Ann Inst Stat Math (2008) 60:813–842
DOI 10.1007/s10463-007-0123-7

Local empirical processes near boundaries of convex bodies

Estate Khmaladze · Wolfgang Weil

Received: 20 April 2005 / Revised: 30 October 2006 / Published online: 24 March 2007

Abstract We investigate the behaviour of Poisson point processes in the neighbourhood of the boundary ∂K of a convex body K in $\mathbb{R}^d, d \geq 2$. Making use of the geometry of K, we show various limit results as the intensity of the Poisson process increases and the neighbourhood shrinks to ∂K. As we shall see, the limit processes live on a cylinder generated by the normal bundle of K and have intensity measures expressed in terms of the support measures of K. We apply our limit results to a spatial version of the classical change-point problem, in which random point patterns are considered which have different distributions inside and outside a fixed, but unknown convex body K.

Keywords Poisson point process · Convex body · Empirical process · Support measure · Normal cylinder · Change-set problem · Limit process

1 Introduction

Let Ψ_n be a Poisson point process in finite-dimensional Euclidean space $\mathbb{R}^d$ which has two homogeneous components; there is a compact set $K \subset \mathbb{R}^d$ such that the intensity of Ψ_n inside K is nc_- while it is nc_+ outside K, $c_+, c_- \geq 0$.

E. Khmaladze (✉)
School of Mathematics, Statistics and Computer Science,
Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand
e-mail: Estate.Khmaladze@mcs.vuw.ac.nz

W. Weil
Institut für Algebra und Geometrie, Universität Karlsruhe (TH),
Englerstr. 2, 76128 Karlsruhe, Germany
e-mail: weil@math.uka.de

We are interested in the limit behaviour of the processes Ψ_n, as $n \to \infty$, in the neighbourhood of the boundary ∂K of K. To be more precise, we let

$$(\partial K)_\varepsilon = \left\{ z \in \mathbb{R}^d : \min_{x \in \partial K} \|z - x\| \le \varepsilon \right\}$$

be the ε-neighbourhood of ∂K and consider the restriction of Ψ_n to $(\partial K)_\varepsilon$. We call this a *local Poisson process* (in the neighbourhood of K). Assume now that the neighborhood shrinks, as $n \to \infty$, in such a way that $n\varepsilon \to const$. Since the intensity measures of the local Poisson processes stay bounded, after a suitable rescaling, we can expect that the rescaled local processes converge to a limit which should be a Poisson process as well. The main questions are: Where does this limit process live and what is its intensity measure?

In the following, we consider this situation for Poisson processes Ψ_n with rather general intensity measures and we study the corresponding local processes in more detail. We concentrate on the case where K is a *convex body* (a compact convex set with interior points). For the description of points in the neighbourhood of ∂K, the *generalized normal bundle* $\mathrm{Nor}(K)$ of K is then a natural choice. It consists of pairs (x, u), where $x \in \partial K$ and u is an outer normal vector to K at x. As we shall show, the natural space where the limit processes of Ψ_n live is the *normal cylinder* $\mathbb{R} \times \mathrm{Nor}(K)$. A basic tool, both for the existence of limit processes as well as for the form of their intensity measures, is the generalized *Steiner formula* which yields a decomposition of the (d-dimensional) Lebesgue measure μ_d on $\mathbb{R}^d$ in terms of the *support measures* $\Theta_j(K; \cdot)$ of K, $j = 0, \ldots, d-1$. The latter are finite measures on $\mathrm{Nor}(K)$ (see Schneider, 1993, for these and other notions from convex geometry which we shall use).

After collecting the basic geometric notions and results in Sect. 2, we show in Sects. 3 and 4 various limit theorems for local processes in total variation. In these results, the intensity measures of the limit processes Ψ are generated by the support measure $\Theta_{d-1}(K; \cdot)$ of K. In the final Sect. 5 we extract certain higher order components of Ψ_n, which are asymptotically driven by the support measures $\Theta_j(K; \cdot), j = 0, \ldots, d-2$, and determine the rate of their standard deviation.

The concentration on convex bodies K has a long tradition in statistics (see, for example, the paper by Ripley and Rasson, 1977). However, in the present context it may seem to be restrictive. In fact, a recent extension of the Steiner formula to arbitrary closed sets $K \subset \mathbb{R}^d$ (see Hug et al., 2004) would allow to transfer some of the results to this more general setting. However, for the limit results in Sect. 5 we need the non-negativity of $\Theta_j(K; \cdot)$ as well as the fact that the Steiner formula yields a polynomial expansion. For closed sets K, the support measures $\Theta_j(K; \cdot)$, $j \in \{0, \ldots, d-2\}$, are signed measures and the expansion given by the Steiner formula is not of polynomial type, in general. In fact, recent results in Heveling et al. (2004) and Hug et al. (2006) show that the convexity of K is a natural condition, if the polynomial behaviour is important.

The motivation for this work arose from the statistical aspects of local empirical processes. For random points in $\mathbb{R}$, the local empirical processes in the neighbourhood of a given point (or in the neighbourhood of ∞) form a very old and important object of statistical theory. The theory of extremal processes and exceedences, estimation of end-points, behaviour of moduli of continuity of empirical and related processes are all strongly connected with the theory of local empirical processes (we refer to Cooke, 1979; Dekkers et al., 1989; Resnick, 1986; Stute, 1982, among many other papers). Higher dimensional analogues, i.e. local empirical processes for random points in $\mathbb{R}^d$, but again in the neighbourhood of a given point (see, for example, the papers by Deheuvels and Mason, 1994, 1995; Einmahl, 1997; and Khmaladze, 1998) or in the neighbourhood of ∞ (see, for example, Einmahl et al. (2001)), have been studied relatively recently. Local empirical processes in the neighbourhood of a set K seem to be new objects and were not studied previously, to the best of our knowledge. However, as far as a set is an immensely richer object than a point, the investigation of local point processes in the neighbourhood of a set must be much more diverse and fruitful. In this paper, we hope to make a first step towards it.

Local empirical processes play also a prominent role in the *change-set problem*, a natural generalization of the classical change-point problem. Consider the two-level process Ψ_n, introduced at the beginning of this Introduction, and assume that the subset K, the change-set, is unknown to the statistician. If K and K' are two possible change-sets, the logarithm of likelihood ratio of the process Ψ_n under K and K' can be easily seen to be

$$\begin{aligned} L_n(K';K) = \ln \frac{c_-}{c_+}\,&[\Psi_n(K'\backslash K) - \Psi_n(K\backslash K')] \\ &-n(c_- - c_+)[\mu_d(K'\backslash K) - \mu_d(K\backslash K')]\,, \end{aligned} \tag{1}$$

where μ_d denotes the Lebesgue measure in $\mathbb{R}^d$. The natural way of looking at this likelihood ratio is to consider the change-set as a parameter of interest and to study (1) as a process indexed by K'.

For both, parametric and semi-parametric situations, when the parameter of interest is a vector or a function (respectively), the asymptotic theory of statistical inference is essentially connected with the local behaviour of the likelihood ratio. It is equaly desirable to develop a similar "local" theory for the likelihood ratio when the parameter is a set. The stochastic part of the likelihood (1) is $\Psi_n(K'\backslash K) - \Psi_n(K\backslash K')$ which, as soon as the symmetric difference $K'\triangle K$ satisfies $K'\triangle K \subset (\partial K)_\varepsilon$, is simply another version of the local Poisson process connected with Ψ_n. Therefore, the local analysis of (1) requires the analysis of local point processes. As an application of our limit theorems, we will use a result of Khmaladze et al. (2006), for a class of convex change-sets, to describe in Corollary 3 the limit distribution of the random variable $n\mu_d(\hat{K}_n \Delta K)$, where $\hat{K}_n$ is the maximum likelihood estimator.

2 Basic definitions from geometry

Let $K \subset \mathbb{R}^d$ be a *convex body*, a compact convex set in d-dimensional Euclidean space $\mathbb{R}^d$ with interior points. For $\varepsilon > 0$, let

$$K_\varepsilon = \{z \in \mathbb{R}^d \setminus K : \|z - p(z)\| \le \varepsilon\} \cup K = K + \varepsilon B^d$$

be the *outer parallel body* of K. Here, $B^d = B(0,1)$ denotes the unit ball in $\mathbb{R}^d$ and $p(z) = p(\partial K, z)$ is the *metric projection* of z onto the boundary ∂K of K, that is, $p(z)$ is the nearest point to z from ∂K. For $z \in \mathbb{R}^d \setminus K$ this point $p(z)$ is unique. For almost all points $z \in K$ the metric projection onto ∂K is also unique. Namely, if we consider the set S_K of points in K which have more than one metric projection onto ∂K, then $\mu_d(S_K) = 0$ (see, for example, Hug et al., 2004). S_K is called the *(inner) skeleton* of K. In order to use the function p on the whole space $\mathbb{R}^d$, we define $p(z)$, for $z \in K$, as the projection point which is smallest in the lexicographic order.

For $\varepsilon \ge 0$, the set

$$K_{-\varepsilon} = \{z \in K : z + \varepsilon B^d \subset K\}$$

is the *inner parallel body* of K (note that $K_0 = K_{-0} = K$). The set $K_{-\varepsilon}$ can also be represented as $K_{-\varepsilon} = \{z \in K : \|z - p(z)\| \ge \varepsilon\}$. Note that the set $(\partial K)_\varepsilon$ of the Introduction is just $(\partial K)_\varepsilon = K_\varepsilon \setminus K_{-\varepsilon}$.

The largest value r^* of ε such that $K_{-\varepsilon}$ is non-empty is the *inradius* of K. Note that $(K_{-\varepsilon})_\varepsilon \subseteq K$, for all $\varepsilon \ge 0$. The largest value $r = r(K) \ge 0$, for which $(K_{-r})_r = K$, that is

$$K_{-r} + rB^d = K,$$

is called the *interior reach* of K. If $r(K) > 0$, we say that K is of positive interior reach. If K is of interior reach r, then $(K_{-\varepsilon})_\varepsilon = K$, for all $0 \le \varepsilon \le r$. The set K is of positive interior reach r if ∂K is of positive reach r in the usual sense (see Federer, 1959). It is possible that $r(K) = 0$, for example if K is a polytope.

Now we introduce a local reach function for general convex bodies K. For $x \in \partial K$, we define the *local* (*interior*) *reach* $r(x) = r(K,x)$ as the largest $r' \ge 0$ such that x is in the boundary of a ball $B(y,r')$ (with center y and radius r') and $B(y,r') \subset K$ (here $r(x) = 0$ means that there is no such ball). We remark that the inradius of K fulfills $r^* = \max_{x \in \partial K} r(x)$ whereas $r(K) = \min_{x \in \partial K} r(x)$ is, again, the interior reach. For each $z \in K \setminus (\partial K \cup S_K)$, we have $\|z - p(z)\| < r(p(z))$.

Let S^{d-1} be the unit sphere in $\mathbb{R}^d$. For a point $x \in \partial K$, a unit vector $u \in S^{d-1}$ is an *outer normal* of K at x, if there is some $z \in \mathbb{R}^d \setminus K$, such that $x = p(z)$ and $u = (z - p(z))/\|z - p(z)\|$. The (generalized) *normal bundle* $\mathrm{Nor}(K)$ of K is then defined as

$$\mathrm{Nor}(K) = \{(x,u) : x \in \partial K, u \text{ is an outer normal of } K \text{ at } x\}.$$

Clearly, $\mathrm{Nor}(K) \subset \partial K \times S^{d-1}$. The mapping $z \mapsto (p(z), u(z))$ with $u(z) = (z - p(z))/\|z - p(z)\|$ maps $\mathbb{R}^d \setminus K$ onto the normal bundle $\mathrm{Nor}(K)$. In general, the outer normal u at a point $x \in \partial K$ is not unique. In case of uniqueness, x is called a *regular boundary point* of K and the outer normal is denoted by $u(x)$. It is well known that almost every boundary point (with respect to the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$ on ∂K) is regular. If x has positive local interior reach $r(x) > 0$, then x is regular and we define $u(x) = u(z) = (p(z)-z)/\|z-p(z)\|$, for any $z \in K\backslash(\partial K \cup S_K)$ such that $x = p(z)$. In this way, the mapping $z \mapsto (p(z), u(z))$ extends to $K\backslash(\partial K \cup S_K)$. The unit vector $-u(z) = (z-p(z))/\|z-p(z)\|$ may be called an *inner normal*. We remark that this also provides us with an alternative representation of the inner skeleton S_K, namely as $S_K = \{x - r(x)u(x) : x \in \partial K, r(x) > 0\}$. Although there may be regular boundary points x with $r(x) = 0$, we have

$$\mathcal{H}^{d-1}(\{x \in \partial K : x \text{ regular}\}) = \mathcal{H}^{d-1}(\{x \in \partial K : r(x) > 0\}) = \mathcal{H}^{d-1}(\partial K)$$

(see Schneider, 1993). Also, we denote by $d(z) = d(\partial K, z) = \|z-p(z)\|$, $z \in \mathbb{R}^d$, the distance of z from ∂K.

We use these notions now to introduce local outer and inner parallel sets. For a Borel set $A \subset \partial K \times S^{d-1}$ and $\varepsilon > 0$, let

$$A_\varepsilon = \{z \in K_\varepsilon \backslash K : (p(z), u(z)) \in A\}$$

and

$$A_{-\varepsilon} = \{z \in K\backslash(K_{-\varepsilon} \cup S_K) : (p(z), u(z)) \in A\}.$$

For completeness, we also define A_0 as the projection of A onto the first coordinate. For small $\varepsilon > 0$, we interpret the sets A_ε and $A_{-\varepsilon}$ as small neighbourhoods of the subset $A \subset \partial K \times S^{d-1}$ (see an example of A_ε in Fig. 1). In the next section, we will consider the behaviour of a point process on these and various other small shells around a set A and therefore we need to clarify the behaviour of the Lebesgue measure on the sets A_ε and $A_{-\varepsilon}$.

It is well known and easy to see that, for a Borel set $A \subset \partial K \times S^{d-1}$, the sets A_ε and $A_{-\varepsilon}$ are Borel sets in $\mathbb{R}^d$ and that $A \mapsto \mu_d(A_\varepsilon)$ and $A \mapsto \mu_d(A_{-\varepsilon})$ define measures on $\partial K \times S^{d-1}$. A local version of the classical Steiner formula (see Schneider, 1993) shows that $\mu_d(A_\varepsilon)$ is a polynomial in $\varepsilon \geq 0$ of degree at most d,

$$\mu_d(A_\varepsilon) = \frac{1}{d}\sum_{j=1}^{d}\binom{d}{j}\varepsilon^j \Theta_{d-j}(K; A), \tag{2}$$

where the coefficients $\Theta_i(K; A)$ (abbreviated in the next sections to $\Theta_i(A)$) define finite Borel measures $\Theta_0(K;\cdot), \ldots, \Theta_{d-1}(K;\cdot)$ on $\partial K \times S^{d-1}$ (and concentrated on $\mathrm{Nor}(K)$). They are called the *support measures* of K and carry most of the geometric information about K. In particular, the projection of $\Theta_{d-1}(K;\cdot)$ onto ∂K is the Hausdorff measure $\mathcal{H}^{d-1}$ on ∂K. We refer to Schneider (1993), for further properties of support measures and their relations to curvature measures and surface area measures of convex bodies. Extensions of these measures to more general set classes (like finite unions of convex sets and others) have been studied by various authors (Hug, 1999; Hug and Last, 2000;

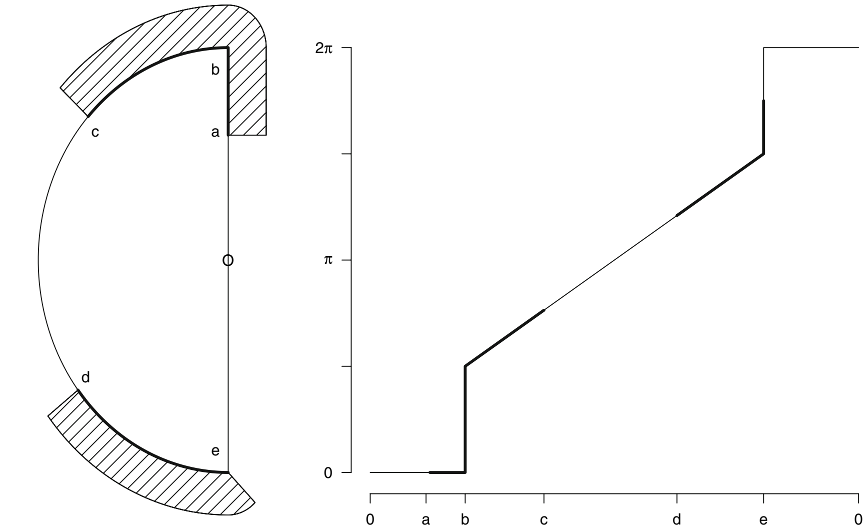

Fig. 1 The set K, on the *left*, is a semicircle in $\mathbb{R}^2$. The *shaded set* is A_ε. *The bold part* of the boundary is the projection A_0 of A onto ∂K. *The polygonal line in the right figure* describes the normal bundle $\mathrm{Nor}(K)$ and *the bold parts* represent the subset A. Here, ∂K has been identified with the interval $[0, 0]$ on the x-axis and the unit sphere S^1 with the interval $[0, 2\pi]$ on the y-axis

Rataj and Zähle, 2001; Schneider, 1993; Zähle, 1986), a quite general result was recently obtained in Hug et al. (2004).

For a body K with interior reach r, one can show that, for $0 \le \varepsilon \le r$, $\mu_d(A_{-\varepsilon})$ has the same polynomial devselopment (2) with ε^j replaced by $(-1)^{j-1}\varepsilon^j$. For arbitrary convex bodies K, (2) generalizes in the form of an integral involving the local interior reach function. The following theorem gives the most general version of such a Steiner formula for convex bodies. It has been proved (in more general form, for closed sets) in Hug et al. (2004), a special case was obtained earlier by Sangwine-Yager (1994).

Theorem 1 *For a convex body K and a μ_d-integrable real function f on $\mathbb{R}^d$, we have*

$$\int_{\mathbb{R}^d} f(z)\,\mu_d(\mathrm{d}z) = \sum_{j=1}^{d} \binom{d-1}{j-1} \int_{\mathrm{Nor}(K)} \int_{-r(x)}^{\infty} t^{j-1} f(x+tu)\,\mathrm{d}t\,\Theta_{d-j}(K;\mathrm{d}(x,u)).$$

In particular, we get

$$\begin{aligned} \mu_d(A_{-\varepsilon}) &= \sum_{j=1}^{d} \binom{d-1}{j-1} \int_A \int_{-\min(r(x),\varepsilon)}^{0} t^{j-1}\mathrm{d}t\,\Theta_{d-j}(K;\mathrm{d}(x,u)) \\ &= \frac{1}{d}\sum_{j=1}^{d} \binom{d}{j} (-1)^{j-1} \int_A (\min(r(x),\varepsilon))^j\,\Theta_{d-j}(K;\mathrm{d}(x,u)). \qquad (3) \end{aligned}$$

The support measure $\Theta_{d-1}(K;\cdot)$ is concentrated on the pairs (x,u), where x is a regular boundary point of K (see Schneider, 1993, Theorem 2.2.4). As we already mentioned, $\mathcal{H}^{d-1}$-almost all points $x \in \partial K$ are regular, and thus the mapping $N : \partial K \to \mathrm{Nor}(K), x \mapsto (x, u(x))$, is defined almost everywhere and is measurable. As it turns out, $\Theta_{d-1}(K;\cdot)$ is the image measure of the *curvature measure* $C_{d-1}(K;\cdot)$ (which equals the restriction of $\mathcal{H}^{d-1}$ to ∂K) under N (conversely, $C_{d-1}(K;\cdot)$ is the projection of $\Theta_{d-1}(K;\cdot)$ onto the first coordinate). More generally, the curvature measure $C_{d-j}(K;\cdot)$ is the projection of $\Theta_{d-j}(K;\cdot)$ onto the first coordinate, $j = 1,\ldots,d$. Hence, if $r(x) > 0$, then x is a regular boundary point and therefore

$$\int_{\mathrm{Nor}(K)} \int_{-r(x)}^{0} t^{j-1} f(x+tu)\,\mathrm{d}t\,\Theta_{d-j}(K;\mathrm{d}(x,u))$$
$$= \int_{\partial K} \int_{-r(x)}^{0} t^{j-1} f(x+tu(x)) \quad \mathrm{d}t\, C^{\mathsf{c}}_{d-j}(K;\mathrm{d}x), \quad j = 1,\ldots,d,$$

where $C^{\mathsf{c}}_{d-j}(K;\cdot)$ denotes the absolutely continuous part of $C_{d-j}(K;\cdot)$ with respect to $C_{d-1}(K;\cdot)$.

There is a useful relationship between the support measures of K and of K_t, $t \in \mathbb{R}$. In order to describe this relationship, we consider the homeomorphism $T_t : \mathbb{R}^d \times S^{d-1} \to \mathbb{R}^d \times S^{d-1}$, defined by $T_t(x,u) = (x+tu, u)$, and let T_tA denote the image of a Borel set $A \subset \mathbb{R}^d \times S^{d-1}$ (T_tA is again a Borel set). For $t \geq -r(K)$, T_t maps $\mathrm{Nor}(K)$ one-to-one onto $\mathrm{Nor}(K_t)$.

Lemma 1 *For a Borel set $A \subset \mathrm{Nor}(K)$, $t \geq -r(K)$ and $j = 1,\ldots,d$, we have*

$$\Theta_{d-j}(K_t;T_tA) = \sum_{i=j}^{d} \binom{d-j}{i-j} t^{i-j}\,\Theta_{d-i}(K;A) \tag{4}$$

and

$$\Theta_{d-j}(K;A) = \sum_{i=j}^{d} \binom{d-j}{i-j} (-t)^{i-j}\,\Theta_{d-i}(K_t;T_tA). \tag{5}$$

Proof For $t \geq 0$, (4) is a simple consequence of (2) (see Theorem 4.2.2 in Schneider, 1993) and (5) follows from (4) by inversion.

It remains to consider the case where K has positive interior reach $r(K)$ and $-r(K) \leq t < 0$. Then, we have $(K_t)_{-t} = K$ and $T_{-t}(T_tA) = A$. Therefore, the Steiner formula (2) applied to K_t and T_tA yields

$$\mu_d(A_t) = \frac{1}{d} \sum_{j=1}^{d} \binom{d}{j} (-1)^{\,j-1} t^j\,\Theta_{d-j}(K_t;T_tA).$$

We consider A_δ with $\delta > 0$. According to (2), $\mu_d(A_\delta)$ has an expansion in terms of $\Theta_{d-j}(K;\cdot)$, but also in terms of $\Theta_{d-j}(K_t;\cdot), j = 1,\dots,d$. Namely,

$$\mu_d(A_\delta) = \frac{1}{d}\sum_{j=1}^{d}\binom{d}{j}\delta^j\Theta_{d-j}(K;A)$$
$$= \frac{1}{d}\sum_{j=1}^{d}\binom{d}{j}[(-t+\delta)^j - (-t)^j]\Theta_{d-j}(K_t;T_tA).$$

Comparing coefficients of δ^j yields (5), and (4) follows by inversion. □

For general $t \le 0$, (4) extends in form of an integral relation,

$$\Theta_{d-j}(K_t;T_tA) = \sum_{i=j}^{d}\binom{d-j}{i-j}t^{i-j}\int_A \mathbf{1}_{\{y:t\ge -r(y)\}}(x)\Theta_{d-i}(K;\mathrm{d}(x,u)),$$

as follows from Corollary 4.4 in Hug et al. (2004). Of course, both sides vanish for $t < -r^*$.

For $j = 1$ and $t \ge 0$, (4) yields

$$\Theta_{d-1}(K_t;T_tA) = \sum_{i=1}^{d}\binom{d-1}{i-1}t^{i-1}\Theta_{d-i}(K;A). \tag{6}$$

Choosing d different values $t_0,\dots,t_{d-1}$, we can invert this system of linear equations and obtain $\Theta_{d-i}(K;A)$, for $i = 1,\dots,d$, as a linear combination of $\Theta_{d-1}(K_{t_0};T_{t_0}A),\dots,\Theta_{d-1}(K_{t_{d-1}};T_{t_{d-1}}A)$. In view of our later applications, we will give now explicit expressions for such inversion formulas for equally spaced $t_0,\dots,t_{d-1}$.

For a function $t \mapsto \varphi(t)$, the operator $\Delta_s = \Delta_s^1$ denotes the (first) forward difference with step s,

$$\Delta_s\varphi(t) = \varphi(t+s) - \varphi(t),$$

and $\Delta_s^k = \Delta_s(\Delta_s^{k-1}), k = 2,3,\dots$ (as well as $\Delta_s^0\varphi(t) = \varphi(t)$). Let S_n^m denote the Stirling numbers of the first kind. Recall that they are defined as the coefficients of the representation

$$x^{(n)} = x(x-1)\cdots(x-n+1) = \sum_{m=0}^{n} S_n^m x^m$$

(see, for example, Hildebrand, 1986, Sect. 1.2 and p. 116).

Lemma 2 *For a Borel set $A \subset \mathrm{Nor}(K)$, $s > 0$ and $i = 1, \ldots, d$, we have*

$$\binom{d-1}{i-1}\Theta_{d-i}(K;A) = \frac{1}{s^{i-1}} \sum_{k=i-1}^{d-1} \frac{1}{k!} S_k^{i-1} \Delta_s^k \Theta_{d-1}(K_r; T_r A)|_{r=0} \tag{7}$$

and

$$\frac{1}{d}\binom{d}{i}\Theta_{d-i}(K;A) = \frac{1}{s^{i-1}} \sum_{k=i}^{d} \frac{1}{k!} S_k^i \Delta_s^k \mu_d(A_r)|_{r=0}. \tag{8}$$

Proof From (6), we see that $\Theta_{d-1}(K_t; T_t A)$ is a polynomial in t, for $t \geq 0$. According to Newton's formula,

$$\Theta_{d-1}(K_t; T_t A) = \sum_{k=0}^{d-1} \frac{1}{k!} \left(\frac{t}{s}\right)^{(k)} \Delta_s^k \Theta_{d-1}(K_r; T_r A)|_{r=0}.$$

Let us replace the factorial moments by the sums using Stirling numbers,

$$\left(\frac{t}{s}\right)^{(k)} = \sum_{i=0}^{k} S_k^i \left(\frac{t}{s}\right)^i,$$

and change the order of summation. This leads to

$$\Theta_{d-1}(K_t; T_t A) = \sum_{i=0}^{d-1} \left(\frac{t}{s}\right)^i \sum_{k=i}^{d-1} \frac{1}{k!} S_k^i \Delta_s^k \Theta_{d-1}(K_r; T_r A)|_{r=0}. \tag{9}$$

Equating coefficients of t^i in (6) and (9), we obtain (7). Similarly, the Steiner formula for $\mu_d(A_t)$ gives the coefficients of t^i explicitly while Newton's formula for it uses coefficients of factorial moments $t^{(k-1)}$. Expressing the latter through t^i and Stirling numbers leads to (8). □

We remark that (7) and (8) can also be written in the form

$$\Theta_{d-i}(K;A) = \frac{1}{s^{i-1}} \sum_{m=0}^{d-1} a_{mi} \Theta_{d-1}(K_{ms}; T_{ms} A)\,, \quad i = 1, \ldots, d, \tag{10}$$

and

$$\Theta_{d-i}(K;A) = \frac{1}{s^{i-1}} \sum_{m=0}^{d-1} \tilde{a}_{mi} \mu_d(A_{ms})\,, \quad i = 1, \ldots, d,$$

with certain coefficients $a_{mi}, \tilde{a}_{mi}$, which can be given explicitly e.g.

$$a_{mi} = \sum_{k=\max(i,m)}^{d-1} (-1)^{k-m} \binom{k}{m} \frac{S_k^i}{k!} .$$

In this paper, we will mostly be interested in point processes concentrated on the neighbourhood $K_{\varepsilon T} \setminus K_{-\varepsilon T}, T > 0$, of ∂K and their limit behaviour as the neighbourhood shrinks, with $\varepsilon \to 0$. To describe the space where the limiting process will live we need to rescale the width of $K_{\varepsilon T} \setminus K_{-\varepsilon T}$ by ε^{-1}. This stretching will actually add an extra coordinate to $\mathrm{Nor}(K)$ and we will obtain the cylinder $\Sigma_T = [-T, T] \times \mathrm{Nor}(K)$. Since we intend to let T also be variable, we work with the infinite cylinder $\Sigma = \mathbb{R} \times \mathrm{Nor}(K)$ and let $\Sigma_+ = [0, \infty) \times \mathrm{Nor}(K)$ and $\Sigma_- = (-\infty, 0] \times \mathrm{Nor}(K)$. For $\varepsilon > 0$, we consider the map $\tau_\varepsilon : \mathbb{R}^d \setminus S_K \to \Sigma$, defined as

$$\tau_\varepsilon z = \left(\frac{d(z)}{\varepsilon}, p(z), u(z) \right) \quad \text{if } z \in \mathbb{R}^d \setminus K,$$

and as

$$\tau_\varepsilon z = \left(-\frac{d(z)}{\varepsilon}, p(z), u(z) \right) \quad \text{if } z \in K \setminus S_K.$$

τ_ε is injective and maps the outside of K onto Σ_+. On the other side, the image $\tau_\varepsilon(K \setminus S_K) \subset \Sigma_-$ is bounded and in general not of rectangular form. Figure 2 shows the effect of τ_ε on inside and outside deformations of ∂K in a simple situation (where only regular boundary points are involved).

The mapping τ_ε is measurable. On $\mathbb{R}^d \setminus K$, measurability follows simply from continuity; on $K \setminus S_K$, it can be seen as follows. The set $\mathrm{reg}(K)$ of regular boundary points is a Borel set (see Schneider, 1993, or Hug et al., 2004). The metric

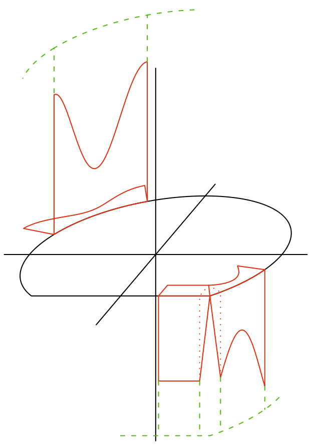

Fig. 2 Here K is a "chopped" circle in $\mathbb{R}^2$. The picture shows small inside and outside deformations of ∂K along with their images under τ_ε. Since only regular boundary points are involved, the projections on the cylinder $\mathbb{R} \times \partial K$ are shown

projection $p(K,\cdot)$ is continuous on $\mathbb{R}^d \setminus K$. The interior reach function $r(\cdot)$ is measurable (see, for example, Sangwine-Yager, 1994). Hence

$$K^* = \{z \in (\mathbb{R}^d \setminus K) \cup \partial K : p(z) \in \mathrm{reg}(K), 0 \leq d(z) < r(p(z))\}$$

is measurable (K^* is the interior of K, without the inner skeleton S_K, folded to the outside). The mapping $\varphi : z \mapsto 2p(z) - z$ is also continuous on $\mathbb{R}^d \setminus K$ and one-to-one from K^* to $K \setminus S_K$. Since τ_ε is continuous on $\mathbb{R}^d \setminus K$, it is measurable on K^*. On $K \setminus S_K$, τ_ε is the composition of φ^{-1}, τ_ε (restricted to K^*) and the reflection $(t,x,u) \mapsto (-t,x,u)$, which proves measurability of τ_ε.

We now introduce the measure $\mu_{(\varepsilon)}$ on Σ as the image of the Lebesgue measure μ_d under τ_ε. For rectangular Borel sets $[t_1,t_2] \times A$ in Σ, with $t_2 \geq t_1 \geq -r(K)$, we thus get

$$\begin{aligned}\mu_{(\varepsilon)}([t_1,t_2] \times A) &= \mu_d(\{z \in K_{\varepsilon t_2} \setminus K_{\varepsilon t_1} : (p(z),u(z)) \in A\}) \\ &= \frac{1}{d}\sum_{j=1}^{d}\binom{d}{j}[(\varepsilon t_2)^j - (\varepsilon t_1)^j]\Theta_{d-j}(K;A). \end{aligned} \tag{11}$$

In general, $\mu_{(\varepsilon)}$ is supported by a set $\Sigma(\varepsilon) \subset \Sigma$ which depends on ε and is no longer of rectangular form,

$$\begin{aligned}\Sigma(\varepsilon) &= \Sigma_+ \cup \tau_\varepsilon(K \setminus S_K) \\ &= \Sigma_+ \cup \left\{\left(-\frac{d(z)}{\varepsilon}, p(z), u(z)\right) : z \in K \setminus S_K\right\}.\end{aligned}$$

The analogue of (11), for $t_2 = 0$ and arbitrary $t_1 = -t < 0$, can be obtained directly from (3) and we get

$$\mu_{(\varepsilon)}([-t,0] \times A) = \frac{1}{d}\sum_{j=1}^{d}\binom{d}{j}(-1)^{j-1}\int_A (\min\{t\varepsilon, r(x)\})^j\, \Theta^c_{d-j}(K; \mathrm{d}(x,u)).$$

In the rest of the paper we will shorten the notation for support measures from $\Theta_{d-j}(K;\cdot)$ to $\Theta_{d-j}(\cdot)$.

3 Poisson processes near the boundary

Suppose Ψ_n is a Poisson process on $\mathbb{R}^d$ with intensity measure Λ_n. We consider Ψ_n in a neighbourhood of ∂K which shrinks to ∂K as $\varepsilon \to 0$, and study the limit behaviour of this point process as $n \to \infty$ and $\varepsilon \to 0$ simultaneously. In this section, we prove first order limit theorems in the total variation norm. In the final section, we will be interested in components of Ψ_n which are asymptotically driven by each of the support measures of K.

Although most of the following results can be proved under suitable conditions on the decay of the singular part (with respect to μ_d) of Λ_n near the boundary of K, we want to simplify the presentation and assume that Λ_n is absolutely continuous in the following. Hence

$$\Lambda_n(\cdot) = \int_{(\cdot)} n f_n \mathrm{d}\mu_d, \tag{12}$$

with density nf_n, where f_n is integrable on every bounded Borel subset in $\mathbb{R}^d$. We additionally assume asymptotic L^1-convergence of f_n near ∂K, in the sense that there are measurable functions $f_+, f_- \geq 0$ on $\mathrm{Nor}(K)$ such that

$$\begin{aligned} \frac{1}{\varepsilon}\int_{K_{\varepsilon T}\setminus K} |f_n(z) - f_+(p(z), u(z))|\, \mu_d(\mathrm{d}z) \to 0, \\ \frac{1}{\varepsilon}\int_{K\setminus K_{-\varepsilon T}} |f_n(z) - f_-(p(z), u(z))|\, \mu_d(\mathrm{d}z) \to 0. \end{aligned} \tag{13}$$

for each $T > 0$, as $n \to \infty, \varepsilon \to 0$, with $n\varepsilon \to 1$. As a simple example of such a sequence $\{\Psi_n\}, n \in \mathbb{N}$, we may consider the process with intensity measure

$$\mathbb{E}\Psi_n(C) = nc_+\mu_d(C\setminus K) + nc_-\mu_d(C \cap K)\,, \tag{14}$$

where $c_+, c_- \geq 0$ and $C \subset \mathbb{R}^d$ varies through the Borel sets.

From the above assumptions and the integrability of f_n it follows that

$$\int_{K_{\varepsilon T}\setminus K} f_+(p(z), u(z))\, \mu_d(\mathrm{d}z) \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Applying Theorem 1, we see that

$$\begin{aligned} &\int_{K_{\varepsilon T}\setminus K} f_+(p(z), u(z))\, \mu_d(\mathrm{d}z) \\ &\quad = \frac{1}{d}\sum_{j=1}^{d}\binom{d}{j}(\varepsilon T)^j \int_{\mathrm{Nor}(K)} f_+(x,u)\, \Theta_{d-j}(\mathrm{d}(x,u)), \end{aligned} \tag{15}$$

and therefore all integrals on the right-hand side of (15) are finite. The same conclusion can be obtained for the function f_-.

Now let $\Psi_{n,\varepsilon}$ be the image of Ψ_n under τ_ε and let $\Lambda_{n,\varepsilon}$ be its intensity measure. $\Psi_{n,\varepsilon}$ is a Poisson process on Σ. For $t > 0$ and a Borel set $A \subset \mathrm{Nor}(K)$, we have

$$\Psi_{n,\varepsilon}([0,t] \times A) = \Psi_n(A_{t\varepsilon})$$

and

$$\Psi_{n,\varepsilon}([-t,0] \times A) = \Psi_n(A_{-t\varepsilon}).$$

In order to prove the convergence $\Phi_n \to \Phi$ in total variation, for Poisson processes Φ_n, Φ on Σ with finite intensity measures, it is sufficient that the intensity measures converge in total variation (see, for example, Reiss, 1993, (3.8)). However, since $\Lambda_{n,\varepsilon}(\Sigma)$ may be infinite, we will have to work with local variants of $\Psi_{n,\varepsilon}$. Therefore, we say that $\Psi_{n,\varepsilon}$ converges in *locally total variation* (LTV) to a Poisson process Ψ on Σ, if the distribution of $\Psi_{n,\varepsilon}(\cdot \cap \Sigma_T)$ converges in total variation to the distribution of $\Psi(\cdot \cap \Sigma_T)$, for all $T > 0$.

Theorem 2 *Let K be a convex body and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes fulfilling (12) and (13). If $n\varepsilon \to 1$, as $n \to \infty, \varepsilon \to 0$, then $\Psi_{n,\varepsilon}$ converges in LTV to a Poisson process Ψ on Σ with intensity measure $\Lambda = \mathbb{E}\Psi$ given by*

$$
\begin{aligned}
\Lambda(B) = \int_{\mathrm{Nor}(K)} \int_{-\infty}^{\infty} \big(&\mathbf{1}_{B\cap\Sigma_+}(t,x,u) f_+(x,u) \\
&+\mathbf{1}_{B\cap\Sigma_-}(t,x,u) f_-(x,u)\big)\, \mathrm{d}t\, \Theta_{d-1}(\mathrm{d}(x,u)), \quad B \subset \Sigma.
\end{aligned}
$$

Proof Let $T > 0$ be given. It is sufficient to show that $\Lambda_{n,\varepsilon}(\cdot \cap \Sigma_T) \to \Lambda(\cdot \cap \Sigma_T)$, in total variation.

For a Borel set $B \subset \Sigma_T \cap \Sigma_+$, we combine (15) with

$$
\begin{aligned}
&\left| \int_{\tau_\varepsilon^{-1}(B)} n f_n(z)\, \mu_d(\mathrm{d}z) - \Lambda(B) \right| \\
&\quad \le \left| \int_{\tau_\varepsilon^{-1}(B)} n f_n(z)\, \mu_d(\mathrm{d}z) - \frac{1}{\varepsilon} \int_{\tau_\varepsilon^{-1}(B)} f_+(p(z), u(z))\, \mu_d(\mathrm{d}z) \right| \\
&\qquad + \left| \frac{1}{\varepsilon} \int_{\tau_\varepsilon^{-1}(B)} f_+(p(z), u(z))\, \mu_d(\mathrm{d}z) \right. \\
&\qquad \left. - \int_{\mathrm{Nor}(K)} \int_0^T \mathbf{1}_B(t,x,u) f_+(x,u)\, \mathrm{d}t\, \Theta_{d-1}(\mathrm{d}(x,u)) \right|.
\end{aligned}
$$

The first summand is asymptotically bounded by

$$
\begin{aligned}
&\left| \frac{1}{\varepsilon} \int_{\tau_\varepsilon^{-1}(B)} f_n(z)\, \mu_d(\mathrm{d}z) - \frac{1}{\varepsilon} \int_{\tau_\varepsilon^{-1}(B)} f_+(p(z), u(z))\, \mu_d(\mathrm{d}z) \right| \\
&\quad \le \frac{1}{\varepsilon} \int_{K_{\varepsilon T} \setminus K} |f_n(z) - f_+(p(z), u(z))|\, \mu_d(\mathrm{d}z),
\end{aligned}
$$

and therefore tends to 0, uniformly in B, by assumption (13).

For the second summand, we use Theorem 1 and get

$$
\begin{aligned}
&\int_{\tau_\varepsilon^{-1}(B)} f_+(p(z),u(z))\,\mu_d(\mathrm{d}z)\\
&\quad= \sum_{j=1}^{d}\binom{d-1}{j-1}\int_{\mathrm{Nor}(K)}\int_0^{\varepsilon T}\mathbf{1}_B(\varepsilon^{-1}t,x,u)t^{j-1}f_+(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\\
&\quad= \sum_{j=1}^{d}\binom{d-1}{j-1}\varepsilon^j\int_{\mathrm{Nor}(K)}\int_0^{T}\mathbf{1}_B(t,x,u)t^{j-1}f_+(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u)),
\end{aligned}
$$

and hence

$$
\begin{aligned}
&\left|\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(B)} f_+(p(z),u(z))\,\mu_d(\mathrm{d}z)\right.\\
&\quad\left.-\int_{\mathrm{Nor}(K)}\int_0^{T}\mathbf{1}_B(t,x,u)f_+(x,u)\,\mathrm{d}t\,\Theta_{d-1}(\mathrm{d}(x,u))\right|\\
&\qquad\le \sum_{j=2}^{d}\binom{d-1}{j-1}\varepsilon^{j-1}\int_{\mathrm{Nor}(K)}\int_0^{T}\mathbf{1}_B(t,x,u)t^{j-1}f_+(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\\
&\qquad\le \frac{1}{d}\sum_{j=2}^{d}\binom{d}{j}\varepsilon^{j-1}T^j\int_{\mathrm{Nor}(K)} f_+(x,u)\,\Theta_{d-j}(\mathrm{d}(x,u)).
\end{aligned}
$$

Since all integrals of f_+ are finite, the second summand also tends to 0, uniformly in B, as $\varepsilon \to 0$.

The case $B \subset \Sigma_T \cap \Sigma_-$ is treated analogously, with f_+ replaced by f_-. The only difference here is that Theorem 1 gives us

$$
\begin{aligned}
&\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(B)} f_-(p(z),u(z))\,\mu_d(\mathrm{d}z)\\
&\quad= \sum_{j=1}^{d}\binom{d-1}{j-1}\int_{\mathrm{Nor}(K)}\int_{-\min\{\varepsilon T,r(x)\}}^{0}\mathbf{1}_B(\varepsilon^{-1}t,x,u)t^{j-1}\\
&\qquad\times f_-(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\\
&\quad= \sum_{j=1}^{d}\binom{d-1}{j-1}\varepsilon^j\int_{\mathrm{Nor}(K)}\int_{-\min\{T,\varepsilon^{-1}r(x)\}}^{0}\mathbf{1}_B(t,x,u)t^{j-1}\\
&\qquad\times f_-(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u)),
\end{aligned}
$$

and hence

$$\begin{aligned}
&\left|\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(B)} f_-(p(z),u(z))\,\mu_d(\mathrm{d}z)\right.\\
&\qquad\left.-\int_{\mathrm{Nor}(K)}\int_{-T}^{0}\mathbf{1}_B(t,x,u)f_-(x,u)\,\mathrm{d}t\,\Theta_{d-1}(\mathrm{d}(x,u))\right|\\
&\qquad\le \int_{\mathrm{Nor}(K)}\int_{-T}^{-\min\{T,\varepsilon^{-1}r(x)\}}\mathbf{1}_B(t,x,u)f_-(x,u)\,\mathrm{d}t\,\Theta_{d-1}(\mathrm{d}(x,u))\\
&\qquad\quad+\left|\sum_{j=2}^{d}\binom{d-1}{j-1}\varepsilon^{j-1}\int_{\mathrm{Nor}(K)}\int_{-\min\{T,\varepsilon^{-1}r(x)\}}^{0}\mathbf{1}_B(t,x,u)t^{j-1}\right.\\
&\qquad\quad\left.\times f_-(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\right|.
\end{aligned}$$

The first summand tends to 0 by monotone convergence (uniformly in B). For the second sum we can use the same estimation as above,

$$\begin{aligned}
&\sum_{j=2}^{d}\binom{d-1}{j-1}\varepsilon^{j-1}\left|\int_{\mathrm{Nor}(K)}\int_{-\min\{T,\varepsilon^{-1}r(x)\}}^{0}\mathbf{1}_B(t,x,u)t^{j-1}f_-(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\right|\\
&\qquad\le\frac{1}{d}\sum_{j=2}^{d}\binom{d}{j}\varepsilon^{j-1}T^j\int_{\mathrm{Nor}(K)}f_-(x,u)\,\Theta_{d-j}(\mathrm{d}(x,u))
\end{aligned}$$

and, again, since all integrals on the right hand side are finite, the sum tends to 0 as $\varepsilon\to 0$. □

Corollary 1 *Let K be a convex body and $\{\Psi_n\}, n\in\mathbb{N}$, a sequence of Poisson processes with intensity measure (14). Define the measure ν on Borel sets $B\subset\Sigma$ as*

$$\nu(B)=\iint \mathbf{1}_B(t,x,u)\,\mathrm{d}t\,\Theta_{d-1}(\mathrm{d}(x,u)).$$

If $n\varepsilon\to 1$, as $n\to\infty,\varepsilon\to 0$, then $\Psi_{n,\varepsilon}$ converges in LTV to a Poisson process Ψ on Σ with intensity measure

$$\Lambda(B)=c_+\nu(B\cap\Sigma_+)+c_-\nu(B\cap\Sigma_-),\quad B\subset\Sigma.$$

Since Θ_{d-1} is supported by the regular boundary points $x\in\partial K$, the limit processes Ψ in Theorem 2 and Corollary 1 actually live on the cylinder $\mathbb{R}\times\mathrm{reg}(K)$ which we can think of as being embedded into Σ by $(t,x)\mapsto(t,x,u(x))$. The question may arise whether it would have been sufficient to consider corresponding limit theorems on the cylinder $\mathbb{R} \times \partial K$ and, consequently, start with functions f_+ and f_- depending only on x or on u (recall that the normal u is unique for $x \in \mathrm{reg}(K)$). The reason why this would not be enough and why we need to consider $\mathbb{R} \times \mathrm{Nor}(K)$ is that we want to describe the asymptotic behaviour of point processes near the whole boundary of the convex body K. This not only involves points $z \notin \partial K$ with $p(z) \in \mathrm{reg}(K)$ but also points z which project onto irregular boundary points. Although those points z will not contribute to the intensity measure of the limit process Ψ, they may very well affect the existence of Ψ and hence the convergence of Ψ_n has to be established on these parts of $\mathbb{R}^d \setminus \partial K$ as well. This is guaranteed by the integrability assumption (13).

Let $\tilde{\mathcal{B}}$ be a class of the Borel σ-algebra on Σ and consider the set-indexed stochastic processes

$$(\Psi_{n,\varepsilon}(B))_{B \in \tilde{\mathcal{B}}} \quad \text{and} \quad (\Psi(B))_{B \in \tilde{\mathcal{B}}}.$$

Then Theorem 2 shows that the distribution of $(\Psi_{n,\varepsilon}(B))_{B \in \tilde{\mathcal{B}}}$ converges in total variation to the distribution of $(\Psi(B))_{B \in \tilde{\mathcal{B}}}$, provided $\tilde{\mathcal{B}}$ consists of suitably bounded sets in Σ. However, if we consider the processes $(\Psi_{n,\varepsilon}(B))_{B \in \tilde{\mathcal{B}}_\varepsilon}$ on classes which now may change with ε, the convergence in total variation will not be true anymore, even if the classes $\tilde{\mathcal{B}}_\varepsilon$ converge to the limiting class $\tilde{\mathcal{B}}$ in quite a strong sense.

There is, however, a simple and practically convenient way to avoid this complication.

Corollary 2 *Let K be a convex body and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes fulfilling (12) and (13). For each $\varepsilon > 0$, let $\mathcal{B}_\varepsilon$ be a class of Borel sets in $\Sigma_T, T > 0$. Then, the total variation between the distributions of the processes*

$$(\Psi_{n,\varepsilon}(B))_{B \in \mathcal{B}_\varepsilon} \quad \textit{and} \quad (\Psi(B))_{B \in \mathcal{B}_\varepsilon}$$

converges to 0*, as* $n \to \infty, \varepsilon \to 0$ *with* $n\varepsilon \to 1$.

Corollary 2 allows us to state an analogue of an important theorem of Ibragimov and Has'minski (see Ibragimov and Has'minskii, 1981, Chap. I, Theorem 0.1). This theorem, loosely speaking, states the following. Let $\{\mathbb{P}_\theta^{(n)}, \theta \in \Theta\}$, $n \in \mathbb{N}$, be a sequence of families of distributions indexed by a finite-dimensional parameter θ (taking values in an open set Θ). If the logarithm of the likelihood ratio

$$L_n(\theta + \varepsilon_n u, \theta) = \ln \frac{\mathrm{d}\mathbb{P}^{(n)}_{\theta+\varepsilon_n u}}{\mathrm{d}\mathbb{P}^{(n)}_\theta},$$

as a process in u, with some normalization factor $\varepsilon_n \to 0$, converges in distribution to some process $\xi(u)$, then the ML estimator $\hat{\theta}$ fulfills

$$\frac{1}{\varepsilon_n}(\hat{\theta} - \theta) \xrightarrow{d} \hat{u}, \quad \hat{u} = \arg\max \xi(u).$$

Thus, we have a description of the limiting random variable for the normalized discrepancy $\hat{\theta} - \theta$ (although the distribution of $\hat{u}$ may not be easy to evaluate).

We consider the version of the change-set problem, as described in the Introduction (see (1)), in which case the parameter is a set K'. Let $\mathcal{C}$ be a class of convex bodies in $\mathbb{R}^d$, which we interpret as the class of all apriori possible change-sets, and, for given $K \in \mathcal{C}$ and all $\varepsilon > 0$, let $\mathcal{C}_\varepsilon = \mathcal{C}_\varepsilon(K)$ denote the subclass consisting of the sets $K' \in \mathcal{C}$ such that $\gamma^s(K', K) = \mu_d(K' \triangle K) \leq \varepsilon$. Since the symmetric difference metric γ^s and the Hausdorff metric generate the same topology on $\mathcal{C}$ (see, for example, Shephard and Webster, 1965), it can be seen that

$$\mathcal{C}_\varepsilon = \{C \in \mathcal{C} : K_{-c\varepsilon} \subset C \subset K_{c\varepsilon}\}$$

with some constant c (which may depend on K).

Let $\hat{K}_n$ be the ML estimator for the "true" change-set K,

$$\hat{K}_n = \arg\max_{K' \in \mathcal{C}} L_n(K'; K), \tag{16}$$

where L_n is given by (1). The rate of convergence of $\hat{K}_n$ to K is very sensitive to the "richness" of the class $\mathcal{C}$ and, more specifically, to its local "richness" at K. We will assume that the class $\mathcal{C}$ is a *locally simple VC-class* at K, that is, $\mathcal{C}$ is totally bounded (with respect to γ^s) and the covering number of $\mathcal{C}_{\varepsilon t}$ satisfies the inequality

$$N_{\delta\varepsilon}(\mathcal{C}_{\varepsilon t}) \leq c\left(\frac{t}{\delta}\right)^m \tag{17}$$

for some $m \in \mathbb{N}$ and for all $t \geq 0, \delta > 0$, and all sufficiently small $\varepsilon > 0$ and with some constant c independent of ε, δ and t (but maybe dependent on K and $\mathcal{C}$). One can show (see Khmaladze et al. 2006, Corollary 2.1) that the assumption (17) guarantees that $\gamma^s(\hat{K}_n, K)$ is of order $1/n$ in probability.

Note that the covering number of a neighbourhood of a point in a finite dimensional (m dimensional) space satisfies the inequality (17). Therefore, classes of sets which can be continuously labeled by a finite dimensional parameter will also satisfy this inequality. As an example one can consider the class of all polytopes in some fixed bounded subset of $\mathbb{R}^d$ with $k < \infty$ edges (or at most k edges) or the class of all ellipsoids in a given bounded set. One can make a step further and consider all intersections of no more than $l < \infty$ polytopes and ellipsoids from the above classes (that is, one can intersect polytopes with ellipsoids), etc.

Notice that the class $\mathcal{C}$ of all convex bodies (in some fixed bounded subset of $\mathbb{R}^d$) does not fulfill the conditions above. Although this class is uniformly bounded, its covering number increases with δ much faster (see, for example, Bronstein, 1976), and consequently $\hat{K}_n$ converges to K with a rate much slower than $1/n$. Therefore, for this class, the MLE will asymptotically lie much further away from K than the range of Poisson convergence of the likelihood process.

Also notice that the statement on the rate of $\gamma^s(\hat{K}_n, K)$ is not specific to the particular form of the change-set problem: under the same condition (17) the rate of convergence of $\gamma^s(\hat{K}_n, K)$ can again be shown to be $1/n$ in probability in the second change-set problem (see the Introduction). We would only need the condition that the score function $\ln(\mathrm{d}P_1/\mathrm{d}P_0)$ is bounded (or truncated on some fixed level), see Khmaladze et al. (2006).

In order to make use of the previous theorems, we relate now $\mathcal{C}$ to the cylinder Σ. Namely, we choose $T > 0$ such for all sufficiently small $\varepsilon > 0$, we have

$$\tau_\varepsilon(K' \triangle K) \subset \Sigma_T,$$

for all $K' \in \mathcal{C}_\varepsilon(K)$, and we denote by $\mathcal{D}_\varepsilon$ the corresponding class of sets,

$$\mathcal{D}_\varepsilon = \{\tau_\varepsilon(K' \triangle K) : K' \in \mathcal{C}_{\varepsilon T}(K)\}.$$

In the following corollary, Ψ is a Poisson process on Σ with intensity measure Λ as defined in Corollary 1.

Corollary 3 (Ibragimov-Has'minskii theorem for the change-set problem) *For the change-set problem described in* (1)*, let the class $\mathcal{C}$ of all possible change-sets be totally bounded with respect to γ^s. Let $\hat{K}_n$ be the ML estimator of the change-set defined by* (16)*, let $K \in \mathcal{C}$ denote the true change-set and let*

$$\hat{D}_\varepsilon = \arg\max_{D \in \mathcal{D}_\varepsilon} \left[\ln \frac{c_-}{c_+} \cdot \Psi(D) - \Lambda(D) \right].$$

If $\mathcal{C}$ is a locally simple VC-class at K then, as $n \to \infty, \varepsilon \to 0$ with $n\varepsilon \to 1$, the total variation between the distributions of $n\gamma^s(\hat{K}_n, K)$ and $\Lambda(\hat{D}_\varepsilon)$ converges to 0.

Proof As we said, see Khmaladze et al. (2006), Sect. 2, for details, the local entropy condition (17) implies that, for arbitrarily small $\delta > 0$, there exists T large enough such that $\mathbb{P}\{\gamma^s(\hat{K}_n, K) > T/n\} \leq \delta$ or

$$\mathbb{P}\left\{ \max_{K' \in \mathcal{C}_{T/n}} L_n(K', K) = \max_{K' \in \mathcal{C}} L_n(K', K) \right\} \geq 1 - \delta.$$

In other words, for

$$\hat{K}_{T,n} = \arg\max_{K' \in \mathcal{C}_{T/n}} L_n(K'; K),$$

we have

$$\mathbb{P}\left\{\gamma^s(\hat{K}_n, K) = \gamma^s(\hat{K}_{T,n}, K)\right\} \geq \mathbb{P}\left\{\hat{K}_n = \hat{K}_{T,n}\right\} \geq 1 - \delta.$$

On the other hand, for any $x \leq T$,

$$\mathbb{P}\left\{\gamma^s(\hat{K}_{T,n}, K) \leq \frac{x}{n}\right\} = \mathbb{P}\left\{\max_{K' \in \mathcal{C}_{x/n}} L_n(K', K) = \max_{K' \in \mathcal{C}_{T/n}} L_n(K', K)\right\},$$

and both maxima are measurable functionals of the process $\Psi_{n,\varepsilon}$, $\varepsilon = T/n$, on $\mathcal{D}_\varepsilon$. Therefore, $\gamma^s(\hat{K}_{T,n}, K)$ is also a measurable functional of this process and Corollary 2 implies the result. □

4 Unbounded perturbations of K

Theorem 2 and Corollary 1 yield instruments to control the convergence of Ψ_n on the symmetric difference $K' \triangle K$, as long as the collection of sets K' is uniformly bounded, in the sense that $K_{-\varepsilon T} \subset K' \subset K_{\varepsilon T}$, for all K' and some $T > 0$. However, one can envisage applications where it is interesting and useful to consider small but not necessarily bounded deviations from K. In Theorem 3 we will show that we indeed may also consider classes of unbounded sets K'. For this purpose, we make use of the fact that perturbations of K can, in a natural way, be described by functions on $\mathrm{Nor}(K)$. Namely, for a function g on $\mathrm{Nor}(K)$, we call

$$g_{\mathrm{sub}} = \{(t, x, u) \in \Sigma_+ : 0 \leq t \leq g(x, u)\} \cup \{(t, x, u) \in \Sigma_- : g(x, u) \leq t \leq 0\}$$

the *subgraph* of g and we define $K' = K(g)$ as

$$\begin{aligned} K' = &\operatorname{cl}\{z \in \mathbb{R}^d \backslash K : \tau_1(z) \in g_{\mathrm{sub}}\} \\ &\cup \operatorname{cl}\{z \in K \backslash (\partial K \cup S_K) : \tau_1(z) \notin g_{\mathrm{sub}}\} \end{aligned}$$

(here $\operatorname{cl} A$ denotes the closure of a set A). If, for example, g is continuous, then K' is compact and fulfills

$$K' \backslash K = \{z \in \mathbb{R}^d : 0 < d(z) \leq g(p(z), u(z))\} \tag{18}$$

and

$$K \backslash K' = \{z \in K \backslash S_K : 0 < d(z) \leq -g(p(z), u(z))\}. \tag{19}$$

Moreover, if we replace g by εg, $\varepsilon > 0$, then K' is close to K if ε is small and g has appropriate integrability properties. More generally, if g_1, g_2 belong to the subspace of measurable functions g, for which the integrals

$$\int_{\mathrm{Nor}(K)} |g(x,u)|^j \Theta_{d-j}(\mathrm{d}(x,u)), \quad j=1,\dots,d,$$

are finite, and $K_1 = K(\varepsilon g_1)$ and $K_2 = K(\varepsilon g_2)$ are corresponding perturbed sets, then using Theorem 1 one can show that there is asymptotic isometry between the classes of subgraphs and the corresponding functions given by

$$\frac{1}{\varepsilon}\gamma^s(K_1,K_2) = \int_{\mathrm{Nor}(K)} |g_1(x,u) - g_2(x,u))|\Theta_{d-1}(\mathrm{d}(x,u)) + O(\varepsilon),$$

for $\varepsilon \to 0$; recall that $\gamma^s(K_1,K_2) = \mu_d(K_1 \triangle K_2)$ is the symmetric difference metric. In general $K' = K(g)$ will not be convex and, conversely, not every convex body K' near to K can be described in this way. Nevertheless, we get a reasonably large class of sets which seems sufficient for many purposes. This would not be the case if we only consider functions g on ∂K in (18) and (19). For example, if K is a polytope, the boundary of K' would then have to be of spherical shape on all regions outside K which project onto the same point of k-faces of K, $0 \le k \le d-2$. For a similar reason, it is not sufficient to consider only functions g on S^{d-1}. For such functions, the corresponding sets K' would have flat boundary parts on all regions which project onto the same $(d-1)$-face of K.

In order now to extend Theorem 2 to subgraphs of not necessarily bounded functions g on $\mathrm{Nor}(K)$, we will need additional assumptions on the growth of f_n (with respect to f_+ and f_-), namely

$$\begin{aligned} &\sup_{n\in\mathbb{N}} \sup_{z\in\mathbb{R}^d\setminus K} |f_n(z) - f_+(p(z),u(z))| < \infty, \\ &\sup_{n\in\mathbb{N}} \sup_{z\in K\setminus S_K} |f_n(z) - f_-(p(z),u(z))| < \infty. \end{aligned} \tag{20}$$

Of course, both conditions are trivially fulfilled for the process with intensity measure (14).

Theorem 3 *Let K be a convex body and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes fulfilling (12), (13) and (20). Let g be a measurable function on* $\mathrm{Nor}(K)$, *such that*

$$\int_{\mathrm{Nor}(K)} \max(f_\pm(x,u),1)|g(x,u)|^j\, \Theta_{d-j}(\mathrm{d}(x,u)) < \infty, \quad j=1,\dots,d. \tag{21}$$

Then, the distribution of $\Psi_{n,\varepsilon}(\cdot \cap g_{\mathrm{sub}})$ converges in total variation to the distribution of $\Psi(\cdot \cap g_{\mathrm{sub}})$, as $n \to \infty, \varepsilon \to 0$ with $n\varepsilon \to 1$.

Proof We have to show that

$$|\Lambda_{n,\varepsilon}(B) - \Lambda(B)| \to 0,$$

uniformly in $B \subset g_{\mathrm{sub}}$.

We first prove that (13) and (20) imply

$$\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(g_{\mathrm{sub}}\cap\Sigma_+)} |f_n(z)-f_+(p(z),u(z))|\,\mu_d(\mathrm{d}z)\to 0 \tag{22}$$

(and similarly

$$\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(g_{\mathrm{sub}}\cap\Sigma_-)} |f_n(z)-f_-(p(z),u(z))|\,\mu_d(\mathrm{d}z)\to 0).$$

In fact, we have for $T>0$,

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(g_{\mathrm{sub}}\cap\Sigma_+)} |f_n(z)-f_+(p(z),u(z))|\,\mu_d(\mathrm{d}z)\\
&\quad\le \frac{1}{\varepsilon}\int_{K_{\varepsilon T}\setminus K} |f_n(z)-f_+(p(z),u(z))|\,\mu_d(\mathrm{d}z)\\
&\qquad+\frac{c}{\varepsilon}\int \mathbf{1}\{\varepsilon T<d(z)\le \varepsilon g(p(z),u(z))\}\,\mu_d(\mathrm{d}z),
\end{aligned}$$

where

$$c=\sup_{n\in\mathbb{N}}\sup_{z\in\mathbb{R}^d\setminus K}|f_n(z)-f_+(p(z),u(z))|.$$

By Theorem 1, the second integral on the right-hand side equals

$$\begin{aligned}
&\frac{c}{\varepsilon}\int \mathbf{1}\{\varepsilon T<d(z)\le \varepsilon g(p(z),u(z))\}\,\mu_d(\mathrm{d}z)\\
&\quad=\frac{c}{\varepsilon}\sum_{j=1}^{d}\binom{d-1}{j-1}\int_{\mathrm{Nor}(K)}\int_{\min(\varepsilon T,\varepsilon g(x,u))}^{\varepsilon g(x,u)} t^{j-1}\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\\
&\quad\le \frac{c}{d}\sum_{j=1}^{d}\binom{d}{j}\varepsilon^{j-1}\int_{g>T}(g(x,u))^j\,\Theta_{d-j}(\mathrm{d}(x,u))
\end{aligned}$$

and hence can be made arbitrarily small for T large enough, as follows from the integrability condition (21). For this T, the first integral tends to 0, as $n\to\infty,\varepsilon\to 0$ with $n\varepsilon\to 1$ (here we use (13)). The second assertion above follows similarly.

We now proceed as in the proof of Theorem 2. We get for $B\subset g_{\mathrm{sub}}\cap\Sigma_+$,

$$|\Lambda_{n,\varepsilon}(B)-\Lambda(B)|=\left|\int_{\tau_\varepsilon^{-1}(B)} nf_n(z)\,\mu_d(\mathrm{d}z)-\Lambda(B)\right|$$

which is asymptotically bounded from above by

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(g_{\mathrm{sub}})} |f_n(z)-f_+(p(z),u(z))|\,\mu_d(\mathrm{d}z)\\
&\quad+\left|\frac{1}{\varepsilon}\int_{\tau_\varepsilon^{-1}(g_{\mathrm{sub}})} f_+(p(z),u(z))\,\mu_d(\mathrm{d}z)\right.\\
&\quad\left.-\int_{\mathrm{Nor}(K)}\int_0^{g(x,u)} f_+(x,u)\,\mathrm{d}t\,\Theta_{d-1}(\mathrm{d}(x,u))\right|.
\end{aligned}$$

The first summand tends to 0, by (22). The second summand equals

$$\begin{aligned}
&\sum_{j=2}^{d}\binom{d-1}{j-1}\varepsilon^{j-1}\int_{\mathrm{Nor}(K)}\int_0^{g(x,u)} t^{j-1}f_+(x,u)\,\mathrm{d}t\,\Theta_{d-j}(\mathrm{d}(x,u))\\
&\leq \frac{1}{d}\sum_{j=2}^{d}\binom{d}{j}\varepsilon^{j-1}\int_{\mathrm{Nor}(K)}(g(x,u))^j f_+(x,u)\,\Theta_{d-j}(\mathrm{d}(x,u))\\
&= c'\varepsilon,
\end{aligned}$$

in view of (21).

The case $B \subset g_{\mathrm{sub}} \cap \Sigma_-$ follows the same path and can be left to the reader. □

For two functions on $\mathrm{Nor}(K)$, $g_1 \geq 0$ and $g_2 \leq 0$, the union $(g_1)_{\mathrm{sub}} \cup (g_2)_{\mathrm{sub}}$ does not have to be a subgraph of any function (on $\mathrm{Nor}(K)$). However, Theorem 3 immediately generalizes to finite collections of processes $\Psi_{n,\varepsilon}(\cdot \cap (g_i)_{\mathrm{sub}})$, $i = 1, \dots, m$, and hence a limit theorem for

$$\Psi_{n\varepsilon}\left(\cdot \cap \left(\bigcup_{i=1}^{m}(g_i)_{\mathrm{sub}}\right)\right)$$

follows.

Note that for $g \leq 0$ the integrability condition (21) can be weakened, namely we can replace the integration over $\mathrm{Nor}(K)$ with respect to Θ_{d-j} by an integration over the regular boundary points with respect to the absolutely continuous part of the curvature measure C_{d-j}.

5 Higher order components and consistent estimators for support measures

Let us now extract higher order components of Ψ_n near ∂K which are asymptotically driven by the curvature measures $\Theta_{d-j}, j = 2, \dots, d$, and consider limit theorems for them. A first possibility to do this is connected with the simple idea of thinning. This requires to change the limit behaviour of n and ε so as to have more points to choose from. Results in this direction can be found in Khmaladze and Weil (2005).

A second, perhaps, more effective possibility to extract curvature driven components of Ψ_n is associated with the representation (7) and uses narrow shells of width ε near ∂K, at some distance s apart. If we choose $\varepsilon \sim 1/n$, we will be able to use Theorem 2 on each shell. But then it is necessary to keep s fixed. Although general results similar to Theorem 2 would be possible here (but would require some technical assumptions), we concentrate in Theorem 4 on processes Ψ_n which fulfill (14) and consider only the situation outside K.

The idea behind our construction is the following. For any fixed $s > 0$ and $m = 0, \dots, d-1$, we use the Poisson process Ψ_n on the shell $K_{ms+\varepsilon} \setminus K_{ms}$ to define a Poisson process $\Psi_{n,m,\varepsilon}$ on $\mathrm{Nor}(K)$,

$$\Psi_{n,m,\varepsilon}(A) = \sum_{i=1}^{\infty} \mathbf{1}\{0 \le d(Z_i) - ms \le \varepsilon, (p(Z_i), u(Z_i)) \in A\}.$$

The intensity measure of this process is

$$\begin{aligned} & nc_+\mu_d(\{z \in \mathbb{R}^d \setminus K : 0 \le d(Z_i) - ms \le \varepsilon, (p(Z_i), u(Z_i)) \in A\}) \\ & = nc_+\mu_d(A_{ms+\varepsilon} \setminus A_{ms}) = \frac{nc_+}{d} \sum_{j=1}^{d} \binom{d}{j} \varepsilon^j \Theta_{d-j}(K_{ms}; T_{ms}A) \\ & = nc_+\varepsilon\Theta_{d-1}(K_{ms}; T_{ms}A) + o(1) \end{aligned}$$

as $n \to \infty$, $n\varepsilon \to 1$. But (10) connects the measures $\Theta_{d-1}(K_{ms}; T_{ms}A)$, $m = 0, \dots, d-1$, with the lower order support measures $\Theta_{d-i}(A)$. Therefore, if we replace $\Theta_{d-1}(K_{ms}; T_{ms}A)$ on the right-hand side of (10) by $\Psi_{n,m,\varepsilon}(A)$, we obtain the random measure:

$$\Psi_{n,\varepsilon}^{(i)} = \tfrac{1}{s^{i-1}} \sum_{m=0}^{d-1} a_{mi} \Psi_{n,m,\varepsilon}$$

on $\mathrm{Nor}(K)$ with measure $nc_+\varepsilon\Theta_{d-i}(K; \cdot) + o(1)$ as the expected value.

Theorem 4 *Let K be a convex body, $s > 0$, $i \in \{1, \dots, d\}$ and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes with intensity measure* (14)*. If $n\varepsilon \to 1$, as $n \to \infty, \varepsilon \to 0$ and s is fixed, then $\Psi_{n,\varepsilon}^{(i)}$ converges in total variation to a random measure $\Psi^{(i)}$ on* $\mathrm{Nor}(K)$ *with expectation measure $\Lambda^{(i)}$ given by*

$$\Lambda^{(i)} = c_+\Theta_{d-i}(K; \cdot).$$

Proof The map τ_ε maps $\mathbb{R}^d \setminus K$ to Σ_+. For $m = 0, \dots, d-1$, let

$$\tau_{\varepsilon,m} : z \mapsto \left(\frac{d(K_{ms}, z)}{\varepsilon}, p(K_{ms}, z), u(K_{ms}, z) \right)$$

be the corresponding map from $\mathbb{R}^d \setminus K_{ms}$ to $\Sigma_+^{(m)} = [0, \infty) \times \mathrm{Nor}(K_{ms})$ (which can be identified with $[ms, \infty) \times \mathrm{Nor}(K)$). Here $\tau_{\varepsilon,0} = \tau_\varepsilon$ and $\Sigma_+^{(0)} = \Sigma_+$.

Applying Corollary 1 to the process Ψ_n, restricted to $K_{ms+\varepsilon} \setminus K_{ms}$ and mapped to $\Sigma_+^{(m)}$ by $\tau_{\varepsilon,m}$, we obtain a process $\tilde{\Psi}_{n,m,\varepsilon}$ which converges in total variation to a Poisson process $\tilde{\Psi}_m$ on $[0,1] \times \mathrm{Nor}(K_{ms})$, with intensity measure $c_+\mu_1 \times \Theta_{d-1}(K_{ms}; \cdot)$. Moreover, $\tilde{\Psi}_0, \ldots, \tilde{\Psi}_{d-1}$ are independent. Making use of the above-mentioned identification, the projection of $\tilde{\Psi}_{n,m,\varepsilon}$ onto $\mathrm{Nor}(K)$ yields $\Psi_{n,m,\varepsilon}$ and we denote the corresponding projection of $\tilde{\Psi}_m$ by Ψ_m. Thus the Poisson processes $\Psi_{n,m,\varepsilon}$ converge in total variation to the Poisson processes Ψ_m, $m = 0, \ldots, d-1$, the latter are independent and they have intensity measure $c_+\Theta_{d-1}(K_{ms}; \cdot)$. The random measure $\Psi_{n,\varepsilon}^{(i)}$ then converges in total variation to the random measure

$$\Psi^{(i)} = \tfrac{1}{s^{i-1}} \textstyle\sum_{m=0}^{d-1} a_{mi} \Psi_m,$$

and (10) shows that the latter has expectation measure $c_+\Theta_{d-i}(K; \cdot)$. □

The curvature driven processes $\Psi_{n,\varepsilon}^{(k)}, k = 1, \ldots, d$, involve asymptotically small shells (of width $\varepsilon \to 0$) taken at multiple distances ms, $m = 0, \ldots, d-1$, for some fixed s, and the rate we considered was $\varepsilon \sim 1/n$. This rate, however, is in no sense obligatory, necessary or "most natural" in problems concerning the support measures of ∂K. If, for example, we consider the problem of estimation of Θ_{d-k}, $k = 1, \ldots, d$, we will see that much slower rates of ε will appear and it is possible and interesting to consider even fixed ε. Below we clarify this situation and derive the mean square rates of appropriate estimators of Θ_{d-k}, $k = 1, \ldots, d$. Again, we assume that Ψ_n has intensity measure (14) and, in order to simplify the formulas slightly, we choose $c_+ = 1$. In the following, we use the notation $\varepsilon \propto a_n$, if $\varepsilon/a_n \to \text{const}$, $0 < \text{const} < \infty$, as $n \to \infty$. Also we write $\varepsilon = \infty(a_n)$, if $\varepsilon/a_n \to \infty$ and we use the abbreviation

$$\Delta_\varepsilon^k \varphi(A_0) = \Delta_\varepsilon^k \varphi(A_s) \mid_{s=0},$$

for measures φ on $\mathbb{R}^d$. As an estimator of the support measure $\Theta_{d-k}(K; A)$, we consider first

$$\frac{\Delta_\varepsilon^k \Psi_n(A_0)}{n\varepsilon^k c_k}$$

with

$$c_k = \frac{1}{d}\binom{d}{k} \Delta_1^k t^k \mid_{t=0} = \frac{(d-1)!}{(d-k)!}.$$

This estimator is simple and of a clear and intuitive nature.

In our first statement, we consider a class $\mathcal{A}$ of subsets of $\mathrm{Nor}(K)$. We say that $\mathcal{A}$ has *finite bracketing entropy*, if for any $\delta > 0$ there are finitely many pairs of subsets $(\underline{B}_{i,\delta}, \overline{B}_{i,\delta})$, $i = 1, \dots, N_\delta$, called brackets and not necessarily belonging to $\mathcal{A}$, such that any $A \in \mathcal{A}$ can be placed in some bracket, i.e. for any A there is a pair $(\underline{B}_{i,\delta}, \overline{B}_{i,\delta})$ such that $\underline{B}_{i,\delta} \subset A \subset \overline{B}_{i,\delta}$ and $\mu_d(\overline{B}_{i,\delta} \setminus \underline{B}_{i,\delta}) \le \delta$.

Theorem 5 *Let K be a convex body and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes with intensity measure (14). Let $\mathcal{A}$ be a class of Borel subsets of* $\mathrm{Nor}(K)$ *with finite bracketing entropy. For $k \in \{1, \dots, d-1\}$, assume that $\varepsilon \to 0$ and $n\varepsilon^{2k-1} = \infty(\ln n)$. Then*

$$\sup_{A \in \mathcal{A}} \left| \frac{\Delta_\varepsilon^k \Psi_n(A_0)}{n\varepsilon^k c_k} - \Theta_{d-k}(K; A) \right| \to 0 \quad \text{a.s.}$$

For $k = d$, the condition $\varepsilon \to 0$ is not necessary.

Proof $\Delta_\varepsilon^k \Psi_n(A_s) \mid_{s=0}$ is a linear combination of increments

$$\Delta_\varepsilon \Psi_n(A_{j\varepsilon}) = \Psi_n(A_{(j+1)\varepsilon}) - \Psi_n(A_{j\varepsilon})$$

with $j = 0, \dots, k-1$. Each of these increments is a Poisson random variable with expectation $n\Delta_\varepsilon \mu_d(A_{j\varepsilon})$. We first show that, for any A,

$$\frac{\Delta_\varepsilon \Psi_n(A_{j\varepsilon}) - n\Delta_\varepsilon \mu_d(A_{j\varepsilon})}{n\varepsilon^k} \to 0 \quad \text{a.s.} \tag{23}$$

Indeed, from the estimation of the tail of the Poisson distribution we obtain

$$\mathbb{P}(\{|\Delta_\varepsilon \Psi_n(A_{j\varepsilon}) - n\Delta_\varepsilon \mu_d(A_{j\varepsilon})| > cn\varepsilon^k\}) \le \exp\left(-\frac{c^2 n\varepsilon^{2k}}{2\Delta_\varepsilon \mu_d(A_{j\varepsilon})} \right).$$

The increment $\Delta_\varepsilon \mu_d(A_{j\varepsilon})$ is bounded from above by the volume of $K_{(j+1)\varepsilon} \setminus K_{j\varepsilon}$, hence, for all sufficiently small ε, it is not greater than $\varepsilon\vartheta$, for a suitable constant ϑ. Therefore,

$$\exp\left(-\frac{c^2 n\varepsilon^{2k}}{2\Delta_\varepsilon \mu_d(A_{j\varepsilon})} \right) \le \exp\left(-\frac{c^2}{2\vartheta} n\varepsilon^{2k-1} \right) \le \exp\left(-\frac{c^2}{2\vartheta} \beta_n \ln n \right)$$

with

$$\beta_n = \frac{n\varepsilon^{2k-1}}{\ln n} \to \infty.$$

Then,

$$\mathbb{P}(\{|\Delta_\varepsilon \Psi_n(A_{j\varepsilon}) - n\Delta_\varepsilon \mu_d(A_{j\varepsilon})| > cn\varepsilon^k\}) \le \exp\left(-\frac{\beta_n c^2}{2\vartheta} \ln n \right)$$

for any fixed c, and (23) follows from the Borel-Cantelli Lemma. However, since both $A \mapsto \Delta_\varepsilon \Psi_n(A_{j\varepsilon})$ and $A \mapsto n\Delta_\varepsilon \mu_d(A_{j\varepsilon})$ are positive measures on $\mathrm{Nor}(K)$, we can use our assumption of finite bracketing entropy and hence the Glivenko-Cantelli statement also follows for $\mathcal{A}$ (see van der Vaart and Wellner, 1996). Therefore, we proved that

$$\sup_{A\in\mathcal{A}} \left| \frac{\Delta_\varepsilon^k \Psi_n(A_0) - n\Delta_\varepsilon^k \mu_d(A_0)}{n\varepsilon^k} \right| \to 0 \quad \text{a.s.}$$

Now consider

$$\frac{\Delta_\varepsilon^k \mu_d(A_s)|_{s=0}}{\varepsilon^k} - c_k \Theta_{d-k}(K; A)\,. \tag{24}$$

According to the Steiner formula (2), if $k = d$, this difference is equal to 0 for any fixed ε and all A and the statement follows without any further requirements on ε. If $k < d$, then (24) is of order not larger than $\varepsilon\Theta_{d-k-1}(K; A)$ uniformly in A and, hence, tends to 0 as $\varepsilon \to 0$. □

Concerning the rate of convergence of our estimators, we will see that it depends not only on k but also on the nature of the set A, for a fixed k. Namely, for given A and $k \le d-1$, let $l \in \{1, \ldots, d\}$ be such that $\Theta_{d-l}(K; \cdot)$ is the first support measure (in decreasing order) which is non-zero on A,

$$\Theta_{d-j}(K; A) = 0, \quad j = 1, \ldots, l-1, \quad \Theta_{d-l}(K; A) \neq 0, \tag{25}$$

and let $\Theta_{d-m}(K; \cdot), k+1 \le m \le d$, be the first support measure "after" $\Theta_{d-k}(K; \cdot)$ which is non-zero on A,

$$\Theta_{d-j}(K; A) = 0, \quad j = k+1, \ldots, m-1, \quad \Theta_{d-m}(K; A) \neq 0. \tag{26}$$

In the following, we only consider the case where l exists, i.e. not all support measures of K vanish on A. However, (26) need not be satisfied. Note that the case $l = m$ is possible (and implies $\Theta_{d-k}(K; A) = 0$, i.e. our estimator tends to 0). It may also happen that $l < k < m$ and still $\Theta_{d-k}(K; A) = 0$. The case $l = 1$ and $m = k+1$ typically occurs if the boundary ∂K of K is smooth. It can be viewed as the worst case for the rate of convergence. Cases of sets A with $l \neq 1$ and $m \neq k+1$ occur, for example, if K is a polytope.

Theorem 6 *Let K be a convex body and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes with intensity measure (14). Let $A \subset \mathrm{Nor}(K)$ be a Borel set such that (25) is satisfied.*

If also (26) holds with $m \in \{k+1, \ldots, d\}$, then, for

$$\varepsilon = \varepsilon_n \propto n^{-1/(2m-l)},$$

we have

$$n^{(m-k)/(2m-l)}\mathbb{E}\left[\frac{\Delta_\varepsilon^k\Psi_n(A_0)}{n\varepsilon^k c_k}-\Theta_{d-k}(K;A)\right]^2\le C, \tag{27}$$

for some constant C *and all* n*, so that the rate of convergence of the mean square error is* $n^{-(m-k)/(2m-l)}$*. For any other choice of* ε_n*, the left-hand side of* (27) *tends to* ∞*.*

If (26) *is not fulfilled for any* $m\in\{k+1,\dots,d\}$*, then we have, uniformly in* $0\le\varepsilon\le\mathrm{const}$,

$$n\varepsilon^{2k-l}\mathbb{E}\left[\frac{\Delta_\varepsilon^k\Psi_n(A_0)}{n\varepsilon^k c_k}-\Theta_{d-k}(K;A)\right]^2\le C,$$

so that the choice $\varepsilon=\mathrm{const}$ *is possible and leads to the fastest rate* n^{-1} *of convergence for the mean square error.*

Proof Let us decompose the expectation in (27) into variance and square of bias,

$$\begin{aligned}&\mathbb{E}\left[\frac{\Delta_\varepsilon^k\Psi_n(A_0)}{n\varepsilon^k c_k}-\Theta_{d-k}(K;A)\right]^2\\&\quad=\mathbb{E}\left[\frac{\Delta_\varepsilon^k\Psi_n(A_0)}{n\varepsilon^k c_k}-\frac{\Delta_\varepsilon^k\mu_d(A_0)}{\varepsilon^k c_k}\right]^2+\left[\frac{\Delta_\varepsilon^k\mu_d(A_0)}{\varepsilon^k c_k}-\Theta_{d-k}(K;A)\right]^2.\end{aligned} \tag{28}$$

The variance in (28) is the linear combination of variances

$$\mathbb{E}\left[\frac{\Delta_\varepsilon\Psi_n(A_{j\varepsilon})-n\Delta_\varepsilon\mu_d(A_{j\varepsilon})}{n\varepsilon^k c_k}\right]^2=\frac{n\Delta_\varepsilon\mu_d(A_{j\varepsilon})}{n^2\varepsilon^{2k}c_k^2}$$

calculated on each individual shell $A_{(j+1)\varepsilon}\setminus A_{j\varepsilon}$, $j=0,\dots,d-1$. Under our assumptions on the set A, we have

$$\frac{n\Delta_\varepsilon\mu_d(A_{j\varepsilon})}{n^2\varepsilon^{2k}c_k^2}\propto\frac{\varepsilon^l}{n\varepsilon^{2k}} \tag{29}$$

and the variance in (28) is of the same order. At the same time, we have, under our conditions on A,

$$\frac{\Delta_\varepsilon^k\mu_d(A_0)}{\varepsilon^k c_k}-\Theta_{d-k}(K;A)\propto\varepsilon^{m-k}. \tag{30}$$

The optimal rate is obtained if the orders on the right-hand sides of (29) and (30) are the same, hence

$$n^{-1}\varepsilon^{l-2k} = \varepsilon^{2(m-k)}.$$

This yields

$$\varepsilon_n \propto n^{-1/(2m-l)}$$

and gives us the rate

$$\frac{\varepsilon^l}{n\varepsilon^{2k}} \propto \varepsilon_n^{2(m-k)} \propto n^{-2(m-k)/(2m-l)}.$$

If $\varepsilon = o(n^{-1/(2m-l)})$ or $\varepsilon = \infty(n^{-(m-l)/(2m-l)})$ either the variance or the bias will become $\infty(n^{-(m-l)/(2m-l)})$ and the mean square error, normalized by $n^{\frac{m-l}{2m-l}}$, will diverge to ∞.

If (26) is not fulfilled for any $m \in \{k+1, \ldots, d\}$, then

$$\frac{\Delta_\varepsilon^k \mu_d(A_0)}{\varepsilon^k c_k} - \Theta_{d-k}(K; A) = 0\,,$$

which means that there is no bias term, and the rate is determined by the variance alone. In this case, the mean square error is of order $n^{-1}\varepsilon^{l-2k}$. □

We see that for typical sets A where the support measure $\Theta_{d-1}(K; A)$ as well as all lower order support measures are non-zero, the rate of convergence of our estimator is only $n^{\frac{1}{2k+1}}$. This situation, however, can be easily improved by using estimators based on (8), namely

$$\hat{\Theta}_{d-k}(A) = \frac{1}{n\varepsilon^k} \sum_{j=k}^{d} \tilde{b}_{jk} \Delta_\varepsilon^j \Psi_n(A_0).$$

For example, for $d = 3$ we obtain

$$\begin{aligned}
\hat{\Theta}_0(A) &= \frac{3}{2} \frac{\Delta_\varepsilon^3 \Psi_n(A_0)}{n\varepsilon^3}, \\
\hat{\Theta}_1(A) &= \frac{1}{2} \frac{\Delta_\varepsilon^2 \Psi_n(A_0)}{n\varepsilon^2} - \frac{9}{2} \frac{\Delta_\varepsilon^3 \Psi_n(A_0)}{n\varepsilon^2}, \\
\hat{\Theta}_2(A) &= \frac{\Delta_\varepsilon \Psi_n(A_0)}{n\varepsilon} - \frac{1}{2} \frac{\Delta_\varepsilon^2 \Psi_n(A_0)}{n\varepsilon} + 4 \frac{\Delta_\varepsilon^3 \Psi_n(A_0)}{n\varepsilon}.
\end{aligned}$$

For these estimators, we get immediately the following statement.

Theorem 7 *Let K be a convex body and $\{\Psi_n\}, n \in \mathbb{N}$, a sequence of Poisson processes with intensity measure* (14). *Let $A \subset \mathrm{Nor}(K)$ be a Borel set such that* (25) *is satisfied. Then we have, for any bounded $\varepsilon = \varepsilon_n$,*

$$n\varepsilon^{2k-l}\mathbb{E}\left[\hat{\Theta}_{d-k}(A) - \Theta_{d-k}(A)\right]^2 \leq C,$$

for some constant C and all n. Hence, under these assumptions, the rate of convergence of $\tilde{\Theta}_{d-k}(A)$ in the mean square sense is n^{-1}.

We see that for small but fixed ε the uniform rate of convergence over A is now achieved. This rate is essentially higher than that of $\Delta_\varepsilon^k \Psi_n(A_0)/n\varepsilon^k c_k$. For $\varepsilon \to 0$ (the case we are still interested in), the rate of convergence of $\hat{\Theta}_{d-k}(A)$ may still be different, for different types of A, but the differences have been clarified now.

References

Bronstein, E. M. (1976). ε-Entropy of convex sets and functions. (Russian) *Sibirski Mathematiceski Zurnal*, *17*, 508–514; English translation: *Siberian Mathematical Journal*, *17*, 393–398.

Cooke, P. (1979). Statistical inference for bounds of random variables. *Biometrika*, *66*, 367–374.

Deheuvels, P., Mason, D. M. (1994). Functional laws of the iterated logarithm for local empirical processes indexed by sets. *Annals of Probability*, *22*, 1619–1661.

Deheuvels, P., Mason, D. M. (1995). Nonstandard local empirical processes indexed by sets. *Journal of Statistical Planning and Inference*, *45*, 91–112.

Dekkers, A. L. M., Einmahl, J. H. J., de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics*, *17*, 1833–1855.

Einmahl, J. H. J. (1997). Poisson and Gaussian approximation of weighted local empirical processes. *Stochastic Processes and Applications*, *70*, 31–58.

Einmahl, J. H. J., de Haan, L., Piterbarg, V. I. (2001). Nonparametric estimation of the spectral measure of an extreme value distribution. *Annals of Statistics*, *29*, 1401–1423.

Federer, H. (1959). Curvature measures. *Transactions of the American Mathematical Society*, *93*, 418–491.

Heveling, M., Hug, D., Last, G. (2004). Does polynomial parallel volume imply convexity? *Mathematische Annalen*, *328*, 469–479.

Hildebrand, F. B. (1968). *Finite-difference equations and simulations.* Englewood Cliffs: Prentice-Hall.

Hug, D. (1999). *Measures, curvatures and currents in convex geometry*. Habilitationsschrift, Freiburg: Universität Freiburg.

Hug, D., Last, G. (2000). On support measures in Minkowski spaces and contact distributions in stochastic geometry. *Annals of Probability*, *28*, 796–850.

Hug, D., Last, G., Weil, W. (2004). A local Steiner-type formula for general closed sets and applications. *Mathematische Zeitschrift*, *246*, 237–272.

Hug, D., Last, G., Weil, W. (2006). Polynomial parallel volume, convexity and contact distributions of random sets. *Probability Theory and Related Fields*, *135*, 169–200.

Ibragimov, I. A., Has'minskii, R. Z. (1981). *Statistical estimation. Asymptotic theory*. New York: Springer.

Khmaladze, E. (1998). Goodness of fit tests for "chimeric" alternatives. *Statistica Neerlandica*, *52*, 90–111.

Khmaladze, E., Mnatsakanov, R., Toronjadze, N. (2006). The change-set problem for Vapnik-Červonenkis classes. *Mathematical Methods of Statistics*, *15*, 224–231.

Khmaladze, E., Weil, W. (2005). Local empirical processes near boundaries of convex bodies. *Preprint Series, Department of Mathematics, University of Karlsruhe*, 05/13.

Rataj, J., Zähle, M. (2001). Curvature and currents for unions of sets with positive reach II. *Annals of Global Analysis and Geometry*, *20*, 1–21.
Reiss, D. (1993). *A course on point processes*. New York: Springer.
Resnick, S. (1986). *Extreme values, regular variation, and point processes*. New York: Springer.
Ripley, B. D., Rasson, J.-P. (1977). Finding the edge of a Poisson forest. *Journal of Applied Probability*, *14*, 483–491.
Sangwine-Yager, J. R. (1994). A representation for volume involving interior reach. *Mathematische Annalen*, *298*, 1–5.
Schneider, R. (1993). *Convex bodies: the Brunn-Minkowski theory*. *Encyclopedia of mathematics and its applications* vol. 44. Cambridge: Cambridge University Press.
Shephard, G. C., Webster, R. J. (1965). Metrics for sets of convex bodies. *Mathematika*, *12*, 73–88.
Stute, W. (1982). The oscillation behavior of empirical processes. *Annals of Probability*, *10*, 86–107.
van der Vaart, A. W., Wellner, J. A. (1996). *Weak convergence and empirical processes*. New York: Springer.
Zähle, M. (1986). Integral and current representation of Federer's curvature measures. *Archiv der Mathematik*, *46*, 557–567.