



HAL
open science

Set estimation under dependence

Catherine Aaron, Paul Doukhan, Laurence Rebol

► **To cite this version:**

Catherine Aaron, Paul Doukhan, Laurence Rebol. Set estimation under dependence. 2024. hal-04375650v1

HAL Id: hal-04375650

<https://hal.science/hal-04375650v1>

Preprint submitted on 5 Jan 2024 (v1), last revised 27 Sep 2024 (v3)

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.

Set estimation under dependence

C. Aaron^a, P. Doukhan^b, L. Reboul^c

^aUniversity Clermont Auvergne

^bCY Cergy Paris University, AGM Mathematics, 2 Bd. Adolphe Chauvin, 95000 Cergy-Pontoise, AGM-UMR 8080

^cInstitute of Mathematics of Marseille, 13288 Marseille Cedex 9, UMR 7373 and CY Cergy Paris University, AGM Mathematics

Abstract

In this paper, we survey and generalize results on the R -convex hull of the points of a random sample. We provide explicit convergence rates for the R -convex hull and its natural plug-in, as estimates (estimators ?) of the support S , its volume and perimeter, and level-sets of a density $f \in \mathbb{R}^d$, when the underlying sequence is strictly stationary and admits some kind of weak dependence condition and when the density satisfies $f(x) \geq cd(x, \partial S)^\alpha$ for all $x \in S$. We obtain convergence rates that exceed those obtained in the i.i.d. case which is of order $(\ln n/n)^{\frac{2}{d+1+2\alpha}}$. The loss depends on the dependency condition, it can be an extra power of $\ln n$ or an extra power of n in the worst cases.

Keywords: Strictly stationary process, Linear processes, density, support

1. Introduction

Set estimation has many applications. In ecology for instance, it may be used to study the home-range or the core-area of species [19], based on the observation of the spatial distribution of species or the movement of one or more representative members. Assimilating the position of a member to a random variable with given distribution $f : \mathbb{R}^d \rightarrow \mathbb{R}$, the home-range corresponds to the support, $S = \overline{\{x, f(x) > 0\}}$, of f and the core-area to a level set of f , defined, for a given threshold $t > 0$ as the set $L_t = \{x, f(x) \geq t\}$. Setting $\int_{L_t} f(x)dx = \alpha$, L_t can be seen as a multidimensional α -quantile.

The estimation of a density support S and of its related quantities -level sets, boundary, volume and perimeter- has been the focus of several works for iid data sets. The historical Devroye-Wise estimator of S has been studied by [10, 12] (see also references therein). It has universal properties but non-optimal rates under reasonable regularity hypotheses. When S is convex, optimal rates are achieved by the convex hull of the sample. More generally, under "convexity type" assumptions on S , optimal rates may be achieved by estimators defined as some extensions of the convex hull. This is the case of the R -convex hull, studied in [2, 5, 26–28, 31], among other (other estimators of S have also been proposed, as in [1] or [20]). The R -convex hull of a set S is a generalization of its convex hull, defined as follows :

$$C_R(S) = \left(\bigcup_{\dot{B}(x,R) \cap S = \emptyset} \dot{B}(x,R) \right)^c,$$

where E^c and \dot{E} respectively denote the complement and interior of a set E and $B(x,R)$ is the closed ball of radius R centered at x . The use of the R -convex hull to estimate the support of a density f has an heuristic justification: when the set S is regular enough, we have that $C_R(S) = S$ so that a natural estimator of S is $\widehat{S}_n = C_R(\mathbb{X}_n)$, with $\mathbb{X}_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$ is a sample of f and $C_R(\mathbb{X}_n) = \left(\bigcup_{\dot{B}(x,R) \cap \mathbb{X}_n = \emptyset} \dot{B}(x,R) \right)^c$.

The study of \widehat{S}_n has been introduced by [31] and [26]. They derive convergence rates of the Hausdorff distance between S and its estimator ($d_H(\widehat{S}_n, S)$) towards 0 and obtain the minimax rate achieved by [20] under regularity

Email addresses: catherine.aaron@uca.fr (C. Aaron), doukhan@cyu.fr (P. Doukhan), laurence.reboul@univ-amu.fr (L. Reboul)

assumptions on S and when f is bounded below by a positive constant (similar rates are obtained for $d_H(\partial\widehat{S}_n, S)$). In [28], a data-driven way to tune the parameter R is proposed and applied to home-range estimation of the extent of occurrence of invasive plant species in the Azores archipelago. The R -convex hull may also be used to estimate related quantities, such as the volume (see [5] and [6]) and perimeter (see [2]) of S . In [27], the authors propose an estimator of level sets derived from the R -convex hull of conveniently chosen sample points with a data-driven R -parameter.

In all those papers, the sample \mathbb{X}_n is i.i.d. and the density is bounded below on S by some positive constant. However, for real life applications, the independence hypothesis is generally unrealistic. This is the case for instance when a member's location is observed via GPS or when the sample consists of the trajectory of one or more members observed over a time period. A more realistic hypothesis should be to consider the strict stationarity of the sequence \mathbb{X}_n as well as some kind of short memory condition. Set estimation under dependence has been little studied and the two related works are to our knowledge that of [8], who considers trajectories of a reflected Brownian motion and the recent paper of [21] in which the authors extend results on reconstruction of probabilistic supports of random i.i.d variables to supports of dependent stationary \mathbb{R}^d -valued random variables with the use of the Devroye-Wise estimator.

In this article, we study the asymptotic properties of \widehat{S}_n , its related quantities and estimators of and level sets estimators derived from \widehat{S}_n in a non i.i.d. framework, extending the results of [2, 5, 6, 26, 27, 31]. Namely, we derive convergence rates in Hausdorff distance of estimators of support, its boundary, level sets and their boundary. We also derive rates for related estimators of the volume and perimeter of the support and its level sets. Finally we show that, with probability one and for large enough n the estimated support (resp. level sets) is homeomorphic to the support (resp. the level set). Regarding the dependency, we assume that \mathbb{X}_n is a strictly stationary sequence that admits some kind of short memory property. We consider here strong mixing and α -dependence, introduced respectively by [29] and [11].

The rest of the paper is organized as follows. Section 2 presents some notation and general results that will be used in the sequel. Section 3 concerns asymptotic results on the R -convex hull and its natural plug-in as estimators of the density support and its related quantities, volume and perimeter. Section 4 deals with the estimation of level-sets. The proofs are postponed to Section 5.

2. General framework

In this paper, we are interested in several geometrical nonparametric estimation problems, based on a strictly stationary weak dependent sequence $\mathbb{X}_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$. In this section, we first give some notation and preliminary material that will be useful to settle and prove our main results.

2.1. Notation

- $B(x, r) \subset \mathbb{R}^d$ denotes the closed ball of radius $r \geq 0$ centered at x and $S(x, r)$ is its boundary, i.e. the sphere of radius r centered at x .
- For each set $E \subset \mathbb{R}^d$, we denote by $|E|_d$ its d -dimensional volume and $|\partial E|_{d-1}$ the $(d-1)$ -dimensional volume of its boundary.

For the sake of simplicity, we set $\omega_d = |B(0, 1)|_d$ the volume of $B(0, 1) \subset \mathbb{R}^d$ and $\sigma_d = |S(0, 1)|_{d-1}$ its surface (i.e. the $(d-1)$ -dimensional volume of its boundary).

- For each set $E \subset \mathbb{R}^d$ and positive number r , we denote by $E \oplus rB$ and $E \ominus rB$ the Minkowski sum and difference, namely $E \oplus rB = \{x \in \mathbb{R}^d, d(x, E) \leq r\} = \bigcup_{a \in E} B(a, r)$, $F \ominus rB = \{x, B(x, r) \subset F\}$.
- For two subsets E and F of \mathbb{R}^d , we denote by $d_H(E, F)$ their Hausdorff distance:

$$d_H(E, F) = \max(\sup_{a \in E} d(a, F), \sup_{b \in F} d(b, E)).$$

- We set $E \approx F$ if and only if E and F are homeomorphic.

- For two non-negative sequences (u_n) and (v_n) in $\mathbb{R}^{\mathbb{N}}$, we set $u_n \lesssim v_n$ if and only if there exists $c > 0$ such that, for all $n \geq 0$, $u_n \leq cv_n$.
- Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a differentiable function, we denote by ∇_f its gradient and by \mathcal{H}_f its Hessian matrix if it is twice differentiable.
- We denote by $\|\cdot\|_{op}$ the operator norm on the set of $n \times p$ matrices.
- For $k \in \mathbb{N}$, we denote by C^k the set of k -times continuously differentiable functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$.
- Let $\rho \geq 1$. We say that K is a kernel of order ρ if and only if $\int P(x)K(x)dx = P(0)$ for each polynomial of degree less than or equal to ρ .

2.2. Dependence background

In the sequel, we make use of a strictly stationary short memory process $\mathbf{X} = (X_t)_{t \in \mathbb{Z}}$. We consider $\mathbb{X}_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$ a sample of \mathbf{X} with marginal density $f : \mathbb{R}^d \rightarrow \mathbb{R}$.

We first recall below the short memory properties considered in the sequel, namely the strong mixing and θ -dependence conditions as well as some covariance properties and density estimation results for each of them.

2.2.1. Weak dependence conditions

Let $\mathbf{X} = (X_t)_{t \in \mathbb{Z}}$ be a strictly stationary sequence on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $-\infty \leq j \leq l \leq +\infty$, let the σ -field $\mathcal{F}_j^l = \sigma(X_k, 1 \leq k \leq l, k \in \mathbb{Z})$ of events generated by the random variables $(X_k)_{j \leq k \leq l}$. In the sequel, we consider the following dependence conditions for \mathbf{X} .

Strong or α -mixing. Let $(\alpha(q))_{q>0}$ be the decreasing sequence of strong mixing coefficients of $\mathbf{X} = (X_t)_{t \in \mathbb{Z}}$, defined as

$$\alpha(q) = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_q^{\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

Quote that the above expression also writes as the supremum of $|\text{cov}(\mathbf{1}_A, \mathbf{1}_B)|$. We say that \mathbf{X} is strongly mixing (or α -mixing) if $\lim_{q \rightarrow \infty} \alpha(q) = 0$.

θ -weak dependence. Let $\nu \geq 1$ and let us define the Lipschitz modulus of a function $h : (\mathbb{R}^d)^\nu \rightarrow \mathbb{R}$ by

$$\text{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x_1 - y_1\| + \dots + \|x_\nu - y_\nu\|},$$

where $\|\cdot\|$ is a norm of \mathbb{R}^d . Let W be any random variable such that $\|W\|_\infty \leq 1$ measurable with respect to $\mathcal{F}_{-\infty}^0$. We say that \mathbf{X} is θ -weakly dependent if there exists a decreasing sequence $(\theta(q))_{q>0}$ with $\lim_{q \rightarrow \infty} \theta(q) = 0$ and a set of indices $i_1, \dots, i_\nu \geq 1$ such that $|\text{cov}(W, h(X_{i_1}, \dots, X_{i_\nu}))| \leq \nu \text{Lip } h \cdot \theta(q)$. Notice that for $\nu = 1$ we have $|\text{cov}(W, h(X_q))| \leq \text{Lip } h \cdot \theta(q)$.

Examples. A large number of classical models which satisfy strong mixing conditions is provided in [13]; in particular the class of linear processes based on innovations with an absolutely continuous density function, such as strictly stationary ARMA processes and also certain bilinear, ARCH or GARCH models. Another example is that of a stationary Gaussian process with a continuous positive spectral density. Notice also that this property remains valid through images by measurable functions.

Some classical processes are not mixing. A famous counterexample is in [4], in which \mathbf{X} is the stationary solution of

$$X_t = \frac{1}{2}(X_{t-1} + \varepsilon_t),$$

where X_0 is independent of $(\varepsilon_t)_{t>0}$, which is a sequence of i.i.d. random variables with Bernoulli distribution of parameter $1/2$. This example and many others, such as linear processes with discrete innovations [4], on the other hand, also fit the θ -weak dependence condition. Notice also that this property remains valid through images by bounded Lipschitz function.

2.2.2. A covariance based result

Let E be a subset of \mathbb{R}^d with $\mathbb{P}(X_t \in E) = p > 0$. In each of the above-mentioned dependence conditions, under covariance based arguments we derive a useful upper-bound for $\mathbb{P}(\mathbb{X}_n \cap E = \emptyset)$.

Proposition 1. *Let $(X_t)_{t \in \mathbb{Z}}$ be a strongly mixing sequence with mixing coefficients $(\alpha(q))_{q>0}$. Then, setting $\mathcal{E}_E(q) = \alpha(q)$, we have*

$$\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \leq (1-p)^{\frac{n}{q}} + \frac{\mathcal{E}_E(q)}{p}.$$

Proposition 2. *Let $(X_t)_{t \in \mathbb{Z}}$ be a θ -weakly dependent sequence with dependence coefficients $(\theta(q))_{q>0}$. Assume moreover that there exists $c > 0$ such that for each measurable set E :*

$$\mathbb{P}(X_t \in E) \leq c|E|_d. \quad (1)$$

If now,

1. If $E = B(x, r)$ we have that

$$\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \leq (1-p)^{\frac{n}{q}} + \frac{\mathcal{E}_E(q)}{p} \quad \text{with} \quad \mathcal{E}_E(q) \leq \inf_{\varepsilon \in (0, r)} \left(\frac{\theta(q)}{\varepsilon} + c\omega_d 2^d \varepsilon r^{d-1} \right) \quad (2)$$

2. If $E = B(x, r_x) \cap B(y, r_y)$ with $r_x + r_y - \|x - y\| = \ell$, and $\ell \leq r_x \wedge r_y$, we have that

$$\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \leq (1-p)^{\frac{n}{q}} + \frac{\mathcal{E}_E(q)}{p} \quad \text{with} \quad \mathcal{E}_E(q) \leq \inf_{\varepsilon \in (0, \ell/2)} \left(\frac{2\theta(q)}{\varepsilon} + c^c \mathcal{V}(r_x, r_y, \ell, \varepsilon) \right) \quad (3)$$

$$\text{and} \quad \mathcal{V}(r_x, r_y, \ell, \varepsilon) \underset{\ell \rightarrow 0}{\sim} 4c\omega_{d-1} \varepsilon \left(\frac{\ell 2 r_x r_y}{r_x + r_y} \right)^{\frac{d-1}{2}}. \quad (4)$$

Remark 1. Notice that condition (1) holds for instance if the distribution of X_0 admits a bounded density f wrt Lebesgue measure.

Remark 2. Our results may be extended to alternative dependence structures, yielding other values of $\mathcal{E}_E(q)$, possibly involving other methods of proofs, such as coupling.

- An alternative to the above result involves couplings. For instance if \mathbf{X} is absolutely regular, see [13], Berbee's coupling inequalities ([7]) yield $\mathcal{E}_E(q) \leq \beta(q)$.
- Another case of interest is that of negatively associated processes, see [30]. Considering the proof, we set in this case $q = 1$ and $k = n$ and here

$$\text{cov}(Z_1 \cdots Z_{k-1}, Z_k) \leq 0.$$

In this case blocks are useless (here $q = 1$) and $p_n \geq p^n$ and the results are conducted as in the classical independent case.

The above upper bounds for $\mathbb{P}(\mathbb{X}_n \cap E = \emptyset)$ will allow to derive convergence rates for the estimation of S , ∂S , $|S|_d$ and $|\partial S|_{d-1}$ in Section 3. In order to get explicit rates, we will propose two scenarii

- *Exponential decays*, $\alpha(m) \leq Bb^{-m}$ or $\theta(m) \leq Bb^{-m}$, for some $b > 1$.
- *Riemannian decays*, $\alpha(m) \leq Bm^{-b}$ or $\theta(m) \leq Bm^{-b}$, for some $b > 0$.

2.2.3. Uniform consistency of kernel density estimators under dependence

To obtain results on level set estimation, we need an upper bound on

$$\epsilon'_n = \sup_{x \in \mathbb{R}^d} |\widehat{f}_n(x) - f(x)|,$$

where $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is the marginal density of the strictly stationary \mathbb{R}^d -valued process \mathbf{X} and \widehat{f}_n is a kernel density estimator of f :

$$\widehat{f}_n(x) = \frac{1}{nh_n^d} \sum_{t=1}^n K\left(\frac{X_t - x}{h_n}\right). \quad (5)$$

We provide below bounds for ϵ'_n in the α -mixing and θ -weakly dependent cases. Proposition 3 concerns exponential decays and Proposition 4 concerns Riemannian decays. For that task, we first introduce some regularity assumptions on f and \widehat{f}_n :

- (K1) $K : \mathbb{R}^d \rightarrow \mathbb{R}$ is a bounded square-integrable density with compact support.
- (K2) $K : \mathbb{R}^d \rightarrow \mathbb{R}$ is of order $(k - 1)$.
- (M1) For some $k \geq 2$, $f \in C^k$ and $f^{(k)}$ is bounded.
- (M2) The marginal density f is bounded on \mathbb{R}^d .
- (M3) The joint densities f_i of (X_0, X_i) are bounded, uniformly with respect to $i > 0$.
- (M4) The marginal density f tends to zero as n tends to infinity.
- (M5) $\mathbb{E}\|X_0\|^c < \infty$ for some $c > 2$.

We thus have

Proposition 3. *Let \mathbb{X}_n be a strictly stationary sequence from \mathbf{X} and \widehat{f}_n be defined by (5). Assume that (K1), (M2), (M3), (M4) and (M5) hold, and that (K2) and (M1) hold for some $k \geq 2$.*

(A1) *If \mathbf{X} is strong mixing with dependence coefficients $\alpha(m) = O(b^{-m})$ for some $b > 1$, thus*

1. *If $k > d + 3$, the choice $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1}{d+2k}}$ yields $\epsilon'_n = O_{a.s.}\left(\left(\frac{\log n}{n}\right)^{k/(2k+d)}\right)$.*
2. *If $k \leq d + 3$, any choice of $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1-\gamma}{3d+6}}$ with $\gamma \in (0, 1)$ yields $\epsilon'_n = O_{a.s.}\left(\left(\frac{\log n}{n}\right)^{k(1-\gamma)/(3d+6)}\right)$.*

(T1) *If \mathbf{X} is θ -weakly dependent with dependence coefficients $\theta(m) = O(b^{-m})$ for some $b > 1$, thus*

1. *If $k > d + 2$ the choice $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1}{d+2k}}$ yields $\epsilon'_n = O_{a.s.}\left(\left(\frac{\log n}{n}\right)^{k/(2k+d)}\right)$.*
2. *If $k \leq d + 2$ any choice of $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1-\gamma}{3d+4}}$ with $\gamma \in (0, 1)$ yields $\epsilon'_n = O_{a.s.}\left(\left(\frac{\log n}{n}\right)^{k(1-\gamma)/(3d+4)}\right)$.*

Now, in case of (plus forte) dependence with coefficient with Riemannian decay we have the following result.

Proposition 4. *Let \mathbb{X}_n be a strictly stationary sequence from \mathbf{X} and \widehat{f}_n be defined by (5). Assume that (K1), (M2), (M3), (M4) and (M5) hold, and that (K2) and (M1) hold for some the same $k \geq 2$.*

(A2) *If \mathbf{X} is strong mixing with dependence coefficients $\alpha(m) = O(m^{-b})$ for some $b > \max(2 + 3/d, d + 2)$. Let*

$$\gamma_0 = \frac{d+2 - \frac{d(d-2)}{2k+d}}{b+1 - \frac{d(d-2)}{2k+d}}. \text{ Then the choice of } h_n = O\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\frac{1}{d+2k}}, \text{ for any } \gamma \in (\gamma_0, 1) \text{ yields } \epsilon'_n = O_{a.s.}\left(\left(\frac{\log n}{n^{1-\gamma}}\right)^{k/(2k+d)}\right).$$

(T2) *If \mathbf{X} is θ -weak dependent with dependence coefficients with $\theta(m) = O(m^{-b})$ for some $b > 3(d + 2)$, let $\gamma_0 =$*

$$\frac{d+2 - \frac{d(d-2)}{2k+d}}{b/3 + 1 - \frac{d(d-2)}{2k+d}} \text{ then the choice of } h_n = O\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\frac{1}{d+2k}}, \text{ for any } \gamma \in (\gamma_0, 1) \text{ yields } \epsilon'_n = O_{a.s.}\left(\left(\frac{\log n}{n^{1-\gamma}}\right)^{k/(2k+d)}\right).$$

3. Support Estimation

In this section we give the convergence rates of the R -convex hull estimator of S and that of their boundaries in the context of strictly stationary dependent strong mixing or θ -weak dependent sequences \mathbb{X}_n .

3.1. Distributional Background

In order to obtain a consistent support estimator $C_R(\mathbb{X}_n)$ of S we will ask the distribution to have regular enough compact support S to have $C_R(S) = S$. This will be ensured by sets satisfying the rolling ball property defined as follow

DEFINITION (ROLLING BALL PROPERTY). A closed compact set $S \subset \mathbb{R}^d$ satisfies the R_o -outside (resp. R_i -inside) rolling ball property if and only if $\forall x \in \partial S, \exists O_x^o \in \mathbb{R}^d, \|x - O_x^o\| = R_o$ and $\mathring{B}(O_x^o, R_o) \subset \mathring{S}^c$. (resp. $\exists O_x^i \in \mathbb{R}^d, \|x - O_x^i\| = R_i$ and $B(O_x^i, R_i) \subset S$).

When S satisfies the R_o -outside (R_i -inside) rolling ball property, we also say that balls of radius R_o (resp. R_i) roll freely outside (resp. inside) S . Such sets have nice properties due to the generalization of the Blaschke's rolling theorem (cf [31]'Theorem 1). As aforementioned sets with rolling ball property satisfy that $S_{R_o}(S) = S$, this ensures the convergence of the plugin support estimator $S_{R_o}(\mathbb{X}_n)$ since $d_H(\mathbb{X}_n, S) \rightarrow 0$. In order to derive explicit convergence rates, we have to complete this assumption with another on the density. Specifically we will require the density to have an a polynomial decay ($f(x) \geq f_0 d(x, \partial S)^\alpha$ for all $x \in S$) as proposed in [20].

DEFINITION. A density f belongs to the class $\mathcal{M}_{R_o, R_i, f_0, \alpha}$ if it has a compact support S that satisfies the R_o -outside rolling ball property and the R_i -inside rolling ball property. Also, for all $x \in S$ we have $f(x) \geq f_0 d(x, \partial S)^\alpha$.

Notice that the special case of $\alpha = 0$ corresponds to the classical class of lower bounded densities on S , known as "close to uniform" hypothesis.

3.2. Main Theoretical results

In this section we give explicit convergence rates. This rates generalize those obtained in previous works on the R -convex hull in the independent case with $\alpha = 0$. Namely, [26] give rates for $d_H(\partial C_R(\mathbb{X}_n), \partial S)$ and $d_H(C_R(\mathbb{X}_n), S)$, while [2] gives rates for $|\partial C_R(\mathbb{X}_n)|_{d-1} - |\partial S|_{d-1}$ and $\partial C_R(\mathbb{X}_n) \approx \partial S$ e.a.s.

Also notice that, in these works author choose to work with sets S satisfying the \bar{R} inside and outside rolling ball property. As in [27] we propose to emphasize the different roles of the outside (resp. inside) radius R_o (resp. R_i). The tuning parameter R must be less than R_o to have the convergence and the convergence rate depends on R_i .

The following results recall that obtained in the aforementioned works. Our contribution consists in generalizing these results to $\alpha \geq 0$. Nevertheless it seems important to present the independent reference rate.

Theorem 1. *Let R_i, R_o and f_0 be positive and α be nonnegative. Let $\mathbb{X}_n = \{X_1, \dots, X_n\}$ be an iid n -sample with density $f \in \mathcal{M}_{R_o, R_i, f_0, \alpha}$ supported by S and $R < R_o$. There exists $A(d, \alpha, f_0, R, R_i)$ such that for n large enough, one has with probability 1*

$$\begin{aligned} d(\partial C_R(\mathbb{X}_n), \partial S) &\leq \varepsilon_n & , & & d(C_R(\mathbb{X}_n), S) &\leq \varepsilon_n \\ \partial C_R(\mathbb{X}_n) &\approx \partial S & , & & C_R(\mathbb{X}_n) &\approx S \\ \|\partial C_R(\mathbb{X}_n)\|_{d-1} - \|\partial S\|_{d-1} &\lesssim \varepsilon_n & , & & \|C_R(\mathbb{X}_n)\|_d - \|S\|_d &\lesssim \varepsilon_n. \end{aligned}$$

with

$$\varepsilon_n = A(d, \alpha, f_0, R, R_i) \left(\frac{\ln n}{n} \right)^{\frac{2}{d+1+2\alpha}}$$

Remark that in case of negatively associated process as in [30] we can obtain the same rates as in the iid case.

In Theorem 2, which deals with exponential decay rates, we obtain, up to an additional power of $\ln n$, the same rates as in the independent case whatever the dependence structure (strong mixing or θ -dependence) is. Indeed the dependence structure only impacts constants.

Theorem 2. Let R_i, R_o and f_0 be positive and α be nonnegative. Let $(X_t)_{t \in \mathbb{Z}}$ be a strictly stationary sequence with marginal density $f \in \mathcal{M}_{R_o, R_i, f_0, \alpha}$ supported by S . Let $R < R_o$. Assume that $(X_t)_{t \in \mathbb{Z}}$ is strong mixing or θ -weakly dependent sequence with an **exponential decay** of the sequence of dependence coefficients. Then, there exists $A(d, \alpha, f_0, R, R_i, b)$ such that for large enough n , one has with probability 1

$$\begin{aligned} d(\partial C_R(\mathbb{X}_n), \partial S) &\leq \varepsilon_n, & d(C_R(\mathbb{X}_n), S) &\leq \varepsilon_n \\ \partial C_R(\mathbb{X}_n) &\approx \partial S, & C_R(\mathbb{X}_n) &\approx S \\ \|\partial C_R(\mathbb{X}_n)\|_{d-1} - \|\partial S\|_{d-1} &\lesssim \varepsilon_n, & \|C_R(\mathbb{X}_n)\|_d - \|S\|_d &\lesssim \varepsilon_n. \end{aligned}$$

with

$$\varepsilon_n = A(d, \alpha, f_0, R, R_i, b) \left(\frac{(\ln n)^2}{n} \right)^{\frac{2}{d+1+2\alpha}}$$

In Theorem 3 below, we provide rates in the case of Riemannian decay of the dependence coefficients. This time, the loss in convergence rate consists of an additional power of n . The rate and the loss also depend on the dependency structure.

Theorem 3. Let R_i, R_o and f_0 be positive and α be nonnegative. Let $(X_t)_{t \in \mathbb{Z}}$ be a strictly stationary sequence with marginal density $f \in \mathcal{M}_{R_o, R_i, f_0, \alpha}$ supported by S . Let $R < R_o$. Assume that $(X_t)_{t \in \mathbb{Z}}$ is strong mixing or θ -weakly dependent sequence with an **Riemannian decay** of the sequence of dependence coefficients. Then, there exists b_0 such that, if $b > b_0$ then, for large enough n , one has with probability 1

$$\begin{aligned} d(\partial C_R(\mathbb{X}_n), \partial S) &\leq \varepsilon_n, & d(C_R(\mathbb{X}_n), S) &\leq \varepsilon_n \\ \partial C_R(\mathbb{X}_n) &\approx \partial S, & C_R(\mathbb{X}_n) &\approx S \\ \|\partial C_R(\mathbb{X}_n)\|_{d-1} - \|\partial S\|_{d-1} &\lesssim \varepsilon_n, & \|C_R(\mathbb{X}_n)\|_d - \|S\|_d &\lesssim \varepsilon_n. \end{aligned}$$

With $\varepsilon_n = ((\ln n)^a / n^{1-\gamma})^{2/(d+1+2\alpha)}$ explicit values of b_0 , a and γ are

1. Strong mixing $b_0 = 1$, $\lambda = \frac{3d+1+2\alpha}{d+1+2\alpha}$, $a = \frac{b-2-d}{b-\lambda}$ and $\gamma = \frac{\lambda+1}{\lambda+b}$.
2. θ -weak dependence $b_0 = 2$, $\lambda = \frac{3d+1+2\alpha}{d+1+2\alpha}$, $a = \frac{b/2-2-d}{b/2-\lambda}$ and $\gamma = \frac{\lambda+1}{\lambda+b/2}$.

4. Level set estimation

We now shift our focus to level set estimation. Recall that for a given $t \in \mathbb{R}$, the level set L_t is defined as $L_t = \{z \in \mathbb{R}^d, f(z) \geq t\}$ and its associated level line ℓ_t is defined by $\ell_t = \{z, f(z) = t\}$. Given a suitable density estimator \widehat{f}_n of f , [27] introduced the following level set estimator

$$\widehat{L}_{t,R} = C_R(\mathbb{X}_n \cap \{z, \widehat{f}(z) \geq t\})$$

and studied its asymptotic properties in the i.i.d. case. We aim at generalize the rates to dependence setup.

4.1. Distributional background

When dealing with level set estimation, the distributional hypotheses are somewhat different. The support no longer needs to be compact since the set $L_t = \{z, f(z) > t\}$ is compact when $t > 0$ (when $t = 0$ we are in the support estimation framework). Also, the rate of decay of f near the boundary is no longer important. As a counterpart, we require that $f \in C^k$ for some $k \geq 2$ which has two important consequences. Firstly, it allows the density estimator to converge with suitable rates. Secondly, with implicit function theorem, it enables L_t to satisfy the inside and outside rolling ball property for suitable values of t . The explicit condition to have suitable level is the following

DEFINITION. [Regular levels] $t > 0$ is a regular level if there exists $\Delta_t > 0$ and $m_t > 0$ such that

$$L_{t-\Delta_t} \subset \overset{\circ}{S} \text{ and } \min_{z \in L_{t-\Delta_t} \setminus \overset{\circ}{L}_{t+\Delta_t}} \|\nabla f(z)\| = m_t > 0.$$

When t is a regular level and f is C^k ($k \geq 2$), then by Theorem 2 in [31], the level set L_t naturally inherits of the inside and outside rolling ball property.

Proposition 5. Corollary of Theorem 2 in [31] *If $f \in C^2$ and $\sup_z \|\mathcal{H}_f\| = f_{\max}^{(2)} < +\infty$ and if t is a regular level, then there exists δ_t such that, for all $t' \in]t - \delta_t, t + \delta_t[$, $L_{t'}$ has the inside and outside rolling ball property for balls of radius $\frac{m_t}{f_{\max}^{(2)}}$.*

4.2. Main Theoretical results

Hereafter, we establish the asymptotic behavior of $\widehat{L}_{t,R}$ under weak dependence *Assumptions* and deduce the asymptotics of its boundary and volume. By incorporating tools provided in [2], we also obtain topological guarantees and perimeter estimation rates. We give explicit convergence rates in the particular cases of exponential and Riemannian decay of the weak dependence coefficients. We show that the convergence rate can be decomposed into a sum $\varepsilon_n + \varepsilon'_n$, where ε_n is a set estimation error and ε'_n is a density estimation loss which depends on the dependence structure and coefficients, the regularity of the density and the dimension d (see section 2.2.3). We focus in this section on the geometric part ε_n . Some explicit rates taking into account all the parameters of the model are given in following section.

Theorem 4. *Let $\mathbb{X}_n = \{X_1, \dots, X_n\}$ be a n -sample of a strictly stationary sequence of a strong mixing or θ -weak dependent sequence **with exponential decay rates** and with marginal class C^2 density f on its support S . Assume that $\sup_S \|\nabla f\| = f_{\max}^{(1)} < +\infty$, $\sup_S \|\mathcal{H}_f\|_{op} = f_{\max}^{(2)} < +\infty$. Let $t > 0$ be a regular level. Let R be a constant such that $0 < R < m_t / f_{\max}^{(2)}$ and define $\mathbb{X}_n^+(t) = \mathbb{X}_n \cap \{f(z) \geq t\}$ and $\widehat{L}_{t,R} = C_r(\mathbb{X}_n^+(t))$. Then, with probability 1 and for n large enough, one has*

$$\begin{aligned} d(\widehat{\partial L}_{t,R}, \partial L_t) &\lesssim \varepsilon_n + \varepsilon'_n, & d(\widehat{L}_{t,R}, L_t) &\lesssim \varepsilon_n + \varepsilon'_n \\ \widehat{\partial L}_{t,R} &\approx \partial L_t, & \widehat{L}_{t,R} &\approx L_t \\ \|\widehat{L}_{t,R}\|_d - |L_t|_d &\lesssim \varepsilon_n + \varepsilon'_n, & \|\widehat{\partial L}_{t,R}\|_{d-1} - |\partial L_t|_{d-1} &\lesssim \varepsilon_n + \varepsilon'_n. \end{aligned}$$

$$\text{with } \varepsilon_n \lesssim ((\ln n)^2 / n)^{\frac{1}{d+1}}$$

We now focus on the case of the Riemannian decay. As for support estimation, Riemannian decay slows down the convergence and convergence rates now depend on the dependence structure.

Theorem 5. *Let $\mathbb{X}_n = \{X_1, \dots, X_n\}$ be a n -sample of a strictly stationary sequence of a strong mixing or θ -weak dependent sequence **with Riemannian decay rates** and with marginal class C^2 density f on its support S . Assume that $\sup_S \|\nabla f\| = f_{\max}^{(1)} < +\infty$, $\sup_S \|\mathcal{H}_f\|_{op} = f_{\max}^{(2)} < +\infty$. Let $t > 0$ be a regular level. Let R be a constant such that $0 < R < m_t / f_{\max}^{(2)}$ and define $\mathbb{X}_n^+(t) = \mathbb{X}_n \cap \{f(z) \geq t\}$ and $\widehat{L}_{t,R} = C_r(\mathbb{X}_n^+(t))$. When $b > b_0$, with probability 1 and for n large enough, one has*

$$\begin{aligned} d(\widehat{\partial L}_{t,R}, \partial L_t) &\lesssim \varepsilon_n + \varepsilon'_n, & d(\widehat{L}_{t,R}, L_t) &\lesssim \varepsilon_n + \varepsilon'_n \\ \widehat{\partial L}_{t,R} &\approx \partial L_t, & \widehat{L}_{t,R} &\approx L_t \\ \|\widehat{L}_{t,R}\|_d - |L_t|_d &\lesssim \varepsilon_n + \varepsilon'_n, & \|\widehat{\partial L}_{t,R}\|_{d-1} - |\partial L_t|_{d-1} &\lesssim \varepsilon_n + \varepsilon'_n. \end{aligned}$$

With $\varepsilon_n = ((\ln n)^a / n^{1-\gamma})^{2/(d+1)}$ explicit values of b_0 , γ_1 and γ_2 are

1. Strong mixing $b_0 = 1$, $a = \frac{b-2-d}{b-\lambda}$ and $\gamma = \frac{\lambda+1}{\lambda+b}$ with $\lambda = \frac{3d+1}{d+1}$.
2. θ -weak dependence $b_0 = 2$ and $a = 2 \frac{2-d+b/2}{b/2-\lambda}$ and $\gamma = \frac{\lambda+1}{\lambda+b/2}$ with $\lambda = \frac{5d-1}{2(d+1)}$.

5. Proofs

5.1. Proofs of Propositions 1 and 2

Set first $k = \lfloor n/q \rfloor$ and $t_i = iq$ for $1 \leq i \leq k$. We have $Z_i = \mathbf{1}_{\{X_i \in E\}}$ and we define $p_k = \mathbb{E}(Z_1 \cdots Z_k)$, then

$$p_k = \mathbb{E}(Z_1 \cdots Z_{k-1})\mathbb{E}(Z_k) + \text{cov}(Z_1 \cdots Z_{k-1}, Z_k).$$

The propositions will result from proving that $\text{cov}(Z_1 \cdots Z_{k-1}, Z_k)$ can be bounded above by some $\mathcal{E}_E(q)$ then applying a recursion on $p_k \leq (1-p)p_{k-1} + \mathcal{E}_E(q)$.

Proof of Proposition 1. From heredity of strong mixing the sequence $(Z_t)_t$ is still strong mixing with mixing coefficients less than $\alpha(q)$ and the fact that $Z_1 \cdots Z_{k-1}$ writes as the indicator function of $\bigcap_{1 \leq i \leq n/q} (X_{t_i} \notin E)$ implies $|\text{cov}(Z_1 \cdots Z_{k-1}, Z_k)| \leq \alpha(q)$, we thus have $p_k \leq (1-p)p_{k-1} + \alpha(q)$ hence a recursion yields

$$p_k \leq (1-p)^k + \frac{\alpha(q)}{p}.$$

Indeed if the above relationship holds with $k-1$ then

$$p_k \leq (1-p) \left(p_{k-1} + \frac{\alpha(q)}{p} \right) + \alpha(q) \leq (1-p)^k + \frac{\alpha(q)}{p}.$$

□

Proof of Proposition 2. Under θ -weak dependence we have by definition $|\text{cov}(Z_1 \cdots Z_{k-1}, h(X_k))| \leq \theta(q) \text{Lip}(h)$ for any Lipschitz function h and we specialize the function for the two cases considered in the Proposition.

1. Case $E = B(x, r)$.

Now as in the Lemma 4.1 page 68 of [11] we consider $H_{x,r,\varepsilon}^-$ and $H_{x,r,\varepsilon}^+$ two Lipschitz approximations $\mathbb{R}^d \rightarrow \mathbb{R}$ of $\mathbf{1}_{\{u \in E\}}$, with

$$H_{x,r,\varepsilon}^-(u) \leq \mathbf{1}_{\{u \notin E\}} \leq H_{x,r,\varepsilon}^+(u),$$

then

$$\begin{aligned} \text{cov}(Z_1 \cdots Z_{k-1}, Z_k) &\leq \mathbb{E}(Z_1 \cdots Z_{k-1} H_{x,r,\varepsilon}^+(X_k)) - \mathbb{E}(Z_1 \cdots Z_{k-1}) \mathbb{E}(H_{x,r,\varepsilon}^+(X_k)) \\ &\leq \text{cov}(Z_1 \cdots Z_{k-1}, H_{x,r,\varepsilon}^+(X_k)) + \mathbb{E}(H_{x,r,\varepsilon}^+(X_k)) - \mathbb{E}(H_{x,r,\varepsilon}^-(X_k)) \end{aligned}$$

Now we build such Lipschitz functions with $\text{Lip}(H_{x,r,\varepsilon}^\pm) \lesssim 1/\varepsilon$. More precisely, we introduce intermediary functions:

$$m_{r,\varepsilon}^-(s) = \begin{cases} 1, & s < (r - \varepsilon) \\ \frac{r-s}{\varepsilon}, & s \in [r - \varepsilon, r] \\ 0, & s > r \end{cases} \quad \text{and} \quad m_{r,\varepsilon}^+(s) = \begin{cases} 1, & s < r \\ 1 - \frac{s-r}{\varepsilon}, & s \in [r, r + \varepsilon] \\ 0, & s > r + \varepsilon \end{cases}$$

Then the functions $m_{r,\varepsilon}^\pm$ are piecewise linear and satisfy $\text{Lip} m_{r,\varepsilon}^\pm = 1/\varepsilon$. Now we define $H_{x,r,\varepsilon}^+(u) = m_{r,\varepsilon}^+(\|u - x\|)$ and $H_{x,r,\varepsilon}^-(u) = m_{r,\varepsilon}^-(\|u - x\|)$, thus we get:

- $H_{x,r,\varepsilon}^+(u) = 1$ if $u \notin B(x, r)$ and $H_{x,r,\varepsilon}^+(u) = 0$ if $u \in B(x, (r + \varepsilon))$.
- $H_{x,r,\varepsilon}^-(u) = 1$ if $u \notin B(x, (r - \varepsilon)^+)$ and $H_{x,r,\varepsilon}^-(u) = 0$ if $u \in B(x, r)$.

Those functions inherit Lipschitz property $\text{Lip}(H_{x,r,\varepsilon}^\pm) = 1/\varepsilon$. Moreover, because $H_{x,r,\varepsilon}^+ - H_{x,r,\varepsilon}^-$ is bounded above by the indicator function of an annulus with radius r and thickness ε .

$$\begin{aligned} \mathbb{E}(H_{x,r,\varepsilon}^+(X_k) - H_{x,r,\varepsilon}^-(X_k)) &\leq c\omega_d \left((r + \varepsilon)^d - ((r - \varepsilon)^+)^d \right) \\ &\leq c\omega_d r^d \left((1 + \varepsilon/r)^d - ((1 - \varepsilon/r)^+)^d \right) \\ &\leq c\omega_d 2^{d+1} r^{d-1} \varepsilon \end{aligned}$$

for all $\varepsilon \leq r$, since a binomial expansion entails $(1+x)^d - (1-x)^d \leq 2x \sum_{k=1}^d \binom{d}{k} \leq 2^{d+1} x$ if $x \in [0, 1]$. Thus, setting $A' = c\omega_d 2^{d+1}$, we derive for all $\varepsilon \leq r$: $\text{cov}(Z_1 \cdots Z_{k-1}, Z_k) \leq \frac{1}{\varepsilon} \theta(q) + A' \varepsilon r^{d-1}$. For the reverse inequality, we obtain analogously for all $\varepsilon \in [0, r]$ $\text{cov}(Z_1 \cdots Z_{k-1}, Z_k) \geq -\frac{1}{\varepsilon} \theta(q) - A' \varepsilon r^{d-1}$. Indeed

$$\begin{aligned} \text{cov}(Z_1 \cdots Z_{k-1}, Z_k) &\geq \mathbb{E}(Z_1 \cdots Z_{k-1} H_{x,r,\varepsilon}^-(X_k)) - \mathbb{E}(Z_1 \cdots Z_{k-1}) \mathbb{E}(H_{x,r,\varepsilon}^-(X_k)) \\ &\geq \text{cov}(Z_1 \cdots Z_{k-1}, H_{x,r,\varepsilon}^-(X_k)) - \mathbb{E}(H_{x,r,\varepsilon}^+(X_k)) + \mathbb{E}(H_{x,r,\varepsilon}^-(X_k)) \end{aligned}$$

and thus:

$$\mathcal{E}_E(q) \leq \min_{\varepsilon \in [0, r]} \frac{\theta(q)}{\varepsilon} + A' \varepsilon r^{d-1}$$

2. Case $E = B(x, r_x) \cap B(y, r_y)$ with $r_x + r_y - \|x - y\| = \ell$.

Consider the two $2/\varepsilon$ -Lipschitz functions: $\overline{H}^+(u) = H_{x, r_x, \varepsilon}^+(u) \wedge H_{y, r_y, \varepsilon}^+(u)$ and $\overline{H}^-(u) = H_{x, r_x, \varepsilon}^-(u) \wedge H_{y, r_y, \varepsilon}^-(u)$. By the concentration inequality (1), we have for some $c > 0$:

$$\mathbb{E} \left(\overline{H}^+(X_k) - \overline{H}^-(X_k) \right) \leq c \left| B(x, r_x + \varepsilon) \cap B(y, r_y + \varepsilon) \setminus (B(x, r_x - \varepsilon) \cap B(y, r_y - \varepsilon)) \right|_d$$

Thus according to Proposition 9 in Appendix A, if $\ell \leq r_x \wedge r_y$ and $\varepsilon \leq r_x \wedge r_y$: $\mathbb{E}(\overline{H}^+(X_k) - \overline{H}^-(X_k)) \leq c\mathcal{V}(r_x, r_y, \ell, \varepsilon)$. As in case 1, this allows to conclude that $\text{cov}(Z_1 \cdots Z_{k-1}, Z_k) \leq \frac{2}{\varepsilon}\theta(q) + c\mathcal{V}(r_x, r_y, \ell, \varepsilon)$. As in the end of the proof for case 1 we infer $|\text{cov}(Z_1 \cdots Z_{k-1}, Z_k)| \leq \frac{2}{\varepsilon}\theta(q) + c\mathcal{V}(r_x, r_y, \ell, \varepsilon)$, which concludes the proof of point 2. in Proposition 2. \square

5.2. Proofs of Propositions 3 and 4

In the following we set $\epsilon'_n = A_n + B_n$ with $A_n = \sup_{x \in \mathbb{R}^d} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))|$ and $B_n = \sup_{x \in \mathbb{R}^d} |\mathbb{E}(\widehat{f}_n(x)) - f(x)|$. Propositions 3 and 4 are direct corollaries of Propositions 6 and 7 below. Proposition 6 gives a control of B_n , while item 1 and item 2 of Proposition 7 give a control of A_n in the geometric and Riemaniann decay cases respectively.

Proposition 6. *Let \mathbb{X}_n be a strictly stationary sequence from \mathbf{X} with marginal density f and \widehat{f}_n be defined by (5), with $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$ as n goes to ∞ . Assume that (K1), (M1) and (K2) hold for some $k \geq 2$. Thus, $B_n = O(h_n^k)$*

Proposition 7. *Let \mathbb{X}_n be a strictly stationary sequence from \mathbf{X} and \widehat{f}_n be defined by (5).*

1. Assume that (K1), (M2), (M3), (M4) and (M5) hold and that one of the conditions below is fulfilled:

(A1) \mathbf{X} is strong mixing with dependence coefficients $\alpha(m) = O(b^{-m})$ for some $b > 1$ and $h_n = O\left(\left(\frac{\log n}{n}\right)^\beta\right)$ for some $\beta < 1/(3d + 6)$.

(T1) \mathbf{X} is θ -weakly dependent with dependence coefficients $\theta(m) = O(b^{-m})$ for some $b > 1$ and $h_n = O\left(\left(\frac{\log n}{n}\right)^\beta\right)$ for some $\beta < 1/(3d + 4)$.

Then

$$A_n = O_{a.s.} \left(\sqrt{\frac{\log n}{nh_n^d}} \right).$$

2. Assume that (K1), (M2), (M3), (M4), (M5) and (K2) hold for some $k \geq 2$ and that one of the conditions below is fulfilled:

(A2) \mathbf{X} is strongly mixing with $\alpha(m) = O(m^{-b})$ and $h_n = O\left(\left(\frac{\log n}{n^{1-\gamma}}\right)^\beta\right)$ for some $0 < \beta < 1/d$, some $0 \leq \gamma < \gamma_0 = \frac{d+2-\beta d(d-2)}{q-\beta d(d-2)}$ and some $q > d+2$ and $b > \max((2d+3)/d, q-1)$.

(T2) \mathbf{X} is θ -weakly dependent with $\theta(m) = O(m^{-b})$ and $h_n = O\left(\left(\frac{\log n}{n^{1-\gamma}}\right)^\beta\right)$ for some $0 < \beta < 1/d$, some $0 \leq \gamma < \gamma_0 = \frac{d+2-\beta d(d-2)}{q-\beta d(d-2)}$ and some even integer $q > d+2$ and $b > \max(2(d+1)/d, 3(q-1))$.

Then

$$A_n = O_{a.s.} \left(\sqrt{\frac{\log n}{n^{1-\gamma} h_n^d}} \right).$$

Remark 3. In the literature, classical results assume in place of (M1) and (K2) the following assumptions to control the bias term (see e.g. Proposition 3.3.1. in [14] extended to the multidimensional case):

- f belongs to the set of Hölder ρ -regular functions

$$C_\rho = \{f : f \text{ is } \lfloor \rho \rfloor\text{-differentiable, } \exists L > 0, \forall (x, y) \in S, |f^{\lfloor \rho \rfloor}(y) - f^{\lfloor \rho \rfloor}(x)| \leq L|y - x|^c\},$$

for some $0 < c \leq 1$ where $\rho \geq 2$ and S is any compact set of \mathbb{R}^d

- K is a kernel of order ρ .

In this case $B_n = O(h_n^\rho)$. These assumptions are somehow more restrictive than (M1) and (K2). Indeed, if $f \in C^k$ and $f^{(k)}$ is c -Hölder then $f \in C_{k+c}$ but the converse does not seem to hold even if $c = 1$ and for an integer k , $C^k \neq C_k$.

Proof of Proposition 3. Let us prove (A1) (the case (T1) can be handled in the same way). By Proposition 6 and item 1 of Proposition 7,

$$\epsilon'_n = O_{a.s.} \left(\left(\frac{\log n}{n} \right)^{\beta k} + \left(\frac{\log n}{n} \right)^{(1-\beta d)/2} \right),$$

with $\beta < 1/(3d + 6)$.

- If $k > d + 3$, thus $1/(2k + d) < 1/(3d + 6)$ and an optimal choice for β is $\beta = 1/(2k + d)$ so that $\epsilon'_n = O_{a.s.} \left(\frac{\log n}{n} \right)^{k/(2k+d)}$.
- If $k \leq d + 3$, thus $1/(2k + d) \geq 1/(3d + 6)$ and $\epsilon'_n = O_{a.s.} \left(\left(\frac{\log n}{n} \right)^{\beta k} \right)$ and the result obtains setting $\beta = (1 - \gamma)/(3d + 6)$ for some $\gamma \in (0, 1)$.

□

Proof of Proposition 4. By Proposition 6 and item 2 of Proposition 7,

$$\epsilon'_n = O_{a.s.} \left(\left(\frac{\log n}{n^{1-\gamma}} \right)^{\beta k} + \left(\frac{\log n}{n^{1-\gamma}} \right)^{(1-\beta d)/2} \right),$$

for some $\gamma > \gamma_0$. To prove (A2) notice that we have $d + 2 < q < b + 1$. Proposition 4 obtains by setting $q = b + 1$ in γ_0 with $b > \max(d + 1, 2 + 3/d)$ in item (A2) of Proposition 7. In case (T2) we have $d + 2 < q < b/3 + 1$ where q is an even integer. Proposition 4 obtains by setting $b/3 + 1$ in place of q in γ_0 and we must have $b/3 + 1 > d + 3$ so $b > 3(d + 2)$ in item (T2) of Proposition 7.

□

Proof of Proposition 6.

$$\mathbb{E}(\widehat{f}_n(x)) = \int_{\mathbb{R}^d} K(u) f(x - hu) du.$$

Since $f \in C^k$, $k \geq 2$ and $\|f^{(k)}\|_\infty < M$ by (M1), we can use an order k Taylor-Lagrange expansion of f around x and using the fact that K is an order $(k - 1)$ kernel by (K2),

$$|\mathbb{E}(\widehat{f}_n(x)) - f(x)| \leq M \frac{h_n^k}{k!} \int_{\mathbb{R}^d} \|u\|^k |K(u)| du.$$

Therefore,

$$\sup_{x \in \mathbb{R}^d} |\mathbb{E}(\widehat{f}_n(x)) - f(x)| = O(h_n^k).$$

Notice that the result still hold when $k = 1$ as soon as

$$\int_{\mathbb{R}^d} \|u\| |K(u)| du \neq 0.$$

□

Proof of Proposition 7. We first start by providing in each case (item 1 and 2 of the proposition), a bound for A_n on a compact set of \mathbb{R}^d of the form

$$\sup_{\|x\| \leq n^\nu} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| = O_{a.s.} \left(\frac{\lambda(n)}{\sqrt{nh^d}} \right), \quad (6)$$

with $\nu > 2/c$ (c defined in (M5)), for some appropriate $\lambda(n) \rightarrow \infty$ such that $\lambda(n) = o(\sqrt{nh^d})$. Then, to state the convergence on the whole set \mathbb{R}^d , we have to show that

$$L_n = \sup_{\|x\| \geq n^\nu} \frac{\sqrt{nh^d}}{\lambda(n)} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| \rightarrow 0 \quad (7)$$

Since K has a compact support, say included in the ball $B_d(0, C)$ of \mathbb{R}^d then if for each $i \leq n$, $\|X_i - x\| > Ch_n$ we have $L_n = \sup_{\|x\| > n^\nu} \mathbb{E}(\widehat{f}_n(x))$. Now since $h_n \leq 1$,

$$|\mathbb{E}(\widehat{f}_n(x))| = \left| \int_{\mathbb{R}^d} K(u) f(x + uh_n) du \right| \leq \sup_{\|x-y\| \leq 1} f(y),$$

thus beeping in mind that $n^\nu - 1 \geq n^\nu/2$ for n large enough, and now:

$$\sup_{\|x\| > n^\nu} |\mathbb{E}(\widehat{f}_n(x))| \leq \sup_{\|y\| \geq n^\nu - 1} f(y)$$

which tends to zero by (M4). To prove that L_n converges a.s. to 0, we have to assume (M5). We thus have

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\max_{1 \leq i \leq n} \|X_i\| > n^\nu - Ch_n \right) \leq \sum_{n=1}^{\infty} 1 \wedge \frac{\mathbb{E}\|X_1\|^c + \dots + \mathbb{E}\|X_n\|^c}{(n^\nu - Ch_n)^c} < \infty,$$

for ν such that $\nu > 2/c$ so that the series converges by (M5). This with the Borel–Cantelli entail that $\sup_{\|x\| \geq n^\nu} |\widehat{f}_n(x)|$ tends to 0 as n goes to infinity a.s. as fast as needed. Indeed $\|X_i - x\| > Ch_n$ a.s. uniformly over $1 \leq i \leq n$ and $\|x\| > n^\nu$ and thus the kernel vanishes.

Let us now prove (6). Set for all $x \in \mathbb{R}^d$,

$$\widehat{G}_n(x) = \sum_{i=1}^n Y_i(x) = \frac{1}{nh^d} \sum_{i=1}^n U_i(x), \quad \text{with } U_i(x) = K \left(\frac{x - X_i}{h_n} \right) - \mathbb{E}K \left(\frac{x - X_i}{h_n} \right). \quad (8)$$

To prove (6), we use the chaining argument of [22]. Let $I = [-M, M]^d$ be a cube of \mathbb{R}^d . Since I is compact, it can be covered by δ^d cubes I_1, \dots, I_{δ^d} with centers x_1, \dots, x_{δ^d} and sides of length $2M/\delta$, where δ is such that $\delta h/M \rightarrow \infty$. Here, notice that h , δ and M are allowed to vary with n but it is omitted in the notation to make the reading easier. Setting $\widehat{G}_n(x) = \widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))$ we have for all j

$$|\widehat{G}_n(x)| \leq |\widehat{f}_n(x) - \widehat{f}_n(x_j)| + |\widehat{f}_n(x_j) - \mathbb{E}\widehat{f}_n(x_j)| + |\mathbb{E}\widehat{f}_n(x_j) - \mathbb{E}\widehat{f}_n(x)|.$$

Under (K1), [22] proves that for some positive constant L and for all $x \in I_j$,

$$\left| K \left(\frac{x - X_i}{h} \right) - K \left(\frac{x_j - X_i}{h} \right) \right| \leq \frac{2MLip(K)}{\delta h} \mathbf{1}_{\|x_j - X_i\| \leq Lh}.$$

Using this fact it is easy to deduce that there exists another kernel density estimator \widetilde{f} for f , obtained by replacing K in (5) by the piecewise linear and compactly supported kernel \widetilde{K} defined by $\widetilde{K}(x) = 1$ if $\|x\| \leq L$ and $\widetilde{K}(x) = 0$ otherwise, such that for some $n \geq n_0$ (see [22] p 78),

$$\sup_{x \in I_j} |\widehat{G}_n(x)| \leq |\widehat{G}_n(x_j)| + \frac{C_1 M}{h\delta} (|\widetilde{G}_n(x_j)| + |\mathbb{E}\widetilde{f}_n(x_j)|)$$

with $C_1 = 2\text{Lip}(K)$ and $\widetilde{G}_n(x) = \widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))$. Therefore,

$$\sup_{\|x\| \leq M^d} |\widehat{G}_n(x)| \leq \max_{1 \leq j \leq \delta^d} |\widehat{G}_n(x_j)| + \frac{C_1 M}{h\delta} \left(\max_{1 \leq j \leq \delta^d} |\widetilde{G}_n(x_j)| + \max_{1 \leq j \leq \delta^d} |\mathbb{E}\widetilde{f}_n(x_j)| \right),$$

so that we can write as soon as $\sup_{\|x\| \leq M^d} |\mathbb{E}\widetilde{f}_n(x)| \leq th\delta/3MC_1$ (which is true when $th\delta/M > C_2$ for some $C_2 > 0$),

$$\begin{aligned} \mathbb{P} \left(\sup_{\|x\| \leq M^d} |\widehat{G}_n(x)| > t \right) &\leq \mathbb{P} \left(\max_{1 \leq j \leq \delta^d} |\widehat{G}_n(x_j)| > \frac{t}{3} \right) + \mathbb{P} \left(\max_{1 \leq j \leq \delta^d} |\widetilde{G}_n(x_j)| > \frac{th\delta}{3MC_1} \right) \\ &\leq \delta^d \sup_{\|x\| \leq M^d} \left\{ \mathbb{P} \left(|\widehat{G}_n(x)| > \frac{t}{3} \right) + \mathbb{P} \left(|\widetilde{G}_n(x)| > \frac{th\delta}{3MC_1} \right) \right\} \end{aligned} \quad (9)$$

We then set $t = C \frac{\lambda(n)}{\sqrt{nh^d}}$ for some appropriate $\lambda(n) \rightarrow \infty$ such that $\lambda(n) = o(\sqrt{nh^d})$ and $M = n^{v/d}$ and have to control each term at the right hand side of (9). If the right hand side is the general term of a convergent series then using Borel-Cantelli lemma we obtain (6).

Therefore, the main goal of the proof of Proposition 7 consists now of finding in each case (A1), (T1), (A2), and (T2) sequences $\lambda(n)$ and $\delta(n)$ such that the right hand side of (9) is the general term of a convergent series. We shall need the following

Lemma 6. *Let \mathbf{X} be strictly stationary sequence with density f and kernel estimator \widehat{f}_n defined by (5). Assume that (M3) holds and that (K1) holds with $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$, and that we have one of the following conditions*

(A1) \mathbf{X} is strongly mixing with $\alpha(m) = O(b^{-m})$ for some $b > 1$.

(T1) \mathbf{X} is θ -weakly dependent with $\theta(m) = O(b^{-m})$ for some $b > 1$.

(A2) \mathbf{X} is strongly mixing with $\alpha(m) = O(m^{-b})$ for some $b > 2 + 3/d$.

(T2) \mathbf{X} is θ -weakly dependent with $\theta(m) = O(m^{-b})$ for some $b > 2 + 2/d$.

Thus one has

$$\sigma_n^2(x) = \text{var}\widehat{f}_n(x) = \frac{1}{nh^d} f(x) \int K^2(u) du + o\left(\frac{1}{nh^d}\right).$$

Proof of item 1 of Proposition 7. In cases of geometric decay rate of the dependence coefficients, we may use a Bernstein's type inequality for $\widehat{G}_n(x)$ and $\widetilde{G}_n(x)$ in order to bound the right hand side of (9). Namely we will use Theorem 1 and Proposition 8 of [17]. Let us recall them below in our special setting

Lemma 7. *Let \mathbb{Y}_n be a n -sample of a \mathbb{R} -valued strictly stationary sequence \mathbf{Y} such that $\mathbb{E}(Y_i) = 0$ and $\|Y\|_\infty \leq H$ in probability. Assume that there exists some $0 < c < 1$ and some $L > 0$ such that for all u -tuples (s_1, \dots, s_u) and all v -tuples (t_1, \dots, t_v) with $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$,*

$$\text{cov}(Y_{s_1} \dots Y_{s_u}, Y_{t_1} \dots Y_{t_v}) \leq H^{u+v-2} L^2 v c^{(t_1 - s_u)}. \quad (10)$$

Then, for all $t > 0$,

$$\mathbb{P} \left(\sum_{i=1}^n Y_i > t \right) \leq \exp \left(-\frac{1}{2} \frac{t^2}{\sigma_n^2 + G_n^{1/3} t^{5/3}} \right),$$

with

$$\sigma_n^2 = \text{var} \left(\sum_{i=1}^n Y_i \right), \quad G_n \sim (L \vee H) \left(\frac{nL^2}{\sigma_n^2} \vee 1 \right).$$

Now, let $Y_i \equiv Y_i(x)$ be defined by (8). Notice that $Y_i(x) = F_n(X_i)$, where F_n is a bounded Lipschitz function with $\|Y_i(x)\|_\infty \leq H = \|F_n\|_\infty \leq 2\|K\|_\infty/nh^d$ and $LipF_n \leq LipK/(nh^{d+1})$. Assume that

$$\sigma_n^2 = O_{a.s.} \left(\frac{1}{nh^d} \right) \quad (11)$$

- *Strong mixing case.* If \mathbf{X} is strong mixing with $\alpha(m) = O(b^{-m})$, $b > 1$, thus \mathbf{Y} is strong mixing with $\alpha_Y(m) \leq \alpha(m)$ by measurability of F_n . Moreover, by [15]'s Lemma 6 and [17]'s Remark 7, (10) is satisfied with $L^2 = (LipF_n)^2$ and $c = 1/b$. Therefore Lemma 7 applies with $H \sim 1/nh^d$, $L^2 \sim 1/(n^2h^{2d+2})$ and $G_n \sim 1/(nh^{2d+3})$. Setting $t = C\sqrt{\log n/nh^d}$ for some convenient constant $C > 0$, we get

$$\mathbb{P} \left(|\widehat{G}_n| > C \sqrt{\frac{\log n}{nh^d}} \right) \leq n^{-C^2/2}$$

as soon as $u_n = nh^d G_n^{1/3} t^{5/3} \rightarrow 0$, which is true as soon as $\beta < 1/(3d+6)$ with $h_n = O((\log n/n)^\beta)$.

- *θ -weak dependent case.* If \mathbf{X} is θ -weakly dependent with $\theta(m) = O(b^{-m})$, $b > 1$, thus \mathbf{Y} is θ -weakly dependent since F_n is a Lipschitz function, by heredity of this property through Lipschitz functions, with $\theta_Y(m) \leq \theta(m)$. Moreover, using Remark 9 and Proposition 8 of [17], (10) is satisfied with $L^2 = HLipF$ and $c = 1/b$. Therefore Lemma 7 applies with $H \sim 1/nh^d$, $L^2 \sim 1/(n^2h^{2d+1})$ and $G_n \sim 1/(nh^{2d+2})$. Setting $t = C\sqrt{\log n/nh^d}$ for some convenient constant $C > 0$, we get

$$\mathbb{P} \left(|\widehat{G}_n| > C \sqrt{\frac{\log n}{nh^d}} \right) \leq n^{-C^2/18}$$

as soon as $u_n = nh^d G_n^{1/3} t^{5/3} \rightarrow 0$, which is true as soon as $\beta < 1/(3d+4)$ with $h_n = O((\log n/n)^\beta)$.

Therefore, we obtain in both cases a bound for the first term at the right hand side of (9). The second term can be handled in the same way, with identical values of L^2 , H and G_n setting $\delta = MC_1/h^\varepsilon$, $\varepsilon > 1$.

Finally, one obtains for (9) setting $M = n^{\nu/d}$,

$$\mathbb{P} \left(\sup_{\|x\| \leq n^\nu} |\widehat{G}_n(x)| > t \right) \leq \delta^d n^{-D} = A n^{\nu+\varepsilon\beta d-D} (\log n)^{-\beta\epsilon d}$$

for some constant $A > 0$ and some conveniently chosen C in the expression of t such that $D > 1 + \nu + \beta\epsilon d$ so that

$$\sup_{\|x\| \leq n^\nu} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| = O_{a.s.} \left(\sqrt{\frac{\log n}{nh_n^d}} \right)$$

using Borel–Cantelli. It remains to show that (11) holds, which is true by Lemma 6 under (K1), (M2) and (M3).

Proof of item 2 of Proposition 7. The proof relies on the following

Lemma 8. Let \mathbb{X}_n be a n -sample of a strictly stationary sequence \mathbf{X} and \widehat{f}_n be defined by (5).

- (i) Assume (K1) with $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$. Assume moreover that there exists some $c > 0$ and some $q \geq 2$ such that for $n > n_0$

$$\sup_{x \in \mathbb{R}^d} \|\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))\|_q \leq \frac{c}{\sqrt{nh_n^d}}. \quad (12)$$

Then, for every sequence (λ_n) such that $\lambda_n \rightarrow \infty$ and $\lambda_n = O(\sqrt{nh_n^d})$ as n tends to ∞ and every $\nu > 0$ satisfying

$$\sum_{n=1}^{\infty} \left(\frac{\sqrt{nh_n^d}}{h_n} \right)^d \frac{n^\nu}{\lambda_n^q} < \infty, \quad (13)$$

$$\sup_{\|x\| \leq n^\nu} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| = O_{a.s.} \left(\frac{\lambda_n}{\sqrt{nh_n^d}} \right).$$

(ii) Assume that (12) holds and $h_n = O\left(\left(\frac{\log n}{n^{1-\gamma}}\right)^\beta\right)$ for some $0 < \beta < 1/d$ and $0 < \gamma < 1$. Then, under (K1) one has

$$\sup_{\|x\| \leq n^\gamma} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| = O_{a.s.} \left(\sqrt{\frac{\log n}{n^{1-\gamma} h_n^d}} \right)$$

as soon as γ, β and q are such that

$$\gamma q + (1 - \gamma)\beta d(d - 2) > d + 2. \quad (14)$$

In the following lemma, we give the conditions under which (12) holds in Riemmanian decay cases (A2) and (T2). The point (ii) of this lemma together with 8 yield item 2 of Proposition 7.

Lemma 9. Let \mathbb{X}_n be a n -sample of a strictly stationary sequence \mathbf{X} and \widehat{f}_n be defined by (5).

(A2) Assume that \mathbf{X} is α -mixing with $\alpha(m) = o(m^{-b})$ for some $b > 0$. Assume moreover that (M2), (M3), (K1) and (K2) hold for some $k \geq 2$, with $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$. Thus (12) obtains for every $q > 2$ and $b > \max(2 + 3/d, q - 1)$.

(T2) Assume that \mathbf{X} is θ -weakly dependent with $\theta(m) = o(m^{-b})$ for some $b > 0$. Assume moreover that (M2), (M3) and (K1) hold, with $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$. Thus (12) obtains for every even $q \geq 2$ and $b > \max(2 + 2/d, 3(q - 1))$. □

Proof of Lemma 6. Let \check{f}_n be the kernel estimator analogue to \widehat{f}_n but built on independent copies of X_0 . A well known result (see for instance [24]) is

$$\text{var} \check{f}_n(x) = \frac{1}{nh^d} f(x) \int K^2(u) du + o\left(\frac{1}{nh^d}\right).$$

We have

$$\text{var} \widehat{f}_n(x) = \text{var} \check{f}_n(x) + \frac{2}{n^2 h^{2d}} \sum_{1 \leq i < j \leq n} \text{cov}(U_i(x), U_j(x)).$$

So, by stationarity of the process $(U_i(x))$,

$$\Delta_n(x) = |\text{var} \widehat{f}_n(x) - \text{var} \check{f}_n(x)| \leq \frac{2}{nh^{2d}} \sum_{m=1}^{n-1} |\text{cov}(U_0(x), U_m(x))|.$$

On the one hand,

$$\begin{aligned} |\text{cov}(U_0(x), U_m(x))| &= h^{2d} \iint (K(s) - \mathbb{E}(K(X_0))) (K(t) - \mathbb{E}(K(X_0))) f_{0,m}(x - hs, x - ht) ds dt \\ &\leq Ch^{2d} \end{aligned}$$

by (K1) and (M3).

On the other hand one has the property (10) with $u = v = 1$, $s_u - t_1 = m$, M and L^2 depending on the dependence structure of \mathbf{X} . Namely,

- If \mathbf{X} is θ -weakly dependent then by (10)

$$|\text{cov}(U_0(x), U_m(x))| \leq 2\|K\|_\infty \frac{\text{Lip} K}{h} \theta(m).$$

Therefore,

$$\frac{1}{nh^{2d}} |\text{cov}(U_0(x), U_m(x))| \leq \frac{D}{nh^d} \left(\frac{\theta(m)}{h^{d+1}} \wedge h^d \right) \leq \frac{D}{nh^d} h^{d-\alpha(2d+1)} \theta(m)^\alpha$$

for all $0 \leq \alpha \leq 1$ by relation (12.18) of [14] so that setting $\alpha = d/(2d + 1)$ Δ_n tends to zero as soon as

$$\sum \alpha(m)^{\frac{d}{2d+1}} < \infty.$$

When $\theta(m) \sim m^{-b}$, this holds for all $b > 2 + 2/d$. In case of geometric decay $\theta(m) \sim b^{-m}$, $b > 1$, this is always true.

- If \mathbf{X} is strong mixing, we have by (10)

$$|\text{cov}(U_0(x), U_m(x))| \leq \frac{(\text{Lip}K)^2}{h^2} \alpha(m)$$

so

$$\frac{1}{nh^{2d}} |\text{cov}(U_0(x), U_m(x))| \leq \frac{D}{nh^d} \left(\frac{\alpha(m)}{h^{d+2}} \wedge h^d \right) \leq \frac{D}{nh^d} h^{d-\alpha(2d+2)} \alpha(m)^\alpha$$

for all $0 \leq \alpha \leq 1$ so that setting $\alpha = d/(2d+2)$ Δ_n tends to zero as soon as

$$\sum_m \alpha(m)^{\frac{d}{2d+2}} < \infty.$$

When $\alpha(m) \sim m^{-b}$, this holds for all $b > 2 + 3/d$. In case of geometric decay $\theta(m) \sim b^{-m}$, $b > 1$, this is always true. □

Proof of Lemma 8. • *Proof of (i)* Let $\lambda(n) \rightarrow \infty$ such that $\lambda(n) = o(\sqrt{nh_n^d})$. We deduce from (12) by Markov inequality that there exists some $C_3 > 0$ such that

$$\mathbb{P} \left(\sup_{\|x\| \leq M_n^d} |\widehat{G}_n(x)| > \frac{\lambda(n)}{\sqrt{nh_n^d}} \right) \leq \frac{C_3 \delta^d}{\lambda(n)^q}$$

as soon as there is some $C_4 > 0$ such that $\delta h_n \geq C_4 M_n \sqrt{nh_n^d}$. For that task, set $t = \frac{\lambda(n)}{\sqrt{nh_n^d}}$ in (9). One has

$$\mathbb{P} \left(|\widehat{G}_n(x)| > \frac{\lambda(n)}{3\sqrt{nh_n^d}} \right) \leq \frac{3^q}{\lambda(n)^q} \left(\sup_{x \in \mathbb{R}^d} \sqrt{nh_n^d} \|\widehat{G}_n(x)\|_q \right)^q \leq \frac{C_5}{\lambda(n)^q}$$

with $C_5 = (3c)^q$. Moreover, $\|\widetilde{G}_n(x)\|_q$ is uniformly bounded so as soon as $\delta h_n \geq C_4 M \sqrt{nh_n^d}$,

$$\mathbb{P} \left(|\widetilde{G}_n(x)| > \frac{\lambda(n) h_n \delta}{3C_4 M_n \sqrt{nh_n^d}} \right) \leq \frac{C_5}{\lambda(n)^q},$$

with $C_5 = (3C_1 C_4)^{-1}$. Set

$$\delta = \frac{M \sqrt{nh_n^d}}{h_n}.$$

We have

$$\sup_{\|x\| \leq M_n^d} |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| = O_{a.s.} \left(\frac{\lambda(n)}{\sqrt{nh_n^d}} \right). \quad (15)$$

This holds from Borel-Cantelli lemma as soon as

$$\sum_n \left(\frac{M_n \sqrt{nh_n^d}}{h_n} \right)^d \frac{1}{\lambda(n)^q} < \infty. \quad (16)$$

The result of Proposition 8 obtains setting $M_n = n^{\nu/d}$.

- *Proof of (ii)*

With the conditions on h_n we have $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$. Set $\lambda(n) = \sqrt{n^\gamma \log n}$. Since $0 < \gamma < 1$ and $0 < \beta < 1/d$, we have $\lambda(n) \rightarrow \infty$ and $\lambda(n)/\sqrt{nh_n^d} \rightarrow 0$. To achieve the proof we need to ensure that we can find $\nu > 0$ such that (16) converges. Set u_n the general term of (16). With the assumptions on h_n and $\lambda(n)$,

$$u_n = O \left(\frac{1}{n^c \log n^a} \right),$$

with $2c = \gamma q - 2\nu - d + \beta d(1 - \gamma)(d - 2)$ and $a = (q - \beta d(d - 2))/2$. The series of general term u_n is a Bertrand's series that converges if and only if $c > 1$ or $c = 1$ and $a > 0$. One thus need that there exists some $\nu > 0$ such that $2\nu \leq \varepsilon$ with $0 < \varepsilon = \gamma q - 2 - d + \beta d(1 - \gamma)(d - 2)$ which holds as soon as $\gamma q - 2 - d + \beta d(1 - \gamma)(d - 2) > 0$. \square

Proof of Lemma 9. • *Proof of (A2)* In the strong mixing case, we apply, for given $x \in \mathbb{R}^d$, [25]'s theorem 6.3 to the sequence $U(x) = (U_i(x))_{1 \leq i \leq n}$. Notice that it is strong mixing with $\alpha_{U(x)}(m) \leq \alpha(m)$ by heredity of sequences expressed as measurable functions of strong-mixing sequences. Set

$$S_n(x) = \sum_{i=1}^n U_i(x) = nh^d \widehat{G}_n(x).$$

For all $q > 2$, $U_i(x)$ has an order q moment since $\|U_i(x)\| \leq 2\|K\|_\infty$. Thus, applying Theorem 6.3 yields for all $n > n_0$,

$$\|S_n(x)\|_q^q \leq as_n(x)^q + nc \int_0^1 (\alpha^{-1}(u))^{q-1} Q_x^q(u) du,$$

where a and c are positive constants that only depend on q , $\alpha^{-1}(u) = \sum_{n>1} \mathbf{1}_{u < \alpha(n)}$, Q_x is the generalized inverse of the tail function $t \mapsto P(|U_0(x)| > t)$ and

$$s_n(x)^2 = \sum_i \sum_j |\text{cov}(U_i(x), U_j(x))|.$$

Using [25]'s inequality C.3 p 157 for some $r > q$ there exists $c' > 0$ that only depends on q, r and $\|K\|_\infty$ such that

$$\begin{aligned} \|S_n(x)\|_q^q &= (nh_n^d)^q \|\widehat{G}_n(x)\|_q^q \leq as_n(x)^q + c'n \|U_0(x)\|_r^q \left(\sum_{i \geq 0} (i+1)^{\frac{qr-2r+q}{r-q}} \alpha(i) \right)^{1-q/r} \\ &\leq as_n(x)^q + c'nh_n^{dq/r} \left(\sum_{i \geq 0} (i+1)^{\frac{qr-2r+q}{r-q}} \alpha(i) \right)^{1-q/r}. \end{aligned}$$

Using Lemma 6 and setting $h_n = O((\log n/n^{1-\gamma})^\beta)$, we obtain (12) under (K2) for some $k > 1$ as soon as (M2) and the following conditions are fulfilled for some $r > q$:

- (1) $\frac{h_n^{dq(1/r-1/2)}}{n^{q/2-1}} \rightarrow 0$.
- (2) $\sum_{i \geq 0} (i+1)^{\frac{qr-2r+q}{r-q}} \alpha(i) < \infty$.
- (3) $b > 2 + 3/d$.

If $q(1 - (1 - \gamma)\beta d) \geq 2$, Condition (1) is fulfilled for all $r > q$. On the other hand, condition (2) is fulfilled as soon as

$$b > \frac{r(q-1)}{r-q} = u(r).$$

Since u is decreasing with $\lim_{r \rightarrow \infty} u(r) = q - 1$, we can take $b > \max(q - 1, 2 + 3/d)$ to meet the conditions (1) to (3).

If $q(1 - (1 - \gamma)\beta d) < 2$, obvious calculations entails that we have to choose r such that

$$q < r \leq r_1 = \frac{2(1 - \gamma)\beta dq}{2 - q(1 - (1 - \gamma)\beta d)}$$

In this case, one can choose

$$b > u(r_1) = \frac{2(1-\gamma)\beta d(q-1)}{(q-2)(1-(1-\gamma)\beta d)}$$

to meet (2). Since $q(1-\beta d) < 2$, $b > u(r_1) > q-1$ so that $b > \max(q-1, 2+3/d)$ meet the conditions (1) to (3).

Finally, (12) obtains as soon as $b > \max(q-1, 2+3/d)$.

• *Proof of (T2)*

In the θ -weakly dependent case, it straightforwardly follows from [16]'s theorem 1: under the conditions (M2), (M3), (K1) with $h_n \rightarrow 0$ and $nh_n^d \rightarrow \infty$ (12) obtains for any even $q \geq 2$ as soon as

$$\sigma_n^2 = O(1/nh_n^d) \quad \text{and} \quad \sum_{m \geq 0} (m+1)^{q-2} \theta(m)^{1/3} < \infty.$$

When $\theta(m) = o(m^{-b})$, the first condition implies $b > 2 + 2/d$ by Lemma 6 and the second obtains with $b > 3(q-1)$. □

5.3. Preliminary results for the proofs of Section 3

In this section, we give some technical tools from [31], [26] and [2] on the R -convex hull and the rolling ball condition used to derive the results of Section 3. The proofs are omitted or only sketched. The presentation of these deterministic results allow to make the paper self contained and to highlight where the main differences lie between the dependent and independent settings

For a set $S \in \mathbb{R}^d$, recall that its R -convex hull $C_R(S)$ is defined as:

$$C_R(S) = \left(\bigcup_{\dot{B}(x,R) \cap S = \emptyset} \dot{B}(x,R) \right)^c.$$

Hereafter, we recall some obvious properties of the R -convex hull:

Property 1. *Let $C_R(S)^c$ be the complement of $C_R(S)$ on \mathbb{R}^d . One has*

$$C_R(S)^c = \{x \in \mathbb{R}^d, x \in \dot{B}(O,R) \text{ and } S \cap \dot{B}(O,R) = \emptyset\}$$

so that

1. *If $R < R'$ then $C_R(S) \subset C_{R'}(S)$.*
2. *if $S' \subset S$ then $C_R(S') \subset C_R(S)$*
3. *Let $H(S)$ be the convex hull of S . One has*

$$\lim_{R \rightarrow +\infty} C_R(S) = C_\infty(S) = H(S).$$

Sets satisfying the inside and outside rolling ball property have nice properties detailed in [32] which are recalled in Property 2 and illustrated in Figure 1.

Property 2. *If S satisfies the inside and outside rolling ball property for some positive R_i and R_o then*

1. *For all $R < R_o$, $C_R(S) = S$.*
2. *For all points $z \in \partial S$ the "unit outward normal" vector (i.e. normal to ∂S and pointing outside S), denoted η_z is well defined as $\eta_z = \frac{O_z^c - z}{R_o}$.*
3. *For all $x \in S$ with $d(x, \partial S) \leq R_i$, its projection $\pi_{\partial S}(x)$ onto ∂S is well defined and $x = \pi_{\partial S}(x) - d(x, \partial S)\eta_{\pi_{\partial S}(x)}$, $O_x^i = \pi_{\partial S}(x) - R_i\eta_{\pi_{\partial S}(x)}$.*

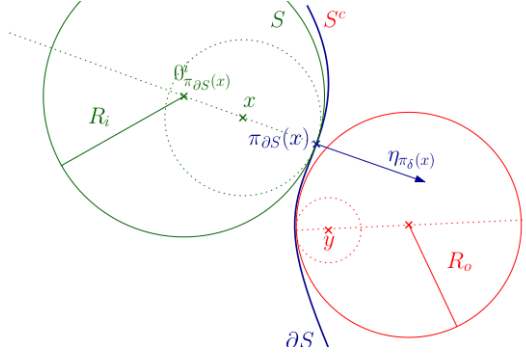


Fig. 1: All you have to have in mind concerning the rolling ball property

4. Conversely, for all $y \in S^c$ with $d(y, \partial S) \leq R_o$, its projection $\pi_{\partial S}(y)$ onto ∂S is well defined and $y = \pi_{\partial S}(y) + d(y, \partial S)\eta_{\pi_{\partial S}(y)}$, $O_y^o = \pi_{\partial S}(y) + R_o\eta_{\pi_{\partial S}(y)}$.

Following the arguments of [26], we can prove the following result which will be a key point for the study of the asymptotic properties of the R -convex hull estimator.

Proposition 8. Let $S' \subset S \subset \mathbb{R}^d$. Assume that balls of radius R_o (resp. R_i) roll freely outside (resp. inside) S . Let $R < R_o$ and $\varepsilon < \min(R_i, R)$. If, for all $O \in \mathbb{R}^d$ such that $d(O, S) \leq R - \varepsilon$ we have that $B(O, R) \cap S' \neq \emptyset$, then

$$S \ominus \varepsilon B = \{x \in S, d(x, \partial S) \geq \varepsilon\} \subset C_R(S') \subset S$$

Proof. Proposition 8. The second inclusion is a direct consequence of $S' \subset S \subset \mathbb{R}^d$ thus $C_R(S') \subset C_R(S)$ then, because $R \leq R_o$ and due to the outside rolling ball condition we have $C_R(S) = S$.

Now let prove the first inclusion by contradiction. Suppose that there exists $x \in S$ with $d(x, \partial S) = \varepsilon' \geq \varepsilon$ and $x \in C_R(S')^c$. As $x \in C_R(S')^c$ there exists O , $\|O - x\| < R$ with $B(O, R) \cap S' = \emptyset$. Introduce $z = [O, x] \cap S(x, \varepsilon')$ we have $z \in S$ then, because $\|O - x\| \leq R$ we have $d(O, S) \leq R - \varepsilon'$ that contradicts $B(O, R) \cap S' = \emptyset$. \square

From Proposition 8 and using theorem 4.1 in [2], we can derive general deterministic results on the R -convex hull summarized in the following

Corollary 1. Let $S' \subset S \subset \mathbb{R}^d$ be a finite set. Assume that balls of radius R_o (resp. R_i) roll freely outside (resp. inside) S . Let $R < R_o$ and $\varepsilon < \min(R_i, R)$. If, for all $O \in \mathbb{R}^d$ such that $d(O, S) \leq R - \varepsilon$ we have that $B(O, R) \cap S' \neq \emptyset$, then we have

$$d(\partial C_R(S'), \partial S) \leq \varepsilon, \quad d(C_R(S'), S) \leq \varepsilon, \\ |C_R(S')|_d - |S|_d \leq |\partial S|_{d-1} \varepsilon (1 + o(1)).$$

Moreover if $d(S', S) \leq \frac{1}{2} \frac{R \min(R_i, R_o)}{R + \min(R_i, R_o)}$ and ε is small enough with regards to d , R and $\min(R_i, R_o)$, one has

$$\partial C_R(S') \approx \partial S, \quad C_R(S') \approx S$$

and there exists some constant $C_{d,R,R_o,R_i} > 0$ depending on d, R, R_o and R_i such that

$$\|\partial C_R(S')\|_{d-1} - \|\partial S\|_{d-1} \lesssim C_{d,R,R_o,R_i} \varepsilon (1 + o(1))$$

when $\varepsilon \rightarrow 0$.

Proof. Corollary 1

Here we only sketch the proof of Corollary 1, giving main arguments. By Proposition 8 we have $S \ominus \varepsilon B \subset C_R(S') \subset S$. Thus

1. About $d(\partial C_R(S'), \partial S)$
 - Since $S \ominus \varepsilon B \subset C_R(S')$, then, for all $x \in \partial C_R(S')$ exists $d(x, \partial S) \leq \varepsilon$.
 - For all $x \in \partial S$ consider $y = x - \varepsilon \eta_x$, by inside rolling ball property of S we have $y \in S \ominus \varepsilon B$ and thus the segment (x, y) intersects $\partial C_R(S')$ and $d(x, \partial C_R(S')) \leq \varepsilon$.
2. About $d(C_R(S'), S)$ we clearly have by inclusions that $d(C_R(S'), S) \leq d(S \ominus \varepsilon B, S) = \max_{x \in S} d(x, S \ominus \varepsilon B) = \max_{x \in S \setminus (S \ominus \varepsilon B)} d(x, S \ominus \varepsilon B)$. Now take $x \in S \setminus (S \ominus \varepsilon B)$, there exists $y \in \partial S$ with $\|y - x\| < \varepsilon$ thus $\|\pi_{\partial S}(x) - x\| < \varepsilon$. We have $x = \pi_{\partial S}(x) - \|x - \pi_{\partial S}(x)\| \eta_{\pi_{\partial S}(x)}$. Introduce $x' = \pi_{\partial S}(x) - \varepsilon \eta_{\pi_{\partial S}(x)}$ which belongs to $S \ominus \varepsilon B$, $\|x - x'\| \leq \varepsilon$ that proves $d(C_R(S'), S) \leq \varepsilon$.
3. $|C_R(S')|_d \leq |S \setminus (S \ominus \varepsilon B)|_d \approx \varepsilon |\partial S|_{d-1}$ indeed it is the inner Minkowski content of a set with positive reach, its asymptotic can be deduced from the Minkowski content of a set with positive reach (see [18]) and its outer part see ([3]).
4. $\partial C_R(S') \approx \partial S$ and $|\partial C_R(S')|_{d-1} - |\partial S|_{d-1} \lesssim C_{d,R,R_0,R_i} \varepsilon (1 + o(1))$ are direct consequences of Theorem 4.1 in [2]. The homeomorphism between $\partial C_R(S')$ and ∂S being $\pi_{\partial S}$ it is easy, and let to the reader, to prove that, since ε is small enough $\varphi : C_R(S') \mapsto S$ define as

$$\begin{cases} \varphi(x) = x & \text{if } x \in S \ominus 2\varepsilon B \\ \varphi(x) = X_0(x) + 2\varepsilon \frac{\|x - X_0(x)\|}{\|X_1(x) - X_0(x)\|} \eta_{\partial S}(x) & \text{if } x \in C_R(S') \setminus (S \ominus 2\varepsilon B) \end{cases}$$

where $X_0(x) = \pi_{\partial S}(x) - 2\varepsilon \eta_{\pi_{\partial S}(x)}$ and $X_1(x) = (X_0(x), \pi_{\partial S}(x)) \cap \partial C_R(S')$ is a homeomorphism and thus that $C_R(S') \approx S$

□

A way to derive asymptotic rates for the R -convex hull of a set of point \mathbb{X}_n clearly consists in finding sequences ε_n such that

$$\sum_n \mathbb{P}(\exists O, d(O, S) < R - \varepsilon_n, B(O, R) \cap \mathbb{X}_n = \emptyset) < +\infty.$$

Also we have to prove that $d(\mathbb{X}_n, S) \rightarrow 0$ almost surely to apply the last part of Corollary 1 and deduce the topological guarantees and the convergence for the measure of the boundary.

Dependence has to be taken into account when dealing with $\mathbb{P}(B(O, R) \cap \mathbb{X}_n = \emptyset)$ where Propositions 1 or 2 are applied (instead of a trivial product in independent case).

Lemma 10. *Assume that we are under the assumptions of Propositions 1 or 2. Assume moreover that the marginal density f belongs to the model $\mathcal{M}_{R_0, R_i, f_0, \alpha}$. Let $\mathcal{E}'_r(q) = \alpha(q)$ in case of strong mixing and $\mathcal{E}'_r(q) = \min_{t \in [0, r]} \left(\frac{\theta(q)}{t} + c\omega_d 2^d t r^{d-1} \right)$ in case of θ -weak dependence). Introduce*

$$P_n = \mathbb{P}(\exists O \in S \oplus (R - \varepsilon_n)B, B(O, R) \cap \mathbb{X}_n = \emptyset)$$

We have the existence of a constant C (depending on model parameters, f_0, α, d, R_i and the parameter R) such that

$$P_n \lesssim (\ln n)^d \varepsilon_n^{-d} \left(\exp\left(-C \frac{n}{q} \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n}\right)^{\frac{d+1+2\alpha}{2}}\right) + \frac{\mathcal{E}'_{R - \frac{\varepsilon_n}{\ln n}}(q)}{C \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n}\right)^{\frac{d+1+2\alpha}{2}}} \right) \quad (17)$$

We also have that

$$\mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) \lesssim \left(\frac{\ln n}{r_n}\right)^d \left(\exp\left(-C' \frac{n}{q} \left(r_n - \frac{r_n}{\ln n}\right)^{d+\alpha}\right) + \frac{\mathcal{E}'_{r_n - \frac{r_n}{\ln n}}(q)}{\left(r_n - \frac{r_n}{\ln n}\right)^{d+\alpha}} \right) \quad (18)$$

Proof. Let cover $S \oplus (R - \varepsilon_n)B$ with $\nu_n \leq C_d \frac{\text{diam}(S) + R_0}{\varepsilon_n} (\ln n)^d$ **deterministic** balls of radius $\varepsilon_n / \ln n$ and centered at $x_i \in S \oplus (R - \varepsilon_n)B$ we have by Propositions 1 or 2 and by the triangular inequality:

$$P_n \leq \sum_{i=1}^{v_n} \mathbb{P}(B(x_i, R - \frac{\varepsilon_n}{\ln n}) \cap \mathbb{X}_n = \emptyset) \leq \sum_{i=1}^{v_n} \left((1 - p_i)^{\frac{n}{q}} + \frac{\mathcal{E}'_{R - \frac{\varepsilon_n}{\ln n}}(q)}{p_i} \right)$$

where $p_i = \mathbb{P}(X_1 \in B(x_i, R - \frac{\varepsilon_n}{\ln n}))$. By Proposition 10 it comes that

$$p_i \geq f_0 C''_{d,\alpha} \left(\frac{\min((R - \frac{\varepsilon_n}{\ln n}), R_i)}{2} \right)^{\frac{d-1}{2}} \left\{ \min \left(\left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right), \frac{\min((R - \frac{\varepsilon_n}{\ln n}), R_i)}{2} \right) \right\}^{\frac{d+1+2\alpha}{2}}$$

Thus as ε_n converges to 0 for any $C < f_0 C''_{d,\alpha} \left(\frac{\min(R, R_i)}{2} \right)^{\frac{d-1}{2}}$ for n large enough we finally obtain:

$$P_n \lesssim (\ln n)^d \varepsilon_n^{-d} \left(\left(1 - C \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right)^{\frac{d+1+2\alpha}{2}} \right)^{\frac{n}{q}} + \frac{\mathcal{E}'_{R - \frac{\varepsilon_n}{\ln n}}(q)}{C \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right)^{\frac{d+1+2\alpha}{2}}} \right)$$

and thus

$$P_n \lesssim (\ln n)^d \varepsilon_n^{-d} \left(\exp \left(-C \frac{n}{q} \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right)^{\frac{d+1+2\alpha}{2}} \right) + \frac{\mathcal{E}'_{R - \frac{\varepsilon_n}{\ln n}}(q)}{C \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right)^{\frac{d+1+2\alpha}{2}}} \right)$$

that concludes the proof of Equation (17).

Suppose Now that $d(\mathbb{X}_n, S) \geq r_n$ with $r_n \leq R_i$ then there exist $x \in S$ with $B(x, r_n) \cap \mathbb{X}_n = \emptyset$. Define Ω as :

$$\Omega = \begin{cases} x & \text{if } d(x, \partial S) \geq 1/3 \\ \pi_{\partial S}(x) - \frac{2r_n}{3} \eta_{\pi_{\partial S}(x)} & \text{if } d(x, \partial S) < 1/3 \end{cases}$$

We have that $B(\Omega, r_n/3) \subset S$ and $B(\Omega, r_n/3) \subset B(x, r_n)$.

As a first conclusion if $d(\mathbb{X}_n, S) \geq r_n$ then there exists $\Omega \in S \ominus \frac{r_n}{3} B$ with $B(\Omega, r_n/3) \cap \mathbb{X}_n = \emptyset$. The sequel of the proof consist in bounding the probability of existence of such a point. Cover $S \ominus \frac{r_n}{3} B$ with balls of radius $\frac{r_n}{3 \ln n}$ centered at y_1, \dots, y_N in $S \ominus \frac{2r_n}{3} B$ with $N \leq C_d \frac{(\text{diam}(S)(\ln n)^d 3^d)}{r_n^d}$.

Introduce

$$\mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) = \mathbb{P} \left(\exists \Omega \in S \ominus \frac{2r_n}{3} B, B \left(\Omega, \frac{r_n}{3} \right) \cap \mathbb{X}_n = \emptyset \right).$$

By the triangular inequality

$$\mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) \leq \sum_{i=1}^N \mathbb{P} \left(B \left(y_i, \frac{r_n}{3} \left(1 - \frac{1}{\ln n} \right) \right) \cap \mathbb{X}_n = \emptyset \right).$$

Let now $q_i = \mathbb{P} \left(B \left(y_i, \frac{r_n}{3} \left(1 - \frac{1}{\ln n} \right) \right) \cap \mathbb{X}_n = \emptyset \right)$ by proposition 10 we have $q_i \geq f_0 C_{d,\alpha} \left(\frac{r_n}{6} \left(1 - \frac{1}{\ln n} \right) \right)^{d+\alpha}$ and thus, setting $C' = f_0 C_{d,\alpha} / 3^{d+\alpha}$

$$\mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) \lesssim \left(\frac{\ln n}{r_n} \right)^d \left(\exp \left(-C' \frac{n}{q} \left(r_n - \frac{r_n}{\ln n} \right)^{d+\alpha} \right) + \frac{\mathcal{E}'_{r_n - \frac{r_n}{\ln n}}(q)}{\left(r_n - \frac{r_n}{\ln n} \right)^{d+\alpha}} \right) \quad (19)$$

□

5.4. Proof for theorems on support estimation

5.4.1. Independence

First Let us focus on the iid case, that is, for all $q \in \mathbb{N}^*$ and $R \in \mathbb{R}$ $\mathcal{E}'_R(q) = 0$.

Proof. of Theorem 1. Apply Lemma 10 Equation (17) with $q = 1$ and $\varepsilon_n = \left(\frac{3}{C} \frac{\ln n}{n}\right)^{\frac{2}{d+1+2\alpha}}$ it comes that $P_n \lesssim (\ln n)^d n^{-1-2\frac{1+2\alpha}{d+1+2\alpha}+o(1)}$ thus $\sum_n P_n < +\infty$.

Also by Equation (18) with $q = 1$ and $r_n = \left(\frac{3d+2\alpha}{C'(d+\alpha)} \frac{\ln n}{n}\right)^{\frac{1}{d+\alpha}}$ we obtain $\mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) \lesssim (\ln n)^d n^{-2+o(1)}$.
From Borrel-Cantelli and Corollary 1 we infer the proof of Theorem 1. \square

When independence is not assumed, the proofs are similar. We just have to take into account the \mathcal{E}' term and find suitable sequences ε_n and q_n in order to obtain convergent series and apply Borrel-Cantelli and Corollary 1.

5.4.2. Exponential decay

First we deal with the case of Exponential decay i.e. $\alpha(q)$ or $\theta(q)$ bounded by Bb^q with $0 < b < 1$

Proof. of Theorem 2.

1. Strong mixing: for all $R \in \mathbb{R}$ we have $\mathcal{E}'_R(q) \leq Bb^q$.

(a) Choose $\varepsilon_n = (\gamma(\ln n)^2/n)^{\frac{2}{d+1+2\alpha}}$ with $\gamma > \gamma_0^{ed} = \frac{2(2d+1+2\alpha)(3d+1+2\alpha)}{(d+1+2\alpha)^2 C \ln(b^{-1})}$, let write $\gamma = \gamma_0^{ed}(1+u)$.

Choose $q_n = \beta_1 \ln n$ with $\beta_1 = \frac{2(2d+1+2\alpha)}{(d+1+2\alpha)\ln(b^{-1})} \left(1 + \frac{u}{2}\right)$ finally introduce $v = \frac{u}{2+u}$. By Equation (17)

$$P_n \lesssim (\ln n)^d n^{-\frac{2d}{d+1+2\alpha}} \left(\exp\left(-\frac{C\gamma \ln n}{\beta_1} \left(1 - \frac{1}{\ln n}\right)^{\frac{d+1+2\alpha}{2}}\right) + \frac{n^{1+\beta_1 \ln(b)}}{(\ln n)^2} \right)$$

Because $d \geq 1$ and $\alpha \geq 0$ then $\frac{d+1+2\alpha}{2} \geq 1$ we have $\left(1 - \frac{1}{\ln n}\right)^{\frac{d+1+2\alpha}{2}} \geq 1 - \frac{d+1+2\alpha}{2} \frac{1}{\ln n}$ and thus

$$P_n \lesssim (\ln n)^d \left(n^{-1-\frac{3d+1+2\alpha}{d+1+2\alpha}v+o(1)} + n^{-1-\frac{2d+1+2\alpha}{d+1+2\alpha}u} \right) \quad (20)$$

we have $\sum_n P_n < +\infty$.

(b) Choose $r_n = \left(c \frac{(\ln n)^2}{n}\right)^{\frac{1}{d+\alpha}}$ and $q_n = \beta \ln n$ by Equation (18):

$$\mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) \lesssim (\ln n)^d n^{\frac{d}{d+\alpha}} \left(n^{-\frac{C'c}{\beta}+o(1)} + n^{\beta \ln(b)+1} \right)$$

choose first β large enough to have $\beta \ln(b) + 1 + \frac{d}{d+\alpha} \leq -2$ and second c large enough to have $-\frac{C'c}{\beta} + \frac{d}{d+\alpha} \leq -2$. We then obtain $\sum \mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) < +\infty$ (and $r_n \rightarrow 0$).

2. θ -weak dependence.

(a) Now $\mathcal{E}'_{R-\varepsilon_n/\ln n}(q_n) = \min_{t \in [0, R-\varepsilon_n/\ln n]} \left(\frac{Bb^{q_n}}{t} + c\omega_d 2^d t R^{d-1} \right)$. For any choice of sequence $\varepsilon_n \rightarrow 0$, for n large enough, $t_n = \sqrt{b^{q_n}}$ is in $[0, R-\varepsilon_n/\ln n]$ thus $\mathcal{E}'_{R-\varepsilon_n/\ln n}(q_n) \leq B' \exp(-q_n \ln(b^{-1})/2)$ thus as previously any $\varepsilon_n = (\gamma(\ln n)^2/n)^{\frac{2}{d+1+2\alpha}}$ with $\gamma > 2\gamma_0^{ed}$ gives sequence $\sum P_n$, with $\sum P_n < +\infty$.

(b) Choose again $r_n = \left(c \frac{(\ln n)^2}{n}\right)^{\frac{1}{d+\alpha}}$. Here also the choice of $t_n = \sqrt{b^{q_n}}$ is in $[0, r_n - r_n/\ln n]$ for n large enough. Thus $\mathcal{E}'_{r_n-r_n/\ln n}(q_n) \leq B' \exp(-q_n \ln(b^{-1})/2)$ and similar to 1.(b) we obtain $\sum \mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) < +\infty$ (and $r_n \rightarrow 0$) for suitable constants c .

To conclude, in all the cases we found $\varepsilon_n = O\left(\left(\frac{(\ln n)^2}{n}\right)^{\frac{2}{2+1+\alpha}}\right)$ such that $\sum P_n < +\infty$ and sequences $r_n \rightarrow 0$ with $\sum \mathbb{P}(d(\mathbb{X}_n, S) \geq r_n) < +\infty$ thus by Borrel Cantelli and Corollary 1 we achieve the proof. \square

Now we finish this section by the proof of Theorem 3 that deals with the Riemannian Decay.

5.4.3. Riemaniann decays

Proof. Theorem 3

1. Strong mixing: $\mathcal{E}'_R(q_n) \leq Bq_n^{-b}$ for a $b > 1$. Set $\lambda = \frac{3d+1+2\alpha}{d+1+2\alpha}$, $a = \frac{b-2-d}{b-\lambda}$ and $\gamma = \frac{\lambda+1}{\lambda+b}$. Choose $\varepsilon_n = \left(A \frac{(\ln n)^\alpha}{n^{1-\gamma}}\right)^{\frac{2}{d+1+2\alpha}}$ and $q_n = \frac{n^\gamma}{(\ln n)^{\alpha-1}}$. From Equation (17) it comes that

$$P_n \lesssim \left(\frac{\ln n}{\varepsilon_n}\right)^d n^{-AC} + \frac{1}{n(\ln n)^2}$$

So that, for A large enough $\sum P_n < +\infty$.

Similarly, Set $\lambda' = \frac{2d+\alpha}{d+\alpha}$, $a' = \frac{b-2-d}{b-\lambda'}$ and $\gamma' = \frac{\lambda'+1}{\lambda'+b}$, choose $r_n = \left(A' \frac{(\ln n)^{\alpha'}}{n^{1-\gamma'}}\right)^{\frac{1}{d+\alpha}}$ and $q_n = \frac{n^{\gamma'}}{(\ln n)^{\alpha'-1}}$. From Equation (17) it comes that

$$P(d_H(\mathbb{X}_n, S) \geq r_n) \lesssim \left(\frac{\ln n}{\varepsilon_n}\right)^d n^{-A'C'} + \frac{1}{n(\ln n)^2}$$

that, again allows to obtain summable series since A' is large enough.

2. θ -weak dependence

3. Now $\mathcal{E}'_{R-\varepsilon_n/\ln n}(q_n) = \min_{t \in [0, R-\varepsilon_n/\ln n]} \left(\frac{Bq_n^{-b}}{t} + c\omega_d 2^d t R^{d-1}\right)$. For any choice of sequence $\varepsilon_n \rightarrow 0$, for n large enough, $t_n = \sqrt{q_n^{-b}}$ is in $[0, R - \varepsilon_n/\ln n]$ thus $\mathcal{E}'_{R-\varepsilon_n/\ln n}(q_n) \leq B'q_n^{-b/2}$ and we can use the sequences exhibited in the strong mixing for changing b into $b/2$.

Similarly, Set $\lambda' = \frac{2d+\alpha}{d+\alpha}$, $a' = \frac{b/2-2-d}{b-\lambda'}$ and $\gamma' = \frac{\lambda'+1}{\lambda'+b/2}$, choose $r_n = \left(A' \frac{(\ln n)^{\alpha'}}{n^{1-\gamma'}}\right)^{\frac{1}{d+\alpha}}$ and $q_n = \frac{n^{\gamma'}}{(\ln n)^{\alpha'-1}}$. and $t_n = \sqrt{q_n^{-b}}$. Check that $t_n < \varepsilon_n$ From Equation (17) it comes that

$$P(d_H(\mathbb{X}_n, S) \geq r_n) \lesssim \left(\frac{\ln n}{\varepsilon_n}\right)^d n^{-A'C'} + \frac{1}{n(\ln n)^2}$$

that, again allows to obtain summable series since A' is large enough.

□

5.5. Proofs concerning level sets

Proof. Theorem 4

Let $R < m_t/f_{\max}^{(2)}$ be a positive constant, we first prove that, for n large enough:

$$\widehat{L}_{t,R} \subset L_t \oplus \frac{2\varepsilon'_n}{m_t} B \quad (21)$$

First by a trivial chain of inclusion, $\widehat{L}_{t,R} = C_R(\{\widehat{f}_n(X_i) \geq t\}) \subset C_R(\{X_i, f(X_i) \geq t - \varepsilon'_n\}) \subset C_R(L_{t-\varepsilon'_n})$. In the sequel we consider n large enough to have $\varepsilon'_n \leq \max(\Delta_t, t - t_0)$ so that Proposition 5 ensures that $C_R(L_{t-\varepsilon'_n}) = L_{t-\varepsilon'_n}$ thus we have $\widehat{L}_{t,R} \subset L_{t-\varepsilon'_n}$. Now for all $z' \in L_{t-\varepsilon'_n}$, introduce $z = z' + 2\frac{\varepsilon'_n}{m_t} \frac{\nabla f(z')}{\|\nabla f(z')\|}$ we have that

$$f(z) \geq f(z') + 2\frac{\varepsilon'_n}{m_t} \|\nabla f(z')\| - 2\left(\frac{\varepsilon'_n}{m_t}\right)^2 f_{\max}^{(2)} \geq t + \varepsilon'_n \left(1 - \frac{2\varepsilon'_n}{m_t^2} f_{\max}^{(2)}\right)$$

thus, since n large enough to also have $\varepsilon'_n < \frac{m_t^2}{2f_{\max}^{(2)}}$ we have $f(z) \geq t$ thus $L_{t-\varepsilon'_n} \subset L_t \oplus \frac{2\varepsilon'_n}{m_t} B$. that achieves the proof of Inclusion (21).

We now aim at proving that, with probability one for n large enough

$$L_t \subset \widehat{L}_{t,R} \oplus \left(\varepsilon_n + 2 \frac{\varepsilon'_n}{m_t} \right) B \quad (22)$$

Notice that $\mathbb{X}_n \cap L_{t+\varepsilon'_n} \subset \mathbb{X}_n^+(t)$ thus $C_R(\mathbb{X}_n \cap L_{t+\varepsilon'_n}) \subset \widehat{L}_{t,R}$. Let first prove that $C_R(\mathbb{X}_n \cap L_{t+\varepsilon'_n})$ contains $L_{t+\varepsilon'_n} \ominus \varepsilon_n B$ with probability one for n large enough. Proceeding by contradiction suppose that there exists $x \in L_{t+\varepsilon'_n}$ with $d(x, \partial L_{t+\varepsilon'_n}) \geq \varepsilon_n$ and $x \notin \widehat{L}_{t,R}$. Because $x \notin \widehat{L}_{t,R}$ there exists O with $\|x - O\| < R$ and $B(O, R) \cap \mathbb{X}_n^+(t) = \emptyset$ thus $B(O, R) \cap \mathbb{X}_n \cap L_{t+\varepsilon'_n} = \emptyset$. Moreover because $d(x, \partial L_{t+\varepsilon'_n}) \geq \varepsilon_n$ we have that $d(O, L_{t+\varepsilon'_n}) \leq R - \varepsilon_n$.

Introduce

$$P_n = \mathbb{P} \left(\exists O \in L_{t+\varepsilon'_n} \oplus (R - \varepsilon_n) B, B(O, R) \cap L_{t+\varepsilon'_n} \cap \mathbb{X}_n = \emptyset \right).$$

Cover $L_{t+\varepsilon'_n} \oplus (R - \varepsilon_n) B$ with v_n **deterministic** balls of radius $\varepsilon_n / \ln n$ and centered at $x_i \in L_{t+\varepsilon'_n} \oplus (R - \varepsilon_n) B$. Notice that we clearly have $|L_{t+\varepsilon'_n}| \leq t^{-1}$ (because the density on $L_{t+\varepsilon'_n}$ is greater than t). Thus by bound on packing and covering numbers it is possible with $v_n \leq \frac{2^d}{t \omega_d} \left(\frac{\ln n}{\varepsilon_n} \right)^d$.

$$P_n \leq \sum_{i=1}^{v_n} \mathbb{P} \left(B \left(x_i, R - \frac{\varepsilon_n}{\ln n} \right) \cap L_{t+\varepsilon'_n} \cap \mathbb{X}_n = \emptyset \right).$$

Now by proposition 5, $L_{t+\varepsilon'_n}$ has the $m_t/f_{\max}^{(2)}$ -inside and outside rolling ball property) and thus, let define $x_i^* = \pi_{\partial L_{t+\varepsilon'_n}}(x_i)$ and $y_i = x_i^* + \frac{m_t}{f_{\max}^{(2)}} \frac{x_i - x_i^*}{\|x_i - x_i^*\|}$ we have $B(y_i, m_t/f_{\max}^{(2)}) \subset L_{t+\varepsilon'_n}$ and $\|x_i - y_i\| \leq R - \varepsilon_n + m_t/f_{\max}^{(2)}$. Then Let $E_i = B(x_i, R - \varepsilon_n / \ln n) \cap B(y_i, m_t/f_{\max}^{(2)})$

$$P_n \leq \sum_{i=1}^{v_n} \left((1 - p_i)^{\frac{n}{q}} + \frac{\mathcal{E}_{E_i}(q)}{p_i} \right)$$

where $p_i = \mathbb{P}(X_1 \in E_i)$.

By Proposition 10 (with $f_0 = t, \alpha = 0$) it comes that

$$p_i \geq t C''_{d,0} \left(\frac{\min \left(R - \frac{\varepsilon_n}{\ln n}, \frac{m_t}{f_{\max}^{(2)}} \right)}{2} \right)^{\frac{d-1}{2}} \left\{ \min \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n}, \frac{\min \left(R - \frac{\varepsilon_n}{\ln n}, \frac{m_t}{f_{\max}^{(2)}} \right)}{2} \right) \right\}^{\frac{d+1}{2}}$$

Thus as ε_n converges to 0 for any $C < t C''_{d,0} \left(\frac{\min \left(R, \frac{m_t}{f_{\max}^{(2)}} \right)}{2} \right)^{\frac{d-1}{2}}$ for n large enough we finally obtain:

$$P_n \lesssim \left(\frac{\ln n}{\varepsilon_n} \right)^d \left(\exp \left(-C \frac{n}{q} \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right)^{\frac{d+1}{2}} \right) + \frac{\max_i \mathcal{E}_{E_i}(q)}{C \left(\varepsilon_n - \frac{\varepsilon_n}{\ln n} \right)^{\frac{d+1}{2}}} \right)$$

This equation is highly similar to Equation 17 with $\alpha = 0$. The only difference consist in \mathcal{E}_{E_i} which is now related to intersections of balls instead of Ball.

For strong mixing \mathcal{E}_E do not depend on the set E there is no change from the proof of Theorem 2, the given choices of ε_n (with $\alpha = 0$) ensures $\sum_n P_n < +\infty$.

Considering θ -weak dependence \mathcal{E}_E depend on the shape of the set E , nevertheless, due to Equations (3) and (4) we have that

$$\mathcal{E}_{E_i}(q) \leq \inf_{t_n \in (0, \varepsilon_n/2)} \left(A \frac{\theta(q)}{t_n} + B t_n \varepsilon_n^{\frac{d-1}{2}} \right),$$

for some constants A and B .

1. In case of exponential decay one can easily check that there exists suitable constants a_1, a_2 and a_3 such that $\varepsilon_n = (a_1 (\ln n)^2 / n)^{(d+1)/2}$, $t_n = \sqrt{a_2 \theta(q_n) \cdot (\varepsilon_n)^{(d-1)/2}}$ and $q_n = a_3 \ln n$ allows to have sequences such that $t_n \in (0, \varepsilon_n)$ and P_n satisfying $\sum P_n < +\infty$.

2. In case of Riemannian decay with $b > 2$, set $\lambda = \frac{5d-1}{2(d+1)}$, $\gamma = \frac{1+\lambda}{\frac{b}{2}+\lambda}$ and $a = 2^{\frac{2-d+b/2}{b-\lambda}}$ one can easily check that there exists a suitable constant a_1 such that $\varepsilon_n = (a_1(\ln n)^a/n^{1-\gamma})^{(d+1)/2}$, $q_n = n^\gamma/((\ln n)^{a-1})$, $t_n = \sqrt{a_2\theta(q_n).(\varepsilon_n)^{(d-1)/2}}$ and $q_n = a_3 \ln n$ allows to have sequences such that $t_n \in (0, \varepsilon_n)$ and P_n satisfying $\sum P_n < +\infty$.

That achieves the proof of

$$L_{t+\varepsilon'_n} \ominus \varepsilon_n B \subset C_R(\mathbb{X}_n \cap L_{t+\varepsilon'_n}) \subset \widehat{L}_{t,R}. \quad (23)$$

Now, considering $L_{t+\varepsilon'_n}$, as in the proof of Inclusion (21) we easily obtain $L_t \ominus \frac{\varepsilon'_n}{2m_t} B \subset L_{t+\varepsilon'_n}$ for n large enough to have $\varepsilon'_n < \frac{m_t^2}{2f_{\max}^{(2)}}$. Indeed, consider $\in L_t \ominus 2\frac{\varepsilon'_n}{m_t} B$ we have that for all vector u with $\|u\| \leq 1$, $f\left(z + 2\frac{\varepsilon'_n}{m_t} u\right) \geq t$ consider now the special choice of $u = -\frac{\nabla_f(z)}{\|\nabla_f(z)\|}$ we have

$$f(z) - 2\frac{\varepsilon'_n}{m_t} \|\nabla_f(z)\| + 2\left(\frac{\varepsilon'_n}{m_t}\right)^2 f_{\max}^{(2)} \geq f\left(z + \frac{\varepsilon'_n}{2m_t} u\right) \geq t$$

Thus

$$f(z) \geq t + 2\frac{\varepsilon'_n}{m_t} \|\nabla_f(z)\| - 2\left(\frac{\varepsilon'_n}{m_t}\right)^2 f_{\max}^{(2)} \geq t + \varepsilon'_n + \varepsilon'_n \left(1 - \frac{2\varepsilon'_n}{m_t^2} f_{\max}^{(2)}\right) \geq t + \varepsilon'_n.$$

We thus obtain, with Inclusion 23 that $(L_t \ominus 2\frac{\varepsilon'_n}{m_t} B) \ominus \varepsilon_n B \subset \widehat{L}_{t,R}$

Now by [32] and [18] because L_t has the inside and outside m_t/f_{\max}^2 rolling ball property, when $\varepsilon_n + 2\frac{\varepsilon'_n}{m_t} \leq m_t/f_{\max}^2$ it comes that

$$L_t \subset \widehat{L}_{t,R} \oplus \left(\varepsilon_n + 2\frac{\varepsilon'_n}{m_t}\right) B$$

Which concludes the proof of Inclusion (22) which together with Inclusion (21) allows to apply Corollary 1 and thus concludes the proof of $d_H(\widehat{L}_t, L_t) \lesssim \varepsilon_n + \varepsilon'_n$, $d_H(\partial\widehat{L}_t, \partial L_t) \lesssim \varepsilon_n + \varepsilon'_n$ and $\|\widehat{L}_t|_d, |L_t|_d\| \lesssim \varepsilon_n + \varepsilon'_n$ in Theorems 4 and 5.

To obtain homeomorphism and convergence for the perimeter one have to also prove that $d_H(L_t, \mathbb{X}_n \cap \{z, \widehat{f}(z) \geq t\}) \xrightarrow{a.s.} 0$ which as easily obtained by use technics extensively used in this paper is let for the reader. \square

Acknowledgements. This work was funded by CY Initiative of Excellence (grant "Investissements d'Avenir" ANR-16-IDEX-0008), Project "EcoDep" PSI-AAP2020-0000000013.

We also thank Guillaume Franchi (ENSAI) for his help wrt data analysis and modelling section.

References

- [1] C. Aaron and O. Bodart. Local convex hull support and boundary estimation. *J. Multivar. Anal.*, 147:82–101, 2016.
- [2] C. Aaron, A. Cholaquidis, and R. Fraiman. Estimation of surface area. *Electronic Journal of Statistics*, 2020.
- [3] L. Ambrosio, A. Colesanti, and E. Villa. Outer minkowski content for some classes of closed sets. *Mathematische Annalen*, 342:727–748, 2008.
- [4] D. W. K. Andrews. Non strong mixing autoregressive processes. *Journal of Applied Probability*, 21:930–934, 1984.
- [5] E. Arias-Castro, B. Pateiro-López, and A. Rodríguez-Casal. Minimax estimation of the volume of a set with smooth boundary. *arXiv: Statistics Theory*, 2016.
- [6] N. Baldin. The wrapping hull and a unified framework for estimating the volume of a body. *arXiv: Statistics Theory*, 2017.
- [7] H. Berbee. Chains with infinite connections: uniqueness and Markov representation. *Probab. Theory Related Fields*, 76(2):243–253, 1987.
- [8] A. Cholaquidis, R. Fraiman, G. Lugosi, and B. Pateiro-López. Set estimation from reflected brownian motion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78, 2014.
- [9] A. Cuevas. Set estimation: another bridge between statistics and geometry. *Bol. Estad. Investig. Oper.*, 25:71–85, 2009.
- [10] A. Cuevas and A. Rodríguez-Casal. On boundary estimation. *Advances in Applied Probability*, 36(2):340–354, 2004.
- [11] J. Dedecker, P. Doukhan, G. Lang, J. R. Leon, S. Louhichi, and C. Priour. Weak dependence: With examples and applications. In *Lecture Notes in Statistics*, volume 190. Springer, 2007.
- [12] L. Devroye and G. L. Wise. Detection of abnormal behavior via nonparametric estimation of the support. *SIAM Journal on Applied Mathematics*, 38(3):480–488, 1980.

- [13] P. Doukhan. *Mixing, properties and models*, volume 85 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1994.
- [14] P. Doukhan. *Stochastic models for time series*. Springer, 2018.
- [15] P. Doukhan and S. Louhichi. A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, 84:313–342, 1999.
- [16] P. Doukhan and S. Louhichi. Functional estimation of a density under a new weak dependence condition. *Scandinavian Journal of Statistics*, 28:325–341, 2001.
- [17] P. Doukhan and M.H. Neumann. Probability and moment inequalities for sums of weakly dependent random variables, with applications. *Stochastic Processes and their Applications*, 117:878–903, 2007.
- [18] H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.
- [19] W. M. Getz and C. C. Wilmers. A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*, 27:489–505, 2004.
- [20] W. Hardle, B.U. Park, and A.B. Tsybakov. Estimation of non-sharp support boundaries. *Journal of Multivariate Analysis*, 55(2):205–218, 1995.
- [21] S. Kallel and S. Louhichi. Topological reconstruction of compact supports of dependent stationary random variables. [arXiv:2307.11674v1 \[math.PR\]](https://arxiv.org/abs/2307.11674v1), 2023.
- [22] E. Liesbscher. Strong convergence of sums of α -mixing random variables with applications to density estimation. *Stochast. Process. Appl.*, 65:69–80, 1996.
- [23] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39:419–441, 2008.
- [24] R. Prakasa. *Nonparametric functional estimation*. Academic press, New York, 1983.
- [25] E. Rio. Théorie asymptotique des processus aléatoires faiblement dépendants. In *Mathématiques et Applications*, volume 31, Berlin, 2000. Springer.
- [26] A. Rodríguez Casal. Set estimation under convexity type assumptions. *Annales De L Institut Henri Poincare-probabilites Et Statistiques*, 43:763–774, 2007.
- [27] A. Rodríguez-Casal and P. Saavedra-Nieves. A data-adaptive method for estimating density level sets under shape conditions. *The Annals of Statistics*, 50(3):1653 – 1668, 2022.
- [28] A. Rodríguez-Casal and P. Saavedra-Nieves. Spatial distribution of invasive species: an extent of occurrence approach. *TEST*, 31:416–441, 2022.
- [29] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA*, 42:43–47, 1956.
- [30] G. G. Roussas. Asymptotical normality of positively or negatively associated processes. *Journal of Mult. Analysis*, 50:152–173, 1994.
- [31] G. Walther. Granulometric smoothing. *The Annals of Statistics*, 25(6):2273 – 2299, 1997.
- [32] Guenther Walther. On a generalization of blaschke’s rolling theorem and the smoothing of surfaces. *Mathematical Methods in The Applied Sciences*, 22:301–316, 1999.

Appendix A. Some bounds for the measure of the intersections

Proposition 9. *let $x, y \in \mathbb{R}^d$, and r_x, r_y, ℓ be positive numbers such that $\|x - y\| = r_x + r_y - \ell$, $\ell \leq \min(r_x, r_y)$ and $h \leq \min(r_x, r_y)$. let $\mathcal{V}(r_x, r_y, \ell, h) = |(B(x, r_x + h) \cap B(y, r_y + h)) \setminus (B(x, r_x - h) \cap B(y, r_y - h))|$, we have that, when $\ell \rightarrow 0$*

$$\mathcal{V}(r_x, r_y, \ell, h) \leq \omega_{d-1} \left(\frac{\ell 2r_x r_y}{r_x + r_y} \right)^{\frac{d-1}{2}} 4h(1 + o(1))$$

Proof. Let u_1 be a unit vector of \mathbb{R}^d we can define the portion of ball $B(O, r) \cap \{z, \langle z, u_1 \rangle \geq a\}$ and we have

$$V(r, a) = |B(O, r) \cap \{z, \langle z, u_1 \rangle \geq a\}|_d \omega_{d-1} \int_0^a (2rx - x^2)^{\frac{d-1}{2}} dx \quad (\text{A.1})$$

Thus by derivation

$$\begin{cases} 0 \leq \frac{\partial V(r, a)}{\partial a} \leq \omega_{d-1} (2ra)^{\frac{d-1}{2}} \\ 0 \leq \frac{\partial V(r, a)}{\partial r} \leq \omega_{d-1} \frac{2(d-1)}{d+1} (2r)^{\frac{d-3}{2}} (a)^{\frac{d+1}{2}} \end{cases} \quad (\text{A.2})$$

See the Figure ?? to be convinced that indeed:

$$|(B(x, r_x + h) \cap B(y, r_y + h))| = V\left(r_x + h, \frac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)}\right) + V\left(r_y + h, \frac{(\ell + 2h)(2r_x - \ell)}{2(r_x + r_y - 2h - 2\ell)}\right)$$

Subtracting $|B(O_x, r_x - h) \cap B(O_y, r_y - h)|$ we obtain the upper-bound.

$$\begin{aligned} & V\left(r_x + h, \frac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)}\right) + V\left(r_y + h, \frac{(\ell + 2h)(2r_x - \ell)}{2(r_x + r_y - 2h - 2\ell)}\right) \\ & - V\left(r_x - h, \frac{(\ell - 2h)(2r_y - \ell)}{2(r_x + r_y + 2h - 2\ell)}\right) - V\left(r_y - h, \frac{(\ell - 2h)(2r_x - \ell)}{2(r_x + r_y + 2h - 2\ell)}\right) \end{aligned}$$

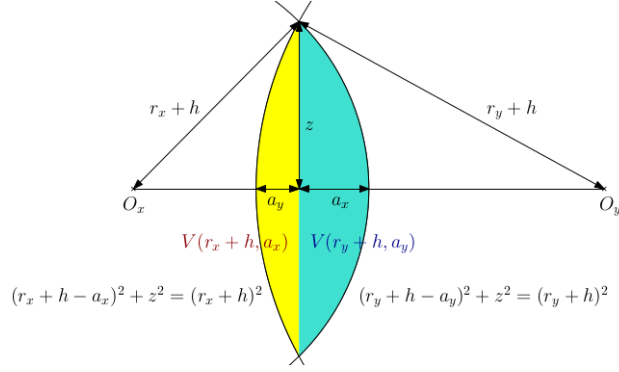


Fig. A.2: The convex body is $B(O_x, r_x + h) \cap B(O_y, r_y + h)$.

$$\text{Solving } \begin{cases} (r_x - h - a_x)^2 + z^2 = (r_x - h)^2 \\ (r_y - h - a_y)^2 + z^2 = (r_y - h)^2 \\ a_x + a_y = \ell + 2h \end{cases} \text{ yields } \begin{cases} a_x = \frac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)} \\ a_y = \frac{(\ell + 2h)(2r_x - \ell)}{2(r_x + r_y - 2h - 2\ell)} \end{cases}$$

as $\ell \rightarrow 0$, with $h \leq \ell/2$, let $a = \frac{\ell r_y}{r_x + r_y}$, $\Delta^+ = \frac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)} - a$ and $\Delta^- = a - \frac{(\ell - 2h)(2r_x - \ell)}{2(r_x + r_y + 2h - 2\ell)} - a$ (thus $\Delta^+ + \Delta^- \leq \frac{4hr_y}{r_x + r_y}(1 + o(1))$) we have

$$\begin{aligned} & V\left(r_x + h, \frac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)}\right) - V\left(r_x - h, \frac{(\ell - 2h)(2r_x - \ell)}{2(r_x + r_y + 2h - 2\ell)}\right) = V(r_x + h, \Delta^+) - V(r_x - h, a - \Delta^-) \\ & \leq 2h \frac{\partial V}{\partial r}(r_x, a) + (\Delta^+ + \Delta^-) \frac{\partial V}{\partial a}(r_x, a) \leq 2h\omega_{d-1} \frac{2(d-1)}{d+1} (2r_x)^{\frac{d-3}{2}} \left(\frac{2\ell r_y}{2r_x + r_y}\right)^{\frac{d+1}{2}} + \frac{4hr_y}{r_x + r_y}(1 + o(1))\omega_{d-1} \left(\frac{\ell 2r_x r_y}{r_x + r_y}\right)^{\frac{d-1}{2}}. \end{aligned}$$

We finally obtain that

$$\mathcal{V}(r_x, r_y, \ell, h) \leq \omega_{d-1} \left(\frac{\ell 2r_x r_y}{r_x + r_y}\right)^{\frac{d-1}{2}} 4h(1 + o(1))$$

which concludes the proof. \square

Proposition 10. Suppose that balls of radius R_i roll inside S . Suppose that the density is such that $f(x) \geq f_0 d(x, \partial S)^\alpha$. If O is such that $d(O, S) = R - \varepsilon$ with $0 < \varepsilon \leq R$ then, exists positive constants $C_{d,\alpha}$ and $C'_{d,\alpha}$ such that

$$\begin{cases} \mathbb{P}(B(O, R)) \geq f_0 C_{d,\alpha} \left(\frac{RR_i}{R+R_i}\right)^{\frac{d-1}{2}} \varepsilon^{\frac{d+1+2\alpha}{2}} \text{ if } O \notin S \text{ and } \varepsilon \leq \min(R, R_i)/4 \\ \mathbb{P}(B(O, R)) \geq f_0 C'_{d,\alpha} \left(\frac{\min(R_i, R)}{2}\right)^{d+\alpha} \text{ otherwise} \end{cases}$$

Proof. In all the proof we introduce O^* a point of ∂S such that $d(O, O^*) = d(O, \partial S)$ (not necessarily uniquely defines) and $O_i = O^* - R_i \eta_{O^*}$. By inside rolling ball condition $B(O_i, R_i) \subset S$ **First case $O \notin S$ and $\varepsilon \leq \min(R, R_i)/4$**

$$\mathbb{P}(B(O, R)) \geq \mathbb{P}(B(O_i, R_i) \cap B(O, R))$$

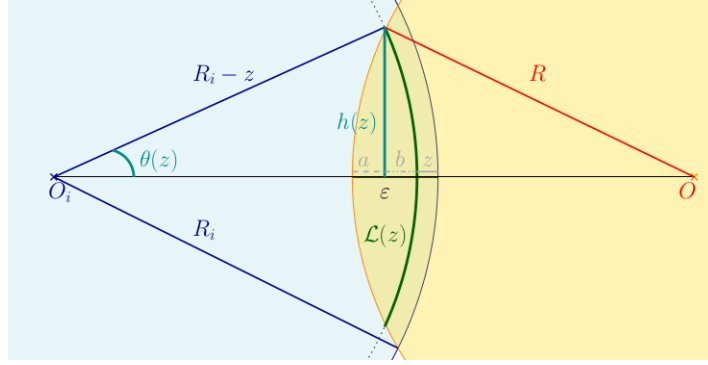
Decompose the calculus on small "lens" $\mathcal{L}(z) = S(O_i, R - z) \cap B(O, R)$ on witch $d(x, \partial S) \geq z$ (due to inside rolling ball properties), see Figure 1

it comes that

$$\mathbb{P}(B(O, R)) \geq \int_{z=0}^{\varepsilon} f_0 z^\alpha |\mathcal{L}(z)|_{d-1} dz$$

Now $|\mathcal{L}(z)|_{d-1} = \int_{\theta=0}^{\theta(z)} \sigma_{d-2}(R_i - z)^{d-1} \sin(\alpha)^{d-2} d\alpha$. By use of $\sin(\alpha) \geq \frac{2}{\pi}\alpha$ when $0 \leq \alpha \leq \pi/2$ (condition $\varepsilon < \min(R, R_i)$ allows that) we obtain that $|\mathcal{L}(z)|_{d-1} \geq \frac{\sigma_{d-2}}{d-1} \left(\frac{2\theta(z)}{\pi}\right)^{d-1} (R_i - z)^{d-1}$. Now

$$\theta(z) = \arcsin\left(\frac{h(z)}{R_i - z}\right) > \frac{h(z)}{R_i - z}$$



Thus

$$\mathbb{P}(B(O, R)) \geq f_0 \frac{\sigma_{d-2}}{d-1} \left(\frac{2}{\pi}\right)^{d-1} \int_{z=0}^{\varepsilon} h(z)^{d-1} dz$$

$$h(z) = \sqrt{\frac{(\varepsilon - z)(2R - \varepsilon + z)}{2(R_i + R - \varepsilon)} \left(2(R_i - z) - \frac{(\varepsilon - z)(2R - \varepsilon + z)}{2(R_i + R - \varepsilon)}\right)}$$

The condition $\varepsilon \leq \min(R_i, R)/4 \leq (R + R_i)/8$ we obtain that

$$h(z) \geq \sqrt{\frac{7 \times 17}{16 \times 14} \frac{RR_i}{R + R_i} (\varepsilon - z)} \geq \sqrt{\frac{1}{2} \frac{RR_i}{R + R_i} (\varepsilon - z)}$$

that finally gives:

$$\mathbb{P}(B(O, R)) \geq f_0 \frac{\sigma_{d-2}}{d-1} \left(\frac{2}{\pi}\right)^{d-1} \left(\frac{1}{2} \frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \int_0^{\varepsilon} z^{\alpha} (\varepsilon - z)^{\frac{d-1}{2}} dz$$

$$\mathbb{P}(B(O, R)) \geq f_0 \frac{\sigma_{d-2}}{d-1} \left(\frac{2}{\pi}\right)^{d-1} \left(\frac{1}{2} \frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \varepsilon^{\frac{d+1+2\alpha}{2}} B\left(\alpha + 1, \frac{d+1}{2}\right) \geq f_0 C_{d,\alpha} \left(\frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \varepsilon^{\frac{d+1+2\alpha}{2}}$$

Second case $O \notin S$ and $\varepsilon > \min(R, R_i)/4$

$$\mathbb{P}(B(O, R)) \geq \mathbb{P}(B(O_i, R_i) \cap B(O, R))$$

Now define $\Omega = O + (\varepsilon - \min(R, R_i)/4)\eta_{O^*}$ we have $\mathbb{P}(B(O, R)) \geq \mathbb{P}(B(O_i, R_i) \cap B(O, R)) \geq \mathbb{P}(B(O_i, R_i) \cap B(\Omega, R))$ and we can apply previous calculus

$$\mathbb{P}(B(O, R)) \geq f_0 C_{d,\alpha} \left(\frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \left(\frac{\min(R, R_i)}{4}\right)^{\frac{d+1+2\alpha}{2}}$$

$$\mathbb{P}(B(O, R)) \geq f_0 \frac{C_{d,\alpha}}{2^{\frac{3d+1+4\alpha}{2}}} (\min(R, R_i))^{d+\alpha}$$

Final case $O \in S$

1. if $d(O, \partial S) > R$ then $\mathbb{P}(B(O, R)) \geq \mathbb{P}(B(O, R/2)) \geq f_0 \left(\frac{R}{2}\right)^{\alpha} \omega_d \left(\frac{R}{2}\right)^d$
2. if $d(O, \partial S) < R$ and $R < R_i$, define $\Omega = O - \frac{R}{2}\eta_{O^*}$, then $\mathbb{P}(B(O, R)) \geq \mathbb{P}(B(\Omega, R/2)) \geq f_0 \left(\frac{R}{2}\right)^{\alpha} \omega_d \left(\frac{R}{2}\right)^d$
3. if $d(O, \partial S) < R$ and $R > R_i$, define $\Omega = O_i - \frac{R_i}{2}\eta_{O^*}$, then $\mathbb{P}(B(O, R)) \geq \mathbb{P}(B(\Omega, R_i/2)) \geq f_0 \left(\frac{R_i}{2}\right)^{\alpha} \omega_d \left(\frac{R_i}{2}\right)^d$

As a conclusion,

$$\mathbb{P}(B(O, R)) \geq f_0 \omega_d \left(\frac{\min(R_i, R)}{2}\right)^{d+\alpha}$$

□