Sharp adaptive estimation for d-dimensional classes

Karine Bertin

Sharp adaptive estimation in sup-norm for d-dimensional Hölder classes

KARINE BERTIN

Departamento de Estadística, Universidad de Valparaíso, Casilla 5030, Avenida Gran Bretaña 1091, Valparaíso, Chile. Tel: 0056322508324 . Email: karine.bertin@uv.cl

Abstract: In this paper, our aim is to estimate an anisotropic function in the framework of the Gaussian white noise model. We suppose that the function belongs to an unknown d-dimensional Hölder class. We evaluate the performance of an estimator by the sup-norm risk. We find the exact asymptotics of the adaptive minimax risk and construct asymptotically adaptive estimators.

Key Words: anisotropic Hölder class, minimax exact constant, uniform norm, white noise Model.

1 Introduction

Minimax adaptive estimation of a non-parametric function f from noisy data is the subject of many papers. Assuming that f belongs to a smoothness class Σ_β, where β is an unknown smoothness parameter, the aim is to find an estimator of f independent of β and which attains asymptotically optimal behaviour on all the classes Σ_β, for β given in a known set B. In this article, we study the problem of adaptive estimation in the Gaussian white noise model in sup-norm, assuming that the function f satisfies a Hölder condition. We observe $\left\{Y_t, t \in \mathbb{R}^d\right\}$, where Y_t is a random process defined by the stochastic differential equation

$$dY_t = f(t)dt + \frac{\sigma}{\sqrt{n}}dW_t, \quad t \in \mathbb{R}^d, \tag{1}$$

where f is an unknown function, $n \in \mathbb{N}$, $\sigma > 0$ is known and W is a standard Brownian sheet on $[0,1]^d$. We want to estimate the function f on $\mathbb{R}^d$ given a realization $y = \left\{Y_t, t \in \mathbb{R}^d\right\}$. We suppose that f belongs to a d-dimensional anisotropic Hölder class $\Sigma(\beta, L)$ where $L = (L_1, \dots, L_d) \in (0, +\infty)^d$ is known, $\beta = (\beta_1, \dots, \beta_d) \in (0,1]^d$ is unknown and belongs to a finite set $B \subset (0,1]^d$ known. The class $\Sigma(\beta, L)$ is defined by

$$\Sigma(\beta, L) = \left\{ f : \mathbb{R}^d \to \mathbb{R} : |f(x) - f(y)| \le L_1|x_1 - y_1|^{\beta_1} + \cdots + L_d|x_d - y_d|^{\beta_d}, \quad x, y \in \mathbb{R}^d \right\},$$

where $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$. In the following, $\mathbb{P}_f$ is the distribution of y under the model (1) and $\mathbb{E}_f$ is the corresponding expectation. An estimator θ_n of f is a random function on $[0,1]^d$ taking its values in $\mathbb{R}$, measurable with respect to y. We will evaluate the quality of an estimator θ_n by the maximal risk in sup-norm on B

$$R_n(\theta_n) = \sup_{\beta \in B} \sup_{f \in \Sigma(\beta, L)} \mathbb{E}_f \left\{ \left(\frac{\|\theta_n - f\|_\infty}{\psi_n(\beta)} \right)^p \right\}, \tag{2}$$

where $\|g\|_\infty = \sup_{t\in[0,1]^d} |g(t)|$, $p > 0$ and $\psi_n(\beta)$ has the form $\psi_n(\beta) = C_\beta \left(\frac{\log n}{n}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}}$ for $\beta = (\beta_1, \dots, \beta_d) \in B$, $\overline{\beta} = \left(\sum_{i=1}^d 1/\beta_i\right)^{-1}$, C_β being a constant depending on β. In the non-adaptive case, i.e. when B contains only one vector, it has been proved that $\psi_n(\beta)$ is the minimax rate of convergence for sup-norm estimation: for $d = 1$, it was done by Ibragimov and Hasminskii (1981); for multidimensional case, this fact was shown by Stone (1985) and Nussbaum (1986) for isotropic setting ($\beta_1 = \cdots = \beta_d$) and in Bertin (2004) for anisotropic setting considered here. In other set-ups, some results on rate of convergence can be found in Barron et al. (1999), Neumann and von Sachs (1997) and Kerkyacharian et al. (2001). In an adaptive set-up, Lepski (1992) proved that $\psi_n(\beta)$ is the adaptive rate of convergence (cf. Tsybakov (1998) for precise definition of adaptive rate of convergence) for the problem considered here when $d = 1$.

Our goal is to study the asymptotics of the minimax risk in sup-norm on B (i.e. the adaptive minimax risk), in others words to study the asymptotics of

$$\inf_{\theta_n} R_n(\theta_n).$$

We want to prove that there exist optimal rate adaptive estimators on the scale of classes $\{\Sigma(\beta, L)\}_{\beta\in B}$ for the L_∞ norm and to find an estimator $\widetilde{f}_n$ and the constant C_β with $\psi_n(\beta) = C_\beta \left(\frac{\log n}{n}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}}$, such that we have

$$\lim_{n\to\infty} \inf_{\theta_n} R_n(\theta_n) = \lim_{n\to\infty} R_n(\widetilde{f}_n) = 1, \tag{3}$$

where $\inf_{\theta_n}$ stands for the infimum over all the estimators. To obtain that there exist optimal rate adaptive estimators, it is enough to have that there exist an estimator $\tilde{\theta}_n$ and a positive constant C such that

$$\limsup_{n\to\infty} R_n(\tilde{\theta}_n) \leq C.$$

An estimator $\widetilde{f}_n$ that satisfies (3) is called *asymptotically exact adaptive estimator on the scale of classes* $\{\Sigma(\beta, L)\}_{\beta\in B}$ *for the* L_∞ *norm*, and C_β is called exact adaptive constant.

In the non-adaptive case (cf. Bertin (2004)), the relation (3) is satisfied by a constant $C_0(\beta)$ which depends on β, L and σ^2, and for $\hat{f}_\beta$ a kernel estimator with kernel close to the kernel K_β. The kernel K_β is defined for $u = (u_1, \dots, u_d) \in \mathbb{R}^d$ by

$$K_\beta(u_1, \dots, u_d) = \frac{\overline{\beta}+1}{\alpha(\beta)\overline{\beta}^2}(1 - \sum_{i=1}^d |u_i|^{\beta_i})_+,$$

with

$$\alpha(\beta) = \frac{2^d \prod_{i=1}^d \Gamma(\frac{1}{\beta_i})}{\Gamma(\frac{1}{\overline{\beta}}) \prod_{i=1}^d \beta_i},$$

Γ denotes the gamma function and $(x)_+ = \max(0, x)$. The constant $C_0(\beta)$ satisfies

$$C_0(\beta) = \left(\sigma^{2\overline{\beta}} L_*(\beta) \left(\frac{\overline{\beta}+1}{\alpha(\beta)\overline{\beta}^3}\right)^{\overline{\beta}}\right)^{\frac{1}{2\overline{\beta}+1}}$$

and

$$L_*(\beta) = \left(\prod_{i=1}^{d} L_j^{1/\beta_j} \right)^{\overline{\beta}}.$$

In this article, we prove that there exist optimal rate adaptive estimators on $\{\Sigma(\beta, L)\}_{\beta \in B}$ for the L_∞ norm. We give precise upper and lower bounds for any finite set B. When B contains only two vectors, we improve the upper bound.

For particular forms of the set B (including sets B of isotropic classes), we prove it exist an estimator $\tilde{f}_n$ and a constant C_β satisfying (3). The constant C_β we have found is equal to the constant $C_0(\beta)$(which was the solution of (3) in Bertin (2004)) multiplied by a constant larger than 1, depending on the set B and p. As a consequence, for this case, we lose efficiency in the constant under adaptation. We will see that the estimator satisfying (3) is obtained using the Lepski method. This method is introduced in Lepski (1992) for the problem considered here with $d = 1$ and it has been studied by many authors. In particular, for $d > 1$, Klemelä and Tsybakov (2001) have used it for the estimation of linear functionals in isotropic settings. Tsybakov (1998) used this method to study pointwise and sup-norm estimation on Sobolev classes.

For more general forms of the set B, we construct optimal rate adaptive estimators using o generalization of the Lepski method proposed in Lepski and Levit (1998).

2 Main results

In this section, we introduce the notation and the assumptions about the set B, we define three families of estimators and we give our results.

2.1 The set B

We suppose that the set $B = \left\{\beta^{(1)}, \dots, \beta^{(l)}\right\}$ contains l vectors belonging to $(0, 1]^d$. The coordinates of $\beta^{(i)}$ are denoted by $\beta_j^{(i)}$, $j = 1, \dots, d$:

$$\beta^{(i)} = (\beta_1^{(i)}, \dots, \beta_d^{(i)}) \in (0, 1]^d. \tag{4}$$

We define the real $\overline{\beta}^{(i)}$ by

$$\overline{\beta}^{(i)} = \left(\sum_{j=1}^{d} \frac{1}{\beta_j^{(i)}} \right)^{-1} \tag{5}$$

and we denote by $\overline{B}$ the set:

$$\overline{B} = \left\{\overline{\beta}^{(1)}, \dots, \overline{\beta}^{(l)}\right\}.$$

We suppose that $\overline{\beta}^{(i)} \neq \overline{\beta}^{(j)}$ for all $i, j \in \{1, \dots, l\}$ such that $i \neq j$. As a consequence, a $\overline{\beta} \in \overline{B}$ is matched to a unique $\beta \in B$ via the relations (4) and (5). We define the following relation of order in B: the vectors β and γ satisfy

$$\beta \leq \gamma \quad \text{if and only if} \quad \overline{\beta} \leq \overline{\gamma}.$$

This is a relation of total order in B, thus the notion of maximum and minimum in B are well-defined. We denote $\beta_{max} = \max\{\beta^{(i)}, i = 1, \cdots, l\}$ (respectively $\beta_{min} = \min\{\beta^{(i)}, i = 1, \cdots, l\}$) and $\overline{\beta}_{max}$ (respectively $\overline{\beta}_{min}$) the associated real number in $\overline{B}$ via (5).

Remark. In the isotropic setting (i.e. for all $\beta = (\beta_1, \dots, \beta_d) \in B$, $\beta_1 = \dots = \beta_d$), this order is the order $\beta = (\beta_1, \dots, \beta_d) \leq \gamma = (\gamma_1, \dots, \gamma_d)$ if and only if $\beta_1 \leq \gamma_1$.

2.2 Three families of estimators

Here we define three families of kernel estimators $(\hat{f}_{\beta,1})_{\beta \in B}$, $(\hat{f}_{\beta,2})_{\beta \in B}$ and $(\hat{f}_{\beta,3})_{\beta \in B}$. They are close to the asymptotically exact estimator of Bertin (2004), but the kernel is somewhat different at the boundary. This is a consequence of the fact that the observations here are for $t \in \mathbb{R}^d$ whereas they were for $t \in [0,1]^d$ in Bertin (2004). The estimators are defined in the following way. For $\beta \in B$ and $j \in \{1,2,3\}$,

$$\hat{f}_{\beta,j}(t) = \int_{\mathbb{R}^d} \overline{K}_{\beta,j}\,(t-u)\,dY_u,$$

defined for $t = (t_1, \dots, t_d) \in [0,1]^d$, where for $u = (u_1, \dots, u_d) \in \mathbb{R}^d$

$$\overline{K}_{\beta,j}\,(u) = \frac{1}{h_{1,j}(\beta) \cdots h_{d,j}(\beta)} K_\beta \left(\frac{u_1}{h_{1,j}(\beta)}, \dots, \frac{u_d}{h_{d,j}(\beta)} \right).$$

For $j \in \{1,2,3\}$, the bandwidth $h_j = (h_{1,j}(\beta), \dots, h_{d,j}(\beta))$ satisfies for $i \in \{1, \dots, d\}$

$$h_{i,j}(\beta) = \left(\frac{C_\beta \lambda_j(\beta)}{L_i} \left(\frac{\log n}{n} \right)^{\overline{\beta}/(2\overline{\beta}+1)} \right)^{1/\beta_i},$$

where

$$\lambda_j(\beta) = \begin{cases} 1 & \text{for } j = 1, \\ 2^{-\frac{2\overline{\beta}}{2\overline{\beta}+1}} \left(\frac{c_2(\beta)}{c_1(\beta)} \right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}} & \text{for } j = 2, \\ 2^{-\frac{2\overline{\beta}}{2\overline{\beta}+1}} & \text{for } j = 3, \end{cases}$$

$$C_\beta = \left(\sigma^{2\overline{\beta}} L_*(\beta) \left(\frac{c_1(\beta)(\overline{\beta}+1)}{\alpha(\beta)\overline{\beta}^3} \right)^{\overline{\beta}} \right)^{\frac{1}{2\overline{\beta}+1}}, \tag{6}$$

$$c_1(\beta) = 1 + \frac{p(\overline{\beta}_{max} - \overline{\beta})}{2\overline{\beta}_{max} + 1},$$

$$c_2(\beta) = 1 + p\mu(\beta)(2\overline{\beta}+1)$$

and

$$\mu(\beta) = \max_{\lambda > \beta} \left(\frac{\overline{\lambda}}{2\overline{\lambda}+1} - \min_{\gamma < \lambda, i=1,\dots,d} \left\{ \frac{\overline{\gamma}}{2\overline{\gamma}+1}, \frac{\lambda_i \overline{\gamma}}{\gamma_i(2\overline{\gamma}+1)} \right\} \right),$$

where the maximum is taken for $\lambda = (\lambda_1, \dots, \lambda_d) \in B$, and the minimum for $\gamma = (\gamma_1, \dots, \gamma_d) \in B$.

Lemma 1. *The kernel K_β satisfies $\int_{[-1,1]^d} K_\beta(u)du = 1$ and $\int_{[-1,1]^d} K_\beta^2(u)du = \frac{2(\overline{\beta}+1)}{\overline{\beta}\alpha(\beta)(2\overline{\beta}+1)}$.*

This a consequence of formula of Gradshteyn and Ryzhik (1965) (cf Appendix of Bertin (2004)) For $t \in [0,1]^d$, $j \in \{1,2,3\}$ and for a function f belonging to one of the classes $\Sigma(\beta, L)$, define the bias of $\hat{f}_{\beta,j}$

$$b_{\beta,j}(t,f) = \mathbb{E}_f(\hat{f}_{\beta,j}(t)) - f(t) = \overline{K}_{\beta,j} * f(t) - f(t)$$

and the stochastic term of the error

$$Z_{\beta,j}(t) = \hat{f}_{\beta,j}(t) - \mathbb{E}_f(\hat{f}_{\beta,j}(t)) = \frac{\sigma}{\sqrt{n}} \int \overline{K}_{\beta,j}(t-u) dW_u.$$

The symbol $*$ stands here the convolution of functions defined on $\mathbb{R}^d$. The quantity $\overline{K}_{\beta,j} * f$ is well-defined since $\overline{K}_{\beta,j}$ is continuous with compact support and f is continuous.

2.3 A lower bound for anisotropic classes

In the following, we will use the maximal risk in sup-norm on B defined in (2) with $\psi_n(\beta) = C_\beta \left(\frac{\log n}{n}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}}$, C_β defined in (6) and $p > 0$. We have the following lower bound which is valid for any finite set B satisfying the conditions of Subsection 2.1.

Theorem 1. *The minimax risk in sup-norm on B satisfies*

$$\liminf_{n\to\infty} \inf_{\theta_n} R_n(\theta_n) \geq 1.$$

2.4 Exact asymptotics for particular forms of the set B

For $\beta \in B$ and $j \in \{1,2,3\}$, we consider

$$\eta_j(\beta) = \left(\frac{2\|K_\beta\|_2^2 \sigma^2 c_j(\beta)}{n \prod_{i=1}^d h_{i,j}(\beta)(2\overline{\beta}+1)} \log n \right)^{1/2},$$

where $c_3(\beta) = c_1(\beta)$. We select the vector $\hat{\beta}_{(p)} \in B$ defined by

$$\hat{\beta}_{(p)} = \max\left\{\beta \in B : \quad \forall \gamma < \beta, \quad \|\hat{f}_{\beta,1} - \hat{f}_{\gamma,1}\|_\infty \leq \eta_1(\gamma)\right\}$$

We have the following theorem for particular forms of the set B, which include the sets B of isotropic classes.

Theorem 2. *We suppose that the set B satisfies the property (P):*

(P) *For all $\beta = (\beta_1, \ldots, \beta_d), \gamma = (\gamma_1, \ldots, \gamma_d) \in B$, if $\beta \leq \gamma$ then for all $i = 1, \ldots, d$ $\beta_i < \gamma_i$.*

Then, the estimator $\widetilde{f}_{(p)} = \hat{f}_{\hat{\beta}_{(p)},1}$ is asymptotically exact adaptive, i.e. it satisfies the condition (3) with the constant C_β defined in (6):

$$\lim_{n\to\infty} \inf_{\theta_n} R_n(\theta_n) = \lim_{n\to\infty} R_n(\widetilde{f}_{(p)}) = 1$$

Remark. The estimator $\widetilde{f}_{(p)}$ is obtained using the method of Lepski. The constant C_β has the form of the constant obtained by Lepski (1992) in the case $d = 1$ where we replace β by $\overline{\beta}$.

2.5 Upper bounds for anisotropic classes

If the set B does not satisfy Condition (P), Theorem 2 is not true anymore. In this subsection, Theorem 3 gives an upper bound for any finite set $B \subset (0,1]^d$ and Theorem 4 improves this upper bound when B contains only two vectors.

We consider new estimators defined for β and $\gamma \in B$ and $t \in [0,1]^d$ by

$$\hat{f}_{\beta*\gamma,2}(t) = \int K_{\beta*\gamma,2}(t-u)dY_u$$

where

$$K_{\beta*\gamma,2} = \overline{K}_{\beta,2} * \overline{K}_{\gamma,2}.$$

We consider the vector $\hat{\beta}_{ani} \in B$ defined by

$$\hat{\beta}_{ani} = \max\left\{\beta \in B : \quad \forall \gamma < \beta, \quad \|\hat{f}_{\beta*\gamma,2} - \hat{f}_{\gamma,2}\|_\infty \le \eta_2(\gamma)\right\}.$$

Theorem 3. *The estimator $\widetilde{f}_{ani} = \hat{f}_{\hat{\beta}_{ani},2}$ satisfies for $p > 0$*

$$\limsup_{n\to\infty} \sup_{\beta \in B} \sup_{f \in \Sigma(\beta,L)} \mathbb{E}_f \left\{\|\widetilde{f}_{ani} - f\|_\infty^p \left(\psi_n(\beta) M_2(\beta)\right)^{-p}\right\} \le 1, \tag{7}$$

where, for $j \in \{2,3\}$ and $\beta \in B$,

$$M_j(\beta) = \left(\frac{c_j(\beta)}{c_1(\beta)}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}} \left(\frac{2^{1/(2\overline{\beta}+1)} + \overline{\beta} 2^{2\overline{\beta}/(2\overline{\beta}+1)}}{2\overline{\beta}+1}\right).$$

The relation (7) imply that $\widetilde{f}_{ani}$ is an optimal rate adaptive estimator on $\{\Sigma(\beta,L)\}_{\beta\in B}$ for the L_∞ norm, for any finite set $B \subset (0,1]^d$.

For $l = 2$, we have a better result. We suppose that $B = \{\gamma, \beta\}$ with $\gamma < \beta$. We consider a new estimator defined for $t \in [0,1]^d$ by

$$\hat{f}_{\beta*\gamma,1}(t) = \int K_{\beta*\gamma,1}(t-u)dY_u$$

where

$$K_{\beta*\gamma,1} = \overline{K}_{\beta,1} * \overline{K}_{\gamma,3}.$$

We select the estimator $\widetilde{f}_{ani2}$ defined by

$$\widetilde{f}_{ani2} = \begin{cases} \hat{f}_{\beta,1} & \text{if } \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\gamma,3}\|_\infty \le \eta_3(\gamma) \text{ or } \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\beta,1}\|_\infty \ge \frac{\lambda_3(\gamma)\psi_n(\gamma)}{2\overline{\gamma}+1}(1+\rho_n) \\ \hat{f}_{\gamma,3} & \text{otherwise,} \end{cases}$$

where $\rho_n = \frac{\psi_n((\beta+\gamma)/2)}{\psi_n(\gamma)}$ and $\psi_n(\lambda)$, for $\lambda = (\lambda_1, \dots, \lambda_d) \in (0,1]^d \cap B^c$, is defined by $\psi_n(\lambda) = \left(\frac{\log n}{n}\right)^{\frac{\overline{\lambda}}{2\overline{\lambda}+1}}$ with $\overline{\lambda} = \left(\sum_{i=1}^d 1/\lambda_i\right)^{-1}$. Here B^c denotes the complement of the set B. Then we have the following theorem

Theorem 4. *The estimator $\widetilde{f}_{ani2}$ satisfies for $p > 0$*

$$\limsup_{n\to\infty} \sup_{f \in \Sigma(\beta,L)} \mathbb{E}_f \left\{\|\widetilde{f}_{ani2} - f\|_\infty^p \psi_n^{-p}(\beta)\right\} \le 1 \tag{8}$$

and

$$\limsup_{n\to\infty} \sup_{f \in \Sigma(\gamma,L)} \mathbb{E}_f \left\{\|\widetilde{f}_{ani2} - f\|_\infty^p \psi_n^{-p}(\gamma)\right\} \le (M_3(\gamma))^p. \tag{9}$$

2.6 Some remarks

1. In the model here, we choose to have observations in $\mathbb{R}^d$ to simplify the calculations. To obtain the results, it is enough to have observations in a neighborhood in $[0,1]^d$.

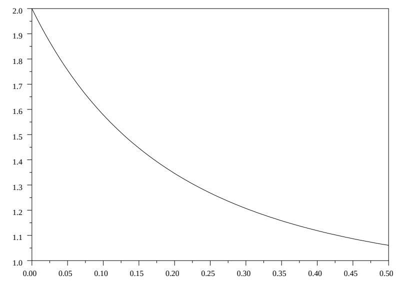
2. These results can be generalized to others models such regression with regular design. This can be done by following the same proofs with minor modifications.

3. In this article, we have considered a finite set B such that its cardinality, $\mathrm{card}(B)$, does not depend on n. We could have considered a set B such that $\mathrm{card}(B)$ grows as $n \to \infty$ and there exists $M > 0$ such that, for all $\beta \in B$, $\overline{\beta} > M$. In Theorems 2 and 4, following the same proofs, it is possible to have the same results taking B such $\mathrm{card}(B) \leq (\log n)^{\frac{1}{2(2\beta_{max}+1)}}$ (The last condition is necessary to have $\lim_{n\to\infty} R_1(\beta) = 0$ in Section 5 and the relation (49)). A weaker condition is possible on B. If, for $j \in \{1,2,3\}$, we replace $c_j(\beta)$ by $c_j(\beta) + \frac{1}{\log\log n}$ in $\eta_j(\beta)$, the results of Theorem 2, Theorem 3 and Theorem 4 can be obtained for B such that $\mathrm{card}(B) \leq (\log n)^a$ with $a > 0$.

4. For $d = 1$, in the Lepski method, for a given family of estimators $(\hat{f}_\beta)_{\beta\in B}$, we choose $\hat{\beta}$, the largest $\beta \in B \subset \mathbb{R}_+$ such that $\|\hat{f}_\beta - \hat{f}_\gamma\|_\infty \leq c\psi_n(\gamma)$ for all $\gamma \leq \beta$, with a constant $c > 0$. This choice is based on the fact that, if $f \in \Sigma(\beta_v, L)$ and $\gamma \leq \beta \leq \beta_v$, the bias of $\hat{f}_\beta - \hat{f}_\gamma$ is upper-bounded by a term of order $\psi_n(\gamma)$. This property is still valid for anisotropic settings when B satisfies Condition (P), but it does not work for general anisotropic settings. Indeed, for anisotropic settings, we do not have such a property on the bias of $\hat{f}_\beta - \hat{f}_\gamma$ and this bias can be very large. Kerkyacharian et al. (2001) give a new criteria which permits to obtain results in anisotropic Besov classes. Rather than comparing $\hat{f}_\beta$ to $\hat{f}_\gamma$ for $\gamma \leq \beta$, they compare, for $\gamma \leq \beta$, $\hat{f}_\gamma$ to a kernel estimator $\hat{f}_{\beta\gamma}$ with bandwidth $(h_1(\beta,\gamma), \ldots, h_d(\beta,\gamma))$ where $h_i(\beta,\gamma) = max(h_i(\beta), h_i(\gamma))$ and $(h_1(\beta), \ldots, h_d(\beta))$, respectively $(h_1(\gamma), \ldots, h_d(\gamma))$, is the bandwidth of $\hat{f}_\beta$, respectively $\hat{f}_\gamma$. This comparison permits to have a new criteria for selecting $\hat{\beta}$. In Theorem 3 and Theorem 4, we use another estimator $\hat{f}_{\beta*\gamma,j}$ to do this kind of comparison. The different choice of $\hat{f}_{\beta*\gamma,j}$ is motivated by our goal to have better constants.

5. Our results are adaptation with respect to β and we suppose that L is fixed and known. To use our method of estimation, the statistician need to know L. Here we suppose the statistician does not know L. We suppose that β belongs to B and $L \in \{L^{(1)}, \ldots, L^{(l)}\} \subset]0, +\infty[^d$. We note for $j = 1, \ldots, l$, $L^{(j)} = (L_1^{(j)}, \ldots, L_d^{(j)})$. Our method can be applied to select $\beta \in B$ with $\tilde{L} = (\max_{j=1,\ldots,d} L_1^{(j)}, \ldots, \max_{j=1,\ldots,d} L_d^{(j)})$.

6. Theorem 2 gives a result of exact estimation. In Theorem 3 and in Theorem 4 in the case $f \in \Sigma(\gamma, L)$, the difference between the lower bound and the upper bound is the factor $(M_2(\beta))^p$ and $(M_3(\gamma))^p$ respectively. The following plot represents the quantity $M_3(\beta)$(on the vertical axis) as a function of $\beta \in (0, 1/2]$.



We can see that, for $\beta \in (0, 1/2]$, $1.06 \leq M_3(\beta) \leq 2$ and then it implies that

$$1.06 \leq 1.06 \left(\frac{c_2(\beta)}{c_1(\beta)} \right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}} \leq M_2(\beta) \leq 2 \left(\frac{c_2(\beta)}{c_1(\beta)} \right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}}.$$

The following sections are devoted to preliminary results and to the proofs of the theorems.

3 Some preliminary results

In the following, D_i, with $i = 1, 2, \ldots$, denote positive constants, except otherwise mentioned. These constants can depend on $\beta \in B$ but we do not indicate explicitly the dependence on β. This does not have consequences on the proofs since B is a finite set. The quantity $\eta_j(\beta)$ satisfies the following lemma which will be proved in Section 8.

Lemma 2. *For $\beta \in B$ and $j \in \{1, 2\}$, we have*

$$\eta_j(\beta) + \frac{j\psi_n(\beta)\lambda_j(\beta)}{2\overline{\beta} + 1} = M_j(\beta)\psi_n(\beta),$$

where $M_1(\beta) = 1$. Moreover, for $\beta \in B$, we have

$$\eta_3(\beta) + \frac{2\psi_n(\beta)\lambda_3(\beta)}{2\overline{\beta} + 1} = M_3(\beta)\psi_n(\beta).$$

We have the following results for the families of estimators $(\hat{f}_{\beta,1})_{\beta \in B}$, $(\hat{f}_{\beta,2})_{\beta \in B}$ and $(\hat{f}_{\beta,3})_{\beta \in B}$.

Proposition 1. *For $\beta \in B$ and $j \in \{1, 2, 3\}$, we have*

$$\sup_{f \in \Sigma(\beta, L)} \|b_{\beta,j}(\cdot, f)\|_\infty \leq \frac{\psi_n(\beta)\lambda_j(\beta)}{2\overline{\beta} + 1}.$$

Proposition 2. *For $\beta \in B$, $j \in \{1, 2, 3\}$ and $t \in [0, 1]^d$, we have*

$$\mathbb{E}_f \left(Z^2_{\beta,j}(t) \right) \leq \frac{\sigma^2 \|K_\beta\|_2^2}{n\tilde{h}_j(\beta)},$$

and for $b \geq \dfrac{D_0}{\sqrt{n\tilde{h}_j(\beta)|\log(\tilde{h}_j(\beta))|}}$, *we have*

$$\mathbb{P}_f\left[\|Z_{\beta,j}\|_\infty \geq b\right] \leq \frac{D_1}{\tilde{h}_j(\beta)} \exp\left\{-\frac{b^2 n\tilde{h}_j(\beta)}{2\|K_\beta\|_2^2\sigma^2}\right\} \exp\left\{\frac{D_2(\beta) b\sqrt{n\tilde{h}_j(\beta)}}{(\log(\tilde{h}_j(\beta)))^{1/2}}\right\},$$

where $\tilde{h}_j(\beta) = \prod_{i=1}^d h_{i,j}(\beta)$.

Proposition 3. *For* $p > 0$, $\beta \in B$ *and* $j \in \{1, 2, 3\}$, *we have*

$$\limsup_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f\left[\left(\|\hat{f}_{\beta,j} - f\|_\infty (M_j(\beta)\psi_n(\beta))^{-1}\right)^p\right] \leq 1 \tag{10}$$

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f\left[\left(\|\hat{f}_{\beta,j} - f\|_\infty \psi_n^{-1}(\beta)\right)^p I_{\{\|\hat{f}_{\beta,j}-f\|_\infty \geq (1+\varepsilon(n))M_j(\beta)\psi_n(\beta)\}}\right] = 0, \tag{11}$$

where $\varepsilon(n)$ *satisfies* $\varepsilon(n) \geq (\log n)^{-1/4}$, *and* I_A *denotes the indicator function of a set* A.

Proposition 1, respectively Proposition 2, can be obtained following the proof of Proposition 1, respectively Lemma 4, of Bertin (2004). The proof of Proposition 3 can be deduced from the proofs of Bertin (2004) and we add some elements of proof in Section 8.

Define for $t \in [0, 1]^d$, $j \in \{1, 2\}$ and a function f the bias term of $\hat{f}_{\beta*\gamma,j}$

$$b_{\beta*\gamma,j}(t, f) = \mathbb{E}_f(\hat{f}_{\beta*\gamma,j}(t)) - f(t),$$

and the stochastic term of $\hat{f}_{\beta*\gamma,j}$

$$Z_{\beta*\gamma,j}(t) = \hat{f}_{\beta*\gamma,j}(t) - \mathbb{E}_f(\hat{f}_{\beta*\gamma,j}(t)) = \frac{\sigma}{\sqrt{n}} \int K_{\beta*\gamma}(t - u) dW_u.$$

The estimator $\hat{f}_{\beta*\gamma,j}$ satisfies the two lemmas.

Lemma 3. *For* $\beta, \gamma \in B$, *we have*

$$\sup_{f\in\Sigma(\beta,L)} \|b_{\beta*\gamma,2}(\cdot, f) - b_{\gamma,2}(\cdot, f)\|_\infty \leq \sup_{f\in\Sigma(\beta,L)} \|b_{\beta,2}(\cdot, f)\|_\infty \leq \frac{\psi_n(\beta)\lambda_2(\beta)}{2\overline{\beta} + 1}.$$

When $B = \{\gamma, \beta\}$ *with* $\gamma < \beta$ *we have*

$$\sup_{f\in\Sigma(\beta,L)} \|b_{\beta*\gamma,1}(\cdot, f) - b_{\gamma,3}(\cdot, f)\|_\infty \leq \sup_{f\in\Sigma(\beta,L)} \|b_{\beta,1}(\cdot, f)\|_\infty \leq \frac{\psi_n(\beta)}{2\overline{\beta} + 1}$$

and

$$\sup_{f\in\Sigma(\gamma,L)} \|b_{\beta*\gamma,1}(\cdot, f) - b_{\beta,1}(\cdot, f)\|_\infty \leq \sup_{f\in\Sigma(\gamma,L)} \|b_{\gamma,3}(\cdot, f)\|_\infty \leq \frac{\psi_n(\gamma)\lambda_3(\gamma)}{2\overline{\gamma} + 1}.$$

Lemma 4. *For $\beta, \gamma \in B$ such that $\gamma < \beta$ and $j \in \{1,2\}$, we have for all $t \in [0,1]^d$*

$$\mathbb{E}_f \left((Z_{\beta*\gamma,j})^2 (t) \right) \leq \frac{\sigma^2 \|K_\beta\|_2^2}{n\tilde{h}_j(\beta)},$$

and for $b \geq \dfrac{D_3}{\sqrt{n\tilde{h}_j(\beta)|\log(\tilde{h}_j(\beta))|}}$

$$\mathbb{P}_f \left[\|Z_{\beta*\gamma,j}\|_\infty \geq b \right] \leq \frac{D_4}{\tilde{h}_j(\beta)} \exp\left\{ -\frac{b^2 n\tilde{h}_j(\beta)}{2\|K_\beta\|_2^2 \sigma^2} \right\} \exp\left\{ \frac{D_5 b \sqrt{n\tilde{h}_j(\beta)}}{(\log(\tilde{h}_j(\beta)))^{1/2}} \right\}. \quad (12)$$

Lemma 5. *Let $j \in \{1,2,3\}$. We have, for $\gamma, \beta \in B$ such $\gamma < \beta$,*

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|Z_{\gamma,j}\|_\infty^p \psi_n^{-p}(\beta) I_{\{\|Z_{\gamma,j}\|_\infty > \tau_{n,j}(\gamma)\}} \right\} = 0,$$

where

$$\tau_{n,j}(\gamma) = \left(\frac{2\|K_\gamma\|_2^2 \sigma^2 (1 + p\overline{\beta}_{max}) \log n}{n \prod_{i=1}^d h_{i,j}(\gamma)(2\overline{\gamma}+1)} \right)^{1/2}.$$

Remark: $\tau_{n,j}(\gamma)$ is of the same order of $\psi_n(\gamma)$.

These three lemmas will be proved in Section 8.

4 Proof of Theorem 1

The proof of the lower bound is similar to the proof of lower bounds in Tsybakov (1998)The difficulties consist in finding a good subclass of $\Sigma(\beta, L)$ and good parameters to apply Theorem 6 of Tsybakov (1998). We have

$$\Delta_n = \inf_{T_n} \sup_{\beta\in B} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|T_n - f\|_\infty^p \psi_n^{-p}(\beta) \right\} \geq \max_{\beta\in B} \{\Delta_n(\beta)\},$$

where

$$\Delta_n(\beta) = \inf_{T_n} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|T_n - f\|_\infty^p \psi_n^{-p}(\beta) \right\}.$$

By Theorem 1 of Bertin (2004), we know that $\liminf_{n\to\infty} \Delta_n(\beta_{max}) \geq 1$. This comes from the fact that the loss function $w(x) = x^p$ satisfies the conditions required in this theorem (i.e. w is a non-decreasing function which admits a polynomial majorant and such that $w(0) = 0$), and the exact constant $C_{\beta_{max}}$ in the rate of convergence is the same as the exact constant in the non-adaptive case. Now, in order to prove the theorem, it is enough to prove that $\liminf_{n\to\infty} \Delta_n(\beta) \geq 1$ for all $\beta \in B_{\setminus\beta_{max}}$.

Let $0 < \varepsilon < 1/2$. Let $\beta \in B_{\setminus\beta_{max}}$. In the following several quantities depend on β, but we do not indicate this dependence to simplify the notations. We consider the set of functions $f_{j,\beta}(\cdot)$ defined, for $j \in \{0, \cdots, M\}$, by

$$\begin{cases} f_{j,\beta} = \psi_n(\beta)(1-\varepsilon)\left(1 - \sum_{i=1}^d \left|\frac{t_i - a_{j,i}}{h_{i,1}(\beta)}\right|^{\beta_i}\right)_+, & j = 1, \ldots, M \\ f_{0,\beta} = 0, \end{cases}$$

where the $a_j = (a_{j,1}, \dots, a_{j,d})$ form a grid of points in $[0,1]^d$. This grid is defined in the following manner. For

$$m_i = \left[\frac{1}{2h_{i,1}(2^{1/\beta} + 1)} - 1 \right]$$

with $[x]$ the integer part of x and $M = \prod_{i=1}^d m_i$, we consider the points $a(l_1, \dots, l_d) \in [0,1]^d$ for $l_i \in \{1, \dots, m_i\}$ and $i \in \{1, \dots, d\}$, such that:

$$a(l_1, \dots, l_d) = 2(2^{\frac{1}{\beta}} + 1)\,(h_{1,1}l_1, \dots, h_{d,1}l_d)\,.$$

To simplify the notation, we denote these points $a_1, \dots, a_M$ and each a_j takes the form:

$$a_j = (a_{j,1}, \dots, a_{j,d}).$$

The functions $f_{j,\beta}$ satisfy the following lemma which can be proved as in Bertin (2004).

Lemma 6.
1- $f_{j,\beta} \in \Sigma(\beta, L)$,
2- $\|f_{j,\beta}\|_2^2 = \frac{\sigma^2 c_1(\beta) \log n (1-\varepsilon)^2}{n(2\beta+1)}$
3- the functions $f_{j,\beta}$ *have disjoint support.*

Here we come back to the study of Δ_n. Let $j \in \{1, \dots, M\}$ and T_n an estimator. We have, since $f_{j,\beta}(a_k) = (1-\varepsilon)\psi_n(\beta)\delta_{j,k}$,

$$\begin{aligned} \psi_n^{-1}(\beta)\|T_n - f_{j,\beta}\|_\infty \geq & \psi_n^{-1}(\beta) \max_{1\leq k\leq M} |T_n(a_k) - f_{j,\beta}(a_k)| \\ \geq & (1-\varepsilon) \max_{1\leq k\leq M} \left|\hat{\theta}_k - \delta_{j,k}\right|, \end{aligned}$$

where $\delta_{j,k}$ is the Kronecker delta and $\hat{\theta}_k = \frac{T_n(a_k)\psi_n^{-1}(\beta)}{1-\varepsilon}$. As a consequence

$$\psi_n^{-1}(\beta)\|T_n - f_{j,\beta}\|_\infty \geq d(\hat{\theta}, \theta_j),$$

where $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_M)$, $\theta_j = (\delta_{1,j}, \dots, \delta_{M,j})$, and $d(u,v) = (1-\varepsilon)\max_{1\leq k\leq M} |u_k - v_k|$ for two vectors $u = (u_1, \dots, u_M)$ and $v = (v_1, \dots, v_M)$ of $\mathbb{R}^M$.
In the same way

$$\begin{aligned} \psi_n^{-1}(\beta)\|T_n - f_{0,\beta}\|_\infty \geq & \psi_n^{-1}(\beta) \max_{1\leq k\leq M} |T_n(a_k)| \\ \geq & \psi_n^{-1}(\beta_{max}) \max_{1\leq k\leq M} |T_n(a_k)| \\ \geq & \frac{\psi_n(\beta)}{\psi_n(\beta_{max})} d(\hat{\theta}, \theta_0), \end{aligned}$$

where θ_0 is the null-vector of $\mathbb{R}^M$. Then we have

$$\begin{aligned} \Delta_n(\beta) &\geq \inf_{T_n} \max_{0\leq k\leq M} E_k\left(\|T_n - f_{k,\beta}\|_\infty^p \psi_n^{-p}(\beta)\right) \\ &\geq \inf_{\hat{\theta}\in\mathbb{R}^d} \max\left\{ E_0\left(\left(\frac{\psi_n(\beta)}{\psi_n(\beta_{max})}\right)^p d^p(\hat{\theta}, \theta_0)\right), \max_{1\leq k\leq M} E_k\left(d^p(\hat{\theta}, \theta_k)\right)\right\}, \end{aligned} \tag{13}$$

where E_k is the expectation $E_k = \mathbb{E}_{f_{k,\beta}}$. We denote also $P_k = \mathbb{P}_{f_{k,\beta}}$.
Here we apply Theorem 6 of Tsybakov (1998) to the relation (13). We consider the $M+1$ parameters $\{\theta_0, \ldots, \theta_M\} \subset [0,1]^M$, the family of probability measures $\{P_{\theta_j} = P_j\}$, the loss function $w(x) = x^p$ and the distance d previously defined. The parameters θ_j satisfy $d(\theta_i, \theta_k) \geq 1 - \varepsilon$ for $i, k \in \{0, \ldots, M\}$ and $i \neq k$. We are now going to prove that there exists α_n, with $0 \leq \alpha_n \leq 1$ and $\lim_{n\to\infty} \alpha_n = 0$, such that

$$Q\left(\frac{dP_0}{dQ} \geq \tau_n\right) \geq 1 - \alpha_n, \tag{14}$$

where $Q = \frac{1}{M}\sum_{k=1}^{M} P_k$, $\tau_n = n^{-\nu}$, $\nu = \frac{p\overline{\beta}_{max}}{2\overline{\beta}_{max}+1} - \frac{p\overline{\beta}}{2\overline{\beta}+1} - \frac{\varepsilon c_1(\beta)}{2\overline{\beta}+1}(1-\varepsilon/2)$ and ε is chosen small enough to have $\nu > 0$.
Indeed, if we prove (14), Theorem 6 of Tsybakov implies that

$$\Delta_n(\beta) \geq \frac{(1-\alpha_n)\tau_n(1-2\varepsilon)^p \left(\frac{\psi_n(\beta)\varepsilon}{\psi_n(\beta_{max})}\right)^p}{(1-2\varepsilon)^p + \tau_n \left(\frac{\psi_n(\beta)\varepsilon}{\psi_n(\beta_{max})}\right)^p}.$$

We have $\lim_{n\to\infty} \tau_n \left(\frac{\psi_n(\beta)\varepsilon}{\psi_n(\beta_{max})}\right)^p = +\infty$. Then

$$\liminf_{n\to\infty} \Delta_n(\beta) \geq (1-2\varepsilon)^p.$$

Since ε can be arbitrarily small, we obtain that $\liminf_{n\to\infty} \Delta_n(\beta) \geq 1$ and it proves the theorem.

Here we prove (14). In the same way as Tsybakov (1998), we have that under P_i

$$\frac{dP_k}{dP_0} = \begin{cases} \exp\left\{\xi_k v_n + v_n^2\right\} & k = i \\ \exp\left\{\xi_k v_n - v_n^2\right\} & k \neq i, \end{cases}$$

where the ξ_k are i.i.d. $\mathcal{N}(0,1)$ variables and $v_n^2 = \frac{n}{\sigma^2}\|f_{k,\beta}\|_2^2 = \frac{2\log n}{2\overline{\beta}+1} c_1(\beta)(1-\varepsilon)^2$. Now, by the independence of the ξ_i, we have

$$P_i\left(\frac{1}{M}\sum_{k=1}^{M}\frac{dP_k}{dP_0} < 1/\tau_n\right) \geq P_i\left(\frac{1}{M}\sum_{k=1,k\neq i}^{M}\frac{dP_k}{dP_0} < \frac{1}{2\tau_n}\right) P_i\left(\frac{1}{M}\frac{dP_i}{dP_0} < \frac{1}{2\tau_n}\right). \tag{15}$$

The probabilities above satisfy, since the ξ_k are $\mathcal{N}(0,1)$ variables,

$$\begin{aligned} P_i\left(\frac{1}{M}\sum_{k=1,k\neq i}^{M}\frac{dP_k}{dP_0} < \frac{1}{2\tau_n}\right) =& P_i\left(\frac{1}{M}\sum_{k=1,k\neq i}^{M}\exp\left\{\xi_k v_n - v_n^2\right\} < \frac{1}{2\tau_n}\right) \\ =& 1 - P_i\left(\sum_{k=1,k\neq i}^{M}\exp\left\{\xi_k v_n - v_n^2\right\} \geq \frac{M}{2\tau_n}\right) \qquad (16) \\ \geq& 1 - \frac{2\tau_n(M-1)}{M}E_i\left[\exp\left\{\xi_k v_n - v_n^2\right\}\right] \\ =& 1 - \frac{2\tau_n(M-1)}{M}, \qquad (17) \end{aligned}$$

and

$$P_i\left(\frac{1}{M}\frac{dP_i}{dP_0}<\frac{1}{2\tau_n}\right) = P_i\left(\exp\left\{\xi_i v_n + v_n^2\right\}<\frac{M}{2\tau_n}\right)$$
$$=1-\Phi\left(\frac{1}{v_n}\left(-\log M+\log(2\tau_n)+v_n^2/2\right)\right), \tag{18}$$

where Φ is the Gaussian c.d.f. By replacing M, τ_n and v_n by their values in terms of n and β, since $\frac{c_1(\beta)}{2\overline{\beta}+1}=\frac{1}{2\overline{\beta}+1}+\frac{p\overline{\beta}_{max}}{2\overline{\beta}_{max}+1}-\frac{p\overline{\beta}}{2\overline{\beta}+1}$, we obtain that

$$\frac{1}{v_n}\left(-\log M+\log(2\tau_n)+v_n^2/2\right)$$
$$=\frac{\log n}{v_n}\left(-\frac{1}{2\overline{\beta}+1}-\frac{p\overline{\beta}_{max}}{2\overline{\beta}_{max}+1}+\frac{p\overline{\beta}}{2\overline{\beta}+1}+\frac{\varepsilon c_1(\beta)}{2\overline{\beta}+1}(1-\varepsilon/2)+\frac{c_1(\beta)}{2\overline{\beta}+1}(1-2\varepsilon+\varepsilon^2)+o(1)\right)$$
$$=-\frac{\log n}{v_n}\left(\frac{c_1(\beta)\varepsilon}{2\overline{\beta}+1}(1-\varepsilon/2)+o(1)\right). \tag{19}$$

Using the relations (15), (17), (18), (19), noting that $\frac{2\tau_n(M-1)}{M}$ tends to 0 and that $-\frac{\log n}{v_n}\left(\frac{c_1(\beta)\varepsilon}{2\overline{\beta}+1}(1-\varepsilon/2)+o(1)\right)$ tends to $-\infty$, as n tends to ∞, we deduce that

$$\lim_{n\to\infty} Q\left(\frac{dP_0}{dQ}\geq\tau_n\right)=1.$$

This implies the existence of a sequence α_n that satisfies (14) and $\lim_{n\to\infty}\alpha_n=0$, which concludes the proof of the lower bound.

5 Proof of Theorem 2

Since Theorem 1 is true, in particular for set B satisfying Condition (P), to prove Theorem 2, it is enough to show that

$$\limsup_{n\to\infty}\sup_{\beta\in B}\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left\{\|\widetilde{f}_{(p)}-f\|_\infty^p\psi_n^{-p}(\beta)\right\}\leq 1. \tag{20}$$

Here we prove (20). Since the cardinal of B is finite, we are going to prove that for all $\beta\in B$,

$$\limsup_{n\to\infty}\Delta(\beta,n)\leq 1,$$

where

$$\Delta(\beta,n)=\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left\{\|\widetilde{f}_{(p)}-f\|_\infty^p\psi_n^{-p}(\beta)\right\}\leq R_{1,n}(\beta)+R_{2,n}(\beta),$$

with

$$R_{1,n}(\beta)=\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left\{\|\widetilde{f}_{(p)}-f\|_\infty^p\psi_n^{-p}(\beta)I_{\{\hat{\beta}_{(p)}<\beta\}}\right\}$$

and

$$R_{2,n}(\beta)=\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left\{\|\widetilde{f}_{(p)}-f\|_\infty^p\psi_n^{-p}(\beta)I_{\{\hat{\beta}_{(p)}\geq\beta\}}\right\}.$$

This will be obtained by proving that

$$\lim_{n\to\infty} R_{1,n}(\beta) = 0 \tag{21}$$

and that

$$\limsup_{n\to\infty} R_{2,n}(\beta) \le 1. \tag{22}$$

•Proof of (21)

Let $\beta \in B$. We have

$$R_{1,n}(\beta) \le \sum_{\gamma\in B,\gamma<\beta} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,1} - f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\hat{\beta}_{(p)}=\gamma\}} \right\}.$$

We fix now $\gamma \in B$ with $\gamma < \beta$. The event $\left\{\hat{\beta}_{(p)} = \gamma\right\}$ satisfies

$$\left\{\hat{\beta}_{(p)} = \gamma\right\} \subset \bigcup_{\beta'\in B,\beta'<\gamma'} \left\{ \|\hat{f}_{\gamma',1} - \hat{f}_{\beta',1}\|_\infty > \eta_1(\beta') \right\},$$

where $\gamma' = \min\{\lambda \in B \text{ such that } \lambda > \gamma\}$. Then we have

$$\sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,1} - f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\hat{\beta}_{(p)}=\gamma\}} \right\} \le \sum_{\beta'\in B,\beta'<\gamma'} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,1} - f\|_\infty^p \psi_n^{-p}(\beta) I_{A_{1,n}(\beta')} \right\},$$

with

$$A_{1,n}(\beta') = \left\{ \|\hat{f}_{\gamma',1} - \hat{f}_{\beta',1}\|_\infty > \eta_1(\beta') \right\}.$$

We do not indicate the dependence of the set $A_{1,n}(\beta')$ on γ' because we argue for fixed γ in B and then for fixed γ'. To prove the result (21), since the cardinal of B is finite, it is enough to prove that the following quantity, for $\beta' \in B$ with $\beta' < \gamma'$,

$$\sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,1} - f\|_\infty^p \psi_n^{-p}(\beta) I_{A_{1,n}(\beta')} \right\}, \tag{23}$$

tends to 0 as $n \to \infty$.

Here we study the quantity (23). Let $f \in \Sigma(\beta, L)$. We fix $\beta' \in B$ such that $\beta' < \gamma'$. We have

$$\mathbb{E}_f \left\{ \|\hat{f}_{\gamma,1} - f\|_\infty^p \psi_n^{-p}(\beta) I_{A_{1,n}(\beta')} \right\} \le 2^p \psi_n^{-p}(\beta) \left(\|b_{\gamma,1}(\cdot, f)\|_\infty^p \mathbb{P}_f(A_{1,n}(\beta')) + \mathbb{E}_f \left\{ \|Z_{\gamma,1}\|_\infty^p I_{A_{1,n}(\beta')} \right\} \right).$$

To study the quantity above, we need the following lemma which will be proved in Section 8.

Lemma 7. *We have for n large enough*

$$\sup_{f\in\Sigma(\beta,L)} \mathbb{P}_f \left[A_{1,n}(\beta') \right] \le D_6 (\log n)^{-\frac{1}{2\overline{\beta}'+1}} n^{-\frac{p(\overline{\beta}_{max}-\overline{\beta}')}{(2\overline{\beta}_{max}+1)(2\overline{\beta}'+1)}}.$$

The bias $b_{\gamma,1}$ satisfies, for all $f \in \Sigma(\beta, L)$,

$$\|b_{\gamma,1}(\cdot, f)\|_\infty \leq \int K_\gamma(u) \sum_{i=1}^{d} L_i |u_i h_{i,1}(\gamma)|^{\beta_i} du.$$

Since for all $i \in \{1, \cdots, d\}$, $\gamma_i < \beta_i$, we have that

$$\sup_{f \in \Sigma(\beta,L)} \|b_{\gamma,1}(\cdot, f)\| = o(\psi_n(\gamma)),$$

as $n \to \infty$. As a consequence, we deduce from Lemma 7 that

$$\sup_{f \in \Sigma(\beta,L)} \psi_n^{-p}(\beta) \|b_{\gamma,1}(\cdot, f)\|_\infty^p \mathbb{P}_f(A_{1,n}(\beta')) \leq D_7 (\log n)^{-\frac{1}{2\overline{\beta}'+1} + \frac{p\overline{\gamma}}{2\overline{\gamma}+1} - \frac{p\overline{\beta}}{2\overline{\beta}+1}} n^{-\frac{p\overline{\beta}_{max}}{2\overline{\beta}_{max}+1} + \frac{p\overline{\beta}'}{2\overline{\beta}'+1} - \frac{p\overline{\gamma}}{2\overline{\gamma}+1} + \frac{p\overline{\beta}}{2\overline{\beta}+1}}. \tag{24}$$

Moreover,

$$\begin{aligned}\mathbb{E}_f \left\{ \psi_n^{-p}(\beta) \|Z_{\gamma,1}\|_\infty^p I_{A_{1,n}(\beta')} \right\} &\leq (\tau_{n,1}(\gamma))^p \psi_n^{-p}(\beta) \mathbb{P}_f(A_{1,n}(\beta')) \\ &\quad + \mathbb{E}_f \left\{ \|Z_{\gamma,1}\|_\infty^p \psi_n^{-p}(\beta) I_{\{\|Z_{\gamma,1}\|_\infty > \tau_{n,1}(\gamma)\}} \right\}. \end{aligned} \tag{25}$$

From Lemma 7, we know that $\mathbb{P}_f(A_{1,n}(\beta'))$ is at most of order $(\log n)^{-\frac{1}{2\overline{\beta}'+1}} n^{-\frac{p\overline{\beta}_{max}}{2\overline{\beta}_{max}+1} + \frac{p\overline{\beta}'}{2\overline{\beta}'+1}}$ thus

$$(\tau_{n,1}(\gamma))^p \psi_n^{-p}(\beta) \mathbb{P}_f(A_{1,n}(\beta')) \leq D_8 (\log n)^{-\frac{1}{2\overline{\beta}'+1} + \frac{p\overline{\gamma}}{2\overline{\gamma}+1} - \frac{p\overline{\beta}}{2\overline{\beta}+1}} n^{-\frac{p\overline{\beta}_{max}}{2\overline{\beta}_{max}+1} + \frac{p\overline{\beta}'}{2\overline{\beta}'+1} - \frac{p\overline{\gamma}}{2\overline{\gamma}+1} + \frac{p\overline{\beta}}{2\overline{\beta}+1}}.$$

Since $\beta' \leq \gamma < \beta \leq \beta_{max}$, we obtain that

$$\lim_{n \to \infty} \sup_{f \in \Sigma(\beta,L)} \psi_n^{-p}(\beta) \|b_{\gamma,1}(\cdot, f)\|_\infty^p \mathbb{P}_f(A_{1,n}(\beta')) = 0 \tag{26}$$

and

$$\lim_{n \to \infty} \sup_{f \in \Sigma(\beta,L)} (\tau_{n,1}(\gamma))^p \psi_n^{-p}(\beta) \mathbb{P}_f(A_{1,n}(\beta')) = 0.$$

Using Lemma 5, we deduce that

$$\lim_{n \to \infty} \sup_{f \in \Sigma(\beta,L)} \mathbb{E}_f \left\{ \psi_n^{-p}(\beta) \|Z_{\gamma,1}\|_\infty^p I_{A_{1,n}(\beta')} \right\} = 0. \tag{27}$$

From (26) and (27), we conclude that the quantity (23) tends to 0 as $n \to \infty$.

•Proof of (22)

Let $\delta_n = \left(\frac{1}{\log n}\right)^{1/4}$ and $\beta \in B$. We have

$$\begin{aligned} R_{2,n}(\beta) &\leq (1+\delta_n)^p + \sup_{f \in \Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\widetilde{f}_{(p)} - f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\hat{\beta}_{(p)} \geq \beta\} \cap \{\|\widetilde{f}_{(p)} - f\|_\infty \psi_n^{-1}(\beta) > 1+\delta_n\}} \right\} \\ &\leq (1+\delta_n)^p + \sum_{\gamma \in B, \gamma \geq \beta} \sup_{f \in \Sigma(\beta,L)} Q_{2,n}(\beta, \gamma, f), \end{aligned}$$

where

$$Q_{2,n}(\beta,\gamma,f)=\mathbb{E}_f\left\{\|\hat f_{\gamma,1}-f\|_\infty^p\psi_n^{-p}(\beta)I_{\{\hat\beta_{(p)}=\gamma\}\cap\{\|\hat f_{\gamma,1}-f\|_\infty\psi_n^{-1}(\beta)>1+\delta_n\}}\right\}.$$

By Proposition 3, we have $\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)}Q_{2,n}(\beta,\beta,f)=0$. To prove (22), since the cardinal of B is finite and $\lim_{n\to\infty}\delta_n=0$, it is enough to prove, for $\gamma\in B$ with $\gamma>\beta$, that

$$\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)}Q_{2,n}(\beta,\gamma,f)=0. \tag{28}$$

Here we prove the result (28). Let $\gamma\in B$ such that $\gamma>\beta$ and $f\in\Sigma(\beta,L)$. If $\hat\beta_{(p)}=\gamma$, since $\gamma>\beta$, we have $\|\hat f_{\gamma,1}-\hat f_{\beta,1}\|_\infty\le\eta_1(\beta)$. Then,

$$\begin{aligned}\|\hat f_{\gamma,1}-f\|_\infty I_{\{\hat\beta_{(p)}=\gamma\}}&\le\left(\|\hat f_{\gamma,1}-\hat f_{\beta,1}\|_\infty+\|\hat f_{\beta,1}-f\|_\infty\right)I_{\{\hat\beta_{(p)}=\gamma\}}\\&\le\left(\eta_1(\beta)+\|\hat f_{\beta,1}-f\|_\infty\right)I_{\{\hat\beta_{(p)}=\gamma\}}.\end{aligned}\tag{29}$$

The quantity $Q_{2,n}(\beta,\gamma,f)$ is upper bounded by

$$\left(\mathbb{E}_f\left\{\|\hat f_{\gamma,1}-f\|_\infty^{2p}\psi_n^{-2p}(\beta)I_{\{\hat\beta_{(p)}=\gamma\}}\right\}\right)^{1/2}\left(\mathbb{P}_f\left(\{\hat\beta_{(p)}=\gamma\}\cap\{\|\hat f_{\gamma,1}-f\|_\infty\psi_n^{-1}(\beta)>1+\delta_n\}\right)\right)^{1/2}.$$

Then, (29) and Proposition 3 imply that

$$\sup_{f\in\Sigma(\beta,L)}\left(\mathbb{E}_f\left\{\|\hat f_{\gamma,1}-f\|_\infty^{2p}\psi_n^{-2p}(\beta)I_{\{\hat\beta_{(p)}=\gamma\}}\right\}\right)^{1/2}$$

is bounded above by a positive constant for n large enough and we deduce that

$$\sup_{f\in\Sigma(\beta,L)}Q_{2,n}(\beta,\gamma,f)\le D_9\left[\mathbb{P}_f\left(\{\hat\beta_{(p)}=\gamma\}\cap\{\|\hat f_{\gamma,1}-f\|_\infty\psi_n^{-1}(\beta)>1+\delta_n\}\right)\right]^{1/2}, \tag{30}$$

for n large enough. Now we are going to prove that the right hand side of inequality (30) tends to 0 as $n\to\infty$. We have

$$\begin{aligned}\|\hat f_{\gamma,1}-f\|_\infty I_{\{\hat\beta_{(p)}=\gamma\}}&\le\left(\|b_{\gamma,1}(\cdot,f)-b_{\beta,1}(\cdot,f)\|_\infty+\|b_{\beta,1}(\cdot,f)\|_\infty+\|Z_{\gamma,1}\|_\infty\right)I_{\{\hat\beta_{(p)}=\gamma\}}\\&\le\left(\|b_{\gamma,1}(\cdot,f)-b_{\beta,1}(\cdot,f)\|_\infty+\frac{\psi_n(\beta)}{2\overline\beta+1}+\|Z_{\gamma,1}\|_\infty\right)I_{\{\hat\beta_{(p)}=\gamma\}},\end{aligned}\tag{31}$$

where the last line is a consequence of Proposition 1.
We have $\|b_{\gamma,1}(\cdot,f)-b_{\beta,1}(\cdot,f)\|_\infty=\left\|\mathbb{E}_f\left(\hat f_{\gamma,1}\right)-\mathbb{E}_f\left(\hat f_{\beta,1}\right)\right\|_\infty$. The function $\phi:t\longmapsto\mathbb{E}_f\left(\hat f_{\gamma,1}(t)\right)-\mathbb{E}_f\left(\hat f_{\beta,1}(t)\right)$ is a continuous function on $[0,1]^d$ which admits a non-random maximum x_0 satisfying

$$\begin{aligned}\|b_{\gamma,1}(\cdot,f)-b_{\beta,1}(\cdot,f)\|_\infty&=\phi(x_0)\\&\le\left|\hat f_{\gamma,1}(x_0)-\hat f_{\beta,1}(x_0)\right|+|\varphi_1|\\&\le\|\hat f_{\gamma,1}-\hat f_{\beta,1}\|_\infty+|\varphi_1|\le\eta_1(\beta)+|\varphi_1|,\end{aligned}\tag{32}$$

where $\varphi_1 = \hat{f}_{\gamma,1}(x_0) - \hat{f}_{\beta,1}(x_0) - \mathbb{E}_f\left[\hat{f}_{\gamma,1}(x_0) - \hat{f}_{\beta,1}(x_0)\right]$. Since φ_1 is a $\mathcal{N}(0,\pi_n^2)$ variable, by Proposition 2, we deduce that its variance π_n^2 satisfies

$$\pi_n^2 \leq 2\mathbb{E}_f\left((Z_{\beta,1}(x_0))^2\right) + 2\mathbb{E}_f\left((Z_{\gamma,1}(x_0))^2\right) \leq \frac{2\|K_\beta\|_2^2\sigma^2(1+o(1))}{n\tilde{h}_1(\beta)}. \tag{33}$$

Then using (31) and (32) we deduce that

$$\begin{aligned}\|\hat{f}_{\gamma,1} - f\|_\infty I_{\{\hat{\beta}_{(p)}=\gamma\}} &\leq \left[\eta_1(\beta) + \frac{\psi_n(\beta)}{2\overline{\beta}+1} + \|Z_{\gamma,1}\|_\infty + |\varphi_1|\right] I_{\{\hat{\beta}_{(p)}=\gamma\}} \\ &\leq \left(\psi_n(\beta) + \|Z_{\gamma,1}\|_\infty + |\varphi_1|\right) I_{\{\hat{\beta}_{(p)}=\gamma\}},\end{aligned}$$

the last line being a consequence of Lemma 2. Thus,

$$\begin{aligned}&\mathbb{P}_f\left(\{\hat{\beta}_{(p)} = \gamma\} \cap \{\|\hat{f}_{\gamma,1} - f\|_\infty > (1+\delta_n)\psi_n(\beta)\}\right) \\ &\leq \mathbb{P}_f\left(\|Z_{\gamma,1}\|_\infty + |\varphi_1| > \delta_n\psi_n(\beta)\right) \\ &\leq \mathbb{P}_f\left(\|Z_{\gamma,1}\|_\infty > \frac{\delta_n}{2}\psi_n(\beta)\right) + \mathbb{P}_f\left(|\varphi_1| > \frac{\delta_n}{2}\psi_n(\beta)\right).\end{aligned} \tag{34}$$

As φ_1 is a $\mathcal{N}(0,\pi_n^2)$ variable and using (33), we have

$$\mathbb{P}_f\left(|\varphi_1| > \frac{\delta_n}{2}\psi_n(\beta)\right) \leq \exp\left\{-\frac{\psi_n^2(\beta)\delta_n^2 n\tilde{h}_1(\beta)}{16\|K_\beta\|_2^2\sigma^2}\right\}.$$

The quantity $\psi_n^2(\beta)\delta_n^2 n\tilde{h}_1(\beta)$ is of order $\sqrt{\log n}$, and then

$$\lim_{n\to\infty} \mathbb{P}_f\left(|\varphi_1| > \frac{\delta_n}{2}\psi_n(\beta)\right) = 0. \tag{35}$$

Using Proposition 2, we have

$$\mathbb{P}_f\left(\|Z_{\gamma,1}\|_\infty > \frac{\delta_n}{2}\psi_n(\beta)\right) \leq \frac{D_1}{\tilde{h}_1(\gamma)} \exp\left\{-\frac{\psi_n^2(\beta)\delta_n^2 n\tilde{h}_1(\gamma)}{8\|K_\beta\|_2^2\sigma^2}\right\} \exp\left\{\frac{D_2\delta_n\psi_n(\beta)\sqrt{n\tilde{h}_1(\gamma)}}{2(\log \tilde{h}_1(\gamma))^{1/2}}\right\}.$$

The quantity $\psi_n^2(\beta)\delta_n^2 n\tilde{h}_1(\gamma)$ is of order $\psi_n^2(\beta)\psi_n^{-2}(\gamma)(\log n)^{D_{10}}$ with some $D_{10} \in \mathbb{R}$ and $\tilde{h}_1(\gamma)$ is of order $\left(\frac{\log n}{n}\right)^{1/(2\overline{\gamma}+1)}$. Hence, since $\beta < \gamma$, this implies that

$$\lim_{n\to\infty} \mathbb{P}_f\left(\|Z_{\gamma,1}\|_\infty > \frac{\delta_n}{2}\psi_n(\beta)\right) = 0. \tag{36}$$

Using (35), (36) and that the fact that the right hand side of inequality (34) does not depend on f, we deduce that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{P}_f\left(\{\hat{\beta}_{(p)} = \gamma\} \cap \{\|\hat{f}_{\gamma,1} - f\|_\infty > (1+\delta_n)\psi_n(\beta)\}\right) = 0.$$

Then we obtain the result (28) which implies (22).

6 Proof of Theorem 3

The scheme of proof is similar to the proof of relation (20) in the proof of Theorem 2. To prove Theorem 3, we will prove that, for all $\beta \in B$,

$$\lim_{n\to\infty} R_{3,n}(\beta) = 0 \tag{37}$$

and

$$\limsup_{n\to\infty} R_{4,n}(\beta) \leq (M_2(\beta))^p, \tag{38}$$

where

$$R_{3,n}(\beta) = \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\widetilde{f}_{ani} - f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\hat{\beta}_{ani}<\beta\}} \right\},$$

$$R_{4,n}(\beta) = \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\widetilde{f}_{ani} - f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\hat{\beta}_{ani}\geq\beta\}} \right\}.$$

•Proof of (37)

Let $\beta, \gamma \in B$ such that $\gamma < \beta$.
The event $\left\{\hat{\beta}_{ani} = \gamma\right\}$ satisfies

$$\left\{\hat{\beta}_{ani} = \gamma\right\} \subset \bigcup_{\beta'\in B, \beta'<\beta} A_{2,n}(\beta'),$$

where

$$A_{2,n}(\beta') = \left\{ \|\hat{f}_{\beta*\beta',2} - \hat{f}_{\beta',2}\|_\infty > \eta_2(\beta') \right\}.$$

Following the same reasoning as in the proof of relation (20), to prove (37), it is enough to prove that, for all $\beta' \in B$ with $\beta' < \beta$,

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,2} - f\|_\infty^p \psi_n^{-p}(\beta) I_{A_{2,n}(\beta')} \right\} = 0. \tag{39}$$

The event $A_{2,n}(\beta')$ satisfies the following lemma:

Lemma 8. *We have for n large enough*

$$\sup_{f\in\Sigma(\beta,L)} \mathbb{P}_f \left[A_{2,n}(\beta')\right] \leq D_{11} (\log n)^{-\frac{1}{2\overline{\beta}'+1}} n^{-p\mu(\beta')}.$$

The bias $b_{\gamma,2}$ satisfies for all $f \in \Sigma(\beta, L)$

$$\|b_{\gamma,2}(\cdot, f)\|_\infty \leq \int K_\gamma(u) \sum_{i=1}^d L_i |u_i h_{i,2}(\gamma)|^{\beta_i} du,$$

then $\sup_{f\in\Sigma(\beta,L)} \|b_{\gamma,2}(\cdot, f)\|_\infty$ is at most of order

$$\left(\frac{\log n}{n}\right)^{\min\left(\frac{\beta_i \overline{\gamma}}{\gamma_i(2\overline{\gamma}+1)}\right)},$$

where the minimum is taken on $i = 1, \dots, d$. Therefore

$$\sup_{f\in\Sigma(\beta,L)} \psi_n^{-p}(\beta)\|b_{\gamma,2}(\cdot,f)\|_\infty^p \mathbb{P}_f(A_{2,n}(\beta'))$$

is upper bounded by

$$D_{12}(\log n)^{-p\omega(\beta,\gamma)-\frac{1}{2\overline{\beta}'+1}} n^{-p(\mu(\beta')-\omega(\beta,\gamma))}$$

where

$$\omega(\beta,\gamma) = \frac{\overline{\beta}}{2\overline{\beta}+1} - \min_{i=1,\dots,d} \frac{\beta_i\overline{\gamma}}{\gamma_i(2\overline{\gamma}+1)}.$$

If B satisfies Condition (P), then

$$\omega(\beta,\gamma) < \frac{\overline{\beta}}{2\overline{\beta}+1} - \frac{\overline{\gamma}}{2\overline{\gamma}+1}.$$

Since $\overline{\beta} > \overline{\gamma}$ and $\mu(\beta') > 0$, this implies that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \psi_n^{-p}(\beta)\|b_{\gamma,2}(\cdot,f)\|_\infty^p \mathbb{P}_f(A_{2,n}(\beta')) = 0. \tag{40}$$

If B do not satisfy Condition (P), there exists $i \in \{1,\dots,d\}$ such that $\beta_i/\gamma_i \le 1$, then

$$\mu(\beta') \ge \omega(\beta,\gamma) \ge \frac{\overline{\beta}}{2\overline{\beta}+1} - \frac{\overline{\gamma}}{2\overline{\gamma}+1} > 0.$$

This implies that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \psi_n^{-p}(\beta)\|b_{\gamma,2}(\cdot,f)\|_\infty^p \mathbb{P}_f(A_{2,n}(\beta')) = 0. \tag{41}$$

Reasoning as in the proof of (21), using the decomposition (25), Lemma 5, Lemma 8 and that

$$\frac{\overline{\beta}}{2\overline{\beta}+1} - \frac{\overline{\gamma}}{2\overline{\gamma}+1} \le \mu(\beta'),$$

we deduce that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f\left\{\psi_n^{-p}(\beta)\|Z_{\gamma,2}\|_\infty^p I_{A_{2,n}(\beta')}\right\} = 0. \tag{42}$$

From (41), (40) and (42), we obtain (37).

•Proof of (38)

The proof will be similar to the proof in the proof of (22). We fix $\delta_n = \left(\frac{1}{\log n}\right)^{1/4}$. We have

$$R_{4,n}(\beta) \le (1+\delta_n)^p (M_2(\beta))^p + \sum_{\gamma\in B,\gamma\ge\beta} \sup_{f\in\Sigma(\beta,L)} Q_{4,n}(\beta,\gamma,f),$$

where

$$Q_{4,n}(\beta,\gamma,f) = \mathbb{E}_f\left\{\|\hat{f}_{\gamma,2}-f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\hat{\beta}_{ani}=\gamma\}\cap C_{1,n}}\right\},$$

and

$$C_{1,n} = \{\|\hat{f}_{\gamma,2}-f\|_\infty \psi_n^{-1}(\beta) > M_2(\beta)(1+\delta_n)\}.$$

By Proposition 3, we have $\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)} Q_{4,n}(\beta,\beta,f) = 0$. To prove (38), it is enough to prove that, for all $\gamma \in B$ such that $\gamma > \beta$,

$$\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)} Q_{4,n}(\beta,\gamma,f) = 0. \tag{43}$$

Let $\gamma \in B$, such that $\gamma > \beta$, and $f \in \Sigma(\beta, L)$. Following the same reasoning as in the proof of (22) from (31) to (33), using Proposition 1, Lemma 2 and Lemma 3, we deduce that

$$\begin{aligned}
\|\hat f_{\gamma,2} - f\|_\infty I_{\{\hat\beta_{ani}=\gamma\}} \leq & (\|b_{\beta*\gamma,2}(\cdot,f) - b_{\gamma,2}(\cdot,f)\|_\infty + \|b_{\beta*\gamma,2}(\cdot,f) - b_{\beta,2}(\cdot,f)\|_\infty \\
& + \|b_{\beta,2}(\cdot,f)\|_\infty + \|Z_{\gamma,2}\|_\infty) I_{\{\hat\beta_{ani}=\gamma\}} \\
\leq & \left(\frac{2\psi_n(\beta)\lambda_2(\beta)}{2\bar\beta+1} + \eta_2(\beta) + \|Z_{\gamma,2}\|_\infty + |\varphi_2|\right) I_{\{\hat\beta_{ani}=\gamma\}}, \\
\leq & (M_2(\beta) + \|Z_{\gamma,2}\|_\infty + |\varphi_2|)\, I_{\{\hat\beta_{ani}=\gamma\}}
\end{aligned} \tag{44}$$

where $\varphi_2 = \hat f_{\beta*\gamma,2}(x_0) - \hat f_{\beta,2}(x_0) - \mathbb{E}_f\left[\hat f_{\beta*\gamma,2}(x_0) - \hat f_{\beta,2}(x_0)\right]$, with some $x_0 \in [0,1]^d$. Using Lemma 4 and Proposition 2, we have that φ_2 is a $\mathcal{N}(0,\pi_n^2)$ variable, with variance π_n^2 satisfying

$$\pi_n^2 \leq \frac{2\|K_\beta\|_2^2\sigma^2(1+o(1))}{n\tilde h_1(\beta)}.$$

We have

$$\mathbb{E}_f\left[\|Z_{\gamma,2}\|_\infty^{2p}\psi_n^{-2p}(\beta)\right] \leq (\tau_{n,2}(\gamma))^{2p}\psi_n^{-2p}(\beta) + \psi_n^{-2p}(\beta)\mathbb{E}_f\left[\|Z_{\gamma,2}\|_\infty^{2p} I_{\{\|Z_{\gamma,2}\|_\infty > \tau_{n,2}(\gamma)\}}\right].$$

Reasoning as in the proof of Lemma 5, we have that

$$\lim_{n\to\infty}(\psi_n(\beta))^{-2p}\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left[\|Z_{\gamma,2}\|_\infty^{2p} I_{\{\|Z_{\gamma,2}\|_\infty > \tau_{n,2}(\gamma)\}}\right] = 0.$$

Since $\gamma > \beta$, we deduce that

$$\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left[\|Z_{\gamma,2}\|_\infty^{2p}\psi_n^{-2p}(\beta)\right] = 0. \tag{45}$$

Moreover, the variable φ_3 satisfies the properties

$$\mathbb{P}_f\left[|\varphi_2| > \delta_n\psi_n(\beta)t\right] \leq \exp\left\{-D_{13}t^2\sqrt{\log n}\right\}, t \geq 0, \tag{46}$$

$$\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)}\mathbb{E}_f\left[|\varphi_2|^{2p}\psi_n^{-2p}(\beta)\right] = 0. \tag{47}$$

The property (46) comes from the fact that φ_2 is a $\mathcal{N}(0,\pi_n^2)$ variable, and the property (47) can be proved as (45) using (46) and the proof of Lemma 5.

The relations (45) and (47) and Proposition 3 imply that

$$\sup_{f\in\Sigma(\beta,L)}\left(\mathbb{E}_f\left\{\|\hat f_{\gamma,2} - f\|_\infty^{2p}\psi_n^{-2p}(\beta) I_{\{\hat\beta_{ani}=\gamma\}}\right\}\right)^{1/2}$$

is bounded above by a positive constant for n large enough. Now to have (43), it is enough to prove that

$$\lim_{n\to\infty}\sup_{f\in\Sigma(\beta,L)}\mathbb{P}_f\left[C_{1,n}\cap\{\hat\beta_{ani}=\gamma\}\right] = 0. \tag{48}$$

Using the relations (44) and (46), and following the proof of (22) from (34) to (36), we deduce (48), which finishes the proof.

7 Proof of Theorem 4

•Proof of (8)

By Proposition 3, we have

$$\limsup_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta,1} - f\|_\infty^p \psi_n^{-p}(\beta) \right\} \leq 1$$

Then to have (8), it is enough to prove that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,3} - f\|_\infty^p \psi_n^{-p}(\beta) I_{\{\widetilde{f}_{ani2}=\hat{f}_{\gamma,3}\}} \right\} = 0 \tag{49}$$

The event $\left\{\widetilde{f}_{ani2} = \hat{f}_{\gamma,3}\right\}$ satisfies $\left\{\widetilde{f}_{ani2} = \hat{f}_{\gamma,3}\right\} \subset A_{3,n} \cap A_{4,n}$, where

$$A_{3,n} = \left\{ \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\gamma,3}\|_\infty > \eta_3(\gamma) \right\},$$

$$A_{4,n} = \left\{ \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\beta,1}\|_\infty < \frac{\psi_n(\gamma)\lambda_3(\gamma)}{2\overline{\gamma}+1}(1+\rho_n) \right\},$$

The event $A_{3,n}$ satisfies the lemma:

Lemma 9. *We have for n large enough*

$$\sup_{f\in\Sigma(\beta,L)} \mathbb{P}_f \left[A_{3,n}\right] \leq D_{14} (\log n)^{-\frac{1}{2\overline{\gamma}+1}} n^{-\frac{p(\overline{\beta}-\overline{\gamma})}{(2\overline{\beta}+1)(2\overline{\gamma}+1)}}.$$

The proof of Lemma 9 is similar to the proof of Lemma 7 and 8. Let $f \in \Sigma(\beta, L)$. We have

$$\|\hat{f}_{\gamma,3} - f\|_\infty I_{A_{3,n}\cap A_{4,n}} \leq \left(\|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\gamma,3}\|_\infty + \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\beta,1}\|_\infty + \|\hat{f}_{\beta,1} - f\|_\infty \right) I_{A_{3,n}\cap A_{4,n}}$$

Using Proposition 3 and Lemma 9 we deduce that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta,1} - f\|_\infty^p \psi_n^{-p}(\beta) I_{A_{3,n}\cap A_{4,n}} \right\} = 0,$$

Moreover, using the definition of $A_{4,n}$, we have

$$\mathbb{E}_f \left\{ \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\beta,1}\|_\infty^p \psi_n^{-p}(\beta) I_{A_{3,n}\cap A_{4,n}} \right\} \leq \left(\frac{\lambda_3(\gamma)\psi_n(\gamma)\psi_n^{-1}(\beta)}{2\overline{\gamma}+1} \right)^p (1+\rho_n)^p \mathbb{P}_f(A_{3,n}).$$

Using Lemma 9, we deduce that, for n large enough, $\psi_n^{-p}(\beta)\psi_n^p(\gamma)(1+\rho_n)^p\mathbb{P}_f(A_{3,n})$ is at most of order $(\log n)^{-\frac{1}{2\overline{\gamma}+1}+\frac{p\overline{\gamma}}{2\overline{\gamma}+1}-\frac{p\overline{\beta}}{2\overline{\beta}+1}}$, therefore, since $\gamma < \beta$,

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\beta,1}\|_\infty^p \psi_n^{-p}(\beta) I_{A_{3,n}\cap A_{4,n}} \right\} = 0$$

To prove (49), it remains to prove that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\gamma,3}\|_\infty^p \psi_n^{-p}(\beta) I_{A_{3,n}\cap A_{4,n}} \right\} = 0.$$

We have

$$\|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\gamma,3}\|_\infty I_{A_{3,n}\cap A_{4,n}} \leq \left(\|b_{\beta*\gamma,1}(\cdot,f) - b_{\gamma,3}(\cdot,f)\|_\infty + \|Z_{\gamma,3}\|_\infty + \|Z_{\beta*\gamma,1}\|_\infty\right) I_{A_{3,n}\cap A_{4,n}}.$$

Reasoning as in the proof of (21), using the decomposition (25) and applying Lemma 5 to $Z_{\gamma,3}$, we deduce that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ (\|Z_{\gamma,3}\|_\infty)^p \, \psi_n^{-p}(\beta) I_{A_{3,n}\cap A_{4,n}} \right\} = 0.$$

Since $Z_{\beta*\gamma,1}$ satisfies Lemma 4, a similar reasoning permits to have

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left\{ (\|Z_{\beta*\gamma,3}\|_\infty)^p \, \psi_n^{-p}(\beta) I_{A_{3,n}\cap A_{4,n}} \right\} = 0.$$

Appying Lemma 3, we obtain that the quantity $\sup_{f\in\Sigma(\beta,L)} \|b_{\beta*\gamma,1}(\cdot,f) - b_{\gamma,3}(\cdot,f)\|_\infty$ is of order $\psi_n(\beta)$, and finally we have by Lemma 9 that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \|b_{\beta*\gamma,1}(\cdot,f) - b_{\gamma,3}(\cdot,f)\|_\infty \psi_n^{-p}(\beta) \mathbb{P}_f\left(A_{3,n}\cap A_{4,n}\right) = 0,$$

which prove (49) and then (8).

•Proof of (9)

The proof will be similar to the proof in the proof of (22). Let $f \in \Sigma(\gamma, L)$. We fix $\delta_n = \left(\frac{1}{\log n}\right)^{1/4}$. We have

$$\begin{aligned}
\mathbb{E}_f \left\{ \|\widetilde{f}_{ani2} - f\|_\infty^p \psi_n^{-p}(\gamma) \right\} \leq & (1+\delta_n)^p (M_3(\gamma))^p + \mathbb{E}_f \left\{ \|\widetilde{f}_{ani2} - f\|_\infty^p \psi_n^{-p}(\gamma) I_{C_{2,n}} \right\}, \\
\leq & (1+\delta_n)^p (M_3(\gamma))^p + \mathbb{E}_f \left\{ \|\hat{f}_{\beta,1} - f\|_\infty^p \psi_n^{-p}(\gamma) I_{C_{2,n}\cap A_n} \right\} \\
& + \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,3} - f\|_\infty^p \psi_n^{-p}(\gamma) I_{C_{2,n}\cap A_n^c} \right\},
\end{aligned}$$

where

$$C_{2,n} = \left\{ \|\widetilde{f}_{ani2} - f\|_\infty \psi_n^{-1}(\gamma) > (1+\delta_n) M_3(\gamma) \right\},$$

and $A_n = (A_{3,n} \cap A_{4,n})^c$. By Proposition 3, we have

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\gamma,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\gamma,3} - f\|_\infty^p \psi_n^p(\gamma) I_{C_{2,n}\cap A^c} \right\} = 0.$$

The event $A_{4,n}$ satisfies the lemma

Lemma 10. *For n large enough, we have*

$$\sup_{f\in\Sigma(\gamma,L)} \mathbb{P}_f\left[A_{4,n}^c\right] \leq \frac{D_{15}}{\tilde{h}_1(\beta)} \exp\left\{-D_{16} (\log n)^{D_{17}} n^{D_{18}}\right\} \exp\left\{-D_{19}\sqrt{(\log n)^{D_{17}} n^{D_{18}}}\right\},$$

with $D_{17} \in \mathbb{R}$.

Lemma 10 implies that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\gamma,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta,1} - f\|_\infty^p \psi_n^{-p}(\gamma) I_{C_{2,n}\cap A_{4,n}^c} \right\} = 0.$$

Then to have (9), it is enough to prove that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\gamma,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta,1} - f\|_\infty^p \psi_n^{-p}(\gamma) I_{A_{5,n}} \right\} = 0, \tag{50}$$

where $A_{5,n} = C_{2,n} \cap A_{3,n}^c \cap A_{4,n}$ On $A_{3,n}^c \cap A_{4,n}$, we have

$$\begin{aligned} \|\hat{f}_{\beta,1} - f\|_\infty \leq & \|\hat{f}_{\beta,1} - \hat{f}_{\beta*\gamma,1}\|_\infty + \|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\gamma,3}\|_\infty + \|\hat{f}_{\gamma,3} - f\|_\infty \\ \leq & \eta_3(\gamma) + \frac{\psi_n(\gamma)\lambda_3(\gamma)(1+\rho_n)}{2\overline{\gamma}+1} + \|\hat{f}_{\gamma,3} - f\|_\infty. \end{aligned}$$

Using Proposition 3, we deduce that, for n large enough,

$$\sup_{f\in\Sigma(\gamma,L)} \mathbb{E}_f \left\{ \|\hat{f}_{\beta,1} - f\|_\infty^{2p} \psi_n^{-2p}(\gamma) I_{A_{5,n}} \right\}$$

is upper bounded by a constant. Now to have (50), it is enough to prove that

$$\lim_{n\to\infty} \sup_{f\in\Sigma(\gamma,L)} \mathbb{P}_f \left[A_{5,n}\right] = 0. \tag{51}$$

We have by Lemma 2 and Proposition 1,

$$\begin{aligned} \|\hat{f}_{\beta,1} - f\|_\infty I_{A_{5,n}} \leq & \Big(\|\hat{f}_{\beta,1} - \hat{f}_{\beta*\gamma,1}\|_\infty + \|b_{\beta*\gamma,1}(\cdot,f) - b_{\gamma,3}(\cdot,f)\|_\infty + \|b_{\gamma,3}(\cdot,f)\|_\infty + \|Z_{\beta*\gamma,1}\|_\infty \Big) I_{A_{5,n}} \\ \leq & \|b_{\beta*\gamma,1}(\cdot,f) - b_{\gamma,3}(\cdot,f)\|_\infty + \frac{2\lambda_3(\gamma)\psi_n(\gamma)}{2\overline{\gamma}+1} + \|Z_{\beta*\gamma,1}\|_\infty. \end{aligned}$$

Following the same argument as in the proof of (22) from (31) to (33), we deduce that

$$\|b_{\beta*\gamma,1}(\cdot,f) - b_{\gamma,3}(\cdot,f)\|_\infty \leq \eta_3(\gamma) + |\varphi_3| \tag{52}$$

where $\varphi_3 = \hat{f}_{\beta*\gamma,1}(x_0) - \hat{f}_{\gamma,3}(x_0) - \mathbb{E}_f \left[\hat{f}_{\beta*\gamma,1}(x_0) - \hat{f}_{\gamma,3}(x_0) \right]$, with some $x_0 \in [0,1]^d$. Using Lemma 4, we have that φ_3 is a random variable $\mathcal{N}(0,\pi_n^2)$, where π_n^2 satisfies

$$\pi_n^2 \leq \frac{2\|K_\gamma\|_2^2 \sigma^2 (1+o(1))}{n\tilde{h}_1(\gamma)}.$$

Using Lemma 3 and Lemma 4, we obtain finally that

$$\begin{aligned} \|\hat{f}_{\beta,1} - f\|_\infty I_{A_{5,n}} & \leq \left(\eta_3(\gamma) + \frac{2\lambda_3(\gamma)\psi_n(\gamma)}{2\overline{\gamma}+1} + \|Z_{\beta*\gamma,1}\|_\infty + |\varphi_3| \right) I_{A_{5,n}} \\ & \leq \left(M_3(\gamma)\psi_n(\gamma) + \|Z_{\beta*\gamma,1}\|_\infty + |\varphi_3| \right) I_{A_{5,n}} \end{aligned}$$

Now following the proof of (22) from (34) to (36), we deduce (51), which finishes the proof of (50).

8 Proofs of the lemmas and propositions

Proof of Lemma 2

Since $\|K_\beta\|_2^2 = \frac{2(\overline{\beta}+1)}{\overline{\beta}\alpha(\beta)(2\overline{\beta}+1)}$, we have for $j \in \{1,2,3\}$,

$$\eta_j(\beta) = \left(\frac{\log n}{n}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}} C_\beta^{-\frac{1}{2\overline{\beta}}} \left(\frac{2c_j(\beta)\|K_\beta\|_2^2\sigma^2 L_*^{1/\overline{\beta}}}{2\overline{\beta}+1}\right)^{1/2} (\lambda_j(\beta))^{-\frac{1}{2\overline{\beta}}},$$

$$= \left(\frac{\log n}{n}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}} C_\beta^{-\frac{1}{2\overline{\beta}}} \left(\frac{4\overline{\beta}^2 c_j(\beta)(\overline{\beta}+1)\sigma^2 L_*^{1/\overline{\beta}}}{(2\overline{\beta}+1)^2\alpha(\beta)\overline{\beta}^3}\right)^{1/2} (\lambda_j(\beta))^{-\frac{1}{2\overline{\beta}}}.$$

For j=1, we deduce that

$$\eta_1(\beta) = \left(\frac{\log n}{n}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}} C_\beta^{-\frac{1}{2\overline{\beta}}} C_\beta^{\frac{2\overline{\beta}+1}{2\overline{\beta}}} \frac{2\overline{\beta}}{2\overline{\beta}+1} = \frac{2\overline{\beta}}{2\overline{\beta}+1}\psi_n(\beta).$$

This implies the result of the lemma for $j=1$. For $j \in \{2,3\}$, we have

$$\eta_j(\beta) = \frac{2\overline{\beta}}{2\overline{\beta}+1}\psi_n(\beta)2^{-1/(2\overline{\beta}+1)}\left(\frac{c_j(\beta)}{c_1(\beta)}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}},$$

which permits to have the result for $j \in \{2,3\}$.

Proof of Proposition 3

Following the proof of Theorem 1 of Bertin (2004) with the loss function $w(x) = x^p$, studying the bias term and the stochastic term of $\hat{f}_{\beta,j}$, we can deduce, for $j \in \{1,2,3\}$, that

$$\limsup_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{P}_f\left[\|\hat{f}_{\beta,j} - f\|_\infty > (1+\varepsilon_n)\psi_n(\beta)\left(\frac{\lambda_j(\beta)}{2\overline{\beta}+1} + \frac{\eta_j(\beta)}{\psi_n(\beta)(c_j(\beta))^{1/2}}\right)\right] = 0$$

and

$$\limsup_{n\to\infty} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f\left[\|\hat{f}_{\beta,j} - f\|_\infty^p \psi_n^{-p}(\beta)\right] \le \left(\frac{\lambda_j(\beta)}{2\overline{\beta}+1} + \frac{\eta_j(\beta)}{\psi_n(\beta)c_j(\beta)}\right)^p.$$

For $j=1$, using that $\eta_1(\beta) = \frac{2\overline{\beta}}{2\overline{\beta}+1}\psi_n(\beta)$ (cf. proof of Lemma 2), and that $c_1(\beta) \ge 1$, we have that

$$\frac{\lambda_1(\beta)}{2\overline{\beta}+1} + \frac{\eta_1(\beta)}{\psi_n(\beta)c_1(\beta)} = \frac{1}{2\overline{\beta}+1} + \frac{2\overline{\beta}}{(2\overline{\beta}+1)(c_1(\beta))^{1/2}} \le 1.$$

For $j \in \{2,3\}$, using that $\eta_j(\beta) = \frac{2\overline{\beta}}{2\overline{\beta}+1}\psi_n(\beta)2^{-1/(2\overline{\beta}+1)}\left(\frac{c_j(\beta)}{c_1(\beta)}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}}$ (cf. proof of Lemma 2), and that $c_j(\beta) \ge 1$, we have that

$$\frac{\lambda_j(\beta)}{2\overline{\beta}+1} + \frac{\eta_j(\beta)}{\psi_n(\beta)c_j(\beta)} = \left(\frac{c_j(\beta)}{c_1(\beta)}\right)^{\frac{\overline{\beta}}{2\overline{\beta}+1}}\left(\frac{2^{-\frac{2\overline{\beta}}{2\overline{\beta}+1}}}{2\overline{\beta}+1} + \frac{2\overline{\beta}2^{-\frac{1}{2\overline{\beta}+1}}}{(2\overline{\beta}+1)(c_2(\beta))^{1/2}}\right) \le M_j(\beta).$$

The above lines imply the results of Proposition 3.

Proof of Lemma 3

Here we prove the first result of the lemma. Let $\beta = (\beta_1, \dots, \beta_d) \in B$, $\gamma \in B$, $f \in \Sigma(\beta, L)$ and $x = (x_1, \dots, x_d), y = (y_1, \dots, y_d) \in \mathbb{R}^d$. We have

$$
\begin{aligned}
&\left|\overline{K}_{\gamma,2} * f(x) - \overline{K}_{\gamma,2} * f(y)\right| \\
=&\left|\int K_\gamma(u)\left\{f(x_1 - u_1 h_{1,2}(\gamma), \dots, x_d - u_d h_{d,2}(\gamma)) - f(y_1 - u_1 h_{1,2}(\gamma), \dots, y_d - u_d h_{d,2}(\gamma))\right\} du\right| \\
\le&\left(\int K_\gamma(u)du\right)\sum_{i=1}^d L_i|x_i - y_i|^{\beta_i} = \sum_{i=1}^d L_i|x_i - y_i|^{\beta_i}
\end{aligned}
$$

Then, $\overline{K}_{\gamma,2} * f$ belongs to $\Sigma(\beta, L)$. As a consequence, we have

$$
\|b_{\beta*\gamma,2}(\cdot, f) - b_{\gamma,2}(\cdot, f)\|_\infty = \|b_{\beta,2}(\cdot, \overline{K}_{\gamma,2} * f)\|_\infty \le \sup_{f \in \Sigma(\beta,L)} \|b_{\beta,2}(\cdot, f)\|_\infty.
$$

Thus, from Proposition 1, we deduce that

$$
\sup_{f \in \Sigma(\beta,L)} \|b_{\beta*\gamma,2}(\cdot, f) - b_{\gamma,2}(\cdot, f)\|_\infty \le \frac{\psi_n(\beta)\lambda_2(\beta)}{2\overline{\beta} + 1}.
$$

The two other results can be proved exactly in the same way.

Proof of Lemma 4

We prove here the result for $j = 2$. The result for $j = 1$ can be proved exactly in the same way. Let $\beta = (\beta_1, \dots, \beta_d), \gamma \in B$ such that $\gamma < \beta$ and $t \in [0, 1]^d$. We have

$$
\mathbb{E}_f\left[Z^2_{\beta*\gamma,2}(t)\right] = \frac{\sigma^2}{n\tilde{h}_2^2(\beta)\tilde{h}_2^2(\gamma)} \int \left(\int K_\beta\left(\frac{t - u - v}{h_2(\beta)}\right) K_\gamma\left(\frac{v}{h_2(\gamma)}\right) dv\right)^2 du,
$$

where the notation $\frac{u}{t}$, for two vectors $u = (u_1, \dots, u_d)$ and $t = (t_1, \dots, t_d)$, represents the vector $(u_1/t_1, \dots, u_d/t_d)$. By the generalized Minkowskii inequality (cf. Appendix), we deduce that

$$
\begin{aligned}
\mathbb{E}_f\left[Z^2_{\beta*\gamma,2}(t)\right] \le& \frac{\sigma^2}{n\tilde{h}_2^2(\beta)\tilde{h}_2^2(\gamma)} \left(\int \left(\int K^2_\beta\left(\frac{t - u - v}{h_2(\beta)}\right) K^2_\gamma\left(\frac{v}{h_2(\gamma)}\right) du\right)^{1/2} dv\right)^2 \\
\le& \frac{\sigma^2\|K_\beta\|_2^2}{n\tilde{h}_2^2(\beta)\tilde{h}_2^2(\gamma)} \left(\int K_\gamma\left(\frac{v}{h_2(\gamma)}\right) dv\right)^2 = \frac{\sigma^2\|K_\beta\|_2^2}{n\tilde{h}_2(\beta)}.
\end{aligned}
$$

Now the proof of (12) is similar to the proof of Proposition 3, and then to that of Lemma 4 of Bertin (2004). The process $Z_{\beta*\gamma}$ satisfies as Z_β, $\mathbb{E}_f\left[Z^2_{\beta*\gamma}(t)\right] \le \frac{\sigma^2\|K_\beta\|_2^2}{n\tilde{h}_2(\beta)}$. Furthermore, to apply that argument, we need to bound the quantity $\mathbb{E}_f\left[\left(Z^2_{\beta*\gamma}(t) - Z^2_{\beta*\gamma}(s)\right)^2\right]$ by a multiple of

$\frac{1}{n\tilde{h}_2(\beta)}\sum_{i=1}^d \left|\frac{s_i-t_i}{h_{i,2}(\beta)}\right|^{\beta_i}$ for $s=(s_1,\ldots,s_d)$, $t=(t_1,\ldots,t_d)$ in $[0,1]^d$. Here we look at this quantity. Let $s,t\in[0,1]^d$, we have

$$\begin{aligned}
&\mathbb{E}_f\left[\left(Z^2_{\beta*\gamma}(t)-Z^2_{\beta*\gamma}(s)\right)^2\right]\\
&=\frac{\sigma^2}{n\tilde{h}_2^2(\beta)\tilde{h}_2^2(\gamma)}\int\left(\int K_\gamma\left(\frac{v}{h_2(\gamma)}\right)\left(K_\beta\left(\frac{t-u-v}{h_2(\beta)}\right)-K_\beta\left(\frac{s-u-v}{h_2(\beta)}\right)\right)dv\right)^2du\\
&\le\frac{\sigma^2}{n\tilde{h}_2^2(\beta)\tilde{h}_2^2(\gamma)}\left(\int\left(\int K^2_\gamma\left(\frac{v}{h_2(\gamma)}\right)\left(K_\beta\left(\frac{t-u-v}{h_2(\beta)}\right)-K_\beta\left(\frac{s-u-v}{h_2(\beta)}\right)\right)^2du\right)^{1/2}dv\right)^2\\
&\le\frac{\sigma^2D_{20}}{n\tilde{h}_2(\beta)}\left(\frac{1}{\tilde{h}_2(\gamma)}\int K_\gamma\left(\frac{v}{h_2(\gamma)}\right)dv\right)^2\sum_{i=1}^d\left|\frac{s_i-t_i}{h_{i,2}(\beta)}\right|^{\beta_i}=\frac{\sigma^2D_{20}}{n\tilde{h}_2(\beta)}\sum_{i=1}^d\left|\frac{s_i-t_i}{h_{i,2}(\beta)}\right|^{\beta_i},
\end{aligned}$$

the second line being obtained by the generalized Minkowskii inequality (cf. Appendix) and the third line coming from the fact K_β satisfies an Hölder condition of order β.

As a consequence, following the proof of Lemma 4 of Bertin (2004), we will obtain the result for $j=2$ of Lemma 4.

Proof of Lemma 5

Let $f\in\Sigma(\beta,L)$ and $j\in\{1,2,3\}$. We have, by a change of variables,

$$\begin{aligned}
&\mathbb{E}_f\left\{\|Z_{\gamma,j}\|^p_\infty\psi_n^{-p}(\beta)I_{\{\|Z_{\gamma,j}\|_\infty>\tau_{n,j}(\gamma)\}}\right\}=\psi_n^{-p}(\beta)\int_0^{+\infty}\mathbb{P}_f\left(\|Z_{\gamma,j}\|^p_\infty I_{\{\|Z_{\gamma,j}\|_\infty>\tau_{n,j}(\gamma)\}}>t\right)dt\\
=&\psi_n^{-p}(\beta)\left(\tau_{n,j}(\gamma)\right)^p\int_0^{+\infty}\mathbb{P}_f\left(\|Z_{\gamma,j}\|^p_\infty I_{\{\|Z_{\gamma,j}\|_\infty>\tau_{n,j}(\gamma)\}}>t\left(\tau_{n,j}(\gamma)\right)^p\right)dt\\
=&\ \psi_n^{-p}(\beta)\left(\tau_{n,j}(\gamma)\right)^p\mathbb{P}_f\left(\|Z_{\gamma,j}\|_\infty>\tau_{n,j}(\gamma)\right)+\psi_n^{-p}(\beta)\left(\tau_{n,j}(\gamma)\right)^p\int_1^{+\infty}\mathbb{P}_f\left(\|Z_{\gamma,j}\|^p_\infty>t\left(\tau_{n,j}(\gamma)\right)^p\right)dt.
\end{aligned}\tag{53}$$

We are going to prove that the two elements of the sum (53) tend to 0 as n tends to ∞. We obtain by Proposition 2 that

$$\mathbb{P}_f\left(\|Z_{\gamma,j}\|^p_\infty>t\left(\tau_{n,j}(\gamma)\right)^p\right)\le D_1\left(\frac{n}{\log n}\right)^{\frac{1}{2\overline{\gamma}+1}}\exp\left\{-\frac{t^{2/p}(1+p\overline{\beta}_{max})}{2\overline{\gamma}+1}\log n\right\}\exp\left\{t^{1/p}D_{21}\sqrt{\log n}\right\}.\tag{54}$$

Using the formula (54) with $t=1$, we deduce that $\psi_n^{-p}(\beta)\left(\tau_{n,j}(\gamma)\right)^p\mathbb{P}_f\left(\|Z_{\gamma,j}\|_\infty>\tau_{n,j}(\gamma)\right)$ is of order $n^{p(\frac{\overline{\beta}}{2\overline{\beta}+1}-\frac{\overline{\gamma}}{2\overline{\gamma}+1}-\frac{\overline{\beta}_{max}}{2\overline{\gamma}+1})}(\log n)^{D_{22}}$ with some $D_{22}\in\mathbb{R}$, and then it tends to 0 as n tends to ∞ since $\gamma\le\beta\le\beta_{max}$. Moreover by (54) we have, after a change of variables,

$$\int_1^{+\infty}\mathbb{P}_f\left(\|Z_{\gamma,j}\|^p_\infty>t\left(\tau_{n,j}(\gamma)\right)^p\right)dt$$

is bounded above by

$$D_{23}\left(\frac{n}{\log n}\right)^{\frac{1}{2\overline{\gamma}+1}}\int_1^{+\infty}\exp\left\{-\frac{t^2}{2\overline{\gamma}+1}(1+p\overline{\beta}_{max})\log n\right\}\exp\left\{D_{21}t\sqrt{\log n}\right\}t^{p-1}dt.$$

By using several integrations by parts, one can find that this integral is at most of order

$$\left(\frac{n}{\log n}\right)^{\frac{1}{2\overline{\gamma}+1}} \exp\left\{-\frac{(1+p\overline{\beta}_{max})}{2\overline{\gamma}+1}\log n\right\} = n^{-\frac{p\overline{\beta}_{max}}{2\overline{\gamma}+1}} (\log n)^{-\frac{1}{2\overline{\gamma}+1}}.$$

Hence $\psi_n^{-p}(\beta)\,(\tau_{n,j}(\gamma))^p \int_1^{+\infty} \mathbb{P}_f\left(\|Z_{\gamma,j}\|_\infty^p > t\,(\tau_{n,j}(\gamma))^p\right) dt$ tends to 0 as n tends to ∞.

Proof of Lemma 7

Let $f \in \Sigma(\beta, L)$. Using the same argument as in the proof of (21), we can deduce that $\|b_{\gamma',1}(\cdot, f)\|_\infty = o(\psi_n(\gamma'))$ and $\|b_{\beta',1}(\cdot, f)\|_\infty = o(\psi_n(\beta'))$ for all $f \in \Sigma(\beta, L)$. As a consequence for all $f \in \Sigma(\beta, L)$ $\|b_{\gamma',1}(\cdot, f) - b_{\beta',1}(\cdot, f)\|_\infty = o(\psi_n(\beta'))$, since $\beta' < \gamma'$, and therefore

$$\mathbb{P}_f\left[\|\hat{f}_{\gamma',1} - \hat{f}_{\beta',1}\|_\infty > \eta_1\left(\beta'\right)\right] \le \mathbb{P}_f\left[\|Z_{\gamma',1}\|_\infty + \|Z_{\beta',1}\|_\infty > \eta_1\left(\beta'\right)(1+\kappa_n)\right] \le P_1(n) + P_2(n),$$

where κ_n is of order $n^{-\delta}$ with a $\delta > 0$ and

$$P_1(n) = \mathbb{P}_f\left[\|Z_{\gamma',1}\|_\infty > \psi_n((\beta'+\gamma')/2)(1+\kappa_n)\right],$$

$$P_2(n) = \mathbb{P}_f\left[\|Z_{\beta',1}\|_\infty > \eta_1\left(\beta'\right)(1+\kappa_n)\left(1 - \frac{\psi_n(\frac{\beta'+\gamma'}{2})}{\eta_1\left(\beta'\right)}\right)\right].$$

Using Proposition 2, since $\frac{\eta_1^2(\beta') n \tilde{h}_1(\beta')}{2\|K_{\beta'}\|_2^2\sigma^2} = \frac{1}{2\overline{\beta}'+1}\left(1 + p\frac{\overline{\beta}_{max}-\overline{\beta}'}{2\overline{\beta}_{max}+1}\right)\log n$, we obtain that for n large enough

$$P_2(n) \le D_{24}\,(\log n)^{-\frac{1}{2\overline{\beta}'+1}} \exp\left\{-\frac{p}{2\overline{\beta}'+1}\left(\frac{\overline{\beta}_{max}-\overline{\beta}'}{2\overline{\beta}_{max}+1}\right)\log n\right\}. \tag{55}$$

Using Proposition 2, it can be proved that $P_1(n)$ is negligible with respect to $P_2(n)$ as $n \to \infty$. The relation (55) implies the lemma.

Proof of Lemma 8

Let $f \in \Sigma(\beta, L)$. By Lemma 3, we have that $\|b_{\beta' * \beta,2}(\cdot, f) - b_{\beta',2}(\cdot, f)\|_\infty$ is at most of order $\psi_n(\beta)$. Since $\beta' < \beta$, $\psi_n(\beta)$ is negligible with respect to $\eta_2(\beta')$ and therefore

$$\mathbb{P}_f\left[\|\hat{f}_{\beta*\beta',2} - \hat{f}_{\beta',2}\|_\infty > \eta_2\left(\beta'\right)\right] \le \mathbb{P}_f\left[\|Z_{\beta*\beta',2}\|_\infty + \|Z_{\beta',2}\|_\infty > \eta_2\left(\beta'\right)(1+\kappa_n)\right] \le P_3(n) + P_4(n),$$

where κ_n is of order $\psi_n(\beta)/\eta_2(\beta')$,

$$P_3(n) = \mathbb{P}_f\left[\|Z_{\beta*\beta',2}\|_\infty > \psi_n((\beta'+\beta)/2)(1+\kappa_n)\right],$$

and

$$P_4(n) = \mathbb{P}_f\left[\|Z_{\beta',2}\|_\infty > \eta_2\left(\beta'\right)(1+\kappa_n)\left(1 - \frac{\psi_n(\frac{\beta'+\beta}{2})}{\eta_2\left(\beta'\right)}\right)\right].$$

Using Proposition 2, since

$$\frac{\eta_2^2(\beta')\prod_{i=1}^d h_{i,2}(\beta')}{2\|K_{\beta'}\|_2^2\sigma^2} = \left(\frac{1}{2\overline{\beta}'+1} + p\mu(\beta')\right)\log n,$$

we obtain that for n large enough

$$P_4(n) \leq D_{11} \left(\log n\right)^{-\frac{1}{2\overline{\beta}'+1}} \exp\left\{-\left(p\mu(\beta')\right)\log n\right\}. \tag{56}$$

Using Lemma 4, it can be proved that $P_3(n)$ is negligible with respect to $P_4(n)$ as $n \to \infty$. The relation (56) implies the lemma.

Proof of Lemma 10

Let $f \in \Sigma(\gamma, L)$. By Lemma 3, we have that

$$\|b_{\beta*\gamma,1}(\cdot,f) - b_{\beta,1}(\cdot,f)\|_\infty \leq \frac{\psi_n(\gamma)\lambda_3(\gamma)}{2\overline{\gamma}+1}.$$

Then

$$\mathbb{P}_f\left[\|\hat{f}_{\beta*\gamma,1} - \hat{f}_{\beta,1}\|_\infty \geq \frac{\psi_n(\gamma)\lambda_3(\gamma)}{2\overline{\gamma}+1}(1+\rho_n)\right] \leq P_5(n) + P_6(n),$$

where

$$P_5(n) = \mathbb{P}_f\left[\|Z_{\beta*\gamma,1}\|_\infty > \frac{\psi_n(\gamma)\lambda_3(\gamma)\rho_n}{2(2\overline{\gamma}+1)}\right],$$

and

$$P_6(n) = \mathbb{P}_f\left[\|Z_{\beta,1}\|_\infty > \frac{\psi_n(\gamma)\lambda_3(\gamma)\rho_n}{2(2\overline{\gamma}+1)}\right].$$

Using Proposition 2, since $\rho_n = \frac{\psi_n((\beta+\gamma)/2)}{\psi_n(\gamma)}$, we obtain that

$$P_6(n) \leq \frac{D_1}{\tilde{h}_1(\beta)} \exp\left\{-\frac{\lambda_3^2(\gamma)\psi_n^2((\beta+\gamma)/2)n\tilde{h}_1(\beta)}{8\|K_\beta\|_2^2\sigma^2(2\overline{\gamma}+1)^2}\right\} \exp\left\{-\frac{D_2\lambda_3(\gamma)\psi_n((\beta+\gamma)/2)\sqrt{n\tilde{h}_1(\beta)}}{2(2\overline{\gamma}+1)\sqrt{\log\tilde{h}_1(\beta)}}\right\}. \tag{57}$$

Using Lemma 4, we have that $P_5(n)$ satisfies the same inequality as (57) but with different constants D_1 and D_2. Now, since $\beta > \gamma$, we have

$$\frac{\lambda_3^2(\gamma)\psi_n^2((\beta+\gamma)/2)n\tilde{h}_1(\beta)}{8\|K_\beta\|_2^2\sigma^2(2\overline{\gamma}+1)^2} = D_{16}(\log n)^{D_{17}} n^{D_{18}},$$

with $D_{17} \in \mathbb{R}$, which includes the lemma.

Appendix: Minkowskii inequality

For all function g measurable on $\mathbb{R}^d \times \mathbb{R}^d$, we have

$$\int\left(\int g(u,x)du\right)^2 dx \leq \left(\int\left(\int g^2(u,x)dx\right)^{1/2} du\right).$$

(cf. Besov et al. (1978) for a proof)

References

Barron, A., Birgé, L., and Massart, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields*, 113(3):301–413.

Bertin, K. (2004). Asymptotically exact minimax estimation in sup-norm for anisotropic Hölder classes. *Bernoulli*, 10(5):873–888.

Besov, O. V., Ilin, V. P., and Nikolskii, S. M. (1978). *Integral representations of functions and imbedding theorems. Vol. I.* V. H. Winston & Sons, Washington, D.C. Scripta Series in Mathematics, Edited by Mitchell H. Taibleson.

Gradshteyn, I. S. and Ryzhik, I. M. (1965). *Table of integrals, series, and products.* Fourth edition prepared by Ju. V. Geronimus and M. Ju. Ceĭtlin. Translated from the Russian by Scripta Technica, Inc. Translation edited by Alan Jeffrey. Academic Press, New York.

Ibragimov, I. A. and Hasminskii, R. Z. (1981). *Statistical estimation*, volume 16 of *Applications of Mathematics.* Springer-Verlag, New York. Asymptotic theory, Translated from the Russian by Samuel Kotz.

Kerkyacharian, G., Lepski, O., and Picard, D. (2001). Nonlinear estimation in anisotropic multi-index denoising. *Probab. Theory Related Fields*, 121(2):137–170.

Klemelä, J. and Tsybakov, A. B. (2001). Sharp adaptive estimation of linear functionals. *Ann. Statist.*, 29(6):1567–1600.

Lepski, O. V. (1992). On problems of adaptive estimation in white Gaussian noise. In *Topics in nonparametric estimation*, volume 12 of *Adv. Soviet Math.*, pages 87–106. Amer. Math. Soc., Providence, RI.

Lepski, O. V. and Levit, B. Y. (1998). Adaptive minimax estimation of infinitely differentiable functions. *Math. Methods Statist.*, 7(2):123–156.

Neumann, M. H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.*, 25(1):38–76.

Nussbaum, M. (1986). On the nonparametric estimation of regression functions that are smooth in a domain in $\mathbf{R}^k$. *Theory Probab. Appl.*, 31(1):118–125.

Stone, C. J. (1985). Additive regression and other nonparametric models. *Ann. Statist.*, 13(2):689–705.

Tsybakov, A. B. (1998). Pointwise and sup-norm sharp adaptive estimation of functions on the Sobolev classes. *Ann. Statist.*, 26(6):2420–2469.