Set estimation under dependence

Catherine Aaron, Paul Doukhan, Laurence Reboul

Set estimation under dependence

C. Aaron[a], P. Doukhan[b], L. Reboul[c]

[a] *University Clermont Auvergne*
[b] *CY Cergy Paris University, AGM Mathematics, 2 Bd. Adolphe Chauvin, 95000 Cergy-Pontoise, AGM-UMR 8080*
[c] *Institute of Mathematics of Marseille, 13288 Marseille Cedex 9, UMR 7373 and CY Cergy Paris University, AGM Mathematics*

Abstract

In this paper, we extend estimation results on the R-convex hull of the points of a random sample from independence to some dependent cases. Explicit convergence rates are obtained in that case; we deal both with R-convex hulls and with natural plug-in estimators of the support S, of its volume, its perimeter and the level-sets of the marginal density f on $\mathbb{R}^d$, of a strictly stationary process. For this we assume some weak dependence conditions. Convenient assumptions of f are needed on the boundary of its support. The loss with respect to the classical independent case is at at least a an extra power of $\ln n$.

Keywords: Strictly stationary process, strong mixing, weak dependence, multidimensional density, density estimation, density support estimation, set estimation

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 [21], 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 a given density distribution $f : \mathbb{R}^d \to \mathbb{R}$, 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 which is defined as the set $L_t = \overline{\{x, f(x) \geqslant t\}}$ for a given threshold $t > 0$. If $\int_{L_t} f(x)dx = \alpha$ then the level set L_t is also interpretated as a multidimensional $\alpha-$quantile, see [10, 14].
The estimation of a density support and of its related quantities -level sets, boundary, volume and perimeter- has been the focus of several works for i.i.d. data sets. The historical Devroye-Wise estimator of S in [11, 13] (see also references therein) has universal properties but non-optimal rates. If S is convex then optimal rates are achieved by the convex hull estimator 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 for example in [2, 5, 27–29, 32] (other estimators of S have also been proposed, as in [1] or [22]).
The R-convex hull of a set S is defined as :

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

where E^c and $\mathring{E}$ respectively denote the complement and the interior of the 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 the heuristic justification:
if S is regular enough, we have $C_R(S) = S$ so that a natural estimator of S is $\widehat{S}_n = C_R(\mathbb{X}_n)$, with $\mathbb{X}_n = \{X_1, \ldots, X_n\} \subset \mathbb{R}^d$ a sample whose marginal density f admits the support S and $C_R(\mathbb{X}_n) = \left(\bigcup_{\mathring{B}(x,R)\cap\mathbb{X}_n=\emptyset} \mathring{B}(x,R)\right)^c$.
The study of $\widehat{S}_n$ has been introduced by [32] and [27]. They derive convergence rates of the Hausdorff distance

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

$d_H(\widehat{S}_n, S)$ towards 0 and obtain the minimax rate of [22] under regularity assumptions on S. Similar rates are obtained for $d_H(\partial\widehat{S}_n, \partial S)$. Functionals of the R-convex hull may also be used to estimate related quantities such as the volume [5, 6], the perimeter [2] and level sets [28] of S.
In the above results, the density f is bounded below on S by some positive constant and the sample $\mathbb{X}_n$ is i.i.d. 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. Our hypothesis is that $\mathbb{X}_n$ is a $n-$sample from a strictly stationary process $\mathbf{X}$. Additionally, we assume some dependence condition: we consider here strong mixing and θ-weak dependence, introduced respectively in [30] and [12].
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 [23] in which the authors extend results on reconstruction of density supports of i.i.d random variables to supports of dependent stationary $\mathbb{R}^d$-valued random processes with the use of the Devroye-Wise estimator.
In this article, we study the asymptotic properties of $\widehat{S}_n$ and some of its natural plug-in as estimators of the support (and boundary, perimeter, volume) in a non-independent framework: thus we extend the results of [2, 5, 6, 27, 28, 32] to dependent cases. Plug-in estimates also lead to the case of level sets estimation. Namely, we derive convergence rates in Hausdorff distance of those estimators. Finally we show that, with probability one for n large enough, the estimated support (resp. level sets) is homeomorphic to the support (resp. to the level sets).
The rest of the paper is organized as follows. Section 2 presents some notation and general useful results. Section 3 dedicates to asymptotic results for the R-convex hull and its natural plug-in as the estimators of the density support and its related quantities, volume and perimeter. Section 4 deals with the estimation of level-sets. As a conclusion, Section 5 proposes some perspectives for data-driven tuning of the geometric parameter R. The proofs are postponed to Section 6.

2. General framework

In this paper, we are interested in several geometric non-parametric estimation problems, based on a $n-$sample $\mathbb{X}_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$ of strictly stationary weakly dependent sequence $\mathbf{X}$. 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

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

  For the sake of simplicity, we set $\omega_d = |B(0,1)|_d$ the volume of $B(0,1) \subset \mathbb{R}^d$ and $\sigma_d = |\mathcal{S}(0,1)|_{d-1}$ its surface.
- For each $E \subset \mathbb{R}^d$ and each positive number r, we denote by $E \oplus rB$ and $E \ominus rB$ the Minkowski sum and difference:

$$E \oplus rB = \{x \in \mathbb{R}^d,\, d(x, E) \leqslant r\} = \bigcup_{a \in E} B(a, r) \qquad \text{and} \qquad E \ominus rB = \{x, B(x, r) \subset E\}.$$

- For two subsets E and F of $\mathbb{R}^d$, we denote by $d_H(E, F)$ their Hausdorff distance:

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

- 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 $n \geqslant 0$, $u_n \leqslant c v_n$. If $u_n \lesssim v_n$ and $v_n \lesssim u_n$ then we set $u_n \approx v_n$.

- For $k \in \mathbb{N}$, we denote by C^k the set of k-times continuously differentiable functions $\mathbb{R}^d \to \mathbb{R}$.
  For $f \in C^1$ we denote by ∇_f its gradient and in case $f \in C^2$ we denote by $\mathcal{H}_f$ its Hessian matrix.
- We denote by $\|\cdot\|_{op}$ the operator norm on the set of $n \times p$ matrices.
- Let $\rho \geqslant 1$. 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 assume that $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ is a strictly stationary short memory process. We consider $\mathbb{X}_n = \{X_1, \dots X_n\} \subset \mathbb{R}^d$ a sample of $\mathbf{X}$ with marginal density $f : \mathbb{R}^d \to \mathbb{R}$.

We first recall below the short memory properties considered in the sequel, namely the strong mixing and θ-weak 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 $-\infty \leqslant j \leqslant l \leqslant +\infty$, let $\mathcal{F}_j^l = \sigma(X_k, 1 \leqslant k \leqslant l,\ k \in \mathbb{Z})$ be the σ-field of events generated by the random variables $(X_k)_{j\leqslant k\leqslant l}$. 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 $\mathbf{X}$, 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 $|\mathrm{cov}(\mathbf{1}_A, \mathbf{1}_B)|$. We say that $\mathbf{X}$ is strongly mixing (or α-mixing) if $\lim_{q\to\infty} \alpha(q) = 0$.

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

$$Lip(h) = \sup_{x\neq y} \frac{|h(x) - h(y)|}{\|x_1 - y_1\| + \cdots + \|x_v - y_v\|},$$

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 $\lim_{q\to\infty}\theta(q) = 0$ and a set of indices $i_1, \dots, i_v \geqslant 1$ such that $|\mathrm{cov}(W, h(X_{i_1}, \dots, X_{i_v}))| \leq v Lip\, h \cdot \theta(q)$. Notice that for $v = 1$ we have $|\mathrm{cov}(W, h(X_q))| \leq Lip\, h \cdot \theta(q)$.

Examples. A large number of classical models which satisfy strong mixing conditions is provided in [15]; 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 strongly mixing. A famous counterexample is in [4], in which $\mathbf{X}$ is the stationary solution of the recursion

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

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 is hereditary through Lipschitz functions.

2.2.2. A covariance based result

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

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

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

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

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

$$\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \leq (1-p)^{\frac{n}{q}} + \frac{\mathcal{E}_E(q)}{p}, \quad \textit{with} \tag{3}$$

$$\mathcal{E}_E(q) \leqslant \inf_{\varepsilon\in(0,\ell/2)}\left(\frac{2\theta(q)}{\varepsilon} + c\mathcal{V}(r_x, r_y, \ell, \varepsilon)\right), \quad \textit{and} \quad \mathcal{V}(r_x, r_y, \ell, \varepsilon) \underset{\ell\to 0}{\sim} 4c\omega_{d-1}\varepsilon\left(\frac{\ell 2 r_x r_y}{r_x + r_y}\right)^{\frac{d-1}{2}}. \tag{4}$$

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

Another important dependent case is that of negatively associated processes, see [31]. Considering the proof, we set in this case $q = 1$ and $k = n$ and here

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

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 and we obtain:

Proposition 3. *Let $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ be a negatively associated. Then, setting $\mathcal{E}_E(q) = \alpha(q)$, we have:*

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

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 is if $\mathbf{X}$ is absolutely regular, [15], Berbee's coupling inequalities ([7]) yields $\mathcal{E}_E(q) \leq \beta(q)$.

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,* $\alpha(q) \leqslant Bb^{-q}$ or $\theta(q) \leqslant Bb^{-q}$, for some $b > 1$.
- *Riemannian decays,* $\alpha(q) \leqslant Bq^{-b}$ or $\theta(q) \leqslant Bq^{-b}$, 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

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

where $f : \mathbb{R}^d \to \mathbb{R}$ as above is the marginal density of the strictly stationary $\mathbb{R}^d$-valued process $\mathbf{X}$ and $\widehat{f_n}$ is its kernel density estimator:

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

We provide below bounds for ε'_n in the α-mixing and θ-weakly dependent cases. Proposition 4 concerns exponential decays and Proposition 5 concerns Riemannian decays. For that task, we fist introduce some regularity assumptions on f and K:

(K1) $K : \mathbb{R}^d \to \mathbb{R}$ is a bounded density with compact support.

(K2) $K : \mathbb{R}^d \to \mathbb{R}$ is of order $(k-1)$.

(M1) For some $k \geqslant 2$, $f \in C^k$ 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 $c > 2$.

Proposition 4. *Let $\mathbb{X}_n$ be a n-sample of a strictly stationary sequence $\mathbf{X}$ 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 $k \geqslant 2$.*

(A1) *If $\mathbf{X}$ is strongly mixing with $\alpha(q) = O(b^{-q})$ 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 $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k}{2k+d}}\right)$.*
2. *If $k \leqslant d+3$, any choice $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1-\gamma}{3d+6}}$ with $\gamma \in (0,1)$ yields $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k(1-\gamma)}{3d+6}}\right)$.*

(T1) *If $\mathbf{X}$ is θ-weakly dependent with $\theta(q) = O(b^{-q})$ 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 $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k}{2k+d}}\right)$.*
2. *If $k \leqslant d+2$, any choice $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1-\gamma}{3d+4}}$ with $\gamma \in (0,1)$ yields $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k(1-\gamma)}{3d+4}}\right)$.*

Proposition 5. *Let $\mathbb{X}_n$ be a n-sample of a strictly stationary sequence $\mathbf{X}$ 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 $k \geqslant 2$.*

(A2) *Let $\mathbf{X}$ be strongly mixing with $\alpha(q) = O(q^{-b})$ for some $b > \max(2+3/d, d+2)$. Let $\gamma_0 = \frac{d+2-\frac{d(d-2)}{2k+d}}{b+1-\frac{d(d-2)}{2k+d}}$.*

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

(T2) *Let $\mathbf{X}$ be θ-weakly dependent with $\theta(q) = O(q^{-b})$ for some $b > 3(d+2)$. Let $\gamma_0 = \dfrac{d+2-\frac{d(d-2)}{2k+d}}{2\lfloor\frac{b+3}{6}\rfloor - \frac{d(d-2)}{2k+d}}$.*

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

3. Support estimation

In this section, we derive convergence rates for the R-convex hull estimator of S and that of their boundaries in the context of strictly stationary strong-mixing or θ-weakly dependent sequences $\mathbf{X}$. We first give regularity conditions on S under which are derived our main theoretical results.

3.1. Distributional background

In order to obtain a consistent estimator $C_R(\mathbb{X}_n)$ of the support S we will assume that S is compact and regular enough in order that $C_R(S) = S$. This will fulfilled if the following rolling ball property holds:

Definition 1 (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:*

$$\forall x \in \partial S,\ \exists O_x^o \in \mathbb{R}^d, \|x - O_x^o\| = R_o, \ \text{ and } \ \mathring{B}(O_x^o, R_o) \subset \mathring{S}^c. \ (resp.\ \exists\, O_x^i \in \mathbb{R}^d,\ \|x - O_x^i\| = R_i, \text{ and } B(O_x^i, R_i) \subset S).$$

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 Blaschke's rolling ball's theorem ([32]'s theorem 1). In particular, they satisfy $S_{R_o}(S) = S$, which ensures the convergence of the plug-in support estimator $S_{R_o}(\mathbb{X}_n)$ since $d_H(\mathbb{X}_n, S) \to 0$.
To derive explicit convergence rates we have moreover to assume that the density decreases regularly enough. Specifically we will assume a power decay as in [22]:

Definition 2. *A density f belongs to the class $\mathcal{M}_{R_o,R_i,f_0,\alpha}$ if its support S is compact and satisfies the R_o-outside rolling ball property and the R_i-inside rolling ball property and that*

$$f(x) \geqslant f_0\, d(x, \partial S)^\alpha, \quad \forall x \in S.$$

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

3.2. Main Theoretical results

Hereafter, we provide explicit convergence rates that generalize to a dependent framework and to $\alpha \geqslant 0$ in Definition 2 the previous works of [2, 27] on the R-convex hull, obtained in the independent case and when $\alpha = 0$. As a reference point, we first extend these previous results to the i.i.d. case but with $\alpha \geqslant 0$:

Theorem 1. *Let R_i, R_o and f_0 be positive constants and $\alpha \geqslant 0$. Let $\mathbb{X}_n = \{X_1, \ldots, X_n\}$ be an i.i.d. $n-$sample with density $f \in \mathcal{M}_{R_o,R_i,f_0,\alpha}$ supported by S and $R < R_o$. Thus, there exists a constant $A(d, \alpha, f_0, R, R_i)$ such that for n large enough, one has with probability* 1

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

with

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

Remark 3 (Case of negatively associated processes). In case of negatively associated process as in [31] we obtain the same rates as in the i.i.d. case.

The two theorems below extend Theorem 1 to samples $\mathbb{X}_n$ of a strictly stationary strong mixing or $\theta-$weakly dependent sequence. Theorem 2 deals with exponential decay, and Theorem 3 focuses on the Riemannian decay of the dependence coefficients.

Theorem 2. *Let R_i, R_o and f_0 be positive constants and let $\alpha \geqslant 0$. Let $\mathbb{X}_n = \{X_1, \dots, X_n\}$ be a n−sample of a strictly stationary process $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ with marginal density $f \in \mathcal{M}_{R_o,R_i,f_0,\alpha}$ supported by S, with $R < R_o$. Assume that $\mathbf{X}$ is strongly mixing or θ-weakly dependent with an **exponential decay** rate of its sequence of dependence coefficients. Then, there exists a constant $A(d, \alpha, f_0, R, R_i, b)$ such that for large enough n, one has with probability* 1

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

with

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

Theorem 3. *Let R_i, R_o and f_0 be positive constants and $\alpha \geqslant 0$. Let $\mathbb{X}_n = \{X_1, \dots, X_n\}$ be a n−sample of a strictly stationary process $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ with marginal density $f \in \mathcal{M}_{R_o,R_i,f_0,\alpha}$ supported by S, with $R < R_o$. Assume that $\mathbf{X}$ is strongly mixing or θ-weakly dependent with **Riemannian decay** rates. Then, there exists b_0 such that for all $b > b_0$ and large enough n one has, with probability* 1

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

with

$$\varepsilon_n = \left(\frac{\ln^a n}{n^{1-\gamma}}\right)^{\frac{2}{d+1+2\alpha}}.$$

Explicit values of b_0, a and γ are given by

1. $b_0 = 1$, $\lambda = \dfrac{3d+1+2\alpha}{d+1+2\alpha}$, $a = \dfrac{b-2-d}{b-\lambda}$ *and* $\gamma = \dfrac{\lambda+1}{\lambda+b}$, *in the strongly mixing case,*
2. $b_0 = 2$, $\lambda = \dfrac{3d+1+2\alpha}{d+1+2\alpha}$, $a = \dfrac{b/2-2-d}{b/2-\lambda}$ *and* $\gamma = \dfrac{\lambda+1}{\lambda+b/2}$, *in the θ-weakly dependent case.*

Theorem 2 highlights that in the exponential decay case, the estimation error achieves, up to an additional power of $\ln n$, the same rate than in the independent setting whatever the dependence structure is (strong mixing or $\theta-$ weak dependence). In the exponential decay case, the dependence structure only impacts constants. Under Riemannian decays, Theorem 3, the loss in convergence rate with respect to the independent case is of the order of an additional power of n. Moreover, contrary to the exponential decay case, both the error and the loss depend on the dependence structure.

Moreover, as in [28] our results highlight the different roles of the outside and inside radius R_o and R_i: in both cases, the tuning parameter R must be less than R_o to achieve the convergence and the convergence rate depends on R_i.

4. Level set estimation

We now focus on level set estimation. For a given $t \in \mathbb{R}$, the level set L_t and its associated level line ℓ_t are defined as

$$L_t = \{z \in \mathbb{R}^d, f(z) \geqslant t\}, \qquad \ell_t = \{z, f(z) = t\}.$$

Let $\widehat{f_n}$ be a suitable density estimator of f. Thus, [28] introduced the following empirical level set estimator

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

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

4.1. Distributional background

When dealing with level set estimation, the distributional hypotheses are somehow different. Namely, the support of the distribution does no longer need to be compact since $L_t = \overline{\{z, f(z) > t\}}$ is compact for all $t > 0$ ($t = 0$ is the support estimation framework). Moreover, 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 $k \geqslant 2$ which has two important consequences. Firstly, it allows the density estimator to converge with suitable rates. Secondly, it can be shown that for suitable values of t (see Definition 3 below), sets L_t satisfy the inside and outside rolling ball property.

Definition 3 (Regular levels)**.** *The level $t > 0$ is a regular level if there exist $\Delta_t > 0$ and $m_t > 0$ such that*

$$L_{t-\Delta_t} \subset \mathring{S}, \quad \text{and} \quad \min_{z \in L_{t-\Delta_t} \setminus \mathring{L}_{t+\Delta_t}} \|\nabla_f(z)\| = m_t > 0.$$

If t is a regular level and f is C^k with $k \geqslant 2$, then by Theorem 2 in [32], the level set L_t naturally inherits of the inside and outside rolling ball property.

Proposition 6 (Corollary of Theorem 2 in [32])**.** *If $f \in C^2$ and $\sup_z \|\mathcal{H}_f\| = f^{(2)}_{\max} < +\infty$ 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 $m_t / f^{(2)}_{\max}$.*

4.2. Main Theoretical results

Hereafter, we establish the asymptotic behavior of $\widehat{L}_{t,R}$ under weak dependence *assumptions* and deduce asymptotics for its boundary and volume. By incorporating tools provided in [2], we also obtain topological guarantees and perimeter estimation rates. We show that the convergence rate can be decomposed as $\varepsilon_n + \varepsilon'_n$, where ε'_n is a density estimation error and ε_n is a set estimation error. The density estimation error depends on the dependence structure, the regularity of the density and the dimension d (see Section 2.2.3). We focus in this first 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* **X** ***with exponential decay rates*** *and with marginal class C^2 density f on its support S. Assume that $\sup_S \|\nabla_f\| = f^{(1)}_{\max} < +\infty$, $\sup_S \|\mathcal{H}_f\|_{op} = f^{(2)}_{\max} < +\infty$. Let $t > 0$ be a regular level. Let R be a constant such that $0 < R < m_t / f^{(2)}_{\max}$ and define $\mathbb{X}^+_n(t) = \mathbb{X}_n \cap \{\widehat{f}(z) \geqslant t\}$ and $\widehat{L}_{t,R} = C_r(\mathbb{X}^+_n(t))$. Then, with probability 1 and for n large enough, one has*

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

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

We now focus on Riemmanian decays. As for the support estimation problem, Riemmanian decays degrade the convergence and the convergence rates precisely 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 $\sup_S \|\nabla_f\| = f^{(1)}_{\max} < +\infty$, $\sup_S \|\mathcal{H}_f\|_{op} = f^{(2)}_{\max} < +\infty$. Let $t > 0$ be a regular level. Let R be a constant such that $0 < R < m_t / f^{(2)}_{\max}$ and define $\mathbb{X}^+_n(t) = \mathbb{X}_n \cap \{\widehat{f}(z) \geqslant t\}$ 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*

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

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

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

4.3. Examples of explicit rates under exponential decays

In case of exponential decays, for small dimensions, $d \leqslant 3$, the convergence rate is the density estimation rate (whatever the regularity of the density is).
For larger dimensions the limiting rate is the set estimation rate when density is regular enough. More precisely, depending on the dependency model, the dimension and the regularity of the density. The rates are:

1. Under strong mixing
   (a) if $d \leqslant 3$, then the convergence rate is limited by the density estimation step.
   Namely for some $a > 0$:
      i. if $k \leqslant d+3$, then for all $k' \in (0,k)$, we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k'}{3d+6}})$.
      ii. if $k > d+3$, then $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k}{d+2k}})$;
   (b) if $d \geqslant 4$,
      i. when $k \leqslant \min\left(d+3, 6\cdot\frac{d+2}{d+1}\right)$, then for all $k' \in (0,k)$, we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k'}{3d+6}})$.
      ii. when $k > \min\left(d+3, 6\cdot\frac{d+2}{d+1}\right)$, then we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{2/(d+1)})$.
2. Under θ-weak dependence then for some $a > 0$:
   (a) if $d \leqslant 3$, the convergence rate is limited by the density estimation step. Namely
      i. when $k \leqslant d+3$, then for all $k' \in (0,k)$, we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k'}{3d+6}})$.
      ii. when $k > d+3$, then $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k}{d+2k}})$;
   (b) if $d = 4$,
      i. when $k < 7$, then for all $k' \in (0,k)$, we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k'}{18}})$.
      ii. when $k = 7$, we have $\varepsilon_n + \varepsilon'_n = O(\ln^a n/n)^{\frac{7}{19}})$.
      iii. when $k \geqslant 8$, then we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{2}{5}})$.
   (c) if $d > 4$,
      i. when $k \leqslant \min\left(d+2, 6\cdot\frac{d+2}{d+1}\right)$, then for all $k' \in (0,k)$, we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{k'}{3d+6}})$.
      ii. when $k > \min\left(d+2, 6\cdot\frac{d+2}{d+1}\right)$, then we have $\varepsilon_n + \varepsilon'_n = O((\ln^a n/n)^{\frac{2}{d+1}})$.

5. Perspectives

If the data are i.i.d, there exists data-driven way to tune the geometric parameter R. When dealing with support estimation, in [29] a method based on the maximal spacings is proposed. This approach is very difficult to adapt to the dependent case. Indeed we can not expect the maximal spacing statistics to have the same limit law if the data are dependent. An other approach based on reach estimation [9] should be more promising. When dealing with level set estimation, a generalization of the proposition in [28] may work since it only relies on a tricky splitting of the sample which could be done in even with dependent data.

6. Proofs

6.1. Proof of Propositions 1, 2, and 3.

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

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

The propositions will result from proving that $\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, Z_k)$ 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 $\alpha(m)$ and the fact that $Z_1 \times \cdots \times Z_{k-1}$ writes as the indicator function of $\bigcap_{1\leq i\leq n/q}(X_{t_i} \notin E)$ implies $|\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, Z_k)| \leq \alpha(q)$, 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 $\theta-$weak dependence we have by definition $|\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, h(X_k))| \leq \theta(q)Lip\,(h)$ 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 [12] we consider $H^-_{x,r,\varepsilon}$ and $H^+_{x,r,\varepsilon}$ two Lipschitz approximations $\mathbb{R}^d \to \mathbb{R}$ 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
   $$\begin{aligned}\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, Z_k) &\leq \mathbb{E}(Z_1 \times \cdots \times Z_{k-1}H^+_{x,r,\varepsilon}(X_k)) - \mathbb{E}(Z_1 \times \cdots \times Z_{k-1})\mathbb{E}(H^-_{x,r,\varepsilon}(X_k)) \\ &\leq \mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, H^+_{x,r,\varepsilon}(X_k)) + \mathbb{E}(H^+_{x,r,\varepsilon}(X_k) - H^-_{x,r,\varepsilon}(X_k))\end{aligned}$$
   Now we build such Lipschitz functions with $Lip\,(H^\pm_{x,r,\varepsilon}) \lesssim 1/\varepsilon$. More precisely, we introduce intermediary functions:
   $$m^-_{r,\varepsilon}(s) = \begin{cases} 1, & s < (r-\varepsilon) \\ \dfrac{r-s}{\varepsilon}, & s \in [r-\varepsilon, r] \\ 0, & s > r \end{cases} \quad \text{and } m^+_{r,\varepsilon}(s) = \begin{cases} 1, & s < r \\ 1 - \dfrac{s-r}{\varepsilon}, & s \in [r, r+\varepsilon] \\ 0, & s > r+\varepsilon \end{cases}$$
   Then the functions $m^\pm_{r,\varepsilon}$ are piecewise linear and satisfy $Lip\, m^\pm_{r,\varepsilon} = 1/\varepsilon$. 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 $Lip\,(H^\pm_{x,r,\varepsilon}) = 1/\varepsilon$. 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.
   $$\begin{aligned}\mathbb{E}\left(H^+_{x,r,\varepsilon}(X_k) - H^-_{x,r,\varepsilon}(X_k)\right) &\leq c \cdot \omega_d\left((r+\varepsilon)^d - ((r-\varepsilon)^+)^d\right) \\ &\leq c \cdot \omega_d r^d\left((1+\varepsilon/r)^d - ((1-\varepsilon/r)^+)^d\right) \\ &\leq c \cdot \omega_d 2^{d+1} r^{d-1}\varepsilon\end{aligned}$$
   for all $\varepsilon \leqslant r$, since a binomial expansion entails $(1+x)^d - (1-x)^d \leqslant 2x\sum_{k=1}^d \binom{d}{k} \leqslant 2^{d+1}x$ if $x \in [0,1]$. Thus, setting $A' = c\omega_d 2^{d+1}$, we derive for all $\varepsilon \leqslant r$: $\mathrm{cov}(Z_1 \cdots Z_{k-1}, Z_k) \leq \frac{1}{\varepsilon}\theta(q) + A'\varepsilon r^{d-1}$ For the reverse inequality, we obtain analogously for all $\varepsilon \in [0, r]$ $\mathrm{cov}(Z_1 \cdots Z_{k-1}, Z_k) \geq -\frac{1}{\varepsilon}\theta(q) - A'\varepsilon r^{d-1}$ Indeed
   $$\begin{aligned}\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, Z_k) &\geq \mathbb{E}(Z_1 \times \cdots \times Z_{k-1}H^-_{x,r,\varepsilon}(X_k)) - \mathbb{E}(Z_1 \cdots Z_{k-1})\mathbb{E}H^+_{x,r,\varepsilon}(X_k) \\ &\geq \mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, H^-_{x,r,\varepsilon}(X_k)) - \mathbb{E}(H^+_{x,r,\varepsilon}(X_k) - H^-_{x,r,\varepsilon}(X_k))\end{aligned}$$
   and thus:
   $$\mathcal{E}_E(q) \leqslant \min_{\varepsilon\in[0,r]} \frac{\theta(q)}{\varepsilon} + A'\varepsilon r^{d-1}\,.$$

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

The proof of Proposition 3 is straightforward and left to the reader. □

6.2. Proofs of Propositions 4 and 5.

In the following we set $\varepsilon'_n = A_n + B_n$ 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 4 and 5 are direct corollaries of Propositions 7 and 8 below. Proposition 7 gives a control of B_n, while item 1 and item 2 of Proposition 8 give a control of A_n in the geometric and Riemannian decay cases respectively.

Proposition 7. *Let $\mathbb{X}_n$ be a sample from a strictly stationary sequence $\mathbf{X}$ with marginal density f and $\widehat{f_n}$ be defined by* (5)*, with $h_n \to 0$ and $nh_n^d \to \infty$ as n goes to ∞. Assume that (K1), (M1) and (K2) hold for some $k \geqslant 2$. Thus, $B_n = O(h_n^k)$*

Proposition 8. *Let $\mathbb{X}_n$ be a sample from a strictly stationary sequence $\mathbf{X}$ 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 $\alpha(q) = O(b^{-q})$ for some $b > 1$ and $h_n = O\left(\left(\frac{\ln n}{n}\right)^\beta\right)$ for some $\beta < 1/(3d + 6)$.*
   or

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

   Then

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

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

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

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

   Then

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

Remark 4. 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 [16] 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 $0 < c \leqslant 1$ where $\rho = \lfloor\rho\rfloor + c \geqslant 2$ 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 4. Let us prove (A1) (the case (T1) can be handled in the same way). By Proposition 7 and item 1 of Proposition 8,

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

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 $\varepsilon_n' = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k}{2k+d}}\right)$.

- If $k \leqslant d+3$, thus $1/(2k+d) \geq 1/(3d+6)$ and $\varepsilon_n' = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\beta k}\right)$ and the result is obtained by setting $\beta = (1-\gamma)/(3d+6)$ for some $\gamma \in (0,1)$.

□

Proof of Proposition 5. To prove item (A2) of Proposition 5 for any $b > \max((2d+3)/d, d+1)$, set $q = b+1-t$ for a $t \in (0, b-d-1)$ and then apply item (A2) of Proposition 8 and Proposition 7 to obtain for all $0 < \beta < 1/d$:

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

for all $\gamma > \frac{d+2-\beta d(d-2)}{b+1-t-\beta d(d-2)}$. then choose $\beta = \frac{1}{d+2k}$ and let $t \to 0$ to end the proof.
To prove item (T2) of Proposition 5 first notice that, to find an even number q larger than $d+2$ and smaller than $b/3+1$ we need the condition $b > 3(d+3)$. Assume now that $b > 3(d+3)$ and introduce $q^* = 2\lfloor\frac{b+3}{6}\rfloor$ the largest even number smaller than $b/3+1$ then apply item (T2) of Proposition 8 and Proposition 7 to obtain, for all $0 < \beta < 1/d$

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

for all $\gamma > \dfrac{d+2-\beta d(d-2)}{q^* - \beta d(d-2)}$. then choose $\beta = \frac{1}{d+2k}$ to achieve the proof.

□

Proof of Proposition 7.

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

Since $f \in C^k$, $k \geqslant 2$ 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),

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

Therefore,

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

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 8. 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

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

with $\nu > 2/c$ (c defined in (M5)), for some appropriate $\lambda(n) \to \infty$ 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

$$L_n = \sup_{\|x\|\geqslant n^\nu} \frac{\sqrt{nh^d}}{\lambda(n)} |\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))| \to 0 \tag{7}$$

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 $h_n \leq 1$,

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

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

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

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

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\max_{1\leqslant t\leqslant n} \|X_t\| > n^\nu - Ch_n\right) \leqslant \sum_{n=1}^{\infty} 1 \wedge \frac{\mathbb{E}\|X_1\|^c + \cdots + \mathbb{E}\|X_n\|^c}{(n^\nu - Ch_n)^c} < \infty,$$

for ν such that $\nu > 2/c$ so that the series converges by (M5). This with the Borel–Cantelli lemma entails that $\sup_{\|x\|\geqslant n^\nu} |\widehat{f_n}(x)|$ tends to 0 as n goes to infinity a.s. as fast as needed. Indeed $\|X_t - x\| > Ch_n$ 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 $x \in \mathbb{R}^d$,

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

To prove (6), we use the chaining argument of [24]. Let $I = [-M, M]^d$ be a cube of $\mathbb{R}^d$. Since I is compact, it can be covered by δ^d cubes $I_1, \ldots, I_{\delta^d}$ with centers $x_1, \ldots, x_{\delta^d}$ and sides of length $2M/\delta$, where δ is such that $\delta h/M \to \infty$. 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), [24] proves that for some positive constant L and for all $x \in I_j$,

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

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 $\|x\| \leqslant L$ and $\widetilde{K}(x) = 0$ otherwise, such that for some $n \geqslant n_0$ (see [24] p 78),

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

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

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

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

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

We then set $t = C\dfrac{\lambda(n)}{\sqrt{nh^d}}$ for some appropriate $\lambda(n) \to \infty$ such that $\lambda(n) = o(\sqrt{nh^d})$ and $M = n^{\nu/d}$ 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 8 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* $h_n \to 0$ *and* $nh_n^d \to \infty$*, 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* $\theta-$*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* $\theta-$*weakly dependent with* $\theta(m) = O(m^{-b})$ *for some* $b > 2 + 2/d$.

Thus one has

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

Proof of item 1 in Proposition 8. In cases of a 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 [19]. Let us recall them below in our special setting

Lemma 7 (Proposition 8 in [19])**.** *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* $\|Y\|_\infty \leqslant H$ *in probability. Assume that there exists some* $0 < c < 1$ *and some* $L > 0$ *such that for all* u*-tuples* $(s_1, ..., s_u)$ *and all* v*-tuples* $(t_1, ..., t_v)$ *with* $1 \leqslant s_1 \leqslant \ldots \leqslant s_u \leqslant t_1 \leqslant \ldots \leqslant t_v \leqslant n$,

$$\left|cov\left(Y_{s_1}\ldots Y_{s_u}, Y_{t_1}\ldots Y_{t_v}\right)\right| \leqslant H^{u+v-2}L^2 vc^{(t_1-s_u)}. \qquad (10)$$

Then, for all $t > 0$,

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

with

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

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 $\|Y_i(x)\|_\infty \leqslant H = \|F_n\|_\infty \leqslant 2\|K\|_\infty/nh^d$ and $LipF_n \leqslant LipK/(nh^{d+1})$. Assume that

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

- *Strong mixing case.* If $\mathbf{X}$ is strong mixing with $\alpha(m) = O(b^{-m})$, $b > 1$, thus $\mathbf{Y}$ is strong mixing with $\alpha_Y(m) \leqslant \alpha(m)$ by measurability of F_n. Moreover, by [17]'s Lemma 6 and [19]'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 $t = C\sqrt{\ln n/nh^d}$ for some convenient constant $C > 0$, we get

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

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

- *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 $\theta_Y(m) \leqslant \theta(m)$. Moreover, using Remark 9 and Proposition 8 of [19], (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 $t = C\sqrt{\ln n/nh^d}$ for some convenient constant $C > 0$, we get

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

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

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 (9) by setting $M = n^{\nu/d}$,

$$\mathbb{P}\left(\sup_{\|x\| \leqslant n^\nu} |\widehat{G}_n(x)| > t\right) \leqslant \delta^d n^{-D} = An^{\nu+\varepsilon\beta d-D}(\ln n)^{-\beta\varepsilon d},$$

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

$$\sup_{\|x\| \leqslant n^\nu} |\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))| = O_{a.s.}\left(\sqrt{\frac{\ln n}{nh_n^d}}\right),$$

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

Proof of Proposition 8, item 2. Our proof relies on the following lemma.

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 $h_n \to 0$ and $nh_n^d \to \infty$. Assume moreover that there exists some $c > 0$ and some $q \geqslant 2$ such that for $n > n_0$*

$$\sup_{x \in \mathbb{R}^d} \|\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))\|_q \leqslant \frac{c}{\sqrt{nh_n^d}}. \tag{12}$$

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

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

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

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

$$\sup_{\|x\| \leqslant n^\nu} |\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))| = O_{a.s.}\left(\sqrt{\frac{\ln n}{n^{1-\gamma}h_n^d}} \right)$$

as soon as γ, β *and* q *are such that*

$$\gamma q + (1-\gamma)\beta d(d-2) > d + 2. \tag{14}$$

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

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* $k \geqslant 2$, *with* $h_n \to 0$ *and* $nh_n^d \to \infty$. *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* $h_n \to 0$ *and* $nh_n^d \to \infty$. *Thus* (12) *obtains for every even* $q \geqslant 2$ *and* $b > \max(2 + 2/d, 3(q-1))$.

□

Proof of Lemma 6. Let $\breve{f}_n$ be the kernel estimator analogue to $\widehat{f_n}$ but built on independent copies of X_0. A well known result (see for instance [25]) is

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

We have

$$\mathrm{var}\widehat{f_n}(x) = \mathrm{var}\breve{f}_n(x) + \frac{2}{n^2h^{2d}} \sum_{1 \leqslant i \leqslant j \leqslant n} \mathrm{cov}(U_i(x), U_j(x))|.$$

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

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

• On the one hand,

$$|\mathrm{cov}(U_0(x), U_m(x))| = h^{2d} \iint (K(s) - \mathbb{E}(K(X_0)))\,(K(t) - \mathbb{E}(K(X_0)))\, f_{0,m}(x - hs, x - ht)dsdt \leqslant Ch^{2d},$$

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)

$$|\mathrm{cov}(U_0(x), U_m(x))| \leqslant 2\|K\|_\infty \frac{LipK}{h} \cdot \theta(m).$$

Therefore,

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

for all $0 \leqslant \alpha \leqslant 1$ by relation (12.18) of [16] so that setting $\alpha = d/(2d+1)$ Δ_n tends to zero as soon as

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

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

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

$$|\mathrm{cov}(U_0(x), U_m(x))| \leqslant \frac{(LipK)^2}{h^2}\alpha(m),$$

so

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

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

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

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

□

Proof of Lemma 8.

- *Proof of (i).* Let $\lambda(n) \to \infty$ 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

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

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

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

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

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

with $C_5 = (3C_1C_4)^{-1}$. Set

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

We have

$$\sup_{\|x\|\leqslant M_n^d} |\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))| = O_{a.s.}\left(\frac{\lambda(n)}{\sqrt{nh^d}}\right). \tag{15}$$

This holds from Borel-Cantelli lemma as soon as

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

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

- *Proof of (ii).* With the conditions on h_n we have $h_n \to 0$ and $nh_n^d \to \infty$. Set $\lambda(n) = \sqrt{n^\gamma \ln n}$. Since $0 < \gamma < 1$ and $0 < \beta < 1/d$, we have $\lambda(n) \to \infty$ and $\lambda(n)/\sqrt{nh^d} \to 0$. 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 $(u_n)_n$ is of the Bertrand type and converges if and only if $c > 1$ or $c = 1$ and $a > 0$. One thus need that there exists some $\nu > 0$ such that $2\nu \leqslant \varepsilon$ with $0 < \varepsilon = \gamma q - 2 - d + \beta d(1-\gamma)(d-2)$ which holds as soon as $\gamma q - 2 - d + \beta d(1-\gamma)(d-2) > 0$.

□

Proof of Lemma 9.

- *Proof of (A2).* In the strong mixing case we apply [26]'s theorem 6.3 to the sequence $U(x) = (U_i(x))_{1\leqslant i\leqslant n}$, for a given $x \in \mathbb{R}^d$. Notice that strong mixing still holds with $\alpha_{U(x)}(m) \leqslant \alpha(m)$ from the heredity properties of strong-mixing. Set

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

For all $q > 2$, $U_i(x)$ admits finite moments with order q since $\|U_i(x)\| \leqslant 2\|K\|_\infty$. Thus, applying Theorem 6.3 yields for all $n > n_0$,

$$\|S_n(x)\|_q^q \leqslant a s_n(x)^q + nc \int_0^1 \left(\alpha^{-1}(u)\right)^{q-1} Q_x^q(u)du,$$

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 $t \mapsto \mathbb{P}(|U_0(x)| > t)$ and

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

Using [26]'s inequality C.3 on page 157 for some $r > q$ there exists $c' > 0$ that only depends on q, r and $\|K\|_\infty$ such that

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

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

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

If $q(1-(1-\gamma)\beta d) \geqslant 2$, 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 $\lim_{r\to\infty} u(r) = q-1$, 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

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

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 [18]'s theorem 1: under the conditions (M2), (M3), (K1) with $h_n \to 0$ and $nh_n^d \to \infty$ (12) obtains for any even $q \geqslant 2$ as soon as

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

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

□

6.3. Preliminary results on the R-convex hull

In this section, we give some technical tools on the R-convex hull and the rolling ball condition. That will allow us to derive asymptotic properties of the R-convex hull estimator for the support S of the density of $\mathbb{X}_n$, as well as that of its related quantities. In this preliminary section we mostly only present rewordings of results obtained in [32], [27] and [2]. Thus proofs might be either left for the reader -as for Properties 1 and 2-, or only sketched -as for Corollary 1. Rewordings are made to have a self content paper, also we aim at presenting deterministic results that emphasize the difference between dependent and independent settings.
For a set $S \in \mathbb{R}^d$, recall that its R-convex hull $C_R(S)$ is defined as:

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

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*

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

so that

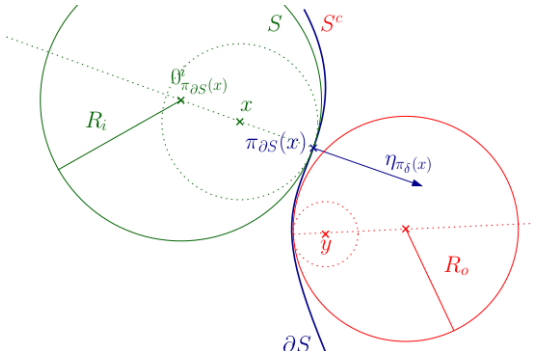

Fig. 1: All you have to have in mind concerning the rolling ball property

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

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

Sets satisfying the inside and outside rolling ball properties have nice properties listed in [33] which we recall in the Property 2 and are 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 $\eta_z = \frac{O^o_z - z}{R_o}$.*
3. *For all $x \in S$ with $d(x, \partial S) \leqslant R_i$, its projection $\pi_{\partial S}(x)$ onto ∂S is well defined and $x = \pi_{\partial S}(x) - d(x, \partial S)\eta_{\pi_{\partial S}(x)}$, $O^i_x = \pi_{\partial S}(x) - R_i \eta_{\pi_{\partial S}(x)}$.*
4. *Conversely, for all $y \in S^c$ with $d(y, \partial S) \leqslant R_o$, its projection $\pi_{\partial S}(y)$ onto ∂S is well defined and $y = \pi_{\partial S}(y) + d(y, \partial S)\eta_{\pi_{\partial S}(y)}$, $O^o_y = \pi_{\partial S}(y) + R_o \eta_{\pi_{\partial S}(y)}$.*

Following the arguments of [27], 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 9. *Let $S' \subset S \subset \mathbb{R}^d$. Assume that balls of radius R_o (resp. R_i) roll freelly outside (resp. inside) S. Let $R < R_o$ and $\varepsilon < \min(R_i, R)$. If, for all $O \in \mathbb{R}^d$ such that $d(O, S) \leqslant R - \varepsilon$ we have that $B(O, R) \cap S' \neq \emptyset$, then*

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

Proof of Proposition 9. 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 $R \leqslant R_o$ 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 $d(x, \partial S) = \varepsilon' \geqslant \varepsilon$ 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 $z = [0, x] \cap \mathcal{S}(x, \varepsilon')$ we have $z \in S$ then, because $\|O - x\| \leqslant R$ we have $d(O, S) \leqslant R - \varepsilon'$ that contradicts $B(O, R) \cap S' = \emptyset$. □

From Proposition 9 and by using theorem 4.1 in [2] we 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 $d(O, S) \leqslant R - \varepsilon$ we have that $B(O, R) \cap S' \neq \emptyset$, then we have*

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

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

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

and there exists some constant $C_{d,R,R_0,R_i} > 0$ *depending on* d, R, R_0 *and* R_i *such that*

$$||\partial C_R(S')|_{d-1} - |\partial S|_{d-1}| \lesssim C_{d,R,R_0,R_i}\varepsilon(1+o(1)), \qquad \text{when } \varepsilon \to 0.$$

Proof of Corollary 1. Here we only sketch the proof of Corollary 1 by giving the main arguments. By Proposition 9 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 $d(x, \partial S) \leqslant \varepsilon$.
   - 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 $d(x, \partial C_R(S')) \leqslant \varepsilon$.
2. About $d(C_R(S'), S)$, we clearly have by inclusions that $d(C_R(S'), S) \leqslant d(S \ominus \varepsilon B, S) = \max_{x\in S} d(x, S \ominus \varepsilon B) = \max_{x\in S\setminus(S\ominus\varepsilon B)} d(x, S \ominus \varepsilon B)$. 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 $S \ominus \varepsilon B$, $\|x - x'\| \leqslant \varepsilon$ that proves $d(C_R(S'), S) \leqslant \varepsilon$.
3. $|C_R(S')|_d \leqslant |S \setminus (S \ominus \varepsilon B)|_d \approx \varepsilon|\partial S|_{d-1}$, indeed it is the inner Minkowski content of a set with positive reach, its asymptotic behaviour can be deduced from the Minkowski content of a set with positive reach, see [20], and for 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 left to the reader to prove that, since ε is small enough $\varphi : C_R(S') \to S$ defined 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 convergence rates for the R-convex hull of a set of points $\mathbb{X}_n$ consists in finding sequences ε_n such that

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

Also we must prove that $d(\mathbb{X}_n, S) \to 0$ almost surely to apply the last part of Corollary 1 and deduce the topological guarantees and the convergence for the measure of the boundary.
The dependence is taken into account when we deal with $\mathbb{P}(B(O, R) \cap \mathbb{X}_n = \emptyset)$ where Propositions 1 or 2 are applied; this replaces a trivial product in the independent case.

Lemma 10. *Consider assumptions of Propositions 1 or 2. Also assume 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*

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

Depending on the parameters of the model f_0, α, d, R_i and on the parameter R such that, we prove the existence of a constant C such that

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

We also get

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

Proof of lemma 10. Let us cover $S \oplus (R - \varepsilon_n)B$ with $\nu_n \leqslant C_d(\frac{diam(S)+R_o}{\varepsilon_n} \ln n)^d$ **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:

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

where $p_i = \mathbb{P}(X_1 \in B(x_i, R - \frac{\varepsilon_n}{\ln n}))$. By Proposition 11 it comes that

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

Thus as ε_n converges to 0 as $n \to \infty$, then for any $C < f_0 C''_{d,\alpha} \left(\frac{\min(R, R_i)}{2} \right)^{\frac{d-1}{2}}$ if n is 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

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

which concludes the proof of equation (17).
Suppose now that $d(\mathbb{X}_n, S) \geqslant r_n$ with $r_n \leqslant R_i$, then there exists $x \in S$ with $B(x, r_n) \cap \mathbb{X}_n = \emptyset$. Define Ω as :

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

Then $B(\Omega, r_n/3) \subset S$ and $B(\Omega, r_n/3) \subset B(x, r_n)$.
As a first conclusion: if $d(\mathbb{X}_n, S) \geqslant r_n$, 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 consists in bounding the probability of the 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 $y_1, \ldots, y_N$ in $S \ominus \frac{2r_n}{3} B$ with $N \leqslant C_d \frac{(3\text{diam}(S) \ln n)^d}{r_n^d}$.
Introduce:

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

By the triangular inequality

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

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 11 we have $q_i \geqslant f_0 C_{d,\alpha}\left(\frac{r_n}{6}\left(1-\frac{1}{\ln n}\right)\right)^{d+\alpha}$ and thus, setting $C' = f_0 C_{d,\alpha}/3^{d+\alpha}$

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

□

6.4. Proof for theorems on support estimation

6.4.1. Proof under independence

First, let us focus on the i.i.d. (or negatively associated) case, which entails, for all $q \in \mathbb{N}^*$ and $R \in \mathbb{R}$, that $\mathcal{E}'_R(q) = 0$.

Proof of Theorem 1. Apply Lemma 10 and equation (17) with $q = 1$ and $\varepsilon_n = \left(\frac{3}{C}\frac{\ln n}{n}\right)^{\frac{2}{d+1+2\alpha}}$, then $P_n \lesssim \ln^d n \cdot n^{-1-2\frac{1+2\alpha}{d+1+2\alpha}+o(1)}$ 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 $\mathbb{P}(d(\mathbb{X}_n, S) \geqslant r_n) \lesssim \ln^d n \cdot n^{-2+o(1)}$.
From the Borel-Cantelli lemma and the Corollary 1, we end the proof of Theorem 1. □

When independence is not assumed, the proofs are similar. We only need to take into account the $\mathcal{E}'$ term and to find suitable sequences ε_n and q_n in order to get convergent series. Then we apply the Borel-Cantelli lemma and the Corollary 1.

6.4.2. Proof under exponential decays.

First we deal with exponential decays i.e. $\alpha(q)$ or $\theta(q)$ is bounded above by Bb^q with $0 < b < 1$
Proof of Theorem 2.

1. Strong mixing. For all $R \in \mathbb{R}$, we have $\mathcal{E}'_R(q) \leqslant Bb^q$.
   (a) Choose $\varepsilon_n = (\gamma(\ln^2 n/n)^{\frac{2}{d+1+2\alpha}}$ 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 $\gamma = \gamma_0^{ed}(1+u)$.
   Set $q_n = \beta_1 \ln n, \beta_1 = \frac{2(2d+1+2\alpha)}{(d+1+2\alpha)\ln(b^{-1})}\left(1+\frac{u}{2}\right)$, we finally introduce $v = \frac{u}{2+u}$. By equation (17)

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

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

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

   thus: $\sum_n P_n < +\infty$.
   (b) Choose $r_n = \left(c\frac{\ln^2 n}{n}\right)^{\frac{1}{d+\alpha}}$ and $q_n = \beta\ln n$ by equation (18):

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

   choose first β large enough that: $\beta\ln b + 1 + \frac{d}{d+\alpha} \leqslant -2$ and second c large enough to have $-\frac{C'c}{\beta} + \frac{d}{d+\alpha} \leqslant -2$.
   We then obtain $\sum_n \mathbb{P}(d(\mathbb{X}_n, S) \geqslant r_n) < +\infty$ (and $r_n \to 0$).
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 $\varepsilon_n \to 0$, for n large enough, $t_n = \sqrt{b^{q_n}}$ is in $[0, R-\varepsilon_n/\ln n]$ thus $\mathcal{E}'_{R-\varepsilon_n/\ln n}(q_n) \leqslant B'\exp(-q_n\ln(b^{-1})/2))$ thus as previously any $\varepsilon_n = (\gamma\ln^2 n/n)^{\frac{2}{d+1+2\alpha}}$ with $\gamma > 2\gamma_0^{ed}$ provides a sequence with $\sum_n P_n < +\infty$.

(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 $\mathcal{E}'_{r_n - r_n/\ln n}(q_n) \leqslant B' \exp(-q_n \ln(b^{-1})/2))$ and similar to 1.(*b*) we obtain $\sum_n \mathbb{P}(d(\mathbb{X}_n, S) \geqslant r_n) < +\infty$ and $r_n \to 0$ for a suitable constant 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 $\sum_n P_n < +\infty$, and sequences $r_n \to 0$ with $\sum_n \mathbb{P}(d(\mathbb{X}_n, S) \geqslant r_n) < +\infty$, thus the Borel-Cantelli lemma and the Corollary 1 imply the result.

Now we finish this section by the proof of Theorem 3 which dedicates to the Riemannian decays cases.

6.4.3. Proof under Riemannian decays.

Proof of Theorem 3. According to the dependence assumption we check that

1. *Under strong mixing:* $\mathcal{E}'_R(q_n) \leqslant Bq_n^{-b}$ 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 $\varepsilon_n = \left(A\frac{(\ln n)^a}{n^{1-\gamma}}\right)^{\frac{2}{d+1+2\alpha}}$ and $q_n = \frac{n^\gamma}{(\ln n)^{a-1}}$. 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 $\sum_n P_n < +\infty$.

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

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

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

2. *Under θ-weak dependence.* 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 $\varepsilon_n \to 0$, for n large enough, $t_n = \sqrt{q_n^{-b}}$ is in $[0, R - \varepsilon_n/\ln n]$ thus $\mathcal{E}'_{R-\varepsilon_n/\ln n}(q_n) \leqslant B' q_n^{-b/2}$ and we can use the sequences exhibited in the strong mixing for changing b into $b/2$.
Similarly, set $\lambda' = \dfrac{2d+\alpha}{d+\alpha}$, $a' = \dfrac{b/2-2-d}{b-\lambda'}$ and $\gamma' = \dfrac{\lambda'+1}{\lambda'+b/2}$, choose $r_n = \left(A'\dfrac{(\ln n)^{a'}}{n^{1-\gamma'}}\right)^{\frac{1}{d+\alpha}}$, $q_n = \dfrac{n^{\gamma'}}{(\ln n)^{a-1}}$, and $t_n = \sqrt{q_n^{-b}}$. Check that $t_n < \varepsilon_n$ From Equation (17) it comes that

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

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

6.5. Proofs concerning level sets

Proof of Theorem 4. Let $R < m_t/f^{(2)}_{\max}$ be a positive constant, we first prove that, for n large enough:

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

First a simple chain of inclusions entails $\widehat{L}_{t,R} = C_R(\{\widehat{f_n}(X_i) \geqslant t\}) \subset C_R(\{X_i, f(X_i) \geqslant t - \varepsilon'_n\}) \subset C_R(L_{t-\varepsilon'_n})$. In the sequel, we consider n large enough to have $\varepsilon'_n \leqslant \max(\Delta_t, t - t_0)$ so that Proposition 6 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

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

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

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

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

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 us 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 $d(x, \partial L_{t+\varepsilon'_n}) \geqslant \varepsilon_n$ 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 $d(x, \partial L_{t+\varepsilon'_n}) \geqslant \varepsilon_n$, and we have that $d(O, L_{t+\varepsilon'_n}) \leqslant R - \varepsilon_n$.
We set now

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

Cover $L_{t+\varepsilon'_n} \oplus (r-\varepsilon_n)B$ with ν_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 $|L_{t+\varepsilon'_n}| \leqslant t^{-1}$ (because the density on $L_{t+\varepsilon'_n}$ is greater than t). Thus by bound on packing and covering numbers it is possible with $\nu_n \leqslant \frac{2^d}{t\omega_d}\left(\frac{\ln n}{\varepsilon_n}\right)^d$.

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

Now by proposition 6, $L_{t+\varepsilon'_n}$ admits both the $(m_t/f_{\max}^{(2)})$-inside and outside rolling ball properties. Thus if we define $x_i^* = \pi_{\partial L_{t+\varepsilon'_n}}(x_i)$ and $y_i = x_i^* + \frac{m_t}{f_{\max}^{(2)}} \frac{x_i^* - x_i}{\|x_i^* - x_i\|}$, then we have $B(y_i, m_t/f_{\max}^{(2)}) \subset L_{t+\varepsilon'_n}$ and $\|x_i - y_i\| \leqslant R - \varepsilon_n + m_t/f_{\max}^{(2)}$. Now with $E_i = B(x_i, R-\varepsilon_n/\ln n) \cap B(y_i, m_t/f_{\max}^{(2)})$, we obtain

$$P_n \leqslant \sum_{i=1}^{\nu_n} \left((1-p_i)^{\frac{n}{q}} + \frac{\mathcal{E}_{E_i}(q)}{p_i}\right), \qquad \text{with } p_i = \mathbb{P}(X_1 \in E_i).$$

By proposition 11, we derive, with the choices $f_0 = t$, $\alpha = 0$:

$$p_i \geqslant tC''_{d,0} \left(\frac{\min\left((R - \frac{\varepsilon_n}{\ln n}), \frac{m_t}{f_{\max}^{(2)}}\right)}{2}\right)^{\frac{d-1}{2}} \left\{\min\left(\left(\varepsilon_n - \frac{\varepsilon_n}{\ln n}\right), \frac{\min\left((R - \frac{\varepsilon_n}{\ln n}), \frac{m_t}{f_{\max}^{(2)}}\right)}{2}\right)\right\}^{\frac{d+1}{2}}.$$

Thus as ε_n converges to 0 as $n \to \infty$, then for any $C < tC''_{d,0}\left(\frac{\min\left(R, \frac{m_t}{f_{\max}^{(2)}}\right)}{2}\right)^{\frac{d-1}{2}}$ if n is 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 consists in $\mathcal{E}_{E_i}$ which is now related to intersections of balls instead of only one ball.
Under strong mixing $\mathcal{E}_E$ does 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, then the value of $\mathcal{E}_E$ depends on the shape of the set E, nevertheless, due to equations (3) and (4), we obtain

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

for some positive constants A and B.

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

This ends the proof of the inclusions

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

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 in order that $\varepsilon'_n < \frac{m_t^2}{2f^{(2)}_{\max}}$. Indeed, consider $z \in L_t \ominus 2\frac{\varepsilon'_n}{m_t} B$, we have that for each vector u with $\|u\| \leqslant 1$, $f\left(z + 2\frac{\varepsilon'_n}{m_t}u\right) \geqslant t$. Consider now the special choice $u = -\frac{\nabla_f(z)}{\|\nabla_f(z)\|}$, then we have

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

Thus

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

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 from [33] and [20], since L_t admits the inside and outside $(m_t/f^2_{\max})-$rolling ball properties, we obtain $\varepsilon_n + 2\frac{\varepsilon'_n}{m_t} \leqslant m_t/f^2_{\max}$ and then,

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

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 the homeomorphism property and the convergence for the perimeter estimation we also prove that

$$d_H(L_t, \mathbb{X}_n \cap \{z, \widehat{f}(z) \geqslant t\}) \xrightarrow{a.s.} 0,$$

this is easily obtained by using the techniques extensively used in this paper and we leave it to the reader.

Aknowledgements. The authors acknowledge support of the Institut Henri Poincaré (UAR 839 CNRS-Sorbonne Université), and LabEx CARMIN (ANR-10-LABX-59-01) this work began during the trimester "Geometry and Statistics in Data Science". This work has been supported by the ANR GeoDSic.
This work was also funded by the chair FIME: `https://fime-lab.org/` and CY-AS ("Investissements d'Avenir" ANR-16-IDEX-0008), "EcoDep" PSI-AAP2020-00000000013.

References

[1] C. Aaron and O. Bodart. Local convex hull support and boundary estimation. Journal of Multivariate Analysis, 147:82–101, 2016.
[2] C. Aaron, A. Cholaquidis, and R. Fraiman. Estimation of surface area. Electronic Journal of Statistics, 16(2):3751–3788, 2022.
[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. Rodr'iguez-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. Cholaquidis, R. Fraiman, and L. Moreno. Universally consistent estimation of the reach. Journal of Statistical Planning and Inference, 2021.
[10] M. Coblenz, R. Dyckerhoff, and Oliver Grothe. Nonparametric estimation of multivariate quantiles. Environmetrics, 29, 2018.
[11] A. Cuevas and A. Rodríguez-Casal. On boundary estimation. Advances in Applied Probability, 36(2):340–354, 2004.
[12] J. Dedecker, P. Doukhan, G. Lang, J. R. Leon, S. Louhichi, and C. Prieur. Weak dependence: With examples and applications. In Lecture Notes in Statistics, volume 190. Springer, 2007.
[13] 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.
[14] E. Di Bernardino, T. Laloë, V. Maume-Deschamps, and C. Prieur. Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. Esaim: Probability and Statistics, 17:236–256, 2013.
[15] P. Doukhan. Mixing, properties and models, volume 85 of Lecture Notes in Statistics. Springer-Verlag, New York, 1994.
[16] P. Doukhan. Stochastic models for time series, volume 80. Springer, 2018.
[17] P. Doukhan and S. Louhichi. A new weak dependence condition and applications to moment inequalities. Stochastic Processes and their Applications, 84:313–342, 1999.
[18] P. Doukhan and S. Louhichi. Functional estimation of a density under a new weak dependence condition. Scandinavian Journal of Statistics, 28:325–341, 2001.
[19] 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.
[20] H. Federer. Curvature measures. Transactions of the American Mathematical Society, 93(3):418–491, 1959.
[21] 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.
[22] W. Hardle, B.U. Park, and A.B. Tsybakov. Estimation of non-sharp support boundaries. Journal of Multivariate Analysis, 55(2):205–218, 1995.
[23] S. Kallel and S. Louhichi. Topological reconstruction of compact supports of dependent stationary random variables. arXiv:2307.11674v1 [math.PR], 2023.
[24] E. Liebscher. Strong convergence of sums of α-mixing random variables with applications to density estimation. Stochast. Process. Appl., 65:69–80, 1996.
[25] B. L. S. Prakasa Rao. Nonparametric functional estimation. Academic press, New York, 1983.
[26] E. Rio. Théorie asymptotique des processus aléatoires faiblement dépendants. In Mathématiques et Applications, volume 31, Berlin, 2000. Springer.
[27] A. Rodríguez Casal. Set estimation under convexity type assumptions. Annales De l'Institut Henri Poincaré: Probabilités et Statistiques, 43:763–774, 2007.
[28] 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.
[29] A. Rodríguez-Casal and P. Saavedra-Nieves. Spatial distribution of invasive species: an extent of occurrence approach. TEST, 31:416–441, 2022.
[30] M. Rosenblatt. A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. USA, 42:43–47, 1956.
[31] G. G. Roussas. Asymptotical normality of positively or negatively associated processes. Journal of Mult. Analysis, 50:152–173, 1994.
[32] G. Walther. Granulometric smoothing. The Annals of Statistics, 25(6):2273 – 2299, 1997.
[33] 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 10. *let $x, y \in \mathbb{R}^d$, and r_x, r_y, ℓ be positive numbers such that $\|x-y\| = r_x + r_y - \ell$, $\ell \leqslant \min(r_x, r_y)$ and $h \leqslant \min(r_x, r_y)$. 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 $\ell \to 0$*

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

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

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

Thus by derivation

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

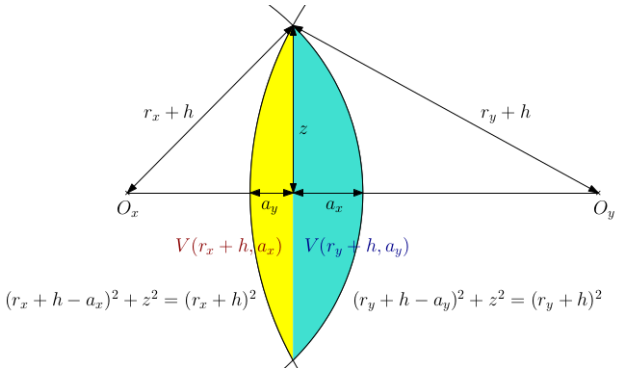

Fig. A.2: The convex body is $B(O_x, r_x + h) \cap B(O_y, r_y + h)$.

Solving

$$\begin{cases} (r_x - h - a_x)^2 + z^2 &= (r_x - h)^2 \\ (r_y - h - a_y)^2 + z^2 &= (r_y - h)^2 \\ a_x + a_y &= \ell + 2h \end{cases},$$

yields

$$\begin{cases} a_x &= \dfrac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)} \\ a_y &= \dfrac{(\ell + 2h)(2r_x - \ell)}{2(r_x + r_y - 2h - 2\ell)} \end{cases}$$

and this complements the Figure A.2 which proves that indeed:

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

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

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

as $\ell \to 0$, with $h \leqslant \ell/2$, let $a = \dfrac{\ell r_y}{r_x + r_y}$, $\Delta^+ = \dfrac{(\ell + 2h)(2r_y - \ell)}{2(r_x + r_y - 2h - 2\ell)} - a$ and $\Delta^- = a - \dfrac{(\ell - 2h)(2r_y - \ell)}{2(r_x + r_y + 2h - 2\ell)} - a$ (thus

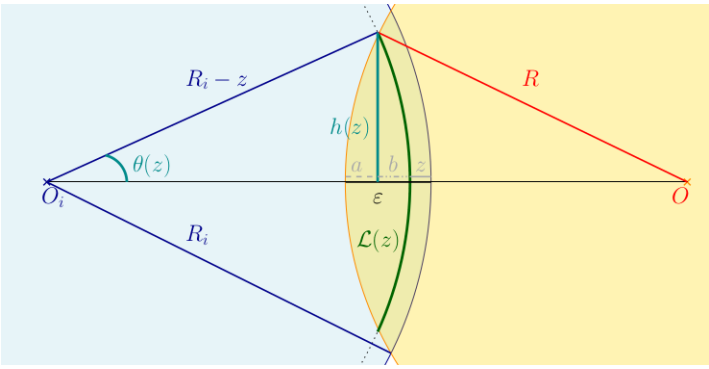

Fig. A.3: Integration

$\Delta^+ + \Delta^- \leqslant \dfrac{4hr_y}{r_x + r_y}(1 + o(1)))$ we have

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

We finally obtain that

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

which concludes the proof. □

Proposition 11. *Suppose that balls of radius R_i roll inside S. Suppose that the density is such that $f(x) \geqslant f_0 d(x, \partial S)^\alpha$. If O is such that $d(O, S) = R - \varepsilon$ with $0 < \varepsilon \leqslant R$ then, exists positive constants $C_{d,\alpha}$ and $C'_{d,\alpha}$ such that*

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

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

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

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

see Figure A.3, it comes that

$$
\mathbb{P}(B(O,R)) \geqslant \int_0^\varepsilon f_0\, z^\alpha \left|\mathcal{L}(z)\right|_{d-1} dz.
$$

Then $|\mathcal{L}(z)|_{d-1} = \int_0^{\theta(z)} \sigma_{d-2}(R_i - z)^{d-1} \sin(\alpha)^{d-2} d\alpha$. From inequality $\sin(\alpha) \geqslant 2\alpha/\pi$ if $0 \leqslant \alpha \leqslant \pi/2$ (condition $\varepsilon < \min(R, R_i)$ allows that) we obtain that $|\mathcal{L}(z)|_{d-1} \geqslant \frac{\sigma_{d-2}}{d-1} \left(\frac{2\theta(z)}{\pi}\right)^{d-1} (R_i - z)^{d-1}$. Now

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

Thus

$$\mathbb{P}(B(O,R)) \geqslant f_0 \cdot \frac{\sigma_{d-2}}{d-1} \left(\frac{2}{\pi}\right)^{d-1} \int_0^{\varepsilon} h(z)^{d-1} dz,$$

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

Under the condition $\varepsilon \leqslant \min(R_i, R)/4 \leqslant (R + R_i)/8$, we obtain that

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

which finally yieds:

$$\begin{aligned}
\mathbb{P}(B(O,R)) &\geqslant f_0 \frac{\sigma_{d-2}}{d-1} \left(\frac{2}{\pi}\right)^{d-1} \left(\frac{1}{2} \frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \int_0^{\varepsilon} z^{\alpha} (\varepsilon - z)^{\frac{d-1}{2}} dz \\
&\geqslant f_0 \frac{\sigma_{d-2}}{d-1} \left(\frac{2}{\pi}\right)^{d-1} \left(\frac{1}{2} \frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \varepsilon^{\frac{d+1+2\alpha}{2}} B\left(\alpha + 1, \frac{d+1}{2}\right) \\
&\geqslant f_0 C_{d,\alpha} \left(\frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \varepsilon^{\frac{d+1+2\alpha}{2}}.
\end{aligned}$$

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

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

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

$$\mathbb{P}(B(O,R)) \geqslant f_0 C_{d,\alpha} \left(\frac{RR_i}{R + R_i}\right)^{\frac{d-1}{2}} \left(\frac{\min(R, R_i)}{4}\right)^{\frac{d+1+2\alpha}{2}} \geqslant f_0 \frac{C_{d,\alpha}}{2^{\frac{3d+1+4\alpha}{2}}} (\min(R, R_i))^{d+\alpha} .$$

- Final case $O \in S$.

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

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