable functions $f : \mathbb{X}\times\mathbf{N}(\mathbb{X}) \to [0,\infty]$.

Besides being an interesting characterization of the Poisson process, the Mecke equation has computational value. Recall the statistics considered in the Campbell formula. There, we consider statistics which sum over $f(x)$ for all $x \in \eta$. As such, each point x contributes to this sum, and each contribution only depends on x, not on the other points. In the Mecke equation, we consider a sum where each contribution not only depends on x but also on all other points of the point process. As a consequence, the Mecke equation may for example be used for studying models which involve a dependent thinning, as highlighted in the example below. A thinning of a point process η is a point process which is obtained by removing points from η according to some possibly random rule.

Example 4 (Matérn I hard-core point process)**.** Let η be a homogeneous Poisson process on $\mathbb{R}^d$ with intensity $\lambda > 0$ and fix $r > 0$. A Matérn I hard-core point process ψ is obtained as a thinning of η as follows. A point $x \in \eta$ belongs to ψ if and only if no other point of η is within distance r of x. That is, we may write: $\psi = \{x \in \eta : \|x-y\| > r\ \forall y \in \eta\setminus\{x\}\}$. It is readily observed that ψ is also given by $\psi = \{x \in \eta : \eta(\bar{B}(x,r)) = 1\}$. Let $B \in \mathcal{B}(\mathbb{R}^d)$, the intensity measure of ψ is given by:

$$\begin{aligned}
\mathbb{E}\left(\psi\left(B\right)\right) &= \mathbb{E}\left(\sum_{x\in\psi} \mathbb{1}_B(x)\right)\\
&= \mathbb{E}\left(\sum_{x\in\eta} \mathbb{1}_B(x)\mathbb{1}\left\{\eta\left(\bar{B}(x,r)\right)=1\right\}\right)\\
&= \lambda\int_B \mathbb{P}\left(\left(\eta+\delta_x\right)\left(\bar{B}(x,r)\right)=1\right)\mathrm{d}x \qquad \text{(Mecke equation)}\\
&= \lambda\int_B \mathbb{P}\left(\eta\left(\bar{B}(x,r)\right)=0\right)\mathrm{d}x.
\end{aligned}$$

Here, we used: $(\eta+\delta_x)(\bar{B}(x,r)) = 1 \iff \eta(\bar{B}(x,r)) = 0$ for all $x \in \mathbb{R}^d$. Because η is a Poisson process, $\eta(\bar{B}(x,r))$ is Poisson distributed with mean $\lambda\nu_d(\bar{B}(x,r)) = \lambda\kappa_d r^d$ for any $x \in \mathbb{R}^d$. Here, κ_d denotes the volume of the unit ball in $\mathbb{R}^d$. Hence, we obtain $\mathbb{P}\left(\eta(\bar{B}(x,r))=0\right) = \exp\left(-\lambda\kappa_d r^d\right)$. As a consequence:

$$\mathbb{E}\left(\psi\left(B\right)\right) = \lambda\exp\left(-\lambda\kappa_d r^d\right)\nu_d(B).$$

Because η is stationary, so is ψ, and ψ has intensity $\lambda \exp\left(-\lambda \kappa_d r^d\right)$.

A Matérn I hard-core process is called a dependent thinning because the probability that a given point of the original Poisson process is removed, depends on the locations of other points in the point process. For independent thinnings of point processes, the probability that a point is removed does not depend on other points in the point process. An independent thinning of a Poisson process is again a Poisson process, but with a different intensity measure than the original Poisson process.

A.2 Marked point processes

An important class of point processes are the so-called marked point processes. Effectively, given a point process one may wish to assign a mark or a label to each point. For example, a point process on $\mathbb{R}^3$ may represent the locations of the centers of the balls in the Wicksell corpuscle problem. We may then assign to each point a random radius, thereby completely describing a random spatial system of balls. In this example, these radii live in the space $(0, \infty)$ which is referred to as the mark space. Formally, a marked point process is defined as follows:

Definition A.9 (Marked point process)**.** A marked point process in $\mathbb{R}^d$ with mark space $\mathbb{M}$ is a simple point process η on $\mathbb{R}^d \times \mathbb{M}$ with intensity measure Λ satisfying:

$$\Lambda(C \times \mathbb{M}) < \infty, \quad \text{for all compact } C \in \mathcal{B}(\mathbb{R}^d).$$

Suppose that $\eta = \{x_1, x_2, \dots\}$ is a homogeneous Poisson process on $\mathbb{R}^d$ with intensity $\lambda > 0$. Let Q be a probability measure on some mark space $\mathbb{M}$ and draw $m_1, m_2, \dots \overset{\text{iid}}{\sim} Q$ independently of η. Then, define the marked point process $\eta_{\mathbb{M}} = \{(x_1, m_1), (x_2, m_2), \dots\}$. More specifically, $\eta_{\mathbb{M}}$ is a so-called independently marked point process as the marks are independent of the point locations. In fact, $\eta_{\mathbb{M}}$ is again a Poisson process on the space $\mathbb{R}^d \times \mathbb{M}$ with intensity measure $\lambda \nu_d \times Q$. As it turns out, this intensity measure is the general form for all stationary marked point processes (even if the marking is not independent):

Theorem A.10 (Theorem 3.5.1. in [86])**.** *Let η be a stationary marked point process on $\mathbb{R}^d$ with mark space $\mathbb{M}$ and intensity measure $\Lambda \neq 0$, then:*

$$\Lambda = \lambda \nu_d \times Q,$$

with $0 < \lambda < \infty$ and Q a probability measure on $\mathbb{M}$.

That being said, we have not yet defined what stationarity means in the context of marked point processes. The notion of stationarity can be adapted to the marked context by only considering shifts with respect to the spatial coordinate. That is, for $v \in \mathbb{R}^d$ let S_v again denote the shift operator. Suppose $\eta = \{(x_1, m_1), (x_2, m_2), \dots\}$ is a marked point process with $x_i \in \mathbb{R}^d$ and $m_i \in \mathbb{M}$. Then, we define $S_v \eta := \{(x_1 - v, m_1), (x_2 - v, m_2), \dots\}$. The marked point process η is called stationary if η and $S_v \eta$ are equal in distribution for all $v \in \mathbb{R}^d$.

The distribution Q in Theorem A.10 is also referred to as the mark distribution, and $\lambda > 0$ represents the intensity. Let η and Λ be as in Theorem A.10, let $B \in \mathcal{B}(\mathbb{R}^d)$ with $\nu_d(B) > 0$ and $A \in \mathcal{B}(\mathbb{M})$. Then, a direct application of Campbell's formula yields the following:

$$Q(A) = \frac{1}{\lambda \nu_d(B)} \mathbb{E}\left(\sum_{(x,m)\in\eta} \mathbb{1}_B(x)\mathbb{1}_A(m) \right). \tag{A.1}$$

Because the behavior of a stationary marked point process is the same in every region of $\mathbb{R}^d$, this expression for $Q(A)$ does not depend on the choice of the region B. Drawing from Q corresponds to picking the mark of a point uniformly at random from all points in the region B. Because B can be made as large as possible, loosely speaking Q describes the distribution of a random mark, when picking a marked point uniformly at random from all points. Mark distributions of the form (A.1) are examples of so-called Palm distributions.

A.3 Particle processes and random tessellations

While we may intuitively tend to think of point processes on $\mathbb{X} = \mathbb{R}^d$, in stochastic geometry, many interesting models are described as point processes on other spaces. Let $C'(\mathbb{R}^d)$ denote the system of non-empty compact sets in $\mathbb{R}^d$. One may for instance equip $C'(\mathbb{R}^d)$ with the Hausdorff metric to obtain a complete separable metric space.

Definition A.11 (Particle process)**.** A particle process is a point process on $\mathbb{X} = C'(\mathbb{R}^d)$.

In a particle process $X = \{K_1, K_2, \dots\}$, each "point" K_i in X is a non-empty compact set, which we call a particle instead of a point. A particle process $X = \{K_1, K_2, \dots\}$ is called stationary if X is equal in distribution to $S_v X = \{K_1 - v, K_2 - v, \dots\}$ for all $v \in \mathbb{R}^d$.

Suppose η is an independently marked homogeneous Poisson process on $\mathbb{R}^3$ with mark space $(0, \infty)$. We may write $\eta = \{(x_1, r_1), (x_2, r_2), \dots\}$ with $x_i \in \mathbb{R}^3$ and $r_i > 0$. A particle process describing a random spatial system of balls may then be defined via $X = \{\bar{B}(x_1, r_1), \bar{B}(x_2, r_2), \dots\}$. This specific type of particle process, which is essentially defined by first taking a point process $\{x_1, x_2, \dots\}$ on $\mathbb{R}^d$, and then inserting particles centered at the x_i's is also referred to as a germ-grain process.

For other types of particle processes it is often still useful to assign a unique center to each particle, this may for example be the center of mass of each particle. Let $c(K)$ denote the center of $K \in C'(\mathbb{R}^d)$. If $X = \{K_1, K_2, \dots\}$ is a particle process of interest, then it is sometimes more convenient to instead study the marked point process $\eta = \{(c(K), K - c(K)) : K \in X\}$, which is a marked point process on $\mathbb{R}^d$ with mark space $C'(\mathbb{R}^d)$. If X is a stationary particle process, then η is also stationary and the mark distribution of η is referred to as the grain distribution, or the distribution of the typical grain. Because the grain distribution is a probability measure on $C'(\mathbb{R}^d)$, it is the distribution of a random closed set, which is concentrated on $C'(\mathbb{R}^d)$. Note that the distribution of the typical grain depends on the choice of center function. Depending on the context, it may be the case that one specific choice of the center function is the most natural. A type of particle process which deserves its own name is the class of random tessellations or random mosaics. Tessellations and random tessellations are respectively defined as follows:

Definition A.12 (Tessellation)**.** A tessellation of $\mathbb{R}^d$ is a countable collection $T = \{C_i : i \in \mathbb{N}\}$, of sets $C_i \subset \mathbb{R}^d$ (the cells of the tessellation) such that:

- $\text{int}(C_i) \cap \text{int}(C_j) = \emptyset$, if $i \neq j$.
- $\cup_{i\in\mathbb{N}} C_i = \mathbb{R}^d$.
- T is locally finite: $\#\{i \in \mathbb{N} : C_i \cap B \neq \emptyset\} < \infty$ for all bounded $B \in \mathcal{B}(\mathbb{R}^d)$.
- Each C_i is a compact and convex set with interior points.

Definition A.13 (Random tessellation)**.** A random tessellation is a particle process which is with probability one a tessellation.

The "points" of a random tessellation are not called particles, but cells. A classical random tessellation model is the Poisson-Voronoi tessellation, which is constructed as follows. Suppose η is a homogeneous Poisson process on $\mathbb{R}^d$ with intensity $\lambda > 0$. Then, the Voronoi cell associated with $x \in \eta$ is a random polytope, which is given by:

$$C(x, \eta) = \{y \in \mathbb{R}^d : \|x - y\| \leq \|x' - y\| \text{ for all } x' \in \eta\}.$$

The collection $V(\eta) = \{C(x, \eta) : x \in \eta\}$ of all cells is a random tessellation, called the Poisson-Voronoi tessellation. In this context, η may also be called the point process of generator points. There is a significant research interest in studying properties of the so-called typical cell of the Poisson-Voronoi tessellation. Viewing a random tessellation as a particle process, the distribution of the typical grain is instead called the distribution of the typical cell. Intuitively, the typical Poisson-Voronoi cell may be thought of as the distribution of the cell which is obtained by picking a cell at random from the Poisson-Voronoi tessellation and centering it at the origin. This can be made precise via a Palm distribution. Formally, one may associate $V(\eta)$ with the marked point process:

$$\psi = \{(x, C(x, \eta) - x) : x \in \eta\}. \tag{A.2}$$

Indeed, in this context it is natural to assign the center x to the Voronoi cell $C(x, \eta)$. Given an individual Voronoi cell its is in general not possible to determine its generator x, meaning that the centering is not obtained via a function $c : C'(\mathbb{R}^d) \to \mathbb{R}^d$. In this case, the underlying center function is an example of a so-called generalized center function. The point process ψ is a stationary marked point process on $\mathbb{R}^d$ with mark space $C'(\mathbb{R}^d)$ and intensity λ. The distribution of the typical Poisson-Voronoi cell is defined as the mark distribution of ψ. Because the shape of a Voronoi cell $C(x, \eta)$ not only depends on x but also depends on other points in η, ψ is not an independently marked point process. A well-known result is the following stochastic representation of the typical cell:

Theorem A.14. *Let Q denote the distribution of the typical Poisson-Voronoi cell, corresponding to a homogeneous Poisson process η on $\mathbb{R}^d$. Then, $C(0, \eta) \sim Q$.*

That is, the Voronoi cell with 0 as its generator is a random closed set which is equal in distribution to the typical Poisson-Voronoi cell. This result can be proven via the Mecke equation, as shown below.

Proof of Theorem A.14. Let $A \in \mathcal{B}(C'(\mathbb{R}^d))$ and $B \in \mathcal{B}(\mathbb{R}^d)$ with $\nu_d(B) > 0$. Let $\lambda > 0$ denote the intensity of η. From the definition of the Voronoi cell it can be seen that for $x \in \mathbb{R}^d$: $C(x, \eta) = C(x, \eta + \delta_x)$. From the definition of the shift operator one can see: $C(x, \eta) - x = C(0, S_x\eta)$. Let $\psi = \{(x_1, K_1), (x_2, K_2), \dots\}$ be as in (A.2), applying the Mecke equation to (A.1) yields:

$$
\begin{aligned}
Q(A) &= \frac{1}{\lambda \nu_d(B)} \mathbb{E}\left(\sum_{(x,K)\in\psi} \mathbb{1}_B(x) \mathbb{1}_A(K) \right) \\
&= \frac{1}{\lambda \nu_d(B)} \mathbb{E}\left(\sum_{x\in\eta} \mathbb{1}_B(x) \mathbb{1}_A(C(x,\eta) - x) \right) \\
&= \frac{1}{\lambda \nu_d(B)} \int_B \mathbb{P}\left(C\left(x, \eta + \delta_x\right) - x \in A\right) \lambda \mathrm{d}x \qquad \text{(Mecke equation)} \\
&= \frac{1}{\lambda \nu_d(B)} \int_B \mathbb{P}\left(C\left(0, S_x\eta\right) \in A\right) \lambda \mathrm{d}x \\
&= \frac{1}{\lambda \nu_d(B)} \int_B \mathbb{P}\left(C\left(0, \eta\right) \in A\right) \lambda \mathrm{d}x \qquad \text{(stationarity of } \eta) \\
&= \mathbb{P}\left(C\left(0, \eta\right) \in A\right).
\end{aligned}
$$

□

Appendix B

The stereological integral equation

In this appendix we present an alternative way of deriving the stereological integral equation (3.3) in Chapter 3. Here, we describe the random system of particles via a stationary particle process, which we intersect with a fixed (deterministic) plane. For an introduction to the terminology used in this appendix we refer to Appendix A. For this derivation we also occasionally require some results obtained in Chapter 2.

First, we introduce some notation such that we can describe the particle process of interest. Let ρ denote the unique rotation invariant probability measure on $\mathrm{SO}(3)$, also known as the Haar measure. Suppose that $\mathbb{H}$ is a probability measure on $(0, \infty)$, and for $\lambda \geq 0$ we denote its CDF by $H(\lambda) := \mathbb{H}((0, \lambda])$. As in Chapter 3 we assume that $\mathbb{H}$ has a finite first moment:

$$\mathbb{E}(\Lambda) = \int_0^\infty \lambda \mathrm{d}H(\lambda) < \infty.$$

Due to technical reasons, we additionally need to assume that $\mathbb{E}(\Lambda^3) < \infty$. That is, by Theorem 4.1.2. in [86], the particle process we are about to define is only well-defined under this assumption as it ensures local finiteness of the intensity measure of the particle process.

We now fix a common shape for the particles by taking a fixed convex body $K \in \mathcal{K}^3$. Draw $M \sim \rho$, and independently draw $\Lambda \sim \mathbb{H}$. Let Q be the probability measure on $\mathcal{K}^3$ induced by the random closed set $C := \Lambda M K$. In words, by drawing from Q, we obtain a randomly rotated and randomly scaled version of K. Let η be a stationary marked point process on $\mathbb{R}^3$ with mark space $\mathcal{K}^3$, mark distribution Q and intensity $N_V > 0$. Then, we may write $\eta = \{(x_1, C_1), (x_2, C_2), \dots\}$ with $x_i \in \mathbb{R}^3$ and $C_i \in \mathcal{K}^3$. Hence, the intensity measure of η is given by $N_V \nu_3 \times Q$.

This means that the particle process $X = \{x_1 + C_1, x_2 + C_2, \dots\}$ describes the random spatial system of particles. For $x \in \mathbb{R}^3$, we write $x = (\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3)$. We now intersect X with the plane $T = \{(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3) \in \mathbb{R}^3 : \mathrm{x}_3 = 0\}$ and study the areas of the particles of X which are hit by T. Let η_T be the marked point process which is obtained as follows. For any $(x, C) \in \eta$ with $(C + x) \cap T \neq \emptyset$ we project x onto T, which we denote by x_T, and as a mark corresponding to this projected point we take the area $A = \nu_2((C + x) \cap T)$. We identify T with $\mathbb{R}^2$ such that $x_T = (\mathrm{x}_1, \mathrm{x}_2)$ and η_T is a marked point process on $\mathbb{R}^2$ with

marks in $(0, \infty)$. Because η is stationary, so is η_T. Note that we may write:

$$\eta_T = \{(x_T, \nu_2((C + x) \cap T)) : (x, C) \in \eta, (C + x) \cap T \neq \emptyset\}.$$

Let F_A denote the CDF of the mark distribution of η_T and let N_A denote the intensity of η_T. Via the Campbell formula we obtain for $a \geq 0$:

$$\begin{aligned}
N_A F_A(a) &= \mathbb{E}\left(\sum_{(x_T, A) \in \eta_T} \mathbb{1}_{[0,1]^2}(x_T) \mathbb{1}_{(0,a]}(A) \right) \\
&= \mathbb{E}\left(\sum_{(x,C) \in \eta} \mathbb{1}_{[0,1]^2}((\mathrm{x}_1, \mathrm{x}_2)) \mathbb{1}\{\nu_2((C + x) \cap T) \leq a\} \mathbb{1}\{(C + x) \cap T \neq \emptyset\} \right) \\
&= N_V \int_0^\infty \int_{\mathrm{SO}(3)} \int_{\mathbb{R}^3} \mathbb{1}_{[0,1]^2}((\mathrm{x}_1, \mathrm{x}_2)) \mathbb{1}\{\nu_2((\lambda m K + x) \cap T) \leq a\} \\
&\quad \cdot \mathbb{1}\{(\lambda m K + x) \cap T \neq \emptyset\} \mathrm{d}x \rho(\mathrm{d}m) \mathrm{d}H(\lambda).
\end{aligned}$$

Recall the notation for a hyperplane $T_{\theta,s} = \{x \in \mathbb{R}^d : \langle \theta, x \rangle = s\}$ for $\theta \in \mathbb{S}^{d-1}$, $s \in \mathbb{R}$. Then, for the hyperplane T we have $T = T_{\theta,0}$ with $\theta = (0, 0, 1)$. By (2.20) and (2.26) we have for any $x \in \mathbb{R}^3$, $m \in \mathrm{SO}(3)$ and $\lambda > 0$:

$$(\lambda m K + x) \cap T_{\theta,0} = (\lambda m K \cap T_{\theta,-\mathrm{x}_3}) + x = \left(m \left(\lambda K \cap T_{m^T \theta,-\mathrm{x}_3}\right)\right) + x.$$

Using the translation- and rotation invariance of Lebesgue measure we obtain:

$$\begin{aligned}
N_A F_A(a) &= N_V \int_0^\infty \int_{\mathrm{SO}(3)} \int_{\mathbb{R}^3} \mathbb{1}_{[0,1]^2}((\mathrm{x}_1, \mathrm{x}_2)) \mathbb{1}\left\{\nu_2\left(\lambda K \cap T_{m^T \theta,-\mathrm{x}_3}\right) \leq a\right\} \\
&\quad \cdot \mathbb{1}\{\lambda K \cap T_{m^T \theta,-\mathrm{x}_3} \neq \emptyset\} \mathrm{d}x \rho(\mathrm{d}m) \mathrm{d}H(\lambda) \\
&= N_V \int_0^\infty \int_{\mathrm{SO}(3)} \int_{\mathbb{R}} \mathbb{1}\left\{\nu_2\left(\lambda K \cap T_{m^T \theta,z}\right) \leq a\right\} \mathbb{1}\{\lambda K \cap T_{m^T \theta,z} \neq \emptyset\} \\
&\quad \mathrm{d}z \rho(\mathrm{d}m) \mathrm{d}H(\lambda).
\end{aligned}$$

In the final equality we substituted $z = -\mathrm{x}_3$. Now, we need that fact that if $M \sim \rho$ then also $M^T \sim \rho$. Additionally, if $M \sim \rho$, then for any fixed $\theta \in \mathbb{S}^2$: $M\theta \sim \mathrm{Unif}(\mathbb{S}^2)$. For our purposes, this means that we have for all bounded integrable functions $g : \mathbb{S}^2 \to \mathbb{R}$:

$$\int_{\mathrm{SO}(3)} g(m\theta) \rho(\mathrm{d}m) = \frac{1}{\sigma_2(\mathbb{S}^2)} \int_{\mathbb{S}^2} g(u) \sigma_2(\mathrm{d}u).$$

Hence, setting $u = m^T \theta$ we obtain:

$$\begin{aligned}
N_A F_A(a) &= \frac{N_V}{4\pi} \int_0^\infty \int_{\mathbb{S}^2} \int_{\mathbb{R}} \mathbb{1}\left\{\nu_2\left(\lambda K \cap T_{u,z}\right) \leq a\right\} \mathbb{1}\{\lambda K \cap T_{u,z} \neq \emptyset\} \mathrm{d}z \mathrm{d}u \mathrm{d}H(\lambda) \\
&= \frac{N_V}{2\pi} \int_0^\infty \left(\int_{\mathbb{S}^2_+} \int_{\mathbb{R}} \mathbb{1}\left\{\nu_2\left(\lambda K \cap T_{u,z}\right) \leq a\right\} \mathbb{1}\{\lambda K \cap T_{u,z} \neq \emptyset\} \mathrm{d}z \mathrm{d}u \right) \mathrm{d}H(\lambda).
\end{aligned} \tag{B.1}$$

The double integral in brackets in equation (B.1) may be recognized to be equal to $G_{\lambda K}(a)\mu([\lambda K])$ (see also Section 2.2). Recall that $\mu([\lambda K]) = 2\pi\bar{b}(\lambda K) = 2\pi\lambda\bar{b}(K)$, such that:

$$N_A F_A(a) = N_V \bar{b}(K) \int_0^\infty G_{\lambda K}(a)\lambda \mathrm{d}H(\lambda). \tag{B.2}$$

Letting $a \to \infty$, and applying the dominated convergence theorem, we obtain the well-known stereological equation:

$$N_A = N_V \bar{\bar{b}} := N_V \bar{b}(K) \int_0^\infty \lambda \mathrm{d}H(\lambda) = N_V \bar{b}(K)\mathbb{E}(\Lambda). \tag{B.3}$$

Combining (B.2) and (B.3) yields the desired expression:

$$F_A(a) = \frac{1}{\mathbb{E}(\Lambda)} \int_0^\infty G_{\lambda K}(a)\lambda \mathrm{d}H(\lambda) = \frac{1}{\mathbb{E}(\Lambda)} \int_0^\infty G_K\left(\frac{a}{\lambda^2}\right)\lambda \mathrm{d}H(\lambda).$$

The final equality follows from point 2 of Lemma 2.2.

If one considers the point process of particle centers marked with the particle sizes restricted to the subset of particles which are hit by the section plane, then deriving the intensity measure of this point process will show that the size distribution of particles which are hit by the section plane is given by the length-biased version of H. This computation involves similar techniques as those used above. Due to the properties of the Poisson process, if η is an independently marked Poisson process, then so is η_T.

Bibliography

[1] N. Allain-Bonasso, F. Wagner, S. Berbenni, and D. P. Field. A study of the heterogeneity of plastic deformation in IF steel by EBSD. *Materials Science and Engineering: A*, 548:56–63, 2012.

[2] A. Alpers, A. Brieden, P. Gritzmann, A. Lyckegaard, and H. F. Poulsen. Generalized balanced power diagrams for 3D representations of polycrystals. *Philosophical Magazine*, 95(9):1016–1028, 2015.

[3] A. Alpers, O. Furat, C. Jung, M. Neumann, C. Redenbach, A. Saken, and V. Schmidt. Comparative analysis of algorithms for the fitting of tessellations to 3D image data. *Arxiv preprint arXiv:2507.14268*, 2025.

[4] R. Armstrong. The influence of polycrystal grain size on several mechanical properties of materials. *Metallurgical and Materials Transactions*, 1(5):1169 – 1176, 1970.

[5] R. Arratia, L. Goldstein, and F. Kochman. Size bias for one and all. *Probability Surveys*, 16:1–61, 2019.

[6] F. Aurenhammer. Power diagrams: Properties, algorithms and applications. *SIAM Journal on Computing*, 16:78–96, 1987.

[7] F. Baccelli and B. Błaszczyszyn. *Stochastic Geometry and Wireless Networks: Volume I Theory*. Now Publishers, 2009.

[8] F. Baccelli, B. Błaszczyszyn, and M. Karray. *Random Measures, Point Processes, and Stochastic Geometry*. Inria, 2024. hal-02460214v2.

[9] A. Baddeley and E. B. V. Jensen. *Stereology for Statisticians*. Chapman and Hall/CRC, 2004.

[10] F. Barlat. Crystallographic texture, anisotropic yield surfaces and forming limits of sheet metals. *Materials Science and Engineering*, 91:55–72, 1987.

[11] V. Beneš and J. Rataj. *Stochastic Geometry: Selected Topics*. Kluwer Academic Publishers, 2004.

[12] S. Benito, C. Cuervo, F. Pöhl, and W. Theisen. Improvements on the recovery of 3D particle size distributions from 2D sections. *Materials Characterization*, 156:109872, 2019.

[13] C. Bhattacharjee and A. Gusakova. Gaussian approximation for extreme points in Laguerre tessellations. *Arxiv preprint arXiv:2510.21665*, 2025.

[14] N. Bissantz and H. Holzmann. Statistical inference for inverse problems. *Inverse Problems*, 24:034009, 2008.

[15] D. P. Bourne, P. J. J. Kok, S. M. Roper, and W. D. T. Spanjer. Laguerre tessellations and polycrystalline microstructures: a fast algorithm for generating grains of given volumes. *Philosophical Magazine*, 100, 2020.

[16] P. L. Butzer and S. Jansche. A Direct Approach to the Mellin Transform. *Journal of Fourier Analysis and Applications*, 3:325–376, 1997.

[17] L. Cavalier. Nonparametric statistical inverse problems. *Inverse Problems*, 24:034004, 2008.

[18] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. John Wiley & Sons, Ltd, 2013.

[19] R. Coleman. Random paths through convex bodies. *Journal of Applied Probability*, 6:430–441, 1969.

[20] L. M. Cruz-Orive. *Stereology*, volume 59. Springer Nature Switzerland, 2024.

[21] J. N. Cuzzi and D. M. Olson. Recovering 3D particle size distributions from 2D sections. *Meteoritics & Planetary Science*, 52:532–545, 2017.

[22] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume I*. Springer-Verlag, 2003.

[23] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume II*. Springer New York, 2008.

[24] M. Dao and M. Li. A micromechanics study on strain-localization-induced fracture initiation in bending using crystal plasticity models. *Philosophical Magazine A*, 81(8):1997–2020, 2001.

[25] P. Davy and R. E. Miles. Sampling Theory for Opaque Spatial Specimens. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39:56–65, 1977.

[26] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39:1–22, 1977.

[27] Q. Duan, D. P. Kroese, T. Brereton, A. Spettl, and V. Schmidt. Inverting Laguerre Tessellations. *The Computer Journal*, 57, 2014.

[28] S. Falco, J. Jiang, F. D. Cola, and N. Petrinic. Generation of 3D polycrystalline microstructures with a conditioned Laguerre-Voronoi tessellation technique. *Computational Materials Science*, 136:20–28, 2017.

[29] 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.

[30] D. Flimmel, Z. Pawlas, and J. E. Yukich. Limit theory for unbiased and consistent estimators of statistics of random tessellations. *Journal of Applied Probability*, 57:679–702, 2020.

[31] R. J. Gardner. The Brunn-Minkowski inequality. *Bulletin of the American Mathematical Society*, 39:355–406, 2002.

[32] J. Gates. Some properties of chord length distributions. *Journal of Applied Probability*, 24:863–874, 1987.

[33] F. Gili, G. Jongbloed, and A. W. van der Vaart. Adaptive and efficient isotonic estimation in Wicksell's problem. *Journal of Nonparametric Statistics*, 37:491–531, 2025.

[34] F. Gili, G. Jongbloed, and A. W. van der Vaart. Semiparametric Bernstein-von Mises Phenomenon via Isotonized Posterior in Wicksell's problem. *To appear in The Annals of Statistics*, 2025+.

[35] P. Groeneboom and G. Jongbloed. Isotonic Estimation and Rates of Convergence in Wicksell's Problem. *The Annals of Statistics*, 23:1518–1542, 1995.

[36] P. Groeneboom and G. Jongbloed. *Nonparametric Estimation under Shape Constraints*. Cambridge University Press, 2014.

[37] P. Groeneboom, G. Jongbloed, and S. Michael. Consistency of maximum likelihood estimators in a large class of deconvolution models. *Canadian Journal of Statistics*, 41:98–110, 2013.

[38] P. Groeneboom and J. A. Wellner. *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Birkhäuser Basel, 1992.

[39] A. Gusakova and M. in Wolde-Lübke. Poisson-Laguerre tessellations. *Electronic Journal of Probability*, 30:1–48, 2025. article no. 56.

[40] A. Gusakova, Z. Kabluchko, and C. Thäle. The β-Delaunay tessellation: Description of the model and geometry of typical cells. *Advances in Applied Probability*, 54:1252–1290, 2022.

[41] A. Gusakova, Z. Kabluchko, and C. Thäle. Sectional Voronoi tessellations: Characterization and high-dimensional limits. *Bernoulli*, 30:1482–1501, 2024.

[42] U. Hahn and U. Lorz. On the precision of some stereological estimators for the model parameter of the spatial Poisson-Voronoi tessellation. *Acta Stereologica*, 13:245–250, 1994.

[43] N. Hansen. Hall–Petch relation and boundary strengthening. *Scripta Materialia*, 51(8):801–806, 2004. Viewpoint set no. 35. Metals and alloys with a structural scale from the micrometer to the atomic dimensions.

[44] H. S. Harutyunyan and V. K. Ohanyan. The chord length distribution function for regular polygons. *Advances in Applied Probability*, 41:358–366, 2009.

[45] G. Jongbloed. The Iterative Convex Minorant Algorithm for Nonparametric Estimation. *Journal of Computational and Graphical Statistics*, 7:310, 1998.

[46] G. Jongbloed. Sieved Maximum Likelihood Estimation in Wicksell's Problem and Related Deconvolution Problems. *Scandinavian Journal of Statistics*, 28:161–183, 2001.

[47] M. Kasemer and P. Dawson. A finite element methodology to incorporate kinematic activation of discrete deformation twins in a crystal plasticity framework. *Computer Methods in Applied Mechanics and Engineering*, 358:112653, 2020.

[48] T. Kawata. *Fourier Analysis in Probability Theory*. Academic Press, 1972.

[49] M. S. Khorrami, J. R. Mianroodi, N. H. Siboni, P. Goyal, B. Svendsen, P. Benner, and D. Raabe. An artificial neural network for surrogate modeling of stress fields in viscoplastic polycrystalline materials. *npj Computational Materials*, 9(1):37, 2023.

[50] D.-K. Kim, W. Woo, W.-W. Park, Y.-T. Im, and A. Rollett. Mesoscopic coupled modeling of texture formation during recrystallization considering stored energy decomposition. *Computational Materials Science*, 129:55–65, 2017.

[51] J. Kisel'ák and G. Baluchová. On Particle-Size Distribution of Convex Similar Bodies in $\mathbb{R}^3$. *Journal of Mathematical Imaging and Vision*, 63:108–119, 2021.

[52] A. Koldobsky. *Fourier Analysis in Convex Geometry*. American Mathematical Society, 2005.

[53] G. Last and M. Penrose. *Lectures on the Poisson Process*. Cambridge University Press, 2017.

[54] C. Lautensack. *Random Laguerre tessellations*. PhD thesis, Universität Karlsruhe, 2007.

[55] C. Lautensack. Fitting three-dimensional Laguerre tessellations to foam structures. *Journal of Applied Statistics*, 35:985–995, 2008.

[56] C. Lautensack and S. Zuyev. Random Laguerre tessellations. *Advances in Applied Probability*, 40:630–650, 2008.

[57] A. Liebscher. Laguerre approximation of random foams. *Philosophical Magazine*, 95:2777–2792, 2015.

[58] M. A. Lopez-Sanchez and S. Llana-Fúnez. An extension of the Saltykov method to quantify 3D grain size distributions in mylonites. *Journal of Structural Geology*, 93, 2016.

[59] R. McAfee and I. Nettleship. The simulation and selection of shapes for the unfolding of grain size distributions. *Acta Materialia*, 51:4603–4610, 2003.

[60] K. S. McGarrity, J. Sietsma, and G. Jongbloed. Nonparametric inference in a stereological model with oriented cylinders applied to dual phase steel. *The Annals of Applied Statistics*, 8, 2014.

[61] K. Mehnert, J. Ohser, and P. Klimanek. Testing stereological methods for the estimation of spatial size distributions by means of computer-simulated grain structures. *Materials Science and Engineering: A*, 246:207–212, 1998.

[62] J. Meijering. Interface area, edge length, and number of vertices in crystal aggregates with random nucleation. *Philips Research Reports*, 8:270–290, 1953.

[63] J. Meyron. Initialization procedures for discrete and semi-discrete optimal transport. *Computer-Aided Design*, 115:13–22, 2019.

[64] J. R. Mianroodi, N. H. Siboni, and D. Raabe. Teaching solid mechanics to artificial intelligence—a fast solver for heterogeneous materials. *npj Computational Materials*, 7(1):99, 2021.

[65] R. E. Miles. The importance of proper model specification in stereology. In R. E. Miles and J. Serra, editors, *Geometrical Probability and Biological Structures: Buffon's 200th Anniversary*, pages 115–136. Springer Berlin Heidelberg, 1978.

[66] I. Molchanov. *Theory of Random Sets*. Springer London, 2017.

[67] J. Møller. *Lectures on Random Voronoi Tessellations*, volume 87. Springer New York, 1994.

[68] J. Ohser and F. Mücklich. Stereology for some classes of polyhedrons. *Advances in Applied Probability*, 27:384–396, 1995.

[69] J. Ohser and F. Mücklich. *Statistical Analysis of Microstructures in Materials Science*. John Wiley & Sons, Inc., 2000.

[70] J. Ohser and M. Nippe. Stereology of cubic particles: various estimators for the size distribution. *Journal of Microscopy*, 187:22–30, 1997.

[71] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, Inc., 2000.

[72] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, 2009.

[73] V. Olsbo. On the correlation between the volumes of the typical Poisson-Voronoi cell and the typical Stienen sphere. *Advances in Applied Probability*, 39:883–892, 2007.

[74] B. Pachpatte. *Inequalities for Differential and Integral Equations*. Mathematics in Science and Technology. Academic Press, 1998.

[75] J. L. Paul. Distribution curves of sectional areas through some families of convex particles. *Journal of Microscopy*, 122:165–172, 1981.

[76] H. Pirgazi. On the alignment of 3D EBSD data collected by serial sectioning technique. *Materials Characterization*, 152:223–229, 2019.

[77] D. Raabe, B. Sun, A. Kwiatkowski Da Silva, B. Gault, H.-W. Yen, K. Sedighiani, P. Thoudden Sukumar, I. R. Souza Filho, S. Katnagallu, E. Jägle, P. Kürnsteiner, N. Kusampudi, L. Stephenson, M. Herbig, C. H. Liebscher, H. Springer, S. Zaefferer, V. Shah, S.-L. Wong, C. Baron, M. Diehl, F. Roters, and D. Ponge. Current Challenges and Opportunities in Microstructure-Related Properties of Advanced High-Strength Steels. *Metallurgical and Materials Transactions A*, 51(11):5517–5586, 2020.

[78] C. Redenbach and C. Jung. Random tessellations: An overview of models. In H. Biermé, editor, *Stochastic Geometry: Percolation, Tesselations, Gaussian Fields and Point Processes*, pages 35–80. Springer Nature Switzerland, 2025.

[79] T. Robertson, F. T. Wright, and R. L. Dykstra. *Order Restricted Statistical Inference*. Wiley, 1988.

[80] F. Roters, M. Diehl, P. Shanthraj, P. Eisenlohr, C. Reuber, S. Wong, T. Maiti, A. Ebrahimi, T. Hochrainer, H.-O. Fabritius, S. Nikolov, M. Friák, N. Fujita, N. Grilli, K. Janssens, N. Jia, P. Kok, D. Ma, F. Meier, E. Werner, M. Stricker, D. Weygand, and D. Raabe. DAMASK – The Düsseldorf Advanced Material Simulation Kit for modeling multi-physics crystal plasticity, thermal, and damage phenomena from the single crystal up to the component scale. *Computational Materials Science*, 158:420–478, 2019.

[81] F. Roters, P. Eisenlohr, L. Hantcherli, D. Tjahjanto, T. Bieler, and D. Raabe. Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: Theory, experiments, applications. *Acta Materialia*, 58(4):1152–1211, 2010.

[82] C. H. Rycroft. VORO++: A three-dimensional Voronoi cell library in C++. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19:041111, 2009.

[83] S. A. Saltikov. The determination of the size distribution of particles in an opaque material from a measurement of the size distribution of their sections. In H. Elias, editor, *Stereology*, pages 163–173. Springer Berlin Heidelberg, 1967.

[84] L. A. Santaló and M. Kac. *Integral Geometry and Geometric Probability*. Cambridge University Press, 2004.

[85] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, 2013.

[86] R. Schneider and W. Weil. *Stochastic and Integral Geometry*. Springer Berlin Heidelberg, 2008.

[87] F. P. Schoenberg. Inverting Dirichlet tessellations. *The Computer Journal*, 46:76–83, 2003.

[88] T. Schreiber. Limit theorems in stochastic geometry. In W. S. Kendall and I. Molchanov, editors, *New perspectives in Stochastic Geometry*, pages 111–144. Oxford University Press, 2010.

[89] E. F. Schuster. Incorporating support constraints into nonparametric estimators of densities. *Communications in Statistics - Theory and Methods*, 14:1123–1136, 1985.

[90] K. Sedighiani, M. Diehl, K. Traka, F. Roters, J. Sietsma, and D. Raabe. An efficient and robust approach to determine material parameters of crystal plasticity constitutive laws from macro-scale stress–strain curves. *International Journal of Plasticity*, 134:102779, 2020.

[91] K. Sedighiani, V. Shah, K. Traka, M. Diehl, F. Roters, J. Sietsma, and D. Raabe. Large-deformation crystal plasticity simulation of microstructure and microtexture evolution through adaptive remeshing. *International Journal of Plasticity*, 146:103078, 2021.

[92] K. Sedighiani, K. Traka, F. Roters, D. Raabe, J. Sietsma, and M. Diehl. Determination and analysis of the constitutive parameters of temperature-dependent dislocation-density-based crystal plasticity models. *Mechanics of Materials*, 164:104117, 2022.

[93] K. Sedighiani, K. Traka, F. Roters, J. Sietsma, D. Raabe, and M. Diehl. Crystal plasticity simulation of in-grain microstructural evolution during large deformation of IF-steel. *Acta Materialia*, 237:118167, 2022.

[94] F. Seitl, J. Møller, and V. Beneš. Fitting three-dimensional Laguerre tessellations by hierarchical marked point process models. *Spatial Statistics*, 51:100658, 2022.

[95] F. Seitl, L. Petrich, J. Staněk, C. E. Krill, V. Schmidt, and V. Beneš. Exploration of Gibbs-Laguerre Tessellations for Three-Dimensional Stochastic Modeling. *Methodology and Computing in Applied Probability*, 23:669–693, 2021.

[96] V. Shah, K. Sedighiani, J. Van Dokkum, C. Bos, F. Roters, and M. Diehl. Coupling crystal plasticity and cellular automaton models to study meta-dynamic recrystallization during hot rolling at high strain rates. *Materials Science and Engineering: A*, 849:143471, 2022.

[97] S. J. Sheather and M. C. Jones. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53:683–690, 1991.

[98] K. Somlo, B. Frodal, C. Funch, K. Poulios, G. Winther, O. Hopperstad, T. Børvik, and C. Niordson. Anisotropic yield surfaces of additively manufactured metals simulated with crystal plasticity. *European Journal of Mechanics - A/Solids*, 94:104506, 2022.

[99] R. P. Srivastav. A Note on Certain Integral Equations of Abel-Type. *Proceedings of the Edinburgh Mathematical Society*, 13:271–272, 1963.

[100] D. Sterio. The unbiased estimation of number and sizes of arbitrary particles using the disector. *Journal of microscopy*, 134(2):127–136, 1984.

[101] D. Stoyan, V. Beneš, and F. Seitl. Dependent radius marks of Laguerre tessellations: a case study. *Australian & New Zealand Journal of Statistics*, 63:19–32, 2021.

[102] C. Tasan, J. Hoefnagels, M. Diehl, D. Yan, F. Roters, and D. Raabe. Strain localization and damage in dual phase steels investigated by coupled in-situ deformation experiments and crystal plasticity simulations. *International Journal of Plasticity*, 63:198–210, 2014. Deformation Tensors in Material Modeling in Honor of Prof. Otto T. Bruhns.

[103] K. Traka, K. Sedighiani, C. Bos, J. Galan Lopez, K. Angenendt, D. Raabe, and J. Sietsma. Topological aspects responsible for recrystallization evolution in an IF-steel sheet – Investigation with cellular-automaton simulations. *Computational Materials Science*, 198:110643, 2021.

[104] R. van de Weygaert. Fragmenting the Universe. 3: The construction and statistics of 3-D Voronoi tessellations. *Astronomy and Astrophysics*, 283:361–406, 1994.

[105] T. van der Jagt, G. Jongbloed, and M. Vittorietti. Existence and approximation of densities of chord length- and cross section area distributions. *Image Analysis and Stereology*, 42:171–184, 2023.

[106] T. van der Jagt, G. Jongbloed, and M. Vittorietti. Stereological determination of particle size distributions for similar convex bodies. *Electronic Journal of Statistics*, 18:742–774, 2024.

[107] T. van der Jagt, G. Jongbloed, and M. Vittorietti. Nonparametric inference for Poisson-Laguerre tessellations. *Scandinavian Journal of Statistics*, 52:1816–1851, 2025.

[108] T. van der Jagt, M. Vittorietti, K. Sedighiani, C. Bos, and G. Jongbloed. Estimation of 3D grain size distributions from 2D sections in real and simulated microstructures. *Computational Materials Science*, 256:Article 113949, 2025.

[109] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

[110] Y. Vardi. Multiplicative Censoring, Renewal Processes, Deconvolution and Decreasing Density: Nonparametric Estimation. *Biometrika*, 76:751–761, 1989.

[111] H. Wadell. Volume, Shape, and Roundness of Quartz Particles. *The Journal of Geology*, 43:250–280, 1935.

[112] L. Wang, E. Mostaed, X. Cao, G. Huang, A. Fabrizi, F. Bonollo, C. Chi, and M. Vedani. Effects of texture and grain size on mechanical properties of AZ80 magnesium alloys at lower temperatures. *Materials & Design*, 89:1–8, 2016.

[113] J. A. Wellner and Y. Zhan. A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. *Journal of the American Statistical Association*, 92:945–959, 1997.

[114] S. D. Wicksell. The Corpuscle Problem. A Mathematical Study of a Biometric Problem. *Biometrika*, 17:84–99, 1925.

[115] Y. Wu, W. Zhou, B. Wang, and F. Yang. Modeling and characterization of two-phase composites by Voronoi diagram in the Laguerre geometry based on random close packing of spheres. *Computational Materials Science*, 47:951–961, 2010.

[116] F. Yin, A. Sakurai, and X. Song. Determination of spatial grain size with the area-weighted grain area distribution of the planar sections in polycrystalline materials. *Metallurgical and Materials Transactions A*, 37:3707–3714, 2006.

[117] T. Zamfirescu. Nearly all convex bodies are smooth and strictly convex. *Monatshefte für Mathematik*, 103:57–62, 1987.

[118] V. M. Zolotarev. Mellin-Stieltjes Transforms in Probability Theory. *Theory of Probability & Its Applications*, 2:433–460, 1957.

Summary

In this thesis we develop statistical methodology for stereological estimation problems which in particular appear in the field of materials science. We study mathematical models which may describe materials microstructures, and we develop statistical methods for estimating the parameter(s) of these models in a stereological setting. That is, typically we consider a 3D model, and instead of directly observing data generated by this model we only observe a 2D planar section. In many cases we even consider $(d-1)$-dimensional sections of a d-dimensional model ($d = 2, 3, \dots$), because it turns out this is often not more complicated than the $d = 3$ case. In order to formally define and study these models we mainly rely on concepts from stochastic geometry and point process theory. This thesis is divided into two parts. In part I we consider a model of randomly positioned and -oriented particles of the same shape, but varying in size, and in part II we consider a space-filling model by studying a random tessellation (mosaic). In these models, the particles of a particle process, or the cells of a random tessellation, may represent the so-called grains of a materials microstructure.

In part I we study a model describing a system of 3D particles, a particle process. Here, each particle has the same known shape, but the particles do not have the same size. Then, the system of particles is intersected with a plane and the areas of the observed 2D section profiles are used to estimate the 3D particle size distribution. The estimation problem we consider here can be seen as a generalization of the classical Wicksell corpuscle problem. Because we consider general convex shapes for the particles instead of the ball (as in Wicksell's corpuscle problem), this model is more flexible, and therefore has the potential to more accurately describe real materials microstructures.

In order to deal with this estimation problem we first study areas of random sections of a fixed convex body in chapter 2. We show that for a large class of convex bodies, the so-called section area / volume distribution is absolutely continuous with respect to Lebesgue measure. Via this result we define a nonparametric likelihood-based estimator for the 3D particle size distribution in chapter 3. We prove strong consistency of this estimator, discuss computational aspects, and we perform a simulation study to empirically study its behavior. Additionally, we show that this estimator can be computed efficiently and we observe that the estimator performs well for a range of particle shapes and underlying 3D particle size distributions. Note that consistent estimators have not yet been derived before in this context. Having introduced this methodology we apply it to a real materials microstructure, and to some commonly considered models for materials microstructures in chapter 4. That is, we empirically investigate how well we can estimate the 3D grain vol-

ume distribution in these (simulated) microstructures via our estimation procedure. Here, we focus in particular on the choice of the assumed grain shape. In the model used for defining an estimator for the 3D particle size distribution the common shape of the particles is assumed to be known. Therefore, in practice one needs to choose this shape. We observe that a few choices of the grain shape tend to yield the most accurate estimates. We also provide some recommendations on how the estimation procedure may be used in practice by materials scientists.

In part II of this thesis we study another model, the so-called Poisson-Laguerre tessellation. This model is a generalization of the classical Poisson-Voronoi tessellation. At the heart of this model lies a Poisson point process which is used for defining the random cells of the tessellation. Loosely speaking, this point process describes a homogeneous pattern of points in space, and each point carries a randomly assigned weight. In this context the points are often called generators. The Poisson-Laguerre tessellation may also be described via a crystallic growth process, and then each weight can also be interpreted as the time at which a cell corresponding to a weighted generator point starts growing. The parameter in this model is a distribution function F, which effectively describes the distribution of the weights, or the distribution of the growth times / arrival times of cells.

In chapter 5 we define nonparametric estimators for this distribution function F. Throughout, only the point process of weighted generators corresponding to observed cells are assumed to be known. As a consequence of dependency in the model, this observed point process is no longer a Poisson point process. We develop statistical methodology for Poisson-Laguerre tessellations in two settings. First, we consider a setting with direct observations. This means that we directly observe a d-dimensional region of the d-dimensional tessellation. Then, we consider a stereological setting. This means that we only observe a 2D section of a 3D Poisson-Laguerre tessellation. In both settings we derive estimators for the underlying distribution function, which we prove to be strongly consistent as the so-called observation window tends to the whole space. Consistent estimators of the parameter(s) of a Poisson-Laguerre tessellation have not yet been derived before. In a simulation study we observe that the estimators perform satisfactorily for various choices of the underlying distribution function F.

In chapter 6 we study an issue associated with statistical inference for random Laguerre tessellations which seems to often be ignored in the literature. That is, for an observed Laguerre tessellation it is in general not possible to uniquely determine the weighted generators of the observed cells. In chapter 6 we characterize the class of all configurations of weighted generators leading to the same Laguerre tessellation, under a set of commonly satisfied regularity conditions. Then, we propose a method to asymptotically obtain a close approximation of the original weighted generators of an observed Poisson-Laguerre tessellation. We conclude the chapter by investigating how well the statistical methodology from chapter 5 performs when it is applied to these approximated weighted generator points instead of the true weighted generators. That is, we study via simulations whether we can still estimate F, when only the observed cells of the tessellation are considered known, without prior knowledge of the weighted generators. Here, we see that overall the estimation procedure still performs well, but the obtained estimates do behave somewhat differently compared to the estimates obtained in the simulation study in chapter 5.

Samenvatting

In dit proefschrift richten wij ons op het ontwikkelen van statistische methodologie voor stereologische schattingsproblemen. Deze schattingsproblemen zijn onder andere relevant voor het vakgebied materiaalkunde. We onderzoeken wiskundige modellen die gebruikt kunnen worden voor het beschrijven van microstructuren van materialen, en we introduceren statistische methoden voor het schatten van de parameter(s) van deze modellen in een stereologische context. Dit betekent dat we data gegenereerd door een gegeven 3D model niet direct observeren, in plaats daarvan zien we alleen een 2D doorsnede. In veel gevallen kijken we naar $(d-1)$-dimensionale doorsneden van d-dimensionale modellen $(d = 2, 3, \dots)$, omdat dit wiskundig gezien niet ingewikkelder is dan de $d = 3$ setting. Om deze wiskundige modellen formeel te kunnen definiëren en bestuderen maken we gebruik van concepten uit de stochastische meetkunde en puntprocestheorie. Dit proefschrift bestaat uit twee delen. In deel I bestuderen we een model voor willekeurig geplaatste en georiënteerde deeltjes van dezelfde vorm, maar verschillend in grootte. In deel II bestuderen we een ruimte-vullend model, omdat we ons hier bezighouden met willekeurige tessellaties (ook wel bekend als mozaïeken of betegelingen). In deze modellen kunnen de deeltjes van een deeltjesproces of de cellen van een willekeurige tessellatie de zogeheten korrels in de microstructuur van een materiaal beschrijven.

In deel I bestuderen we een model voor een stelsel van 3D deeltjes, een deeltjesproces. Elk deeltje heeft dezelfde vorm, maar deze deeltjes hebben verschillende groottes. Dit stelsel van deeltjes doorsnijden we met een vlak, en de oppervlaktes van de 2D deeltjes die we observeren in de doorsnede gebruiken we om de verdeling van de groottes van de 3D deeltjes te schatten. Dit schattingsprobleem is een veralgemenisering van het Wicksell-probleem. Omdat de deeltjes in ons model een convexe vorm naar keuze kunnen hebben in plaats van dat deze bolvormig zijn (zoals in het Wicksell-probleem), is dit model flexibeler, en daarmee geeft dit model potentieel een betere beschrijving van de microstructuur van een echt materiaal.

Om dit schattingsprobleem aan te kunnen pakken bestuderen we in hoofdstuk 2 eerst oppervlaktes van willekeurige doorsneden van een gegeven convex lichaam. We laten zien dat voor een grote klasse van convexe lichamen, de zogeheten doorsnede oppervlakteverdeling absoluut continu is ten opzichte van de Lebesgue-maat. Dit resultaat gebruiken we in hoofdstuk 3 om met behulp van de likelihoodfunctie een niet-parametrische schatter te definiëren voor de verdeling van de groottes van de 3D deeltjes. We tonen aan dat deze schatter sterk consistent is, we laten zien hoe de schatter te berekenen is, en we voeren simulaties uit om op empirische wijze het gedrag van de schatter te bestuderen. Consistente

schatters bestonden voorheen nog niet in deze context. We laten zien dat onze schatter op een efficiënte manier berekend kan worden en we observeren dat de schatter goed presteert voor verscheidene keuzes voor de vorm van de 3D deeltjes en hun grootte-verdeling. Na het introduceren en bestuderen van de schattingsmethode passen we deze in hoofdstuk 4 toe op de microstructuur van een staal, en op een aantal gesimuleerde microstructuren. Dit houdt in dat we op empirische wijze onderzoeken hoe accuraat de verdeling van volumes van de korrels in deze microstructuren geschat kunnen worden via onze schattingsmethode. We focussen hier met name op de keuze van de onderliggende korrelvorm. In het model dat we hebben gebruikt voor het definiëren van een schatter voor de verdeling van de groottes van de 3D deeltjes, is de gedeelde vorm van alle deeltjes aangenomen als bekend. Daarom is het in de praktijk nodig om deze vorm te kiezen. Uit de resultaten blijkt dat een paar keuzes voor de korrelvorm leiden tot de meest accurate schattingen. We geven ook enkele aanbevelingen voor het gebruik van onze schattingsmethode in de praktijk, in de context van materiaalkunde.

In deel II van dit proefschrift onderzoeken we een ander model, de zogenaamde Poisson-Laguerre tessellatie. Dit model is een veralgemenisering van de klassieke Poisson-Voronoi tessellatie. De kern van dit model is een Poisson-puntproces dat wordt gebruikt om de willekeurige cellen van de tessellatie te beschrijven. Het komt er ongeveer op neer dat dit puntproces een homogeen patroon van punten in de ruimte beschrijft, en bij elk punt hoort een willekeurig gewicht. In deze setting worden de punten vaak generatoren genoemd. De Poisson-Laguerre tessellatie kan ook worden beschreven via een kristallisatieproces, dan kan ieder gewicht worden gezien als het tijdstip waarop de corresponderende cel begint met groeien. De parameter in dit model is een verdelingsfunctie F, die praktisch gezien de verdeling van de gewichten beschrijft, ofwel de verdeling van de groeitijden van de cellen.

In hoofdstuk 5 definiëren we niet-parametrische schatters voor deze verdelingsfunctie F. In deze context nemen we aan dat alleen het puntproces van de gewogen generatoren die horen bij geobserveerde cellen bekend zijn. Als gevolg van de afhankelijkheden in het model is dit geobserveerde puntproces geen Poisson-puntproces. We ontwikkelen statistische methodologie voor twee verschillende gevallen. Ten eerste beschouwen we de situatie waarbij we directe observaties hebben. Dit betekent dat we een d-dimensionaal gebied van het d-dimensionale model observeren. Ten tweede beschouwen we een stereologische setting. Dit houdt in dat we alleen een 2D doorsnede van een 3D Poisson-Laguerre tessellatie observeren. In beide gevallen leiden we schatters af voor de onderliggende verdelingsfunctie, en laten we zien dat deze schatters sterk consistent zijn als het zogeheten observatie-raam groeit naar de hele ruimte. Consistente schatters voor de parameter(s) van een Poisson-Laguerre tessellatie waren nog niet eerder afgeleid. Via simulaties zien we dat de schatters naar behoren presteren voor verschillende keuzes van de onderliggende verdelingsfunctie F.

In hoofdstuk 6 bestuderen we een probleem dat gepaard gaat met het schatten van de parameter(s) van willekeurige Laguerre tessellaties. Het lijkt erop dat dit probleem in de literatuur meestal achterwege wordt gelaten. Het probleem is namelijk dat het voor een geobserveerde Laguerre tessellatie niet mogelijk is om op unieke wijze de gewogen generatoren van de geobserveerde cellen te achterhalen. In hoofdstuk 6 geven we een karakterisering van de klasse van alle mogelijke configuraties van gewogen generatoren die leiden tot dezelfde Laguerre tessellatie, onder een aantal voorwaarden die vaak worden

vervuld. Vervolgens stellen we een methode voor om in de limiet een goede benadering te krijgen van de originele gewogen generatoren van een geobserveerde Poisson-Laguerre tessellatie. Aan het einde van het hoofdstuk onderzoeken we hoe de statistische methodologie uit hoofdstuk 5 werkt als deze wordt toegepast op deze benadering van de gewogen generatoren in plaats van de echte gewogen generatoren. In feite onderzoeken we via simulaties of het mogelijk is om F te schatten wanneer alleen de geobserveerde cellen van de tessellatie als bekend worden beschouwd, zonder kennis van de gewogen generatoren aan te nemen. We zien dat de schattingsmethode in deze setting nog steeds leidt tot goede resultaten, al vertonen de resulterende schattingen wel een aantal afwijkende kenmerken ten opzichte van de schattingen die we hebben verkregen via simulaties in hoofdstuk 5.

Acknowledgements

When I started my PhD four years ago, I really did not quite realize what I got myself into. Initially, I started as a researcher in a temporary research position, and as time passed it slowly became clear to me how much I enjoyed the process of doing research. However, this process is not always easy. Over the years I had to deal with all kinds of challenges, and I want to thank everyone who helped me along the way.

I want to start by thanking my supervisors Geurt and Martina, for providing me with the opportunity to do a PhD. Your encouragement, feedback and support throughout these years helped me to stay motivated, to grow my skills as a researcher, and to deal with stressful times. Throughout my PhD we essentially always held meetings with the three of us, and I thank the both of you for always staying involved and checking in with me, I do realize it is not common for a PhD candidate to always have two supervisors to rely on. Geurt, because of your endless enthusiasm it is always a joy to work with you. The positive energy you bring always motivated me to keep going. Thank you Martina, for never hesitating to give advice or to encourage me.

I would like to thank my collaborators, Kees and Karo. You helped me better understand the challenges that materials scientists face, and thanks to your input we managed to develop new statistical tools which are useful for materials scientists. I also want to thank the Materials innovation institute (M2i) for funding the first year of my PhD.

I would like to thank the external members of the doctoral committee; Viktor Beneš, Annoesjka Cabo, Marie-Colette van Lieshout, Maria Santofimia Navarro and Aad van der Vaart, for reading my thesis, for providing constructive feedback and for participating in my PhD ceremony. I would also like to thank Marie-Colette for organizing the reading group ”An introduction to the theory of point processes”. These sessions helped me in gaining a better understanding of point processes, which was crucial for my research. Of course, also thanks to the other members of the reading group; Changqing, Justin, Maike, Robin and Zhuldyzay, for the stimulating discussions.

My time in Delft would not have been the same without my colleagues at the statistics department. Sharing an office with Francesco Gili has been a pleasure, thanks for the interesting conversations about life, talks about math and discussions about our careers. Thanks to my fellow PhDs Andrea, Ardjen, Chris, Dan, Dominique, Francesco Di Giuseppe, Jeffrey, Koen, Lasse, Marc, Máté, Matthias, Naqi, Sebastiano, Shixiang, Thorben and Wieger. Also thanks to my colleagues Alexis, Caroline, Fabian, Jakob, Joris, Hanne, Nestor and Rik. Thanks for the fun coffee- and lunch breaks, and conversations over drinks in the \Pub after the informal statistics seminars.

Thanks to my parents, Linda and Arie, for always supporting me and for being there for me. I could not have done this without you. Thanks to my brother, Sam, and his wife Allison. We cannot see each other very often in person, but I always look forward to seeing you, and whenever we are on different continents I am glad we stay in touch via video calls. Thanks to my family, I know that my work is complicated, but I very much appreciate your effort in trying to understand what I do. A special thanks to my grandfather, Fok, your curiosity and interest always cheers me up.

Thanks to my friends, for being there for me and for providing welcome distractions. A PhD can be challenging at times, and playing tennis helps me to unwind and to take my mind off work. I would like to thank my friends at T.V. D'Oudelande and T.V. Hiaten for all the great memories, on and off the court. Finally, thanks to my good friend and paranymph Frank. Throughout all these years since high school, our friendship has stayed strong, and I am very grateful for that.

Curriculum Vitae

Thomas Folkert Willem van der Jagt

26-07-1995 Born in Rotterdam, The Netherlands.

Education

2014–2018 Bachelor of Science in Applied Mathematics
Delft University of Technology, Delft, The Netherlands

2018–2021 Master of Science in Applied Mathematics
Delft University of Technology, Delft, The Netherlands

2021–2025 PhD in Statistics
Delft University of Technology, Delft, The Netherlands
Thesis: Stereological estimation for particle processes and random tessellations
Promotor: Prof. dr. ir. G. Jongbloed
Copromotor: Dr. M. Vittorietti

Publications

Published

- T. van der Jagt, G. Jongbloed, and M. Vittorietti. Existence and approximation of densities of chord length- and cross section area distributions. *Image Analysis and Stereology*, 42:171–184, 2023
- T. van der Jagt, G. Jongbloed, and M. Vittorietti. Stereological determination of particle size distributions for similar convex bodies. *Electronic Journal of Statistics*, 18:742–774, 2024
- T. van der Jagt, M. Vittorietti, K. Sedighiani, C. Bos, and G. Jongbloed. Estimation of 3D grain size distributions from 2D sections in real and simulated microstructures. *Computational Materials Science*, 256:Article 113949, 2025
- T. van der Jagt, G. Jongbloed, and M. Vittorietti. Nonparametric inference for Poisson-Laguerre tessellations. *Scandinavian Journal of Statistics*, 52:1816–1851, 2025

In preparation

- T. van der Jagt, G.Jongbloed, and M.Vittorietti. Inverting Poisson-Laguerre tessellations. 2026+