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 with R-convex hulls and its natural plug-in to estimate the support S or level-sets together with their volume and perimeter of the marginal density f on $\mathbb{R}^d$, of a strictly stationary process.

For this we assume some weak dependence conditions and different distributional assumption depending on the estimation purpose (support or level sets estimation).

The loss with regard to the classical independent case depends on the current dependence structure.

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 [25], 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 interpreted as a multidimensional $\alpha-$quantile, see [12, 17].

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 [13, 16] (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 [3, 7, 35–37, 40] (other estimators of S have also been proposed, as in [1] or [27]).

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 admits the following heuristic justification.

If S is regular enough, we have $C_R(S) = S$ so that a natural estimator of S is $\widehat{S}_n = C_R(\mathbb{X}_n)$, where $\mathbb{X}_n = \{X_1, \ldots, X_n\} \subset$

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

$\mathbb{R}^d$ *denotes a n-sample with marginal density* f *having a 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 [40] and [35]. They derive convergence rates of the Hausdorff distance $d_H(\widehat{S}_n, S)$ towards 0 and obtain the minimax rate of [27] 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 [7, 8], the perimeter [3] and level sets [36] 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.. Hereafter we extend the study to densities which are not bounded below and to dependent settings. Indeed for real life applications, independence hypothesis is generally unrealistic. This is the case for instance if the location of a member of an animal specie is observed *via GPS* or if the sample consists of the trajectory of one or more members observed over a given 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 both strong mixing and θ-weak dependence conditions, introduced respectively in [38] and [15].

Set estimation under dependence has been little studied. A few related works are those of [10], [11] and [9] who consider continuous time trajectories of a reflected Brownian motion. In the recent paper of [29] 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 [3, 7, 8, 35, 36, 40] 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) of f.

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 estimators of the density support and of related quantities, such as volume and perimeter. Section 4 deals with the estimation of level-sets of f. As a conclusion, Section 6 proposes some perspectives for data-driven tuning of the geometric parameter R. The proofs are postponed to Section 5.

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 a 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 respectively denote by $|E|_d$, ∂E and $|\partial E|_{d-1}$ its d-dimensional volume, its boundary and the $(d-1)$-dimensional volume of ∂E.

  For the sake of simplicity, we sometimes refer to volume and perimeter when talking about $|E|_d$ and $|\partial E|_{d-1}$.

- $B(x,r) \subset \mathbb{R}^d$ denotes the closed ball of radius $r \geq 0$ centered at x and $\mathcal{S}(x,r)$ is the surface area of 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 $\sigma_d = |\mathcal{S}(0,1)|_{d-1}$ the surface of its boundary.

- 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 each compact set $E \subset \mathbb{R}^d$ and each $x \in \mathbb{R}^d$ we set $\underline{\mathrm{d}}\,(x,E) = \min_{y\in E} \|x-y\|^2$, where $\|\cdot\|$ denotes a norm on $\mathbb{R}^d$.

- For two subsets E and F of $\mathbb{R}^d$, we denote by $d_H(E,F)$ their Hausdorff distance:
$$d_H(E,F) = \max\left(\sup_{a\in E} \underline{\mathrm{d}}\,(a,F), \sup_{b\in F} \underline{\mathrm{d}}\,(b,E)\right).$$
- We set $E \approx F$ if and only if E and F are homeomorphic, i.e. there exists a bi-continuous bijection $\psi : E \to F$.
- 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 $n \geqslant 0$, $u_n \leqslant cv_n$. 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 $f : \mathbb{R}^d \to \mathbb{R}$. For $f \in C^1$ we denote by $f^{(i)}$ its gradient and in case $f \in C^2$ we denote by $f^{(2)}$ its Hessian matrix.
- We denote by $\|\cdot\|_{op}$ the operator norm on the set of $n \times p$ matrices.
- Let $\rho \geqslant 2$. 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 with degree (strictly) less than ρ.

2.2. 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 all $-\infty \leqslant j \leqslant l \leqslant +\infty$, we set $\mathcal{F}_j^l = \sigma(X_k, j \leqslant k \leqslant l,\ k \in \mathbb{Z})$ the σ-field of events generated by the random variables $(X_k)_{j\leqslant k\leqslant l}$. We denote by $\mathbb{X}_n = \{X_1, \ldots, X_n\} \subset \mathbb{R}^d$ a n-sample of $\mathbf{X}$ with marginal density $f : \mathbb{R}^d \to \mathbb{R}$.

In this paper, we consider two weak dependence conditions for $\mathbf{X}$. These conditions cover a broad class of classical models of times series as in Remark 1.

Strong or α-mixing. Let $(\alpha(q))_{q>0}$ be the decreasing sequence of strong mixing coefficients of $\mathbf{X}$, defined as
$$\alpha(q) = \sup_{A\in\mathcal{F}_{-\infty}^0,\ B\in\mathcal{F}_q^\infty} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$
Note that the above expression can also be written as the supremum of $|\operatorname{cov}(\mathbf{1}_A, \mathbf{1}_B)|$. We say that $\mathbf{X}$ is strongly mixing, *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 $h : \left(\mathbb{R}^d\right)^v \to \mathbb{R}$ by
$$\operatorname{Lip}(h) = \sup_{x\neq y} \frac{|h(x) - h(y)|}{\|x_1 - y_1\| + \cdots + \|x_v - y_v\|}.$$
The process $\mathbf{X}$ is θ-weakly dependent, [14], if there exists a decreasing sequence $(\theta(q))_{q>0}$ with $\lim_{q\to\infty} \theta(q) = 0$ such that
$$\left|\operatorname{cov}(W, h(X_{i_1}, \ldots, X_{i_v}))\right| \leq v \operatorname{Lip}(h) \cdot \theta(q)\ ,$$
for any $v \in \mathbb{N}^*$ any sequence of indices $i_1, \ldots, i_v \geqslant q$ and any random variable W measurable with respect to $\mathcal{F}_{-\infty}^0$ and such that $\|W\|_\infty \leq 1$.
Notice that for $v = 1$ this simply writes: $\left|\operatorname{cov}(W, h(X_q))\right| \leq \operatorname{Lip}(h) \cdot \theta(q)$.

Remark 1. A large number of classical models which satisfy strong mixing conditions is provided in [18]; 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 bi-linear, 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 displayed in [5], in which $\mathbf{X}$ is the stationary solution of the recursion
$$X_t = \tfrac{1}{2} \cdot \left(X_{t-1} + \varepsilon_t\right), \qquad \forall t > 0,$$
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 fit the θ-weak dependence condition. Notice also that this property is hereditary through Lipschitz functions.

2.3. A key result under dependence

Let E be a measurable subset of $\mathbb{R}^d$ with $\mathbb{P}(X_t \in E) = p > 0$. If $\mathbb{X}_n$ is an i.i.d. sample, then

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

This is a key property to derive results on set estimation. This relation essentially extends for some dependent cases of interest quoted in the following remark.

Remark 2. (Cases related with independence).

- Consider a negatively associated process, see [39]. In the proofs of Subsection 5.1, we let $q = 1$, and for $Z_i = \mathbf{1}_{\{X_i \in E\}}$ $(\forall i \geq 0)$, $\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, Z_k) \leq 0$, hence (1) holds.

The lemmas 7.1 and 7.2 in [30] entail that the same result (1) holds for two other cases of interest

- $\mathbf{X}$ is a $m-$dependent process, $\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \leq (1-p)^{\frac{n}{m}}$,
- $\mathbf{X}$ is a geometrically ergodic Markov process, $\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \lesssim (1-p)^n$.

In the dependence settings of Remark 2, we obtain the same rates as for the independent case when dealing with set estimation.

We relax the independence assumptions to settings such that a weaker version of Equation (1) holds. Namely, we will give some explicit values of $\mathcal{E}_E(q)$ such that for each $q \in \mathbb{N}$ with $1 \leq q \leq n$ we have

$$\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \leq (1-p)^{\frac{n}{q}} + \frac{\mathcal{E}_E(q)}{p}. \tag{2}$$

Property (2) weakens (1) for several dependence structures and for some cases of sets E of interest. This allow to derive convergence rates for the estimation of S, ∂S, $|S|_d$ and $|\partial S|_{d-1}$ in Section 3 and for level sets in Section 4.

Proposition 1. *Let $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ be a stationary strongly mixing sequence with mixing coefficients $(\alpha(q))_{q>0}$. Let $E \subset \mathbb{R}^d$ be an arbitrary measurable set such that $\mathbb{P}(X_t \in E) = p > 0$. Then* (2) *holds for each $1 \leq q \leq n$ with $\mathcal{E}_E(q) = \alpha(q)$.*

Proposition 2. *Let $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ be a stationary $\theta-$weakly dependent sequence with dependence coefficients $(\theta(q))_{q>0}$, then we restrict to two special cases of measurable sets $E \subset \mathbb{R}^d$ of interest.*

1. *For $E = B(x,r)$, then there exists $c > 0$ such that* (2) *holds with*

$$\mathcal{E}_E(q) \leqslant \inf_{\varepsilon\in(0,r)} \left(\frac{\theta(q)}{\varepsilon} + c\omega_d 2^d \varepsilon r^{d-1} \right). \tag{3}$$

2. *For $E = B(x,r_x) \cap B(y,r_y)$ with $r_x + r_y - \|x-y\| = \ell$, and $\ell \leqslant r_x \wedge r_y$, then there exists $c > 0$ such that* (2) *holds with*

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

In order to get explicit rates, we will propose in the sequel two scenarii:

- *Exponential decays,* $\alpha(q) \leqslant Bb^{-q}$, or $\theta(q) \leqslant Bb^{-q}$, for some $b < 1$.
- *Polynomial decays,* $\alpha(q) \leqslant Bq^{-b}$, or $\theta(q) \leqslant Bq^{-b}$, for some $b > 0$.

Remark 3. In Proposition 2, for both cases, the constant $c > 0$ is defined through:

$$\mathbb{P}(X_t \in E) \leq c\,|E|_d. \tag{5}$$

Notice that this condition holds for instance if the distribution of X_0 admits a bounded density f w.r.t. Lebesgue measure.

Remark 4. Our results may be extended to alternative dependence structures, yielding other values of $\mathcal{E}_E(q)$, possibly involving other methods of proofs.

3. Support estimation

In this section, we derive convergence rates for the R-convex hull estimators of the support S, its boundary, and their measures in the context of strictly stationary strong-mixing or θ-weakly dependent sequences $\mathbf{X}$.

We first give in Subsection in 3.1 regularity conditions on S under which are derived our main theoretical results. We then provide in Subsection 3.2 the main lines of the proofs by enlightening the differences between the independent case and the dependence setting for which an additional term for a probability bound appears.

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 $C_R(S) = S$. More precisely, we need that S satisfies the following

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 and only if:*

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

If 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.

Sets satisfying the rolling ball property have nice properties due to the generalization of Blaschke's rolling ball's theorem, see [40]'s theorem 1. In particular, they satisfy $C_{R_o}(S) = S$, which ensures the convergence of the plug-in support estimator $C_{R_o}(\mathbb{X}_n)$ 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 [27]:

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

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

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 [3, 35] 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

$$d_H(\partial C_R(\mathbb{X}_n), \partial S) \leqslant \varepsilon_n, \qquad d_H(C_R(\mathbb{X}_n), S) \leqslant \varepsilon_n, \tag{6}$$
$$\partial C_R(\mathbb{X}_n) \approx \partial S, \qquad C_R(\mathbb{X}_n) \approx S, \tag{7}$$
$$\left|\left|\partial C_R(\mathbb{X}_n)\right|_{d-1} - |\partial S|_{d-1}\right| \lesssim \varepsilon_n, \quad \left|\left|C_R(\mathbb{X}_n)\right|_d - |S|_d\right| \lesssim \varepsilon_n, \tag{8}$$

with

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

Optimal results for several classes of dependent models leading to the same convergence rates as for independence are obtained from Remark 2 in Proposition 3 below:

Proposition 3. *Let $\mathbf{X} = (X_t)_{t\in\mathbb{Z}}$ be a negatively associated or m−dependent or is a geometrically ergodic Markov chain. Then $\mathbb{P}(\mathbb{X}_n \cap E = \emptyset) \lesssim (1-p)^n$ and the conclusions of Theorem 1 also hold.*

Theorem below extend Theorem 1 to samples $\mathbf{X_n}$ of a strictly stationary sequence which is either strong mixing or $\theta-$weakly dependent with either exponential decay or polynomial decay of the dependence coefficients.

Theorem 2. *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. Assume that $\mathbf{X}$ is strongly mixing or θ-weakly dependent. Choose $R < R_o$.*

- *If the sequence of dependence coefficients of $\mathbf{X}$ has an **exponential decay** rate $\alpha(q) \leqslant Bb^{-q}$, or $\theta(q) \leqslant Bb^{-q}$, for some $b < 1$, then, with probability 1 for large enough n, Equations (6), (7) and (8) are satisfied with the following ε_n*

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

  where the constant $A(d, \alpha, f_0, R, R_i, b)$ is explicit.

- *If the sequence of dependence coefficients of $\mathbf{X}$ has a **polynomial decay** rate $\alpha(q) \leqslant Bq^{-b}$ or $\theta(q) \leqslant Bq^{-b}$ with $b > b_0$ then, with probability 1 for large enough n, Equations (6), (7) and (8) are satisfied with the following ε_n*

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

  Explicit values of b_0, a and γ are given by

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

Theorem 2 firstly 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. Secondly in the polynomial decay 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 [36], 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 in order to achieve the convergence and the convergence rate depends on R_i.

3.3. Main arguments for Theorems 1 and 2.

The proofs are split into two parts.

- First, geometric and deterministic arguments allow to prove that:

$$\forall O \in \mathbb{R}^d, \forall R \in (0, R_o), \exists\, \varepsilon > 0 : \qquad \underline{\mathrm{d}}\,(O, S) \leqslant R - \varepsilon \Longrightarrow B(O, R) \cap \mathbb{X}_n \neq \emptyset.$$

  Then for each $R < R_O$ and for a convenient ε, we prove that if volume and perimeters are close of each other

$$d_H(C_R(\mathbb{X}_n), S) \leqslant \varepsilon, \qquad d_H(\partial C_R(\mathbb{X}_n), \partial S) \leqslant \varepsilon,$$

  and the homeomorphism property holds. Those results are stated in Proposition 6 and Corollary 1, which is a part of theorem 4.1 from [3].
  We defer the reader to the fourth point in the proof of Corollary 1 to explicit an homeomorphism $S \to C_R(\mathbb{X}_n)$.

- As a consequence of Proposition 6 and Corollary 1, if there exists a sequence $(\varepsilon_n)_n$ such that

$$\sum_{n=1}^{\infty} p_n < \infty, \quad \text{with} \quad p_n = \mathbb{P}\Big(\exists\, O,\ \underline{\mathrm{d}}\,(O,S) \leqslant R - \varepsilon_n \text{ and } B(O,R) \cap \mathbb{X}_n = \emptyset\Big), \quad \text{for each} \quad n \geq 1,$$

then the Borel-Cantelli lemma allows to conclude.
Seeking suitable sequence $(\varepsilon_n)_n$ is the probabilistic part of the proof and make use of the key inequality (2). A compactness argument allows to exhibit a coverage of $\{x,\ d(x,S) \leqslant R - \varepsilon_n\}$ by a finite number of balls $O_1, \ldots, O_{\nu_n} \in \{x, d(x,S) \leqslant R - \varepsilon_n\}$ centred at O_i and with radius $\varepsilon_n / \log n$:

$$\Big\{x, d(x,S) \leqslant R - \varepsilon_n\Big\} \subset \bigcup_{i=1}^{\nu_n} B\Big(O_i, \frac{\varepsilon_n}{\ln n}\Big), \quad \text{where a dimension argument implies } \nu_n \lesssim \left(\frac{\ln n}{\varepsilon_n}\right)^d.$$

Now for each n and $i \in [1, \nu_n]$ we set

$$E_{i,n} = B\Big(O_i, R - \frac{\varepsilon_n}{\ln n}\Big).$$

From the triangular inequality, we infer

$$p_n \leqslant \sum_{i=1}^{\nu_n} \mathbb{P}\left(E_{i,n} \cap \mathbb{X}_n = \emptyset\right).$$

  - In the independent case:

$$\mathbb{P}\left(E_{i,n} \cap \mathbb{X}_n = \emptyset\right) = \left(1 - \mathbb{P}\left(E_{i,n}\right)\right)^n$$

    and the problem boils down to bound $\mathbb{P}(E_{i,n})$ uniformly, with respect to $i \in [1, \nu_n]$, by using of the inside rolling ball property combined with $\underline{\mathrm{d}}\,(O_i, S) \leqslant R - \varepsilon_n(1 - (\ln n)^{-1})$ and the property on f assumed in Definition 2.
    We thus can exhibit a sequence $(\varepsilon_n)_n$ such that the series $\displaystyle\sum_{n=1}^{\infty} p_n < \infty$ are summable.

  - Under dependence, the key inequality (2) yields an additional term,

$$\mathbb{P}\left(E_{i,n} \cap \mathbb{X}_n = \emptyset\right) = \left(1 - \mathbb{P}\left(E_{i,n}\right)\right)^{\frac{n}{q_n}} + \frac{\mathcal{E}_{E_{i,n}}(q_n)}{\mathbb{P}\left(E_{i,n}\right)}$$

    for each $q_n \in \mathbb{N}$ such that $1 \leqslant q_n \leqslant n$.
    The probability bounds with the additional term for dependent data are set precisely in Subsection 2.3 and they are proved in Subsection 5.1. In Subsection 5.3 we give detailed proofs of Theorems 1 and 2.

4. Level set estimation

We now focus on level set estimation. For a given $t \in \mathbb{R}^+$, the level set L_t is defined as

$$L_t = \overline{\Big\{z \in \mathbb{R}^d,\ f(z) > t\Big\}}.$$

If $t = 0$ we turn to the case of support estimation.
In the i.i.d. case, [36] introduced and studied the following empirical level set estimator

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

where $\widehat{f_n}$ is a suitable density estimator of f.
In this section, we aim at extending [36]'s convergence rates to our dependence setting. We first recall the classical assumptions when dealing with level set estimation. We prove Theorem 3 which explicits the convergence rates for our level set estimator in the dependent case as the sum of two terms

- a set estimation error ε_n, due to the use of the r-convex hull,
- a density estimation error term ε'_n.

Possible values for ε'_n in different contexts of dependence are given in Subsection 4.3. Finally, a sketch of the proof of Theorem 3 is given in Subsection 4.4.

Remark 5. Suitable levels t defined in the sequel as regular levels see Definition 3 satisfy that $L_t = \left\{z \in \mathbb{R}^d,\ f(z) \geqslant t\right\}$ and ℓ_t is a $(d-1)$-dimensional manifold, see [36].

4.1. Distributional background

In the context of level set estimation, we assume different distributional hypotheses than those required in the support estimation problem.
Namely, the compactness of the support of the distribution is useless since $L_t = \overline{\{z,\ f(z) > t\}}$ is compact for all $t > 0$. 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 $k \geqslant 2$ to derive uniform convergence rates for kernel density estimation. Also, for regular levels t, as defined in Definition 3 below, and by Proposition 4, see theorem 2 in [40], we derive that the sets L_t naturally inherit the inside and outside rolling ball properties.

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

$$L_{t-\Delta_t} \subset \mathring{S}, \quad \text{and} \quad \min_{z \in L_{t-\Delta_t} \setminus \mathring{L}_{t+\Delta_t}} \|f^{(1)}(z)\| = m_t > 0.$$

A corollary of theorem 2 in [40] writes:

Proposition 4 ([40])**.** *Let $f \in C^2$, we assume that $f^{(2)}_{\max} = \sup_{z \in S} \|f^{(2)}(z)\|_{op} < \infty$. If $t \in \mathbb{R}^+$ is a regular level, then there exists δ_t such that, for all $t' \in (t-\delta_t, t+\delta_t)$, $L_{t'}$ satisfies the inside and outside rolling ball properties for balls of radius $m_t / f^{(2)}_{\max}$.*

4.2. Main Theoretical results

Hereafter, we study the asymptotic behavior of $\widehat{L}_{t,R}$, its boundary and its volume under weak dependence conditions. We also obtain topological guarantees and perimeter estimation rates using ideas inspired from [3].
We focus here on the geometric part ε_n. Explicit rates taking into account all the parameters of the model are given in Appendix C and Appendix B.

Theorem 3. *Let $\mathbb{X}_n = \{X_1, \dots, X_n\}$ be a n-sample of a strictly stationary sequence of a strong mixing or a θ-weak dependent sequence $\mathbf{X}$ marginal class C^k density f on its support S.*
Assume that $\sup_{z \in S} \|f^{(1)}(z)\| = f^{(1)}_{\max} < +\infty$, and $\sup_{z \in S} \|f^{(2)}\|_{op} = f^{(2)}_{\max} < +\infty$.
Assume that the density estimator $\widehat{f_n}$ satisfies $\sup_x |\widehat{f_n}(x) - f(x)| \leqslant \varepsilon'_n$ for a sequence ε'_n such that $\varepsilon'_n \to 0$.
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

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

With

- *If the sequence of dependence coefficients of $\mathbf{X}$ has an **exponential decay** rate $\alpha(q) \leqslant Bb^{-q}$, or $\theta(q) \leqslant Bb^{-q}$, for some $b < 1$, then:*

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

- *If the sequence of dependence coefficients of* $\mathbf{X}$ *has a* ***polynomial decay*** *rate* $\alpha(q) \leqslant Bq^{-b}$ *or* $\theta(q) \leqslant Bq^{-b}$ *with* $b > b_0$, *then:*

$$\varepsilon_n = (\ln^a n / n^{1-\gamma})^{2/(d+1)}.$$

Explicit values of b_0, γ_1 *and* γ_2 *are respectively:*

- *Strong mixing:* $b_0 = 1$, $a = \dfrac{b-2-d}{b-\lambda}$, *and* $\gamma = \dfrac{\lambda+1}{\lambda+b}$, *with* $\lambda = \dfrac{3d+1}{d+1}$.
- *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. Explicit rates for uniform consistency of kernel density estimators under dependence

Density estimation is an important question and it is a widely investigated problem in the literature. We defer the reader to [26] under independence and e.g. [20] under strongly mixing and [22] for the weakly dependent case. Indeed the density is a $d-$th order partial derivative of the c.d.f. which is easily fitted from the empirical c.d.f. F and describes the marginal distribution of a stationary process. Each functional of F, including f, is thus of a main importance for statistics. Anyway the herein assumptions do not fit the current setting, see Remark 6. The density estimation error relies on the dependence structure, and on the regularity of the density and the dimension d.

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

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

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:

$$\widehat{f_n}(x) = \frac{1}{nh_n^d} \sum_{t=1}^{n} K\left(\frac{X_t - x}{h_n}\right). \tag{9}$$

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

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

(K2) $K : \mathbb{R}^d \to \mathbb{R}$ is of order k for some $k \geqslant 2$.

(M1) For some $k \geqslant 2$, $f \in C^k$ and $\|f^{(k)}\|_{op}$ 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) $\mathbb{E}\|X_0\|^c < \infty$ for some $c > 2$.

Proposition 5. *Let* $\mathbb{X}_n$ *be a n-sample of a strictly stationary sequence* $\mathbf{X}$ *and* $\widehat{f_n}$ *be defined by* (9). *Assume that (K1), (M2), (M3) and (M4) hold, and that (K2) and (M1) hold for some* $k \geqslant 2$.

- *If the sequence of dependence coefficients of* $\mathbf{X}$ *has an* ***exponential decay*** *rate*

(A1) *In the strong mixing case* $\alpha(q) \leqslant Bb^{-q}$, *for some* $b < 1$,

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

(T1) *In the θ-weak dependent case,* $\theta(q) \leqslant Bb^{-q}$, *for some* $b < 1$,

– *If* $k > d+2$, *the choice* $h_n = O\left(\frac{\ln n}{n}\right)^{\frac{1}{d+2k}}$ *yields* $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k}{2k+d}}\right)$.

– *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* $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n}\right)^{\frac{k(1-\gamma)}{3d+4}}\right)$.

- *If the sequence of dependence coefficients of* $\mathbf{X}$ *has a* ***polynomial decay*** *rate then,*

  (A2) *In the strong mixing case* $\alpha(q) \leqslant Bq^{-b}$ *for some* $b > \max(2+3/d, d+2)$.
  Let $\gamma_0 = \dfrac{d+2-\frac{d(d-2)}{2k+d}}{b+1-\frac{d(d-2)}{2k+d}}$, *then the choice* $h_n = O\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\frac{1}{d+2k}}$ *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) *In the* θ*-weak dependent case* $\theta(q) \leqslant Bq^{-b}$ *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* $h_n = O\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\frac{1}{d+2k}}$ *for any* $\gamma \in (\gamma_0, 1)$, *yields* $\varepsilon'_n = O_{a.s.}\left(\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\frac{k}{2k+d}}\right)$.

Remark 6. Comments on the assumptions and on the results.

- The $(\ln n)$-terms may be omitted in the above mentioned rates because the presence of an additional power term make them essentially non optimal.
- In the strong mixing case with polynomial or geometric decreasing rates, [20] achieved optimal asymptotic rates for ε'_n that generalize [26]'s theorem 2.3. Such rates are also achieved in [22] and [23]) in the $\theta-$weakly dependent case with polynomial and geometric decreasing rates of the coefficients. However, the assumptions on f, h_n and the rate of decay of the dependence coefficients under which such rates may be achieved are more demanding than ours. Namely, it is required some Hölder conditions on f. Note that in case of strong mixing coefficients decaying at a polynomial rate to 0 not quickly enough, [6] also provide convergence results.
- To have a control of the error term on the whole set $\mathbb{R}^d$, we need that the marginal density f satisfies (M5) : $\lim_{\|x\|\to\infty} f(x) = 0$. Remark that if (M1) and (M2) hold with some $k \geqslant 2$ thus every derivative of order less or equal to k are bounded and this entails (M5), see [31].
- (M2) and (M3) are necessary in particular to have an equivalence of the variance $\sigma_n^2 = \mathrm{var}(\widehat{f_n})$. (M2) can be dispensed with at the cost of additional technicality. To cancel the singularities, one may for instance consider $g = vf$ with $v(x) = (x_1 \times \cdots \times x_d)^{\varepsilon}$ and an estimator $\widehat{g_n} = v\widehat{f_n}$ of g. Thus the variance term admits the convenient decay rate.

4.4. Main arguments for Theorem 3.

In Subsection 4.3 we obtained explicit decay rates $(\varepsilon'_n)_{n>0}$ such that $\sup_x |\widehat{f_n}(x) - f(x)| \leqslant \varepsilon'_n$ with $\varepsilon'_n \to 0$.
This entails: $L_{t+\varepsilon'_n} \cap \mathbb{X}_n \subset \mathbb{X}_n^+(t) \subset L_{t-\varepsilon'_n}$, thus, for any $R \geqslant 0$ we have $C_R(L_{t+\varepsilon'_n} \cap \mathbb{X}_n) \subset C_R(\mathbb{X}_n^+(t)) \subset C_R(L_{t-\varepsilon'_n})$.

First notice that, using the regularity of the level t we have that, for n large enough

1. By Proposition 4 $L_{t-\varepsilon'_n}$ and $L_{t+\varepsilon'_n}$ admit the R-rolling ball property (in other world $C_R(L_{t-\varepsilon'_n}) = L_{t-\varepsilon'_n}$ and $C_R(L_{t+\varepsilon'_n}) = L_{t+\varepsilon'_n}$
2. By use of the mean value theorem, we also derive $L_{t-\varepsilon'_n} \subset L_t \oplus \frac{2\varepsilon'_n}{m_t}$ and $L_t \ominus \frac{2\varepsilon'_n}{m_t} \subset L_{t+\varepsilon'_n}$.

The proof is then divided into two parts.

1. Firstly notice that we yet have $\widehat{L}_{t,R} = C_R(\mathbb{X}_n^+(t)) \subset C_R(L_{t-\varepsilon'_n}) = L_{t-\varepsilon'_n} \subset L_t \oplus \frac{2\varepsilon'}{m_t} B$.
2. By the r-convex hull properties, as in the section § 3 devoted to support estimation part, we have:

$$\left(\forall O,\ \underline{\mathrm{d}}\,(O, L_{t+\varepsilon'_n}) \leqslant R-\varepsilon \Rightarrow B(O,R) \cap (L_{t+\varepsilon'_n} \cap \mathbb{X}_n) \neq \emptyset\right) \Longrightarrow d_H(L_{t+\varepsilon'_n}, C_R(L_{t+\varepsilon'_n} \cap \mathbb{X}_n)) \leqslant \varepsilon$$

To bound the probability p_n that there exists O with $d(O, L_{t+\varepsilon'_n}) \leqslant R-\varepsilon_n$ and $\mathbb{P}(B(O,R) \cap (L_{t+\varepsilon'_n}) \cap \mathbb{X}_n) = \emptyset$ we morally proceed as in the proof of results for the support estimation with an additional initial step.

(a) By use of the inner rolling ball property we can find O' such that $\|O - O'\| = 2R - \varepsilon_n$ and $B(O', R) \subset L_{t+\varepsilon'_n}$ thus $B(O', R) \cap B(O, R) \cap \mathbb{X}_n = \emptyset$. We thus have

$$p_n \leqslant \mathbb{P}\Big(\exists O, \underline{\mathrm{d}}\,(O, L_{t+\varepsilon'_n}) \leqslant \varepsilon_n, \exists O', B(O', R) \subset L_{t+\varepsilon'_n}, \|O - O'\| \leqslant 2R - \varepsilon_n, B(O', R) \cap B(O, R) \cap \mathbb{X}_n = \emptyset\Big)$$

(b) As for the support estimation, we define coverings $O_1, \ldots, O_{v_n}$ and $O'_1, \ldots, O'_{v'_n}$ on $\mathbb{R}^d$ such that $(L_{t+\varepsilon'_n} \oplus (R - \varepsilon_n)B) \setminus L_{t+\varepsilon'_n}$ and $L_{t+\varepsilon'_n} \ominus RB$ are covered with balls with radius $\varepsilon_n / \ln n$. Let us introduce $B_{i,j}(R, \varepsilon_n) = B(O'_j, R - \varepsilon_n/\ln n) \cap B(O_i, R - \varepsilon_n/\ln n)$, as for the case of support estimation then using (2) then the display in the point (a) may be written as:

$$\begin{aligned} p_n &\leqslant \sum_{i,j} \mathbf{1}_{\left\{\|O_i - O'_j\| \leqslant 2R - \varepsilon_n(1 - 2/\ln n)\right\}} \mathbb{P}\left(B(O'_j, R - \varepsilon_n/\ln n) \cap B(O_i, R - \varepsilon_n/\ln n) \cap \mathbb{X}_n = \emptyset\right) \\ &\leqslant \sum_{i,j} \mathbf{1}_{\left\{\|O_i - O'_j\| \leqslant 2R - \varepsilon_n(1 - 2/\ln n)\right\}} \mathbb{P}\left(B_{i,j}(R, \varepsilon_n) = \emptyset\right) \\ &\leqslant \sum_{i,j} \mathbf{1}_{\left\{\|O_i - O'_j\| \leqslant 2R - \varepsilon_n(1 - 2/\ln n)\right\}} \left\{(1 - \mathbb{P}((B_{i,j}(R, \varepsilon_n)))^{n/q_n} + \frac{\mathcal{E}_{B_{i,j}(R,\varepsilon_n)}(q_n)}{\mathbb{P}(B_{i,j}(R, \varepsilon_n))}\right\} \end{aligned}$$

The additional term $\mathcal{E}_{B_{i,j}(r,\varepsilon_n)}(q_n)$ writes $\alpha(q_n)$ in the strong mixing case and it is given by (4) in Proposition 2 for the θ-weak dependent case.
In order to conclude, we work out sequences ε_n and q_n such that $\sum_n p_n < +\infty$.

(c) With such ε_n we conclude by arguing that $L_{t+\varepsilon'_n} \subset C_R(L_{t+\varepsilon'_n} \cap \mathbb{X}_n) \oplus \varepsilon_n B \subset \widehat{L}_{t,r} \oplus \varepsilon_n B$. Then, the preliminary remark entails $L_t \ominus \frac{2\varepsilon'_n}{m_t} B \subset \widehat{L}_{t,R} \oplus \varepsilon_n B$, that, with regularity of the level finally implies $L_t \subset \widehat{L}_{t,R} \oplus \Big(\varepsilon_n + \frac{2\varepsilon'_n}{m_t}\Big) B$.

5. Detailed Proofs

5.1. Proofs of Propositions 1, 2, 3

Set first $k = \lfloor n/q \rfloor$ and $t_i = iq$ for $1 \leq i \leq k$. Define $Z_i = \mathbf{1}_{\{X_{t_i} \notin E\}}$ and $p_k = \mathbb{E}(Z_1 \times \cdots \times Z_k)$. 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 bounds of $\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, Z_k)$ by some $\mathcal{E}_E(q)$ then applying a recursion on $p_k \leq (1 - p)p_{k-1} + \mathcal{E}_E(q)$. We obtain $p_k \leq (1 - p)^{k-1} p_1 + \sum_{i=1}^{k} (1 - p)^i \mathcal{E}_E(q)$ that is

$$p_k \leqslant (1 - p)^k + \frac{\mathcal{E}_E(q)}{p}$$

Proof of Proposition 1. From the heredity of strong mixing, the sequence $(Z_t)_t$ is still strong mixing with mixing coefficients less than $\alpha(q)$, and the fact that $Z_1 \times \cdots \times Z_{k-1}$ 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 for $k - 1$ 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 $|\mathrm{cov}(Z_1 \times \cdots \times Z_{k-1}, h(X_k))| \leq \theta(q)\,\mathrm{Lip}\,(h)$ 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 [15] we consider $H^-_{x,r,\varepsilon}$ and $H^+_{x,r,\varepsilon}$ two Lipschitz approximations $\mathbb{R}^d \to \mathbb{R}$ of $\mathbf{1}_{\{u\notin E\}}$, with

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

then

$$\begin{aligned}\operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, Z_k) &\le \mathbb{E}(Z_1\times\cdots\times Z_{k-1}H^+_{x,r,\varepsilon}(X_k)) - \mathbb{E}(Z_1\times\cdots\times Z_{k-1})\mathbb{E}(H^-_{x,r,\varepsilon}(X_k))\\ &\le \operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, H^+_{x,r,\varepsilon}(X_k)) + \mathbb{E}(H^+_{x,r,\varepsilon}(X_k) - H^-_{x,r,\varepsilon}(X_k))\end{aligned}$$

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

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

Then the functions $m^\pm_{r,\varepsilon}$ are piecewise linear and satisfy $\operatorname{Lip} m^\pm_{r,\varepsilon} = 1/\varepsilon$.
Now we define $H^+_{x,r,\varepsilon}(u) = 1 - m^-_{r,\varepsilon}(\|u-x\|)$ and $H^-_{x,r,\varepsilon}(u) = 1 - m^- + r, \varepsilon(\|u-x\|)$, 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 $\operatorname{Lip}(H^\pm_{x,r,\varepsilon}) = 1/\varepsilon$. 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.

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

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 $\varepsilon \leqslant r$: $\operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, Z_k) \le \frac{1}{\varepsilon}\theta(q) + A'\varepsilon r^{d-1}$.
For the reverse inequality, we obtain analogously for all $\varepsilon \in [0, r]$ $\operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, Z_k) \ge -\frac{1}{\varepsilon}\theta(q) - A'\varepsilon r^{d-1}$
Indeed

$$\begin{aligned}\operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, Z_k) &\ge \mathbb{E}(Z_1\times\cdots\times Z_{k-1}H^-_{x,r,\varepsilon}(X_k)) - \mathbb{E}(Z_1\times\cdots\times Z_{k-1})\mathbb{E}H^+_{x,r,\varepsilon}(X_k)\\ &\ge \operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, H^-_{x,r,\varepsilon}(X_k)) - \mathbb{E}(H^+_{x,r,\varepsilon}(X_k) - H^-_{x,r,\varepsilon}(X_k))\end{aligned}$$

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 (5), we have for some $c > 0$:

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

Thus according to Proposition 7 in Appendix A, if $\ell \leqslant r_x \wedge r_y$ and $\varepsilon \leqslant r_x \wedge r_y$: $\mathbb{E}(\overline{H}^+(X_k) - \overline{H}^-(X_k)) \le c\mathcal{V}(r_x, r_y, \ell, \varepsilon)$. As in case 1, this allows to conclude that $\operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, Z_k) \le \frac{2}{\varepsilon}\theta(q) + c\mathcal{V}(r_x, r_y, \ell, \varepsilon)$. As in the end of the proof for case 1 we infer $|\operatorname{cov}(Z_1\times\cdots\times Z_{k-1}, Z_k)| \le \frac{2}{\varepsilon}\theta(q) + c\mathcal{V}(r_x, r_y, \ell, \varepsilon)$, which concludes the proof of point 2 in Proposition 2.

□

Proof of Proposition 3. This proof is based on the inequality $p_n \ge p^n$, and it is left to the reader. □

5.2. Preliminary results for 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 $\mathbf{X}_n$, as well as that of its related quantities. In this preliminary section we mostly only present rephrasing of results obtained in [40], [35] and [3]. Thus proofs might be either left for the reader -as for Properties 1 and 2-, or only sketched -as for Corollary 1. To have a self content paper we retate those results, 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 $\mathbb{R}^d$, then $C_R(S)^c = \{x \in \mathbb{R}^d, \exists\, O \in \mathbb{R}^d$ such that $x \in \mathring{B}(O,R)$ and $S \cap \mathring{B}(O,R) = \emptyset\}$ so that*

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

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

w.r.t. the Hausdorff distance.

Sets satisfying the inside and outside rolling ball properties have nice properties listed in theorem 1 of [41]; we recall them below in Property 2, together with some hints; they are illustrated in the 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 $R < \min(R_o, R_i)$ we have $(S \ominus RB) \oplus RB = S = (S \oplus RB) \ominus RB$, in fact the second part of the equality does not involve rolling ball conditions.*
3. *For each $z \in \partial S$, the unit vector η_z normal to ∂S and pointing outside S is well defined and $\eta_z = (O_z^o - z)/R_o$.*
4. *For each $x \in S$ with $\underline{d}(x, \partial S) \leqslant R_i$, the projection $\pi_{\partial S}(x)$ onto ∂S is well defined and $x = \pi_{\partial S}(x) - \underline{d}(x, \partial S)\eta_{\pi_{\partial S}(x)}$, $O_x^i = \pi_{\partial S}(x) - R_i\eta_{\pi_{\partial S}(x)}$.*
5. *For all $y \in S^c$ with $\underline{d}(y, \partial S) \leqslant R_o$, the projection $\pi_{\partial S}(y)$ onto ∂S is well defined and $y = \pi_{\partial S}(y) + \underline{d}(y, \partial S)\eta_{\pi_{\partial S}(y)}$, $O_y^o = \pi_{\partial S}(y) + R_o\eta_{\pi_{\partial S}(y)}$.*
6. *For all $z, z' \in \partial S$, $\|\eta_z - \eta_{z'}\| \leqslant \dfrac{\|x-y\|}{\min(R_0, R_i)}$. Thus ∂S has a positive reach $\min(R_o, R_i)$ and $\pi_{\partial S}|_{\partial S \oplus(\min(R_i,R_o))B}$ is continuous (this follows from theorem 4.8 in [24]).*

Moreover, using the arguments of [35], 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 6. *Let $S' \subset S \subset \mathbb{R}^d$. Assume that balls of radius R_o roll freely outside S and that balls of radius R_i roll freely inside S. Let $R < R_o$ and $\varepsilon < \min(R_i, R)$. If for each $O \in \mathbb{R}^d$ such that $\underline{d}(O, S) \leqslant R - \varepsilon$, we have that $B(O,R) \cap S' \neq \emptyset$, then*

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

Proof of Proposition 6. 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 $\underline{d}(x, \partial S) = \varepsilon' \geqslant \varepsilon$ 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 $\underline{d}(O, S) \leqslant R - \varepsilon'$ that contradicts $B(O,R) \cap S' = \emptyset$. □

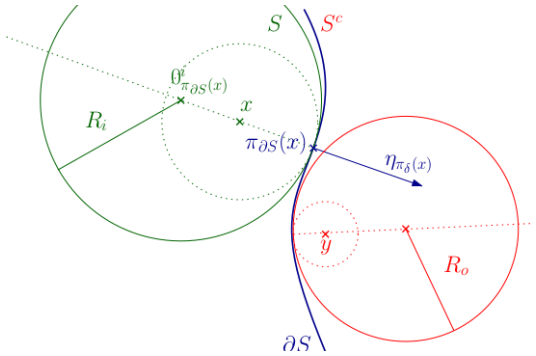

Fig. 1: All you have to have in mind concerning the rolling ball property

From Proposition 6 and by using theorem 4.1 in [3] we derive general deterministic results on the R-convex hull:

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

$$d_H(\partial C_R(S'), \partial S) \leqslant \varepsilon, \qquad d_H(C_R(S'), S) \leqslant \varepsilon,$$
$$\left| |C_R(S')|_d - |S|_d \right| \leqslant |\partial S|_{d-1}\varepsilon(1 + o(1)), \qquad \text{as } \varepsilon \to 0.$$

Moreover if $d_H(S', S) \leqslant \dfrac{1}{2}\dfrac{R\min(R_i, R_o)}{R + \min(R_i, R_o)}$, 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 provide the main arguments. By Proposition 6 we have

$$S \ominus \varepsilon B \subset C_R(S') \subset S \tag{10}$$

- Since $\varepsilon \leqslant \min(R_i, R)$ and $R \leqslant R_o$ we have $\varepsilon \leqslant \min(R_i, R_o)$ and by application of point 2 in proposition 2 $S \subset C_R(S') \oplus \varepsilon B$, the fact that $C_R(S') \subset S \subset S \oplus \varepsilon B$ yields $d_H(C_R(S'), S) \leqslant \varepsilon$.
- To get a upper bound of $d_H(\partial C_R(S'), \partial S)$ we quote that:
  - By inclusion (10), for all $x \in \partial C_R(S')$ we have $x \in S$. Suppose first that $x \in \mathring{S}$ and set $\delta_x = \underline{d}(x, \partial S) > 0$. For each $0 < t < \delta_x$ there exits $x_t \notin C_R(S')$ with $\|x - x_t\| \leqslant t$, thus $x_t \in S$. Therefore, by the first part of inclusion (10), we have $\underline{d}(x_t, \partial S) < \varepsilon$. Thus $\underline{d}(x, \partial S) < \varepsilon + t$ and we have $\underline{d}(x, \partial S) \leqslant \varepsilon$ as $t \to 0$. Secondly, if $x \in \partial S$ then $\underline{d}(x, \partial S) = 0$ so we conclude that for all $x \in \partial C_R(S')$: $\underline{d}(x, \partial S) \leqslant \varepsilon$.
  - For all $x \in \partial S$, let $y = x - \varepsilon\eta_x$ with η_x defined in Property 2, then by the inside rolling ball property of S, we have $y \in S \ominus \varepsilon B$ and thus the segment (x, y) intersects $\partial C_R(S')$ so that $\underline{d}(x, \partial C_R(S')) \leqslant \varepsilon$.
- $|C_R(S')|_d \leqslant \left| S \setminus (S \ominus \varepsilon B) \right|_d \approx \varepsilon|\partial S|_{d-1}$, 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 [24], and for its outer part see [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 [3]. In [3] it is proved that $\psi : \partial C_R(S') \to \partial S$ defined as $x \mapsto \pi_{\partial S}(x)$ is an homeomorphism, as it is mentioned in Property 2 point 4, π_S the projection onto ∂S is well define since $\varepsilon \leqslant R_i$. The Figure 2 supports the construction of $\varphi : C_R(S') \to S$ and the proof that it is an homeomorphism. To get a proper definition of φ, introduce, for $x \in S$, $\underline{\mathrm{d}}(x, \partial S) \leqslant R_i$, $T_0(x) = \pi_{\partial S}(x) - 2\varepsilon\eta_{\pi_{\partial S}(x)}$ where η_z is defined in Property 2 and $T_1(x) = \psi^{-1}(\pi_{\partial S}(x))$. Then $\varphi : C_R(S') \to S$ is defined as

$$\begin{cases} \varphi(x) = x, & \text{if } x \in S \ominus 2\varepsilon B, \\ \varphi(x) = T_0(x) + 2\varepsilon \dfrac{\|x - T_0(x)\|}{\|T_1(x) - T_0(x)\|} \cdot \eta_{\partial S}(x), & \text{if } x \in C_R(S') \setminus (S \ominus 2\varepsilon B). \end{cases}.$$

  Also define $g : S \to C_R(S)$ as follows

$$\begin{cases} g(x) = x, & \text{if } x \in S \ominus 2\varepsilon B, \\ g(x) = T_0(x) + \|T_1(x) - T_0(x)\| \dfrac{\|x - T_0(x)\|}{2\varepsilon} \eta_{\partial S}(x), & \text{if } x \in S \setminus (S \ominus 2\varepsilon B). \end{cases}.$$

  Arguments to prove that φ *is an homeomorphism* for ε small enough, *and thus* $C_R(S') \approx S$, are provided below.

  - $g = \varphi^{-1}$ is easy.
  - The continuity of φ and g on the sets $C_R(S') \setminus \left(S \ominus 2\varepsilon B\right)$ and $S \setminus (S \ominus 2\varepsilon B)$ follows from the continuity of $\pi_{\partial S}$ and from the continuity of $x \mapsto \eta_x$ for $\varepsilon \leqslant \min(R_o, R_i)$: for this we use the last point of Property 2. Then the global continuity of the functions φ and g is a consequence of the fact that $T_0(x) = x$ if $x \in \partial(S \setminus (S \ominus 2\varepsilon B)) = \left\{x \in S,\ d(x, S) = 2\varepsilon\right\}$. This point is thus a consequence of the regularity of the support for ε small enough.

□

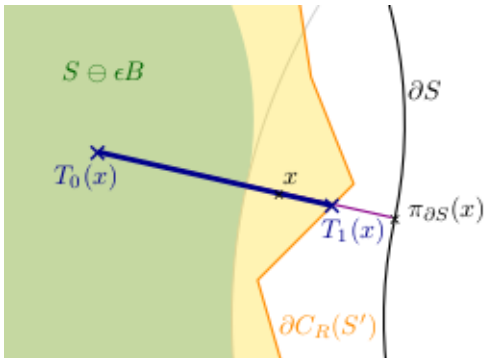

Fig. 2: The white zone is S, the yellow zone is $C_R(S')$ and its boundary is indicated by the thick orange line, moreover the green zone is $S \ominus \varepsilon B$. $\varphi : [T_0(x), T_1(x)] \to [T_0(x), \pi_{\partial S}(x)]$ is an onto homothety with ratio $\dfrac{\|T_0(x) - \pi_{\partial S}(x)\|}{\|T_0(x) - T_1(x)\|} = \dfrac{2\varepsilon}{\|T_0(x) - T_1(x)\|}$, and centered at $T_0(x)$.

To derive convergence rates for the R-convex hull of a set of points $\mathbb{X}_n$, we will exhibit sequences ε_n such that

$$\sum_{n=1}^{\infty} \mathbb{P}(\exists O, \underline{\mathrm{d}}(O, S) < R - \varepsilon_n, B(O, R) \cap \mathbb{X}_n = \emptyset) < +\infty.$$

Also we will prove that $d_H(\mathbb{X}_n, S) \to 0$ almost surely in order to apply the last part of Corollary 1; we will also deduce the necessary topological guarantees and the convergence for the measure of the boundary.

The dependence is taken into account when we deal with $\mathbb{P}(B(O,R)\cap\mathbb{X}_n=\emptyset)$ in which the Propositions 1 or 2 are used; this replaces a trivial product in the independent case.

We now use the notations and results of Propositions 1 and 2, and according to the dependence conditions we obtain under the assumptions of the corresponding proposition that if:

(A) $\mathbf{X}$ is strong mixing then $\mathcal{E}_r(q)=\alpha(q)$,

(T) $\mathbf{X}$ is θ-weak dependent then $\mathcal{E}_r(q)=\min\limits_{t\in[0,r]}\left(\dfrac{\theta(q)}{t}+c\omega_d 2^d t r^{d-1}\right)$.

Lemma 4. *Assume that the marginal density f belongs to the set $\mathcal{M}_{R_0,R_i,f_0,\alpha}$, defined in Definition 2, we set:*

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

Then there exists a constant C depending on the parameters f_0,α,d,R_i,R such that:

$$P_n \lesssim (\ln n)^d\varepsilon_n^{-d}\left(\exp\left(-C\frac{n}{q}\left(\varepsilon_n-\frac{\varepsilon_n}{\ln n}\right)^{\frac{d+1+2\alpha}{2}}\right)+\frac{\mathcal{E}_{R-\frac{\varepsilon_n}{\ln n}}(q)}{C\left(\varepsilon_n-\frac{\varepsilon_n}{\ln n}\right)^{\frac{d+1+2\alpha}{2}}}\right). \tag{11}$$

$$\mathbb{P}(d_H(\mathbb{X}_n,S)\geqslant r_n) \lesssim \left(\frac{\ln n}{r_n}\right)^d\left(\exp\left(-C'\frac{n}{q}\left(r_n-\frac{r_n}{\ln n}\right)^{d+\alpha}\right)+\frac{\mathcal{E}_{r_n-\frac{r_n}{\ln n}}(q)}{\left(r_n-\frac{r_n}{\ln n}\right)^{d+\alpha}}\right). \tag{12}$$

Here, according to the dependence frame, $r\mapsto\mathcal{E}_r(q)$ is defined in conditions (**A**) *and* (**T**) *above from the Assumptions of Propositions 1 and 2 respectively.*

Proof of lemma 4. Let us cover $S\oplus(R-\varepsilon_n)B$ with a finite number of $\nu_n\leqslant C_d\left(\dfrac{\mathrm{diam}(S)+R_o}{\varepsilon_n}\ln n\right)^d$ deterministic balls of radius $\varepsilon_n/\ln n$ and centered at $x_i\in S\oplus(R-\varepsilon_n)B$.
Then the Propositions 1 and 2 together with the triangle inequality imply:

$$P_n\leqslant\sum_{i=1}^{\nu_n}\mathbb{P}\left(B\left(x_i,R-\frac{\varepsilon_n}{\ln n}\right)\cap\mathbb{X}_n=\emptyset\right)\leqslant\sum_{i=1}^{\nu_n}\left((1-p_i)^{\frac{n}{q}}+\frac{\mathcal{E}_{R-\frac{\varepsilon_n}{\ln n}}(q)}{p_i}\right),\quad\text{with } p_i=\mathbb{P}\Big(X_1\in B\Big(x_i,R-\frac{\varepsilon_n}{\ln n}\Big)\Big).$$

By Proposition 8 we obtain

$$p_i\geqslant f_0C''_{d,\alpha}\left(\frac{\min((R-\frac{\varepsilon_n}{\ln n}),R_i)}{2}\right)^{\frac{d-1}{2}}\left\{\min\left(\left(\varepsilon_n-\frac{\varepsilon_n}{\ln n}\right),\frac{\min((R-\frac{\varepsilon_n}{\ln n}),R_i)}{2}\right)\right\}^{\frac{d+1+2\alpha}{2}}$$

Since $\varepsilon_n\to0$ as $n\to\infty$, thus for any $C<f_0C''_{d,\alpha}\left(\frac{\min(R,R_i)}{2}\right)^{\frac{d-1}{2}}$ 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 (11).
Suppose now that $d_H(\mathbb{X}_n,S)\geqslant r_n$ 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 :

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

Then $B(\Omega, r_n/3) \subset S$ and $B(\Omega, r_n/3) \subset B(x, r_n)$.
As a first conclusion: if $d_H(\mathbb{X}_n, S) \geqslant r_n$, 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 the set $S \ominus \frac{r_n}{3}B$ with N balls of radius $r_n/(3\ln n)$, centered at $y_1, \ldots, y_N$ in $S \ominus \dfrac{2r_n}{3}B$ with $N \leqslant C_d \dfrac{\left(3\mathrm{diam}(S)\ln n\right)^d}{r_n^d}$.
Now

$$\begin{aligned}\mathbb{P}(d_H(\mathbb{X}_n, S) \geqslant r_n) &= \mathbb{P}\left(\exists \Omega \in S \ominus \frac{2r_n}{3}B,\ B\left(\Omega, \frac{r_n}{3}\right) \cap \mathbb{X}_n = \emptyset\right) \\ &\leqslant \sum_{i=1}^{N} \mathbb{P}\left(B\left(y_i, \frac{r_n}{3}\left(1 - \frac{1}{\ln n}\right)\right) \cap \mathbb{X}_n = \emptyset\right),\end{aligned}$$

from the triangular inequality.
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 8: $q_i \geqslant f_0 C_{d,\alpha}\left(\frac{r_n}{6}\left(1 - \frac{1}{\ln n}\right)\right)^{d+\alpha}$.
Thus, with $C' = f_0 C_{d,\alpha}/3^{d+\alpha}$ we obtain

$$\mathbb{P}(d_H(\mathbb{X}_n, S) \geqslant r_n) \lesssim \left(\frac{\ln n}{r_n}\right)^d \left(\exp\left(-C'\frac{n}{q}\left(r_n - \frac{r_n}{\ln n}\right)^{d+\alpha}\right) + \frac{\mathcal{E}_{r_n - \frac{r_n}{\ln n}}(q)}{\left(r_n - \frac{r_n}{\ln n}\right)^{d+\alpha}}\right). \tag{13}$$

□

5.3. Proof of Theorem 1

First, let us focus on the i.i.d. and negatively associated cases. Those conditions entail, for all $q \in \mathbb{N}^*$ and $R \in \mathbb{R}$, that $\mathcal{E}_R(q) = 0$.

Apply Lemma 4 and Equation (11) 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 from equation (12) with $q = 1$ and $r_n = \left(\dfrac{3d+2\alpha}{C'(d+\alpha)}\dfrac{\ln n}{n}\right)^{\frac{1}{d+\alpha}}$, we obtain: $\mathbb{P}(d_H(\mathbb{X}_n, S) \geqslant r_n) \lesssim \ln^d n \cdot n^{-2+o(1)}$.

The Borel-Cantelli lemma and the Corollary 1 together allow to conclude the proof of Theorem 1.

5.4. Proof of Theorem 2

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.

Proof under exponential decays.

- In the strong mixing case, $\mathcal{E}_R(q) \leqslant Bb^q$ for all $R \in \mathbb{R}$.
  - Choose $\varepsilon_n = \gamma(\ln^2 n/n)^{\frac{2}{d+1+2\alpha}}$ with $\gamma > \gamma_0^{ed} = \dfrac{2(2d+1+2\alpha)(3d+1+2\alpha)}{(d+1+2\alpha)^2 C\ln(b^{-1})}$, let write $\gamma = \gamma_0^{ed}(1+u)$.
    Set $q_n = \beta_1 \ln n, \beta_1 = \dfrac{2(2d+1+2\alpha)}{(d+1+2\alpha)\ln(b^{-1})}\left(1 + \dfrac{u}{2}\right)$, we finally introduce $v = \dfrac{u}{2+u}$. By Equation (11)

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

    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

$$P_n \lesssim \ln^d n\left(n^{-1-\frac{3d+1+2\alpha}{d+1+2\alpha}v+o(1)} + n^{-1-\frac{2d+1+2\alpha}{d+1+2\alpha}u}\right), \tag{14}$$

    thus $\sum_n P_n < +\infty$.

– Choose $r_n = \left(c\frac{\ln^2 n}{n}\right)^{\frac{1}{d+\alpha}}$ and $q_n = \beta \ln n$ by Equation (12):

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

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 $\sum_n \mathbb{P}(d_H(\mathbb{X}_n, S) \geqslant r_n) < +\infty$ and we let $r_n \to 0$.

- In the θ-weak dependence case, we have

  – Now $\mathcal{E}_{R-\varepsilon_n/\ln n}(q_n) = \min_{t\in[0,R-\varepsilon_n/\ln n]} \left(Bb^{q_n}/t + c\omega_d 2^d t R^{d-1}\right)$. For any choice of sequence $\varepsilon_n \to 0$, and for n large enough, $t_n = \sqrt{b^{q_n}}$ is in $[0, R - \varepsilon_n/\ln n]$ thus $\mathcal{E}_{R-\varepsilon_n/\ln n}(q_n) \leqslant B' \exp(q_n \ln(b)/2))$ 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 $\sum_n P_n < +\infty$.

  – Choose again $r_n = \left(c\frac{\ln^2 n}{n}\right)^{\frac{1}{d+\alpha}}$. Here again, $t_n = \sqrt{b^{q_n}} \in [0, r_n - r_n/\ln n]$, 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 $\sum_n \mathbb{P}(d_H(\mathbb{X}_n, S) \geqslant r_n) < +\infty$ and $r_n \to 0$ for a suitable constant c.

To conclude, in all the cases we found $\varepsilon_n = O\left(\left(\frac{\ln^2 n}{n}\right)^{\frac{2}{2+1+\alpha}}\right)$ 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.

Proof under polynomial decays. According to the dependence assumption we check that

- *In the strong mixing case,* $\mathcal{E}_R(q_n) \leqslant Bq_n^{-b}$, 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 $\varepsilon_n = \left(A\frac{\ln^a n}{n^{1-\gamma}}\right)^{\frac{2}{d+1+2\alpha}}$ and $q_n = \frac{n^\gamma}{(\ln n)^{a-1}}$. From Equation (11) 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 $r_n = \left(A'\frac{\ln^{a'} n}{n^{1-\gamma'}}\right)^{\frac{1}{d+\alpha}}$ and $q_n = n^{\gamma'}/(\ln n)^{a-1}$.
  From Equation (11) it comes that

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

  This yields the summability of the series $\sum_n P_n$ for A large enough.

- *Under θ-weak dependence,* Here $\mathcal{E}_{R-\varepsilon_n/\ln n}(q_n) = \min_{t\in[0,R-\varepsilon_n/\ln n]} \left(\frac{Bq_n^{-b}}{t} + c\omega_d 2^d t R^{d-1}\right)$. For any choice of sequence $\varepsilon_n \to 0$ and 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 $r_n = \left(A'\dfrac{\ln^{a'} n}{n^{1-\gamma'}}\right)^{\frac{1}{d+\alpha}}$, $q_n = \dfrac{n^{\gamma'}}{(\ln n)^{a-1}}$ and $t_n = \sqrt{q_n^{-b}}$. Check that $t_n < \varepsilon_n$ From Equation (11) we obtain

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

  This again yields the summability of the series $\sum_n P_n$ for A' large enough.

5.5. Proof of Theorem 3.

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

$$\widehat{L}_{t,R} \subset L_t \oplus \frac{2\varepsilon'_n}{m_t} B. \tag{15}$$

First a simple chain of inclusions entails $\widehat{L}_{t,R} = C_R(\{X_i, \widehat{f}_n(X_i) \geqslant t\}) \subset C_R(\{X_i, f(X_i) \geqslant t - \varepsilon'_n\}) \subset C_R(L_{t-\varepsilon'_n})$. In the sequel, we consider n large enough to have $\varepsilon'_n \leqslant \max(\Delta_t, t - t_0)$ so that Proposition 4 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 $z' \in L_{t-\varepsilon'_n}$, let $z = z' + 2\dfrac{\varepsilon'_n}{m_t}\dfrac{f^{(1)}(z')}{\|f^{(1)}(z')\|}$:

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

thus, if n is large enough so that $\varepsilon'_n < m_t^2/2f^{(2)}_{\max}$, we have $f(z) \geqslant t$, thus $L_{t-\varepsilon'_n} \subset L_t \oplus \dfrac{2\varepsilon'_n}{m_t} B$ which ends the proof for inclusion (15).

We will now prove that, with probability 1, for n large enough:

$$L_t \subset \widehat{L}_{t,R} \oplus \left(\varepsilon_n + 2\frac{\varepsilon'_n}{m_t}\right) B. \tag{16}$$

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 $x \in L_{t+\varepsilon'_n}$ such that $\underline{\mathrm{d}}\,(x, \partial L_{t+\varepsilon'_n}) \geqslant \varepsilon_n$ 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 $\underline{\mathrm{d}}\,(x, \partial L_{t+\varepsilon'_n}) \geqslant \varepsilon_n$, thus we have $\underline{\mathrm{d}}\,(O, L_{t+\varepsilon'_n}) \leqslant R - \varepsilon_n$.
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 with radius $(\varepsilon_n/\ln n)$ and centered at $x_i \in L_{t+\varepsilon'_n} \oplus (R - \varepsilon_n)B$. Using the fact that the density on $L_{t+\varepsilon'_n}$ is greater than t, we obtain $|L_{t+\varepsilon'_n}| \leqslant t^{-1}$; thus

$$t\,|L_{t+\varepsilon'_n}| \leqslant \int \mathbf{1}_{L_{t+\varepsilon'_n}}(z) f(z) dz \leqslant 1.$$

We obtain

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

Now by Proposition 4, $L_{t+\varepsilon'_n}$ admits both the $(m_t/f^{(2)}_{\max})$-inside and outside rolling ball properties.

Thus if we set $x_i^* = \pi_{\partial L_{t+\varepsilon'_n}}(x_i)$, and $y_i = x_i^* + \dfrac{m_t}{f^{(2)}_{\max}} \cdot \dfrac{x_i^* - x_i}{\|x_i^* - x_i\|}$, then we have
$B(y_i, m_t/f^{(2)}_{\max}) \subset L_{t+\varepsilon'_n}$ and $\|x_i - y_i\| \leqslant R - \varepsilon_n + m_t/f^{(2)}_{\max}$.
Set also $E_i = B(x_i, R - \varepsilon_n/\ln n) \cap B(y_i, m_t/f^{(2)}_{\max})$, we obtain

$$P_n \leqslant \sum_{i=1}^{\nu_n} \left((1 - p_i)^{\frac{n}{q}} + \frac{\mathcal{E}_{E_i}(q)}{p_i}\right), \qquad \text{with} \qquad p_i = \mathbb{P}(X_1 \in E_i). \tag{17}$$

A bound of the covering number is

$$\nu_n \leqslant \frac{2^d}{t\omega_d}\left(\frac{\ln n}{\varepsilon_n}\right)^d. \tag{18}$$

By Proposition 8, we derive, with the choices $f_0 = t$, $\alpha = 0$:

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

Since ε_n converges to 0 as $n \to \infty$, then for any $C < t\,C''_{d,0}\left(\frac{\min\left(R, m_t/f^{(2)}_{\max}\right)}{2}\right)^{\frac{d-1}{2}}$ if n is large enough, we finally obtain by using the key inequality (2) and from (17) and (18):

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

This equation is similar to Equation (11) with $\alpha = 0$. The difference consists in $\mathcal{E}_{E_i}$ since E_i is now the intersection of two balls instead of being a ball.

- Under strong mixing, $\mathcal{E}_{E_i}$ does not depend on the set E_i, there is no change and the proof of Theorem 2 still holds. The given choices of ε_n (with $\alpha = 0$) ensures $\sum_n P_n < +\infty$.
- Under θ-weak dependence, then the value of $\mathcal{E}_{E_i}$ depends on the shape of the set E_i. Nevertheless, due to Equations (2) 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.

Now

- In case of exponential decay, there exist suitable constants a_1, a_2 and a_3 such that with $\varepsilon_n = (a_1 \ln^2 n/n)^{(d+1)/2}$, $t_n = \sqrt{a_2\theta(q_n)\varepsilon_n^{\frac{d-1}{2}}}$ and choosing $q_n = a_3 \ln n$, there exists $t_n \in (0, \varepsilon_n)$ with:

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

- In case of polynomial decay with $b > 2$, set $\lambda = \dfrac{5d-1}{2(d+1)}$, $\gamma = \dfrac{1+\lambda}{\frac{b}{2}+\lambda}$ and $a = 2\cdot\dfrac{2-d+b/2}{b-\lambda}$.
  As above there exist suitable constants a_1, a_2 and a_3 such that with $\varepsilon_n = (a_1 \ln^a n/n^{1-\gamma})^{(d+1)/2}$, $q_n = n^\gamma/(\ln^{a-1} n)$, $t_n = \sqrt{a_2\theta(q_n)\varepsilon_n^{(d-1)/2}}$ and choosing $q_n = a_3 \ln n$, there exists $t_n \in (0, \varepsilon_n)$ with:

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

This ends the proof of the inclusions

$$L_{t+\varepsilon'_n} \ominus \varepsilon_n B \subset C_R(\mathbb{X}_n \cap L_{t+\varepsilon'_n}) \subset \widehat{L_{t,R}}. \tag{19}$$

Now considering $L_{t+\varepsilon'_n}$, as in the proof of inclusion (15) we easily obtain $L_t \ominus \frac{\varepsilon'_n}{2m_t}B \subset L_{t+\varepsilon'_n}$ for n large enough in order that $\varepsilon'_n < m_t^2/(2f^{(2)}_{\max})$.

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

Indeed, consider $z \in L_t \ominus 2\varepsilon'_n/m_t \cdot B$, 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 $u = -f^{(1)}(z)/\|f^{(1)}(z)\|$, then we have

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

Thus

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

We thus obtain, with the inclusion (19) that $(L_t \ominus 2\frac{\varepsilon'_n}{m_t}B) \ominus \varepsilon_n B \subset \widehat{L}_{t,R}$. Now from [41] and [24], since L_t admits the inside and outside $(m_t/f^2_{\max})-$rolling ball properties, we obtain that $\varepsilon_n + 2\frac{\varepsilon'_n}{m_t} \leqslant m_t/f^2_{\max}$ 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 (16); this, together with inclusion (15) allow to apply Corollary 1 and we thus conclude the proof of bounds

$$d_H(\widehat{L}_t, L_t) \lesssim \varepsilon_n + \varepsilon'_n, \quad d_H(\partial\widehat{L}_t, \partial L_t) \lesssim \varepsilon_n + \varepsilon'_n.$$

We also obtain the bound $|\|\widehat{L}_t|_d - |L_t|_d| \lesssim \varepsilon_n + \varepsilon'_n$, and the proof of Theorem 2 is complete.
To obtain the homeomorphism property and the convergence for the perimeter estimator, we may also prove

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

using the same techniques as above.

6. Conclusion and perspectives

In conclusion, we have shown that, as in the independent case, the $R-$convex hull of the observations and that of the observations whose estimated density is greater than a threshold, allow us to find the support and the level sets and to derive convergence rates.
Two points at least remain to be explored: the minimax aspects of those convergence rates and the adjustment or the automatic choice of the parameter R. The most difficult point of these three perspectives seems to be the the automatic choice of the parameter R [36] for the estimation of the support. Indeed, the existing method (cf [35]) is based on the asymptotic law of maximal spacing [28] and [2] which is only known under the assumption of independence.
Finally an important application of the above results is the control of the of the territory of competitive species along the time. For this a prediction result for the level sets of the considered density is of importance.

Acknowledgements. 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-AAP2020-0000000013.

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. A generalization of the maximal-spacings in several dimensions and a convexity test. Extremes, 20:605–634, 2014.
[3] C. Aaron, A. Cholaquidis, and R. Fraiman. Estimation of surface area. Electronic Journal of Statistics, 16(2):3751–3788, 2022.

[4] L. Ambrosio, A. Colesanti, and E. Villa. Outer minkowski content for some classes of closed sets. Mathematische Annalen, 342:727–748, 2008.

[5] D. W. K. Andrews. Non strong mixing autoregressive processes. Journal of Applied Probability, 21:930–934, 1984.

[6] P. Ango Nze and R. Rios. Density estimation in $\ll^{\infty}$-norm for mixing processes. JSPI, 83:75–90, 2000.

[7] 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.

[8] N. Baldin. The wrapping hull and a unified framework for estimating the volume of a body. arXiv: Statistics Theory, 2017.

[9] A. Cholaquidis, Fraiman, and M. Hernández-Banadik. Home-range estimation under a restricted sample scheme. Journal of Nonparametric Statistics, 31:1–20, 2023.

[10] 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.

[11] A. Cholaquidis, R. Fraiman, E. Mordecki, and C. Papalardo. Level set and drift estimation from a reflected brownian motion with drift. Statistica Sinica, 31:29–51, 2021.

[12] M. Coblenz, R. Dyckerhoff, and Oliver Grothe. Nonparametric estimation of multivariate quantiles. Environmetrics, 29, 2018.

[13] A. Cuevas and A. Rodríguez-Casal. On boundary estimation. Advances in Applied Probability, 36(2):340–354, 2004.

[14] J. Dedecker and P. Doukhan. A new covariance inequality and applications. Stochastic Processes and their Applications, 106:63–80, 2003.

[15] 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.

[16] 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.

[17] 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.

[18] P. Doukhan. Mixing, properties and models, volume 85 of Lecture Notes in Statistics. Springer-Verlag, New York, 1994.

[19] P. Doukhan. Stochastic models for time series, volume 80. Springer, 2018.

[20] P. Doukhan and P. Ango Nze. Functional estimation for time series: Uniform convergence properties. Journal of Statistical Planning and Inference, 68:5–29, 1998.

[21] P. Doukhan and S. Louhichi. A new weak dependence condition and applications to moment inequalities. Stochastic Processes and their Applications, 84:313–342, 1999.

[22] P. Doukhan and S. Louhichi. Functional estimation of a density under a new weak dependence condition. Scandinavian Journal of Statistics, 28:325–341, 2001.

[23] 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.

[24] H. Federer. Curvature measures. Transactions of the American Mathematical Society, 93(3):418–491, 1959.

[25] 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.

[26] E. Giné and A. Guillou. Rates of strong uniform consistency for multivariate kernel density estimators. Ann. Inst. H. Poincaré, 38:907–921, 2002.

[27] W. Hardle, B.U. Park, and A.B. Tsybakov. Estimation of non-sharp support boundaries. Journal of Multivariate Analysis, 55(2):205–218, 1995.

[28] Svante Janson. Maximal spacings in several dimensions. Annals of Probability, 15:274–280, 1987.

[29] S. Kallel and S. Louhichi. Topological reconstruction of compact supports of dependent stationary random variables. arXiv:2307.11674v1 [math.PR], 2023.

[30] S. Kallel, S. an Louhichi. Topological reconstruction of compact supports of dependent stationary random variables, `https://hal.science/hal-04366871`). Advances in Applied Probability, 4-2:1–31, 2024.

[31] A. Kolmogorov. On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval. Amer. Math. Soc. Transl., 1–2:233–243, 1949.

[32] E. Liebscher. Strong convergence of sums of α-mixing random variables with applications to density estimation. Stochast. Process. Appl., 65:69–80, 1996.

[33] B. L. S. Prakasa Rao. Nonparametric functional estimation. Academic press, New York, 1983.

[34] E. Rio. Théorie asymptotique des processus aléatoires faiblement dépendants. In Mathématiques et Applications, volume 31, Berlin, 2000. Springer.

[35] A. Rodríguez Casal. Set estimation under convexity type assumptions. Annales De l'Institut Henri Poincaré: Probabilités et Statistiques, 43:763–774, 2007.

[36] 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.

[37] A. Rodríguez-Casal and P. Saavedra-Nieves. Spatial distribution of invasive species: an extent of occurrence approach. TEST, 31:416–441, 2022.

[38] M. Rosenblatt. A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. USA, 42:43–47, 1956.

[39] G. G. Roussas. Asymptotical normality of positively or negatively associated processes. Journal of Mult. Analysis, 50:152–173, 1994.

[40] G. Walther. Granulometric smoothing. The Annals of Statistics, 25(6):2273 – 2299, 1997.

[41] 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 7. *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 $B(O,r) \cap \{z, \langle z.u_1 \rangle \geqslant a\}$ and we have

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

Thus by derivation

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

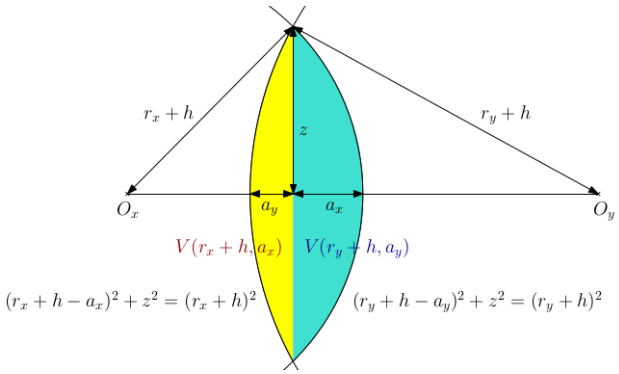

Fig. A.3: The convex body is $B(O_x, r_x + h) \cap B(O_y, r_y + h)$.

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.

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

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 $\Delta^+ + \Delta^- \leqslant \dfrac{4hr_y}{r_x + r_y}(1 + o(1))$) 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 8. *Suppose that balls of radius R_i roll inside S. Suppose that the density is such that $f(x) \geqslant f_0 \underline{d}(x, \partial S)^\alpha$. If O is such that $\underline{d}(O, S) = R - \varepsilon$ 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.4, 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 we have

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

Thus we obtain

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

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

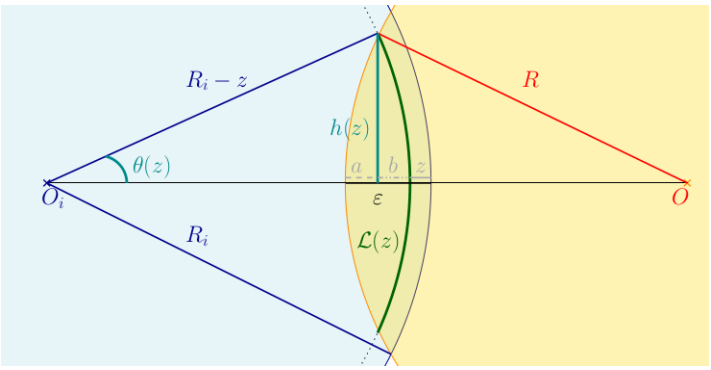

Fig. A.4: Integration

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 $\underline{\mathrm{d}}(O, \partial S) > R$, 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 $\underline{\mathrm{d}}(O, \partial S) < R$ 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 $\underline{\mathrm{d}}(O, \partial S) < R$ 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, $\mathbb{P}(B(O,R)) \geqslant f_0\omega_d\left(\frac{\min(R_i,R)}{2}\right)^{d+\alpha}$. □

Appendix B. *Proof of the Proposition 5 .*

We will first introduce the two preliminary lemmas. We set here $\varepsilon_n' = A_n + B_n$ with $A_n = \sup_{x\in\mathbb{R}^d}|\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))|$ and $B_n = \sup_{x\in\mathbb{R}^d}|\mathbb{E}(\widehat{f_n}(x)) - f(x)|$. Proposition 5 is a direct consequences of Propositions 9 and 10 below. Proposition

9 gives a control of B_n, while the two items of Proposition 10 give a control of A_n in the exponential and polynomial decay cases respectively.

Proposition 9. *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* (9)*, 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 10. *Let $\mathbb{X}_n$ be a sample from a strictly stationary sequence $\mathbf{X}$ and $\widehat{f_n}$ be defined by* (9)*.*

- *Assume that (K1), (M2), (M3) and (M4) 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 $\beta < 1/(3d+6)$.*

  (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

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

- *Assume that (K1), (M2), (M3), (M4) 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 $\gamma_0 = \dfrac{d+2-\beta d(d-2)}{q-\beta d(d-2)}$ 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 $\gamma > \gamma_0 = \dfrac{d+2-\beta d(d-2)}{q-\beta d(d-2)}$ 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 7. In the literature, classical results assume in place of (M1) and (K2) the following assumptions to control the bias term, here the proposition 3.3.1. in [19] is 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, $C^k \neq C_k$.

In the following, except in the statement of propositions and lemmas, we systematically set $h \equiv h_n$ (and h should be understood as h_n) for simplicity sake.

Proof of Proposition 5 under geometric decays. Let us prove (A1) (the case (T1) can be handled in the same way). By Proposition 9 and item 1 of Proposition 10,

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

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 under polynomial decays.

- Proof of item (A2) of Proposition 5. For any $b > \max((2d+3)/d, d+1)$, set $q = b+1-t$ for each $t \in (0, b-d-1)$ and then apply item (A2) of Proposition 10 and Proposition 9 to obtain for all $0 < \beta < 1/d$:

$$\varepsilon_n' = O_{a.s.}\left(\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\beta k} + \left(\frac{\ln n}{n^{1-\gamma}}\right)^{\frac{1-\beta d}{2}}\right), \quad \text{for all} \quad \gamma > \frac{d+2-\beta d(d-2)}{b+1-t-\beta d(d-2)},$$

  then choose $\beta = \frac{1}{d+2k}$ and let $t \to 0$ to end the proof.

- Proof of 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 10 and Proposition 9 to obtain, for all $0 < \beta < 1/d$

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

  then choose $\beta = \frac{1}{d+2k}$ allows to achieve the result.

□

Proof of Proposition 9.

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

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

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

Therefore,

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

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 10. We first start by providing in each case (the two items of the proposition), a bound for A_n on a compact set of $\mathbb{R}^d$ of the form

$$\sup_{\|x\|\leqslant n^\nu} \left|\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))\right| = O_{a.s.}\left(\frac{\lambda(n)}{\sqrt{nh^d}}\right), \tag{B.1}$$

with $\nu > 2/c$ (c defined in (M4)), 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

$$L_n = \sup_{\|x\|\geqslant n^\nu} \frac{\sqrt{nh^d}}{\lambda(n)} \left|\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))\right| \to 0 \tag{B.2}$$

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

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

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

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

which tends to zero by (M1) and (M2) by Remark 6. To prove that L_n converges a.s. to 0, we have to assume (M4). We thus have

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

for ν such that $\nu > 2/c$ so that the series converges by (M4) (quote that here $h \equiv h_n$ indeed depends on n). 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 $\|X_t - x\| > Ch$ a.s. uniformly over $1 \leq i \leq n$ and $\|x\| > n^\nu$ and thus the kernel vanishes.

Let us now prove (B.1). Set for all $x \in \mathbb{R}^d$,

$$\widehat{G_n}(x) = \sum_{i=1}^{n} Y_i(x) = \frac{1}{nh^d} \sum_{i=1}^{n} U_i(x), \quad \text{with} \quad U_i(x) = K\left(\frac{x - X_i}{h}\right) - \mathbb{E}K\left(\frac{x - X_i}{h}\right). \tag{B.3}$$

To prove (B.1), we use the chaining argument of [32]. 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), [32] proves that for some positive constant L and for all $x \in I_j$,

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

Using this fact it is easy to deduce that there exists another kernel density estimator $\widetilde{f}$ for f, obtained by replacing K in (9) 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 [32], page 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 $C_2 > 0$),

$$\begin{aligned}\mathbb{P}\left(\sup_{\|x\|\leqslant M^d} |\widehat{G}_n(x)| > t\right) &\leq \mathbb{P}\left(\max_{1\leqslant j\leqslant \delta^d} |\widehat{G}_n(x_j)| > \frac{t}{3}\right) + \mathbb{P}\left(\max_{1\leqslant j\leqslant \delta^d} |\widetilde{G}_n(x_j)| > \frac{th\delta}{3MC_1}\right)\\ &\leq \delta^d \sup_{\|x\|\leqslant M^d}\left\{\mathbb{P}\left(|\widehat{G}_n(x)| > \frac{t}{3}\right) + \mathbb{P}\left(|\widetilde{G}_n(x)| > \frac{th\delta}{3MC_1}\right)\right\} \qquad \text{(B.4)}\end{aligned}$$

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 (B.4). If the right hand side is the general term of a convergent series then using Borel-Cantelli lemma we obtain (B.1).

Therefore, the main goal of the proof of Proposition 10 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 (B.4) is the general term of a convergent series.
We will need the following

Lemma 5. *Let* $\mathbf{X}$ *be strictly stationary sequence with density* f *and kernel estimator* $\widehat{f_n}$ *defined by* (9)*. Assume that (M3) holds and that (K1) holds with* $h \to 0$ *and* $nh^d \to \infty$*, 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* $b > 2 + 2/d$.

Then:

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

Proof of item 1 in Proposition 10. In cases of exponential decays rates 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 (B.4). Namely we will use Theorem 1 and Proposition 8 of [23]. Let us recall them below in our special setting

Lemma 6 ([23], Proposition 8)**.** *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* $(s_1, \ldots, s_u)$ *and all* v*-tuples* $(t_1, \ldots, t_v)$ *with* $1 \leqslant s_1 \leqslant \cdots \leqslant s_u \leqslant t_1 \leqslant \cdots \leqslant t_v \leqslant n$,

$$\left|cov\left(Y_{s_1} \times \cdots \times Y_{s_u}, Y_{t_1} \times \cdots \times Y_{t_v}\right)\right| \leqslant H^{u+v-2}L^2 v c^{(t_1 - s_u)}. \qquad \text{(B.5)}$$

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 (B.3). 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 $\operatorname{Lip} F_n \leqslant \operatorname{Lip} K/(nh^{d+1})$. Assume that

$$\sigma_n^2 = O_{a.s.}\left(\frac{1}{nh^d}\right). \tag{B.6}$$

- *Strong mixing case.* If $\mathbf{X}$ is strong mixing with $\alpha(m) = O(b^{-m})$, for some $b > 1$, thus $\mathbf{Y}$ is strong mixing with $\alpha_Y(m) \leqslant \alpha(m)$ by measurability of F_n. Moreover, by [21]'s Lemma 6 and [23]'s Remark 7, (B.5) is satisfied with $L^2 = (\operatorname{Lip} F_n)^2$ and $c = 1/b$. Therefore Lemma 6 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 $h = O((\ln n/n)^\beta)$.

- *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 [23], (B.5) is satisfied with $L^2 = H \operatorname{Lip} F$ and $c = 1/b$. Therefore Lemma 6 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 $h = O((\ln n/n)^\beta)$.

Therefore, we obtain in both cases a bound for the first term at the right hand side of (B.4). 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 (B.4) 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

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

using Borel–Cantelli. It remains to show that (B.6) holds, which is true by Lemma 5 under (K1), (M2) and (M3).

Proof of Proposition 10, item 2. Our proof relies on the following lemma.

Lemma 7. *Let $\mathbb{X}_n$ be a n-sample of a strictly stationary sequence $\mathbf{X}$ and $\widehat{f_n}$ be defined by* (9)*.*

(i) *Assume (K1) with $h \to 0$ and $nh^d \to \infty$. Assume moreover that there exists some $c > 0$ and some $q \geqslant 2$ such that for $n > n_0$*

$$\sup_{x\in\mathbb{R}^d} \|\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))\|_q \leqslant \frac{c}{\sqrt{nh^d}}. \tag{B.7}$$

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

$$\sum_{n=1}^{\infty} \left(\frac{\sqrt{nh^d}}{h}\right)^d \frac{n^\nu}{\lambda_n^q} < \infty, \tag{B.8}$$

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

(ii) *Assume that* (B.7) *holds and* $h = O\left(\left(\frac{\ln n}{n^{1-\gamma}}\right)^{\beta}\right)$ *for some* $0 < \beta < 1/d$ *and* $0 < \gamma < 1$. *Then, under (K1) one has*

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

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

$$\gamma q + (1-\gamma)\beta d(d-2) > d + 2. \tag{B.9}$$

In the following lemma, we give the conditions under which (B.7) holds in arithmetic decay cases (A2) and (T2). The point (ii) of this lemma together with (7) yields item 2 of Proposition 10.

Lemma 8. *Let* $\mathbb{X}_n$ *be a n-sample of a strictly stationary sequence* $\mathbf{X}$ *and* $\widehat{f_n}$ *be defined by* (9).

(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* $h \to 0$ *and* $nh^d \to \infty$. *Thus* (B.7) *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* $h \to 0$ *and* $nh^d \to \infty$. *Thus* (B.7) *obtains for every even* $q \geqslant 2$ *and* $b > \max(2 + 2/d, 3(q-1))$.

□

Proof of Lemma 5. Let $\check{f}_n$ be the kernel estimator analogue to $\widehat{f_n}$ but built on independent copies of X_0. A well known result (see for instance [33]) is

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

We have

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

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

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

• On the one hand,

$$|\operatorname{cov}(U_0(x), U_m(x))| = h^{2d} \iint (K(s) - \mathbb{E}(K(X_0)))\,(K(t) - \mathbb{E}(K(X_0)))\, f_{0,m}(x - hs, x - ht)dsdt \leqslant C_o h^{2d}, \tag{B.10}$$

by (K1) and (M3).

• On the other hand one has the property (B.5) 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 (B.5)

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

Therefore with C_o defined through (B.10),

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

for all $0 \leqslant \alpha \leqslant 1$ by relation (12.18) of [19] so that setting $\alpha = d/(2d+1)$ Δ_n tends to zero as soon as

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

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

- If $\mathbf{X}$ is strongly mixing, we have by (B.5)

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

so

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

for all $0 \leqslant \alpha \leqslant 1$ so that with $\alpha = d/(2d+2)$, we obtain 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 exponential decay rate $\theta(m) \sim b^{-m}$, for some $b > 1$ this is always true.

□

Proof of Lemma 7.

- *Proof of (i).* Let $\lambda(n) \to \infty$ such that $\lambda(n) = o(\sqrt{nh^d})$. We deduce Markov inequality in (B.7) that there exists some $C_3 > 0$ such that

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

as soon as there is some $C_4 > 0$ such that $\delta h \geqslant C_4 M_n \sqrt{nh^d}$. For that task, set $t = \dfrac{\lambda(n)}{\sqrt{nh^d}}$ in (B.4). One has

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

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

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

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

$$\delta = \frac{M\sqrt{nh^d}}{h}.$$

We have

$$\sup_{\|x\| \leqslant M_n^d} |\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))| = O_{a.s.}\left(\frac{\lambda(n)}{\sqrt{nh^d}}\right). \tag{B.11}$$

This holds from Borel-Cantelli lemma as soon as

$$\sum_n \left(\frac{M_n\sqrt{nh^d}}{h}\right)^d \frac{1}{\lambda(n)^q} < \infty. \tag{B.12}$$

The result of Proposition 7 is obtained by setting $M_n = n^{\nu/d}$.

- *Proof of (ii).*
With the conditions on h we have $h \to 0$ and $nh^d \to \infty$. 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 (B.12) converges. Set u_n the general term of (B.12). With the assumptions on h 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 8.

- *Proof of (A2).* In the strong mixing case we apply [34]'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, $\alpha^{-1}(u) = \sum_{n>1} \mathbf{1}_{u<\alpha(n)}$, *and* Q_x denotes the generalized inverse of the tail function $t \mapsto \mathbb{P}(|U_0(x)| > t)$ and

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

Using [34]'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

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

Using Lemma 5 and setting $h = O((\ln n/n^{1-\gamma})^\beta)$, we obtain (B.7) under (K2) for some $k > 1$ as soon as (M2) and the following conditions are fulfilled for some $r > q$:

(1) $\dfrac{h^{dq(1/r-1/2)}}{n^{q/2-1}} \to 0$, as $n \to \infty$.
(2) $\displaystyle\sum_{i\geqslant 0} (i+1)^{\frac{qr-2r+q}{r-q}} \alpha(i) < \infty$.
(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, (B.7) obtains as soon as $b > \max(q-1, 2+3/d)$.

- *Proof of (T2).*

  In the θ-weakly dependent case, it straightforwardly follows from [22]'s theorem 1: under the conditions (M2), (M3), (K1) with $h \to 0$ and $nh^d \to \infty$ (B.7) obtains for any even $q \geqslant 2$ as soon as

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

  When $\theta(m) = o(m^{-b})$, the first condition implies $b > 2 + 2/d$ by Lemma 5 and the second condition is obtained with $b > 3(q-1)$.

□

Appendix C. Explicit rates

In this final section we explicit the rates for level set estimation in the case of polynomial dependence decay rates. We show which of the two components involved $\varepsilon_n, \varepsilon'_n$ is the convergence rate. For dimensions $d \leqslant 3$ the convergence rate reduces to the density estimation rate ε'_n whatever is the regularity of the density. For larger dimensions, the rate is the set estimation rate, ε_n, if density is regular enough and under convenient dependence assumptions.
More precisely, we obtain the rates:

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