



Surface and length estimation based on Crofton's formula

Catherine Aaron, Alejandro Cholaquidis, Ricardo Fraiman

► To cite this version:

Catherine Aaron, Alejandro Cholaquidis, Ricardo Fraiman. Surface and length estimation based on Crofton's formula. 2021. hal-02907297v2

HAL Id: hal-02907297

<https://hal.science/hal-02907297v2>

Preprint submitted on 2 Feb 2021 (v2), last revised 24 Mar 2022 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Surface and length estimation based on Crofton's formula

Catherine Aaron, Alejandro Cholaquidis and Ricardo Fraiman

February 2, 2021

Abstract

We study the problem of estimating the surface area of the boundary of a sufficiently smooth set when the available information is only a set of points (random or not) that becomes dense (with respect to Hausdorff distance) in the set or the trajectory of a reflected diffusion. We obtain consistency results in this general setup, and we derive rates of convergence for the iid case or when the data corresponds to the trajectory of a reflected Brownian motion. We propose an algorithm based on Crofton's formula, which estimates the number of intersections of random lines with the boundary of the set by counting, in a suitable way (given by the proposed algorithm), the number of intersections with the boundary of two different estimators: the Devroye-Wise estimator and the α -convex hull of the data.

Key Words : Crofton's formula; Surface estimation; α -convex hull; Devroye-Wise estimator

1 Introduction

Let $S \subset \mathbb{R}^d$ be a compact set, we aim to estimate its surface area, i.e. the $(d - 1)$ - Hausdorff measure of its boundary ∂S . Surface area estimation has been extensively considered in stereology (see Baddeley, Gundersen and Cruz-Orive (1986); Baddeley and Jensen (2005), Gokhale (1990)). It has also been studied as a further step in the theory of nonparametric set estimation (see Pateiro-López and Rodríguez-Casal (2008)), and has practical applications in medical imaging (see Cuevas, Fraiman, and Rodríguez-Casal (2007)). Although the 2-dimensional case has many significant applications, this is also the case where $d = 3$, since surface area is an important biological parameter, in organs such as the lungs. Also surface area estimation is widely used in magnetic resonance imaging (MRI) techniques. From a theoretically point of view, in Penrose (2021), the surface area of the boundary plays a significant role as a parameter of a probability distribution, being able to estimate it allows to apply plug-in methods.

When, as in image analysis, one can observe data points from two distinguishable sets of random data-points (one from inside S and the other one from outside S), the problem of the estimation of the surface area of the boundary has been considered, for any $d \geq 2$ in Cuevas, Fraiman, and Rodríguez-Casal (2007), Pateiro-López and Rodríguez-Casal (2008), Jiménez and Yukich (2011), Cuevas, Fraiman and Györfi (2013) and Thäle and Yukich (2016).

The three- and two-dimensional cases are addressed in Berrendero et al. (2014), where the authors propose parametric estimators when the available data are the distances to S , from a sample outside the set, but at a distance smaller than a given $R > 0$.

We aim to propose surface area estimators, in any dimension, when the available data is only a sample in the set S . With such data points only the two dimensional case has been yet studied.

In dimension 2, under an iid setting, length estimation problem (under convexity assumptions) has been previously addressed in Bräker and Hsing (1998) using Crofton’s formula. Later on, still in dimension 2, under the r -convexity assumption, Arias-Castro And Rodríguez-Casal (2017) obtained the convergence of the α -shape’s perimeter to the perimeter of the support. Still in dimension 2 but when the data comes from reflected Brownian motion, (with and without drift) a consistency result is obtained in Theorem 4 in Cholaquidis et al. (2016). To estimate surface area in any dimension, we propose two consistent estimators that are based on Crofton’s formula.

This well-known formula, proved by Crofton in 1868 for dimension two, and extended to arbitrary dimensions (see Santaló (2004)), states that the surface of ∂S equals the integral of the number of intersections with ∂S of lines in \mathbb{R}^d (see Equations (2) and (3) for explicit Crofton formulas for $d = 2$ and $d \geq 2$, respectively).

We propose to ‘estimate’ the number of intersections with ∂S of lines, by using two different support estimators. First we consider the Devroye–Wise estimator (see Devroye and Wise (1980)), and next the α -convex hull estimator (see Rodríguez-Casal (2007)).

Considering first the Devroye–Wise based estimator, notice that the proposed estimator is not just a plug-in, because in general the number of intersections of a line with ∂S is different from the number of intersections of that line with the boundary of the Devroye–Wise estimator. When we observe $\mathcal{X} \subset S$ our Crofton-based surface estimator attains a rate proportional to $d_H(\mathcal{X}, S)^{1/2}$ (where d_H denotes the Hausdorff distance), this rates being possibly improved to $d_H(\mathcal{X}, S)$ when adding a reasonable assumption. This result can be applied to many deterministic or random situations, to obtain explicit convergence rates. We focus on two random situations: the case $\mathcal{X} = \mathcal{X}_n = \{X_1, \dots, X_n\}$ of iid drawn on S (with a density bounded from below by a positive constant), and the case of random trajectories of reflected diffusions on S . In particular, we provide convergence rates when the trajectory is the result of a reflected Brownian motion (see Cholaquidis et al. (2016, 2021)). This last setting has several applications in ethology, such as home-range estimation, where the trajectory is obtained by recording the location of an animal (or several animals) living in an area S that is called the home range (the territorial range of the animal), and X_t represent the position at time t transmitted by the instrument (see for instance Cholaquidis et al. (2016, 2021), Baíllo and Chacón (2018) and references therein). Using tracking and telemetry technology, such GPS, have allowed to collect location data for animals at an ever-increasing rate and accuracy. The most commonly cited definition of an animal’s home range goes back to Burt (1943), p. 351: “that area traversed by the individual in its normal activities of food gathering, mating and caring for young”.

To use Crofton’s formula when the support estimator is the α -convex hull of a sample \mathcal{X}_n (denoted by $C_\alpha(\mathcal{X}_n)$), we first extend the result in Cuevas, Fraiman and Pateiro-López (2012) and prove that in any dimension the surface area of the hull’s boundary, i.e. $|\partial C_\alpha(\mathcal{X}_n)|_{d-1}$, converges to $|\partial S|_{d-1}$. This result is interesting in itself, but in practice to compute $|\partial C_\alpha(\mathcal{X}_n)|_{d-1}$ is difficult, especially for dimension $d > 2$. However, by means of the Crofton formula, it can easily be estimated via Monte-Carlo method.

The rest of this paper is organized as follows. In Section 2, we introduce the notation and some well-known geometric restrictions. Section 3 aims to present Crofton’s formula, first for dimension two and then for the general case. After that, we introduce the main geometric restrictions required in one of the main theorems. Section 4 introduces the algorithms from a mathematical standpoint, and explains the heuristics behind them. The computational aspects of the algorithms are given in Section 5 and the main results are stated in Section 6, their proofs are given in the Appendix.

2 Some preliminaries

The following notation will be used throughout the paper.

Given a set $S \subset \mathbb{R}^d$, we denote by \mathring{S} , \overline{S} and ∂S the interior, closure and boundary of S , respectively, with respect to the usual topology of \mathbb{R}^d . We also write $\text{diam}(S) = \sup_{(x,y) \in S \times S} \|x - y\|$. The parallel set of S of radius ε is denoted by $B(S, \varepsilon)$, that is, $B(S, \varepsilon) = \{y \in \mathbb{R}^d : \inf_{x \in S} \|y - x\| \leq \varepsilon\}$.

If $A \subset \mathbb{R}^d$ is a Borel set, then $|A|_d$ denotes its d -dimensional Lebesgue measure (when within an integral we will use μ_{d-1}). When $A \subset \mathbb{R}^d$ is a $(d-1)$ -dimensional manifold then $|A|_{d-1}$ denotes its $(d-1)$ -Hausdorff measure.

We denote by $\mathcal{B}(x, \varepsilon)$ the closed ball in \mathbb{R}^d , of radius ε , centred at x , and $\omega_d = |\mathcal{B}_d(x, 1)|_d$. Given two compact non-empty sets $A, C \subset \mathbb{R}^d$, the *Hausdorff distance* or *Hausdorff–Pompei distance* between A and C is defined by

$$d_H(A, C) = \inf\{\varepsilon > 0 : \text{such that } A \subset B(C, \varepsilon) \text{ and } C \subset B(A, \varepsilon)\}.$$

The $(d-1)$ -dimensional sphere in \mathbb{R}^d is denoted by S^{d-1} , while the half-sphere in \mathbb{R}^d is denoted by $(S^+)^{d-1}$, i.e., $(S^+)^{d-1} = (\mathbb{R}^{d-1} \times \mathbb{R}^+) \cap S^{d-1}$. Given M a sufficiently smooth $(d-1)$ -manifold and $x \in M$, we denote by η_x the unit outward normal vector at x . The affine tangent space of M at x is denoted by $T_x M$.

Given a vector $\theta \in (S^+)^{d-1}$ and a point y , $r_{\theta, y}$ denotes the line $\{y + \lambda\theta, \lambda \in \mathbb{R}\}$. If y_1 and y_2 are two points in $r_{\theta, y}$, then $y_i = y + \lambda_i\theta$; with a slight abuse of notation, we write $y_1 < y_2$ when $\lambda_1 < \lambda_2$.

We will now recall some well-known shape restrictions in set estimation.

Definition 1. A set $S \subset \mathbb{R}^d$ is said to be α -convex, for $\alpha > 0$, if $S = C_\alpha(S)$, where

$$C_\alpha(S) = \bigcap_{\{\mathring{\mathcal{B}}(x, \alpha) : \mathring{\mathcal{B}}(x, \alpha) \cap S = \emptyset\}} \left(\mathring{\mathcal{B}}(x, \alpha)\right)^c, \quad (1)$$

is the α -convex hull of S . When S is α -convex, a natural estimator of S from a random sample \mathcal{X}_n of points (drawn from a distribution with support S), is $C_\alpha(\mathcal{X}_n)$.

Definition 2. A set $S \subset \mathbb{R}^d$ is said to satisfy the outside α -rolling condition if for each boundary point $s \in \partial S$ there exists an $x \in S^c$ such that $\mathcal{B}(x, \alpha) \cap \partial S = \{s\}$. A compact set S is said to satisfy the inside α -rolling condition if \overline{S}^c satisfies the outside α -rolling condition at all boundary points.

3 Crofton's formula

Crofton in 1868 proved the following result (see Crofton (1868)): given γ a regular plane curve (i.e. there exists a differentiable parametrization $c : [0, 1] \rightarrow \gamma \subset \mathbb{R}^2$ such that $\|c'(t)\| > 0$ for all t), then its length $|\gamma|_1$ can be computed by

$$|\gamma|_1 = \frac{1}{2} \int_{\theta=0}^{\pi} \int_{p=-\infty}^{+\infty} n_\gamma(\theta, p) dp d\theta, \quad (2)$$

$n_\gamma(\theta, p)$ being the number of intersections of γ with the line $r_{\theta^*, \theta p}$, where $\theta^* \in (S^+)^{d-1}$ is orthogonal to θ , and $dp d\theta$ is 2-dimensional Lebesgue measure, see Figure 1. This result has been generalized to \mathbb{R}^d for any $d > 2$, and also to Lie groups, see Santaló (2004).

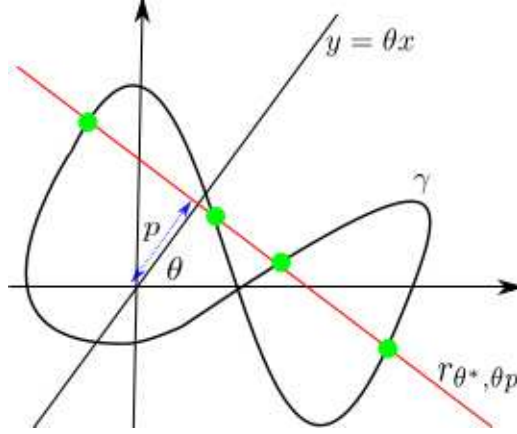


Figure 1: The function n_γ counts the number of intersections of γ with the line $r_{\theta^*, \theta p}$ determined by θ and p with the curve.

To introduce the general Crofton's formula in \mathbb{R}^d for a compact $(d-1)$ -dimensional manifold M , let us define first the constant

$$\beta(d) = \Gamma(d/2)\Gamma((d+1)/2)^{-1}\pi^{-1/2},$$

where Γ stands for the well known Gamma function. Let $\theta \in (\mathcal{S}^+)^{d-1}$, θ determine a $(d-1)$ -dimensional linear space $\theta^\perp = \{v : \langle v, \theta \rangle = 0\}$. Given $y \in \theta^\perp$, let us write $n_M(\theta, y) = \#(r_{\theta, y} \cap M)$, where $\#$ is the cardinality of the set. see Figure 2.

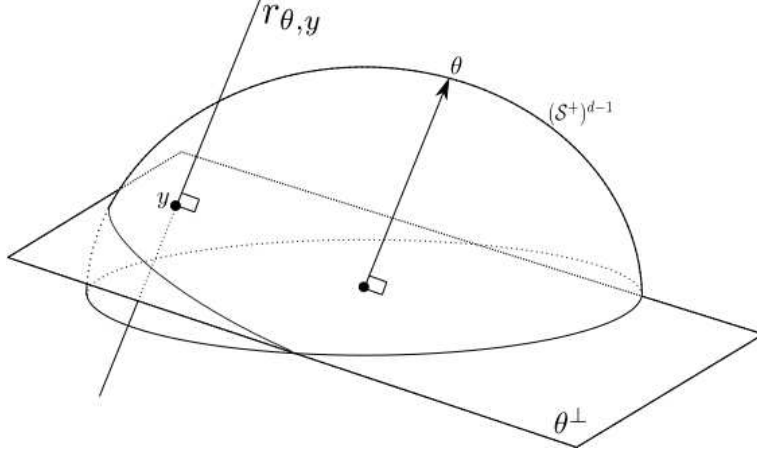


Figure 2: The line $r_{\theta, y} = y + \lambda\theta$ is shown, where $y \in \theta^\perp$ and $\theta \in (\mathcal{S}^+)^{d-1}$.

It is proved in Federer (1969) (see Theorem 3.2.26) that if M is an $(d-1)$ -dimensional rectifiable set, then the integralgeometric measure of M (which will be denote by $I_{d-1}(M)$, and is defined by the right-hand side of 3) equals its $(d-1)$ -dimensional Hausdorff measure, i.e.,

$$|M|_{d-1} = I_{d-1}(M) = \frac{1}{\beta(d)} \int_{\theta \in (\mathcal{S}^+)^{d-1}} \int_{y \in \theta^\perp} n_M(\theta, y) d\mu_{d-1}(y) d\theta. \quad (3)$$

The measure $d\theta$ is the uniform measure on $(\mathbb{S}^+)^{d-1}$ (with total mass 1).

Remark 1. *Throughout this paper we will assume that ∂S is the boundary of a compact set $S \subset \mathbb{R}^d$ such that $S = \overline{\text{int}(S)}$. We will also assume that S fulfills the outside and inside α -rolling condition and then ∂S is rectifiable (see Theorem 1 in Walther, G. (1999)). From this it follows that $I_{d-1}(\partial S) = |\partial S|_{d-1} < \infty$, which implies (by (3)) that, except for a set of measure zero with respect to $d\mu_{d-1}(y) \times d\theta$, any line $r_{\theta,y}$ meets ∂S a finite number of times: $n_{\partial S}(\theta, y) < \infty$. From Theorem 1 in Walther, G. (1999), it also follows that ∂S is a \mathcal{C}^1 manifold, which allows us to consider for all $x \in \partial S$, η_x , the unit outward normal vector.*

For a given θ we will separate the integral with respect to μ_{d-1} in (3), as a sum of two integrals. In the first one, we will consider the lines (defined by $y \in \theta^\perp$) that are *far* (properly defined later as condition $L(\varepsilon)$ in Definition 4) from all of the tangent spaces to ∂S , while in the second integral we will consider those lines that are *close* to some tangent space. To control the measure of these last lines, we need to introduce the following shape restriction.

Definition 3. *Let us define $E_\theta(\partial S) = \{x \in \partial S, \langle \eta_x, \theta \rangle = 0\}$ and F_θ its normal projection onto θ^\perp . Let us define, for $\varepsilon > 0$,*

$$\varphi_\theta(\varepsilon) = |\theta^\perp \cap B(F_\theta, \varepsilon)|_{d-1}.$$

We will say that ∂S is (C, ε_0) -regular if for all θ and all $\varepsilon \in (0, \varepsilon_0)$, $\varphi'_\theta(\varepsilon)$ exists and $\varphi'_\theta(\varepsilon) \leq C$.

When we use the Devroye–Wise estimator we will assume the (C, ε_0) -regular boundary condition. Once the rolling balls condition is imposed, we will show through some examples that the (C, ε_0) -regularity of the boundary is not a too restrictive hypothesis.

For instance, a polyhedron with ‘rounded corners’, such as in Figure 3, satisfies the (C, ε_0) -regularity of the boundary. Under regularity and geometric conditions on ∂S , the (C, ε_0) -regularity is related to the conjecture proposed in Alesker (2018).

To find sets that satisfy the inside and outside α rolling ball properties but without a (C, ε_0) -regular boundary, the only case that we were able to construct is a set with some E_θ having infinitely many connected components, such as the one shown in Figure 6, whose boundary is locally around some boundary point, the hypograph of the function $x^5 \sin(1/x)$.

- (a) The first set, presented in Figure 3, is a square with ‘round angles’, it has a 2-regular boundary.
- (b) The second set, presented in Figure 4, is a 2-dimensional ‘peanut’ made of 4 arcs of circle. It has a 6-regular boundary.
- (c) The third set, presented in Figure 5, is the surface of revolution generated by (b). The number of connected components of E_θ is bounded by 3 and the maximal length of a component is bounded by L , the length of the maximal perimeter (shown in blue in the figure). Thus, it is C -regular with $C \leq 3L$.
- (d) The rolling ball condition is not sufficient to guarantee the (C, ε_0) regularity of the boundary: this happens if, for instance, we replace in the smooth square shown in (a) a flat piece of the boundary by the graph of the function $x^5 \sin(1/x)$. To illustrate this behaviour, Figure 6 shows a set such that the number of connected components of E_θ (with a horizontal θ) is infinite.

For the Devroye–Wise type estimator we will also show that the convergence rate is better when we additionally assume that the number of intersections between any line and ∂S is bounded from above (that exclude the case of a linear part in ∂S).

Definition 4. *Given $S \subset \mathbb{R}^d$, we say that ∂S has a bounded number of linear intersections if there exists N_S such that, for all $\theta \in (\mathbb{S}^+)^{d-1}$ and $y \in \theta^\perp$, $n_{\partial S}(\theta, y) \leq N_S$.*

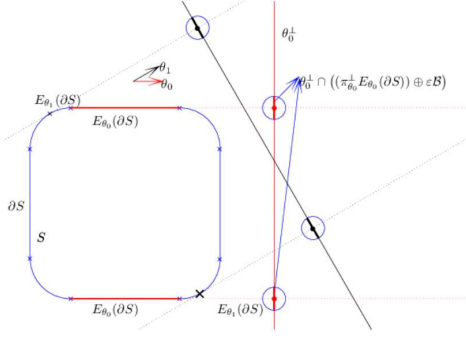


Figure 3: (a) smooth square

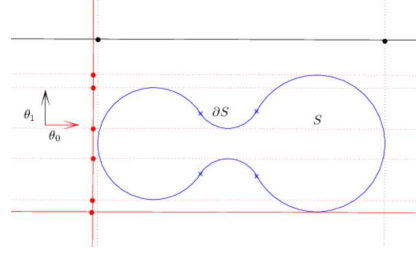


Figure 4: (b) 2D peanut

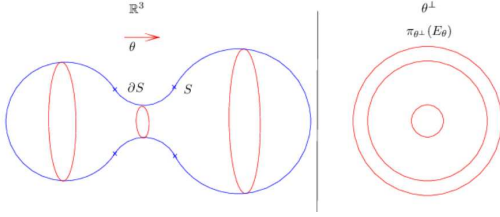


Figure 5: (c) 3D peanut

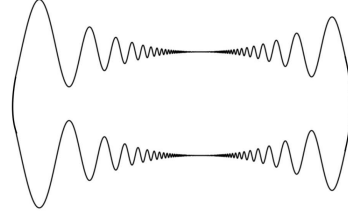


Figure 6: (d) an 'infinite wave' shape

4 Definitions of the estimators

4.1 Devroye–Wise based approach

To estimate $n_{\partial S}(\theta, y)$, note that when $r_{\theta, y}$ is not included in a $(d-1)$ -dimensional affine tangent space (tangent to ∂S), then $n_{\partial S}(\theta, y) = 2k_S(\theta, y)$ where $k_S(\theta, y)$ is the number of connected components of $r_{\theta, y} \cap S$.

Given that in general the set S is unknown, the natural idea is to plug into k_S an estimator of S . There are different kinds of set estimators, depending on the geometric restrictions imposed on S and the structure of the data (see Devroye and Wise (1980), Cholaquidis et al. (2016) and references therein). One of the most studied in the literature, which is also universally consistent, is the Devroye–Wise estimator (see Devroye and Wise (1980)), given by

$$\hat{S}_n(\varepsilon_n) = \bigcup_{i=1}^n \mathcal{B}(X_i, \varepsilon_n),$$

where $\varepsilon_n \rightarrow 0$ is a sequence of positive real numbers. This all-purpose estimator has the advantage that it is quite easy to compute the intersection of a line with its boundary (i.e. the points in the line at a distance of exactly ε_n from the sample). Unfortunately, a direct plug-in estimator does not provide consistency (i.e. $2k_{\hat{S}_n(\varepsilon_n)}(\theta, y)$ does not converges in general to $n_{\partial S}(\theta, y)$). It needs a small adjustment, as we will explain in the following definition.

Definition 5. Consider a line $r_{\theta, y}$. If $\hat{S}_n(\varepsilon_n) \cap r_{\theta, y} = \emptyset$, define $\hat{n}_{\varepsilon_n}(\theta, y) = 0$, otherwise:

- denote by I_1, \dots, I_m the connected components of $\hat{S}_n(\varepsilon_n) \cap r_{\theta, y}$. Order this sequence in such a way that $I_i = (a_i, b_i)$, with $a_1 < b_1 < a_2 < b_2 < \dots < a_m < b_m$.

- If for some consecutive intervals $I_i, I_{i+1}, \dots, I_{i+l-1}$, for all $a_i < t < b_{i+l}$ and $t \in r_{\theta, y}$, $d(t, \mathcal{X}_n) < 4\varepsilon_n$, define $A_i = (a_i, b_{i+l-1})$.
- Let j be the number of disjoint open intervals A_1, \dots, A_j that this process ended with. Then define $\hat{n}_{\varepsilon_n}(\theta, y) = 2j$.

Our first proposed estimator is

$$\hat{I}_{d-1}(\partial S) = \frac{1}{\beta(d)} \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \theta^\perp} \hat{n}_{\varepsilon_n}(\theta, y) d\mu_{d-1}(y) d\theta.$$

Under the assumption that ∂S has a bounded number of linear intersections (see Definition 4) we will consider, for a given $N_0 \geq N_S$,

$$\hat{I}_{d-1}^{N_0}(\partial S) = \frac{1}{\beta(d)} \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \theta^\perp} \min(\hat{n}_{\varepsilon_n}(\theta, y), N_0) d\mu_{d-1}(y) d\theta.$$

4.2 α -convex hull based approach

The α -convex hull of a finite set of points \mathcal{X}_n (defined by (1) with $S = \mathcal{X}_n$), which is also a consistent estimator of S under some regularity conditions (see for instance Rodríguez-Casal (2007)), has the advantage that the $(d-1)$ -dimensional Lebesgue measure of its boundary converges to the $(d-1)$ -dimensional Lebesgue measure of ∂S (see Theorem 3 below). This, together with the fact that $\partial C_\alpha(\mathcal{X}_n)$ is a rectifiable set (see the comment before Remark 1), suggests using Crofton's formula to estimate $|\partial C_\alpha(\mathcal{X}_n)|_{d-1}$. Then our second proposed estimator is

$$\begin{aligned} \check{n}_\alpha(\theta, y) &= n_{\partial C_\alpha(\mathcal{X}_n)}(\theta, y) \\ \check{I}_{d-1}(\partial S) &= \frac{1}{\beta(d)} \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \theta^\perp} \check{n}_\alpha(\theta, y) d\mu_{d-1}(y) d\theta. \end{aligned}$$

In this case, the computation of the intersection of a line with $\partial C_\alpha(\mathcal{X}_n)$ is not as direct as in the Devroye–Wise estimator. However, weaker regularity restrictions on ∂S will be required (see Theorem 2) to get the consistency of $\check{I}_{d-1}(\partial S)$ with a better convergence rate.

5 Computational and practical aspects of the algorithms

The algorithms to compute $\hat{n}_{\varepsilon_n}(\theta, y)$ and $\check{n}_\alpha(\theta, y)$ work for any finite set \mathcal{X}_n (not necessarily random). The general case for stochastic processes indexed by $T \in \mathbb{R}^+$ is obtained by replacing the set \mathcal{X}_n in the algorithm by a discretization of a trajectory of the process observed in $[0, T]$ (which is not restrictive since, the trajectories are always stocked as a finite number of points in a computer).

Let us first describe the algorithms that allows to compute the estimations of $n_{\partial S}(\theta, y)$ for a given (θ, y) .

5.1 Devroye–Wise based approach

To compute $\hat{n}_{\varepsilon_n}(\theta, y)$ for a given (θ, y) we proceed as follows.

1. Identify the centres $\mathcal{Y}_{n'} = \{Y_1, \dots, Y_{n'}\}$ of the boundary balls of $\hat{S}_n(\varepsilon_n)$ (see Aaron, Cholaquidis and Cuevas (2017)), i.e., the points $X_i \in \mathcal{X}_n$ such that

$$\max\{\|x - X_i\| : x \in \text{Vor}(X_i)\} \geq \varepsilon_n,$$

where $\text{Vor}(X_i) = \{x \in \mathbb{R}^d \text{ s.t. for all } j : \|x - X_i\| \leq \|x - X_j\|\}$ denotes the Voronoi cell of X_i .

2. Compute $d_i = d(r_{\theta,y}, Y_i)$.
3. Compute the connected components I_i , of $r_{\theta,y} \cap \hat{S}_n(\mathcal{X}_n)$, according to the following steps: Initialize the list of the extremes of these intervals by $\text{list} = \emptyset$, and then, for $i = 1$ to n' :
 - If $d_i \leq \varepsilon_n$ then compute $\{z_1, z_2\} = \mathcal{B}(Y_i, \varepsilon_n) \cap r_{\theta,y}$.
 - For $j = 1$ to 2: if $d(z_j, \mathcal{X}_n) \geq \varepsilon_n$ do $\text{list} = \text{list} \cup \{z_j\}$.

The a_i and b_i (and so the I_i) introduced in Definition 5 are obtained by a sorting procedure applied to the points z_j .

4. Obtain the a'_i and b'_i such that $I'_i = (a'_i, b'_i)$ are the connected components of $\hat{S}(4\varepsilon_n) \cap r_{\theta,y}$ by using the same procedure.
5. Lastly, compute $\hat{n}_{\varepsilon_n}(\theta, y)$, as follows:

initialization $\hat{n}_{\varepsilon_n}(\theta, y) = m$. For $i = 1$ to $m - 1$

- If there exists k such that $(b_i, a_{i+1}) \subset I'_k$ then: $\hat{n}_{\varepsilon_n}(\theta, y) = \hat{n}_{\varepsilon_n}(\theta, y) - 1$

5.2 α -convex hull based approach

It is much more involved to compute $\tilde{n}_\alpha(\theta, y)$: it requires the computation of the α -convex hull, as well as the convex hull, of the set \mathcal{X}_n . Recall that the convex hull of a sample is equal to the intersection of a finite number of half-spaces. In Edelsbrunner et al. (1983) it is proved, for dimension 2, but mentioned that the generalization is not difficult, that $C_\alpha(\mathcal{X}_n)^c$ is the union of a finite number of balls and the aforementioned half-spaces. The centres O_i of these balls, and their radii r_i , are obtained by computing the Delaunay complex of the points. Let us write $C_\alpha(\mathcal{X}_n)^c = \bigcup_i E_i$, where E_i is either a half-space or a ball. Observe that if the line $r_{\theta,y}$ is chosen at random (w.r.t. $d\mu_{d-1} \times d\theta$), $r_{\theta,y} \cap E_i$ contains fewer than 3 points.

Initialize $\text{list} = \emptyset$. Then:
for all i ,

- compute $r_{\theta,y} \cap \partial E_i$
- For all $z \in r_{\theta,y} \cap \partial E_i$
 1. If for all j $z \notin \overset{\circ}{E}_j$ do $\text{list} = \text{list} \cup \{z\}$

then $\tilde{n} = \#\text{list}$.

5.3 Integralgeometric estimations via a Monte Carlo method

Once we have estimated $n_{\partial S}(\theta, y)$ by $\hat{n}_{\varepsilon_n}(\theta, y)$ for any given (y, θ) , $\hat{I}_{d-1}(\partial S)$ can be calculated via the Monte-Carlo method, as follows. Generate a random sample $\theta_1, \dots, \theta_k$ uniformly distributed on $(S^+)^{d-1}$. For each $i = 1, \dots, k$, build a random sample $\aleph_i = \{y_1^i, \dots, y_\ell^i\}$ uniformly distributed on the $(d-1)$ -dimensional hyper-cube $[-L, L]^{d-1} \subset \theta_i^\perp$, where $L = \max_{j=1, \dots, n} \|X_j\|$, and independent of $\theta_1, \dots, \theta_k$. Then, the estimators are given by

$$\hat{I}_{d-1}^{(\ell, k)}(\partial S) = \frac{(2L)^{d-1}}{\beta(d)} \frac{\ell}{lk} \sum_{i=1}^k \sum_{j=1}^{\ell} \hat{n}_{\varepsilon_n}(\theta_i, y_j^i) \quad (4)$$

$$\hat{I}_{d-1}^{(\ell, k, N_0)}(\partial S) = \frac{(2L)^{d-1}}{\beta(d)} \frac{\ell}{lk} \sum_{i=1}^k \sum_{j=1}^{\ell} \min(\hat{n}_{\varepsilon_n}(\theta_i, y_j^i), N_0) \quad (5)$$

$$\check{I}_{d-1}^{(\ell, k)}(\partial S) = \frac{(2L)^{d-1}}{\beta(d)} \frac{\ell}{lk} \sum_{i=1}^k \sum_{j=1}^{\ell} \check{n}_r(\theta_i, y_j^i). \quad (6)$$

5.4 Parameter Selection

When considering the Devroye-Wise approach we need to choose the parameter ε_n (and possibly also the parameter N_0) while when considering the α -hull approach it is the parameter α that has to be chosen. With regard to ε_n , as mentioned in Cuevas and Rodriguez-Casal (2004), the choice of $2 \max_i \min_j \|X_i - X_j\|$ provides a fully data-driven selection method. An automatic selection method of α is proposed in Rodríguez-Casal and Saavedra-Nieves (2019).

6 Main results

In this section we will state our main results. All proofs are given in the Appendix.

6.1 Convergence rates for the Devroye-Wise based estimator under α -rolling condition and (C, ε_0) -regularity.

Theorem 1. *Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the outside and inside α -rolling conditions. Assume also that S is (C, ε_0) -regular for some positive constants C and ε_0 . Let $\mathcal{X}_n = \{X_1, \dots, X_n\} \subset S$. Let $\varepsilon_n \rightarrow 0$ such that $d_H(\mathcal{X}_n, S) \leq \varepsilon_n$. Then*

$$\hat{I}_{d-1}(\partial S) = |\partial S|_{d-1} + \mathcal{O}(\sqrt{\varepsilon_n}).$$

Moreover, for n large enough,

$$|\mathcal{O}(\sqrt{\varepsilon_n})| \leq \frac{5C \text{diam}(S)}{6\sqrt{\alpha}} \sqrt{\varepsilon_n},$$

C being the constant of the (C, ε_0) -regularity of S .

Remark 2. *Theorem 4 in Cuevas and Rodriguez-Casal (2004) gives some insight into how to choose the parameter ε_n for the case in which $\{X_1, \dots, X_n\}$ is an iid sample of a random vector X supported on S . It states that if $\varepsilon_n = C'(\log(n)/n)^{1/d}$, where C' is a large enough positive constant, then with probability one, for n large enough, $S \subset \hat{S}_n$. In addition,*

$d_H(\partial S, \partial \hat{S}_n(\varepsilon_n)) \rightarrow 0$, and $d_H(S, \hat{S}_n(\varepsilon_n)) \rightarrow 0$. Although this does not imply that $|\partial \hat{S}_n|_{d-1}$ converges to $|\partial S|_{d-1}$, Theorem 1 states that we can consistently estimate the integralgeometric measure of ∂S by means of Crofton's formula.

From Remark 2 and the previous theorem, we can obtain the rate of convergence for the iid case:

Corollary 1. *Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the inside and outside α -rolling conditions. Assume also that S is (C, ε_0) -regular for some positive constants C and ε_0 . Let X_1, \dots, X_n be an iid sample of X with distribution P_X supported on S . Assume that P_X has density f (w.r.t. μ_d) bounded from below by some $c > 0$. Let $\varepsilon_n = C'(\ln(n)/n)^{1/d}$ and $C' > (6/(c\omega_d))^{1/d}$. Then with probability one, for n large enough,*

$$\hat{I}_{d-1}(\partial S) = |\partial S|_{d-1} + \mathcal{O}\left(\left(\frac{\ln(n)}{n}\right)^{\frac{1}{2d}}\right).$$

In a more general setting, the conclusion of Theorem 1 holds when the set of points \mathcal{X}_n is replaced by the trajectory of any stochastic process $\{X_t\}_{t \geq 0}$ included in S , observed in $[0, T]$, such that $d_H(\mathcal{X}_T, S) \rightarrow 0$ as $T \rightarrow \infty$. This is the case (for example) of some reflected diffusions and in particular the reflected Brownian motion (RBM). This has been recently proven in Corollary 1 in Cholaquidis et al. (2016), for RBM without drift (see also Cholaquidis et al. (2021) for RBM with drift). RBM with drift is defined as follows: let D be a bounded domain in \mathbb{R}^d (i.e., a bounded, connected open set), such that ∂D is \mathcal{C}^2 . Given a d -dimensional Brownian motion $\{B_t\}_{t \geq 0}$, departing from $B_0 = 0$ and defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}_x)$, the RBM with drift is the (unique) solution to the following stochastic differential equation on \overline{D} :

$$X_t = X_0 + B_t + \int_0^t g(X_s)ds - \int_0^t \eta_{X_s} \xi(ds), \quad \text{where } X_t \in \overline{D}, \forall t \geq 0,$$

where the drift, $g(x)$, is assumed to be Lipschitz, and $\{\xi_t\}_{t \geq 0}$ is the corresponding *local time*: i.e., a one-dimensional continuous non-decreasing process with $\xi_0 = 0$ that satisfies $\xi_t = \int_0^t \mathbb{I}_{\{X_s \in \partial D\}} d\xi_s$.

From Corollary 1 together with Proposition 3 of Cholaquidis et al. (2016), we have the following result for the RBM without drift:

Corollary 2. *Let $S \subset \mathbb{R}^d$ be a non-empty compact set with connected interior such that $S = \text{int}(S)$, and suppose that S fulfills the outside and inside α -rolling conditions. Assume also that S is (C, ε_0) -regular for some positive constants C and ε_0 . Let $\{B_t\}_{t \geq 0} \subset S$ be an RBM (without drift). Then, with probability one, for T large enough,*

$$\hat{I}_{d-1}(\partial S) = |\partial S|_{d-1} + o\left(\left(\frac{\ln(T)^2}{T}\right)^{\frac{1}{2d}}\right).$$

6.2 Convergence rates for the Devroye-Wise based estimator under α -rolling condition, (C, ε_0) -regularity hypotheses and bounded number of linear intersections.

If the number of linear intersection of ∂S is assumed to be bounded by a constant N_S , the use of $\min(\hat{n}_{\varepsilon_n}, N_0)$ (for any $N_0 \geq N_S$) allows us to obtain better convergence rates.

Theorem 2. Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the outside and inside α -rolling conditions. Assume also that S is (C, ε_0) -regular for some positive constants C and ε_0 and that ∂S has a number of linear intersection bounded by N_S . Let $\mathcal{X}_n = \{X_1, \dots, X_n\} \subset S$. Let $\varepsilon_n \rightarrow 0$ such that $d_H(\mathcal{X}_n, S) \leq \varepsilon_n$ and $N_0 \geq N_S$. Then

$$\hat{I}_{d-1}^{N_0}(\partial S) = |\partial S|_{d-1} + \mathcal{O}(\varepsilon_n).$$

Moreover, for n large enough,

$$|\mathcal{O}(\varepsilon_n)| \leq 2CN_0\varepsilon_n,$$

C being the constant of the (C, ε_0) -regularity of S .

Corollary 3. Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the inside and outside α -rolling conditions. Assume also that S is (C, ε_0) -regular for some positive constants C and ε_0 and that ∂S has a bounded number of linear intersections. Let X_1, \dots, X_n be iid random vectors with distribution P_X , supported on S . Assume that P_X has density f (w.r.t. μ_d) bounded from below by some $c > 0$. Let $\varepsilon_n = C'(\ln(n)/n)^{1/d}$ and $C' > (6/(c\omega_d))^{1/d}$. Then with probability one, for n large enough,

$$\hat{I}_{d-1}^{N_0}(\partial S) = |\partial S|_{d-1} + \mathcal{O}\left(\left(\frac{\ln(n)}{n}\right)^{\frac{1}{d}}\right).$$

Corollary 4. Let $S \subset \mathbb{R}^d$ be a non-empty compact set with connected interior such that $S = \text{int}(S)$, and suppose that S fulfills the outside and inside α -rolling conditions. Assume also that S is (C, ε_0) -regular for some positive constants C and ε_0 and that ∂S has a number of linear intersection bounded by N_S . Let $\{B_t\}_{t>0} \subset S$ be an RBM (without drift). Then, with probability one, for T large enough,

$$\hat{I}_{d-1}^{N_0}(\partial S) = |\partial S|_{d-1} + o\left(\left(\frac{\ln(T)^2}{T}\right)^{\frac{1}{d}}\right).$$

6.3 α' -hull based estimator under α -rolling ball condition

In Arias-Castro And Rodríguez-Casal (2017) it has been proved that, in dimension two, under some regularity assumptions, the length of the boundary of the α -shape of an iid sample converges to the length of the boundary of the set. The α -shape has the very good property that its boundary is very easy to compute, and so its surface measure. Unfortunately we are not sure that the results can be extend to higher dimension. Nevertheless considering the α -convex hull (which is quite close to the α -shape) allows to extend the results on the surface measure for any dimension. The price to pay is the difficulty to obtain an explicit formula for the surface measure of the α -convex hull. We so propose to skip this problem by a Monte-Carlo estimation based on Crofton's formula. The following theorem states that the surface measure of the boundary of the α -convex hull of an iid sample converges to the surface of the boundary of the set. Observe that in this case, with no need for the additional hypothesis of (C, ε_0) -regularity, the convergence rate is far better than the one obtained with the Devroye–Wise estimator, the price to pay being the computational cost.

Theorem 3. Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the inside and outside α -rolling conditions. Assume also that ∂S is of class \mathcal{C}^3 . Let X_1, \dots, X_n be an iid sample of X with distribution P_X supported on S . Assume that P_X has density f (w.r.t. μ_d) bounded from below by some $c > 0$. Suppose $\alpha' \leq \alpha$. Then with probability one, for n large enough,

$$1. \quad ||\partial S|_{d-1} - |\partial C_{\alpha'}(\mathcal{X}_n)|_{d-1}| = \mathcal{O}((\ln(n)/n)^{2/(d+1)}),$$

2. as a consequence

$$\check{I}_{d-1}(\partial S) = |\partial S|_{d-1} + \mathcal{O}\left(\left(\frac{\ln(n)}{n}\right)^{\frac{2}{d+1}}\right).$$

6.4 On the rates of convergence

- Observe that we obtain the same convergence rate as the one provided in Arias-Castro And Rodríguez-Casal (2017) for $d = 2$, where is also conjectured as suboptimal with regard to the result obtained in Korostel'ev and Tsybakov (1993) (see Chapter 8). Indeed, as mentioned in Arias-Castro And Rodríguez-Casal (2017), if the measure of the symmetric difference between S and an estimator \hat{S}_n is bounded by ε_n , we can only expect that plug-in methods allow to estimate $|\partial S|_{d-1}$ with a convergence rate ε_n .
- Thus, in the iid setting, the estimator defined by (6) (respectively (7) to (9)) can be seen as “optimal” relatively to the use of the Devroye–Wise support estimator (respectively the α -convex hull support estimator), since they achieve the best possible convergence rate for those estimators.
- This is nevertheless far from being optimal from a minimax rate. Indeed the minimax rate can be conjecture to be $n^{-\frac{d+3}{2d+2}}$, because it is the minimax rate for the volume estimation (see Arias-Castro, et al. (2017)) and, in Kim and Korostel'ev (2000) it is proved that the minimax rate is the same for the volume estimation and the surface area estimation (in the image setting that usually extend to the iid inside setting). Unfortunately finding a nice bias correction as in Arias-Castro, et al. (2017) for the surface area estimation is much more involved.

7 Appendix

7.1 Proof of Theorems 1 and 2

Sketch of the proof of Theorems 1 and 2

The idea is to consider separately the set of lines that intersect $\partial S_n(\varepsilon_n)$:

1. If a line $r_{\theta,y} = y + \lambda\theta$ is ‘far enough’ (fulfilling condition $L(\varepsilon)$ for some $\varepsilon > 0$, see Definition 6) from the tangent spaces, then our algorithm allows a perfect estimation of $n_{\partial S}(y, \theta)$, see Lemma 4.
2. Considering the set of lines that are not ‘far enough’ from the tangent spaces (denoted by $\mathcal{A}_{\varepsilon_n}(\theta)$), see Definition 6), Corollary 5 states that, under (C, ϵ_0) -regularity, the integral of $\hat{n}_{\varepsilon_n}(\theta, y)$ on $\mathcal{A}_{\varepsilon_n}(\theta)$ is bounded from above by $C'\varepsilon_n^{1/2}$, with C' a positive constant. Theorem 2 states that the previous bound can be improved to $C'\varepsilon_n$, under (C, ϵ_0) -regularity, if ∂S has a bounded number of linear intersections.

7.1.1 Condition $L(\varepsilon)$

Now we define the two sets of lines to be treated separately: The lines that are ‘far’ from an affine tangent space, and the lines that are ‘close to being tangent’ to ∂S . More precisely, assume that ∂S is smooth enough so that for all $x \in \partial S$, the unit outer normal vector η_x at x is well defined. Now we define

$$\mathcal{T}_S = \{x + (\eta_x)^\perp : x \in \partial S\},$$

the collection of all the affine $(d - 1)$ -dimensional tangent spaces.

Definition 6. Let $\varepsilon \geq 0$. A line $r_{\theta,y} = y + \lambda\theta$ fulfills **condition** $L(\varepsilon)$ if y is at a distance larger than 4ε from all the affine hyper-planes $w + \eta^\perp \in \mathcal{T}_S$ satisfying $\langle \eta, \theta \rangle = 0$.

For a given θ , we define

$$\mathcal{A}_\varepsilon(\theta) = \left\{ y \in \theta^\perp : \|y\| \leq \text{diam}(S) \text{ and } r_{\theta,y} \text{ does not satisfy } L(\varepsilon) \right\}.$$

7.1.2 Some useful lemmas

Lemma 1. Let S be a compact set fulfilling the outside and inside α -rolling conditions. Let $r_{\theta,y}$ be a line that fulfills condition $L(0)$ and $r_{\theta,y} \cap \partial S \neq \emptyset$. Then $r_{\theta,y}$ intersects ∂S in a finite number of points.

Proof. Because S fulfills the outside and inside α -rolling conditions, Theorem 1 in Walther, G. (1999) implies that for any $x \in \partial S$, the affine $(d-1)$ -dimensional tangent space $T_x \partial S$ exists. If $r_{\theta,y}$ fulfills $L(0)$, then $r_{\theta,y}$ is not included in any hyper-plane tangent to S . Suppose that $\partial S \cap r_{\theta,y}$ is not finite. Then, by compactness, one can extract a subsequence t'_n that converges to $y' \in \partial S$. Note that for all $(n, p) \in \mathbb{N}^2$ $(t'_n - t'_{n+p}) / \|t'_n - t'_{n+p}\| = \pm\theta$, which implies that $(t'_n - y') / \|t'_n - y'\| = \pm\theta$. Lastly, if $n \rightarrow \infty$, then $\theta \in T_{y'} \partial S$. Considering y' , we have $y' \in \partial S$, $\theta \in T_{y'} \partial S$ and $y' \in r_{\theta,y}$, which contradicts the assumption that $r_{\theta,y}$ is not included in any hyper-plane tangent to S . \square

Lemma 2. Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the outside and inside α -rolling conditions. Let $\varepsilon > 0$ such that $\varepsilon < 4\alpha$ and $\nu = 2\sqrt{2\varepsilon(\alpha - 2\varepsilon)}$. For any line $r_{\theta,y}$ fulfilling condition $L(\varepsilon)$ and $r_{\theta,y} \cap \partial S \neq \emptyset$, we have that $r_{\theta,y}$ meets ∂S at a finite number of points t_1, \dots, t_k , where $t_{i+1} - t_i > 2\nu$ for all $i = 1, \dots, k-1$. Consequently, if $\varepsilon < \alpha/4$, then $k \leq \text{diam}(S)\varepsilon^{-1/2}/(4\sqrt{\alpha})$.

Proof. Note that if a line fulfills condition $L(\varepsilon)$, then it fulfills condition $L(0)$. Consequently, the fact that $r_{\theta,y}$ intersects ∂S in a finite number of points follows from Lemma 1. Let us denote by $t_1 < \dots < t_k$ the intersection of $r_{\theta,y}$ with ∂S . Proceeding by contradiction, assume that for some i , $t_{i+1} - t_i < 2\nu$. Let us denote by η_{t_i} and $\eta_{t_{i+1}}$ the outer normal vectors at t_i and t_{i+1} , respectively. We have two cases: the open interval $(t_i, t_{i+1}) \subset S^c$ or $(t_i, t_{i+1}) \subset \text{int}(S)$. Let us consider the first case (the proof for the second one is similar).

Because $(t_i, t_{i+1}) \subset S^c$ and S fulfills the inside α -rolling condition on t_i , there exists $z \in S$ such that $t_i \in \partial \mathcal{B}(z, \alpha)$ and $\mathcal{B}(z, \alpha) \subset S$. In particular, $\mathcal{B}(z, \alpha) \cap (t_i, t_{i+1}) = \emptyset$, which implies $\langle \eta_{t_i}, \theta \rangle \geq 0$.

Reasoning in the same way but with t_{i+1} , we get $\langle \eta_{t_{i+1}}, \theta \rangle \leq 0$. Given that $r_{\theta,y}$ is not included in any tangent hyperplane, we have that $\langle \eta_{t_i}, \theta \rangle > 0$ and $\langle \eta_{t_{i+1}}, \theta \rangle < 0$. Because S fulfills the inside and outside α -rolling conditions, ∂S is a $(d-1)$ -dimensional \mathcal{C}^1 manifold whose normal vector is Lipschitz (see Theorem 1 in Walther, G. (1999)). By Theorem 3.8 in Colesanti and Manselli (2010), there exists a curve $\gamma : [0, 1] \rightarrow \partial S$ such that $\gamma(0) = t_i$, $\gamma(1) = t_{i+1}$ and $d(\gamma(t), r_{\theta,y}) < 4\varepsilon$ for all t . From $\langle \eta_{t_i}, \theta \rangle > 0$ and $\langle \eta_{t_{i+1}}, \theta \rangle < 0$, it follows that there exists an $s_0 \in (0, 1)$ such that $\langle \eta_{\gamma(s_0)}, \theta \rangle = 0$, which contradicts the hypothesis that y is at a distance larger than 4ε from all the $(d-1)$ -dimensional hyperplanes tangent to S . This proves that $t_{i+1} - t_i > 2\nu$ for all $i = 1, \dots, k-1$. \square

Lemma 3. Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the outside and inside α -rolling ball conditions and with a (C, ε_0) -regular boundary. Then for all $\varepsilon \leq \varepsilon_0$,

$$\int_{\theta \in (S^+)^{d-1}} \int_{y \in \mathcal{A}_\varepsilon(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta \leq C \frac{\text{diam}(S)}{2\sqrt{\alpha}} \sqrt{\varepsilon}.$$

Moreover if ∂S has bounded number of linear intersections then

$$\int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \mathcal{A}_\varepsilon(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta \leq CN_S \varepsilon. \quad (7)$$

Proof. From the proof of the previous lemma, it follows that for any $y \in E_\theta$ with $d(y, F_\theta) = l$, $n_{\partial S}(\theta, y) \leq \text{diam}(S) l^{-1/2} / (4\sqrt{\alpha})$. Hence,

$$\begin{aligned} & \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \mathcal{A}_\varepsilon(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta \\ &= \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{l=0}^\varepsilon \int_{\{y \in \theta^\perp : d(y, F_\theta) = l\}} n_{\partial S}(\theta, y) d\mu_{d-2}(y) dl d\theta \\ &\leq \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{l=0}^\varepsilon \int_{\{y \in \theta^\perp : d(y, F_\theta) = l\}} \frac{1}{4} \text{diam}(S) (\alpha l)^{-1/2} d\mu_{d-2}(y) dl d\theta \\ &\leq \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{l=0}^\varepsilon \frac{1}{4} \text{diam}(S) (\alpha l)^{-1/2} \int_{\{y \in \theta^\perp : d(y, F_\theta) = l\}} d\mu_{d-2}(y) dl d\theta \\ &\leq \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{l=0}^\varepsilon \frac{1}{4} \text{diam}(S) (\alpha l)^{-1/2} |\{y \in \theta^\perp : d(y, F_\theta) = l\}|_{d-2} dl d\theta. \end{aligned}$$

By the definition of φ_θ ,

$$\left| \{y \in \theta^\perp : l \leq d(y, F_\theta) \leq l + dl\} \right|_{d-1} = \varphi_\theta(l + dl) - \varphi_\theta(l).$$

From the (C, ε_0) -regularity of ∂S and the mean value theorem we obtain

$$\left| \{y \in \theta^\perp : d(y, F_\theta) = l\} \right|_{d-2} \leq \sup_{\varepsilon \in (0, \varepsilon_0)} \varphi'_\theta(\varepsilon) \leq C,$$

which implies

$$\begin{aligned} & \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \mathcal{A}_\varepsilon(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta \leq \\ & \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{l=0}^\varepsilon C \frac{1}{4} \text{diam}(S) (\alpha l)^{-1/2} dl d\theta \leq C \frac{\text{diam}(S)}{2\sqrt{\alpha}} \sqrt{\varepsilon}. \end{aligned}$$

Applying exactly the same calculus, under the hypothesis of bounded number of linear intersections for ∂S , we get

$$\int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \mathcal{A}_\varepsilon(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta \leq \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{l=0}^\varepsilon CN_S dl d\theta \leq CN_S \varepsilon.$$

□

Lemma 4. *Let S be a compact set fulfilling the outside and inside α -rolling conditions. Let $\mathcal{X}_n = \{X_1, \dots, X_n\} \subset S$. Let $\varepsilon_n \rightarrow 0$ be such that $d_H(\mathcal{X}_n, S) \leq \varepsilon_n$. Let $r_{\theta, y} = y + \lambda\theta$ be any line fulfilling condition $L(\varepsilon_n)$. Then, for n large enough such that $4\varepsilon_n < \alpha$, $n_{\partial S}(\theta, y) = \hat{n}_{\varepsilon_n}(\theta, y)$.*

Proof. Note that the choice of ε_n ensures that $S \subset \hat{S}_n(\varepsilon_n)$, thus

$$r_{\theta,y} \cap S \subset r_{\theta,y} \cap \hat{S}_n(\varepsilon_n). \quad (8)$$

First, we will prove that

$$\hat{n}_{\varepsilon_n}(\theta, y) \geq n_{\partial S}(\theta, y). \quad (9)$$

Because $\hat{n}_{\varepsilon_n}(\theta, y)$ is not the number of connected components of $r_{\theta,y} \cap \hat{S}_n(\varepsilon_n)$, (9) does not follow directly from (8). If $r_{\theta,y} \cap \partial S = \emptyset$ inequality (9) holds. Assume $r_{\theta,y} \cap \partial S \neq \emptyset$. Let $t_1 < \dots < t_k$ be the intersection of $r_{\theta,y}$ with ∂S (this set is finite due to Lemma 1). Let us prove that

$$\text{if } (t_i, t_{i+1}) \subset S^c, \text{ then: } \exists s \in (t_i, t_{i+1}) \text{ such that } d(s, S) > 4\varepsilon_n. \quad (10)$$

Because S fulfills the inside α -rolling condition on t_i , there exists a $z_i \in S$ such that $t_i \in \partial \mathcal{B}(z, \alpha)$ and $\mathcal{B}(z, \alpha) \subset S$. Since $\mathcal{B}(z, \alpha) \cap (t_i, t_{i+1}) = \emptyset$, it follows that $\langle \eta_{t_i}, \theta \rangle \geq 0$ (recall that $\eta_{t_i} = (t_i - z_i)/\alpha$ and $t_{i+1} - t_i = \|t_{i+1} - t_i\|\theta$). Reasoning in the same way but with t_{i+1} , $\langle \eta_{t_{i+1}}, \theta \rangle \leq 0$. By condition $L(\varepsilon_n)$ we obtain

$$\langle \eta_{t_i}, \theta \rangle > 0 \text{ and } \langle \eta_{t_{i+1}}, \theta \rangle < 0. \quad (11)$$

Suppose that for all $t \in (t_i, t_{i+1})$ we have $d(t, \partial S) \leq 4\varepsilon_n$. Take n large enough such that $4\varepsilon_n < \alpha$. Because ∂S fulfills the outside and inside α -rolling conditions, by Lemma 2.3 in Pateiro-López and Rodríguez-Casal (2009) it has positive reach. Then, by Theorem 4.8 in Federer (1956), $\gamma = \{\gamma(t) = \pi_{\partial S}(t), t \in (t_i, t_{i+1})\}$, the orthogonal projection onto ∂S of the interval (t_i, t_{i+1}) is well defined and is a continuous curve in ∂S . By Theorem 1 in Walther, G. (1999), the map from ∂S to \mathbb{R}^d that sends $\eta_x \in \partial \mathcal{B}(0, 1)$ to $x \in \partial S$ is Lipschitz. Thus, $t \rightarrow \langle \eta_{\gamma(t)}, \theta \rangle$ is a continuous function of t for all $t \in (t_i, t_{i+1})$, which, together with (11), ensures the existence of an $s \in (t_i, t_{i+1})$ such that $d(s, \gamma(s)) \leq 4\varepsilon_n$ and $\theta \in \eta_{\gamma(s)}^\perp$, which contradicts the assumption that $r_{\theta,y}$ fulfills condition $L(\varepsilon_n)$. This proves (10), which implies that

$$\text{if } (t_i, t_{i+1}) \subset S^c, \text{ then: } \exists s \in (t_i, t_{i+1}) \text{ such that } d(s, \mathcal{X}_n) > 4\varepsilon_n$$

and now (9) follows from (8).

Next we will prove the opposite inequality,

$$\hat{n}_{\varepsilon_n}(\theta, y) \leq n_{\partial S}(\theta, y). \quad (12)$$

Assume first $r_{\theta,y} \cap \partial S \neq \emptyset$. Let $\{t_1, \dots, t_k\}$ be the intersection of $r_{\theta,y}$ with ∂S (this set is finite due to Lemma 1).

Consider $t^* \in (t_i, t_{i+1}) \subset S^c$ and $t^* \in \hat{S}_n(\varepsilon_n)$. Equation (12) will be derived from the fact that $(t^*, t_{i+1}) \subset \hat{S}_n(\varepsilon_n) \cap r_{\theta,y}$ or $[t_i, t^*) \subset \hat{S}_n(\varepsilon_n) \cap r_{\theta,y}$.

Introduce $\psi(t) : (t_i, t_{i+1}) \rightarrow \mathbb{R}$ defined by $\psi(t) = d(t, \partial S)$. Consider points t such that $d(t, \partial S) < \alpha$, and let $p_t \in \partial S$ such that $\|p_t - t\| = d(t, \partial S)$. By item (3) in Theorem 4.8 in Federer (1956), $\psi'(t) = \langle \eta_{p_t}, \theta \rangle$.

Let X_j be the closest observation to t^* (recall that because $t^* \in \hat{S}_n(\varepsilon_n)$, we have $\|X_j - t^*\| \leq \varepsilon_n$). Now, because there exists a point $p^* \in [t, X_j] \cap \partial S$, we obtain that $\psi(t^*) \leq \varepsilon_n$ and, because $r_{\theta,y}$ fulfills $L(\varepsilon_n)$, $\langle \eta_{p^*}, \theta \rangle \neq 0$.

Assume that, for instance, $\langle \eta_{p^*}, \theta \rangle < 0$. Then $\psi(t^*) \leq \varepsilon_n$ and $\psi'(t^*) < 0$. Suppose that there exists a $t' \in (t^*, t_{i+1})$ such that $\psi(t') \geq \varepsilon_n$ and consider $t'' = \inf\{t > t^*, \psi(t) \geq \varepsilon_n\}$. Then for all $t \in (t^*, t'')$ we have $\psi(t) \leq \varepsilon_n$, and thus ψ is differentiable on this interval. From the fact that $\psi(t'') \geq \psi(t^*)$ and $\psi'(t^*) < 0$ we deduce that there exists a $\tilde{t} \in (t^*, t'')$ such that $\psi'(\tilde{t}) = 0$,

which contradicts $L(\varepsilon_n)$ because $\psi(\tilde{t}) \leq \varepsilon_n$. To summarize, we have shown that if $\langle \eta_{p_{i^*}}, \theta \rangle < 0$, then $(t^*, t_{i+1}) \subset \hat{S}_n(\varepsilon_n)$. Symmetrically, if $\langle \eta_{p_{i^*}}, \theta \rangle > 0$, then $(t_i, t^*) \subset \hat{S}_n(\varepsilon_n)$, which concludes the proof.

Reasoning in the same way, if $r_{\theta, y} \cap \partial S = \emptyset$ and $\hat{n}_{\varepsilon_n}(\theta, y) > 0$, a contradiction with condition $L(\varepsilon_n)$ is obtained. \square

Lemma 5. *Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the outside and inside α -rolling conditions. Let $\varepsilon_n \rightarrow 0$ be a sequence such that $d_H(\mathcal{X}_n, S) \leq \varepsilon_n$, while $r_{\theta, y} = y + \lambda\theta$ and A_1, \dots, A_k are the sets in Definition 5. Put $A_i = (a_i, b_i)$ for $i = 1, \dots, k$, and suppose that the sets are indexed in such a way that $a_1 < b_1 < a_2 < \dots < b_k$. Then for all $i = 2, \dots, k$, we have that $\|a_i - b_{i-1}\| > 3\sqrt{\varepsilon_n \alpha}$ for n large enough such that $3\sqrt{\alpha \varepsilon_n} < \alpha/2$ and $\varepsilon_n < \alpha/2$, which implies*

$$\hat{n}_{\varepsilon_n}(\theta, y) \leq \frac{\text{diam}(S)}{3\sqrt{\alpha}} \varepsilon_n^{-1/2}.$$

Proof. Assume by contradiction that for some i , $\|a_i - b_{i-1}\| \leq 3\sqrt{\varepsilon_n \alpha}$. By construction, $[b_{i-1}, a_i] \subset \hat{S}_n(\varepsilon_n)^c \subset S^c$. Because a_i and b_i are on $\partial \hat{S}_n(\varepsilon_n)$, we have $d(a_i, \mathcal{X}_n) = d(b_{i-1}, \mathcal{X}_n) = \varepsilon_n$.

The projection $\pi_S : [b_{i-1}, a_i] \rightarrow \partial S$ is uniquely defined because ∂S has reach at least α and $d(t, \partial S) \leq d(t, a_i) + d(a_i, \partial S) \leq \|a_i - b_{i-1}\| + d(a_i, \mathcal{X}_n)$ for all $t \in (b_{i-1}, a_i)$, $\|a_i - b_{i-1}\| \leq 3\sqrt{\varepsilon_n \alpha} < \alpha/2$ and $d(a_i, \partial S) \leq \varepsilon_n \leq \alpha/2$. Moreover, π is a continuous function. Hence $\max_{x \in [b_{i-1}, a_i]} \|x - \pi_S(x)\| \geq \varepsilon_n$, and the maximum is attained at some $x_0 \in [b_{i-1}, a_i]$. We will prove that $\|x_0 - \pi_S(x_0)\| \geq 3\varepsilon_n$, which guarantees that $x_0 \in (b_{i-1}, a_i)$ and that η_0 , the outward unit normal vector to ∂S at $\pi_S(x_0)$, is normal to θ . Indeed, suppose by contradiction that for all $t \in (b_{i-1}, a_i)$, $d(t, \partial S) \leq 3\varepsilon_n$. Then $d(t, \mathcal{X}_n) \leq 4\varepsilon_n$, which contradicts the definition of the points a_i and b_i . Put $z_0 = \pi_S(x_0) + \eta_0 \alpha$. Observe that $d(a_i, S) \leq \varepsilon_n$ and $d(b_{i-1}, S) \leq \varepsilon_n$. From the outside α -rolling condition at $\pi_S(x_0)$, and using the fact that η_0 is normal to θ , we have (see Figure 7)

$$r_{\theta, y} \cap \mathcal{B}(z_0, \alpha - \varepsilon_n) \subset [b_{i-1}, a_i],$$

which implies, see Figure 7, that $\|a_i - b_{i-1}\| \geq 2\sqrt{(\alpha - \varepsilon_n)^2 - (\alpha - l)^2}$, where $l = d(x_0, \pi_S(x_0))$. Therefore,

$$\|a_i - b_{i-1}\| \geq 2\sqrt{(l - \varepsilon_n)(2\alpha - l - \varepsilon_n)}. \quad (13)$$

If we bound $l \geq 3\varepsilon_n$ and use the fact that $l = o(1)$, which follows from $l \leq \|b_{i-1} - a_i\| + \varepsilon_n \leq 3\sqrt{\varepsilon_n \alpha} + \varepsilon_n$, then we get, from (13),

$$\|a_i - b_{i-1}\| \geq 2\sqrt{2\varepsilon_n(2\alpha - l - \varepsilon_n)} = 2\sqrt{4\varepsilon_n \alpha(1 + o(1))} = 4\sqrt{\alpha \varepsilon_n}(1 + o(1)),$$

and for n large enough this contradicts $\|a_i - b_{i-1}\| \leq 3\sqrt{\alpha \varepsilon_n}$.

Lastly, the number of disjoint intervals A_i is bounded from above by $\text{diam}(S)/(3\sqrt{\varepsilon_n \alpha})$. Thus, $\hat{n}_{\varepsilon_n}(\theta, y) \leq \text{diam}(S)/(3\sqrt{\varepsilon_n \alpha})$. \square

Corollary 5. *Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the outside and inside α -rolling conditions and with a (C, ε_0) -regular boundary. For n large enough such that $3\sqrt{\alpha \varepsilon_n} < \min(\alpha/2, \varepsilon_0)$, we have*

$$\int_{\theta} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} \hat{n}_{\varepsilon_n}(\theta, y) d\mu_{d-1}(y) d\theta \leq C \frac{\text{diam}(S)}{3\sqrt{\alpha}} \sqrt{\varepsilon_n}.$$

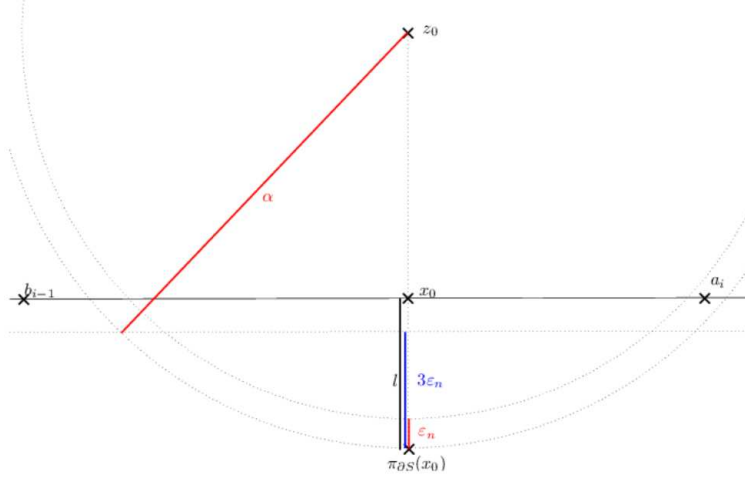


Figure 7: $\|a_i - b_{i-1}\| \geq 2\sqrt{(\alpha - \varepsilon_n)^2 - (\alpha - l)^2}$, where $l = d(x_0, \pi_S(x_0))$.

7.1.3 Proof of Theorem 1

Without loss of generality, we can assume that $0 \in S$. Recall that for $\theta \in (\mathcal{S}^+)^{d-1}$, $\mathcal{A}_{\varepsilon_n}(\theta)$ is the set of all $y \in \theta^\perp$ such that $\|y\| \leq \text{diam}(S)$ and $r_{\theta,y}$ does not fulfill $L(\varepsilon_n)$. First, from Lemma 4, we have

$$|I_{d-1}(\partial S) - \hat{I}_{d-1}(\partial S)| \leq \frac{1}{\beta(d)} \int_{\theta \in (\mathcal{S}^+)^{d-1}} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} |\hat{n}_{\varepsilon_n}(\theta, y) - n_{\partial S}(\theta, y)| d\mu_{d-1}(y) d\theta.$$

So, by the triangle inequality we can bound the difference between the integralgeometric and its estimation by

$$\begin{aligned} \frac{1}{\beta(d)} \int_{\theta \in (\mathcal{S}^+)^{d-1}} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} \hat{n}_{\varepsilon_n}(\theta, y) d\mu_{d-1}(y) d\theta + \\ \frac{1}{\beta(d)} \int_{\theta \in (\mathcal{S}^+)^{d-1}} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta. \end{aligned}$$

Now, by applying Corollaries 5 and Lemma 3, we get that

$$|I_{d-1}(\partial S) - \hat{I}_{d-1}(\partial S)| \leq \frac{5C \text{diam}(S)}{6\beta(d)\sqrt{\alpha}} \sqrt{\varepsilon_n},$$

for n large enough.

7.1.4 Proof of Theorem 2

The proof of Theorem 2 is basically the same than the previous one. Since $N_0 \geq N_S$ Lemma 4 ensures that, for all $r_{y,\theta}$ not in $\mathcal{A}_{\varepsilon_n}(\theta)$, $\min(\hat{n}(\theta, y), N_0) = n_{\partial S}(\theta, y)$, for n large enough such that $4\varepsilon_n < \alpha$ thus we still have, for n large enough,

$$|I_{d-1}(\partial S) - \hat{I}_{d-1}(\partial S)| \leq \frac{1}{\beta(d)} \int_{\theta \in (\mathcal{S}^+)^{d-1}} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} |\hat{n}_{\varepsilon_n}(\theta, y) - n_{\partial S}(\theta, y)| d\mu_{d-1}(y) d\theta.$$

So, by the triangle inequality we can bound the difference between the integralgeometric and its estimation by

$$\frac{1}{\beta(d)} \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} n_{\partial S}(\theta, y) d\mu_{d-1}(y) d\theta + \frac{1}{\beta(d)} \int_{\theta \in (\mathbb{S}^+)^{d-1}} \int_{y \in \mathcal{A}_{\varepsilon_n}(\theta)} \hat{n}_{\varepsilon_n}(\theta, y) d\mu_{d-1}(y) d\theta.$$

Now, by applying (7) for the first part and a similar calculus for the second part we get that

$$|I_{d-1}(\partial S) - \hat{I}_{d-1}(\partial S)| \leq C(N_S + N_0)\varepsilon$$

for n large enough.

7.2 Proof of Theorem 3

Theorem 3 will be obtained from the two following lemmas. The first one states that eventually almost surely, the boundary of the α' -convex hull of an iid sample drawn on a α -convex support has some good geometrical properties.

The second one, which is purely geometric, bounds the difference between the measures of two sets, the first one having a positive reach α (as ∂S) and the second one having the same good geometrical properties as the boundary of $C_\alpha(\mathcal{X}_n)$.

We will introduce some notation. Let A and B be two sub-spaces of \mathbb{R}^d . We denote by $\angle(A, B)$ the operator norm of the difference between the orthogonal projection onto A , π_A , and the projection onto B , π_B , i.e., $\angle(A, B) = \|\pi_A - \pi_B\|_{op}$. If f is a function, then ∇f is its gradient and \mathcal{H}_f is its Hessian matrix. Given a point x in a $(d-1)$ -dimensional manifold E , $N_x E = \{v \in \mathbb{R}^d : \langle v, u \rangle = 0, \forall u \in T_x E\}$ is the 1-dimensional orthogonal subspace. If $A = (a_{i,j})_{i,j}$ is a matrix, $\|A\|_\infty = \max_{i,j} |a_{i,j}|$.

Lemma 6. *Let $S \subset \mathbb{R}^d$ be a compact set fulfilling the inside and outside α -rolling conditions. Let $\{X_1, \dots, X_n\}$ be an iid sample of X with distribution P_X supported on S . Assume that P_X has density f (w.r.t μ_d) bounded from below by some $f_0 > 0$. Then, for each $\alpha' \leq \alpha$, there exists an $a = a(\alpha, \alpha')$ and a $c = c(\alpha, \alpha')$ such that with probability one, for n large enough,*

1. $\partial C_{\alpha'}(\mathcal{X}_n) \cap \partial S = \emptyset$
2. $\partial C_{\alpha'}(\mathcal{X}_n) = \bigcup_{i=1}^m F_i$, where F_i is a compact $(d-1)$ -dimensional \mathcal{C}^2 manifold, for all $i = 1, \dots, m$.
3. $d_H(\partial C_{\alpha'}(\mathcal{X}_n), S) \leq \varepsilon_n^2 < \text{reach}(E)$, with $\varepsilon_n = a(\ln(n)/n)^{1/(d+1)}$.
4. $\pi_{\partial S} : \partial C_{\alpha'}(\mathcal{X}_n) \rightarrow \partial S$ the orthogonal projection onto ∂S is one to one.
5. For all $i = 1, \dots, m$ and all $x \in F_i$, $\angle(N_x F_i, N_{\pi_{\partial S}(x)} \partial S) \leq c\varepsilon_n$.

Proof. 1. Note that $\partial S \cap \partial C_{\alpha'}(\mathcal{X}_n) \neq \emptyset$ implies that $\mathcal{X}_n \cap \partial S \neq \emptyset$, which is an event with null probability, and so

$$\mathbb{P}(\partial S \cap \partial C_{\alpha'}(\mathcal{X}_n) \neq \emptyset) = 0,$$

which proves that condition 1 is fulfilled.

2. Observe that $\partial C_{\alpha'}(\mathcal{X}_n)$ is a finite union of subsets of hyper-spheres of radius α' (this is proven in Edelsbrunner et al. (1983) for dimension 2, and the generalization to any dimension is easy). This proves condition 2.
3. Recall that in Rodríguez-Casal (2007) it is proven that for any $\alpha' \leq \alpha$ there exists an a such that, with probability one for n large enough,

$$d_H(\partial C_{\alpha'}(\mathcal{X}_n), \partial S) \leq a^2 (\ln(n)/n)^{2/(d+1)}. \quad (14)$$

Hence, $d_H(\partial C_{\alpha'}(\mathcal{X}_n), \partial S) < \text{reach}(S) = \alpha$, with probability one for n large enough. This proves condition 3.

4. To prove 4 and 5 let $x \in \partial C_{\alpha'}(\mathcal{X}_n)$, and put $x^* = \pi_{\partial S}(x)$, with $\hat{\eta}_x$ the outward unit normal vector of $\partial C_{\alpha'}(\mathcal{X}_n)$ at x and η_{x^*} the outward unit normal vector of ∂S at x^* . We are going to prove that if equation (14) holds and $a^2 (\ln(n)/n)^{\frac{2}{d+1}} \leq \alpha/2$, then

$$1 - \langle \hat{\eta}_x, \eta_{x^*} \rangle \leq \frac{2(\alpha + \alpha')}{\alpha \alpha'} a^2 \left(\frac{\ln(n)}{n} \right)^{\frac{2}{d+1}}. \quad (15)$$

Put $O = x + \alpha' \hat{\eta}_x$ and $O^* = x^* - \alpha \eta_{x^*}$ (see Figure 8), we will prove that

$$\mathcal{B}(O, \alpha') \subset C_{\alpha'}(\mathcal{X}_n)^c \text{ and } \mathcal{B}(O^*, \alpha) \subset S. \quad (16)$$

To prove the first inclusion, observe that $\partial C_r(\mathcal{X}_n)$ is a union of a finite number of subsets of $\partial B(O_i, r)$ for some centres O_i , such that $B(O_i, r) \subset C_{\alpha'}(\mathcal{X}_n)^c$. Now, if $x \in \partial \mathcal{B}(O, \alpha')$ (with O one of these centres), it follows that $(O - x)/\alpha'$ is the outward unit normal vector of $\partial C_{\alpha'}(\mathcal{X}_n)$ at x , which concludes the proof. The second inclusion is a direct consequence of the inner rolling ball condition.

Write $y^* = [O^*, O] \cap \partial \mathcal{B}(O^*, \alpha)$ and $y = [O^*, O] \cap \partial \mathcal{B}(O, \alpha')$. Then, from the second inclusion in (16), we get $y \in S$, and from the first inclusion in (16) we get $d(y, C_{\alpha'}(\mathcal{X}_n)) \geq \|y - y^*\|$. This fact, combined with (14), implies that $\|y - y^*\| \leq a^2 (\ln(n)/n)^{2/(d+1)}$, which in turn implies

$$\alpha + \alpha' - \|O - O^*\| \leq a^2 \left(\frac{\ln(n)}{n} \right)^{\frac{2}{d+1}}. \quad (17)$$

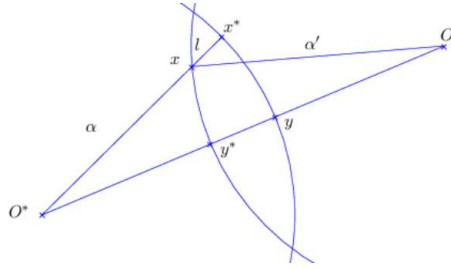


Figure 8: $x \in \partial C_{\alpha'}(\mathcal{X}_n)$, $x^* = \pi_{\partial S}(x)$, $O = x + \alpha' \hat{\eta}_x$ and $O^* = x^* - \alpha \eta_{x^*}$

From $x^* = \pi_{\partial S}(x)$ we get that $x^* = x + l\eta_{x^*}$ where $l = \|x - x^*\| \leq a^2(\ln(n)/n)^{2/(d+1)}$. Then $O = O^* + (\alpha - l)\eta_{x^*} + \alpha'\hat{\eta}_x$ and

$$\begin{aligned} \alpha + \alpha' - \|O - O^*\| &= \alpha + \alpha' - \sqrt{(\alpha')^2 + (\alpha - l)^2 + 2\alpha'(\alpha - l)\langle \hat{\eta}_x, \eta_{x^*} \rangle} \\ &= \alpha + \alpha' - \sqrt{(\alpha' + \alpha - l)^2 - 2\alpha'(\alpha - l)(1 - \langle \hat{\eta}_x, \eta_{x^*} \rangle)} \\ &= \alpha + \alpha' - (\alpha' + \alpha - l) \sqrt{1 - \frac{2\alpha'(\alpha - l)(1 - \langle \hat{\eta}_x, \eta_{x^*} \rangle)}{(\alpha' + \alpha - l)^2}} \\ &\geq l + \frac{\alpha'(\alpha - l)(1 - \langle \hat{\eta}_x, \eta_{x^*} \rangle)}{\alpha + \alpha' - l} \geq \frac{\alpha'\alpha(1 - \langle \hat{\eta}_x, \eta_{x^*} \rangle)}{2(\alpha + \alpha')}, \end{aligned}$$

where in the first inequality of the last line we bounded $\sqrt{1 - 2B/A^2} \leq A(1 - B/A^2) = A - B/A$, and in the last inequality $\alpha - l \geq \alpha/2$.

This combined with (17) proves (15).

Next we show that from (15) it follows that the hypotheses 4) and 5) in Lemma 7 are fulfilled (with probability one for n large enough). The proof of the bijectivity of $\pi_{\partial S}$ restricted to $\partial C_{\alpha'}(\mathcal{X}_n)$ follows the same ideas as those used to prove Theorem 3 in Aaron and Bodart (2016). The surjectivity follows from (14) and the rolling ball conditions, while the injectivity is a consequence of $\langle \hat{\eta}_x, \eta_{x^*} \rangle > 0$. To prove this last assertion, observe that if injectivity is not true, there exists a $y \in \partial S$ such that the half-line $\{y + t\eta_y, t \geq 0\}$ intersects $\partial C_{\alpha'}(\mathcal{X}_n)$ a first time pointing inside $C_{\alpha'}(\mathcal{X}_n)$ and then a second time 'pointing outside $C_{\alpha'}(\mathcal{X}_n)$ ' and at this second point we have $\langle \hat{\eta}_x, \eta_{x^*} \rangle \leq 0$.

Finally, Equation (15) implies that

$$\cos(\angle(\hat{\eta}_x, \eta_{x^*})) \geq 1 - \frac{2(\alpha + \alpha')}{\alpha\alpha'} a^2 \left(\frac{\ln(n)}{n} \right)^{\frac{2}{d+1}},$$

and so

$$\angle(\hat{\eta}_x, \eta_{x^*}) = \angle(N_x \partial C_{\alpha'}(\mathcal{X}_n), N_{\pi_{\partial S}(x)} \partial S) \leq 2a \sqrt{\frac{\alpha + \alpha'}{\alpha\alpha'}} \left(\frac{\ln(n)}{n} \right)^{\frac{1}{d+1}}.$$

□

Lemma 7. Let $E \subset \mathbb{R}^d$ be a compact $(d-1)$ -dimensional \mathcal{C}^3 manifold with positive reach α . Let $\varepsilon > 0$ and $\hat{E} \subset \mathbb{R}^d$ be a set such that

1. $\hat{E} \cap E = \emptyset$.
2. $\hat{E} = \bigcup_{i=1}^m F_i$, where F_i is a compact $(d-1)$ -dimensional \mathcal{C}^2 manifold, for all $i = 1, \dots, m$.
3. $d_H(\hat{E}, E) \leq \varepsilon^2 < \text{reach}(E)$.
4. $\pi_E : \hat{E} \rightarrow E$ the orthogonal projection onto E is one to one.
5. For all $i = 1, \dots, m$ and all $x \in F_i$, $\angle(N_x F_i, N_{\pi_E(x)} E) \leq c\varepsilon$.

Also, assume that ε is small enough to ensure,

$$2\varepsilon^2\alpha + (d-1)\varepsilon^4\alpha^2 + \frac{c^2\varepsilon^2}{(1-c\varepsilon)^2}(1+d\varepsilon^2)^2 < (d-1)^{-3/2}. \quad (18)$$

Then,

$$\left(1 - (d-1)^{\frac{3}{2}}(2\alpha + c^2)\varepsilon^2 + O(\varepsilon^4)\right)^{\frac{d-1}{2}} \leq \frac{|\hat{E}|_{d-1}}{|E|_{d-1}} \leq \left(1 + (d-1)^{\frac{3}{2}}(2\alpha + c^2)\varepsilon^2 + O(\varepsilon^4)\right)^{\frac{d-1}{2}}.$$

Proof. Fix $t > 0$. We will prove first that E can be partitioned into m connected sets G_1, \dots, G_m such that:

1. $|G_i \cap G_j|_{d-1} = 0$ for all $i \neq j$.
2. there exist $I(i) \in \mathbb{N}$ such that $\pi_E^{-1}(G_i) \subset F_{I(i)}$, for each $i = 1, \dots, m$.
3. for each $i = 1, \dots, m$ there exists an orthonormal basis (e_1, \dots, e_d) of \mathbb{R}^d , $H_i \subset \mathbb{R}^{d-1}$, and functions $f_i : H_i \rightarrow \mathbb{R}$, \mathcal{C}^2 such that:

$$G_i = \left\{ (x, f_i(x_1, \dots, x_{d-1})) : x = \sum_{i=1}^{d-1} x_i e_i \in H_i \right\}.$$

4. $\max_i (\max_{x \in G_i} \|\nabla f_i(x)\|_\infty) \leq t$ and $\max_i (\max_{x \in G_i} \|\mathcal{H}_{f_i}(x)\|_{op}) \leq \alpha + t$.

We provide a sketch of the proof, leaving the details to the reader. For any $x \in E$, consider the parametrization $\varphi_x : T_x E \cap \mathcal{B}(x, r_x) \rightarrow E$ such that $\nabla \varphi_x(x) = 0$ and $\mathcal{H}_{\varphi_x}(x)$ is the second fundamental form, which is bounded by α in all directions (see Proposition 6.1 in Niyogi et al. (2008)). The regularity conditions on E allow finding a radius $r_x > 0$ such that for all $y \in \mathcal{B}(x, r_x)$, $\|\nabla \varphi_x(y)\|_\infty < t$, and $\|\mathcal{H}_{\varphi_x}(y)\|_{op} < \alpha + t$. By compactness there exists a finite covering of E by balls $\mathcal{B}(x_1, r_1), \dots, \mathcal{B}(x_m, r_m)$, from which we extract only the Voronoi cells of $\{x_1, \dots, x_m\}$. Let us denote by V_i the Voronoi cell of x_i . Lastly, the family of sets $\{V_i \cap \pi_E(F_j)\}_{i,j}$ is the required partition.

We will now introduce, for $x \in H_i$, $J_i(x)$ the block matrix $J_i(x) = (I_{d-1}, \nabla f_i(x))'$. Observe that this is the Jacobian matrix of the parametrization φ_x . Also $J_i(x)' J_i(x) = I_{d-1} + \nabla f_i(x) \nabla f_i(x)'$. Now if v is any vector orthogonal to $\nabla f_i(x)$, $J_i'(x) J_i(x) v = v$, and it follows that 1 is an eigenvalue of $J_i'(x) J_i(x)$ with multiplicity $d-1$. On the other hand, $J_i'(x) J_i(x) \nabla f_i(x) = (1 + \|\nabla f_i(x)\|^2) \nabla f_i(x) = \|n_x\|^2 \nabla f_i(x)$, where $n_x = (-\nabla f_i(x), 1) \in N_{(x, f(x))} G_i$. Then,

$$|G_i|_{d-1} = \int_{H_i} \sqrt{\det J_i(x)' J_i(x)} dx = \int_{H_i} \|n_x\|_2 dx,$$

from which it follows that

$$|H_i|_{d-1} \leq |G_i|_{d-1} \leq (1+t)|H_i|_{d-1}. \quad (19)$$

Because $d_H(\hat{E}, E) < reach(E)$, by item (3) in Theorem 4.8 in Federer (1956) there exists a function l such that for all $(x, f_i(x)) \in G_i$ and $y = \pi_E^{-1}((x, f_i(x))) \in \hat{E}$, we have that $y = x + f_i(x)e_d + l(x)n_x$ with $|l(x)|/\|n_x\|_2 = d(y, E) > 0$, because $\hat{E} \cap E = \emptyset$. Then $l(x) = \|n_x\|_2 d(y, E)$ or $l(x) = -\|n_x\|_2 d(y, E)$. Since the sets F_j are of class \mathcal{C}^2 , again by item (3) in Theorem 4.8 in Federer (1956) $l(x)$ is of class at least \mathcal{C}^1 . By differentiation, for $j \in \{1, \dots, d-1\}$ let $\hat{t}_j = dy/dx_j$ be the following vector of $T_y \hat{E}$,

$$\hat{t}_j = e_j + \frac{\partial f_i}{\partial x_j}(x)e_d + \frac{\partial l}{\partial x_j}(x)n_x - l(x) \left(\sum_{k=1}^{d-1} \frac{\partial^2 f_i}{\partial x_j \partial x_k}(x)e_k \right). \quad (20)$$

This implies that

$$\|\hat{t}_j\| \leq 1 + t + d(\alpha + t)\varepsilon^2 + \left| \frac{\partial l}{\partial x_j} \right|. \quad (21)$$

Since $\hat{t}_j \in T_y \hat{E}$, $\pi_{N_y F_{I(i)}}(\hat{t}_j) = 0$, thus by Hypothesis 5 we have $\|\pi_{N_x E}(\hat{t}_j)\| \leq c\varepsilon \|\hat{t}_j\|$ that is:

$$\left| \frac{\partial l}{\partial x_j}(x) + \frac{l(x)}{\|n_x\|^2} \left(\sum_{k=1}^{d-1} \frac{\partial^2 f_i}{\partial x_j \partial x_k}(x) \frac{\partial f_i}{\partial x_k}(x) \right) \right| \leq c\varepsilon \|\hat{t}_j\|,$$

which gives that

$$\left| \frac{\partial l}{\partial x_j} \right| \leq c\varepsilon \|\hat{t}_j\| + \varepsilon^2 d(\alpha + t)t \quad (22)$$

Thus, from (21) and (22) we obtain that:

$$\|\nabla l(x)\|_\infty \leq \frac{c\varepsilon(1+t) + d(\alpha+t)\varepsilon^2(t+c\varepsilon)}{1-c\varepsilon} \quad (23)$$

This bound on $\|\nabla l\|_\infty$ allows to bound the surface estimation. Indeed, using a change of variables, it turns out that

$$|\pi_E^{-1}(G_i)|_{d-1} = \int_{H_i} \sqrt{\det(\hat{J}_i(x)' \hat{J}_i(x))} dx, \quad (24)$$

where, from (20),

$$\begin{aligned} \hat{J}_i(x) &= \begin{pmatrix} I_{d-1} - l(x)\mathcal{H}_{f_i}(x) \\ \nabla f_i(x) \end{pmatrix} + n'_x \nabla l(x) \\ &= \begin{pmatrix} I_{d-1} - l(x)\mathcal{H}_{f_i}(x) + (\nabla f(x))' \nabla l(x) \\ \nabla f_i(x) + \nabla l(x) \end{pmatrix} = \begin{pmatrix} I_{d-1} + E(x) \\ u(x) \end{pmatrix}. \end{aligned}$$

thus $\hat{J}_i' \hat{J}_i = I_{d-1} + E' + E + E'E + u'u = I_{d-1} + S_i$ where S_i is diagonalizable and $\|S_i\|_\infty \leq 2\|E\|_\infty + (d-1)\|E\|_\infty^2 + (\|\nabla f_i(x) + \nabla l(x)\|_\infty)^2$ and so, using that $\|\mathcal{H}_{f_i}(x)\|_\infty < \alpha + t$ and $\|\nabla f_i(x)\|_\infty < t$, we get,

$$\|S_i\|_\infty \leq 2(\varepsilon^2(\alpha+t) + t\|\nabla l(x)\|_\infty) + (d-1)(\varepsilon^2(\alpha+t) + t\|\nabla l(x)\|_\infty)^2 + (t + \|\nabla l(x)\|_\infty)^2. \quad (25)$$

If we combine (23) with (18), and choose t small enough to guarantee $\|S_i\|_\infty < (d-1)^{-3/2}$, then $\rho(S_i) \leq (d-1)^{3/2}\|S_i\|_\infty < 1$, $\rho(S_i)$ being the spectral radius of S_i . Indeed, let u be a unit eigenvector associated to the eigen value λ we have $S_i u = \lambda u$, and so $|\lambda|^2 = \|S_i u\|^2 = \sum_k \left(\sum_j S_{k,j} u_j \right)^2 \leq \sum_k \left(\sum_j S_{k,j} \right)^2 \|u\|^2 \leq (d-1)((d-1)\|S_i\|_\infty)^2$. From (24), we get,

$$|H_i|_{d-1} (1 - (d-1)^{\frac{3}{2}} \|S_i\|_\infty)^{\frac{d-1}{2}} \leq |\pi_E^{-1}(G_i)|_{d-1} \leq |H_i|_{d-1} (1 + (d-1)^{\frac{3}{2}} \|S_i\|_\infty)^{\frac{d-1}{2}}.$$

By (19) it follows that,

$$\frac{|G_i|_{d-1}}{1+t} (1 - (d-1)^{\frac{3}{2}} \|S_i\|_\infty)^{\frac{d-1}{2}} \leq |\pi_E^{-1}(G_i)|_{d-1} \leq |G_i|_{d-1} (1 + (d-1)^{\frac{3}{2}} \|S_i\|_\infty)^{\frac{d-1}{2}}.$$

Lastly, if we sum on i theses equations, use the uniform bound on $\|S_i\|_\infty$ obtained in (25), and take $t \rightarrow 0$, we get

$$\left(1 - (d-1)^{3/2} \varepsilon' \right)^{\frac{d-1}{2}} \leq \frac{|\hat{E}|_{d-1}}{|E|_{d-1}} \leq \left(1 + (d-1)^{3/2} \varepsilon' \right)^{\frac{d-1}{2}}.$$

$$\text{Where } \varepsilon' = 2\alpha\varepsilon^2 + (d-1)\varepsilon^4\alpha^2 + \frac{c^2\varepsilon^2}{(1-c\varepsilon)^2}(1+d\varepsilon^2)^2,$$

which concludes the proof of the Lemma. \square

7.2.1 Proof of Theorem 3

Lemma 6 proves that all the hypotheses of Lemma 7 are fulfilled (with probability one for n large enough) with $E = \partial S$, $\hat{E} = \partial C_{\alpha'}(\mathcal{X}_n)$, $\varepsilon_n = a(\ln(n)/n)^{1/(d+1)}$ and $c > 2\sqrt{(\alpha + \alpha')/(\alpha\alpha')}$. Lastly, we obtain that, with probability one, for n large enough,

$$||\partial S|_{d-1} - |\partial C_{\alpha'}(\mathcal{X}_n)|_{d-1}| = O\left(\left(\frac{\ln(n)}{n}\right)^{\frac{2}{d+1}}\right).$$

Conclusion 2 of the theorem is a consequence of Theorem 3.2.26 in Federer (1969), (see page 261).

References

- Aaron, C. and Bodart, O. (2016). Local convex hull support and boundary estimation. *Journal of Multivariate Analysis* **147** 82–101.
- Aaron, C., Cholaquidis, A., and Cuevas, A. (2017). Stochastic detection of low dimensionality and data denoising via set estimation techniques. *Electronic Journal of Statistics* **11**(2) 4596–4628.
- Alesker, S. (2018). Some conjectures on intrinsic volumes of Riemannian manifolds and Alexandrov spaces *Arnold Mathematical Journal* **4**(1) 1–17.
- Arias-Castro, E. And Rodríguez-Casal, A. (2017). On estimating the perimeter using the alpha-shape *Ann. Inst. H. Poincaré Probab. Statist.* **53**(3) 1051–1068.
- Arias-Castro, E., Pateiro-López, B., and Rodríguez-Casal, A. (2018). Minimax estimation of the volume of a set under the rolling ball condition. *Journal of the American Statistical Association*.
- Baíllo, A. and Chacón, J.E. (2018). A survey and a new selection criterion for statistical home range estimation. *preprint arXiv:1804.05129*
- Baddeley, A. J., Gundersen, H. J. G., and Cruz-Orive, L. M. (1986). Estimation of surface area from vertical sections. *J. Microsc.* **142**(3) 259–276.
- Baddeley, A. J. and Jensen, E.B. V. (2004). *Stereology for statisticians*. Chapman and Hall, London, MR2107000
- Berrendero, J.R., Cholaquidis, A., Cuevas, A. and Fraiman, R. (2014). A geometrically motivated parametric model in manifold estimation. *Statistics* **48** 983–1004.
- Bräker, H. and Hsing, T. (1998). On the area and perimeter of a random convex hull in a bounded convex set. *Probability Theory and Related Fields* **111**(4) 517–550.
- Burt, W. H. (1943). Territoriality and home range concepts as applied to mammals. *J. Mammal.*, **24**, 346–352.
- Cholaquidis, A., Fraiman, R., Lugosi, G., and Pateiro-López, B. (2016). Set estimation from reflected Brownian motion. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78**(5) 1057–1078.
- Cholaquidis, A., Fraiman, R., Mordecki, E., Papalardo, C. (2021). Level sets and drift estimation for reflected Brownian motion with drift *Statistical Sinica* 31(1) 29–51.

- Colesanti, A., and Manselli, P. (2010). Geometric and isoperimetric properties of sets of positive reach in E^d . *Atti Semin Mat Fis. Univ. Modena Reggio Emilia* **57** 97–113.
- Crofton, M. W. (1868). On the theory of local probability, applied to straight lines drawn at random in a plane: The methods used being also extended to the proof of certain new theorems in the integral calculus *Philosophical Transactions of the Royal Society of London* **158** 181–199.
- Cuevas, A. and Rodríguez-Casal, A. (2004). On boundary estimation. *Adv. in Appl. Probab.* **36** 340–354.
- Cuevas, A., Fraiman, R., and Rodríguez-Casal, A. (2007). A nonparametric approach to the estimation of lengths and surface areas. *Ann. Statist.* **35**(3) 1031–1051.
- Cuevas, A., Fraiman, R., and Pateiro-López, B. (2012). On statistical properties of sets fulfilling rolling-type conditions. *Adv. in Appl. Probab.* **44** 311–329.
- Cuevas, A., Fraiman, R., and Györfi, L. (2013). Towards a universally consistent estimator of the Minkowski content. *ESAIM: Probability and Statistics* **17** 359–369.
- Devroye, L. and G. L. Wise (1980). Detection of abnormal behavior via nonparametric estimation of the support. *SIAM J. Appl. Math.* **3** 480–489.
- Edelsbrunner, E., Kirkpatrick, D., and Seidel, R. (1983). On the shape of points in the plane. *IEEE Transaction on Information Theory* **29** 551–559.
- Federer, H. (1956). Curvature measures. *Transactions of the American Mathematical Society* **93**(3) 418–491.
- Federer, H. (1969). Geometric Measure Theory *Springer-Verlag*.
- Gokhale, A.M.(1990) Unbiased estimation of curve length in 3D using vertical slices. *J. Microsc.* **195** 133–141.
- Jiménez, R. and Yukich, J.E. (2011). Nonparametric estimation of surface integrals. *Ann. Statist.* **39** 232–260.
- Niyogi, P., Smale, S., and Weinberger, S. (2008). Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.* **39** 419–441.
- Pateiro-López, B. and Rodríguez-Casal, A. (2008). Length and surface area estimation under smoothness restrictions. *Advances in Applied Probability* **40**(2) 348–358.
- Pateiro-López, B. and Rodríguez-Casal, A. (2009). Surface area estimation under convexity type assumptions *Journal of Nonparametric Statistics* **21**(6) 729–741.
- Kim, J.C. and Korostelëv, A. (2000). Estimation of smooth functionals in image models. *Math. Methods Statist.* **9**(2) 140–159.
- Korostelëv, A.P. and Tsybakov, A.B. (1993). Minimax Theory of Image Reconstruction. Lecture Notes in Statistics. *Springer-Verlag, New York*.
- Penrose M.D. (2021). Random Euclidean coverage from within. *preprint arXiv:2101.06306*
- Rodríguez-Casal, A. (2007) Set estimation under convexity type assumptions *Annales de l'Institut Henri Poincaré (B): Probability and Statistics* **43** 763–774.

- Rodríguez-Casal, A. and Saavedra-Nieves, P. (2019). Extent of occurrence reconstruction using a newdata-driven support estimator *preprint arXiv:1907.08627*
- Santaló, L. A. (2004). Integral Geometry and Geometric Probability. *Cambridge University Press*.
- Thäle, C. And Yukich J.E. (2016). Asymptotic theory for statistics of the Poisson–Voronoi approximation *Bernoulli* **22**(4) 2372–2400.
- Walther, G. (1999). On a generalization of Blaschke’s rolling theorem and the smoothing of surfaces, *Math. Meth. Appl. Sci.* **22** 301–316.