Stochastic Geometry and Random Tessellations

Jesper Møller¹ and Dietrich Stoyan²

¹ Department of Mathematical Sciences, Aalborg University, jm@math.aau.dk

² Institute of Stochastics, TU Bergakademie Freiberg, stoyan@math.tu-freiberg.de

1 Introduction

Random tessellations form a very important field in stochastic geometry. They pose many interesting mathematical problems and have numerous important applications in many branches of natural sciences and engineering, as this volume shows. This contribution aims to present mathematical basic facts as well as statistical ideas for random tessellations of the *d*-dimensional Euclidean space \mathbb{R}^d . Having applications in mind, it concentrates on the planar (d = 2) and spatial case (d = 3), while a general *d*-dimensional theory is presented to some extent. The theory of random tessellations uses ideas from various branches of stochastic geometry, for example point process and random set methods. Point processes play a fundamental role both in the construction of random tessellations (the points are cell nuclei) and in their description of 'accompanying structures' such as the point process of cell centres and the point process of vertices.

There are many models and constructions for random tessellations. This contribution concentrates on Voronoi tessellations and tessellations resulting from similar construction principles. For example, tessellations resulting from infinite planes or from crack processes are not discussed here. These topics and more material are presented in the monographs Møller (1994), Stoyan *et al.* (1995), and Okabe *et al.* (2000).

2 Stochastic geometry models

2.1 General

This section briefly presents fundamental models of stochastic geometry which are needed in the theory of random tessellations. These models are either used in the construction of tessellations or in their description; in the latter case one speaks about 'accompanying structures'. Many tessellation models are

defined starting from point processes, the most prominent example is the Voronoi tessellation. Furthermore, each tessellation is accompanied by point processes, for example by the point process of vertices or side face centres. Section 2.2 therefore presents point process theory in some detail.

Random set theory is used when studying the accompanying sets of the settheoretic union of all cell edges (planar case) or faces (spatial case). These sets have Lebesgue measure and volume fraction zero but are nevertheless valuable since set-theoretic characteristics such as contact distribution functions help to describe the size and variability of tessellation cells. The systems of all edges of tessellations form random segment processes, and the edge-length densities L_A (planar case) and L_V (spatial case) are typical parameters in the theory of fibre processes, which includes segment processes as a particular case. The system of all cell faces of a tessellation of \mathbb{R}^3 can be interpreted as a particular surface process, and the specific surface S_V is a parameter of particular interest. Finally, the systems of all edges of tessellations can be considered as particular random networks and analyzed by corresponding methods (Zähle, 1988; Mecke and Stoyan, 2001).

2.2 Point processes

This section provides an introduction to point processes in \mathbb{R}^d , where in applications the planar case d = 2 and the spatial case d = 3 are of particular interest. For more formal treatments, see Stoyan *et al.* (1995), Daley and Vere-Jones (2003), and Møller and Waagepetersen (2003).

Definitions

In the simplest case, a point process \mathbf{X} is a finite random subset of a given bounded region $S \subset \mathbb{R}^d$, and a realization of such a process is a point pattern $\mathbf{x} = \{x_1, \ldots, x_n\}$ of $n \ge 0$ points contained in S. We say that the point process is defined on S. To specify the distribution of \mathbf{X} , one may specify the distribution of the number of points, $n(\mathbf{X})$, and for each $n \ge 1$, conditional on $n(\mathbf{X}) = n$, the joint distribution of the n points in \mathbf{X} . An equivalent approach is to specify the distribution of the count variables $N(B) = n(\mathbf{X}_B)$ for subsets $B \subseteq S$, where $\mathbf{X}_B = \mathbf{X} \cap B$ denotes \mathbf{X} intersected with B.

If it is not known on which region the point process is defined, or if the process extends over a very large region, or if certain invariance assumptions such as stationarity are imposed, then it may be appropriate to consider an infinite point process on \mathbb{R}^d . We define a *point process* \mathbf{X} in \mathbb{R}^d as a locally finite random subset of \mathbb{R}^d , i.e. N(B) is a finite random variable whenever $B \subset \mathbb{R}^d$ is a bounded region. Again the distribution of \mathbf{X} may be specified by the distribution of the count variables N(B) for bounded subsets $B \subseteq \mathbb{R}^d$.

We say that **X** is *stationary* respective *isotropic* if for any bounded region $B \subset \mathbb{R}^d$, N(T(B)) respective N(R(B)) is distributed as N(B) for an arbitrary

translation T respective rotation R about the origin in \mathbb{R}^d . Stationary and isotropic point processes are also called *motion invariant*.

Note that stationarity implies that \mathbf{X} extends all over \mathbb{R}^d in the sense that (with probability one) N(H) > 0 for any closed halfspace $H \subset \mathbb{R}^d$. Stationarity and isotropy may be reasonable assumptions for point processes observed within a homogeneous environment. These assumptions appear commonly in stochastic geometry, and in particular the assumption of stationarity provides the basis for establishing many general results as shown e.g. in Section 3.2.

Fundamental characteristics

Moment measure and intensity

The mean structure of the count variables $N(B), B \subseteq \mathbb{R}^d$, is summarized by the *moment measure*

$$\mu(B) = \mathcal{E}N(B), \quad B \subseteq \mathbb{R}^d$$

(where EN(B) means the mean value of N(B)). Usually, for an arbitrary B, we can write

$$\mu(B) = \int_B \rho(u) \,\mathrm{d}u$$

where ρ is a non-negative function called the *intensity function*. We may interpret $\rho(u) du$ as the probability that precisely one point falls in an infinitesimally small region containing the location u and of size du. If **X** is stationary, then $\rho(u)$ will be constant and is called the *intensity* of **X**.

Covariance structure and pair correlation function

The covariance structure of the count variables is most conveniently given in terms of the second order factorial moment measure $\mu^{(2)}$. This is defined by

$$\mu^{(2)}(A) = \mathbf{E} \sum_{u,v \in \mathbf{X}}^{\neq} \mathbf{1}[(u,v) \in A], \quad A \subseteq \mathbb{R}^d \times \mathbb{R}^d,$$

where \neq at the summation sign means that the sum runs over all pairwise different points u, v in **X**, and $\mathbf{1}[\cdot]$ is the indicator function. For bounded regions $B \subseteq \mathbb{R}^d$ and $C \subseteq \mathbb{R}^d$, the covariance of N(B) and N(C) is expressed in terms of μ and $\mu^{(2)}$ by

$$Cov[N(B), N(C)] = \mu^{(2)}(B \times C) + \mu(B \cap C) - \mu(B)\mu(C).$$

For many important model classes, $\mu^{(2)}$ is given in terms of an explicitly known second order product density $\rho^{(2)}$,

$$\mu^{(2)}(A) = \int \mathbf{1}[(u,v) \in A] \rho^{(2)}(u,v) \, \mathrm{d}u \mathrm{d}v$$

where $\rho^{(2)}(u, v) du dv$ may be interpreted as the probability of observing a point in each of two regions of infinitesimally small sizes du and dv and containing u and v.

In order to characterize the tendency of points to attract or repel each other, while adjusting for the effect of a large or small intensity function, it is useful to consider the *pair correlation function*

$$g(u, v) = \rho^{(2)}(u, v) / (\rho(u)\rho(v))$$

(provided $\rho(u) > 0$ and $\rho(v) > 0$). If the points appear independently of each other, $\rho^{(2)}(u, v) = \rho(u)\rho(v)$ and g(u, v) = 1. When g(u, v) > 1 we interpret this as *attraction* between points of the process at locations u and v, while if g(u, v) < 1 we have *repulsion* at the two locations. In applications it is often assumed that g(u, v) depends only on the distance r = ||u - v||, and we write g(r) for g(u, v). Stationarity and isotropy of X implies this property.

The Poisson process

The Poisson process plays a fundamental role in stochastic geometry, partly because of mathematical tractability, partly since it serves as a reference process when more advanced models are considered, and partly since it is used for constructing more advanced models.

Definition

A Poisson process **X** defined on \mathbb{R}^d and with intensity measure μ and intensity function ρ satisfies for any bounded region $B \subseteq \mathbb{R}^d$ with $\mu(B) > 0$,

- (i) N(B) is Poisson distributed with mean $\mu(B)$,
- (ii) conditional on N(B), the points in \mathbf{X}_B are independent and identically distributed with density proportional to $\rho(u), u \in B$.

Some further definitions and properties

The Poisson process is a model for 'no interaction' or 'complete spatial randomness', since \mathbf{X}_A and \mathbf{X}_B are independent whenever $A, B \subset \mathbb{R}^d$ are disjoint. This property together with the definition above show how to construct the process, using a partition of \mathbb{R}^d into bounded subsets, and it is not hard to verify that this construction is well-defined and unique no matter how we choose the partition. Moreover,

$$\rho^{(2)}(u,v) = \rho(u)\rho(v), \quad g(u,v) \equiv 1,$$

reflecting the lack of interaction.

If $\rho(u)$ is constant for all $u \in \mathbb{R}^d$ with $\rho(u) > 0$, we say that the Poisson process is *homogeneous*. In particular, a stationary Poisson process is homogeneous as well as isotropic.

2.3 Random sets

A random set **X** is a random subset of \mathbb{R}^d (for a formal definition, see e.g. Stoyan *et al.* (1995)). Usually, and also in this text, random closed sets are considered, i.e. sets **X** where the boundary $\partial \mathbf{X}$ belongs to **X**. An example of a random set in the context of three-dimensional tessellations is the set-theoretic union of all cell faces of a given random tessellation of \mathbb{R}^3 .

The distribution of a random closed set **X** is given by its *capacity functional* T_X defined by

$$T_X(K) = P(\mathbf{X} \cap K \neq \emptyset) \quad \text{for} \quad K \in \mathcal{K},$$
 (1)

where \mathcal{K} is the family of all compact subsets K of \mathbb{R}^d (a subset of \mathbb{R}^d is called compact if it is closed and bounded).

A random set **X** is called *stationary* if **X** and the translated set $\mathbf{X}_x = \{y + x : y \in \mathbf{X}\}$ have the same distribution for all $x \in \mathbb{R}^d$. An equivalent condition is

$$T_X(K) = T_X(K_x)$$

for every $K \in \mathcal{K}$ and every $x \in \mathbb{R}^d$, where $K_x = \{k + x : k \in K\}$. Isotropy is defined analogously, where translations (by x) are replaced by rotations around the origin o, and motion invariance means that **X** is both stationary and isotropic.

The capacity functional $T_X(K)$ in (1) is defined for all $K \in \mathcal{K}$, but in practice \mathcal{K} is a too large family of 'test sets', so instead we usually consider a much smaller parameterized family of compact sets. Examples include the family of closed balls b(o, r) of radius r > 0 centred at the origin o, and the family of closed line segment s(o, r) of length r > 0 with one endpoint in o. In the motion invariant case this leads to the spherical and linear contact distribution functions $H_s(r)$ and $H_l(r)$ given by

$$H_s(r) = 1 - \frac{1 - T_X(b(o, r))}{1 - p}, \quad 0 \le r < \infty$$

and

$$H_l(r) = 1 - \frac{1 - T_x(s(o, r))}{1 - p}, \quad 0 \le r < \infty.$$

Here $p = P(o \in \mathbf{X})$ is the volume fraction of \mathbf{X} , and a heuristic interpretation of $H_s(r)$ $(H_l(r))$ is as follows: Consider an arbitrary fixed point $t \in \mathbb{R}^d$, condition on that $t \notin \mathbf{X}$, and consider the random distance from t to the nearest point (to the nearest point in prescribed direction) in \mathbf{X} . Then $H_s(r)$ $(H_l(r))$ is the distribution function of this distance. For the random sets which usually accompany tessellations, we have p = 0, and so $H_s(r) = T_X(b(o, r))$ and $H_l(r) = T_X(s(o, r))$.

Again in the motion invariant case of the random set \mathbf{X} , consider an arbitrary line S in \mathbb{R}^d . Its intersection with \mathbf{X} generates a series of intervals outside of \mathbf{X} . In the case of interest in this paper, where \mathbf{X} is the union of all (d-1)-dimensional cell-faces of a tessellation, these intervals are separated by points which are intersections of S and cell-faces. The chord length distribution function L(r) is defined as the distribution function of the length of an interval which (loosely speaking) is uniformly chosen among these intervals, and the mean chord length \overline{l} is the expected length of this interval (a formal definition is to use Palm distributions as defined later on, and to let the interval be given by the "typical chord outside S"). This is closely related to the linear contact distribution function, since

$$H_l(r) = \frac{1}{\overline{l}} \int_0^r [1 - L(s)] \, \mathrm{d}s \,, \quad 0 \le r < \infty$$

assuming $0 < \overline{l} < \infty$.

2.4 Germ-grain processes

Definitions

If there to each point x in a point process \mathbf{X} is associated a mark in the form of a compact random set $K_x \subset \mathbb{R}^d$, we consider the marked point process $\{(x, K_x) : x \in \mathbf{X}\}$, assuming its distribution is well-defined (see e.g. Stoyan *et al.* (1995)). This is in fact a particular type of a marked point process, which we here call a *germ-grain process* (Hanisch, 1981), where x is the germ and K_x is the grain of a geometric object $E(x) = x + K_x$, the translate of K_x by x. As shown later many aspects of random tessellations can be described by such processes. For example, a *fiber process*, where the fibres are bounded closed curves, can be represented as a germ-grain process, where to each fibre E we have a germ x, given e.g. by the midpoint of E, and a grain $K_x = E - x$. Note that in the stochastic geometry literature, it is usually the random set $\Psi = \bigcup_{x \in \mathbf{X}} E(x)$ given by the union of all geometric objects which is called a germ-grain model, and there may be many natural ways of choosing the germs (see below).

The germ-grain process is *stationary* if its distribution is invariant under translations in \mathbb{R}^d , i.e. if $\{x + y, K_x\} : x \in \mathbf{X}\}$ is distributed as $\{(x, K_x) : x \in \mathbf{X}\}$ for any point $y \in \mathbb{R}^d$. Further, it is *isotropic* if its distribution is invariant under rotations about the origin in \mathbb{R}^d . Stationary and isotropic germ-grain processes are also called *motion invariant*.

Palm distribution

The concept of Palm distributions is useful when defining what is meant by a typical grain. For the purpose of this text, we only define the Palm distribution of the typical grain in the case of a stationary germ-grain process $\{(x, K_x) : x \in \mathbf{X}\}$ when \mathbf{X} has a positive, finite intensity ρ . Then the Palm distribution is defined for all possible events F of a grain by

$$P(F) = \frac{1}{\rho|B|} E\left(\sum_{x \in \mathbf{X}_B} \mathbf{1} \left[K_x \in F\right]\right)$$
(2)

where $B \subset \mathbb{R}^d$ is an arbitrary region with positive, finite content |B| (i.e. |B| is the area of |B| if d = 2, and the volume of |B| if d = 3). Moreover, the typical grain is a random set following this Palm distribution. Since the right hand side in (2) does not depend on the choice of B, and we may let B expand to \mathbb{R}^d , we can interpret the Palm distribution as the distribution of a uniformly selected grain, justifying the terminology 'typical grain'.

As mentioned, there may be many natural ways of choosing the germs for a process E(x) of random sets indexed by a point process **X**. For mathematical reasons, in the stationary case, it is convenient to use so-called *covariant centroids* $c(E) \in \mathbb{R}^d$, meaning that c(T(E)) = T(c(E)) for all possible realizations E of the grains and all translations T. If E is a *convex polytope* (i.e. a finite intersection of closed halfspaces), the centroid c(E) is natural given by the centre of gravity (i.e. the mean of the vertices of E), and this choice is clearly covariant. Another covariant choice is the 'most extreme point' of E in a given fixed direction. No matter the choice it turns out that the geometric properties of the typical grain are the same as long as the centroid is covariant; by 'geometric properties' of e.g. a polyhedron we mean for example its volume, surface area, total length of edges, and number of vertices; see e.g. Møller (1989). Indeed all results presented in the sequel do not depend on the choice of covariant centroid.

3 Random tessellations

3.1 Basic definitions and assumptions

Definition of random tessellations

By a *tessellation* of a given bounded region $S \subset \mathbb{R}^d$ or of the entire space $S = \mathbb{R}^d$, we mean a countable subdivision of S into non-overlapping, compact d-dimensional sets C_i called *cells*. Thus $S = \bigcup_{i \in I} C_i$, where I is a countable index set, each cell C_i is closed and bounded with d-dimensional interior int C_i , and the cells have disjoint interiors: $\operatorname{int} C_i \cap \operatorname{int} C_j = \emptyset$ whenever $i, j \in I$ with $i \neq j$. Further assumptions are usually added, including that the collection of cells is *locally finite* in the sense that an arbitrary bounded region of \mathbb{R}^d is intersected only by a finite number of cells. Often tessellations with *convex cells* (more precisely, the cells are bounded d-dimensional convex polytopes) have been studied in stochastic geometry, though tessellations with non-convex cells

also play some importance and will be considered to some extent also in this contribution.

Suppose that we have specified the joint distribution of a sequence C_i , $i \in I$, of random closed sets so that the sequence is a tessellation of $S = \bigcup_{i \in I} C_i$ (or at least with probability one, it is a tessellation). This is called a *random* tessellation of S.

Intersections between cells

Consider a random tessellation C_i , $i \in I$. For most random tessellation models studied in stochastic geometry, if d = 2, each boundary ∂C_i splits into a number of edges and vertices given by non-void intersections between C_i and one or more other cells C_j . Similarly, if d = 3, non-void intersections between two or more cells result in general in 0, 1, and 2 dimensional connected sets called vertices, edges, and faces, respectively. These concepts can be made more precise under the assumption in the next paragraph.

Henceforth, we restrict attention to the following case (which may actually only happen with probability one). Suppose that each cell is a connected set, which is only intersected by a finite number of other cells, and any non-void intersection between k > 1 cells is a finite union of (maximal) connected components of the same dimension m = m(k), say. The intersection is then called an *m*-facet and each of its connected components is called an *m*-face. Note that a 0-face is nothing but a point or *vertex* of the tessellation, a 1-face is a closed curve called an *edge* of the tessellation, and if d = 3, a 2-face is what we above just called a face. It is convenient to call a cell a *d*-facet or *d*-face (recalling that a cell is assumed to be connected).

If for any k = 2, ..., d+1 and any non-void intersection between k cells we always have that m(k) = d-k+1, we say that the tessellation is *normal*, since many real-life non-artificial tessellations possess this property. For instance, a planar tessellation is normal if the edges and vertices are given by the non-void intersections between pairs respective triplets of cells.

For normal tessellations with convex cells, the sets of *m*-facets and *m*-faces coincide, and they possess many desirable geometrical and topological properties:

- (i) For k = 1, ..., d, any (d k)-face is a bounded convex polytope of dimension d k, and it lies in the relative boundaries of (k+1)!/((l+1)!(k-l)!) intersecting (k l)-faces, $0 \le l \le k \le d$.
- (ii) The set of *m*-faces is locally finite in the sense that the number of *m*-faces intersecting a given bounded subset of R^d is finite.

Accompanying structures

The *m*-faces and *m*-facets of a random tessellation define accompanying structures of random sets, e.g. the point processes of vertices (m = 0) and cell centres (m = d), the fiber process of edges (m = 1), and the surface process of 2-faces (m = 2 and d = 3). Such structures can be represented as germgrain processes, and they are helpful for the description of particular aspects of random tessellations as demonstrated in the following section.

3.2 Stationary tessellations

A random tessellation with cells C_i , $i \in I$, is stationary if its distribution is invariant under translations in \mathbb{R}^d , i.e. if the collection $\{T(C_i), i \in I\}$ of all cells transformed by an arbitrary translation T is distributed as $\{C_i, i \in I\}$. This implies that $\{C_i, i \in I\}$ is necessarily a tessellation of the entire space \mathbb{R}^d . Moreover, *isotropy* of the random tessellation means that its distribution is invariant under rotations about the origin in \mathbb{R}^d . Stationary and isotropic tessellations are also called *motion invariant*.

Stationarity allow us to define the concepts of the typical cell and the typical *m*-face, and to establish various mean value relations for the geometrical characteristics of the accompanying structures of the tessellation. The planar (d = 2) and spatial (d = 3) cases have been studied in Mecke (1980, 1984), Radecke (1980), Møller (1994), and Stoyan *et al.* (1995); general definitions and results in $d \ge 1$ dimensions are given in Møller (1989). These results allow us to parameterize all mean values considered by only a few of them, which in turn often may be easily estimated by non-parametric statistical methods.

In the sequel we concentrate on mean value relations for the 'practical' planar (d = 2) and spatial (d = 3) cases of a stationary tessellation with convex cells. For the accompanying structures represented as germ-grain processes, we use covariant centroids (see Section 2.4).

Mean value relations in the planar case

Consider a planar (d = 2) stationary tessellation with convex cells, and introduce the following notation.

- (i) The point processes of vertices, edge midpoints, and cell centroids are denoted X₀, X₁, X₂, respectively. These are stationary with intensities ρ₀, ρ₁, ρ₂, which are assumed to be positive and finite.
- (ii) For each vertex $x \in \mathbf{X}_0$, consider the edges emanating from x, and denote $N_{0,2}(x)$ the number and $L_0(x)$ the total length of these edges.
- (iii) For each $x \in \mathbf{X}_1$, let $L_1(x)$ denote the length of the associated edge.
- (iv) For each $x \in \mathbf{X}_2$ and the associated cell, denote $N_{2,0}(x)$ the number of edges (or equivalently number of vertices), $L_2(x)$ the length of the perimeter, and $A_2(x)$ the area of the cell.
- (v) Let L_A denote the intensity of the fiber process of edges, i.e. $L_A|B|$ is the mean length of the union of all edges intersected with an arbitrary region $B \subset \mathbb{R}^2$.

Note that ρ_m is the intensity of *m*-faces. In each case of (ii)-(iv) we have obviously an underlying stationary germ-grain process, and we can thereby

define mean values for the geometric characteristics in (ii)-(iv) with respect to the typical vertex, edge or cell. We denote these mean values by $N_{0,2}$, L_0 , L_1 , $N_{2,0}$, L_2 , A_2 ; e.g. $N_{0,2}$ is the mean number of edges emanating from the typical vertex.

Now, mean value relations expressing the parameters ρ_0 , ρ_1 , ρ_2 , $N_{0,2}$, L_0 , L_1 , $N_{2,0}(x)$, A_2 , and L_A in terms of only three parameters, namely ρ_0 , ρ_2 (or $N_{0,2}$), and L_A , have been established, see e.g. Stoyan *et al.* (1995, Section 10.3). Note that ρ_0 , ρ_2 (or $N_{0,2}$), and L_A may be easily estimated in practice (see Section 5). Normality of the tessellation is equivalent to that $N_{0,2} = 3$, in which case we only need to determine two parameters, e.g. ρ_2 and L_A .

Mean value relations in the spatial case

Consider a spatial (d = 3) stationary tessellation with convex cells, and in a similar way as in the planar case, define

- (i) ρ_m , the intensity of *m*-faces (m = 0, 1, 2, 3), which is assumed to be positive and finite;
- (ii) $N_{k,l}$, the mean number of *l*-faces adjacent to the typical *k*-face $(k, l \in \{0, 1, 2, 3\})$;
- (iii) L_1 , the mean length of the typical edge;
- (iv) L_2 and A_2 , the mean perimeter and mean area of the typical 2-face (or just 'face');
- (v) L_3 , S_3 , V_3 , and B_3 , the mean total edge length, the mean surface area, the mean volume, and mean of the average breadth of the typical cell;
- (vi) L_V , the intensity of the fiber process of edges, i.e. $L_V|B|$ is the mean length of the union of all edges intersected with an arbitrary region $B \subset \mathbb{R}^3$;
- (vii) S_V , the intensity of the surface process of 2-faces, i.e. $S_V|B|$ is the mean area of the union of all 2-faces intersected with an arbitrary region $B \subset \mathbb{R}^3$;
- (viii) T_V , the intensity of vertices weighted according to the numbers of adjacent cells, i.e. $T_V|B|$ is the mean of the sum of all such weights for vertices within an arbitrary region $B \subset \mathbb{R}^3$;
 - (ix) Z_V , the intensity of the fibre process of edges weighted according to the numbers of adjacent cells, i.e. $Z_V|B|$ is the mean of the sum over all edges where each term in the sum is obtained by multiplying the number of adjacent cells of an edge by the length of the intersection of that edge with an arbitrary region $B \subset \mathbb{R}^3$.

All the parameters in (i)-(ix) can be expressed by seven parameters, namely ρ_0 , ρ_3 , $\rho_1 + \rho_2$, L_V , S_V , T_V , and Z_V , which may be easily estimated in practice (see Section 5). In the case of a normal tessellation we only need to determine three parameters, e.g. ρ_0 , ρ_3 , and L_V . For details, see e.g. Stoyan *et al.* (1995, Section 10.4).

The typical cell and the point-sampled cell

We return now to the general setting of a stationary tessellation in \mathbb{R}^d , without assuming convexity of the cells. Denote C_* the cell containing the origin 0 (or any other arbitrary fixed point in \mathbb{R}^d); sometimes it is said to be a *point* sampled cell. The point process of cell centroids is stationarity with an intensity ρ_d . Assuming $0 < \rho_d < \infty$, let \mathcal{C} denote the *typical cell* of the tessellation. Intuitively, C_* is larger than \mathcal{C} ; this can be formalized as shown in Mecke (1999).

4 Voronoi and Delaunay tessellations

4.1 General definitions, assumptions, and properties

Definition and some general properties of Voronoi tessellation

Given a realization of point process $\mathbf{X} = \{x_i\}$ in \mathbb{R}^d , the Voronoi tessellation of \mathbb{R}^d with nuclei $\{x_i\}$ has cells

$$C_i = \{ y \in \mathbb{R}^d : ||x_i - y|| \le ||x_j - y|| \text{ for all } j \neq i \},\$$

where $\|\cdot\|$ is the usual Euclidean distance. In other words, the Voronoi cell C_i consists of all points closer to the nucleus x_i than to any other nucleus. Clearly, the Voronoi cells have disjoint interiors, and they are convex, closed sets. Since **X** is locally finite, each point in \mathbb{R}^d belongs to finitely many Voronoi cells, so the Voronoi cells are space-filling ($\mathbb{R}^d = \bigcup_i C_i$). Further conditions are needed to ensure boundness and other desirable properties of the Voronoi cells. For example, stationarity of **X** implies that the Voronoi cells are bounded and hence compact. Furthermore, for subsets $S \subset \mathbb{R}^d$, we have a *Voronoi tessellation of S* which is simply given by the restriction of the Voronoi cells to S.

In the sequel, like in most of the stochastic geometry literature on Voronoi tessellations, we assume the nuclei to be in general quadratic position, that is (with probability one) no k + 1 nuclei lie on a (k - 1)-dimensional affine subspace of \mathbb{R}^d , $k = 2, \ldots, d$, and no d + 2 nuclei lie on the boundary of a sphere. Then the Voronoi cells are compact d-dimensional polytopes, and any bounded region of \mathbb{R}^d is intersected by only finitely many Voronoi cells, so that in the sense of Section 3.1, the Voronoi tessellation is in fact a tessellation of \mathbb{R}^d . Moreover, the Voronoi tessellation is normal. Consequently, in the stationary case, the general simple mean value relations for normal stationary tessellations apply, cf. Section 3.2.

Definition and some general properties of Delaunay tessellation

Since the Voronoi tessellation is normal, each of its vertices is given by an intersection of exactly d + 1 Voronoi cells. The corresponding d + 1 nuclei define a *Delaunay cell*; this is the closed *d*-dimensional simplex with vertices given by these nuclei, i.e. a closed triangle if d = 2, a closed tetrahedron if d = 3, and so on. The Delaunay cells constitute a tessellation of \mathbb{R}^d , the

Delaunay tessellation. The two types of tessellations are said to be dual, since the vertices of the Voronoi tessellation correspond to the Delaunay cells. Note that each k-facet of the Delaunay tessellation is a k-dimensional closed simplex with vertices given by k + 1 nuclei whose Voronoi cells share a (d - k)-facet of the Voronoi tessellation. For this and other reasons, although the Delaunay tessellation is not normal, it is more tractable for mathematical analysis than the Voronoi tessellation.

Consider again the stationary case; clearly, stationarity of the nuclei is equivalent to stationarity of the Voronoi or Delaunay tessellation. By duality, i.e. since the k-faces of the Delaunay tessellation are in one-to-one correspondence with the (d-k)-faces of the Voronoi tessellation, we easily obtain simple mean value relations for the Delaunay tessellation. For instance, the intensity ρ_{d-k} of (d-k)-faces of the Voronoi tessellation equals the intensity of k-faces of the Delaunay tessellation.

4.2 Poisson-Voronoi tessellations

Stochastic geometry models for the point process of nuclei generating a Voronoi tessellation are very different from the highly regular patterns of nuclei used in the seminal work of Dirichlet (1850) and Voronoi (1908). An important particular case is the Voronoi tessellation with respect to a homogeneous Poisson process. For this stationary Poisson-Voronoi tessellation numerous useful results can be established.

Throughout this section, the point process **X** of nuclei is assumed to be a stationary Poisson process with positive and finite intensity ρ . This implies that the nuclei are in general quadratic position.

First order mean value results

As the nuclei are clearly covariant, the intensity of Voronoi cells is $\rho_d = \rho$. Furthermore,

$$L_A = 2\sqrt{\rho} \quad \text{when } d = 2, \tag{3}$$

while if d = 3 we have

$$\lambda_0 = \frac{24\pi^2 \rho}{35} \approx 6.768\rho, \quad L_V = \frac{16}{15} \left(\frac{3}{4}\right)^{1/3} \pi^{5/3} \Gamma\left(\frac{4}{3}\right) \rho^{2/3} \approx 5.832 \rho^{2/3}, \quad (4)$$

where Γ is the Gamma-function. These results follow from a general result in arbitrary dimensions d for the *density of m-faces*, i.e., the intensity of the random set given by the union of *m*-faces of a stationary Poisson-Voronoi tessellation. Specifically, denoting the density of *m*-faces by $\lambda_{d,m}$ and setting k = d - m, we have

$$\lambda_{d,m} = \frac{\lambda^{k/d} 2^{k+1} \pi^{k/2} \Gamma(\frac{dk+d-k+1}{2}) \Gamma(\frac{d}{2}+1)^{k+((d-k)/d)} \Gamma(k+\frac{d-k}{d})}{(k+1)! \Gamma(\frac{dk+d-k}{2}) \Gamma(\frac{d+1}{2})^k \Gamma(\frac{d-k+1}{2}) d}.$$
 (5)

13

Equation (5) is obtained by combining the extremely useful Slivnyak-Mecke formula (considered in more detail in Section 4.4) with an integral geometric result known as the Blaschke-Petkantschin formula, see Miles (1974) and Møller (1989, 1994, 1999). Indeed, (5) when $d \leq 3$ was already established by Meijering (1953).

Recall that for the mean value relations discussed in Section 3.2 where $d \leq 3$, we need only to know $\rho_d = \rho$ and the parameters in (3)–(4) in the planar and spatial cases. If d = 2, we have

$$\rho_0 = 2\rho, \quad \rho_1 = 3\rho, \quad \rho_2 = \rho,$$

$$N_{0,2} = 3, \quad L_0 = 2/\sqrt{\rho}, \quad L_1 = 2/(3\sqrt{\rho}),$$

$$N_{2,0} = 6, \quad L_2 = 4/\sqrt{\rho}, \quad A_2 = 1/\rho, \quad L_A = 2\sqrt{\rho}.$$

The results for d = 3 are given in e.g. Stoyan *et al.* (1995).

Other results

Higher order moments and many other properties of the Poisson-Voronoi tessellation than those considered above are often more complicated to derive and require in general either numerical or Monte Carlo methods. Such results are briefly discussed below; see also Okabe *et al.* (2000).

Higher order mean values: For $d \leq 3$, Gilbert (1962) derived an integral expression for the variance of the size of the typical Poisson-Voronoi cell and the typical sectional cells obtained by considering the tessellation given by the intersection of the stationary Poisson-Voronoi tessellation with a line or plane. Covariances and variances of other kind of characteristics can be calculated as well; in general the details are intricate. This is demonstrated in two unpublished papers by Brakke (1987a, 1987b). Tables showing Gilbert's and Brakke's results can be found in Møller (1994, Section 4.2).

The typical edge: Figures of Brakke's numerical results for the density functions of the length of the typical planar (d = 2) and spatial (d = 3) Poisson-Voronoi edge are shown in Møller (1994); see also Muche (1993, 1996b) and Schlather (2000). Recent work by Muche (2005) and Baumstark and Last (2006) establish a complete description of the joint distribution of the typical Poison-Voronoi edge and the accompanying point process of the d nuclei of the Voronoi cells containing the typical Poison-Voronoi edge.

The typical cell: If we naturally choose the centroids of cells to be given by the nuclei \mathbf{X} , the distribution of the typical Poisson-Voronoi cell is by the Slivnyak-Mecke formula the same as the Voronoi cell with nucleus 0 if we consider the Voronoi tessellation with nuclei $\mathbf{X} \cup \{0\}$. This can formally be expressed by the following distributional equality

$$\mathcal{C} = \{ y \in \mathbb{R}^d : \|y\| \le \|y - x\| \text{ for all } x \in \mathbf{X} \},$$
(6)

which is useful for simulating realizations of the typical Poisson-Voronoi cell, see Section 7.

Table 10.4 in Stoyan *et al.* (1995), based on Monte Carlo simulations from Miles and Maillardet (1982), shows the distribution of the *number of vertices* in the *typical planar Poisson-Voronoi cell*. Calka (2002, 2003) studied size and shape of planar Voronoi cells. Hug *et al.* (2004) showed that large Voronoi cells tend to be spherical. Hayen and Quine (2002) calculated the first three moments of the area of the typical Poisson-Voronoi cell in the plane.

Full Voronoi neighbours: Two nuclei are said to be full Voronoi neighbours if the intersection between the corresponding two Voronoi cells and the line segment with endpoints given by the two nuclei is non-empty, and the full Voronoi neighbouring graph is called the Gabriel graph (the Gabriel graph is a connected subgraph of the Delaunay graph given by Delaunay edges). The mean number of Gabriel (or full) neighbours to the typical Poisson-Voronoi cell is 2^d , and if L is the length of the typical Gabriel edge, then L^d follows an exponential distribution with mean $\Gamma(1 + d/2)/((\pi/4)^{d/2}\rho)$ (Møller, 1994).

Contact distribution functions: Muche and Stoyan (1992) derived integral formulae for $H_l(r)$ and $H_s(r)$ in the case of the random set **X** given by the union of all cell faces (d=3) or edges (d=2). The formulae are numerically tractable and lead to formulae for the chord length distribution function L(r). The corresponding density functions l(r) are shown in Figure 1 for the planar and spatial case for a Poisson process of unit intensity.

Heinrich (1998) studied contact and chord length distributions for Voronoi tessellations with respect to some non-Poisson processes in \mathbb{R}^d . Numerical results were given for Poisson cluster and Gibbs processes.

Angular distributions: The distributions of various angles (e.g. at the typical point or dihedral angles at vertices) are given for the spatial case in Muche (1998, 2005) and Schlather (2000).

Pair correlation function of point process of vertices: Heinrich et al. (1998) studied the pair correlation function of the accompanying point process of vertices and derived numerically tractable formulae. Figure 2 shows this function for the spatial (d=3) case. There is a pole of order one at r=0 and a small local maximum at r=1.5. It can be shown that the pole results from very short edges. Such poles have been also observed for vertex processes of Voronoi tessellations with respect to point processes that are more regular than Poisson processes.

Gamma type results: Møller and Zuyev (1996) derived various gamma-type results and conditional independence results between size and shape of different geometric characteristics determined by a stationary Poisson process. One example is the fundamental region Δ of the typical Poisson-Voronoi cell as defined by the union of balls with centers at the vertices of C and containing 0 in their boundaries. Under the condition that C has $n (\geq d + 1)$ faces, $|\Delta|$ is conditionally independent of the shape and orientation of Δ , and $|\Delta|$



Fig. 1. The chord-length probability densities for Poisson-Voronoi tessellations when d = 2 and d = 3, where in both cases the generating Poisson process is of unit intensity.

follows a gamma distribution with shape parameter n and scale parameter $1/\rho$; see also Miles and Maillardet (1982), Zuyev (1992) and Møller (1994). Other examples of results of this type will be considered in Section 4.3.

4.3 Poisson-Delaunay tessellations

Assume still that the point process **X** of nuclei is a stationary Poisson process with positive and finite intensity ρ . The *typical Poisson-Delaunay cell* is denoted \mathcal{D} ; it is (almost surely) a *d*-dimensional simplex centred at the origin 0 (corresponding to a typical vertex of the Voronoi tessellation).

Miles (1974) determined the distribution of \mathcal{D} (see also Møller (1989, Theorem 7.5) for a simple proof): Let RU_0, \ldots, RU_d denote the d+1 vertices of \mathcal{D} , where R > 0 is the *typical vertex-nucleus distance* and U_0, \ldots, U_d are unit vectors. Then R is independent of the directions (U_0, \ldots, U_d) , and R^d is gamma distributed with shape parameter d and scale parameter $\Gamma(1 + d/2)/(\lambda \pi^{d/2})$.



Fig. 2. The pair correlation function g(r) for the point process of vertices of the spatial Poisson-Voronoi when the generating Poisson process is of unit intensity.

Further, the joint density function of (U_0, \ldots, U_d) is proportional to the *d*-content of the (d + 1)-simplex determined by these unit vectors (here, the density is with respect to the uniform distribution on the product space of d+1 unit spheres in \mathbb{R}^d ; the constant of proportionality is also known). Thereby, the moments of $|\mathcal{D}|$ can be derived:

$$\mathbf{E}\left[|\mathcal{D}|^k\right] = \frac{(d+k-1)!\Gamma(\frac{d^2}{2})\Gamma(\frac{d^2+dk+k+1}{2})\Gamma(\frac{d+1}{2})^{d-k+1}\prod_{i=2}^{d+1}\Gamma(\frac{k+i}{2})/\Gamma(\frac{i}{2})}{(d-1)!\Gamma(\frac{d^2+1}{2})\Gamma(\frac{d^2+dk}{2})\Gamma(\frac{d+k+1}{2})^{d+1}(2^d\pi^{(d-1)/2}\rho)^k}$$

for $k = 1, 2, \dots$ Specific results for d = 2, 3 are given in Møller (1994).

Miles' result plays a fundamental role in the statistical theory of shape (Kendall, 1989). Rathie (1992) has used the result for deriving the distribution of the area (when d = 2) and the volume (d = 3) of the typical Poisson-Delaunay cell. The density for the planar case d = 2 involves a modified Bessel function; the expression for d = 3 becomes more complicated. Also the distributions of the typical angle (when d = 2) and the typical edge (when d = 2, 3) of the Poisson-Delaunay tessellation have been determined: See Collins

(1968), Miles (1970), Sibson (1980), Møller (1994), and Muche (1996a). Muche (1999) studied the distributions of surface area, total edge length and mean breadth of the typical Delaunay cell in the spatial case. Recently, Baumstark and Last (2006) extended Miles' result to a complete description of the Palm distribution describing the nuclei as seen from a typical point on a k-face of the Voronoi tessellation.

4.4 Generalizations of random Voronoi tessellations

So far we have mostly considered Voronoi tessellations with nuclei from a stationary Poisson process. This section reviews some extensions, where the nuclei are specified by a another kind of point process model, or where the construction of Voronoi cells is modified.

$Non-Poisson\ models$

It is basically the Slivnyak-Mecke formula which makes the Poisson-Voronoi tessellation relatively tractable for mathematical analysis. This formula can be extended to characterize *Gibbs point processes* (Georgii, 1976; Nguyen and Zessin, 1979): For disjoint point patterns \mathbf{x} and $\{x_0, \ldots, x_k\}$ in \mathbb{R}^d , let $\lambda(\{x_0, \ldots, x_k\}; \mathbf{x})$ denote the conditional intensity of \mathbf{X} at the locations x_0, \ldots, x_k . Intuitively, $\lambda(\{x_0, \ldots, x_k\}; \mathbf{x}) dx_0 \cdots dx_k$ is the conditional probability that \mathbf{X} has a point in each of infinitesimally small regions around the points x_0, \ldots, x_k of content dx_0, \ldots, dx_k when we condition on that \mathbf{X} agrees with \mathbf{x} outside these regions. In the special case of a Poisson process with intensity function ρ , we have that $\lambda(\{x_0, \ldots, x_k\}; \mathbf{x}) = \rho(x_0) \cdots \rho(x_k)$. Now, \mathbf{X} is a *Gibbs point process* with conditional intensity $\lambda(\{x_0, \ldots, x_k\}; \mathbf{x})$ if

$$(k+1)! \operatorname{E} \sum_{\{x_0, \dots, x_k\} \subset \mathbf{X}} f(\mathbf{X} \setminus \{x_0, \dots, x_k\}, \{x_0, \dots, x_k\})$$
(7)
= $\int \cdots \int \operatorname{E} \left[\lambda(\{x_0, \dots, x_k\}; \mathbf{X}) f(\mathbf{X}, \{x_0, \dots, x_k\})\right] dx_0 \cdots dx_k$

for any integer $k \ge 0$ and non-negative function f. Equation (7) is called the (extended) *Georgii-Nguyen-Zessin formula* (or the GNZ-formula); the specification of $\lambda(\{x_0, \ldots, x_k\}; \mathbf{X})$ on the right side of (7) is arbitrary if $\{x_0, \ldots, x_k\}$ and \mathbf{X} are not disjoint. The GNZ formula reduces to the Slivnyak-Mecke formula in the special case of a Poisson process.

As an interesting example of a Gibbs point process, consider the hard core point process. This has conditional intensity $\lambda(\{x_0, \ldots, x_k\}; \mathbf{x})$ given by

$$\beta^{k+1}\mathbf{1}[\|\xi - \eta\| \ge \delta \text{ for distinct points } \{\xi, \eta\} \subset \mathbf{x} \cup \{x_0, \dots, x_k\}]$$
(8)

where $\beta > 0$ is a parameter controlling the intensity of the process, and $\delta > 0$ is a so-called hard core parameter. For the accompanying Voronoi tessellation,

each cell contains the ball of diameter δ centered at its nucleus. Thus, as β increases, we get more and more regular Voronoi cells.

Another possibly even more interesting model is obtained by replacing the hard core condition $\|\xi - \eta\| \ge \delta$ in (8) by a hard core condition on the size of the Voronoi cells, thereby obtaining *Ord's process*; see Baddeley and Møller (1989). Ord's process and many other examples of Gibbs models specified in terms of Voronoi tessellations are studied in Baddeley and Møller (1989), Kendall (1990), and Bertin *et al.* (1999a, 1999b).

Though Gibbs models may be more realistic for applications than Poisson models, and the GNZ-formula (7) makes it possible to obtain various estimation equations for the model parameters as well as the characteristics of the accompanying Voronoi tessellation, it remains to study such models in more detail.

Since most Gibbs point processes are only well-defined in the case where the points repel each other, the accompanying Voronoi tessellations will usually have more regular cells as compared to a Poisson-Voronoi tessellation. The same is true for the point processes of sphere centres in random packing of hard identical spheres as discussed in Lochmann *et al.* (2006a).

Also point processes where the points aggregate have been considered as models for the nuclei of a Voronoi tessellation. In particular *Poisson cluster processes* **X** have been used. Such a process is given by a union of 'offspring' point processes translated by a 'mother' point process; specifically, $\mathbf{X} = \bigcup_{y \in \mathbf{Y}} (y + K_y)$, where $\{(y, K_y) : y \in \mathbf{Y}\}$ is a germ-grain process, **Y** is a Poisson process of 'mother' points, and the grains K_y are finite 'offspring' point processes, which are independent and identically distributed and independent of **Y**. For example, in a *Matérn cluster process*, K_y is a homogeneous Poisson process defined on a ball with center 0. See Hermann *et al.* (1989), Møller *et al.* (1989), Lorz (1990), Lorz and Hahn (1993), Møller (1994, 1995), Saxl and Ponižil (2002) and Van de Weygaert (1994).

Tessellation constructions related to the Voronoi tessellation

The construction of a Voronoi tessellation has been generalized in various ways as exemplified below.

Generalized Voronoi tessellation: This kind of tessellation is also called a *near*est order *n* diagram. Given a point process **X** of nuclei, each cell of the generalized Voronoi tessellation is specified by a point configuration of *n* nuclei $\{x_1, \ldots, x_n\} \subset \mathbf{X}$. The cell consists of all points in \mathbb{R}^d at least as close to x_1, \ldots, x_n as to any other nuclei in **X**. Some probabilistic results when **X** is a stationary Poisson process for such tessellation have been established in Miles (1970) and Miles and Maillardet (1982).

Johnson-Mehl tessellation: Considering the Voronoi tessellation as the result of growing nuclei (with same speed and start of growth) one can generalize the construction to obtain the Johnson-Mehl tessellation (Johnson and Mehl, 1939), where nuclei starts to growth at different times. This tessellation has non-convex cells, and assuming stationarity, the Slivnyak-Mecke formula can be used to obtain a general expression for the density of faces of a *Poisson-Johnson-Mehl tessellation* (and for sectional Poisson-Johnson-Mehl tessellations as well), whereby further characteristics can be evaluated, see Møller (1992, 1995).

Laguerre tessellation: This kind of tessellation is also called a power tessellation, sectional Dirichlet tessellation or radical tessellation, see Okabe et al. (2000). It is generated with respect to a set \mathbf{X} of balls b(x, r) with centres xcalled nuclei and radii r. The Laguerre cell corresponding to b(x, r) is defined as

$$C(x,r) = \{ y \in \mathbb{R}^d : pow(y,(x,r)) \le pow(v,(x',r')) \text{ for all } b(x',r') \in \mathbf{X} \}$$

where $pow(y, (x, r)) = ||y-x||^2 - r^2$. These cells are closed convex polytopes. In the special case where all balls in **X** have equal radii, the Laguerre tessellation is just a Voronoi tessellation. If the radii are not equal, then in contrast to the Voronoi tessellation the Laguerre cells can be empty or a nucleus may be outside of its cell. If also the balls in **X** are non-overlapping (so-called hard balls), then each ball in **X** is contained in one Laguerre cell. This makes this tessellation interesting for the analysis and description of hard sphere systems, see Lochmann *et al.* (2006b). Figure 3 shows a Laguerre tessellation with respect to a system of random balls.

The Laguerre tessellation is also useful when studying and simulating interaction processes for balls specified in terms of geometric properties of the unions of balls (Møller and Helisova, 2007).

Similarly as in the case of a Voronoi tessellation, also Laguerre-Delaunay tessellations can be defined.

Probabilistic analysis of Laguerre tessellations is rather complicated, even in the case where **X** is an independently marked Poisson process. Lautensack (2007) derived integral formulae for many interesting tessellation characteristics, which can be numerically exploited. Examples in the spatial case (d = 3) are the cell volume distribution, the parameters S_V and L_V , and the intensity of the sub point process of Poisson process points with empty cells. Lautensack also shows that Laguerre tessellations are very good models for various cellular materials.

Anisotropic growth: Yet another generalization is to replace the Euclidean distance used in the definition of Voronoi cells with another Euclidean metric so that the growth is anisotropic. Scheike (1994) derived mean value relations for such tessellations.

5 Statistical inference

So far most research on random tessellations has focused rather on mathematical modelling and analysis than statistical aspects. This section considers first



Fig. 3. The Laguerre tessellation with respect to a random system of hard spheres.

non-parametric estimation of summary characteristics of tessellations and second parameter estimation for stationary Poisson-Voronoi tessellations. Third, recent progress with more complicated tessellation models, using a Bayesian simulation-based approach to inference, is considered.

5.1 Non-parametric estimation of summary characteristics of tessellations

When a sample of a stationary tessellation in \mathbb{R}^d is given within a *d*-dimensional observation window W, estimation of the corresponding summary characteristics is not difficult. The established methods of spatial statistics for point processes, fibre processes, surface processes and random sets can be applied, as sketched in Stoyan *et al.* (1995), p. 334, for the planar case.

For example, estimation of ρ_o means estimation of the intensity of a point process (that of the vertices), while the estimation of ρ_d means estimation of the mean number of grains per volume unit, where the grains are the cells. Further, L_A and L_V are line densities of fibre processes and can be estimated

by the statistical methods for these processes. Furthermore, $H_l(r)$ and $H_s(r)$ can be estimated by the standard methods for random sets, where the random set is here the set theoretic union of all cell boundaries (the union of all edges for d = 2 and of all 2-faces for d = 3).

However, in many applications these methods only apply for planar tessellations, while observation of three-dimensional tessellations are often only given by planar sections which leads to stereological problems as discussed in Section 6. Of course, for simulated d-dimensional tessellations, the methods more easily apply.

5.2 Parameter estimation for Poisson-Voronoi tessellations and related models

Parameter estimation for stationary Poisson-Voronoi tessellations is quite easy: there is only one parameter, the intensity ρ of the cell centre point process, and we can exploit that fundamental summary characteristics are expressed in terms of ρ . In the planar case (the spatial case is similar) there are three natural approaches:

(a) Estimating ρ_o , the intensity of the vertex point process, and then

$$\hat{\rho} = \hat{\rho}_o/2$$
.

(b) Estimating L_A , the line density of the system of edges, and then

$$\hat{\rho} = \hat{L}_{A}^{2}/2$$
.

(c) Estimating ρ_2 , the mean number of cells per area unit, and then

$$\hat{\rho} = \hat{\rho}_2$$
.

The best method is (a) since there are no edge-problems when estimating ρ_o and simple point counting suffices. If all three methods are carried out, comparison of the three estimates of ρ may lead to some impression on the validity of the stationary Poisson-Voronoi model assumption.

Also stereological methods lead in an elegant way to estimates of ρ for spatial tessellations, see Section 6.

For other models statistical analysis is rather complicated, in particular if no formulas for summary characteristics are available. A natural approach is the minimum contrast method, see Gloaguen *et al.* (2006) and Lautensack (2007). There the distributional difference between the tessellation data and model tessellations is characterized by contrast characteristics and these are then minimized. For example, Lautensack (2007) used in the context of a spatial stationary Laguerre tessellations the contrast

$$d = \sum_{i=1}^{8} \left(\frac{\hat{c}_i - c_i}{c_i}\right)^2$$

with

 c_1 (c_2) = mean (variance) of cell volume, c_3 (c_4) = mean (variance) of cell surface, c_5 (c_6) = mean (variance) of average cell width, c_7 (c_8) = mean (variance) of number of faces per cell.

The c_i are the model characteristics and the \hat{c}_i the empirical characteristics. If the c_i can be obtained only by simulation, the Nelder-Mead simplex algorithm may be used for the minimization.

5.3 Bayesian reconstruction of tessellations

In recent years, various papers fitting tessellation models to actual data, using parametric statistical models and a Bayesian Markov chain Monte Carlo (MCMC) approach to inference have appeared. A Bayesian approach is both natural and very useful for many statistical applications of random tessellations, partly because of the complicated structures and models used and partly because some prior knowledge is often available. In contrast a classical/frequentist maximum likelihood approach is in general computationally infeasible.

Some examples of reconstructing unobserved tessellations

Below we consider briefly the work by Blackwell and Møller (2003) where vertices of a Voronoi tessellation are pertubated, and the work by Skare *et al.* (2007) where points of a point process defined on the edge of a Voronoi tessellation are pertubated such that the edges are not directly given. In both papers, the tessellations are unobservable and have to be reconstructed using a Bayesian MCMC approach. These papers consider only application examples of planar and rather small samples, but the ideas used there can be applied also to three-dimensional and much larger samples.

Figure 4 shows an example of a hidden tessellation in a noisy image obtained from a cross-section through a sample of metal; the micro-crystalline structure of the metal may be modelled by a tessellation. Figure 5 shows two point patterns, where the larger circles indicate locations of badger setts and the smaller dots indicate locations of badger latrines, which play a role in the demarcation of badger territories; these territories may be modelled by the cells of a tessellation, where the latrines tend to occur close to the edges of the tessellation. Using a Bayesian MCMC-based approach to inference, Blackwell and Møller (2003) show how to reconstruct the unobserved tessellations in Figures 4 and 5, and how to indicate the uncertainty in the reconstruction. Their approach is sketched below.



Fig. 4. A grey-scale image of a cross-section of the austenite grain structure of a steel sample obtained by light microscopy. The pixel size is about $0.5 \,\mu$ m. We use this rather blurred image to show the potential of the reconstruction method.

First, in Blackwell and Møller (2003), the unobserved tessellation is a priori modelled by a deformed tessellation obtained by random pertubations of the vertices of a planar Voronoi tessellation.

Second, conditional on the deformed tessellation, the data are modelled; this is the likelihood term.

Third, certain priors are imposed on the unknown parameters of the likelihood and of the deformed tessellation model.

Fourth, an MCMC algorithm is constructed to sample from the posterior distribution, which contains information about the unobserved deformed tessellation, unobserved nuclei of the Voronoi tessellation, and all remaining unknown parameters. Since MCMC methods are used for estimating the posterior distribution, we only need to specify the posterior density up to proportionality. It is proportional to the likelihood term times the joint prior density for the unobserved deformed tessellation, the nuclei, and the remaining unknown parameters. A major element of the MCMC algorithm is the



Fig. 5. Badger setts and latrines.

reconstruction of the deformed tessellation after a proposed local change of the tessellation.

Fifth, various illuminating graphical representations of the posterior distribution are shown, including how to reconstruct the deformed tessellation and how to determine the uncertainty in the reconstruction.

In the badger example (Figure 5), the point patterns are observed in a rectangular window W, and in order to account for edge effects, we may consider a larger region $S \supset W$, see Figure 6. The nuclei of the Voronoi tessellation are given by the badger setts defined on S, where the observed badger setts are treated as a fixed point pattern, and the unobserved badger setts are modelled by a homogeneous Poisson process on $S \setminus W$. An example of a typical reconstruction of the deformed tessellation is shown in Figure 6, where the non-convex cells are due to the pertubations of the vertices of the underlying Voronoi tessellation.

The Bayesian approach is able to produce many other such reconstructions and helps so to understand the uncertainty in the reconstruction, which is a great advantage of the Bayesian approach; other reconstruction methods usually just provides one tessellation as the final estimate. Figure 7 shows this uncertainty in the form of the posterior edge intensity of the Voronoi tessellation within the observation window.



Fig. 6. Badgers data: An example of a reconstruction allowing for edge effects, where the small rectangle indicates the observation window W and the larger rectangle indicates the region S. The non-convex cells are due to pertubations of the vertices of the underlying Voronoi tessellation.

Returning to the sample of metal in Figure 4, although a planar crosssection through a 3D Voronoi tessellation is not precisely a 2D Voronoi tessellation (Chiu *et al.*, 1996), it makes nevertheless sense to try to reconstruct the (unobserved) true structure of the grains in the cross-section using a deformed Voronoi tessellation. A Bayesian reconstruction of the deformed tessellation is shown in Figure 8; again this estimate could be supplied with a plot indicating the uncertainty in the reconstruction.

In Skare *et al.* (2007), another point process model with high intensity near the edges of a homogeneous Poisson-Voronoi tessellation is constructed. Given the Voronoi tessellation, the point process is generated by random pertubations of the points of an unobserved homogeneous Poisson process defined on the edges of the tessellation. The point process turns out to be an inhomogeneous Poisson process, and priors on the nuclei of the Voronoi tessellation and other model parameters are imposed. Thereby the model can be analyzed in a rather similar Bayesian fashion as in Blackwell and Møller (2003), using an MCMC algorithm to sample from the posterior, which contains information about the unobserved Voronoi tessellation and the model parameters. Fur-



Fig. 7. Badgers data: Locations of the badger latrines together with a gray scale plot of the posterior edge intensity.



Fig. 8. Sample of metal: The posterior modal reconstruction.

27

ther, it is demonstrated how posterior predictive distributions can be used for model control. Moreover, a simulation study, the 2D application of the badger dataset considered above, and a 3D application in material science (alumina grain structure) are presented.

Further related work

In Green (1995), reversible jump Markov chain Monte Carlo has been developed and applied to image segmentation (subdivision of a digital image into homogeneous regions) via Voronoi tessellations. Heikkinen and Arjas (1998, 1999) studied a non-parametric Bayesian modelling framework for inhomogeneous Poisson processes where the intensity function is piecewise constant on Voronoi cells. In Møller and Skare (2001), Voronoi tessellations have been used in a Bayesian setting for reservoir modelling. Blackwell (2001) considered a Bayesian setting with Voronoi tessellations for modelling animal territories.

Other related statistical work uses the Voronoi tessellation as a way of defining a correlation structure in a spatial model while allowing discontinuities and anisotropy; see e.g. Denison *et al.* (2002). Such 'partition models' have been widely used in e.g. spatial epidemiology. In contrast with the models above, partition models typically assume that mean response does not depend on location within a cell, and the tessellation itself does not necessarily have any direct interpretation within a specific application. In a very different context, the pertubed vertices of a Voronoi tessellation generated by a Poisson process in 3 dimensions have been used as a point process to model the locations of galaxies, see Snethlage *et al.* (2002). The interest there is in the pertubated vertices themselves, however, not in any pertubated tessellation nor in statistical inference.

The abovementioned papers are all related to particular applications, and there is indeed scope for a further development of Bayesian MCMC methods for random tessellations.

6 Stereology for tessellations

In this section, we consider stationary three-dimensional tessellations. Stereology is a 'toolbox' of methods for obtaining three-dimensional information from one- or two-dimensional data, obtained for example by beams or planar sections. Many three-dimensional tessellation characteristics can be obtained by means of stereological methods but the most fundamental ones, ρ_3 and V_3 , are notoriously difficult to estimate from planar sections. Given such data, one approach is to use the formula

$$N_A = \rho_3 B_3,$$

where N_A is the mean number of cell profiles per unit area and B_3 is the mean average breadth. In general B_3 is impossible to estimate from planar sections,

so the formula and other similar approaches based on stereological results can only be used for systems of spherical particles and other particles of fixed shape, and they lead to ill-posed integral equations, see Stoyan *et al.* (1995). For random tessellation characteristics, this approach is useless. Alternative approaches are based on serial sections (Baddeley and Jensen, 2005; Howard and Reed, 2004; Liu *et al.*, 1994) and modern measurement techniques using computerized tomography, but difficult technical problems appear.

6.1 One-dimensional samples

Consider the intersection between an arbitrary line and a motion invariant tessellation. The lengths of intervals given by the intersection points of the line and the cell faces determine what is called the chord length distribution (a formal definition involves the use of Palm measures). The mean chord length \bar{l} , appears in important formulae:

$$\rho_3 = \frac{4}{S_3 \cdot \overline{l}}, \quad S_V = 2/\overline{l}.$$

In practice usually a system of lines is used to provide chord length data, whereby \overline{l} can be estimated. Furthermore, for the particular case of a Poisson-Voronoi tessellation,

$$\bar{l} = 0.687 \rho^{-\frac{1}{3}}$$

which can be used to estimate ρ from the estimate of \overline{l} .

6.2 Planar sections

The result of a planar section of a spatial tessellation is a planar tessellation. Its first order characteristics are N_A , L_A (= mean cell edge length per unit area) and P_A (= mean number of profile vertices per unit area). These satisfy the following fundamental stereological formulae in the motion invariant case:

$$N_A = \rho_3 B_3$$
, $L_A = \frac{\pi}{4} S_V$, $P_A = \frac{1}{2} L_V$.

Thus S_V and L_V can be conveniently determined by means of planar information, whereas ρ_3 can usually not be obtained from planar sections since B_3 is unknown.

In the particular case of a Poisson-Voronoi tessellation the formulae above can be replaced by expressions which contain only the tessellation parameter ρ ,

$$N_A = 1.46\rho^{2/3}$$
, $L_A = 2.29\rho^{1/3}$, $P_A = 2.92\rho^{2/3}$.

These formulas lead to estimators of ρ , where that based on P_A might be preferred. Information on these estimators and on related tests of the Poisson-Voronoi hypothesis can be found in Stoyan *et al.* (1995), p. 374. There it is recommended to estimate ρ_3 for general stationary tessellations by means of

$$\hat{\rho}_3 = 0.566 \hat{N}_A^{3/2}$$
,

using the working assumption that the Poisson-Voronoi tessellation can be used as an approximation.

In metallography more precise methods have been developed. According to U.S. standards (the last version is ASTM Standard E-112) for steel grains

$$\hat{
ho}_3 = 0.8 \hat{N}_A^{\frac{3}{2}}, \quad \hat{
ho}_3 = 0.5659 \overline{l}^{-3}.$$

Horálek (1988) showed that these estimators are unbiased for a particular class of Poisson-Johnson-Mehl tessellation models. Saxl & Ponižil (2001) studied the grain-size estimation problem in materials in more detail and developed the so-called w-s diagram, which leads to fair approximate estimates of mean cell volume V_3 based on \hat{N}_A, \bar{l} and the coefficients of variation of section cell area and chord length.

The papers Schwertel and Stamm (1997) and Coster et al. (2005) are worth reading case studies for the application of stereological methods for tessellations in the context of materials science. There image-processing methods are first used to obtain from rough data images of a quality suitable for statistical methods. Next Poisson-Voronoi and Poisson-Johnson-Mehl tessellations are fitted to these refined data, using the methods discussed above and in Section 5.2.

7 Simulation procedures

In previous sections, mathematically tractable properties of Poisson-Voronoi and Poisson-Delaunay tessellations were outlined. Further analysis of these and other kind of random tessellation models requires Monte Carlo studies. In the case of a tessellation defined on an unbounded region, it is important to account for edge effects, i.e. not to forget that what happens outside a bounded simulation window may effect what happens within the window.

Most simulation studies of Voronoi tessellations are concerned with the Poisson case. In a large scale study, Hinde and Miles (1980) approximated the polygonal characteristics of a typical planar Poisson-Voronoi cell by Monte Carlo methods. A more efficient method is based on combining (6) with the radial simulation algorithm for the stationary Poisson process in \mathbb{R}^d as introduced in Quine and Watson (1984). This algorithm is also tailormade for simulating a Voronoi tessellation within a ball when the nuclei come from a stationary Poisson process or a related point process obtained by thinning or clustering such as Matérn hard core processes and Poisson cluster processes with uniformly bounded clusters (including the Matérn cluster process). For such processes the problem with edge effects can be avoided, while for for Poisson cluster processes with not necessarily uniformly bounded clusters, results in Møller (2003) are useful for evaluating the edge effects. Moreover, Quine

29

and Watson's algorithm can be extended to apply for Poisson-Johnson-Mehl tessellations, see Møller (1995). Several simulated results are shown in Hinde and Miles (1980), Quine and Watson (1984), Hermann *et al.* (1989), Møller *et al.* (1989), Lorz (1990), Lorz and Hahn (1993), Møller (1994, 1995), Van de Weygaert (1994), and Okabe *et al.* (2000).

The Bayesian MCMC algorithms used in the papers mentioned in Section 5.3 depend much on the specific models and problems. Briefly, they are hybrid Metropolis-Hastings algorithms, with separate types of updates for the various kind of parameters. For the point process updates, a birth-deathmove algorithm is used (Geyer and Møller, 1994), and for the other kind of parameters Gibbs or Metropolis random walk updates are used.

Acknowledgments

We thank Lutz Muche for valuable remarks on an earlier version of this paper and for providing Figures 1 and 2, and Joachim Ohser for providing Figure 4. JM was supported by the Danish Natural Science Research Council, grant no. 272-06-0442 ('Point process modelling and statistical inference').

References

- 1. A. Baddeley, E. B. Vedel Jensen: *Stereology of Statisticians* (Chapman & Hall/CRC, Boca Raton, 2005)
- 2. A. Baddeley, J. Møller: Int. Statist. Rev. 2, 89 (1989)
- 3. V. Baumstark, G. Last: Adv. Appl. Probab. 37, 279 (2006)
- 4. P.G. Blackwell: Biometrics 57, 502 (2001)
- 5. P.G. Blackwell, J. Møller: Adv. Appl. Probab. 35, 4 (2003)
- 6. E. Bertin, J.-M. Billiot, R. Drouilhet: Stochastic Models 15, 181 (1999a)
- 7. E. Bertin, J.-M. Billiot, R. Drouilhet: Adv. Appl. Probab. 31, 895 (1999b)
- 8. K.A. Brakke: Statistics of random plane Voronoi tessellations, Department of Mathematical Sciences, Susquehanna University (Manuscript 1987a)
- K.A. Brakke: Statistics of three dimensional random Voronoi tessellations, Department of Mathematical Sciences, Susquehanna University (Manuscript 1987b)
- 10. P. Calka: Adv. Appl. Probab. 34, 702 (2002)
- 11. P. Calka: Adv. Appl. Probab. 35, 551 (2003)
- 12. S.N. Chiu, R. van de Weygaert, D. Stoyan: Adv. Appl. Probab. 28, 356 (1996)
- 13. R. Collins: J. Phys. C 1, 1461 (1968)
- M. Coster, X. Arnould, J.-L. Chermant, A. E. Moataz, T. Chartier: Image Anal. Stereol. 24, 105 (2005)
- 15. D.J. Daley, D. Vere-Jones: An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods, 2nd edn (Springer, New York 2003)
- D.G.T. Denison, C.C. Holmes, B.K. Mallick et al: Bayesian Methods for Nonlinear Classification and Regression (Wiley, New York, 2002)
- 17. G.L. Dirichlet: J. Reine und Angew. Math. 40, 209 (1850)
- 18. H.-O. Georgii: Comm. Math. Phys. 48, 31 (1976)

31

- 19. E.N. Gilbert: Ann. Math. Statist. **33**, 958 (1962)
- C. Gloaguen, F. Fleischer, H. Schmidt, V. Schmidt: Telecommunciation Systems 31, 353 (2006)
- 21. C.J. Geyer, J. Møller: Scand. J. Statist. 21, 359 (1994)
- 22. P.J. Green: Biometrika 82, 711 (1995)
- 23. M. P. Hayen, M. P. Quine: Adv. Appl. Probab. 34, 281 (2002)
- 24. J. Heikkinen, E. Arjas: Scand. J. Statist. 25, 435 (1998)
- 25. J. Heikkinen, E. Arjas: Biometrics 55, 738 (1999)
- 26. L. Heinrich: Adv. Appl. Probab. 30, 603 (1998)
- 27. L. Heinrich, R. Körner, N. Mehlhorn, L. Muche: Statistics 31, 235 (1998)
- 28. H. Hermann, H. Wendrock, D. Stoyan: Metallography 23, 189 (1989)
- 29. A.L. Hinde, R.E. Miles: J. Statist. Comput. Simul. 10, 205 (1980)
- 30. V. Horálek: Adv. Appl. Probab. 20, 684 (1988)
- C. V. Howard, M. G. Reed: Unbiased Stereology: Three-dimensional Measurement in Microscopy, 2nd ed. Bios Scientific Publ., Oxford, 2004
- 32. M. Hug, M. Reitzner, R. Schneider: Adv. Appl. Probab. 36, 667 (2004)
- 33. Y. Isokawa: Adv. Appl. Probab. **32**, 648 (2000)
- 34. W.A. Johnson, R.F. Mehl: Trans. Amer. Inst. Min. Engrs. 135, 416, (1939)
- 35. D.G. Kendall: Statist. Sci. 4, 87 (1989)
- 36. W.S. Kendall: J. Appl. Probab. 28, 767 (1990)
- 37. C. Lautensack: *Random Laguerre Tessellations* (Verlag Lautensack, Weiler bei Bingen 2007)
- 38. G. Liu, H. Wu, W. Li: Acta Stereologica 13, 281 (1994)
- K. Lochmann, A. Anikeenko, A. Elsner, N. Medvedev, D. Stoyan: Eur. Phys. J. B. 53, 67 (2006a)
- 40. K. Lochmann, L. Oger, D. Stoyan: Solid State Sciences 8, 1397 (2006b)
- 41. U. Lorz: Materials Char. 3, 297 (1990)
- 42. U. Lorz, U. Hahn: Preprint 93-05, Fachbereich Mathematik, Technische Universität Bergakademie Freiberg (1993)
- 43. J. Mecke: Palm methods for stationary random mosaics. In: Combinatorial Principles in Stochastic Geometry, ed by R.V. Ambartzumian (Armenian Academy of Science Publishing House, Erevan 1980) pp 124–132
- 44. J. Mecke: Math. Operationsf. Statist. Ser. Statist. 15, 437 (1984)
- 45. J. Mecke: Pattern Recognition 232, 1645 (1999)
- 46. J. Mecke, D. Stoyan: Adv. Appl. Probab. 33, 576 (2001)
- 47. J.L. Meijering: Philips Research Reports 8, 270 (1953)
- 48. R.E. Miles: Math. Biosci. 6, 85 (1970)
- 49. R.E. Miles: On the homogeneous planar Poisson point process. In: *Stochastic Geometry*, ed by E.F. Harding, D.G. Kendall (Wiley, London 1974) pp 202–227
- 50. R.E. Miles, R.L. Maillardet: J. Appl. Probab. 19A, 97, (1982)
- 51. L. Muche: Acta Stereologica **12**, 125 (1993)
- 52. L. Muche: J. Statist. Physics 84, 147 (1996a)
- 53. L. Muche: Math. Nachr. **178**, 125 (1996b)
- 54. L. Muche: Math. Nachr. 191, 247 (1998)
- 55. L. Muche: Adv. Appl. Probab. 37, 279 (2005)
- 56. L. Muche: Delaunay and Voronoi tessellations: Minkowski functionals and edges. In: Proc. S4G (Internat. Conf. Stereology, Spatial Statist, Stoch Geom; Prague, June 1999), ed by V. Beneš, J. Janaček, I. Saxl (Union Czech Mathematicians and Physicists, Prague 1999) pp 21-30

- 32 Jesper Møller and Dietrich Stoyan
- 57. L. Muche, D. Stoyan: Adv. Appl. Probab. 29, 467 (1992)
- 58. J. Møller: Adv. Appl. Probab. 21, 73 (1989)
- 59. J. Møller: Adv. Appl. Probab. 24, 814 (1992)
- J. Møller: Lectures on Random Voronoi Tessellations (Lecture Notes in Statistics 87, Springer, New York 1994)
- 61. J. Møller: Adv. Appl. Probab. 27, 367 (1995)
- 62. J. Møller: Topics in Voronoi and Johnson-Mehl tessellations. In: Stochastic Geometry: Likelihood and Computation, ed by O.E. Barndorff-Nielsen, W.S. Kendall, M.N.M. van Lieshout (Chapman and Hall/CRC, Boca Raton 1999) pp 173–198
- 63. J. Møller: Adv. Appl. Probab. 35, 4 (2003)
- J. Møller, E.B. Jensen, J. S. Petersen et al: Research Report 182, Department of Theoretical Statistics, University of Aarhus (1989)
- J. Møller, K. Helisova: Research Report R-2207-15, Department of Mathematical Sciences, Aalborg University (2007)
- 66. J. Møller, Ø. Skare: Statist. Modelling 1, 213 (2001)
- J. Møller, R.P. Waagepetersen: Statistical Inference and Simulation for Spatial Point Processes (Chapman and Hall/CRC, Boca Raton 2003)
- 68. J. Møller, S. Zuyev: Adv. Appl. Probab. 28, 662 (1996)
- 69. X.X. Nguyen and H. Zessin: Math. Nachr. 88, 105 (1979)
- 70. A. Okabe, B. Boots, K. Sugihara et al: Spatial Tessellations. Concepts and Applications of Voronoi Diagrams, 2nd edn (Wiley, Chichester 2000)
- 71. M.P. Quine, D.F. Watson: J. Appl. Probab. **21**, 548 (1984)
- 72. W. Radecke: Math. Nachr. **97**, 203 (1980)
- 73. P.N. Rathie: J. Appl. Probab. **29**, 740 (1992)
- 74. I. Saxl and P. Ponižil: Materials Characterization 46, 113 (2001)
- 75. T.H. Scheike: Adv. Appl. Probab. 26, 43 (1994)
- 76. M. Schlather: Math. Nachr. 214, 113 (2000)
- 77. J. Schwertel, H. Stamm: J. Microcopy 186, 198 (1997)
- 78. R. Sibson: Scand. J. Statist. 7, 14 (1980)
- M. Snethlage, V. J. Martinez, D. Stoyan, E. Saar: Astron. Astrophys. 388, 758 (2002)
- 80. Ø. Skare, J. Møller, E.B.V. Jensen: Statist. Comput. 17, 369 (2007)
- D. Stoyan, W.S. Kendall, J. Mecke: Stochastic Geometry and its Applications, 2nd edn (Wiley, Chichester 1995)
- 82. R. van de Weygaert: Astron. Astrophys. 283, 361 (1994)
- 83. G. Voronoi: J. Reine und Angew. Math. 134, 198 (1908)
- 84. M. Zähle: Ann. Probab. 16, 1742 (1988)
- 85. S.A. Zuyev: Random Struct. Alg. 3, 149 (1992)