

# FIRST CONTACT DISTRIBUTIONS FOR SPATIAL PATTERNS: REGULARITY AND ESTIMATION

MARTIN B. HANSEN, <sup>1</sup>*Aalborg University*  
ADRIAN BADDELEY, <sup>2</sup>*University of Western Australia*  
RICHARD D. GILL, <sup>3</sup>*University of Utrecht*

## Abstract

For applications in spatial statistics, an important property of a random set  $X$  in  $\mathbb{R}^k$  is its first contact distribution. This is the distribution of the distance from a fixed point  $0$  to the nearest point of  $X$ , where distance is measured using scalar dilations of a fixed test set  $B$ . We show that, if  $B$  is convex and contains a neighbourhood of  $0$ , the first contact distribution function  $F_B$  is absolutely continuous. We give two explicit representations of  $F_B$ , and additional regularity conditions under which  $F_B$  is continuously differentiable. A Kaplan-Meier estimator of  $F_B$  is introduced and its basic properties examined.

EMPTY SPACE FUNCTION; CONVEX TEST SETS; RANDOM CLOSED SETS; STOCHASTIC GEOMETRY; SPATIAL STATISTICS; COAREA FORMULA; PRODUCT LIMIT ESTIMATES; HAZARD RATE

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60D05, 62H11  
SECONDARY 62G05

## 1 Introduction

The statistical analysis of a spatial pattern often involves treating the pattern as a realisation of a stationary random set  $X$  in  $\mathbb{R}^k$ , observed within a bounded window  $W$ , and estimating properties of  $X$ . A particularly useful property is the *first contact distribution function*  $F_B$  of  $X$  with respect to a test set  $B \subset \mathbb{R}^k$ , defined as

$$F_B(r) = \mathbb{P}\{X \cap rB \neq \emptyset\}, \quad r \geq 0,$$

where  $rB$  denotes the scalar dilation of  $B$  by the factor  $r$ . In other words this is the distribution function of the “distance” from  $0$  to  $X$ ,

$$\rho_B(0, X) = \inf\{r > 0 : X \cap rB \neq \emptyset\}.$$

---

<sup>1</sup>Postal address: Department of Mathematics, Aalborg University, Frederik Bajersvej 7 E, 9220 Aalborg Ø, Denmark.

<sup>2</sup>Postal address: Department of Mathematics, University of Western Australia, Nedlands WA 6907, Australia.

<sup>3</sup>Postal address: Mathematical Institute, University of Utrecht, Budapestlaan 6, 3584 CD Utrecht, The Netherlands.

If  $B$  is the unit ball in  $\mathbb{R}^k$ , then  $\rho_B(0, X)$  is the usual Euclidean distance from 0 to the nearest point of  $X$ , and  $F_B$  is called the spherical contact distribution function or *empty space function*. If  $B$  is a line segment of unit length with one endpoint at 0, then  $\rho_B(0, X)$  is the shortest distance from 0 to  $X$  along the line drawn through the segment  $B$ , and  $F_B$  is known as the *linear contact distribution*. These two special cases of  $B$  have been used extensively in applications. The exact form of  $F_B$  is known for certain stochastic models, enabling parameter estimation and (Monte Carlo) hypothesis testing. A qualitative summary of the pattern is available by plotting the estimated  $F_B$  against the theoretical  $F_B$  for an appropriate Poisson random set model. Isotropy (rotational invariance) of the pattern can be tested by comparing the linear contact distributions obtained from line segments at different orientations. See e.g. Cressie (1991), Diggle (1983), Matheron(1975), Ripley (1988), Serra (1982), Stoyan et al. (1995).

There are many reasons for studying  $F_B$  with more general test sets  $B$ . Firstly, in practice, digital image analysers do not estimate the spherical contact function but rather a discrete approximation to it, in which the disc  $B$  is replaced by a polygon. Secondly, the linear contact function is degenerate or undefined for point processes. To investigate isotropy of a point process one might estimate  $F_B$  for (say) ellipses  $B$  of different orientations and eccentricities; this can be computed by applying linear transformations to the pattern and estimating the spherical contact function. Other sets  $B$  are also of interest. Thirdly, the spherical and linear contact distributions do not completely characterise the random set  $X$ . The distribution of  $X$  is, however, determined by the values of

$$T(K) = \mathbb{P}\{X \cap K \neq \emptyset\}$$

for all compact subsets  $K$  of  $\mathbb{R}^k$  (Matheron, 1975). The contact distributions  $F_B(r) = T(rB)$  evaluate  $T$  on certain subclasses of compact sets. Enlarging the class of sets  $K$  for which  $T(K)$  is known should provide better information about  $X$ .

A form of censoring occurs when  $X$  is observed only inside a bounded window  $W$ . The distance  $T = \rho_B(0, X)$  is not observable, since the nearest point of  $X$  may be outside  $W$ . Effectively  $T$  is right-censored by  $C = \rho_B(0, \partial W)$ , the distance to the boundary  $\partial W$  of the window. This interpretation can be used to derive a Kaplan-Meier type estimator for  $F_B$ .

Baddeley and Gill (1993, 1997) studied the spherical contact distribution of a stationary *point process*  $X$ , developed the censoring analogy, showed that  $F_B$  is continuously differentiable, constructed a Kaplan-Meier estimator  $\hat{F}_B$  of  $F_B$ , and established some basic properties of  $\hat{F}_B$ . Hansen et al. (1996) treated the linear contact distribution of a stationary *random set*, showed that  $F_B$  is absolutely continuous on  $(0, \infty)$ , constructed the analogous Kaplan-Meier estimator and established the same basic properties.

In the present paper we treat the case of a general test set  $B$  which is convex and contains a neighbourhood of 0. This encompasses the spherical case but not the linear case treated in Hansen et al. (1996).

The plan of the paper is as follows. Section 2 includes some preliminary notation and results from spatial statistics. In Section 3, it is proved that the first contact distribution function is absolutely continuous for test sets in the class of convex bodies and under further regularity conditions on the

random set  $X$  it can be shown that  $F$  is continuously differentiable. Section 4 contains some remarks on the estimation of  $F_B$  by a Kaplan-Meier type estimator. Some comments are collected in Section 5.

## 2 Preliminaries

Henceforth let  $B$  be a fixed, compact set in  $\mathbb{R}^k$  called the *test set*. Write  $rB = \{rb : b \in B\}$  for the scalar dilation of  $B$  by  $r \in \mathbb{R}$ , and for  $x \in \mathbb{R}^k$  let

$$B_x = \{x + b : b \in B\}$$

be the translation of  $B$  by  $x$ .

**Definition 1** For  $x, y \in \mathbb{R}^k$  define

$$\rho_B(x, y) = \inf\{r \geq 0 : y \in (rB)_x\}$$

to be the ‘distance’ from  $x$  to  $y$  with respect to the test set  $B$ . For  $A \subset \mathbb{R}^k$  define

$$\begin{aligned} \rho_B(x, A) &= \inf\{\rho_B(x, a) : a \in A\} \\ &= \inf\{r \geq 0 : (rB)_x \cap A \neq \emptyset\} \end{aligned}$$

the shortest ‘distance’ from  $x$  to  $A$ .

For example if  $B$  is the unit ball in  $\mathbb{R}^k$  then  $\rho_B(x, y) = \|x - y\|$  and  $\rho_B(x, A)$  is the shortest Euclidean distance from  $x$  to  $A$ . If  $B$  is convex, symmetric about the origin 0, and contains a neighbourhood of 0, then  $\rho_B$  is a metric equivalent to the Euclidean metric, and  $B$  is its unit ball. The concept of the generalised distance  $\rho_B(x, X)$  is illustrated in Figure 1.

**Definition 2** Let  $X$  be a stationary random set in  $\mathbb{R}^k$ . Define the coverage fraction

$$p_X = \mathbb{P}\{0 \in X\}$$

and the first contact distribution

$$F_B(r) = \mathbb{P}\{\rho_B(0, X) \leq r\}, \quad r \geq 0.$$

By Fubini (Robbins’ Theorem, see Kendall and Moran (1963, pp. 109-111))

$$p_X = \frac{\mathbb{E} |X \cap Z|_k}{|Z|_k}$$

for any measurable set  $Z \subset \mathbb{R}^k$  with  $|Z|_k > 0$  where  $|\cdot|_k$  denotes Lebesgue volume. The following constructions are useful in interpreting  $F_B$ .

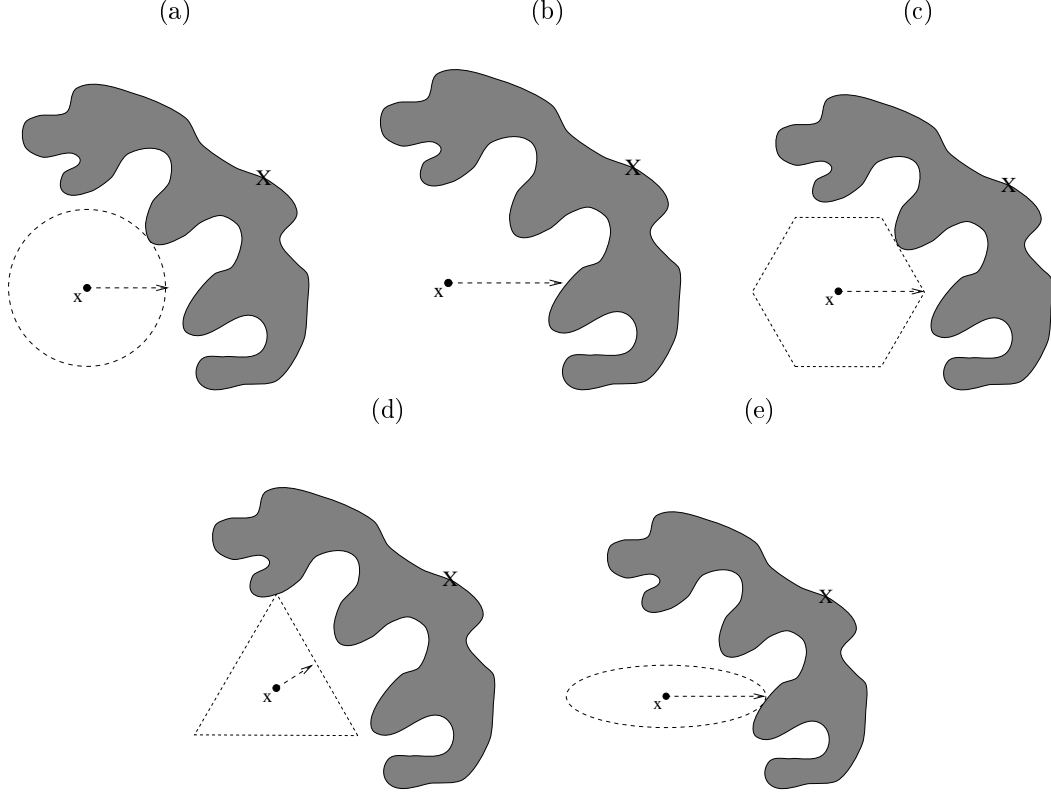


Figure 1: Drawings which illustrate the concept of the generalised distance  $\rho_B(x, X)$ . (a) spherical case; (b) linear case; (c) hexagonal case; (d) triangular case; (e) elliptical case.

**Definition 3** Let  $A, B \subset \mathbb{R}^k$ . Write

$$\check{A} = (-1)A = \{-a : a \in A\}$$

for the reflection of  $A$  through the origin. Define Minkowski addition  $\oplus$  and subtraction  $\ominus$  by

$$\begin{aligned} A \oplus B &= \{x + y : x \in A, y \in B\} \\ A \ominus B &= (A^c \oplus B)^c \end{aligned}$$

where  $^c$  denotes complement.

It can be shown that, if  $A \subset \mathbb{R}^k$  is closed,

$$\begin{aligned} \{x \in \mathbb{R}^k : \rho_B(x, A) \leq r\} &= A \oplus r\check{B}, \\ \{x \in \mathbb{R}^k : \rho_B(x, A^c) > r\} &= A \ominus r\check{B}. \end{aligned}$$

It follows by Robbins' Theorem as for  $p_X$  that

$$F_B(r) = \mathbb{P}\{\rho_B(0, X) \leq r\} = \mathbb{P}\{0 \in X \oplus r\check{B}\} = \frac{\mathbb{E}|Z \cap (X \oplus r\check{B})|_k}{|Z|_k}. \quad (1)$$

In the important spherical case with  $B = B(0, 1)$ , where  $B(x_0, r)$  denotes the closed Euclidean ball in  $\mathbb{R}^k$  centered at  $x_0 \in \mathbb{R}^k$  and with radius  $r \in \mathbb{R}_+$ , we have  $\check{B} = B$  and use the abbreviated notation  $A_{\oplus r}, A_{\ominus r}$  and  $F(r)$  for  $A \oplus r\check{B}, A \ominus r\check{B}$  and  $F_B$ , respectively.

A popular and convenient model of random sets is the Boolean model. It is used as a basic model for amorphous structure in many applications, see Stoyan et al. (1995) and references therein. Let  $\Phi_\mu$  be a stationary Poisson point process in the Euclidean space  $\mathbb{R}^k$  with intensity  $\mu$ , and let  $X_0, X_1, \dots$  be i.i.d. random compact sets independent of  $\Phi_\mu$ . Then

$$X = \cup_{x_i \in \Phi_\mu} (x_i \oplus X_i)$$

is said to be a Boolean model. As shown in Matheron (1975) the capacity functional of the Boolean model  $X$  can be evaluated as

$$T(K) = 1 - \exp[-\mu \mathbb{E}|X_0 \oplus \check{K}|_k].$$

For the Boolean model described above the expression for  $F_B$  becomes analytically tractable if  $B$  is convex and  $X_0$  is taken to be convex with probability one, and  $X_0$  has a distribution invariant with respect to rotations around the origin. It can namely be shown by the generalized Steiner formula (Stoyan et al., 1995, pp. 173) that

$$F_B(r) = 1 - \exp \left[ -\frac{\mu}{\omega_k} \sum_{d=0}^k \binom{k}{d} r^d \mathbb{E} (W_d(X_0)) W_{k-d}(B) \right], \quad (2)$$

where  $\omega_k$  is the volume of the unit ball in  $\mathbb{R}^k$  and  $W_d$  denotes the  $d$ 'th Minkowski functional.

From (2) it is seen that if  $X$  is a Boolean model built up with a.s. convex  $X_i$ 's, then  $F_B$  is absolutely continuous, continuously differentiable and has the hazard rate

$$\begin{aligned} \lambda_B(r) &= -\frac{d}{dr} \log(1 - F_B(r)) = \frac{f_B(r)}{1 - F_B(r)} \\ &= \frac{\mu}{\omega_k} \sum_{d=1}^k d \binom{k}{d} r^{d-1} \mathbb{E} (W_d(X_0)) W_{k-d}(B). \end{aligned}$$

The aim of this paper is to study similar properties for a general stationary random set.

### 3 Regularity properties of $F_B$

The aim of this section is to prove  $F_B$  is absolutely continuous. We shall use  $\mathcal{H}^{k-1}$  and  $|\cdot|_{k-1}$  to denote  $k-1$  dimensional surface area (Hausdorff) measure.

**Theorem 1** *Let  $X$  be a stationary random closed set and  $B$  a convex compact set containing a neighbourhood of the origin. Then*

1. *the first contact distribution function  $F_B$  for  $X$  is absolutely continuous for  $r > 0$  and has an atom at 0 of mass  $p_X$ .*

2. The density of  $F_B$  equals

$$f_B(r) = \mathbb{E} \int_{W \cap \partial(X \oplus r\tilde{B})} J_1 \rho_B(x, X)^{-1} \mathcal{H}^{k-1}(dx) \quad (3)$$

for almost every  $r > 0$  and any compact window  $W \subset \mathbb{R}^k$ , with  $|W|_k > 0$ .

Later in the paper we base estimators of  $F_B$  on the hazard rate, which from the result in Theorem 1 and (1) is given by

$$\lambda_B(r) = \frac{1}{\mathbb{E} |W \setminus (X \oplus r\tilde{B})|_k} \mathbb{E} \int_{W \cap \partial(X \oplus r\tilde{B})} J_1 \rho_B(x, X)^{-1} \mathcal{H}^{k-1}(dx). \quad (4)$$

To prove Theorem 1 we use some basic techniques of geometric measure theory with special emphasis on the coarea formula (see Federer (1969) and Morgan (1988); see also Simon (1984), Hardt and Simon (1986), and Lawrence and Gariepy (1992)). An intuitive explanation of the coarea formula (Federer, 1969, Theorem 3.2.22; Morgan, 1988, 3.13) can be given as follows: According to Fubini's Theorem the area of a two-dimensional set can be found by dividing the set into thin strips parallel to one of the axes and integrating the areas of the strips. The coarea formula generalises this idea. The strips are now determined by the level sets of some function  $f : \mathbb{R}^k \rightarrow \mathbb{R}$  which may be curved or even worse only Lipschitz instead of straight lines parallel to the axis. The Jacobian of  $f$  is a corrective factor needed to capture the change in elementary area at each point of the level sets. One should notice that we possibly get Jacobians of mappings to lower (or higher) dimensional spaces. Therefore to apply the coarea formula we need a Lipschitzian property of the function  $\rho_B(\cdot, A) : \mathbb{R}^k \rightarrow \mathbb{R}_+$ , and a lower bound on its Jacobian, which are given by Lemma 1 and 2, below.

**Lemma 1** *Let  $B$  be a convex compact set in  $\mathbb{R}^k$  containing a neighbourhood of the origin and choose any  $u > 0$  such that  $B(0, u) \subset B$ . Then  $\rho_B(\cdot, A) : \mathbb{R}^k \rightarrow \mathbb{R}_+$  is Lipschitz with constant  $u^{-1}$ , for any set  $A \subset \mathbb{R}^k$ .*

**Proof** Let  $x, y, z \in \mathbb{R}^k$  and define

$$\begin{aligned} r &= \rho_B(x, y), \\ s &= \rho_B(y, z). \end{aligned}$$

Then there exist  $b_r, b_s \in B$  such that

$$\begin{aligned} x + rb_r &= y \\ y + sb_s &= z. \end{aligned}$$

Hence

$$x + (rb_r + sb_s) = z,$$

and

$$x + (r + s)\left(\frac{rb_r}{r + s} + \frac{sb_s}{r + s}\right) = z.$$

It follows that

$$\rho_B(x, z) \leq r + s = \rho_B(x, y) + \rho_B(y, z),$$

so  $\rho_B(\cdot, \cdot) : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}_+$  satisfies the triangle inequality. From this we get

$$|\rho_B(x, y) - \rho_B(z, y)| \leq \max\{\rho_B(x, z); \rho_B(z, x)\},$$

and from the choice of  $u > 0$  such that  $B(0, u) \subset B$  we get that

$$\rho_B(x, y) \leq \frac{\|x - y\|}{u}.$$

Now, if we fix  $y \in \mathbb{R}^k$ , then

$$\begin{aligned} |\rho_B(x, y) - \rho_B(z, y)| &\leq \max\{\rho_B(x, z), \rho_B(z, x)\} \\ &\leq \frac{\|x - z\|}{u}. \end{aligned}$$

Moreover

$$\begin{aligned} \rho_B(x, A) &= \inf_{a \in A} \rho_B(x, a) \\ &\leq \inf_{a \in A} [\rho_B(x, z) + \rho_B(z, a)] \\ &\leq \rho_B(x, z) + \rho_B(z, A), \end{aligned}$$

from which the Lipschitz property follows

$$|\rho_B(x, A) - \rho_B(z, A)| \leq \frac{\|x - z\|}{u}.$$

□

Notice that, for example if  $A = \{y\} \in \mathbb{R}^2$  and  $B$  is a unit cube, then  $\rho_B(\cdot, A)$  is not differentiable along the diagonals of  $B$  but is differentiable at all other points. For any closed set  $A$ ,  $\rho_B(\cdot, A)$  is differentiable at almost every points by Lemma 1 and Rademachers Theorem (Federer, 1969, 3.1.6; Morgan, 1988, 3.2). If  $\rho_B(\cdot, A)$  is differentiable at  $x \in A^c$  we define the 1-dimensional Jacobian (Morgan, 1988, 3.6) as

$$J_1 \rho_B(x, A) = \sqrt{D_1 \rho_B(x, A) D_1 \rho_B(x, A)^*}, \quad (5)$$

where  $D_1 \rho_B(x, A) = (\partial \rho_B(x, A) / \partial x_1, \dots, \partial \rho_B(x, A) / \partial x_k)$  is the gradient of  $\rho_B(x, A)$  and  $*$  denotes transpose.

**Lemma 2** *Let  $B$  be a convex compact set in  $\mathbb{R}^k$  containing a neighbourhood of the origin and choose any  $U > 0$  such that  $B \subset B(0, U)$ . Then for any closed set  $A \subset \mathbb{R}^k$ ,*

$$J_1 \rho_B(x, A) \geq \frac{1}{U} > 0, \quad \text{for almost all } x \in A^c.$$

**Proof** Let us first define

$$c_v = \sup\{t \geq 0 : t \frac{v}{|v|} \in B\}, \quad \text{for } v \in \mathbb{R}^k.$$

Assume  $t = \rho_B(x, A)$  and  $z \in A$  is chosen such that  $\rho_B(x, z) = t$ . Then given  $x \in A^c$ , we want to show that for  $\epsilon \leq |z - x|$  we have

$$(tB)_x \supseteq \left[ \left( 1 - \frac{\epsilon}{c_{z-x}t} \right) tB \right]_{x+\epsilon w}, \quad (6)$$

$$z \in \left[ \left( 1 - \frac{\epsilon}{c_{z-x}t} \right) tB \right]_{x+\epsilon w}, \quad (7)$$

where  $w = (z - x)/|z - x|$ . First notice that

$$t = \frac{|z - x|}{c_{z-x}}. \quad (8)$$

To achieve (6), we want to prove that

$$b \in B \Rightarrow \left( 1 - \frac{\epsilon}{c_{z-x}t} \right) tb + x + \epsilon w \in (tB)_x,$$

which is equivalent to

$$\left( 1 - \frac{\epsilon}{c_{z-x}t} \right) b + \frac{\epsilon}{t} w \in B.$$

Using (8) we get  $c_{z-x} = |z - x|/t$ . Hence we want to prove

$$\left( 1 - \frac{\epsilon}{|z - x|} \right) b + \frac{\epsilon}{|z - x|} (wc_{z-x}) \in B.$$

As  $b \in B$  and  $wc_{z-x} \in B$ , we get for  $\epsilon \leq |z - x|$  a convex combination of elements in  $B$ , so (6) holds. Moreover taking  $b = wc_{z-x} \in B$  we get

$$\left( 1 - \frac{\epsilon}{c_{z-x}t} \right) tb + x + \epsilon w = z,$$

whereby (7) follows. By (6) and (7), we have for  $t$  and  $z$  defined as above

$$\begin{aligned} J_1 \rho_B(x, A) &= \| D_1 \rho_B(x, A) \| = \limsup_{y \rightarrow x} \frac{|\rho_B(y, A) - \rho_B(x, A)|}{|y - x|} \\ &\geq \limsup_{\epsilon \rightarrow 0} \frac{|\rho_B(x + \epsilon w, A) - \rho_B(x, A)|}{\epsilon} \\ &= \limsup_{\epsilon \rightarrow 0} \frac{|t - \frac{\epsilon}{c_{z-x}} - t|}{\epsilon} = \frac{1}{c_{z-x}} \geq \frac{1}{U} \\ &> 0. \end{aligned}$$

□



Note that Lemma 2 and equation (3) imply that the boundary of  $X \oplus r\check{B}$  has finite expected measure in any window,

$$\mathbb{E}|W \cap \partial(X \oplus r\check{B})|_{k-1} < \infty$$

for almost all  $r > 0$ . This is not guaranteed for **all**  $r$  and so for example it is still possible that  $\partial X$  may be a fractal with Hausdorff dimension greater than  $k-1$ . The theorem implies that these exceptional  $r$  values form a set of measure zero.

The proof of Theorem 1 rests on the following application of geometric measure theory (Federer, 1969; Morgan, 1988).

**Lemma 3** *Let  $Z \subset \mathbb{R}^k$  be a measurable set with  $0 < |Z|_k < \infty$ ,  $B$  a convex compact set containing a neighbourhood of the origin, and let  $A \subset \mathbb{R}^k$  be any closed set. Then the function  $r \mapsto |Z \cap (A \oplus r\check{B})|_k$  is nondecreasing, continuous and absolutely continuous with*

$$|Z \cap (A \oplus r\check{B})|_k = |Z \cap A|_k + \int_0^r \int_{Z \cap \partial(A \oplus s\check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx) ds. \quad (9)$$

*In particular the integrand is measurable and integrable.*

**Proof** The intuitive interpretation is as follows. The  $k$ -dimensional volume is broken into strips defined by the level sets of  $\rho_B(\cdot, A)$  (equivalent to the boundaries of  $\partial(A \oplus s\check{B})$ ,  $0 \leq s \leq r$ ), going from  $\partial A$  to  $\partial(A \oplus r\check{B})$ .

Formally the function  $\rho_B(\cdot, A) : \mathbb{R}^k \rightarrow \mathbb{R}_+$  is Lipschitz by Lemma 1, and hence a.e. differentiable with positive 1-dimensional Jacobian by Lemma 2,

$$J_1 \rho_B(x, A) = \|D_1 \rho_B(x, A)\| > 0, \quad x \in A^c.$$

Since  $Z$  is assumed measurable, we use the coarea formula (Federer, 1969, Theorem 3.2.22) and the observation that up to a null set  $\partial(A \oplus s\check{B}) = \rho_B^{-1}(s, A)$ , where  $\rho_B^{-1}(s, A) = \{x \in \mathbb{R}^k : \rho(x, A) = s\}$  denotes a level set of  $\rho_B(\cdot, A)$ , to get

$$\begin{aligned} |Z \cap (A \oplus r\check{B})|_k &= |Z \cap A|_k + |Z \cap ((A \oplus r\check{B}) \setminus A)|_k \\ &= |Z \cap A|_k + \int_Z \frac{1_{\{x \in A \oplus r\check{B} \setminus A\}}}{J_1 \rho_B(x, A)} J_1 \rho_B(x, A) dx \\ &= |Z \cap A|_k + \int_0^r \int_{Z \cap \partial(A \oplus s\check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx) ds. \end{aligned}$$

Monotonicity follows from (9). □

**Proof of Theorem 1** As  $W$  is compact Lemma 3 gives a.s.

$$|W \cap (X \oplus r\check{B})|_k = |W \cap X|_k + \int_0^r \int_{W \cap \partial(X \oplus s\check{B})} J_1 \rho_B(x, X)^{-1} \mathcal{H}^{k-1}(dx) ds.$$

Since the left side is integrable, Fubini's theorem gives

$$\mathbb{E}|W \cap (X \oplus r\check{B})|_k = \mathbb{E}|W \cap X|_k + \int_0^r \mathbb{E} \int_{W \cap \partial(X \oplus s\check{B})} J_1 \rho_B(x, X)^{-1} \mathcal{H}^{k-1}(dx) ds.$$

The conclusions follow using equation (1).  $\square$

In the rest of this subsection we record some stronger statements about the regularity of  $F_B$ . However, for this we need the following result about the uniform boundedness of  $|Z \cap \partial(\Phi \oplus r\check{B})|_{k-1}$ . This proposition is a generalization of Lemma 2 in Baddeley and Gill (1993), to the case of a general convex body as test set.

**Proposition 1** *Let  $\Phi$  be a locally finite point process on  $\mathbb{R}^k$ . Then for any compact set  $Z \subset \mathbb{R}^k$  and convex compact body  $B$  with  $u, U > 0$  chosen such that  $B(0, u) \subset B \subset B(0, U)$ ,*

$$|Z \cap \partial(\Phi \oplus r\check{B})|_{k-1} \leq \frac{k}{ur} |Z \oplus r\check{B}|_k \wedge k\omega_k(Ur)^{k-1} \Phi(Z \oplus r\check{B}) \quad (10)$$

where  $\#$  denotes the cardinality of a set, and  $\wedge$  denotes minimum.

**Proof** As the surface area of the boundary of a convex set is increasing with respect to set inclusion and the additivity of the measure  $\mathcal{H}^{k-1}$ , the second term on the right of (10) follows since  $k\omega_k(Ur)^{k-1} = |\partial B(0, Ur)|_{k-1}$ , i.e. the surface area of a ball circumscribing the expanded structuring element  $r\check{B}$ . For the first term, fix a realization of  $\Phi$  and let  $x_i$ ,  $i = 1, \dots, m$  be the almost surely distinct points in  $\Phi \cap (Z \oplus r\check{B})$ . Now let  $H_i = (r\check{B})_{x_i}$ ,  $i = 1, \dots, m$ , and make the following disjoint partitioning of the union of boundaries  $\cup_{i=1}^m \partial H_i$  in the following way

$$\begin{aligned} D_1 &= \partial H_1 \\ D_i &= (\cap_{j=1}^i H_j^c) \cap \partial H_i, \quad i = 2, \dots, m. \end{aligned}$$

By noticing that  $Z \cap \partial(\Phi \oplus r\check{B}) \subset Z \cap (\cup_{i=1}^m D_i)$ , the argument is as follows: any line segment joining  $x_i$  to a point on the corresponding surface piece  $D_i \cap Z$  is contained entirely within  $(r\check{B})_{x_i}$  since  $B$  is convex. Moreover, two line segments joining  $x_i$  to  $D_i \cap Z$  and  $x_j$  to  $D_j \cap Z$ , for  $i \neq j$ , cannot intersect according to the definition of  $D_i$ . Hence the unions  $F_i$  of these segments are disjoint, so the sum of the volumes  $|F_i|_k$  is bounded by the volume of  $Z \oplus r\check{B}$ :

$$\sum_{i=1}^m |F_i|_k \leq |Z \oplus r\check{B}|_k. \quad (11)$$

Now to provide a bound on  $|D_i|_{k-1}$  define  $f : \mathbb{R}^k \rightarrow S^{k-1}$  by  $f(x) = x/|x|$ , where  $S^{k-1}$  denotes the  $k-1$ -dimensional boundary of the unit sphere in  $\mathbb{R}^k$ . The  $k$ -dimensional Jacobian of  $f$  is found by representing the tangent space of  $\mathbb{R}^k$  at  $x$  as the orthogonal sum  $T = H + V$  where  $H = \{y \in \mathbb{R}^k : x \cdot y = 0\}$  is the hyperplane tangent at  $x$  to the sphere of radius  $|x|$ , centre 0, and  $V = \{rx : r \in \mathbb{R}\}$  is the ray from 0 through  $x$ . Take an orthonormal basis  $u_1, \dots, u_k$  for  $T$  consisting of an orthonormal basis  $u_1, \dots, u_{k-1}$  for  $H$  and a unit vector in  $V$ . By taking directional derivatives  $\lim_{\epsilon \rightarrow 0} (f(x + \epsilon u_j) - f(x))/\epsilon$  we find the differential of  $f$  at  $x$  has

$$Df(x)(u_1) = u_1/|x|$$

$$\begin{aligned}
& \vdots \\
Df(x)(u_{k-1}) &= u_{k-1}/|x| \\
Df(x)(u_k) &= u_k
\end{aligned}$$

i.e.  $Df(x)$  is a  $k$ -dimensional matrix with determinant  $1/|x|^{k-1}$ . So the Jacobian is  $J_k f(x) = |x|^{1-k}$ . By the coarea formula we get for  $A = f(F_i \cap Z) = f(D_i \cap Z)$

$$\begin{aligned}
|D_i \cap Z|_{k-1} &= \int_{D_i \cap Z} \mathcal{H}^{k-1}(dx) \\
&= \int_A \int_{f^{-1}(z) \cap D_i \cap Z} (J_k f(y))^{-1} \mathcal{H}^0(dy) \mathcal{H}^{k-1}(dz) \\
&= \int_A |f^{-1}(z) \cap D_i \cap Z|^{k-1} \mathcal{H}^{k-1}(dz)
\end{aligned}$$

and

$$\begin{aligned}
|F_i \cap Z|_k &= \int_{F_i \cap Z} \mathcal{H}^{k-1}(dx) \\
&= \int_A \int_{f^{-1}(z) \cap F_i \cap Z} (J_k f(y))^{-1} \mathcal{H}^1(dy) \mathcal{H}^{k-1}(dz) \\
&= \frac{1}{k} \int_A |f^{-1}(z) \cap F_i \cap Z|_1^k \mathcal{H}^{k-1}(dz) \\
&= \frac{1}{k} \int_A |f^{-1}(z) \cap D_i \cap Z|^k \mathcal{H}^{k-1}(dz). \\
&\geq \frac{ur}{k} \int_A \mathcal{H}^{k-1}(dz).
\end{aligned}$$

Hence

$$|F_i \cap Z|_k \geq \frac{u}{k} r |D_i \cap Z|_{k-1}. \quad (12)$$

Combining (11) and (12) gives

$$\begin{aligned}
& |Z \cap \partial(\Phi \oplus r\check{B})|_{k-1} \\
& \leq \sum_{i=1}^m |D_i \cap Z|_{k-1} \leq \frac{k}{ur} \sum_{i=1}^m |F_i \cap Z|_k \leq \frac{k}{ur} \sum_{i=1}^m |F_i|_k \\
& \leq \frac{k}{ur} |Z \oplus r\check{B}|_k.
\end{aligned}$$

□

**Proposition 2** *Let  $Z \subset \mathbb{R}^k$  be any compact set,  $B$  a convex compact set containing a neighbourhood of the origin and  $u, U > 0$  chosen such that  $B(0, u) \subset B \subset B(0, U)$ . Then for any closed set  $A$*

$$\begin{aligned}
\int_{Z \cap \partial(A \oplus r\check{B})} J_1 \rho_B(x, A) \mathcal{H}^{k-1}(dx) &\leq U |Z \cap \partial(A \oplus r\check{B})|_{k-1} \\
&\leq \frac{U^2 k}{ur} |Z \oplus r\check{B}|_k, \quad \text{for a.e. } r > 0.
\end{aligned}$$

**Proof** The first inequality follows directly from Lemma 2. For the second, let  $\epsilon > 0$  be given. Since  $A$  is compact there is a finite set  $C = \{x_1, \dots, x_n\} \subset A$  such that

$$\sup_{x \in A} \min_i \rho_B(x_i, x) \leq \epsilon.$$

Equivalently  $A \subset C \oplus \epsilon \check{B}$ . Hence by monotonicity and semigroup properties of dilation (Matheron, 1975)

$$C \oplus r \check{B} \subseteq A \oplus r \check{B} \subseteq C \oplus ((r + \epsilon) \check{B})$$

for any  $r > 0$ . Thus for  $r_2 > r_1 \geq 0$

$$(A \oplus r_2 \check{B}) \setminus (A \oplus r_1 \check{B}) \subseteq (C \oplus ((r_2 + \epsilon) \check{B})) \setminus (C \oplus r_1 \check{B})$$

giving

$$|((A \oplus r_2 \check{B}) \setminus (A \oplus r_1 \check{B})) \cap Z|_k \leq |(C \oplus ((r_2 + \epsilon) \check{B})) \setminus (C \oplus r_1 \check{B})) \cap Z|_k.$$

An application of Lemma 3 yields

$$\int_{r_1}^{r_2} \int_{Z \cap \partial(A \oplus s \check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx) ds \leq \int_{r_1}^{r_2 + \epsilon} \int_{Z \cap \partial(C \oplus s \check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx) ds.$$

Now use the fact that  $U^{-1} \leq J_1 \rho_B(x, A) \leq u^{-1}$ , (which follows from Lemma 1 and 2), to get

$$\begin{aligned} \int_{r_1}^{r_2} |Z \cap \partial(A \oplus s \check{B})|_{k-1} ds &\leq \int_{r_1}^{r_2} \int_{Z \cap \partial(A \oplus s \check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx) ds \\ &\leq \int_{r_1}^{r_2 + \epsilon} \int_{Z \cap \partial(C \oplus s \check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx) ds \\ &\leq U \int_{r_1}^{r_2 + \epsilon} |Z \cap \partial(C \oplus s \check{B})|_{k-1} ds. \end{aligned}$$

Now apply Proposition 1 to bound the integrand on the right hand side by  $\frac{kU}{su} |Z \oplus s \check{B}|_k$ . Since  $\epsilon$  is arbitrary and  $s \rightarrow |Z \oplus s \check{B}|_k$  is continuous and nondecreasing we get

$$\int_{r_1}^{r_2} |Z \cap \partial(A \oplus s \check{B})|_{k-1} ds \leq U \int_{r_1}^{r_2} \frac{kU}{us} |Z \oplus s \check{B}|_k ds.$$

Since  $r_1$  and  $r_2$  are arbitrary the conclusion follows from the Radon-Nikodým theorem.  $\square$

**Proposition 3** *If the map  $r \mapsto \int_{Z \cap \partial(A \oplus r \check{B})} J_1 \rho_B(x, A)^{-1} \mathcal{H}^{k-1}(dx)$  is a.s. continuous, then  $F_B$  is continuously differentiable.*

For example, the condition is satisfied by any point process or line segment process, or more generally when every realization of  $X$  is a locally finite union of convex compact sets (e.g. a Boolean model).

**Proof** Once the uniform bound in Proposition 2 is established, the result follows by invoking the dominated convergence theorem.  $\square$

*Remark 1 (The spherical case)* Consider the spherical case,  $B = B(0, 1)$ ; we have up to a null set  $A_{\oplus r} = \rho_B^{-1}([0, r], A)$  and  $\rho_B(\cdot, A)$  is Lipschitz with constant 1, hence almost everywhere differentiable, with Jacobian bounded by 1. Under the conditions of Proposition 3 the Jacobian is a.s. equal to 1, and the density and hazard rate of  $F$  simplifies to

$$f(r) = \mathbb{E} |W \cap \partial X_{\oplus r}|_{k-1}.$$

$$\lambda(r) = \frac{1}{\mathbb{E} |W \setminus X_{\oplus r}|_k} \mathbb{E} |W \cap \partial X_{\oplus r}|_{k-1}.$$

*Remark 2 (Importance of the Jacobian)* To emphasize the importance of the Jacobian in Lemma 3 we consider the following simple example in  $\mathbb{R}^2$ . Let  $Z = B(0, 3)$ ,  $A = \{0\}$ , and  $B = \{x \in \mathbb{R}^2 : \|x\|_1 \leq 1\}$ , where  $\|\cdot\|_1$  is the  $L^1$ -norm, see Figure 2. Then we get

$$|Z \cap (A \oplus r\check{B})|_k = (\sqrt{2}r)^2 = 2r^2, \quad 0 \leq r \leq 3.$$

But if we set the Jacobian to 1 as in the spherical case shown above (Baddeley and Gill, 1993, 1997) then

$$\begin{aligned} \int_0^r \int_{Z \cap \partial(A \oplus s\check{B})} \mathcal{H}^1(dx) ds &= \int_0^r |Z \cap \partial(A \oplus s\check{B})|_1 ds \\ &= \int_0^r 4\sqrt{2}s ds \\ &= 2\sqrt{2}r^2, \quad 0 \leq r \leq 3. \end{aligned}$$

This example is fairly innocuous in the sense that the error introduces only a constant factor  $\sqrt{2}$ . However, by e.g. shifting  $B$  to make the origin eccentric, the Jacobian becomes non constant and even more vital to be included.

*Remark 3 (A simplified expression of the Jacobian)* Note that the Jacobian can be interpreted as the rate of increase of  $X \oplus r\check{B}$  perpendicular to its boundary, see Baddeley (1977). Using this interpretation it is possible to derive a simplified expression of the Jacobian  $J_1 \rho_B(x, X)$  along the following sketch. Let  $h_B : S^{k-1} \rightarrow \mathbb{R}_+$  be the support function of  $B$  defined by

$$h_B(u) = \sup_{b \in B} u \cdot b,$$

which geometrically is the distance from 0 to the supporting hyperplane in the direction normal to  $u$ . Moreover let  $n(x)$  be the outward unit normal vector to  $\partial(X \oplus r\check{B})$  at  $x$  when  $r = \rho_B(x, X)$ . If we approximate the boundary  $\partial(X \oplus r\check{B})$  in a neighbourhood  $U$  of  $x$  by a line/plane normal to the vector  $n(x)$ , then a simple sketch shows that the boundaries of  $X \oplus r\check{B}$  and  $X \oplus (r+dr)\check{B}$  are approximately

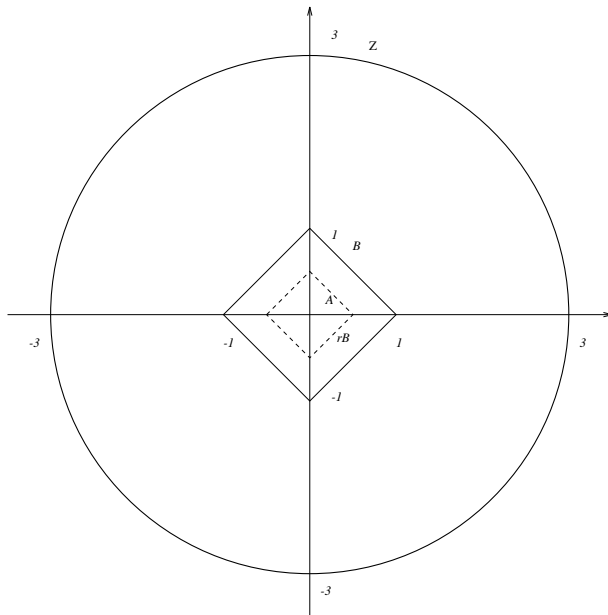


Figure 2: Illustration of a simple test set  $B$  centered at the origin.

parallel, and separated by a distance of approximately  $h_B(n(x))dr$ . So  $U \cap (X \oplus (r+dr)\check{B}) \setminus (X \oplus r\check{B})$  is approximately a slab of height  $h_B(n(x))$ , and

$$J_1 \rho_B(x, X) = h_B(n(x)).$$

## 4 Estimation of $F_B$

The problem considered in this section is to estimate  $F_B$  based on the observable data  $W \cap X$ . To estimate  $F_B$  it would be straightforward to suggest the empirical counterparts to (1), (3), and (4), but to determine  $(X \oplus r\check{B}) \cap W$  would require information outside  $W$ , which is unavailable.

We observe only  $X \cap W$ , and hence we know  $\rho_B(x, X \cap W)$  rather than  $\rho_B(x, X)$  for  $x \in W$ . The true distance  $\rho_B(x, X)$  is in effect censored by the distance to the boundary  $\rho_B(x, \partial W)$ , since

$$\rho_B(x, X) \wedge \rho_B(x, \partial W) = \rho_B(x, X \cap W) \wedge \rho_B(x, \partial W), \quad x \in W. \quad (13)$$

The standard approach to this problem, see Cressie (1991, Chapter 8), Stoyan et al. (1995, pp. 122-131), Serra (1982, pp. 49, 233) and Ripley (1988, Chapter 3), is to restrict attention, when estimating  $F_B(r)$ , to those points  $x$  with  $\rho_B(x, \partial W) \geq r$ , since for such points we have  $\rho_B(x, X) \leq r$  if and only if  $\rho_B(x, X \cap W) \leq r$ . In geometrically terms

$$(X \oplus r\check{B}) \cap (W \ominus r\check{B}) = ((X \cap W) \oplus r\check{B}) \cap (W \ominus r\check{B}) \quad (14)$$

and the right-hand side is computable from data  $X \cap W$ . Now define

$$\begin{aligned}\widehat{Y}(r) &= |(W \ominus r\check{B}) \setminus (X \oplus r\check{B})|_k \\ Y(r) &= \mathbb{E} \widehat{Y}(r).\end{aligned}$$

Thus replacing  $Z$  by  $W \ominus r\check{B}$  in (3), (4), and (1) gives

$$\begin{aligned}f_B(r) &= \mathbb{E} \int_{(W \ominus r\check{B}) \cap \partial(X \oplus r\check{B})} h_B(n(x))^{-1} \mathcal{H}^{k-1}(dx). \\ \lambda_B(r) &= \frac{1}{Y(r)} \mathbb{E} \int_{(W \ominus r\check{B}) \cap \partial(X \oplus r\check{B})} h_B(n(x))^{-1} \mathcal{H}^{k-1}(dx). \\ F_B(r) &= \frac{\mathbb{E} |(W \ominus r\check{B}) \cap (X \oplus r\check{B})|_k}{|W \ominus r\check{B}|_k}\end{aligned}\tag{15}$$

A plug-in estimator for e.g. the hazard rate would then be

$$\hat{\lambda}_B(r) = \frac{1}{\widehat{Y}(r)} \int_{(W \ominus r\check{B}) \cap \partial(X \oplus r\check{B})} h_B(n(x))^{-1} \mathcal{H}^{k-1}(dx)\tag{16}$$

Using (16) directly as an estimator for  $\lambda_B$  seems difficult because of the Jacobian involved. However, another application of the coarea formula gives the following expression for a possible estimator  $\widehat{\Lambda}_B(r)$  of the integrated hazard rate of  $F_B$ , where  $t(x) = \rho_B(x, X \cap W)$  is the distance to  $X \cap W$  and  $c(x) = \rho_B(x, \partial W)$  the distance to the boundary of  $W$

$$\begin{aligned}\widehat{\Lambda}_B(r) &= \int_0^r \hat{\lambda}_B(s) ds \\ &= \int_0^r \int_{\rho_B^{-1}(s)} \frac{1\{t(x) \leq c(x)\} 1\{t(x) \leq r\}}{|(W \ominus t(x)\check{B}) \setminus (X \oplus t(x)\check{B})|_k} h_B(n(x))^{-1} \mathcal{H}^{k-1}(dx) ds \\ &= \int_W \frac{1\{t(x) \leq c(x)\} 1\{t(x) \leq r\}}{|(W \ominus t(x)\check{B}) \setminus (X \oplus t(x)\check{B})|_k} dx\end{aligned}\tag{17}$$

Equation (15) and (17) then motivates the following definition

**Definition 4** *Let  $X$  be a stationary random closed set,  $B$  convex compact and containing a neighbourhood of the origin, and let  $W \subset \mathbb{R}^k$  be a compact set with  $|W|_k > 0$ . The Kaplan-Meier estimator  $\widehat{F}^{km}$  of the first contact distribution function  $F$  of  $X$ , based on data  $W \cap X$  in  $W$ , is defined by*

$$\begin{aligned}\widehat{F}^{km}(r) &= \\ &1 - \frac{|W \setminus X|_k}{|W|_k} \exp \left( - \int_W \frac{1\{t(x) \leq c(x)\} 1\{t(x) \leq r\}}{|(W \ominus t(x)\check{B}) \setminus (X \oplus t(x)\check{B})|_k} dx \right).\end{aligned}\tag{18}$$

The reduced sample estimator  $\widehat{F}^{rs}$  of the first contact distribution function  $F$  of  $X$ , based on data  $X \cap W$ , is defined by

$$\widehat{F}^{rs}(r) = \frac{|(W \ominus r\check{B}) \cap (X \oplus r\check{B})|_k}{|W \ominus r\check{B}|_k}.\tag{19}$$

We see immediately that  $\widehat{F}^{rs}$  is a pointwise unbiased estimator for  $F$ , but not necessarily a distribution function. The expression used for the estimator defined in equation (19) is variously known as “minus sampling” (Miles, 1974) or “border correction” (Ripley, 1988). But under the analogy set up between distance and survival analysis Baddeley and Gill (1993, 1997) applied the notion “reduced sample estimator”, which is standard terminology for an estimator of this type in censoring/survival analysis. Equation (14) is sometimes said to be an instance of the “local knowledge principle” (Serra, 1982, loc. cit.). To see that (17) deserves the epithet Kaplan-Meier we suppose for notational convenience that  $B$  is a sphere and get

$$\begin{aligned}\widehat{F}^{km}(r) &= 1 - \frac{|W \setminus X|_k}{|W|_k} \exp \left( - \int_W \frac{1\{t(x) \leq c(x)\} 1\{t(x) \leq r\}}{|W_{\ominus t(x)} \setminus X_{\oplus t(x)}|_k} dx \right) \\ &= 1 - \frac{|W \setminus X|_k}{|W|_k} \exp \left( - \int_0^r \frac{|\partial(X_{\oplus s}) \cap W_{\ominus s}|_{k-1}}{|W_{\ominus s} \setminus X_{\oplus s}|_k} ds \right)\end{aligned}\quad (20)$$

Now observe that  $t(x)$  and  $c(x)$  are the distance to failure and the censoring distance, respectively, and define the “observed failure distance”

$$\begin{aligned}\tilde{t}(x) &= t(x) \wedge c(x) \\ &= \rho_B(x, X \cap W) \wedge \rho_B(x, \partial W) \\ &= \rho_B(x, X) \wedge \rho_B(x, \partial W)\end{aligned}$$

and the censoring indicator

$$\begin{aligned}d(x) &= 1\{t(x) \leq c(x)\} \\ &= 1\{\rho_B(x, X \cap W) \leq \rho_B(x, \partial W)\} \\ &= 1\{\rho_B(x, X) \leq \rho_B(x, \partial W)\}\end{aligned}$$

where in each case the last line follows by (13). Then the integrand of (20) has denominator

$$|W_{\ominus s} \setminus X_{\oplus s}|_k = |\{x \in W : \tilde{t}(x) \geq s\}|_k,$$

the measure of the set of points “at risk” at distance  $s$ , and numerator

$$|\partial(X_{\oplus s}) \cap W_{\ominus s}|_{k-1} = |\{x \in W : \tilde{t}(x) = s, d(x) = 1\}|_{k-1},$$

the measure of the set of points observed to “fail” at distance  $s$ . Thus (20) is at least intuitively the analogue of the usual Kaplan-Meier estimator for the continuum of data  $\{(\tilde{t}(x), c(x)) : x \in W\}$ .

**Theorem 2** *For a stationary random closed set  $X$ ,  $B$  a convex compact set containing a neighbourhood of the origin, and  $W$  any compact set with  $|W|_k > 0$ , the statistic  $\widehat{F}^{km}$  is a distribution function. It is continuous and absolutely continuous for  $r > 0$ , with hazard rate*

$$\hat{\lambda}(r) = \frac{1}{\widehat{Y}(r)} \int_{(W_{\ominus r} \bar{B}) \cap \partial(X_{\oplus r} \bar{B})} h_B(n(x))^{-1} \mathcal{H}^{k-1}(dx) \quad (21)$$



and an atom at 0 of mass

$$\hat{F}^{km}(0) = \frac{|W \setminus X|_k}{|W|_k} = \hat{\rho}_X = \hat{F}^{rs}(0).$$

The estimator  $\hat{\lambda}(r)$  of  $\lambda(r)$  is ratio unbiased in the sense that

$$\lambda(r) = \frac{\mathbb{E} \int_{(W \ominus r\check{B}) \cap \partial(X \oplus r\check{B})} h_B(n(x))^{-1} \mathcal{H}^{k-1}(dx)}{Y(r)}.$$

**Proof** First we show that (18) is indeed the Kaplan-Meier estimator of  $F$  based on the continuum of observations  $\tilde{t}(x) = \rho_B(x, X) \wedge \rho_B(x, \partial W)$ ,  $d(x) = 1\{\rho_B(x, X) \leq \rho_B(x, \partial W)\}$ . This is defined as the product integral

$$\hat{F}^{km}(r) = 1 - \prod_0^r (1 - d\Lambda)$$

of the measure  $\hat{\Lambda}$  defined by

$$\hat{\Lambda}([0, r]) = \int_0^r \frac{d\hat{N}(s)}{\hat{Y}(s)}$$

where

$$\begin{aligned} \hat{N}(r) &= |\{x \in W : \tilde{t}(x) \leq r, d(x) = 1\}|_k \\ &= |\{x \in W : \rho_B(x, X) \leq (r \wedge \rho_B(x, \partial W))\}|_k \\ \hat{Y}(r) &= |\{x \in W : \tilde{t}(x) \geq r\}|_k \\ &= |(W \ominus r\check{B}) \setminus (X \oplus r\check{B})|_k. \end{aligned}$$

Rewrite  $\hat{N}(r) = |V \cap (X \oplus r\check{B})|_k$  where

$$V = \{x \in W : \rho_B(x, X) \leq \rho_B(x, \partial W)\}. \quad (22)$$

Hence the integral defining  $\hat{\Lambda}([0, r])$  exists and is continuous as a function of  $r$ , and  $\hat{\Lambda}$  has density (21) for a.e.  $r > 0$ . Hence the product integral exists, is continuous in  $r$ , and coincides with (18). The “ratio-unbiasedness” of  $\hat{\lambda}$  is a restatement of equation (4).  $\square$

Paralleling Proposition 3 is the fact that if  $r \mapsto |(W \ominus r\check{B}) \cap \partial(X \oplus r\check{B})|_{k-1}$  is continuous (e.g. if  $X \cap W$  is a finite set of points or a finite union of compact convex sets) then  $\hat{F}^{km}$  is continuously differentiable.

Given an image of a random structure it would be hard to calculate the integrals required in Definition 1 to estimate  $F$ . One would rather use a discretized sampling window  $W$ .

Let  $\mathbb{Z}$  be the set of integers and  $\mathbb{Z}_\epsilon = \{\epsilon m : m \in \mathbb{Z}\}$ ,  $\epsilon > 0$ . Then  $\mathbb{Z}_\epsilon^k = \mathbb{Z}_\epsilon \times \cdots \times \mathbb{Z}_\epsilon$  forms a lattice in  $\mathbb{R}^k$  with mesh  $\epsilon$ . Now, one can calculate for each lattice point  $z_i$  in  $W \cap X$  the following observations:  $t_i = \rho_B(z_i, X \cap W)$ ,  $c_i = \rho_B(z_i, \partial W)$  and  $\tilde{t}_i = t_i \wedge c_i$ ,  $d_i = 1\{t_i \leq c_i\}$ . If the notion of survival analysis is adopted this is a random censorship model, see e.g. Andersen et al. (1993),

$t_i, c_i, \tilde{t}_i$  and  $d_i$  are normally called survival time, censoring time, censored survival time and edge censoring indicator, respectively. Note the immediate difference from standard random censorship models in survival analysis, that  $t_i$  and  $c_i$  are not independent for different  $i$ 's, however a natural possibility is to construct the discrete Kaplan-Meier estimator and look at the behaviour for the lattice mesh tending to zero. Following the lines of Hansen et al. (1996, Theorem 3) we can prove the following properties of the estimator.

**Theorem 3** *Construct the discrete Kaplan-Meier estimator*

$$\hat{F}_\epsilon^{km}(r) = 1 - \prod_{s \leq r} \left( 1 - \frac{\#\{i : \tilde{t}_i = s, d_i = 1\}}{\#\{i : \tilde{t}_i \geq s\}} \right) \quad (23)$$

and the discrete reduced sample estimator

$$\hat{F}_\epsilon^{rs}(r) = \frac{\#\{i : t_i \leq r \leq c_i\}}{\#\{i : c_i \geq r\}}, \quad (24)$$

and assume

$$|\partial V|_k = 0, \quad (25)$$

where  $V$  was defined in (22). Then as the lattice mesh  $\epsilon$  converges to zero,  $\hat{F}_\epsilon^{km}(r) \rightarrow \hat{F}^{km}(r)$  and  $\hat{F}_\epsilon^{rs}(r) \rightarrow \hat{F}^{rs}(r)$  for any  $r < R$ , where

$$R = \inf\{r \geq 0 : (W \ominus r\check{B}) \cap (X \oplus r\check{B}) = \emptyset\}.$$

Moreover the convergence is uniform on any compact interval in  $[0, R)$ .

**Proof** As long as

$$|\partial(W \ominus r\check{B})|_k = 0, \quad (26)$$

$$|\partial((W \ominus r\check{B}) \cap (X \oplus r\check{B}))|_k = 0, \quad (27)$$

and

$$|\partial(V \cap (X \oplus r\check{B}))|_k = 0, \quad (28)$$

holds, the proof is identical to Hansen et al. (1996, Theorem 3) as we note the proof herein does not depend on the structuring element  $B$ . We see that (26) and (27) are guaranteed for almost all  $r$  by the coarea formula. Finally, (28) is guaranteed by (25).  $\square$

It is of practical interest to see under what conditions (25) is true. For example assume that  $B$  is a convex structuring element,  $X$  is a Boolean model with convex grains and  $W$  is a convex observation window. Then (26)–(28) follows from Hansen et al. (1996, Lemma 4) as the proof of this lemma only depends on the convexity of  $B$ . However by taking for instance any point process as an underlying process for the Boolean model (25) is also true, so there seems to be room for more generalisations.

## 5 Final remarks

In this paper we have proved some regularity properties for the first contact distribution function for test sets in the class of convex bodies, by means of the coarea formula. It is difficult to use Theorem 1 directly to derive analytical expressions for  $F_B$  for different random set models because of the Jacobian involved, so in practice one would use Theorem 3 to compute estimates. The main practical implication of Theorem 1 is the absolute continuity of  $F_B$ , signifying that it is sensible to estimate the density and hazard rate. Additionally Theorem 1 is the first step to derivation of the Kaplan-Meier type estimator for  $F_B$  which in some situations has shown to give more efficient estimators, see Baddeley and Gill (1993, 1997) and Hansen et al. (1996).

## Acknowledgments

We thank the referee for a very thorough report on our paper. AJB's research was in part funded by the University of Leiden and Centrum voor Wiskunde en Informatica, Amsterdam. MBH's research was funded by Dina (Danish Informatics Network in the Agricultural Sciences) and Føtek (Centre for Food Research, RVAU).

## References

- Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.
- Baddeley, A. (1977). Integrals on a moving manifold and geometrical probability. *Adv. Appl. Prob.*, 9:588–603.
- Baddeley, A. J. and Gill, R. D. (1993). Kaplan-Meier estimators of interpoint distance distributions for spatial point processes. Research Report BS-R9315, Centrum voor Wiskunde en Informatica.
- Baddeley, A. J. and Gill, R. D. (1997). Kaplan-Meier estimators of interpoint distance distributions for spatial point processes. *Annals of Statistics*.
- Cressie, N. A. C. (1991). *Statistics for Spatial Data*. John Wiley and Sons, New York.
- Diggle, P. (1983). *Statistical Analysis of Spatial Point Patterns*. Academic Press, London.
- Federer, H. (1969). *Geometric Measure Theory*. Springer-Verlag, New York.
- Gill, R. D. (1994). Lectures on survival analysis. In Bernard, P., editor, *Lectures on Probability Theory, Ecole d'Été de Probabilités de Saint-Flour XXII 1992*. Lecture Notes in Mathematics. Springer-Verlag.
- Hansen, M. B., Baddeley, A. J., and Gill, R. D. (1996). Kaplan-Meier type estimators for linear contact distributions. *Scand. J. Statist.*, 23:129–155.

- Hardt, R. and Simon, L. (1986). *Seminar on Geometric Measure Theory*. Birkhäuser, Basel.
- Kendall, M. G. and Moran, P. A. P. (1963). *Geometrical Probability*. Charles Griffin, London. Griffin's Statistical Monographs and Courses no. 10.
- Lawrence, C. E. and Gariepy, R. G. (1992). *Measure Theory and Fine Properties of Functions*. CRC Press.
- Matheron, G. (1975). *Random Sets and Integral Geometry*. John Wiley and Sons, New York.
- Miles, R. G. (1974). On the elimination of edge effects in planar sampling. In Harding, E. F. and Kendall, D. G., editors, *Stochastic geometry: A tribute to the memory of Rollo Davidson*, pages 228–247. John Wiley and Sons, New York.
- Morgan, F. (1988). *Geometric Measure Theory: a beginners guide*. Academic Press, Boston.
- Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press.
- Serra, J. (1982). *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, London.
- Simon, L. M. (1984). Lectures on geometric measure theory. In *Proceedings of the Centre for Mathematical Analysis, Australian National University*, volume 3. Australian National University, Canberra.
- Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and Its Applications*. John Wiley and Sons, Chichester, 2nd edition.