

## ON RESAMPLING SCHEMES FOR POLYTOPES

WEINAN QI\* \*\* AND

MAHMOUD ZAREPOUR,\* \*\*\* *University of Ottawa*

### Abstract

The convex hull of a sample is used to approximate the support of the underlying distribution. This approximation has many practical implications in real life. To approximate the distribution of the functionals of convex hulls, asymptotic theory plays a crucial role. Unfortunately most of the asymptotic results are computationally intractable. To address this computational intractability, we consider consistent bootstrapping schemes for certain cases. Let  $S_n = \{X_i\}_{i=1}^n$  be a sequence of independent and identically distributed random points uniformly distributed on an unknown convex set in  $\mathbb{R}^d$  ( $d \geq 2$ ). We suggest a bootstrapping scheme that relies on resampling uniformly from the convex hull of  $S_n$ . Moreover, the resampling asymptotic consistency of certain functionals of convex hulls is derived under this bootstrapping scheme. In particular, we apply our bootstrapping technique to the Hausdorff distance between the actual convex set and its estimator. For  $d = 2$ , we investigate the asymptotic consistency of the suggested bootstrapping scheme for the area of the symmetric difference and the perimeter difference between the actual convex set and its estimate. In all cases the consistency allows us to rely on the suggested resampling scheme to study the actual distributions, which are not computationally tractable.

*Keywords:* Convex hull; random polytopes; bootstrap; asymptotic

2010 Mathematics Subject Classification: Primary 62G09

Secondary 62G20; 62G32

### 1. Introduction

For any subset  $A \subset \mathbb{R}^d$  ( $d \geq 2$ ), let  $\text{conv}(A)$  be the convex hull of  $A$ , i.e. the smallest convex set containing  $A$ . Most of the discussion below is for  $d = 2$ , but when the results are valid for higher dimensions we use  $d$  for the dimension instead. Suppose  $S_n = \{X_i\}_{i=1}^n$  is a sequence of independent and identically distributed (i.i.d.) random points in  $\mathbb{R}^d$ . The random convex polytopes (random convex hull)  $K_n = \text{conv}(S_n)$  is the smallest convex set containing observations  $\{X_i\}_{i=1}^n$ . Random convex polytopes (in short, random polytopes) have been widely studied. For example, [MacDonald et al. \(1980\)](#) applied  $K_n$  for the estimation of the territorial range of a wildlife species by tagging an individual of this species with a radio transmitter and recording the position as  $X_i$  after release. [Ripley and Rasson \(1977\)](#) applied  $K_n$  for the estimation of the homogeneous planar Poisson process support. Furthermore, in ordered statistics, the method of convex-hull peeling constructs a nested sequence of random polytopes based on the sample points. In this context, the number of convex layers for a sequence of

---

Received 7 August 2018; revision received 2 May 2019.

\* Postal address: Department of Mathematics and Statistics, University of Ottawa, Canada.

\*\* Email address: [wqi049@uottawa.ca](mailto:wqi049@uottawa.ca)

\*\*\* Email address: [zarepour@uottawa.ca](mailto:zarepour@uottawa.ca)

given points is called convex-hull peeling depth, and the convex layers themselves are the depth contours for this notion of data depth. In other words  $\text{conv}(S_n)$  plays the role of the extreme layer, like the maximum of a sample, and the most internal layer can be considered as the multivariate median. For an extensive connection with the concept of data depth we refer the reader to [Liu et al. \(1979\)](#), [Tukey \(1974\)](#), and [Barnett \(1976\)](#).

Throughout this paper we use the following notation. For any Borel set  $A \subset \mathbb{R}^d$  ( $d \geq 2$ ), its volume is denoted by  $|A|$  (the Lebesgue measure of  $A$ ). The sample points,  $S_n = \{X_i\}_{i=1}^n$ , are i.i.d. random elements taking values in  $\mathbb{R}^d$  with the probability distribution  $F$ . Suppose  $F$  has the convex support  $K$ , which in this paper is referred to as the ‘underlying set’. When  $F$  is the uniform distribution, the set  $K_n = \text{conv}(S_n)$  is called the uniform polytope. For functionals of  $K_n$ , denote the volume by  $|K_n|$ , the perimeter  $\partial(K_n)$  (the boundary of  $K_n$ ) by  $L_n$ , and the probability content by  $F(K_n) = P(X_1 \in K_n)$ . In addition, if  $K$  is the underlying set, we denote the set difference  $K \setminus K_n$  by  $D_n$  and the Hausdorff distance between  $K$  and  $K_n$  by  $H(K, K_n)$ . In the following section we provide a short list of some well-known results, which serve as the theoretical basis for this paper.

### 2. Preliminaries

In this section we review some existing results. Let  $F$  be the uniform distribution on the underlying convex set  $K$ . Suppose the underlying set  $K \subset \mathbb{R}^d$  is a convex polygon with  $r$  ( $\geq 3$ ) vertices. Problems related to convex hulls have received much attention from many authors. For  $d = 2$ , according to [Cabo and Groeneboom \(1994\)](#), we have, as  $n \rightarrow \infty$ ,

$$\frac{|D_n|/|K| - \beta_n}{\alpha_n} \xrightarrow{D} \mathcal{N}(0, 1), \tag{2.1}$$

where  $\alpha_n = \frac{1}{n} \sqrt{\frac{28}{27} r \log n}$ ,  $\beta_n = \frac{2}{3n} r \log n$ , and ‘ $\xrightarrow{D}$ ’ is convergence in distribution. For any convex set  $K \subset \mathbb{R}^d$ ,

$$\lim_{n \rightarrow \infty} n^{2/(d+1)} E(|D_n|) = c(d)|K|^{2/(d+1)} \int_{\partial K} (\kappa(z))^{1/(d+1)} dz, \tag{2.2}$$

where  $\kappa$  is the generalized Gaussian curvature and  $c(d)$  is a constant only depending on  $d$ . See [Schütt \(1994\)](#) for details.

Let  $K$  be a convex subset of  $\mathbb{R}^2$ . Suppose the boundary of  $K$ , i.e.  $\partial K$ , has curvature bounded away from 0 and  $\infty$ , and  $\partial K \in C^2(\mathbb{R})$  (the set of all functions with two continuous derivatives with domain  $\mathbb{R}$ ). For definitions and more details, see the Appendix A. Let  $L$  and  $L_n$  be the perimeters of  $K$  and  $K_n$  respectively. [Bräker and Hsing \(1998\)](#) showed that, as  $n \rightarrow \infty$ ,

$$n^{5/6}(|K_n| - E(|K_n|), L_n - E(L_n)) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \Sigma), \tag{2.3}$$

where  $\Sigma$  is a constant matrix. Convergence of expectations is given by

$$\lim_{n \rightarrow \infty} n^{2/3} E(L - L_n) = c_1$$

and

$$\lim_{n \rightarrow \infty} n^{2/3} E(|D_n|) = c_2, \tag{2.4}$$

where  $c_1$  and  $c_2$  are constants.

Let  $d (\geq 2)$  and  $s (\leq d - 1)$  be two nonnegative integers. Let  $K \subset \mathbb{R}^d$  be a convex set with  $\partial K \in \mathcal{C}^2(\mathbb{R}^{d-1})$  and positive Gaussian curvature bounded away from 0 and  $\infty$ . Let  $\Phi$  be the cumulative distribution function of the standard normal distribution. Then Reitzner (2005) proved that there exists a sequence of constants  $C_n$ , bounded between two positive constants, depending only on  $K$ , and another constant  $c$ , such that

$$\left| \mathbb{P}\left(\frac{|K_n| - \mathbb{E}(|K_n|)}{\sqrt{C_n n^{-1-2/(d+1)}}} \leq t\right) - \Phi(t) \right| \leq c\varepsilon(n), \tag{2.5}$$

where

$$\varepsilon(n) = n^{-1/(2(d+1))}(\log n)^{2+2/(d+1)}.$$

Let  $K$  be a convex polygon with interior angles  $\theta_1, \theta_2, \dots, \theta_r$ . Bräker et al. (1998) derived that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}H(K, K_n) \leq x) = G_1(x), \tag{2.6}$$

where

$$G_1(x) := \prod_{k=1}^r (1 - p_k(x))$$

with

$$p_k(x) = \begin{cases} \int_0^{\theta_k} h_k(x, \theta) \, d\theta + \exp\left\{-\frac{x^2}{2|K|} \tan \theta_k\right\}, & 0 < \theta_k < \frac{\pi}{2}, \\ \int_{\theta_k - \pi/2}^{\pi/2} h_k(x, \theta) \, d\theta, & \frac{\pi}{2} \leq \theta_k < \pi, \end{cases} \tag{2.7}$$

and

$$h_k(x, \theta) = \exp\left\{-\frac{x^2}{2|K|}(\tan \theta_k + \tan(\theta_k - \theta))\right\} \frac{x^2}{2|K|} \tan^2 \theta.$$

Bräker et al. (1998) also derived the limit theory for uniform polytopes, when the underlying set  $K$  satisfies a certain smoothness condition. Let  $K$  be a bounded convex set  $K \subset \mathbb{R}^2$ . Suppose the boundary of  $K$  has length  $L$  and we parameterize it (positively oriented) as  $t \rightarrow c(t)$ , with  $t$  the arc length between  $c(0)$  and  $c(t)$ . Suppose the curvature  $\mathcal{K}(t) = |\ddot{c}(t)|$  is well defined, bounded away from 0 and  $\infty$ , and has a bounded derivative. Define the function

$$\lambda(t) = |K| \sqrt{\mathcal{K}(t)}, \quad 0 \leq t < L,$$

and let  $\lambda_0 := \max_{t \in [0, L]} \lambda(t)$ . Suppose that there exists some bounded sequence of nonnegative constants  $v_n$  and positive constant  $\mu$  such that, as  $n \rightarrow \infty$ ,

$$\frac{(\log n)^{v_n}}{|K|} \int_0^L \exp\left\{-\gamma_n \left(\frac{\lambda_0}{\lambda(t)} - 1\right) \log n\right\} dt \rightarrow \mu \in (0, \infty), \tag{2.8}$$

where

$$\gamma_n = \frac{1}{3} + \left(\frac{2}{3} - v_n\right) \frac{\log \log n}{\log n}.$$

Denote

$$c_n = \frac{3\sqrt{2}}{8} \lambda_0 \gamma_n, \quad a_n = n^{-2/3} (\log n)^{-1/3}, \quad \text{and } b_n = n^{-2/3} (c_n \log n)^{2/3}.$$

Then

$$G_2(x) := \lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{H(K, K_n) - b_n}{a_n} \leq x \right) = \exp\{-d_1 e^{-d_2 x}\}, \quad x \in \mathbb{R}, \quad (2.9)$$

for some constants  $d_1$  and  $d_2$ .

In practice, the statistical properties of different functionals of random polytopes are hard to use. For example, the limiting results in (2.6) are impossible to use. To assess the distributions of these functionals, a feasible idea is to apply the bootstrap method. In a general framework, let  $S_n = \{X_i\}_{i=1}^n$  be a sequence of i.i.d. random points drawn from the underlying distribution  $F$ . Let  $\phi_n(S_n; F)$  be a real-valued functional of interest. To draw a precise inference for  $\phi_n$ , we need to know the exact form of the underlying distribution  $F$ . The plug-in principle suggests using some other known approximate distribution to replace the underlying distribution  $F$ . Efron (1979) suggested replacing  $F$  with the empirical distribution  $F_n$ , where

$$F_n(-\infty, x] = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_i \leq x),$$

with  $\mathbf{1}$  the indicator function.

However, this plug-in principle will still need to be examined for consistency for every target functional. Specifically, suppose

$$\mathbb{P} \left( \frac{\phi_n(S_n; F) - b_n}{a_n} \leq x \right) \rightarrow G(x) \quad \text{as } n \rightarrow \infty$$

for any continuity point  $x \in \mathbb{R}$  of  $G$ . Conditional on  $S_n$ , draw a bootstrap sample of points  $S_{n,m}^* = \{X_{n,i}^*\}_{i=1}^m$  from  $F_n$ . For convenience, the regular bootstrap takes  $m = n$ ; however, for generality we do not necessarily assume  $m = n$  in the forthcoming sections. We would like to have

$$\mathbb{P} \left( \frac{\phi_m(S_{n,m}^*; F_n) - b_m}{a_m} \leq x \mid S_n \right) \xrightarrow{\mathbb{P}} G(x),$$

where ' $\xrightarrow{\mathbb{P}}$ ' is convergence in probability.

However, the asymptotic failure of the regular bootstrap ( $m = n$ ) for extremes is well known (see counterexamples in Bickel and Freedman (1981)). Obviously, this asymptotic failure will still be the case for random polytopes (see Zarepour (1999)). When the regular bootstrap fails, some authors suggest applying the  $m$  out of  $n$  resampling method, i.e. setting the resample size as  $m = o(n)$  ( $m/n \rightarrow 0$  as  $n \rightarrow \infty$ ). By assuming  $m = o(n)$ , Zarepour (1999) studied the bootstrapping point processes and proved the bootstrap consistency if the underlying distribution  $F$  satisfies the regularly varying condition:  $nF(a_n^{-1}X_1 \in \cdot) \xrightarrow{V} \mu(\cdot)$ , where  $\mu$  is a Radon measure and ' $\xrightarrow{V}$ ' is vague convergence (Kallenberg (1983)). By the continuous mapping theorem, the conclusion can also be applied to the regular bootstrap for convex hulls when the underlying distribution has regularly varying tails.

In this paper we investigate the consistency of some bootstrapping schemes for convex hulls. In Section 3 we examine the consistency of the regular bootstrap for polygons. In Section 4, a semi-parametric bootstrapping scheme is introduced and with two simple examples (in which the underlying set  $K$  is either a circle or a rectangle) we discuss how this approach works. This new semi-parametric bootstrapping scheme is developed in more detail for more general underlying sets in Section 5. The bootstrapping consistency of the Hausdorff distance on the uniform polytopes is proved in this section. Moreover, we also derive results for the symmetric

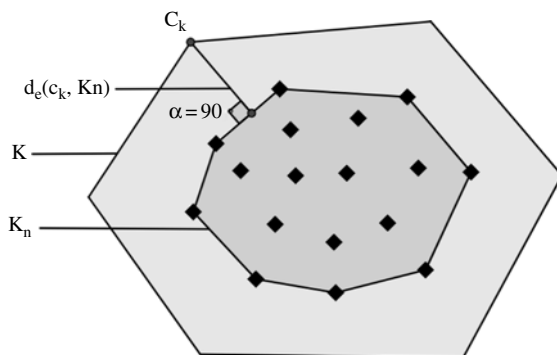


FIGURE 1: A representation of  $d_e(c_k, K_n)$ , where the sample points are drawn uniformly from  $K$ , and  $K_n$  is the convex hull of the sample.

difference and the perimeter difference (see the forthcoming definitions) in Section 6. Since the technique is similar to Section 5, only concise proofs are given for these two examples.

### 3. The consistency of the regular bootstrap on the Euclidean distance

Suppose  $S_n = \{X_i\}_{i=1}^n$  is a sequence of i.i.d. random points uniformly drawn from a polygon  $K$  with vertices  $c_1, c_2, \dots, c_r$  and interior angles  $\theta_1, \theta_2, \dots, \theta_r$ . Let  $d_e$  be the Euclidean metric. Define the distance between a point  $x \in \mathbb{R}^2$  and a set  $A \subset \mathbb{R}^2$  by  $d_e(x, A) := \inf\{d_e(x, y) : y \in A\}$ , and as usual take  $K_n = \text{conv}(S_n)$ . Given the vertex  $c_k$ , Bräker et al. (1998) proved that

$$\lim_{n \rightarrow \infty} P(\sqrt{n}d_e(c_k, K_n) \leq x) = 1 - p_k(x),$$

where  $p_k$  is defined in (2.7). See Figure 1 for a graphical description of  $d_e(c_k, K_n)$ . Now, conditionally on  $S_n$ , the bootstrapping points  $S_{n,m}^* = \{X_{n,i}^*\}_{i=1}^m$  are drawn from the empirical distribution  $F_n(-\infty, x] = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_i \leq x)$ . Define  $K_{n,m}^* := \text{conv}(S_{n,m}^*)$ . We would like to know whether bootstrapping is asymptotically valid for  $d_e(c_k, K_{n,m}^*)$ . To prove consistency, we will combine the point-process techniques in Zarepour (1999) and the results in Bräker et al. (1998). Using the previously introduced notation, the following theorem is stated as follows.

**Theorem 3.1.** *Under the assumptions above, suppose  $m = o(n)$ . Then for any integer  $k$ , where  $0 \leq k \leq r$ , we have*

$$P(\sqrt{m}d_e(c_k, K_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} 1 - p_k(x)$$

as  $n, m \rightarrow \infty$ , where  $p_k$  is defined in (2.7).

*Proof.* We divide our proof into three steps as follows.

*Step 1: Construction of the original point process.* Here and later, the additions and subtractions between a set and a point are Minkowski sum and difference. In other words, if  $K$  is a nonempty subset of  $\mathbb{R}^d$  and  $\mathbf{x} \in \mathbb{R}^d$  then

$$K \pm \mathbf{x} = \{k \pm x : k \in K\}.$$

Let  $\lambda$  be Lebesgue measure. For any measurable subset  $A \subset \mathbb{R}^2$ , we have

$$\begin{aligned} nP(\sqrt{n}(X_1 - \mathbf{c}_k) \in A) &= nP\left(X_1 \in \left(\frac{A}{\sqrt{n}} + \mathbf{c}_k\right)\right) \\ &= \frac{n}{|K|} \lambda\left(\frac{A}{\sqrt{n}} \cap (K - \mathbf{c}_k)\right) \\ &= \frac{1}{|K|} \lambda(A \cap \sqrt{n}(K - \mathbf{c}_k)) \\ &\rightarrow \frac{1}{|K|} \lambda(A) \quad \text{as } n \rightarrow \infty. \end{aligned}$$

This is due to the fact that  $\sqrt{n}(K - \mathbf{c}_k) \uparrow \mathbb{R}^2$ . Therefore, according to Resnick (1987) we have

$$nP(\sqrt{n}(X_1 - \mathbf{c}_k) \in \cdot) \xrightarrow{v} \frac{1}{|K|} \lambda(\cdot).$$

Define the point process  $\xi_{n,k} = \sum_{i=1}^n \delta_{\sqrt{n}(X_i - \mathbf{c}_k)}$ . Here,  $\xi_{n,k}$  can be regarded as a random element in  $M_p(\text{Cone } \mathbf{c}_k)$  (the collection of all nonnegative point measures on the cone  $\mathbf{c}_k$ ; see Resnick (1987)) endowed with vague topology. The Laplace functional of  $\xi_{n,k}$  will be

$$\begin{aligned} \Phi_{\xi_{n,k}}(f) &= \left( \int_{\mathbb{R}^2} \exp\{-f(\sqrt{n}(\mathbf{x} - \mathbf{c}_k))\} \frac{1}{|K|} \mu(d\mathbf{x}) \right)^n \quad \left( \mu = \lambda(\cdot \cap K) \right) \\ &= \left( 1 - \frac{\int_{\mathbb{R}^2} (1 - e^{f(\mathbf{y})}) \tilde{\mu}(d\mathbf{y})}{n} \right)^n \quad (\mathbf{y} = \sqrt{n}(\mathbf{x} - \mathbf{c}_k)) \\ &\rightarrow \exp \left\{ - \int_{\mathbb{R}^2} (1 - e^{f(\mathbf{y})}) \frac{1}{|K|} \lambda(d\mathbf{y}) \right\}, \end{aligned}$$

where

$$\tilde{\mu}(A) = \frac{n}{|K|} \lambda\left(\left(\frac{A}{\sqrt{n}} + \mathbf{c}_k\right) \cap K\right) = \frac{1}{|K|} \lambda(A \cap \sqrt{n}(K - \mathbf{c}_k)) \rightarrow \frac{1}{|K|} \lambda(A) \quad (n \rightarrow \infty)$$

for any Borel set  $A \subset \mathbb{R}^2$ . Hence, given the vertex  $\mathbf{c}_k$ , we have  $\xi_{n,k} \xrightarrow{D} \xi$  as  $n \rightarrow \infty$ , where  $\xi$  is a Poisson point process with intensity measure  $\frac{1}{|K|} \lambda$ .

*Step 2: Construction of the bootstrapping point process.* Conditionally on  $S_n$ , let  $\xi_{n,m,k}^* = \sum_{i=1}^m \delta_{\sqrt{m}(X_{n,i}^* - \mathbf{c}_k)}$ , which is also a random element in  $M_p(\text{Cone } \mathbf{c}_k)$ . The Laplace functional of  $\xi_{n,m,k}^*$  is

$$\begin{aligned} \Phi_{\xi_{n,m,k}^*}(f) &= \left( \int_{\mathbb{R}^2} e^{-f(\sqrt{m}(\mathbf{x} - \mathbf{c}_k))} F_n(d\mathbf{x}) \right)^m \\ &= \left( \int_{\mathbb{R}^2} e^{-f(\sqrt{m}(\mathbf{x} - \mathbf{c}_k))} \frac{1}{n} \sum_{i=1}^n \delta_{X_i}(d\mathbf{x}) \right)^m \\ &= \left( \frac{1}{n} \sum_{i=1}^n e^{f(\sqrt{m}(X_i - \mathbf{c}_k))} \right)^m \end{aligned}$$

$$\begin{aligned}
 &= \left( \int_{\mathbb{R}^2} e^{f(\mathbf{y})} \frac{1}{n} \sum_{i=1}^n \delta_{\sqrt{m}(X_i - \mathbf{c}_k)}(\mathbf{d}\mathbf{y}) \right)^m \\
 &= \left( 1 - \frac{\int_{\mathbb{R}^2} (1 - e^{-f(\mathbf{y})}) \frac{m}{n} \sum_{i=1}^n \delta_{\sqrt{m}(X_i - \mathbf{c}_k)}(\mathbf{d}\mathbf{y})}{m} \right)^m.
 \end{aligned} \tag{3.1}$$

If

$$\eta_{n,m,k} := \frac{m}{n} \sum_{i=1}^n \delta_{\sqrt{m}(X_i - \mathbf{c}_k)},$$

then (3.1) becomes

$$\Phi_{\xi_{n,m,k}^*}(f) = \left( 1 - \frac{\int_{\mathbb{R}^2} (1 - e^{-f(\mathbf{y})}) \eta_{n,m,k}(\mathbf{d}\mathbf{y})}{m} \right)^m. \tag{3.2}$$

Since  $m\mathbb{P}(\sqrt{m}(X_1 - \mathbf{c}_k) \in \cdot) \xrightarrow{V} \frac{1}{|K|} \lambda(\cdot)$  and  $(e^x - 1)/x \rightarrow 1$  as  $x \rightarrow 0$ , we have

$$\begin{aligned}
 &E(\exp\{t\eta_{n,m,k}(A)\}) \\
 &= (e^{tm/n} \mathbb{P}(\sqrt{m}(X_1 - \mathbf{c}_k) \in A) + 1 - \mathbb{P}(\sqrt{m}(X_1 - \mathbf{c}_k) \in A))^n \\
 &= \left( \frac{m\mathbb{P}(\sqrt{m}(X_1 - \mathbf{c}_k) \in A) (\frac{n}{m}(e^{tm/n} - 1))t}{n} + 1 \right)^n \\
 &\rightarrow \exp \left\{ \frac{t}{|K|} \lambda(A) \right\} \quad \text{as } n, m \rightarrow \infty.
 \end{aligned}$$

Therefore, from [Kallenberg \(1983\)](#), Theorem 4.2, page 32, we have

$$\eta_{n,m,k} \xrightarrow{P} \frac{1}{|K|} \lambda \quad \text{as } n, m \rightarrow \infty. \tag{3.3}$$

See also [Zarepour \(1999\)](#). Combining (3.3) with (3.2) implies that, as  $n, m \rightarrow \infty$ ,

$$\xi_{n,m,k}^* \xrightarrow{D} \xi \quad \text{in probability.} \tag{3.4}$$

*Step 3: The continuous mapping theorem.* For any measure  $\eta \in M_p(\text{Cone } \mathbf{c}_k)$ , suppose  $f_k$  maps  $\eta$  to the smallest distance between the origin and the convex hull of the points of  $\eta$ . It is easy to find that

$$f_k(\xi_{n,m,k}^*) = \sqrt{m}d_e(\mathbf{c}_k, K_{n,m}^*).$$

Now apply Skorokhod’s representation theorem and the continuous mapping theorem on (3.4) to get

$$\mathbb{P}(\sqrt{m}d_e(\mathbf{c}_k, K_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} \mathbb{P}(f_k(\xi) \leq x)$$

as  $n \rightarrow \infty$ , noticing that  $m/n \rightarrow 0$ . From [Bräker et al. \(1998\)](#), we have

$$\mathbb{P}(f_k(\xi) \leq x) = 1 - p_k(x),$$

completing the proof. □

This proves that the result in [Bräker et al. \(1998\)](#), as stated in Theorem 3.1, is valid for the regular bootstrap when  $m/n \rightarrow 0$ .

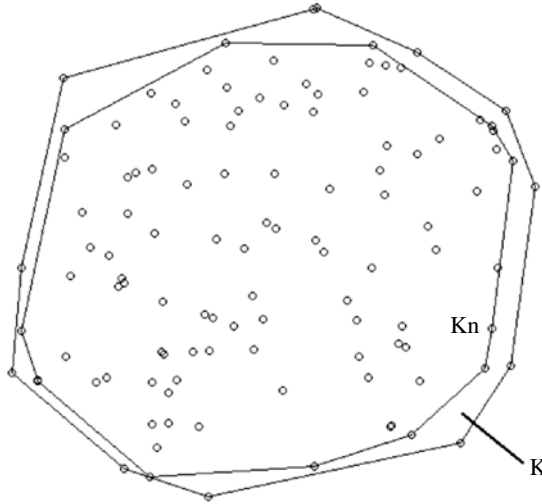


FIGURE 2:  $K$  is a convex polygon. The sample size  $n = 100$ .  $K_n$  is the convex hull of the sample points.

### 4. New bootstrapping scheme on uniform samples

Here we introduce the maximum likelihood estimator of the underlying set as follows.

**Definition 4.1.** Suppose  $S_n = \{X_i\}_{i=1}^n$  is a sequence of i.i.d. random points drawn from  $F$ , where  $F$  has a nonempty, convex support  $K \subset \mathbb{R}^d$  and  $K$  is assumed to be unknown. We denote the maximum likelihood estimator of  $K$  by  $ML(S_n)$ .

Since  $ML(S_n)$  is the maximum likelihood estimator for  $K$ , the geometric features of  $K$  determine the geometric features of  $ML(S_n)$ . For example, if we only know that  $K$  is a convex set, then  $ML(S_n) = \text{conv}(S_n) = K_n$  (see Figure 2).

Based on the maximum likelihood estimator of the underlying set, we introduce the following semi-parametric bootstrap method when we know  $F$  to a certain degree. Let  $F$  be the underlying distribution with nonempty convex support. Let  $F_A$  be the restriction of  $F$  on a set  $A \subset \mathbb{R}^d$ , i.e. for any Borel set  $B \subset \mathbb{R}^d$ ,

$$F_A(B) = \frac{F(B \cap A)}{F(A)}.$$

Then in our *semi-parametric bootstrapping scheme*, points  $S_{n,m}^* = \{X_{n,i}^*\}_{i=1}^m$  are independently drawn from  $F_{ML(S_n)}$  when the sample points  $S_n$  are given.

In the following two examples we consider two simple cases to explain how our approach works. In the first example we assume that the underlying set is a disk with an unknown radius  $R$  centered at the origin; in the second, we consider a rectangle as the underlying set. Additionally, the distribution on both sets is assumed to be uniform.

**Example 4.1.** Let  $d_e$  be the Euclidean metric on  $\mathbb{R}^2$ . Define the Euclidean norm  $\|x\| = d_e(\mathbf{0}, x)$  for all  $x \in \mathbb{R}^2$ . Denote  $B(\mathbf{0}, r) := \{x \in \mathbb{R}^2 : \|x\| \leq r\}$ . Let  $S_n = \{X_i\}_{i=1}^n$  be the i.i.d. random points uniformly drawn from the disk  $C := B(\mathbf{0}, R)$ . Let  $C_n := B(\mathbf{0}, R_n)$  where

$$R_n = \max_{1 \leq i \leq n} \|X_i\|,$$



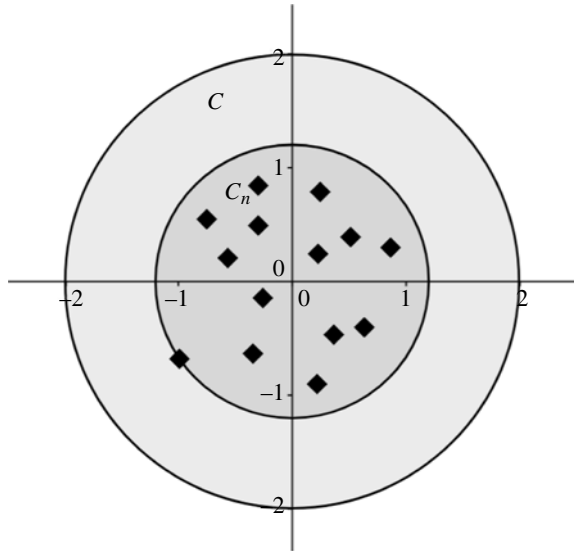


FIGURE 3:  $C$  is a disk centered at the origin.

i.e.  $C_n$  is the smallest disk with its center at the origin and containing all the sample points (see Figure 3). It is easy to check that  $C_n$  is the maximum likelihood estimator for  $C$ , i.e.  $ML(S_n) = C_n$ . Since the sample points are drawn independently and uniformly from  $C$ , we have

$$\begin{aligned} P(|R_n - R| > \varepsilon) &= P(R_n < R - \varepsilon) \\ &= \left( \frac{\pi(R - \varepsilon)^2}{\pi R^2} \right)^n \rightarrow 0 \quad \text{as } n \rightarrow \infty, \end{aligned}$$

which implies  $R_n \xrightarrow{P} R$ . Indeed,

$$\text{var } R_n = \frac{nR^2}{(n + 1)(2n + 1)^2} = \Theta(n^{-2}).$$

Therefore,

$$\sum_{n=1}^{\infty} P(|R_n - R| > \varepsilon) \leq \varepsilon^{-2} \sum_{n=1}^{\infty} \text{var } R_n < \infty,$$

which implies  $R_n$  converges to  $R$  completely (i.e.  $R_n \xrightarrow{C} R$ , which is stronger than almost sure convergence). See [Hsu and Robins \(1974\)](#) for the definition.

Now we can evaluate the asymptotic validity of a functional using our bootstrapping scheme on this sample. Given  $S_n$ , define

$$R_{n,m}^* := \max_{1 \leq i \leq m} \|X_{n,i}^*\|,$$

where  $X_{n,i}^*$ ,  $1 \leq i \leq m$ , are drawn uniformly from  $C_n$ . Notice that as  $n \rightarrow \infty$  we have

$$\begin{aligned} P(n(R - R_n) > x) &= P(R_n < R - x/n) \\ &= \left( \frac{\pi(R - x/n)^2}{\pi R^2} \right)^n \rightarrow \exp\{-2x/R\}, \end{aligned} \tag{4.1}$$

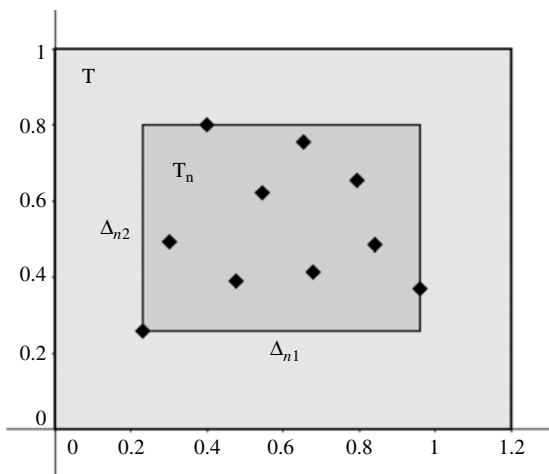


FIGURE 4:  $T$  is a rectangle with its edges parallel to the coordinate axes.

and as  $n, m \rightarrow \infty$  we have

$$\begin{aligned}
 P(m(R_n - R_{n,m}^*) > x \mid C_n) &= P(R_{n,m}^* < R_n - x/m \mid C_n) \\
 &= \left( \frac{\pi(R_n - x/m)^2}{\pi R_n^2} \right)^m \xrightarrow{\text{a.s.}} \exp\{-2x/R\}.
 \end{aligned}
 \tag{4.2}$$

From (4.1) and (4.2), it can be seen that the bootstrap approximation is valid in this example. Notice that the standard condition of  $m = o(n)$  ( $m$  out of  $n$  bootstrapping) is *not* required.

**Example 4.2.** Let  $(X_i, Y_i)$ ,  $1 \leq i \leq n$ , be a sequence of i.i.d. random points uniformly distributed on a rectangle  $T = [0, a] \times [0, b]$ , where  $a$  and  $b$  are two unknown positive constants. Let  $T_n := [\min_{1 \leq i \leq n} X_i, \max_{1 \leq i \leq n} X_i] \times [\min_{1 \leq i \leq n} Y_i, \max_{1 \leq i \leq n} Y_i]$ , i.e. the smallest rectangle containing all the sample points, and with all edges parallel to the coordinate axes (see Figure 4). Since  $T_n$  is the maximum likelihood estimator for  $T$ , we have  $ML(S_n) = T_n$ . Here,  $\partial T$  and  $\partial T_n$  consist of the four edges of  $T$  and  $T_n$ , respectively. For any  $A, B \subset \mathbb{R}^2$ , let  $h(A, B)$  be the shortest horizontal or vertical distance between  $A$  and  $B$ , i.e.

$$h(A, B) = \min\{x_0, y_0\},$$

where

$$\begin{aligned}
 x_0 &= \inf\{|x_1 - x_2| : (x_1, y) \in A, (x_2, y) \in B, y \in \mathbb{R}\}, \\
 y_0 &= \inf\{|y_1 - y_2| : (x, y_1) \in A, (x, y_2) \in B, x \in \mathbb{R}\}.
 \end{aligned}$$

Therefore, as  $n \rightarrow \infty$  we have

$$\begin{aligned}
 P(nh(\partial T, \partial T_n) > x) &= P(h(\partial T, \partial T_n) > x/n) \\
 &= P^n(h(\partial T, X_i) > x/n) \rightarrow e^{-2x(a^{-1}+b^{-1})}.
 \end{aligned}
 \tag{4.3}$$

Given  $S_n$ , let  $(X_{n,i}^*, Y_{n,i}^*), 1 \leq i \leq m$ , be a sequence of i.i.d. random points uniformly drawn from  $T_n$ . Denote

$$T_{n,m}^* = [\min_{1 \leq i \leq m} X_{n,i}^*, \max_{1 \leq i \leq m} X_{n,i}^*] \times [\min_{1 \leq i \leq m} Y_{n,i}^*, \max_{1 \leq i \leq m} Y_{n,i}^*],$$

$$\Delta_{n,1} := \max_{1 \leq i \leq n} X_i - \min_{1 \leq i \leq n} X_i,$$

$$\Delta_{n,2} := \max_{1 \leq i \leq n} Y_i - \min_{1 \leq i \leq n} Y_i.$$

Since  $\Delta_{n,1}$  and  $\Delta_{n,2}$  are both nondecreasing sequences and  $\Delta_{n,1} \xrightarrow{P} a, \Delta_{n,2} \xrightarrow{P} b$ , then  $\Delta_{n,1} \xrightarrow{\text{a.s.}} a, \Delta_{n,2} \xrightarrow{\text{a.s.}} b$ . Therefore, as  $n, m \rightarrow \infty$ , we have

$$\begin{aligned} P(mh(\partial T_n, \partial T_{n,m}^*) > x \mid S_n) &= P^m(h(\partial T_n, X_{n,1}^*) > x/m \mid S_n) \\ &= \Delta_{n,1}^{-m} \Delta_{n,2}^{-m} \left(\Delta_{n,1} - \frac{2x}{m}\right)^m \left(\Delta_{n,2} - \frac{2x}{m}\right)^m \\ &\xrightarrow{\text{a.s.}} e^{-2x(a^{-1}+b^{-1})}. \end{aligned} \tag{4.4}$$

From (4.3) and (4.4), we can conclude that our bootstrap approximation is valid for the functional  $h$  in this example. Again, the condition  $m = o(n)$  is *not* necessary.

In the following sections we prove the consistency results for the Hausdorff distance when applying our semi-parametric bootstrapping scheme on uniform polytopes. In this case, the underlying distribution  $F$  is the uniform distribution on an unknown convex set  $K$ , which implies that  $F_{ML(S_n)}$  is exactly the uniform distribution on  $ML(S_n) = K_n = \text{conv}(S_n)$ , and the bootstrapping points  $S_{n,m}^* = \{X_{n,i}^*\}_{i=1}^m$  are drawn from  $F_{ML(S_n)}$ .

### 5. The bootstrap consistency of the Hausdorff distance on uniform polytopes with a semi-parametric bootstrapping scheme

#### 5.1. Notation

Here and later the following notation is used. Suppose  $F$  is the uniform distribution on an unknown convex set  $K$ . Let  $S_n := \{X_i\}_{i=1}^n$  be a sequence of i.i.d. random points drawn from  $F$  and  $K_n = \text{conv}(S_n)$ . Let  $S_{n,m}^* := \{X_{n,i}^*\}_{i=1}^m$  be a sequence of i.i.d. random points drawn from  $F_{K_n}$ . Let  $S'_m := \{X'_i\}_{i=1}^m$  be another sequence of i.i.d. random points drawn from  $F$  and independent from  $S_n$ . For any set  $A$ , we denote  $A^n = \underbrace{A \times \dots \times A}_n$ . For simplicity, we also use

the following notation:

$$K'_m = \text{conv}(S'_m),$$

$$K_{n,m}^* = \text{conv}(S_{n,m}^*).$$

The following convergences are equivalent:

$$\begin{aligned} m(1 - F(K_n)) &\xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty, \\ m \log(F(K_n)) &\xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty. \end{aligned} \tag{5.1}$$

Note that  $P(m(1 - F(K_n)) > m(1 - \exp\{-\varepsilon/m\})) \rightarrow 0$  if and only if  $P(|m \log(F(K_n))| > \varepsilon) \rightarrow 0$ , where  $\varepsilon > 0$  and  $m(1 - \exp(-\varepsilon/m)) \rightarrow \varepsilon$  as  $m \rightarrow \infty$ .

### 5.2. The problem

For any functional  $\phi$ , suppose we have

$$\frac{\phi(S_n; F) - \beta_n}{\alpha_n} \xrightarrow{D} Z, \tag{5.2}$$

where  $\alpha_n$  and  $\beta_n$  are real numbers and  $Z$  is a random variable. Since  $F$  is the uniform distribution on  $K$ ,  $F_{K_n}$  is the uniform distribution on  $K_n$ . Then we will investigate the following two types of convergence:

$$\begin{aligned} \text{Type I: } & \mathbb{P}\left(\frac{\phi(S_{n,m}^*; F) - \beta_m}{\alpha_m} \leq x \mid S_n\right) \xrightarrow{D} \mathbb{P}(Z \leq x) \quad \text{as } n, m \rightarrow \infty. \\ \text{Type II: } & \mathbb{P}\left(\frac{\phi(S_{n,m}^*; F_{K_n}) - \beta_m}{\alpha_m} \leq x \mid S_n\right) \xrightarrow{D} \mathbb{P}(Z \leq x) \quad \text{as } n, m \rightarrow \infty. \end{aligned}$$

Type I convergence is used for derivation of type II convergence and also for theoretical interest. Of course, type II convergence establishes that the suggested bootstrap scheme works for the given functional  $\phi$ .

According to (2.6) and (2.9), the convergence in (5.2) holds if we let

$$\phi(S_n; F) = H(K, K_n).$$

In this case, in the next section we will establish both type I convergence where

$$\phi(S_{n,m}^*; F) = H(K, K_{n,m}^*),$$

and type II convergence where

$$\phi(S_{n,m}^*; F_{K_n}) = H(K_n, K_{n,m}^*).$$

### 5.3. Uniform distribution on convex polygons

Suppose  $F$  is the uniform distribution on  $K \subset \mathbb{R}^2$ , where  $K$  is a convex polygon with  $r$  vertices  $c_1, c_2, \dots, c_r$  and interior angles  $\theta_1, \theta_2, \dots, \theta_r$ . From (2.6), we have

$$\sqrt{n}H(K, K_n) \xrightarrow{D} Z_1, \quad \text{as } n \rightarrow \infty,$$

where  $Z_1$  has the distribution function  $G_1$  as in (2.6). Based on this result, we get the following theorem.

**Theorem 5.1.** *Let  $m \log n/n \rightarrow 0$  as  $n \rightarrow \infty$ . Then, as  $n, m \rightarrow \infty$ ,*

$$\mathbb{P}(\sqrt{m}H(K, K_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} G_1(x).$$

*Proof.* Define the event

$$E_{n,m} \triangleq \{S'_m \subset K_n\}.$$

First, we show that, as  $n, m \rightarrow \infty$ ,

$$\mathbb{P}(E_{n,m} \mid S_n) \xrightarrow{P} 1. \tag{5.3}$$

This is equivalent to

$$\log (\mathbb{P}(E_{n,m} \mid S_n)) \xrightarrow{\mathbb{P}} 0.$$

Notice that  $\mathbb{P}(E_{n,m} \mid S_n) = (|K_n|/|K|)^m$ . Then we have

$$\log (\mathbb{P}(E_{n,m} \mid S_n)) = m \log (|K_n|/|K|).$$

Let  $|D_n| := |K \setminus K_n| = |K| - |K_n|$ . Using the equivalence in (5.1), we only need to show that

$$m|D_n|/|K| \xrightarrow{\mathbb{P}} 0.$$

Notice that (2.1) shows that  $(|D_n|/|K| - \beta_n)/\alpha_n$  converges weakly where  $\beta_n = \frac{2}{3n}r \log n$  and  $\alpha_n = \frac{1}{n}\sqrt{\frac{28}{27}r \log n}$ , and also notice that as  $n, m \rightarrow \infty$ ,  $m\alpha_n \rightarrow 0$  and  $m\beta_n \rightarrow 0$  since  $m \log n/n \rightarrow 0$ . Then, as  $n, m \rightarrow \infty$  we have

$$\begin{aligned} m|D_n|/|K| &= m\left(\alpha_n \frac{|D_n|/|K| - \beta_n}{\alpha_n} + \beta_n\right) \\ &= m\alpha_n \left(\frac{|D_n|/|K| - \beta_n}{\alpha_n}\right) + m\beta_n \xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Now we develop our proof for Theorem 5.1. Define the function  $f_n : \mathbb{R}^n \rightarrow \mathbb{R}$  such that

$$f_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sqrt{n}H(K, \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_n\}).$$

From (2.6), we get, as  $n \rightarrow \infty$ ,

$$f_n(X_1, \dots, X_n) \xrightarrow{\mathbb{D}} Z_1,$$

i.e. for any continuity point of  $G_1$ , say  $x \in \mathbb{R}$ , as  $n \rightarrow \infty$ ,

$$\mathbb{P}(f_n(X_1, \dots, X_n) \leq x) \rightarrow G_1(x). \tag{5.4}$$

Let  $A_n := (f_n^{-1}(-\infty, x]) \cap K^n$ , where  $f_n^{-1}(A) = \{u : f_n(u) \in A\}$  for any subset  $A$  of the range of function  $f_n$ . Since, given  $S_n$ , the random vector  $(X_{n,1}^*, \dots, X_{n,m}^*)$  is independent of  $E_{n,m}$ , we have

$$\begin{aligned} \mathbb{P}((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m \mid S_n) &= \mathbb{P}((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m \mid S_n, E_{n,m}) \\ &= \mathbb{P}((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m \cap K_n^m \mid S_n, E_{n,m}) \\ &= \mathbb{P}((X'_1, \dots, X'_m) \in A_m \cap K_n^m \mid S_n, E_{n,m}). \end{aligned} \tag{5.5}$$

The last equation holds because the joint distribution of  $(X'_1, \dots, X'_m)$  is the same as that of  $(X_{n,1}^*, \dots, X_{n,m}^*)$  when its distribution is restricted on  $K_n$ . Applying the multiplication rule to the probability condition on  $S_n$  (Lemma B.1), we have

$$\begin{aligned} \mathbb{P}((X'_1, \dots, X'_m) \in A_m \cap K_n^m \mid S_n, E_{n,m}) &\mathbb{P}(E_{n,m} \mid S_n) \\ &= \mathbb{P}(\{(X'_1, \dots, X'_m) \in A_m \cap K_n^m\} \cap E_{n,m} \mid S_n) \\ &= \mathbb{P}((X'_1, \dots, X'_m) \in A_m \cap K_n^m \mid S_n). \end{aligned} \tag{5.6}$$

From (5.3), we have  $P(E_{n,m} \mid S_n) \xrightarrow{P} 1$  as  $n \rightarrow \infty$ . Therefore, (5.6) implies

$$|P((X'_1, \dots, X'_m) \in A_m \cap K_n^m \mid S_n, E_{n,m}) - P((X'_1, \dots, X'_m) \in A_m \cap K_n^m \mid S_n)| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty. \tag{5.7}$$

Since  $(A_m \cap K_n^m) \subset A_m$  and  $A_m \setminus (A_m \cap K_n^m) \subset D_n^m$ , we have

$$\begin{aligned} &|P((X'_1, \dots, X'_m) \in A_m \cap K_n^m \mid S_n) - P((X'_1, \dots, X'_m) \in A_m \mid S_n)| \\ &\leq P((X'_1, \dots, X'_m) \in D_n^m \mid S_n) \\ &= (|D_n|/|K|)^m \\ &\xrightarrow{\text{a.s.}} 0 \quad \text{as } n, m \rightarrow \infty. \end{aligned} \tag{5.8}$$

(Note that for almost all  $\omega \in \Omega$ ,  $|D_n(\omega)|/|K| < |D_3(\omega)|/|K| < 1$  implies  $(|D_n(\omega)|/|K|)^m \rightarrow 0$ , as  $n, m \rightarrow \infty$ .) Thus, combining (5.5), (5.7), and (5.8) implies that

$$|P((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m \mid S_n) - P((X'_1, \dots, X'_m) \in A_m \mid S_n)| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty.$$

This is equivalent to

$$|P(f_m(X_{n,1}^*, \dots, X_{n,m}^*) \leq x \mid S_n) - P(f_m(X'_1, \dots, X'_m) \leq x \mid S_n)| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty. \tag{5.9}$$

Since  $f_m(X'_1, \dots, X'_m)$  is independent of  $S_n$ , using (5.4) we have

$$P(f_m(X'_1, \dots, X'_m) \leq x \mid S_n) = P(f_m(X'_1, \dots, X'_m) \leq x) \rightarrow G_1(x) \quad \text{as } n, m \rightarrow \infty. \tag{5.10}$$

Finally, (5.9) and (5.10) imply that

$$P(f_m(X_{n,1}^*, \dots, X_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} G_1(x) \quad \text{as } n, m \rightarrow \infty. \quad \square$$

In the following theorem, we extend Theorem 5.1, which is of type I convergence, to type II convergence.

**Theorem 5.2.** *Under the same conditions as Theorem 5.1, we have*

$$P(\sqrt{m}H(K_n, K_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} G_1(x) \quad \text{as } n, m \rightarrow \infty.$$

*Proof.* In the proof of Theorem 5.1, given  $S_n = \{X_i\}_{i=1}^n$ , we reset  $A_m = (g_m^{-1}(-\infty, x]) \cap K^m$ , where

$$g_m(\mathbf{x}_1, \dots, \mathbf{x}_m) = \sqrt{m}H(K_n, \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_n\}).$$

Let  $x$  be a continuity point of  $G_1$ . Then follow the steps of the proof of Theorem 5.1 to get

$$|P((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m \mid S_n) - P((X'_1, \dots, X'_m) \in A_m \mid S_n)| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty.$$

This is equivalent to

$$|\mathbb{P}(g_m(X_{n,1}^*, \dots, X_{n,m}^*) \leq x \mid S_n) - \mathbb{P}(g_m(X'_1, \dots, X'_m) \leq x \mid S_n)| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty.$$

So,

$$|\mathbb{P}(\sqrt{m}H(K_n, K_{n,m}^*) \leq x \mid S_n) - \mathbb{P}(\sqrt{m}H(K_n, K'_m) \leq x \mid S_n)| \xrightarrow{P} 0 \tag{5.11}$$

as  $n, m \rightarrow \infty$ . From Theorem 5.1, we have

$$\mathbb{P}(\sqrt{m}H(K_n, K_{n,m}^*) \leq x \mid S_n) \geq \mathbb{P}(\sqrt{m}H(K, K_{n,m}^*) \leq x \mid S_n) \tag{5.12}$$

and

$$\mathbb{P}(\sqrt{m}H(K, K_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} G_1(x) \tag{5.13}$$

as  $n, m \rightarrow \infty$ . Notice that  $H(K, K_{n,m}^*) \geq H(K_n, K_{n,m}^*)$ . Suppose  $G_{1,m}$  is the distribution function of  $f_m(X'_1, \dots, X'_m)$ , where  $f_m$  is the same as in the proof of Theorem 5.1. Then  $\lim_{m \rightarrow \infty} G_{1,m}(x) = G_1(x)$ . Since the Hausdorff distance satisfies the triangle inequality

$$H(K_n, K'_m) \geq H(K, K'_m) - H(K, K_n),$$

we have

$$\begin{aligned} \mathbb{P}(\sqrt{m}H(K_n, K'_m) \leq x \mid S_n) &\leq \mathbb{P}(\sqrt{m}(H(K, K'_m) - H(K, K_n)) \leq x \mid S_n) \\ &= \mathbb{P}(\sqrt{m}H(K, K'_m) \leq \sqrt{m}H(K, K_n) + x \mid S_n) \\ &= G_{1,m}\left(\sqrt{\frac{m}{n}}(\sqrt{n}H(K, K_n)) + x\right) \\ &\xrightarrow{P} G_1(x) \quad \text{as } n, m \rightarrow \infty, \end{aligned} \tag{5.14}$$

where the last convergence follows from (2.6) and applying Skorokhod’s representation theorem. Then use Lemma B.4 with (5.11), (5.12), (5.13), and (5.14) to get

$$\mathbb{P}(\sqrt{m}H(K_n, K_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} G_1(x) \quad \text{as } n, m \rightarrow \infty. \quad \square$$

### 5.4. Uniform distribution on convex sets with a smooth boundary

A similar technique can also be applied to the smooth case. Let  $K \subset \mathbb{R}^2$  be a convex set with  $\partial K \subset \mathcal{C}^2(\mathbb{R})$ . Suppose the curvature of  $\partial K$  is bounded away from 0 and  $\infty$  and has a bounded derivative. Suppose our underlying distribution  $F$  is the uniform distribution on  $K$ . Also suppose  $F$  satisfies (2.8). From (2.9), we get

$$\frac{H(K, K_n) - b_n}{a_n} \xrightarrow{D} Z_2 \quad \text{as } n \rightarrow \infty,$$

where  $Z_2$  has the distribution  $G_2$ . Based on this result, we have the following theorem.

**Theorem 5.3.** *Let  $m/n^{2/3} \rightarrow 0$  as  $n \rightarrow \infty$ . Then we have*

$$\mathbb{P}\left(\frac{H(K, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \xrightarrow{P} G_2(x) \quad \text{as } n, m \rightarrow \infty,$$

where  $G_2(x) = \exp\{-d_1 e^{-d_2 x}\}$ , and  $a_m, b_m$ , and  $G_2$  are all defined in (2.9).

*Proof.* The proof follows the same lines as in the case of Theorem 5.1. Define the events  $E_{n,m} \triangleq \{S'_m \subset K_n\}$ . First we need to show that, as  $n, m \rightarrow \infty$ ,

$$P(E_{n,m} \mid S_n) \xrightarrow{P} 1. \tag{5.15}$$

Suppose  $D_n = K \setminus K_n$ . Since  $P(E_{n,m} \mid S_n) = (|K_n|/|K|)^m$ , when taking logs of both sides and using the equivalence in (5.1), (5.15) becomes

$$m|D_n| \xrightarrow{P} 0.$$

Notice that

$$\begin{aligned} m|D_n| &= m(|K| - |K_n|) \\ &= (mn^{-2/3})(n^{2/3}E(|D_n|)) - (mn^{-5/6})(n^{5/6}(|K_n| - E(|K_n|))), \end{aligned}$$

where  $mn^{-2/3} \rightarrow 0$  and  $mn^{-5/6} \rightarrow 0$  according to our assumption about  $m$ . Finally, using (2.3) and (2.4), we complete the proof for  $P(E_{n,m} \mid S_n) \xrightarrow{P} 1$ .

Now, define the function  $f_n : \mathbb{R}^n \rightarrow \mathbb{R}$  such that

$$f_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = \frac{H(K, \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) - b_n}{a_n}.$$

We only need to follow the steps in the proof of Theorem 5.1 and use (2.9). Therefore, for any continuity point  $x$  of  $G_2$ , we get

$$P(f_m(X_{n,1}^*, \dots, X_{n,m}^*) \leq x \mid S_n) \xrightarrow{P} G_2(x) \quad \text{as } n, m \rightarrow \infty. \quad \square$$

Again, in the following theorem we extend Theorem 5.3 (which is of type I convergence) to type II convergence.

**Theorem 5.4.** *Under the same conditions as Theorem 5.3, we have*

$$P\left(\frac{H(K_n, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \xrightarrow{P} G_2(x) \quad \text{as } n, m \rightarrow \infty.$$

*Proof.* Here we use the same technique as for Theorem 5.1. Given  $S_n = \{X_i\}_{i=1}^n$ , we reset  $A_m = (g_m^{-1}(-\infty, x]) \cap K^m$ , where

$$g_m(\mathbf{x}_1, \dots, \mathbf{x}_m) = \frac{H(K_n, \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_m\}) - b_m}{a_m}.$$

Let  $x$  be a continuity point of  $G_2$ . Then follow the steps in the proof of Theorem 5.1 to get

$$\begin{aligned} &|P((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m \mid S_n) - P((X'_1, \dots, X'_m) \in A_m \mid S_n)| \\ &\xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty. \end{aligned}$$

This is equivalent to

$$\begin{aligned} &|P(g_m(X_{n,1}^*, \dots, X_{n,m}^*) \leq x \mid S_n) - P(g_m(X'_1, \dots, X'_m) \leq x \mid S_n)| \\ &\xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty. \end{aligned}$$



Therefore, we have

$$\left| \mathbb{P}\left(\frac{H(K_n, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) - \mathbb{P}\left(\frac{H(K_n, K'_m) - b_m}{a_m} \leq x \mid S_n\right) \right| \xrightarrow{\mathbb{P}} 0 \tag{5.16}$$

as  $n, m \rightarrow \infty$ . From Theorem 5.3, we have

$$\mathbb{P}\left(\frac{H(K_n, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \geq \mathbb{P}\left(\frac{H(K, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \tag{5.17}$$

and

$$\mathbb{P}\left(\frac{H(K, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \xrightarrow{\mathbb{P}} G_2(x) \tag{5.18}$$

as  $n, m \rightarrow \infty$ . Suppose  $G_{2,m}$  is the distribution function of  $f_m(X'_1, \dots, X'_m)$ , where  $f_m$  is the same as in the proof of Theorem 5.3. Then  $\lim_{m \rightarrow \infty} G_{2,m}(x) = G_2(x)$ . Since the Hausdorff distance satisfies the triangle inequality

$$H(K_n, K'_m) \geq H(K, K'_m) - H(K, K_n),$$

we have

$$\begin{aligned} \mathbb{P}\left(\frac{H(K_n, K'_m) - b_m}{a_m} \leq x \mid S_n\right) &\leq \mathbb{P}\left(\frac{H(K, K'_m) - H(K, K_n) - b_m}{a_m} \leq x \mid S_n\right) \\ &= \mathbb{P}\left(\frac{H(K, K'_m) - b_m}{a_m} \leq \frac{H(K, K_n)}{a_m} + x \mid S_n\right) \\ &= G_{2,m}\left(\frac{H(K, K_n)}{a_m} + x\right). \end{aligned} \tag{5.19}$$

Since  $a_n/a_m \rightarrow 0$  and  $b_n/a_m \rightarrow 0$ , we get

$$\frac{H(K, K_n)}{a_m} = \frac{a_n}{a_m} \frac{H(K, K_n) - b_n}{a_n} + \frac{b_n}{a_m} \xrightarrow{\mathbb{D}} 0 \quad \text{as } n, m \rightarrow \infty. \tag{5.20}$$

Therefore, according to (5.20) and Skorokhod’s representation theorem, we have

$$G_{2,m}\left(\frac{H(K, K_n)}{a_m} + x\right) \xrightarrow{\mathbb{P}} G_2(x) \quad \text{as } n, m \rightarrow \infty. \tag{5.21}$$

Finally, use Lemma B.4 and (5.16), (5.17), (5.18), (5.19), and (5.21) to get

$$\mathbb{P}\left(\frac{H(K_n, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \xrightarrow{\mathbb{P}} G_2(x) \quad \text{as } n, m \rightarrow \infty. \quad \square$$

**Remark 5.1.** We can also get another bonus convergence result:

$$\frac{H(K_n, K'_m) - b_m}{a_m} \xrightarrow{\mathbb{D}} Z_2 \quad \text{as } m \rightarrow \infty.$$

(This result is only a by-product of our proof and has nothing to do with our bootstrapping scheme.) By Lemma B.3 and (2.9), it is sufficient to prove

$$\frac{H(K, K_n)}{a_m} \xrightarrow{\mathbb{D}} 0 \quad \text{as } m \rightarrow \infty,$$

which is similar to (5.20).

### 6. Two more examples

In addition to the Hausdorff distance, the same technique can also be applied to other functionals. Two noticeable examples are the symmetric difference and the perimeter difference.

The following example is the symmetric difference on  $\mathbb{R}^d$ . Suppose  $d (\geq 2)$  is a fixed integer and  $K$  is a convex set in  $\mathbb{R}^d$  such that  $\partial K \in \mathcal{C}^2(\mathbb{R}^{d-1})$  with positive Gaussian curvature (see Appendix A for details) bounded away from 0 and  $\infty$ . With these assumptions, we can use the results in (2.5) and (2.2).

**Example 6.1.** Let  $\phi_1(A, B)$  be the area of the symmetric difference between  $A$  and  $B$ , i.e.  $\phi_1(A, B) = |(A \setminus B) \cup (B \setminus A)|$ . Since  $K_n \subset K$ , we have  $\phi_1(K, K_n) = |D_n|$ . From (2.5) we have

$$C_n^{-1/2} n^{\frac{1}{2} + \frac{1}{d+1}} (|D_n| - E(|D_n|)) \xrightarrow{D} \mathcal{N}(0, \sigma^2) \quad \text{as } n \rightarrow \infty, \tag{6.1}$$

and (2.2) gives

$$\lim_{n \rightarrow \infty} n^{\frac{2}{d+1}} E(|D_n|) = c,$$

where  $\sigma$  and  $c$  are constants. Rewrite (6.1) using  $\phi_1$  to get

$$\frac{\phi_1(K, K_n) - b_n}{a_n} \xrightarrow{D} \mathcal{N}(0, 1) \quad \text{as } n \rightarrow \infty,$$

where  $a_n = \sigma C_n^{1/2} n^{-\frac{1}{2} - \frac{1}{d+1}}$  and  $b_n = E(|D_n|) = \Theta(n^{-\frac{2}{d+1}})$ . Let  $\Phi$  be the cumulative distribution function for the standard normal distribution. Our goal is to find a proper assumption about  $m$  such that the type I convergence holds for  $\phi_1$ , i.e., as  $n, m \rightarrow \infty$  we have

$$P\left(\frac{\phi_1(K, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \xrightarrow{P} \Phi(x). \tag{6.2}$$

As before, define the events  $E_{n,m} \triangleq \{S'_m \subset K_n\}$  for the positive integers  $m$  and  $n$ . To impose a proper assumption on  $m$ , we must ensure that

$$P(E_{n,m} \mid S_n) = P^m(|K_n|/|K|) \xrightarrow{P} 1 \quad \text{as } n, m \rightarrow \infty.$$

Taking logs of both sides and using the equivalence in (5.1), we have

$$m|D_n| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty.$$

Since

$$\begin{aligned} m|D_n| &= m\phi_1(K, K_n) \\ &= m \left[ a_n \frac{\phi_1(K, K_n) - b_n}{a_n} + b_n \right] \\ &= ma_n \left[ \frac{\phi_1(K, K_n) - b_n}{a_n} \right] + mb_n, \end{aligned}$$

the assumption required for  $m$  is

$$ma_n \rightarrow 0, \quad mb_n \rightarrow 0$$

as  $n, m \rightarrow \infty$ . So, the resample size  $m$  should satisfy  $m = o(n^{\frac{2}{d+1}})$ . We omit the details for getting type I convergence in (6.2) as the proof mimics those for Theorems 5.1 and 5.3.

Furthermore, to prove type II convergence for  $\phi_1$ , i.e.

$$P\left(\frac{\phi_1(K_n, K_{n,m}^*) - b_m}{a_m} \leq x \mid S_n\right) \xrightarrow{D} \Phi(x) \quad \text{as } n, m \rightarrow \infty,$$

notice that

$$\phi_1(K, K_{n,m}^*) \geq \phi_1(K_n, K_{n,m}^*)$$

and

$$\phi_1(K_n, K'_m) \geq \phi_1(K, K'_m) - \phi_1(K, K_n),$$

which are easy to verify. On the other hand, observe that

$$a_n/a_m \rightarrow 0, \quad b_n/a_m \rightarrow 0$$

as  $n, m \rightarrow \infty$ .

Now we can consider the perimeter difference (defined below) on  $\mathbb{R}^2$ . To use the result in (2.3), suppose  $K$  satisfies  $\partial K \in C^2(\mathbb{R})$  with positive curvature bounded away from 0 and  $\infty$ .

**Example 6.2.** Let  $d = 2$  and define  $\phi_2(A, B) = |\text{perimeter of } A - \text{perimeter of } B|$ . From (2.3) we have

$$n^{5/6}(L_n - E(L_n)) \xrightarrow{D} \mathcal{N}(0, (\sigma')^2) \quad \text{as } n \rightarrow \infty \tag{6.8}$$

and

$$\lim_{n \rightarrow \infty} n^{2/3}E(L - L_n) = c',$$

where  $\sigma'$  and  $c'$  are two constants. Rewriting (6.8) in terms of  $\phi_2$ , we have

$$\frac{\phi_2(K, K_n) - b'_n}{a'_n} \xrightarrow{D} \mathcal{N}(0, 1) \quad \text{as } n \rightarrow \infty,$$

where  $a'_n = \sigma' n^{-5/6}$  and  $b'_n = E(L - L_n) = \Theta(n^{-2/3})$ . To prove type I convergence, i.e.

$$P\left(\frac{\phi_2(K, K_{n,m}^*) - b'_m}{a'_m} \leq x \mid S_n\right) \xrightarrow{P} \Phi(x) \quad \text{as } n, m \rightarrow \infty,$$

we need to show that

$$P(E_{n,m} \mid S_n) = P^m(|K_n|/|K|) \xrightarrow{P} 1 \quad \text{as } n, m \rightarrow \infty.$$

As in Example 6.1,  $m$  should satisfy the condition

$$m|D_n| \xrightarrow{P} 0 \quad \text{as } n, m \rightarrow \infty.$$

Thus, our assumption for  $m$  should follow  $m = o(n^{2/3})$ . (Notice that  $d = 2$ .)

To prove type II convergence, i.e.

$$P\left(\frac{\phi_2(K_n, K_{n,m}^*) - b'_m}{a'_m} \leq x \mid S_n\right) \xrightarrow{P} \Phi(x) \quad \text{as } n, m \rightarrow \infty,$$

notice that

$$\phi_2(K, K_{n,m}^*) \geq \phi_2(K_n, K_{n,m}^*)$$

and

$$\phi_2(K_n, K'_m) \geq \phi_2(K, K'_m) - \phi_2(K, K_n),$$

which are easy to verify using Lemma B.2. Obviously,

$$a'_n/a'_m \rightarrow 0, \quad b'_n/a'_m \rightarrow 0$$

as  $n, m \rightarrow \infty$ .

### Appendix A. Basic definitions and notation

#### A.1. Curvature and Gaussian curvature

Let the planar curve  $c(t) = (x(t), y(t))$ , where  $x, y \in C^2(\mathbb{R})$  and  $t$  is the arc length between  $c(0)$  and  $c(t)$ . Then the curvature  $\kappa(t) = |\ddot{c}(t)|$ .

Let  $S$  be a surface in Euclidean space with second fundamental form  $II$  (see Kobayashi and Nomizu (1996)). Given a point  $z$  on  $S$ , suppose  $\{v_1, v_2\}$  is an orthonormal basis of tangent vectors at  $z$ . Suppose  $\kappa_1$  and  $\kappa_2$  are the eigenvalues of the following matrix:

$$\begin{bmatrix} II(v_1, v_1) & II(v_1, v_2) \\ II(v_2, v_1) & II(v_2, v_2) \end{bmatrix}.$$

Then the Gaussian curvature at the point  $z$  is given by  $\kappa = \kappa_1\kappa_2$ .

#### Appendix B. Some lemmas

The following multiplication rule is obvious, but we mention it for completeness.

**Lemma B.1.** *For any events  $A, B$ , and  $C$  we have*

$$P(A \mid B \cap C)P(C \mid B) = P(A \cap C \mid B).$$

**Lemma B.2.** *If  $P \subset K \subset \mathbb{R}^2$  such that  $P$  is a convex polygon and  $K$  is a convex set, we have*

$$L(P) \leq L(K),$$

where  $L(P)$  and  $L(K)$  are the perimeters of  $P$  and  $K$ , respectively.

*Proof.* We refer the reader to Figure 5 in the proof of this lemma. Suppose  $P$  has  $r$  vertices. Let  $e_1, e_2, \dots, e_r$  be the edges of  $P$ . It is enough to show that disjoint segments  $s_1, s_2, \dots, s_r$  on the boundary of  $K$  exist such that

$$e_i \leq s_i, \quad i = 1, 2, \dots, r.$$

Consider the segments  $s_1, s_2, \dots, s_r$  in Figure 5. Since the length of  $e_i$ , for any  $1 \leq i \leq r$ , is the distance between the two parallel lines, the length of  $s_i$  is not shorter than the length of  $e_i$ . Hence,

$$L(P) = \sum_{i=1}^r e_i \leq \sum_{i=1}^r s_i \leq L(K). \quad \square$$

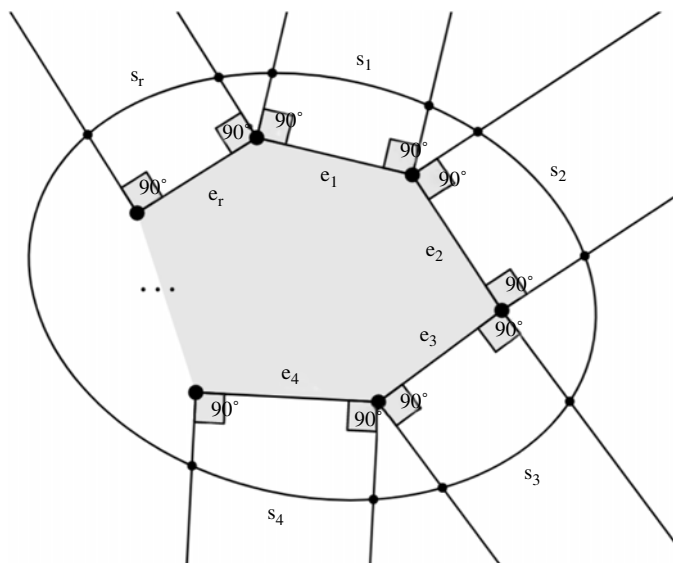


FIGURE 5: The edges  $e_1, e_2, \dots, e_r$  and the segments  $s_1, s_2, \dots, s_r$ .

**Lemma B.3.** (Billingsley (1999), Theorem 3.1, page 27) Suppose  $d$  is the metric on  $S$ . If  $(X_n, Y_n)$  are random elements of  $S \times S$  such that  $X_n \xrightarrow{D} X_0$  and  $d(X_n, Y_n) \xrightarrow{D} 0$ , then  $Y_n \xrightarrow{D} X_0$ .

**Lemma B.4.** If  $|A_n - B_n| \xrightarrow{P} 0$ ,  $A_n \geq C_n$ ,  $C_n \xrightarrow{P} Z$ ,  $B_n \leq D_n$ , and  $D_n \xrightarrow{P} Z$ , then we have  $A_n \xrightarrow{P} Z$  and  $B_n \xrightarrow{P} Z$ .

*Proof.* For any given  $\varepsilon > 0$ , as  $n \rightarrow \infty$ ,

$$\begin{aligned} P(A_n - Z > \varepsilon) &= P(A_n - B_n + B_n - D_n + D_n - Z > \varepsilon) \\ &\leq P(A_n - B_n > \varepsilon/3) + P(B_n - D_n > \varepsilon/3) + P(D_n - Z > \varepsilon/3) \rightarrow 0; \\ P(Z - A_n > \varepsilon) &= P(Z - C_n + C_n - A_n) \\ &\leq P(Z - C_n > \varepsilon/2) + P(C_n - A_n > \varepsilon/2) \rightarrow 0. \end{aligned}$$

□

### Acknowledgements

The authors are greatly indebted to the referees and the associate editors for several corrections and helpful suggestions. Financial support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-04008) is gratefully acknowledged.

### References

BARNETT, V. (1976). The ordering of multivariate data. *J. R. Statist. Soc. A* **139**, 318–355.  
 BICKEL, P. J. AND FREEDMAN, D. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9**, 1196–1217.  
 BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, Chichester.  
 BRÄKER, H., HSING, T. AND BINGHAM, N. H. (1998). On the Hausdorff distance between a convex set and an interior random convex hull. *Adv. Appl. Prob.* **30**, 295–316.  
 BRÄKER, H. AND HSING, T. (1998). On the area and perimeter of a random convex hull in a bounded convex set. *Prob. Theory Relat. Fields* **111**, 517–550.

- CABO, A. J. AND GROENEBOOM, P. (1994). Limit theorems for functionals of convex hulls. *Prob. Theory Relat. Fields* **100**, 31–55.
- EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7**, 1–26.
- HSU, P. L. AND ROBBINS, H. (1974). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA* **33**, 25–31.
- KALLENBERG, O. (1983). *Random Measures*, 3rd edn. Akademie, Berlin.
- KOBAYASHI, S. AND NOMIZU, K. (1996). *Foundations of Differential Geometry*. Wiley-Interscience, New York.
- LIU, R. Y., PARELIUS, J. M. AND SINGH, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion and a rejoinder by Liu and Singh). *Ann. Statist.* **3**, 783–858.
- MACDONALD, D. W., BALL, F. G. AND HOUGH, N. G. (1980). The evaluation of home range size and configuration using radio tracking data. In *A Handbook on Biotelemetry and Radio Tracking*, eds D. W. MacDonald and C. J. Amlaner, Pergamon Press, Oxford.
- REITZNER, M. (2005). Central limit theorems for random polytopes. *Prob. Theory Relat. Fields* **133**, 483–507.
- RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- RIPLEY, B. D. AND RASSON, J. P. (1977). Finding the edge of a Poisson forest. *J. Appl. Prob.* **14**, 483–491.
- SCHÜTT, C. (1994). Random polytopes and affine surface area. *Math. Nachr.* **170**, 227–249.
- TUKEY, J. W. (1974). Mathematics and the picturing of data. In *Proc. Int. Cong. Math.*, Vol. 2, pp. 523–532.
- ZAREPOUR, M. (1999). Bootstrapping convex hulls. *Statist. Prob. Lett.* **45**, 55–63.