Applied Probability Trust (April 24, 2002)

ESTIMATION OF THE DIRECTIONAL MEASURE OF PLANAR RANDOM SETS BY DIGITIZATION

MARKUS KIDERLEN, EVA B. VEDEL JENSEN,*

Abstract

Estimation methods for the directional measure of a stationary planar random set Z, based only on discretized realizations of Z, are discussed. Properties of the discretized set that can be derived by comparing neighbouring grid points are used. Larger grid configurations of more than two grid points are considered. It is shown that the probabilities of observing the various types of configurations can be expressed in terms of the first contact distribution function of Z (with a finite structuring element). An important prerequisite result concerning deterministic dilation areas is also established. The inference on the mean normal measure based on 2×2 configurations is discussed in detail.

MEAN NORMAL MEASURE; ORIENTED ROSE OF NORMAL DIRECTIONS; DIGITIZATION; CONTACT DISTRIBUTION; POINT CONFIGURATION

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 69D05;62F10
SECONDARY 52A10

1. Introduction

The *mean normal measure* $\overline{S}(Z, \cdot)$ of a stationary random closed set Z in $\mathbb{R}^d$ can be used for detecting and quantifying anisotropy of Z. It is defined under suitable regularity conditions on Z. Its normalized version, the so called *directional distribution* or *(oriented) rose of normal directions*, can be interpreted as the distribution of the outer unit normal at a 'typical' boundary point of Z. The mean normal measure, also called *directional measure*, has been introduced in WEIL [16] and [17].

In the present paper we will focus on the planar case ($d = 2$) and discuss estimation methods for $\overline{S}(Z, \cdot)$ using only discretized realizations of Z. A discretization of a set is its intersection with a scaled regular grid in $\mathbb{R}^2$; we will only consider the intersection with the scaled standard grid $t\mathbb{Z}^2$, $t > 0$, as this is often used in applications. Typically, the grid has to be refined ($t \to 0+$), in order to obtain information about the boundary behaviour of Z. We will only use properties of the discretized set that can be derived by comparing neighbouring grid points, because this information is easily accessible in applications by filtering the discretized set (see OHSER & MÜCKLICH [9]).

A classical result (SERRA [14]) states that the information obtained by comparing *pairs* of neighbouring grid points can be used to estimate the mean length of total projection in directions associated to the discretization. This, in turn, yields certain information about the directional measure. Clearly, the intersection of Z with a pair of points can not yield more information than the intersection of Z with the line passing through these points can do. Therefore the above mentioned procedure cannot yield more information about Z than all intersections of Z with lines do. It is well known (see KIDERLEN [3]) that in general the family of intersections $\{Z \cap g \,|\, g \text{ line in } \mathbb{R}^2\}$

only determines the *un-oriented* or *symmetrized* directional measure

$$\frac{1}{2}(\overline{S}(Z,\cdot)+\overline{S}^*(Z,\cdot))$$

but not the directional measure itself. (Here $\overline{S}^*(Z,\cdot)$ is the reflection of $\overline{S}(Z,\cdot)$ at the origin 0.) This shows that comparison of pairs of grid points is not sufficient for the estimation of $S(Z,\cdot)$. We will therefore consider configurations of more than two grid points. Larger configurations, such as grid–squares of size 2×2 or 3×3, have also been used in Ohser & Mücklich [9] and Ohser et al. [8] to estimate the three Quermass densities of Z (area density, length density and density of the Euler number of Z). It appears that the present paper describes for the first time estimation methods of the (oriented) mean normal measure from larger grid neighbourhood configurations. A related result is mentioned by Rataj in [10]: It states that the mean normal measure of Z is determined, and can be estimated, by a suitable three–point test set and its rotations.

In the present paper, we consider larger grid configurations which carry more precise information about the directional measure. Events of the type $tB\subset Z$, $tW\subset Z^C$ are observed, where tB and tW are finite subsets of the scaled standard grid $t\mathbb{Z}^2$. For instance, $tB\cup tW$ may be a grid square of size 2×2 or 3×3. The probabilities of such events can be expressed in terms of the first contact distribution function F_M of Z

$$F_M(t):=\mathbb{P}(Z\cap tM\neq\emptyset),\quad t\geq 0,$$

where $M\subset\mathbb{R}^2$ is a suitably chosen finite structuring element. A formula for the derivative of F_M at the origin is derived, which holds for stationary random sets with values in the extended convex ring. An important prerequisite result concerning deterministic dilation areas is also established.

The paper is organized as follows. Preliminaries concerning convex geometry, random sets and contact distribution functions are given in section 2. Results concerning dilation areas for deterministic sets (which are either convex, or finite unions of convex sets) are presented in section 3. First contact distribution functions are presented in section 4. The main result is stated in Theorem 4 in section 5. It concerns the probabilities of observing different types of grid configurations. In section 6 this result is applied to 2×2 configurations and used to obtain (parametric) estimators of $\overline{S}(Z,\cdot)$. Perspectives on $n\times n$ configurations for $n>2$ are given in section 7.

2. Preliminaries

The standard scalar product in $\mathbb{R}^2$ will be denoted by $\langle\cdot,\cdot\rangle$. $B^2\subset\mathbb{R}^2$ is the unit ball with respect to the induced Euclidean norm $\|\cdot\|$ and S^1 is the unit circle. We recall some basic facts from convex geometry, details can be found in Schneider's book [12]. The set of convex bodies (compact convex subset of $\mathbb{R}^2$) will be denoted by $\mathcal{K}$. The support function $h(K,\cdot)=\max\{\langle x,\cdot\rangle\,|\,x\in K\}$ of a non–empty set $K\in\mathcal{K}$ is a function on the unit circle S^1. Formally, we put $h(\emptyset,\cdot)\equiv 0$ on S^1. The support function is Minkowski additive, i.e.

$$h(K+K',\cdot)=h(K,\cdot)+h(K',\cdot)$$

for any convex bodies K and K' (here and in the following, the addition of subsets in $\mathbb{R}^2$ is understood pointwise). If $M \subset \mathbb{R}^2$ we will write $\mathrm{conv}M$ for its convex hull.

λ_d will always denote Lebesgue measure in $\mathbb{R}^d$. For K and K' in $\mathcal{K}$, the area of the sum $K + tK'$ obeys a generalization of Steiner's formula

$$\lambda_2(K + tK') = \lambda_2(K) + 2tA(K, K') + t^2\lambda_2(K'), \quad t \geq 0. \tag{2.1}$$

The mixed area $A(K, K')$ occurring here as a coefficient, is symmetric, $A(K, K') = A(K', K)$, and has the homogeneity property

$$A(tK, K') = tA(K, K'), \qquad t \geq 0.$$

We will write $S(K, \cdot)$ for the usual surface area measure on S^1 (as it is of order one, it is often denoted by $S_1(K, \cdot)$, but we omit the subscript as no other surface area measures will occur here). The mixed area obeys

$$2A(K, K') = \int_{S^1} h(K', w)\, S(K, dw) = \int_{S^1} h(K, w)\, S(K', dw). \tag{2.2}$$

The line segment $\mathrm{conv}\{x, y\}$ with endpoints x and y in $\mathbb{R}^2$ will be denoted by $[x, y]$, for short. (2.2) implies

$$2A(K, [0, u]) = \lambda_1(K|u^\perp), \tag{2.3}$$

where $u \in S^1$ and $K|u^\perp$ is the orthogonal projection of K on the line through 0, orthogonal to u.

As both, the support function and the surface area measure are additive and continuous on $\mathcal{K}$ (with respect to the Hausdorff metric), they have a unique additive extension to the convex ring $\mathcal{R}$ (the family of finite unions of convex bodies). We will use the same notation for these extensions as we used for the corresponding functionals on $\mathcal{K}$. Using these extensions, (2.2) defines the mixed area on $\mathcal{R} \times \mathcal{R}$, which is additive in each of its arguments.

Let Z be a stationary random set in $\mathbb{R}^2$ with values in the extended convex ring (i.e. the intersection of Z with a convex body is almost surely an element of $\mathcal{R}$). For this and further notions from stochastic geometry we refer to Schneider & Weil [13]. We assume throughout the following that Z satisfies the integrability condition

$$\mathbb{E}2^{N(Z \cap K)} < \infty \tag{2.4}$$

for all $K \in \mathcal{K}$, where $N(L)$ of a nonempty set $L \in \mathcal{R}$ is the minimal number $k \in \mathbb{N}$ such that $L = \bigcup_{i=1}^k K_i$ with $K_i \in \mathcal{K}$ and $N(\emptyset) = 0$.

The mean normal measure of Z is defined by

$$\overline{S}(Z, \cdot) = \lim_{r \to \infty} \frac{S(Z \cap rK, \cdot)}{\lambda_2(rK)}.$$

This definition is independent of the choice of the convex body $K \in \mathcal{K}_0 := \{K \in \mathcal{K} \mid \mathrm{int}K \neq \emptyset\}$. As

$$\int_{S^1} u\, \overline{S}(Z, du) = 0, \tag{2.5}$$

Minkowski's existence Theorem shows that there is a non-empty convex body $B(Z) \in \mathcal{K}$ with

$$\overline{S}(Z, \cdot) = S(B(Z), \cdot). \tag{2.6}$$

Under the additional assumption that this body is centered (i.e. its Steiner point coincides with 0), $B(Z)$ is uniquely determined and called the *mean Blaschke body of* Z. This convex body can be used to visualize the mean normal measure.

A standard example of a random set Z is the *stationary Boolean model.* It is obtained as follows: First, an ordinary Poisson point process (the germ–process) with intensity $\gamma > 0$ is generated. It is characterized by the fact that the number of points in any bounded Borel set $B \subset \mathbb{R}^2$ is Poisson distributed, with mean $\gamma\lambda_2(B)$. Then, convex grains are attached to the germs. They are independent copies of a random convex body K_0 and independent of the germ-process. The distribution of the *typical grain* K_0 is assumed to be concentrated on the family of centered convex bodies (e.g. those with Steiner point in 0). The union set of this particle process is called a Boolean model (with convex grains). Its distribution is uniquely determined by the intensity γ and the typical grain K_0. More geometrically, it is also determined by the capacity functional T_Z, where

$$T_Z(M) := \mathbb{P}(Z \cap M \neq \emptyset), \qquad \text{for compact } M \subset \mathbb{R}^2.$$

Due to the independence properties of a Boolean model, we have

$$T_Z(M) = 1 - \exp\left(-\gamma \mathbb{E}\lambda_2(K_0 + \check{M})\right), \tag{2.7}$$

where $\check{M}$ is the set M reflected at the origin.

The mean normal measure of a Boolean model with intensity γ and typical convex grain K_0 obeys

$$\overline{S}(Z, \cdot) = \gamma(1 - \overline{A}(Z))\mathbb{E}S(K_0, \cdot),$$

where $\overline{A}(Z)$ is the area density of Z. If the *mean typical particle* $\mathbb{E}K_0 \in \mathcal{K}$ (a set valued mean), is defined by

$$S(\mathbb{E}K_0, \cdot) = \mathbb{E}S(K_0, \cdot),$$

with the additional condition that this body is centered, the last equation and (2.6) imply

$$B(Z) = \gamma(1 - \overline{A}(Z))\mathbb{E}K_0, \tag{2.8}$$

which shows that $B(Z)$ equals the mean typical particle up to scaling.

We return to general random sets and recall the definition of the *first contact distribution function* F_M of Z with respect to the compact structuring element M:

$$F_M(t) := \mathbb{P}(Z \cap tM \neq \emptyset) = T_Z(tM), \qquad t \geq 0. \tag{2.9}$$

Under the assumption $0 \in M$, this function is closely related to *the contact distribution function* $H_M(t) := \mathbb{P}(Z \cap tM \neq \emptyset \,|\, 0 \notin Z)$:

$$(1 - F_M(0))(1 - H_M(t)) = 1 - F_M(t), \qquad t \geq 0.$$

If the set M is star–shaped with respect to 0, we have

$$(2.10) \qquad F_M(t) = \mathbb{P}(d_M(Z) \le t), \qquad H_M(t) = \mathbb{P}(d_M(Z) \le t \,|\, 0 \notin Z)$$

and hence F_M and H_M are (usual resp. conditioned) distribution functions of the random variable $d_M(Z) := \inf\{s \ge 0 \,|\, Z \cap sM \ne \emptyset\}$. Usually, M is assumed to be a convex body with $0 \in M$ (see the survey Hug et al. [2] on contact distribution functions). In this setting, (2.10) can be used equivalently as definition. For arbitrary nonempty compact structuring elements M, however, (2.10) does not hold. We have $d_M(Z) = d_{\mathrm{star}M}(Z)$, where

$$\mathrm{star}M := \bigcup_{x \in M} [0, x]$$

is the star–hull of M (with respect to 0). Hence

$$(2.11) \qquad \mathbb{P}(d_M(Z) \le t) = F_{\mathrm{star}M}(t), \qquad \mathbb{P}(d_M(Z) \le t \,|\, 0 \notin Z) = H_{\mathrm{star}M}(t).$$

In present context, we will consider *finite* structuring elements M, more precisely, finite subsets of $\mathbb{Z}^2$. As a consequence, F_M and H_M need not be distribution functions – they need not even be monotonic.

Due to the stationarity of Z, we have for arbitrary compact M and $t \ge 0$

$$F_M(t) = \mathbb{P}(0 \in Z + t\check{M}) = \mathbb{E}\lambda_2((Z + t\check{M}) \cap [0,1]^2).$$

We will therefore examine dilation areas more closely. In the next section we will consider the deterministic setting for elements of $\mathcal{K}$ and $\mathcal{R}$. In section 4, the obtained results will be extended to random sets.

3. The deterministic dilation area

3.1. *Results for convex bodies.* As a basic tool, we will use the *dilation area* of a convex body K which is a function on the family of Borel sets in $\mathbb{R}^2$ given by

$$M \mapsto \lambda_2(K + M).$$

If $M = \{0, x\}$, this function depends on $x \in \mathbb{R}^2$ only and is called the variogram of K. We will be interested in 'directional derivatives' of the dilation area at $\{0\}$ and therefore we introduce the function

$$(3.1) \qquad \psi_K(M, t) := \lambda_2(K + tM), \qquad t \ge 0.$$

Mathéron [5] showed that the directional derivative of the variogram in direction $u \in S^1$ in the point $x = 0$ is equal to $\lambda_1(K|u^\perp)$. Equivalently we have according to (2.1) and (2.3)

$$\psi'_K(\{0, u\}, 0+) = 2A(K, [0, u]) = \psi'_K([0, u], 0+),$$

where $\psi'_K(M, 0+)$ denotes the right sided derivative of $t \mapsto \psi_K(M, t)$ at $t = 0$. This shows that for small t, the function $\psi_K(\{0, u\}, t)$ can approximately be calculated by replacing the two point set $\{0,u\}$ with its convex hull $[0,u]$. In the following, this result will by generalized by replacing the set $\{0,u\}$ with some finite set M. In addition, an error estimate will be given. For its formulation, two constants c_0, c_M associated to a finite set $M \subset \mathbb{R}^2$ with at least two elements are needed. Put

$$c_0 := \min\{\frac{\sqrt{3}}{\mathrm{diam} M}, \frac{2}{l}\},$$

where l is the length of the longest edge of the convex polygon $\mathrm{conv} M$. Furthermore let c_M be the sum of all squared lengths of edges of $\mathrm{conv} M$ (if $\mathrm{conv} M$ is a line segment of length l, we set $c_M := 2l^2$).

Theorem 1 Let $K \in \mathcal{K}$ with $rB^2 \subset K \subset RB^2$, $0 < r \leq R$, and a finite set $M \subset \mathbb{R}^2$ with at least two points be given. Then

$$\psi_K(\mathrm{conv} M, t) - \frac{c_M}{4}\frac{R}{r}t^2 \leq \psi_K(M,t) \leq \psi_K(\mathrm{conv} M, t), \qquad 0 \leq t \leq c_0 r. \tag{3.2}$$

In particular, this implies

$$\psi'_K(M, 0+) = \psi'_K(\mathrm{conv} M, 0+) = 2A(K, \mathrm{conv} M). \tag{3.3}$$

Proof. (3.3) is a direct consequence of (3.2) and (2.1). As $K + tM \subset K + t\mathrm{conv} M$, we only have to show the first inequality in (3.2). The convex polygon $\mathrm{conv} M$ will be denoted by P in the following. As $rB^2 \subset K$ we have $tP \subset K + tM$ as long as $0 \leq t \leq \frac{\sqrt{3}r}{\mathrm{diam} M}$.

We restrict to these t–values from here on excluding moreover the trivial case $t = 0$. Let $\mathrm{Nor} P$ be the set of all outer unit normals of boundary segments of the polygon P. For $u \in \mathrm{Nor} P$ let z_u, $z'_u \in M$ be the endpoints of this boundary segment and H_u the closed half plane with $\mathrm{bd} H_u = \mathrm{aff}\{z_u, z'_u\}$ and $z_u + u \in H_u$. We will now show

$$(K + tP)\backslash(tP) \subset \bigcup_{u \in \mathrm{Nor} P} (tH_u \cap (K + t[z_u, z'_u])). \tag{3.4}$$

If $x \in (K + tP)\backslash(tP)$, then

$$(x + s\check{K}) \cap (tP) \neq \emptyset$$

for $s = 1$. Let $s_0 > 0$ be the minimal number such that this intersection is nonempty and let y be in $(x + s_0\check{K}) \cap (tP)$. The point y is a boundary point of tP. If y lies in the relative interior of some boundary segment of tP, we have $y \in t[z_u, z'_u]$ for a suitable $u \in \mathrm{Nor} P$. It is easily seen that $x \in tH_u$. But $x \in y + s_0 K \subset y + K \subset K + t[z_u, z'_u]$ and (3.4) is shown in this case. If y is a vertex of tP, then $y \in M$ and $x \in y + K$. As s_0 is minimal, one of the normals u of the two boundary segments with endpoint y fulfils $x \in tH_u$. This completes the proof of (3.4).

Using (3.4) we get

$$\lambda_2(K + t\mathrm{conv} M) - \lambda_2(K + tM) \leq \sum_{u \in \mathrm{Nor} P} \lambda_2(D_u). \tag{3.5}$$

where $D_u = tH_u \cap (K + t[z_u, z'_u]) \backslash (K + t\{z_u, z'_u\})$. The set D_u is sketched in Figure 1. Let the point y be an element of the support set of $K + tz_u$ in direction u. Then $y' = y - tz_u + tz'_u$ is an element of the corresponding support set of $K + tz'_u$ in the same direction.

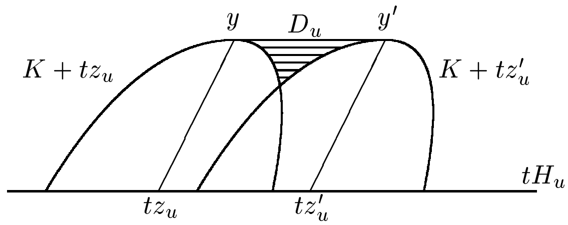

Figure 1: Two translated versions of K and the region D_u (dashed).

As rB^2 is a subset of the convex body K, the triangle $\operatorname{conv}\{tz_u, y, w\}$ lies in $tz_u + K$, where $w := tz_u + r\frac{z'_u - z_u}{\|z'_u - z_u\|}$. Similarly we have $\operatorname{conv}\{tz'_u, y', w'\} \subset tz'_u + K$, where $w' := tz'_u - r\frac{z'_u - z_u}{\|z'_u - z_u\|}$. If $2r \geq t \parallel z_u - z'_u \parallel$, then there is a point $x \in [y, w] \cap [y', w']$ and $D_u \subset T_u$, where the triangle T_u is the convex hull of x, y, y' (see Figure 2).

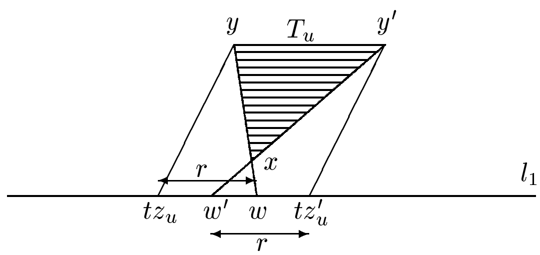

Figure 2: The triangle T_u.

Without loss of generality we can assume that the line $l_1 := t\,\mathrm{bd}H_u$ through tz_u and tz'_u is perpendicular to the line l_2 through y and tz_u. (Otherwise an affine transformation with determinant 1 can be applied which leaves l_1 pointwise fixed and maps l_2 to a line perpendicular to l_1.) Planar geometry now easily yields

$$\lambda_2(D_u) \leq \lambda_2(T_u) = \frac{\parallel y - tz_u \parallel}{4r} \parallel tz_u - tz'_u \parallel^2 \leq t^2 \frac{R}{4r} \parallel z_u - z'_u \parallel^2,$$

where we used $K \subset RB^2$ for the last inequality. Substituting this into (3.5) yields the desired result (3.2).

The quality of the left hand estimate in (3.2) depends on the quotient R/r. Minimization of this quotient naturally leads to the following shape ratio $q(K)$ of a convex body K. To define it, put $R(K) := \min\{t \geq 0 \,|\, K \subset tB^2\}$ and $r(K) := \sup\{t \geq 0 \,|\, tB^2 \subset K\}$ (with $\sup \emptyset := -\infty$). The shape ratio

$$q(K) := \inf\{\frac{R(K-z)}{r(K-z)} \,|\, z \in \operatorname{int}K\} \tag{3.6}$$

($\inf \emptyset := \infty$) describes the deviation of K from circular shape: $1 \le q(K) \le \infty$, with $1 = q(K)$ if and only if K is a circular disc. Note that $q(K) < \infty$ if and only if K has interior points. For sufficiently small $t > 0$, the ratio R/r in (3.2) can be replaced by $q(K)$.

Even if (3.2) is formulated with $q(K)$, the lower bound gives reasonable estimates only in the case where the body K is not too elongated. It is therefore worth mentioning the following estimate in the case where the convex body K is known to be ε-smooth for some $\varepsilon > 0$ (i.e. for every boundary point x of K there is a disc B with radius ε such that $x \in B \subset K$). With the notation of the foregoing proof we fix $u \in \mathrm{Nor} P$ and assume $2\varepsilon \ge t \parallel z_u - z'_u \parallel =: s$. Then D_u is included in a triangle $\tilde{T}_u$ with basis length s and height $\varepsilon - \sqrt{\varepsilon^2 - (s/2)^2}$. This implies

$$\lambda_2(D_u) \le \lambda_2(\tilde{T}_u) = \frac{s}{2}(\varepsilon - \sqrt{\varepsilon^2 - (s/2)^2}) \le \frac{s^2}{4}.$$

Thus,

$$\psi_K(\mathrm{conv} M, t) - \frac{c_M}{4} t^2 \le \psi_K(M, t) \le \psi_K(\mathrm{conv} M, t) \tag{3.7}$$

holds for all $0 \le t \le c_0 \varepsilon$. The constants c_0 and c_M are defined before Theorem 1. The estimate in (3.7) is independent of ε and is therefore better than (3.2). Note, however, that (3.2) holds for arbitrary convex bodies, in particular for polygons.

Due to (2.2), formula (3.3) can be reformulated as follows:

$$\psi'_K(M, 0+) = \int_{S^1} h(\mathrm{conv} M, w)\, S(K, dw). \tag{3.8}$$

3.2. *Extension to the convex ring.* The definition (3.1) of $\psi_K(M, \cdot)$ is also valid for $K \in \mathcal{R}$. We now show that (3.3) extends to the convex ring, using an idea of Rataj [11]. It is clear that (3.3) does not hold for arbitrary M in the case of a line segment K. More generally, an extension of (3.3) to the convex ring can only be true, if K can be written as a finite union of full dimensional convex bodies. The latter condition is equivalent to saying that K is *topologically regular*, which means $K = \mathrm{cl}(\mathrm{int} K)$.

Theorem 2 Let $K \in \mathcal{R}$ be topologically regular and $\emptyset \ne M \subset \mathbb{R}^2$ a finite set. Then

$$\psi'_K(M, 0+) = 2A(K, \mathrm{conv} M). \tag{3.9}$$

Proof. Let $K \in \mathcal{R}$ be topologically regular and M a finite set with at least two elements. Without loss of generality we may assume $0 \in M$. In a first step, we show

$$\lim_{t \to 0+} \frac{1}{t} \lambda_2((K + tL) \setminus K) = 2A(K, L) \tag{3.10}$$

for any convex body L with $0 \in L$. In the case of strictly convex $L \in \mathcal{K}$, (3.10) is a consequence of Theorem 3.3 in [1], which states a local Steiner-type formula for sets in the extended convex ring in Minkowski space. Also, Theorems 3.9 and 2.3 of this work are used, to simplify the obtained expression. Both sides of (3.10) are monotonic in L with respect to set inclusion, so (3.10) can be shown for all convex bodies L with interior points by approximation. To proof (3.10) for line segments L, we may assume $\lambda_1(L) = 1$ and $0 \in L \subset u^\perp$ for $u \in S^1$. Fubini's Theorem implies

$$\frac{1}{t}\lambda_2((K+tL)\backslash K) =$$
$$= \int_{-\infty}^{\infty} 1/t \left[\lambda_1\left(K \cap (su+u^\perp) + tL\right) - \lambda_1\left(K \cap (su+u^\perp)\right)\right] \lambda_1(ds).$$

For every $s \in \mathbb{R}$, the integrand is majorized by, and converges to the Euler characteristic $\chi(K \cap (su+u^\perp))$ of $K \cap (su+u^\perp)$. This shows

$$\begin{aligned} \lim_{t\to 0+} \frac{1}{t}\lambda_2((K+tL)\backslash K) &= \int_{-\infty}^{\infty} \chi(K \cap (su+u^\perp))\lambda_1(ds) \\ &= 2A(K,L). \end{aligned}$$

The last equality follows for convex K from (2.3) and for arbitrary $K \in \mathcal{R}$ by additivity. Thus, (3.10) is shown.

Due to (3.10), it is enough to proof

$$\lambda_2((K + t\mathrm{conv}M)\backslash(K+tM)) = o(t), \qquad t \to 0+. \tag{3.11}$$

Consider a representation $K = \bigcup_{i=1}^n K_i$ of K with topologically regular convex bodies $K_1, \ldots, K_n$. Assume

$$0 < t < \min\{\mathrm{diam}(K_1), \ldots, \mathrm{diam}(K_n)\}/\mathrm{diam}(M),$$

which implies $K_i \not\subset x + t\,\mathrm{conv}\check{M}$ for all $x \in \mathbb{R}^2$ and $i \in \{1, \ldots, n\}$. Consider x in the set of the left hand side of (3.11), which means that $x + t\mathrm{conv}\check{M}$ hits one of the sets K_i, whereas $x + t\check{M}$ hits none of them. Due to the topological regularity of K_i, one of the sides of the polygon $x + t\mathrm{conv}\check{M}$ must be hit twice by $\mathrm{bd}K_i$. Lemma 1 in [11] now yields the assertion, as $\mathrm{bd}K_i$ is H^1–rectifiable and H^1–measurable.

4. First contact distribution with finite structuring element

In this section, we first transfer Theorem 2 to random sets. Afterwards, a stronger result for Boolean models will be presented, which is based on Theorem 1.

Theorem 3 Let Z be an (almost surely topologically regular) stationary random set in $\mathbb{R}^2$ satisfying (2.4). Then, its first contact distribution function with respect to the finite set $M \subset \mathbb{R}^2$ obeys

$$F'_M(0+) = \int_{S^1} h(\mathrm{conv}M, -v)\,\overline{S}(Z, dv).$$

Proof. We may assume that $0 \in M$. We will show for a convex body K with interior points that

$$\mathbb{E}\int_{S^1} h(\mathrm{conv}M, -v) S(Z \cap K, dv) = \lambda_2(K)F'_M(0+) + f(K), \tag{4.1}$$

where $f(rK)/r^2$ tends to 0 as $r \to \infty$. The definition of $\overline{S}(Z,\cdot)$ then yields the assertion.

(2.4) and the inclusion– exclusion formula allow us to interchange limit and expectation in what follows. Using (2.2), Theorem 2 and Fubini's Theorem, we get

$$\begin{aligned}
&\mathbb{E}\int_{S^1} h(\mathrm{conv}M,-v)\,S(Z\cap K,dv) = \\
&= \; 2\mathbb{E}A(Z\cap K,\mathrm{conv}\check{M}) \\
&= \; \mathbb{E}\psi'_{Z\cap K}(\mathrm{conv}\check{M},0+) \\
&= \; \mathbb{E}\lim_{t\to 0+} 1/t\left[\lambda_2((Z\cap K)+t\check{M})-\lambda_2(Z\cap K)\right] \\
&= \; \lim_{t\to 0+} 1/t\int_{\mathbb{R}^2}\left[\mathbb{P}((Z\cap K\cap(tM+x)\neq\emptyset)-\mathbb{P}(x\in Z\cap K)\right]\lambda_2(dx).
\end{aligned}$$

We can assume that the Minkowski-difference

$$L_t(K)=L_t:=\{x\in\mathbb{R}^2 \mid x+t\mathrm{conv}M\subset K\}\subset K$$

of K and $t\mathrm{conv}M$ is nonempty. Note that L_t increases (in the sense of set inclusion) to K, as t decreases to 0. The definition of L_t and the stationarity of Z imply $\mathbb{P}(Z\cap K\cap(tM+x)\neq\emptyset)=F_M(t)$ for $x\in L_t$. Splitting the integral over $\mathbb{R}^2=L_t\cup L_t^C$ gives (4.1) with

$$f(K)=\lim_{t\to 0+} 1/t\int_{L_t^C}\left[\mathbb{P}(Z\cap K\cap(tM+x)\neq\emptyset)-\mathbb{P}(x\in Z\cap K)\right]\lambda_2(dx).$$

The stationarity of Z implies

$$\begin{aligned}
\mathbb{P}(Z\cap K\cap(tM+x)\neq\emptyset) &= \mathbb{P}(Z\cap tM\cap(K-x)\neq\emptyset) \\
&= \mathbb{P}(-x\in(Z\cap tM)+\check{K}) \\
&\leq \mathbb{P}(Z\cap tM\neq\emptyset)1_{-x\in tM+\check{K}} \\
&= F_M(t)1_{x\in K+t\check{M}}.
\end{aligned}$$

Hence,

$$\begin{aligned}
f(K) &\leq \lim_{t\to 0+} 1/t\left[F_M(t)\lambda_2(K+t\check{M}\setminus L_t)-F_M(0)\lambda_2(K\setminus L_t)\right] \\
&= 2A(K,\mathrm{conv}\check{M})F_M(0)+\lim_{t\to 0+} 1/t\lambda_2(K\setminus L_t)\left[F_M(t)-F_M(0)\right].
\end{aligned}$$

Here, again, we used Theorem 2. As $1/t\lambda_2(K\setminus L_t)$ remains bounded (see the 'convexity lemma' in [6]), we find $0\leq f(rK)\leq 2rA(K,\mathrm{conv}\check{M})F_M(0)$, $r\geq 0$, and the assertion is shown.

Now consider a stationary Boolean model Z with intensity γ and convex typical grain K_0. Clearly, Theorem 3 holds for Z. But due to the independence properties of a Boolean model, Theorem 1 yields a stronger result, namely estimates on the error. To compare the first contact distributions of a Boolean model with varying structuring elements $M \subset \mathbb{R}^2$, it is convenient to consider the function

$$I_M(t) := -\frac{1}{\gamma}\log(1-F_M(t)).$$

Note that

$$I_M(t) = \mathbb{E}\lambda_2(K_0 + t\check{M}) = \mathbb{E}\psi_{K_0}(\check{M}, t) \tag{4.2}$$

due to the definition of $F_M(t)$ and (2.7). If M is a convex body, this and (2.1) implies that $I_M(t)$ is a polynomial in $t \geq 0$ of degree at most 2:

$$I_M(t) = \mathbb{E}\lambda_2(K_0) + 2t\mathbb{E}A(K_0, \check{M}) + t^2\lambda_2(M). \tag{4.3}$$

This formula is the basis of the classical *minimum contrast method* (see e.g. MOLCHANOV [7]): From estimates of $F_M(t)$ for several different $t \geq 0$ the coefficients of the quadratic polynomial $\gamma I_M(t)$ can be estimated. This yields estimators for γ, the mean area of the typical particle and $\mathbb{E}A(K_0, \check{M})$ (in the isotropic case, the latter is proportional to the mean boundary length of K_0). If M is replaced by a finite set, (4.3) is no longer true. But the following proposition shows that the deviation of $I_{\mathrm{conv}M}(t)$ from $I_M(t)$ is of quadratic order in t. Thus, in the cases where the constant of the quadratic error term is small, an application of the minimum contrast method with finite structuring element M allows to estimate $\gamma\mathbb{E}\lambda_2(K_0)$ and $\gamma\mathbb{E}A(K_0, \mathrm{conv}\check{M})$.

Proposition 1 Let Z be a stationary Boolean model and $M \subset \mathbb{R}^2$ a finite set of at least two points. Then

$$0 \leq I_{\mathrm{conv}M}(t) - I_M(t) \leq t^2 C_{K_0, M}, \qquad t \geq 0,$$

with

$$C_{K_0, M} := \mathbb{E}q(K_0)\left(\frac{c_M}{4} + \frac{6 + L(\mathrm{conv}M)c_0}{c_0^2}\right) + \lambda_2(\mathrm{conv}M),$$

where c_M and c_0 are defined before Theorem 1 and $L(\mathrm{conv}M)$ is the boundary length of $\mathrm{conv}M$.

Proof. Due to (4.2), it is enough to show the right inequality. Only for this proof, we write $\rho(K) = r(K - z)$, where $z \in \mathrm{int}K$ is the minimizer in the definition (3.6) of $q(K)$ (and, again, $\rho(K) = \infty$, if $\mathrm{int}K = \emptyset$). According to (4.2) and Theorem 1, we have

$$\begin{aligned}
I_M(t) &\geq \mathbb{E}\psi_{K_0}(\check{M}, t)1_{\rho(K_0)\geq t/c_0} \\
&\geq \mathbb{E}\left(\psi_{K_0}(\mathrm{conv}\check{M}, t) - \frac{c_M}{4}q(K_0)t^2\right)1_{\rho(K_0)\geq t/c_0} \\
&\geq I_{\mathrm{conv}M}(t) - g(t) - \frac{c_M}{4}\mathbb{E}q(K_0)t^2,
\end{aligned}$$

where, according to (2.1),

$$\begin{aligned}
g(t) &= \mathbb{E}\left(\psi_{K_0}(\mathrm{conv}\check{M}, t)1_{\rho(K_0)<t/c_0}\right) \\
&\leq \mathbb{E}\left(\lambda_2(K_0)1_{\rho(K_0)<t/c_0}\right) + t\mathbb{E}\left(A(K_0, \mathrm{conv}\check{M})1_{\rho(K_0)<t/c_0}\right) + t^2\lambda_2(\mathrm{conv}M).
\end{aligned}$$

But, due to an elementary geometric argument (which does not yield an optimal bound), $\lambda_2(K_0) \leq 6\rho^2(K_0)q(K_0)$. Evidently, a suitable ball of radius $q(K_0)\rho(K_0)$ contains K_0, and so,

$$g(t) \leq t^2 \left(\frac{6\mathbb{E}q(K_0)}{c_0^2} + \frac{\mathbb{E}q(K_0)L(\mathrm{conv} M)}{c_0} + \lambda_2(\mathrm{conv} M) \right),$$

which gives the assertion.

Note that the estimate of the error in the last Proposition is global in $t \geq 0$, and hence is useful in the cases where $\mathbb{E}q(K_0)$ or an upper bound for this mean is known. In the case where M consists of two points, say $M = \{0, u\}$, $u \in S^1$, Proposition 1 and (2.3) imply

$$0 \leq \mathbb{E}\lambda_2(K_0) + t\mathbb{E}\lambda_1(K_0|u^\perp) - I_{\{0,u\}}(t) \leq 3.1 \cdot \mathbb{E}q(K_0)t^2, \qquad 0 \leq t.$$

This formula can be used to estimate the expected width of K_0 in direction $u^\perp$.

5. General point configurations

Let $t > 0$, the compact set $W \subset \mathbb{R}^2$ and the finite set $B \subset \mathbb{R}^2$ be given. (B stands for 'black' points belonging to the random set Z, W for 'white' points.) By induction on the number of elements of B we obtain

$$\mathbb{P}(tB \subset Z, tW \subset Z^C) = \sum_{M \subset B} (-1)^{|M|}(1 - F_{W \cup M}(t)).$$

For nonempty B this implies

$$\mathbb{P}(tB \subset Z, tW \subset Z^C) = \sum_{M \subset B} (-1)^{|M|+1} F_{W \cup M}(t). \tag{5.1}$$

We therefore obtain the following Theorem as a consequence of Theorem 3.

Theorem 4 Let Z be an (almost surely topologically regular) stationary random set fulfiling (2.4). If B and W are two nonempty finite subsets of $\mathbb{R}^2$, then

$$\lim_{t \to 0+} \frac{1}{t}\mathbb{P}(tB \subset Z, tW \subset Z^C) = \int_{S^1} h(-v)\, \overline{S}(Z, dv), \tag{5.2}$$

where

$$\begin{aligned} h &= [\min_{x \in B}\langle x, \cdot\rangle - \max_{x \in W}\langle x, \cdot\rangle]^+ \\ &= h\left(\mathrm{conv}\left((W + \check{B}) \cup \{o\}\right), \cdot\right) - h\left(\mathrm{conv}(W + \check{B}), \cdot\right) \end{aligned} \tag{5.3}$$

(Here $f^+ := \max\{f, 0\}$ denotes the positive part of a function f). Equivalently

$$\lim_{t \to 0+} \frac{1}{t}\mathbb{P}(tB \subset Z, tW \subset Z^C) = \int_{S^1} h(B(Z), -v)\, \mathbf{S}(dv),$$

where the (signed and discrete) measure $\mathbf{S}$ is given by

$$\mathbf{S} := S(\mathrm{conv}\left((W + \check{B}) \cup \{o\}\right), \cdot) - S(\mathrm{conv}(W + \check{B}), \cdot).$$

Proof. Theorem 3 and (5.1) imply (5.2) with

$$h = \sum_{M \subset B} (-1)^{|M|+1} h(\mathrm{conv}(W \cup M), \cdot).$$

In the following, for fixed $u \in S^1$, an induction argument in the number of points in B will lead to the alternative representations of $h(u)$ according to (5.3): As B is nonempty, there is a support element $x \in B$ in direction u, i.e. $\langle x, u \rangle = h(\mathrm{conv}B, u)$. If $|B| > 1$, then

$$\sum_{M \subset B, x \in M} (-1)^{|M|+1} h(\mathrm{conv}(W \cup M), u) = 0,$$

as all the summands are up to sign equal to $h(\mathrm{conv}(W \cup \{x\}), u)$. If $y \in B$ is a support element of conv$B$ in direction $-u$, $\langle y, -u \rangle = h(\mathrm{conv}B, -u)$, we get by induction

$$h(u) = \sum_{M \subset \{y\}} (-1)^{|M|+1} h(\mathrm{conv}(W \cup M), u) = [-h(\mathrm{conv}B, -u) - h(\mathrm{conv}W, u)]^+.$$

This directly yields the first equality in (5.3) and, as

$$-h(\mathrm{conv}B, -u) - h(\mathrm{conv}W, u) = -h(\mathrm{conv}(W + \check{B}), u),$$

the second equality in (5.3) follows. The reformulation in terms of the mean Blaschke body now is a consequence of (2.6) and (2.2).

The first representation of $h(u)$, $u \in S^1$, in (5.3) has the following geometric interpretation: Let $S(u)$ be the (possibly empty) union of all lines orthogonal to u, separating the sets W and B, in such a way that W lies in the negative half plane with respect to u. Then $h(u)$ is equal to the width of the strip $S(u)$.

6. 2×2 point configurations

We will now apply (5.2) in the case, where B and W are subsets of the unit cell in the standard grid $t\mathbb{Z}^2$. Let e_1, e_2 be the standard basis of $\mathbb{R}^2$ and $V = \{0, e_1, e_2, e_1 + e_2\}$ the set of vertices of the unit cell $[0,1]^2$. For $t > 0$, a 2×2*-point configuration (with scaling-factor t)* is a subset of tV. As an example, consider the configuration

$$\mathcal{C}_t = \begin{bmatrix} \circ & \circ \\ \bullet & \bullet \end{bmatrix}_t := \{0, te_1\}.$$

Note that $Z \cap tV = \mathcal{C}_t$ is the event that $\{0, te_1\} \subset Z$ and $\{te_2, te_1 + te_2\} \subset Z^C$. Similar notations in what follows will be self explaining. (In subsequent sections the notion of a configuration will be extended to all translations of a subset of tV, thus allowing to count the number of configurations of a given type in $Z \cap \mathbb{Z}^2$ in a sampling window.) Configurations different from

$$\begin{bmatrix} \circ & \circ \\ \circ & \circ \end{bmatrix}_t \quad \text{and} \quad \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix}_t$$

will be called *boundary configurations*, as they contain information about the boundary of Z. This will be made precise in the following Corollary of Theorem 4.

Corollary 1 Let Z be an (almost surely topologically regular) stationary random set fulfiling (2.4). Then for every boundary configuration $\mathcal{C}_t$ there are $a, b \in \mathbb{R}^2$ such that

$$\lim_{t\to 0+} \frac{1}{t}\mathbb{P}(Z \cap tV = \mathcal{C}_t) = \int_{S^1} h_{a,b}(v)\, \overline{S}(Z, dv), \tag{6.1}$$

where

$$h_{a,b}(v) := \min\{\langle a, v\rangle^+, \langle b, v\rangle^+\}.$$

Equivalently we have

$$\lim_{t\to 0+} \frac{1}{t}\mathbb{P}(Z \cap tV = \mathcal{C}_t) = h(B(Z), a) + h(B(Z), b) - h(B(Z), a+b). \tag{6.2}$$

The values of a and b are listed in Table 1.

To derive Table 1 from Theorem 4, it is convenient to write $h_{B,W}$ for the function h in (5.3) to indicate its dependence on B and W. As $h_{W,B}(v) = h_{B,W}(-v)$ and $h_{\vartheta B,\vartheta W}(\vartheta v) = h_{B,W}(v)$ for all $v \in S^1$ and all rotations ϑ of $\mathbb{R}^2$ (fixing the origin), the calculations can be reduced to three boundary configurations to be considered, e.g. $\mathcal{C}_t^{(i)}$, $i = 1, 3, 6$, in Table 1 below.

i	$\mathcal{C}_t^{(i)}$	a	b
0	$\begin{bmatrix} \circ & \circ \\ \circ & \circ \end{bmatrix}_t$	–	–
1	$\begin{bmatrix} \circ & \circ \\ \bullet & \circ \end{bmatrix}_t$	e_1	e_2
2	$\begin{bmatrix} \circ & \circ \\ \circ & \bullet \end{bmatrix}_t$	$-e_1$	e_2
3	$\begin{bmatrix} \circ & \circ \\ \bullet & \bullet \end{bmatrix}_t$	$e_1 + e_2$	$-e_1 + e_2$
4	$\begin{bmatrix} \bullet & \circ \\ \circ & \circ \end{bmatrix}_t$	e_1	$-e_2$
5	$\begin{bmatrix} \bullet & \circ \\ \bullet & \circ \end{bmatrix}_t$	$e_1 + e_2$	$e_1 - e_2$
6	$\begin{bmatrix} \bullet & \circ \\ \circ & \bullet \end{bmatrix}_t$	0	0
7	$\begin{bmatrix} \bullet & \circ \\ \bullet & \bullet \end{bmatrix}_t$	e_1	e_2
8	$\begin{bmatrix} \circ & \bullet \\ \circ & \circ \end{bmatrix}_t$	$-e_1$	$-e_2$
9	$\begin{bmatrix} \circ & \bullet \\ \bullet & \circ \end{bmatrix}_t$	0	0
10	$\begin{bmatrix} \circ & \bullet \\ \circ & \bullet \end{bmatrix}_t$	$-e_1 + e_2$	$-e_1 - e_2$
11	$\begin{bmatrix} \circ & \bullet \\ \bullet & \bullet \end{bmatrix}_t$	$-e_1$	e_2
12	$\begin{bmatrix} \bullet & \bullet \\ \circ & \circ \end{bmatrix}_t$	$e_1 - e_2$	$-e_1 - e_2$
13	$\begin{bmatrix} \bullet & \bullet \\ \bullet & \circ \end{bmatrix}_t$	e_1	$-e_2$
14	$\begin{bmatrix} \bullet & \bullet \\ \circ & \bullet \end{bmatrix}_t$	$-e_1$	$-e_2$
15	$\begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix}_t$	–	–

Table 1: The values of a and b for the configuration $\mathcal{C}_t$.

Remark 1 The integrals corresponding to the configurations $\mathcal{C}_t^{(6)}$ and $\mathcal{C}_t^{(9)}$ are 0. The number of these configurations in a digitized image can therefore be used as *an indicator for non-sufficient resolution t. Note that four of the twelve remaining integrals occur twice, so, in fact, the 2×2 point configurations yield at most eight nontrivial different integrals of $\overline{S}(Z,\cdot)$.*

6.1. *Inference on the mean normal measure.* We will now discuss different possibilities to obtain estimators for $\overline{S}(Z,\cdot)$ using Corollary 1. Set $v_i := (\cos(\frac{i}{8}\cdot 2\pi), \sin(\frac{i}{8}\cdot 2\pi))^\top, i \in \mathbb{N}_0$, and let

$$T := \{v_0, v_1, \dots, v_7\}$$

be the set of all directions defined by sides and diagonals of the unit square $[0,e_1] \times [0,e_2]$. For convenience we put

$$I_v := \int_{S^1} h_{a(v),b(v)}(w)\,\overline{S}(Z,dw),$$

where $(a(v),b(v))$ is the pair of neighbours of $v \in T$ in $S^1 \cap T$ (multiplied by $\sqrt{2}$ if $v \in \{\pm e_1, \pm e_2\}$). So, up to normalization, we have $(a(v_i),b(v_i)) = (v_{i-1}, v_{i+1})$, where cyclic indexing is used. Note that the numbers I_{v_i} can be estimated by counting the corresponding configuration in a discretized image, if the resolution t is small enough. See also section 6.2 below.

It is clear that the eight different integrals $(I_{v_0},\dots,I_{v_7})$ do not determine $\overline{S}(Z,\cdot)$ uniquely. We will therefore impose on $\overline{S}(Z,\cdot)$ the additional assumption to belong to some subclass of the space of measures. The choice of the appropriate class in applications must then be adapted to the situation. Let us suppose that the mean normal measure belongs to a parametrized class of measures, parametrized by $\theta \in \Theta$, where Θ is a subset of $\mathbb{R}^p$, say. Let $\overline{S}_\theta(Z,\cdot)$ be the notation for the mean normal measure with parameter θ. The parameter θ is said to be *identifiable in* Θ if the mapping $\theta \mapsto (I_{v_i}(\theta))_{i=0}^7$ is injective on Θ, where

$$I_v(\theta) := \int_{S^1} h_{a(v),b(v)}(w)\,\overline{S}_\theta(Z,dw) \tag{6.3}$$

for $v \in T$. The following examples will illustrate this approach:

Example 1 The discrete case.
Assume that the mean normal measure belongs to the class

$$\mathcal{M}(T) := \{\sum_{i=0}^{7} \theta_i \delta_{v_i} \,|\, \theta_i \in \mathbb{R}_+\}$$

of all measures supported by T, where δ_v is the probability measure supported by $\{v\} \subset S^1$. We have $\Theta = \mathbb{R}_+^8$. As $\overline{S}_\theta(Z,\cdot) = \sum_{i=0}^7 \theta_i \delta_{v_i}$, we obtain

$$I_{v_i}(\theta) = \theta_i \cdot h_{a(v_i),b(v_i)}(v_i),$$

for $i = 0,\dots,7$. Therefore

$$S_\theta(Z,\cdot) = \sum_{i=0}^{7} \frac{I_{v_i}(\theta)}{h_{a(v_i),b(v_i)}(v_i)} \delta_{v_i},$$

and θ is identifiable in Θ.

Example 2 The case of a piecewise constant density.
Assume that the mean normal measure has a piecewise constant density with respect to the ordinary length measure ω_1 *on* S^1 *(spherical Lebesgue measure). More precisely, assume that this density is constant on* D_i *where* $D_i \subset S^1$ *is the arc of length* $\pi/4$ *centered at* v_i *for* $i = 0, \ldots, 7$*. This means*

$$\overline{S}_\theta(Z, \cdot) = \sum_{i=0}^{7} \theta_i \cdot \omega_1(\,\cdot\, \cap D_i)$$

with $\theta = (\theta_0, \theta_1, \ldots, \theta_7)^\top \in \Theta := \mathbb{R}^8_+$*. We have*

$$A \cdot \theta = \begin{pmatrix} I_{v_0}(\theta)/\sqrt{2} \\ I_{v_1}(\theta) \\ I_{v_2}(\theta)/\sqrt{2} \\ I_{v_3}(\theta) \\ \vdots \\ I_{v_6}(\theta)/\sqrt{2} \\ I_{v_7}(\theta) \end{pmatrix}, \tag{6.4}$$

with the matrix

$$A := \begin{pmatrix} 2\gamma & \tau & 0 & 0 & 0 & 0 & 0 & \tau \\ \tau & 2\gamma & \tau & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau & 2\gamma & \tau & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau & 2\gamma & \tau & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau & 2\gamma & \tau & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau & 2\gamma & \tau & 0 \\ 0 & 0 & 0 & 0 & 0 & \tau & 2\gamma & \tau \\ \tau & 0 & 0 & 0 & 0 & 0 & \tau & 2\gamma \end{pmatrix},$$

where $\gamma = \cos(\pi/8) - \cos(\pi/4)$ *and* $\tau = 1 - \cos(\pi/8)$*. As* A *is invertible, the parameter* $\theta \in \mathbb{R}^8_+$ *is identifiable and can be obtained solving the linear system (6.4). The advantage of this model over the discrete model in Example 1 is that it includes the case where* Z *is isotropic (then, the mean normal measure is a multiple of* ω_1*).*

Example 3 Other parametric families.
It is also possible to use standard parametric families in the analysis. The mean normal measure is then parametrized by $\theta = (\theta_t, \theta_d)$ *where* θ_t *is the total mass and* θ_d *is the parameter of a* ω_1-*probability density* $p(\,\cdot\,; \theta_d)$ *on* S^1*. Then,*

$$\overline{S}_\theta(Z, A) = \theta_t \int_A p(v; \theta_d) d\omega_1(v), \quad A \in \mathcal{B}(S^1).$$

For each specific choice of a density family it must be checked whether θ *is identifiable in* Θ*.*

We have already seen that the mean Blaschke body and the mean normal measure of Z are in one-to-one correspondence (see 2.6). For visualization it might therefore be useful to 'estimate' the mean Blaschke body instead of $\overline{S}(Z,\cdot)$. This procedure is well-known in the context of stationary fibre processes, where the associated zonoid (the Steiner compact) is estimated from the rose of intersection. The key observation is that

$$Q := \bigcap_{i=0}^{7} \{\langle\,\cdot\,, v_i\rangle \leq h(B(Z), v_i)\}$$

is a polygonal approximation of $B(Z)$ from outside, only based on the support values of $B(Z)$ in the directions $v_0, \ldots, v_7$. To construct this polygon, we need to determine $(h(B(Z), v_i))_{i=0}^{7}$ from I_{v_i}, $i = 0, \ldots, 7$. In view of (6.2), we have

$$I_{v_i} = h(B(Z), a(v_i)) + h(B(Z), b(v_i)) - h(B(Z), a(v_i) + b(v_i)), \quad i = 0, \ldots, 7.$$

More explicitly,

$$D \cdot \begin{pmatrix} h(B(Z), v_0) \\ \vdots \\ h(B(Z), v_7) \end{pmatrix} = \begin{pmatrix} I_{v_0}/\sqrt{2} \\ I_{v_1} \\ I_{v_2}/\sqrt{2} \\ I_{v_3} \\ \vdots \\ I_{v_6}/\sqrt{2} \\ I_{v_7} \end{pmatrix} \tag{6.5}$$

with

$$D = \begin{pmatrix} -\sqrt{2} & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -\sqrt{2} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\sqrt{2} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\sqrt{2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\sqrt{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\sqrt{2} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -\sqrt{2} & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & -\sqrt{2} \end{pmatrix}.$$

Note that $\operatorname{rank}(D) = 6$. Thus, the vector $\theta = (h(B(Z), v_i))_{i=0}^{7}$ is determined by the integrals I_v, $v \in T$, only up to addition of a vector $(\langle x, v_i\rangle)_{i=0}^{7}$ (which corresponds to a translation of $B(Z)$ by $x \in \mathbb{R}^2$). So, any solution of (6.5) will yield a polygon Q which is an approximation of a translate of $B(Z)$. Note, however that the described procedure is sufficient to get information about the *shape* of $B(Z)$, which determines $\overline{S}(Z,\cdot)$. If the underlying set Z is a Boolean model, then the obtained polygon Q is an approximation of the mean typical grain (up to homothety), due to (2.8).

6.2. *Statistical considerations.* In this section, we will discuss how to estimate the mean normal measure from observations of the different types of 2×2 configurations in a sampling window. Among 2×2 configurations, there are 12 informative configurations (those different from No.s $0, 6, 9$ and 15, see Table 1). We will not distinguish between informative configurations having the same probability of being observed. This leaves us with 8 different types of informative configurations which can be indexed by T. Other configurations than the informative ones will be called type 0 configurations. As in the previous section we assume that the mean normal measure belongs to a parametrized class of measures, parametrized by $\theta \in \Theta$. According to Corollary 1, the probability of observing a configuration of type $v \in T$ is, approximately,

$$p_v(\theta) = tm_v I_v(\theta).$$

Here, m_v is the number of configurations that are combined in type v ($m_v = 1$ for $v \in \{e_1, e_2, -e_1, -e_2\}$ and $m_v = 2$ otherwise) and $I_v(\theta)$ is defined in (6.3). Furthermore, we let

$$p_0(\theta) = 1 - \sum_{v \in T} p_v(\theta)$$

and $\overline{T} = T \cup \{0\}$.

Let us suppose that we have observed n_v configurations of type $v \in \overline{T}$ in a discretization of Z in a sampling window. Let $n = \sum_{v \in \overline{T}} n_v$ be the total number of observed configurations. We suggest to base the estimation of θ on the function ℓ, defined by

$$\ell(\theta) := \sum_{v \in \overline{T}} n_v \ln p_v(\theta), \quad \theta \in \Theta. \tag{6.6}$$

The value $\ell(\theta)$ is closely related to the Kullback-Leibler divergence of the observed frequencies $f = (n_v/n)_{v \in \overline{T}}$ from the probability function $p(\theta) = (p_v(\theta))_{v \in \overline{T}}$. Recall that for general probability functions $p = (p_v)_{v \in \overline{T}}$ and $q = (q_v)_{v \in \overline{T}}$, the divergence of p from q is defined as

$$D(p, q) = \sum_{v \in \overline{T}} p_v \ln \frac{p_v}{q_v},$$

cf. e.g. Lauritzen [4], p. 238-239, and references therein. The divergence fulfils $D(p,q) \geq 0$ and $D(p,q) = 0$ if and only if $p = q$. The mentioned relation between ℓ and D is

$$\ell(\theta) = \sum_{v \in \overline{T}} n_v \ln(\frac{n_v}{n}) - nD(f, p(\theta)). \tag{6.7}$$

Note also that in the case of independent observed configurations $\ell(\theta)$ is the log–likelihood function (up to an additive term which is independent of θ).

As an estimate of θ we will use a value $\hat{\theta} \in \Theta$ at which (6.6) is maximal, if such a value exists. Equivalently, we use a value of θ at which the divergence $D(f, p(\theta))$ is minimal. Usually, $\hat{\theta}$ has to be found by an iterative procedure. However, if there exists $\tilde{\theta} \in \Theta$ such that

$$p_v(\tilde{\theta}) = \frac{n_v}{n}, \quad v \in \overline{T}, \tag{6.8}$$

then, due to (6.7) and the properties of the divergence, $\tilde{\theta}$ is a solution of the maximization problem. Thus, a parameter value for which the theoretical probabilities coincide with observed frequencies is always a solution to the estimation problem. If $\tilde{\theta}$ in (6.8) is unique, $\hat{\theta} = \tilde{\theta}$ is the only solution of (6.6).

This remark is relevant for the discrete case and the case of a piecewise constant density mentioned in section 6.1. In the discrete case, (6.8) is always fulfilled. Here, $\Theta = \mathbb{R}_+^8$ and

$$p_{v_i}(\theta) = tm_{v_i}\theta_i h_{a(v_i),b(v_i))}(v_i), \quad i = 0,\dots,7.$$

Accordingly,

$$\hat\theta_i = \frac{n_{v_i}}{ntm_{v_i}h_{a(v_i),b(v_i)}(v_i)}, \quad i = 0,\dots,7.$$

In the case of a piecewise constant density,

$$p_{v_i}(\theta) = t\tilde m_{v_i}(A\theta)_{v_i}, \quad i = 0,\dots,7,$$

where $\tilde m_{v_i} := 2$ for even i and $\tilde m_{v_i} := \sqrt{2}$ for odd i. Let $\rho = (n_{v_i}/nt\tilde m_{v_i})_{i=0}^7$. If

$$A^{-1}\rho \in \mathbb{R}_+^8,$$

then (6.8) has the unique solution $\hat\theta = \tilde\theta = A^{-1}\rho$. Otherwise, $\hat\theta$ has to be found using an iterative procedure.

Note that if the main interest is in the directional *distribution*

$$\mathcal{R} = \overline{S}(Z,\cdot)/\overline{L}(Z),$$

we may instead use the conditional probability of observing a type v configuration, $v \in T$, given the configuration is informative. If $\mathcal{R}$ is parametrized by θ_d then the conditional probabilities take the form

$$p_{v|T}(\theta_d) = \frac{m_v \int_{S^1} h_{a(v),b(v)}(w)\mathcal{R}_{\theta_d}(dw)}{\sum_{v\in T} m_v \int_{S^1} h_{a(v),b(v)}(w)\mathcal{R}_{\theta_d}(dw)}.$$

In analogy with (6.6) the parameter θ_d may be estimated by maximizing

$$\ell(\theta_d) := \sum_{v\in T} n_v \ln p_{v|T}(\theta_d).$$

6.3. *Related approaches.* Our approach coincides essentially with a known method if $\overline{S}(Z,\cdot)$ is even. In this case we obtain at most four nontrivial different integrals of $\overline{S}(Z,\cdot)$ due to symmetry.

According to Theorem 4, we have for $x \in \mathbb{R}^2$

$$\lim_{t\to 0+} \frac{1}{t}\mathbb{P}(0 \in Z, tx \notin Z) = \frac{1}{2}\int_{S^1} |\langle v,x\rangle|\, \overline{S}(Z,dv), \tag{6.9}$$

where we used (2.5). Formula (6.9) is well known, its right hand side can be interpreted as specific length of total projection of Z on the line $x^\perp$. This result has been used for the analysis of planar digitized images by assuming that x is a point of the grid, see e.g. Ohser et al. [8]. If we restrict attention to a point in the 8-neighbourhood of 0, then $x \in \{\pm e_1, \pm e_2, \pm(e_1+e_2), \pm(e_1-e_2)\}$. As (6.9) is invariant under reflection of x, only four directions have to be considered. For example, if

$x = e_1$, we have to consider the event that $0 \in Z$ but $te_1 \in Z^C$, which can be written more descriptive as

$$Z \cap tV \in \begin{bmatrix} \cdot & \cdot \\ \bullet & \circ \end{bmatrix}_t,$$

if we set

$$\begin{bmatrix} \cdot & \cdot \\ \bullet & \circ \end{bmatrix}_t := \left\{ \begin{bmatrix} \circ & \circ \\ \bullet & \circ \end{bmatrix}_t, \begin{bmatrix} \circ & \bullet \\ \bullet & \circ \end{bmatrix}_t, \begin{bmatrix} \bullet & \circ \\ \bullet & \circ \end{bmatrix}_t, \begin{bmatrix} \bullet & \bullet \\ \bullet & \circ \end{bmatrix}_t \right\}.$$

With a self explaining use of notation, the four directions mentioned above correspond to the four events

$$(6.10) \quad Z \cap tV \in \begin{bmatrix} \cdot & \cdot \\ \bullet & \circ \end{bmatrix}_t, \; Z \cap tV \in \begin{bmatrix} \cdot & \circ \\ \cdot & \bullet \end{bmatrix}_t, \; Z \cap tV \in \begin{bmatrix} \circ & \bullet \\ \cdot & \cdot \end{bmatrix}_t, \; Z \cap tV \in \begin{bmatrix} \bullet & \cdot \\ \circ & \cdot \end{bmatrix}_t.$$

Clearly, the four events in (6.10) can be expressed as (disjoint) unions of events in the family $\{Z \cap tV = \mathcal{C}_t^{(i)} \mid i = 1, \ldots, 14\}$. On the other hand, the probability of any of the boundary configurations can, in the case of a symmetric mean normal measure, be expressed in terms of probabilities of events in (6.10). This shows that the events in (6.10) contain the same information as all boundary configurations. Therefore, in the case of a symmetric mean normal measure, there is no gain of information using our approach. Yet, in the general case, our procedure yields eight different integrals and hence more information on $\overline{S}(Z, \cdot)$.

We also note the following consequence for the determination of the length density $\overline{L}(Z) := \overline{S}(Z, S^1)$ from the integrals in (6.1). Only in this paragraph, we will use the abbreviation

$$\rho_i := \lim_{t \to 0+} \frac{1}{t} \mathbb{P}(Z \cap tV = \mathcal{C}_t^{(i)}), \qquad i = 1, \ldots, 14.$$

If Cauchy's formula for the length density is discretized with the rectangular rule (in the present setting one could equivalently work with the trapezoidal rule), then

$$\tilde{L} := \frac{\pi}{8\sqrt{2}} \left[(1 + \sqrt{2}) \sum_{i=1}^{14} \rho_i + \sum_{i \in \{3,5,10,12\}} \rho_i \right]$$

is obtained as an approximation of $\overline{L}(Z)$ (see Ohser & Mücklich [9], formula (4.9)). If the values ρ_i are replaced by estimators, $\tilde{L}$ becomes a (biased) estimator of $\overline{L}(Z)$. Corollary 1 yields

$$\tilde{L} = \int_{S^1} g(v) \overline{S}(Z, dv),$$

where

$$g = \frac{(1 + \sqrt{2})\pi}{4\sqrt{2}} \left[\min\{|\langle e_1, \cdot \rangle|, |\langle e_2, \cdot \rangle|\} + \frac{1}{\sqrt{2}} \min\{|\langle e_1 + e_2, \cdot \rangle|, |\langle e_1 - e_2, \cdot \rangle|\} \right].$$

Put $u_{\pi/8} := (\cos(\pi/8), \sin(\pi/8))^\top$. We have $g(e_1) \le g \le g(u_{\pi/8})$ on S^1, with

$$g(e_1) = (1 + \sqrt{2}) \frac{\pi}{8} \approx 0.95 \quad \text{and} \quad g(u_{\pi/8}) = g(e_1) \cdot 2\sqrt{2} \sin(\frac{\pi}{8}) \approx 1.03.$$

Hence

$$g(e_1)\overline{L}(Z) \le \tilde{L} \le g(u_{\pi/8})\overline{L}(Z) \tag{6.11}$$

and these estimates are sharp. In particular, this shows that $\tilde{L}$ can deviate maximally 5% from $\overline{L}(Z)$. This maximal deviation is attained if and only if Z is a polygonal random set for which the directions of all boundary segments are parallel to sides or diagonals of the unit square (almost surely). We note that these estimates could as well be derived without application of Corollary 1, using the well known fact that total projections of Z are cosine transforms of $\overline{S}(Z,\cdot)$. Therefore, (6.11) is also valid in the context of general stationary fibre processes, if it is interpreted according to this setting.

7. On $n \times n$ configurations

In section 6 , we have concentrated on 2×2 configurations and shown how they can be used for making inference about the mean normal measure. Clearly, Theorem 4 can be used for developing corresponding results for $n \times n$ configurations with $n > 2$.

The crucial step is to find explicit integral representations for

$$\lim_{t \to 0+} \frac{1}{t}\mathbb{P}(tB \subset Z, tW \subset Z^C) \tag{7.1}$$

for disjoint, non-empty sets B and W such that $tB \cup tW$ is a scaled grid square of size $n \times n$. For positive t small enough, $t \le t_0$, say, the probability $\mathbb{P}(tB \subset Z, tW \subset Z^C)$ can be used as an approximation of t times the limit in (7.1), as previously used for 2×2 configurations. Note, however, that t_0 is expected to depend on n: the larger n, the smaller t_0.

Let $S(u)$ be the union of all lines orthogonal to $u \in S^1$, separating B and W as explained just after the proof of Theorem 4. Let

$$S = \bigcup_{u \in S^1} S(u)$$

be the set of all separating lines. If $S = \emptyset$, such that B and W cannot be separated by a line (in the above sense), then Theorem 4 implies that the limit (7.1) is zero and the corresponding $n \times n$ configuration is not informative. If $S \ne \emptyset$, then it is possible to find a separating line spanned by two points in the $n \times n$ grid square $tB \cup tW$. Since at most $\binom{n^2}{2} = n^2(n^2-1)/2$ such lines exist, there are less than n^4 informative configurations. Thus, although the number of $n \times n$ configurations increases exponentially in n, the estimation of the mean normal measure from $n \times n$ configurations is only a polynomial problem.

The thorough study of these aspects, including an extension to rectangular grids, is part of our future research plans.

Acknowledgement

This work has been supported in part by MaPhySto, funded by a grant from the Danish National Research Foundation.

References

[1] Hug, D., Last, G. (2000) On support measures in Minkowski spaces and contact distributions in stochastic geometry, *Ann. Probab.* **28**, 796-850.

[2] Hug, D., Last, G., Weil, W. (2002+) A survey on contact distribution functions, submitted.

[3] Kiderlen, M. (2002) Determination of the mean normal measure from isotropic means of flat sections, *Adv. Appl. Prob.*, to appear.

[4] Lauritzen, S.L. (1996) *Graphical Models*, Oxford University Press, Oxford.

[5] Matheron, G. (1967) *Eléments pour une théorie des milieux poreux*, Masson et Cie, Paris.

[6] Matheron, G. (1974) Hyperplans poissoniens et compacts de Steiner, *Adv. Appl. Prob.* **6**, 563-579.

[7] Molchanov, I. S. (1997) *Statistics of the Boolean Model for Practioners and Mathematicians*, Wiley, Chichester.

[8] Ohser, J., Steinbach, B., Lang, C. (1998) Efficient texture analysis of binary images, *J. Microsc.* **192**, 20–28.

[9] Ohser, J., Mücklich, F. (2000) *Statistical Analysis of Microstructures in Material Science*. Wiley, New York.

[10] Rataj, J. (1996) Estimation of oriented direction distribution of a planar body, *Adv. Appl. Prob.* **28**, 394-404.

[11] Rataj, J. (2002+) On the set covariance and three-point test sets, submitted.

[12] Schneider, R. (1993) *Convex Bodies: the Brunn-Minkowski Theory*, Cambridge Univ. Press, Cambridge.

[13] Schneider, R., Weil, W. (2000) *Stochastische Geometrie*, Teubner, Stuttgart 2000.

[14] Serra, J. (1982) *Image Analysis and Mathematical Morphology*, Vol. 1. Academic Press, London.

[15] Stoyan, D., Beneš, V. (1991) Anisotropy analysis for particle systems *J. Microsc.* **164**, 159–168.

[16] W. Weil (1997) Mean bodies associated with random closed sets *Suppl. Rend. Circ. Mat. Palermo* (2) **50**, 387–412.

[17] W. Weil (1997) The mean normal distribution of stationary random sets and particle processes. In: *Advances in Theory and Applications of Random Sets, Proc. International Symposium, Oct. 9–11, 1996, Fontainebleau, France,* (D. Jeulin, ed.) World Scientific Publ., Singapore, 21–33.