

EXACT ADAPTIVE POINTWISE ESTIMATION ON SOBOLEV CLASSES OF DENSITIES

CRISTINA BUTUCEA¹

Abstract. The subject of this paper is to estimate adaptively the common probability density of n independent, identically distributed random variables. The estimation is done at a fixed point $x_0 \in \mathbb{R}$, over the density functions that belong to the Sobolev class $W_n(\beta, L)$. We consider the adaptive problem setup, where the regularity parameter β is unknown and varies in a given set B_n . A sharp adaptive estimator is obtained, and the explicit asymptotical constant, associated to its rate of convergence is found.

Mathematics Subject Classification. 62N01, 62N02, 62G20.

Received November 15, 2000. Revised April 9, 2001.

1. INTRODUCTION

Consider n independent, identically distributed random variables X_1, \dots, X_n , having common unknown probability density $f : \mathbb{R} \rightarrow [0, \infty)$. We assume that f belongs to a Sobolev class of densities.

For any $L > 0$ and β positive integer, we define the Sobolev class of densities $W(\beta, L)$, as the set of functions

$$W(\beta, L) = \left\{ f : \mathbb{R} \rightarrow [0, \infty) : \int_{\mathbb{R}} f = 1, \int_{\mathbb{R}} \left(f^{(\beta)}(x) \right)^2 dx \leq L^2 \right\},$$

where $f^{(\beta)}$ will denote from now on the generalized derivative of order β of f . We may define for an absolutely integrable function $f : \mathbb{R} \rightarrow \mathbb{R}$ its Fourier transform $\mathcal{F}(f)(x) = \int_{\mathbb{R}} f(y)e^{-ixy} dy$, for any x in \mathbb{R} . We adopt now a more general definition of the Sobolev class, allowing non-integer values of $\beta > 1/2$

$$W(\beta, L) = \left\{ f : \mathbb{R} \rightarrow [0, \infty) : \int_{\mathbb{R}} f = 1, \int_{\mathbb{R}} |\mathcal{F}(f)(x)|^2 |x|^{2\beta} dx \leq 2\pi L^2 \right\}.$$

Let f_n be an estimator of f based on the sample X_1, \dots, X_n and x_0 a fixed point. The performance of the estimator f_n at the point x_0 is measured by the maximal risk

$$R_{n,\beta}^*(f_n, \varphi_{n,\beta}) = \sup_{f \in W(\beta, L)} E_f \left[\varphi_{n,\beta}^{-q} |f_n(x_0) - f(x_0)|^q \right], \quad (1.1)$$

Keywords and phrases: Density estimation, exact asymptotics, pointwise risk, sharp adaptive estimator.

¹ Université Paris 10, Modal'X, bâtiment G, 200 avenue de la République, 92001 Nanterre, France; e-mail: cbutucea@u-paris10.fr and Université Paris 6, Laboratoire Probabilités et Modèles Aléatoires, 6 rue Clisson, 75013 Paris, France; e-mail: butucea@ccr.jussieu.fr

where $E_f(\cdot)$ is the expectation with respect to the distribution P_f of X_1, \dots, X_n , when the underlying probability density is f , $\varphi_{n,\beta}$ is a given sequence of positive numbers and $q > 0$.

The sequence $\varphi_{n,\beta}$ such that the maximal risk related to (1.1) remains positive for all estimation procedures f_n , asymptotically, and finite for some explicit estimator, asymptotically, is called the optimal rate of convergence for the class $W(\beta, L)$. Using the argument of optimal recovery as in Donoho and Low [10], it is easy to find that the optimal pointwise rate of convergence over the Sobolev class $W(\beta, L)$ is $\varphi_{n,\beta} = (1/n)^{\frac{\beta-1/2}{2\beta}}$. This rate and the estimator attaining this rate depend on the regularity β of the unknown density f . Thus, it is difficult to implement such an estimator in practice. Our goal is to suggest an adaptive estimator $f_n(x_0)$ of $f(x_0)$, $x_0 \in \mathbb{R}$, *i.e.* an estimator independent of the regularity β of f , which is optimal in an exact asymptotic sense.

To define the notion of adaptive optimality we follow the minimax framework applied to the problem of adaptivity by Lepskii [23]. He considered the Gaussian white noise model, rather than density estimation. In this context, he introduced the notions of adaptive rate of convergence and rate adaptive estimator.

Adaptive rates of convergence on different functional classes for the Gaussian white noise model were obtained by Donoho *et al.* [8] (who give a detailed overview of the results in adaptive estimation), by Lepski *et al.* [27], Goldenshluger and Nemirovski [13], Juditsky [19]. Most of these results relate to Besov classes of functions. The latter three papers use the Lepski type of adaptation.

For the same framework of the Gaussian white noise model, exact adaptive results are available for several cases. Exact adaptivity means that not only the rate but the best asymptotic constant associated to it is attained by the proposed adaptive estimator. The first result of this kind in the estimation in \mathbb{L}_2 norm on Sobolev periodic classes belongs to Efromovich and Pinsker [12]. For further developments see Golubev [14, 15], Golubev and Nussbaum [17]. Then Lepskii [23], Lepski and Spokoiny [28] obtained exact adaptive results in \mathbb{L}_∞ and at a fixed point, respectively, on the Hölder classes with $0 < \beta \leq 2$ (see also Lepskii [24] and [25]). Tsybakov [30] proved exact adaptive results for the Gaussian white noise model both in \mathbb{L}_∞ and at a fixed point, on the Sobolev classes. Lepski and Levit [26] gave exact adaptive results in pointwise estimation over a large scale of infinitely differentiable functions.

Similar results on the adaptive rates of convergence exist in density estimation. We refer here to Donoho *et al.* [9], Kerkycharian *et al.* [21] and Juditsky [19] for the general setting of Besov classes and \mathbb{L}_p norm, with $p < \infty$. They applied the wavelet shrinkage in order to construct the adaptive estimator. Barron *et al.* [1], Birgé and Massart [3] propose different adaptive density estimators constructed by the methods of penalization and prove their adaptivity in the \mathbb{L}_2 norm. Devroye and Lugosi [7] obtained similar results for the \mathbb{L}_1 norm using an original adaptation procedure.

The results on exact adaptive estimation of the density of i.i.d. random variables, under the quadratic \mathbb{L}_2 risk, are due to Efromovich [11] and Golubev [16]. Efromovich [11] considered estimation over Sobolev and more general ellipsoids of periodic densities on $[0, 1]$, Golubev [16] estimated a density in Sobolev classes over the real line.

In this paper, we consider the problem of exact adaptive density estimation at a fixed point on the Sobolev classes. Note that, similarly to the results of Lepskii [23], Brown and Low [4] concerning the pointwise estimation in Gaussian white noise model, the adaptive estimation with the optimal rate $\varphi_{n,\beta} = (1/n)^{\frac{\beta-1/2}{2\beta}}$ is not possible and we have to normalize the risk by the adaptive rate of convergence, which is logarithmically worse than the optimal rate. In fact, the adaptive rate of convergence on a slightly modified Sobolev class $W_n(\beta, L)$, described below, is of the order $(\log n/n)^{\frac{\beta-1/2}{2\beta}}$ (see Butucea [5]). The main result of this paper is to find the constant $c(\beta, L, q, f(x_0))$ multiplying this rate of convergence which allows to attain the exact asymptotics of the minimax adaptive risk. We also construct explicitly the adaptive estimator that attains this exact asymptotics. This extends a pointwise adaptive result of Tsybakov [30] to the problem of density estimation.

2. RESULTS

Our results concern the exact adaptive estimation, at a fixed point $x_0 \in \mathbb{R}$, over the Sobolev class of densities, with regularity $\beta \in B_n$, for some set B_n . We introduce the following class of Sobolev densities

$$W_n(\beta, L) = \{f \in W(\beta, L) : f(x_0) \geq \rho_n\},$$

where ρ_n is a sequence of positive real numbers that satisfies

$$\lim_{n \rightarrow \infty} \rho_n = 0 \text{ and } \liminf_{n \rightarrow \infty} (\rho_n \log n) > 0. \quad (2.1)$$

This truncation prevents us from the possible case of a density f that varies with n such that $f(x_0) \rightarrow 0$ too fast as $n \rightarrow \infty$.

An alternative possibility is to consider that our density belongs to a local Sobolev class, where the smoothness on a neighborhood around the estimation point is quantified. In this context only the adaptive rate of convergence is maintained by the estimation approach described here and not the exact constant normalization.

Consider now the following maximal risk over $W_n(\beta, L)$ at fixed x_0 , for $q > 0$

$$R_{n,\beta}(f_n, \psi_{n,\beta}) = \sup_{f \in W_n(\beta, L)} E_f \left[\psi_{n,\beta}^{-q} |f_n(x_0) - f(x_0)|^q \right]. \quad (2.2)$$

We assume that the set B_n of regularities is a discrete set, $B_n = \{\beta_1, \dots, \beta_{N_n}\}$, where $1/2 < \beta_1 < \dots < \beta_{N_n} < +\infty$ are positive integers. Moreover, we suppose that $\beta_1 > 1/2$ is fixed, while $\lim_{n \rightarrow \infty} \beta_{N_n} = +\infty$ and $\{N_n\}_{n \geq 1}$ is a nondecreasing sequence of positive integers. We define $\Delta_n = \max_{i=1 \dots N_n-1} |\beta_{i+1} - \beta_i|$ and assume that it satisfies

$$\limsup_{n \rightarrow \infty} \Delta_n < +\infty \quad (2.3)$$

together with

$$\lim_{n \rightarrow \infty} \frac{\Delta_n \log n}{\beta_{N_n}^2 \log \log n} = \infty. \quad (2.4)$$

Considering a discrete set B_n of regularities is not only a technical matter. Similarly to Lepski's result, we may as well consider a bounded set of regularities $[a, b] \subseteq (1/2, \infty)$. This changes nothing to the adaptive rate of convergence, but a factor $\frac{1}{2\beta} - \frac{1}{2b}$ should appear in the constant (where β is the true underlying regularity). We refer to Klemelä and Tsybakov [22] for such type of sharp normalizations. A continuous set of values $[a, b]$ would reduce to the same techniques by considering a grid of values on this interval which is growing finer with n . From a practical point of view, the estimation method described later works on discretized sets of regularities. Our approach considers a larger union of classes, where β_{N_n} goes to infinity, without loss in the rate.

In particular, conditions (2.3) and (2.4) entail

$$\frac{\log n}{\beta_{N_n}} \xrightarrow[n \rightarrow \infty]{} +\infty \quad (2.5)$$

$$\left(\frac{\beta_{N_n}}{\log n} \right)^{1/2} \log \beta_{N_n} \xrightarrow[n \rightarrow \infty]{} 0 \quad (2.6)$$

$$\left(\frac{\beta_{N_n}}{\log n} \right)^{1/2} \log \frac{1}{\Delta_n} \xrightarrow[n \rightarrow \infty]{} 0. \quad (2.7)$$

This is proved in the beginning of Section 3.

The following definition is a modification for the density problem of the adaptive optimality introduced by Lepskii [23] (see also Tsybakov [30]).

Definition 2.1. The sequence $\psi_{n,\beta}$ is an **adaptive rate of convergence** over the set B_n , if:

1. there exists an estimator f_n^* , independent of β over B_n , which is called **rate adaptive estimator**, such that

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_n} R_{n,\beta}(f_n^*, \psi_{n,\beta}) < \infty; \quad (2.8)$$

2. if there exists another sequence of positive real numbers $\rho_{n,\beta}$ and an estimator f_n^{**} such that

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_n} R_{n,\beta}(f_n^{**}, \rho_{n,\beta}) < \infty$$

and, at some β' in B_n , $\frac{\rho_{n,\beta'}}{\psi_{n,\beta'}} \xrightarrow{n \rightarrow \infty} 0$, then there is another β'' in B_n such that $\frac{\rho_{n,\beta'}}{\psi_{n,\beta'}} \cdot \frac{\rho_{n,\beta''}}{\psi_{n,\beta''}} \xrightarrow{n \rightarrow \infty} +\infty$.

Note that condition (2.8) introduces a wide class of rates. We choose between those rates by a criterion of uniformity over the set B_n , expressed in the second part of Definition 2.1. If some other rate satisfies a condition similar to (2.8) and if this rate is faster at some point β' then the loss at some other point β'' has to be infinitely greater for large sample sizes n .

Let us denote $B_- = B_n \setminus \{\beta_{N_n}\}$ and

$$\psi_{n,\beta} = \begin{cases} c \left(\frac{\log n}{n} \right)^{\frac{\beta-1/2}{2\beta}}, & \text{for } \beta \in B_- \\ \left(\frac{1}{n} \right)^{\frac{\beta-1/2}{2\beta}}, & \text{for } \beta = \beta_{N_n}, \end{cases} \quad (2.9)$$

where the constant c is function of β , L , q and $f(x_0)$, $c = c(\beta, L, q, f(x_0)) > 0$ and satisfies

$$0 < \liminf_{\beta \rightarrow \infty} c(\beta, L, q, f(x_0)) \sqrt{\beta} \leq \limsup_{\beta \rightarrow \infty} c(\beta, L, q, f(x_0)) \sqrt{\beta} < \infty.$$

2.1. Exact adaptive estimation procedure

The estimation procedure contains three steps. First, we consider a preliminary estimator $\hat{f}_n(x_0)$ of $f(x_0)$, as the following kernel estimator

$$\hat{f}_n(x_0) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{X_i - x_0}{h_n}\right), \quad (2.10)$$

where K is a bounded, positive kernel, such that $\int |u| |K(u)| du < \infty$, and the bandwidth is $h_n > 0$. Let us mention that the regularity β of f being greater than $1/2$ the density is regular enough in order to be pointwise evaluable.

The bandwidth satisfies $\lim_{n \rightarrow \infty} h_n = 0$ and $\lim_{n \rightarrow \infty} nh_n = \infty$, which guarantee that $\hat{f}_n(x_0)$ is a consistent estimator. This preliminary estimator does not need to be particularly well performing, but for technical reasons in Lemma 3.2 it must not be too slow. Let us consider a preliminary bandwidth such that

$$h_n = O(n^{-\alpha_0}), \quad (2.11)$$

for some fixed $0 < \alpha_0 < 1/2$. We truncate this estimator at ρ_n that tends to 0 when $n \rightarrow \infty$ such that (2.1) holds and get $\hat{\rho}_n(x_0) = \max\{\hat{f}_n(x_0), \rho_n\}$.

The second step consists in defining a family of kernel estimators, whose bandwidths contain the preliminary estimator as follows

$$\begin{aligned}\widehat{h}_{n,\beta} &= k_\beta \left(\frac{\widehat{\rho}_n(x_0) \log n}{n} \right)^{\frac{1}{2\beta}}, \text{ where } \beta \in B_- \text{ and } k_\beta = \left(\frac{q}{2\beta(2\beta-1)L^2} \right)^{\frac{1}{2\beta}} \\ \widehat{h}_{n,\beta_{N_n}} &= \left(\frac{1}{n} \right)^{\frac{1}{2\beta_{N_n}}}, \text{ if } \beta = \beta_{N_n}.\end{aligned}$$

For $\beta > 1/2$, define a kernel K_β by the expressions

$$K_\beta(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\exp(ixu)}{1+|u|^{2\beta}} du = \frac{1}{\pi} \int_0^\infty \frac{\cos(xu)}{1+|u|^{2\beta}} du. \quad (2.12)$$

Define the kernel estimator depending on β

$$\widehat{f}_{n,\beta}(x_0) = \frac{1}{n\widehat{h}_{n,\beta}} \sum_{i=1}^n K_\beta \left(\frac{X_i - x_0}{\widehat{h}_{n,\beta}} \right).$$

At the third step, we use a version of Lepski's method to approximate β by an estimator $\widehat{\beta}$ and to substitute it in the expression of the kernel estimator $\widehat{f}_{n,\beta}(x_0)$. For any $\beta > 1/2$, we define

$$b_\beta^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{|u|^{2\beta}}{(1+|u|^{2\beta})^2} du, \nu_\beta^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{(1+|u|^{2\beta})^2} du.$$

By simple calculation, we can prove that $\nu_\beta^2 = (2\beta-1)b_\beta^2$. We consider the sequence

$$\widehat{\eta}_{n,\beta} = \nu_\beta \sqrt{\frac{q}{2\beta k_\beta}} \left(\frac{\widehat{\rho}_n(x_0) \log n}{n} \right)^{\frac{\beta-1/2}{2\beta}}$$

and the estimator of β defined as

$$\widehat{\beta} = \max \left\{ \beta \in B_n : \left| \widehat{f}_{n,\gamma}(x_0) - \widehat{f}_{n,\beta}(x_0) \right| \leq \widehat{\eta}_{n,\gamma}, \forall \gamma \in B_n, \gamma \leq \beta \right\}.$$

Finally, we replace β by $\widehat{\beta}$ in the kernel estimator $\widehat{f}_{n,\beta}(x_0)$ in order to get the estimator

$$f_n^*(x_0) = \widehat{f}_{n,\widehat{\beta}}(x_0). \quad (2.13)$$

This estimator will be shown to be exact adaptive *i.e.* to attain the adaptive rate of convergence in our setup, up to a constant, explicitly given.

The key point in this construction is Lepski's algorithm in the third step for the evaluation of the smoothness β of the underlying density, at each point x_0 . The same method was employed by Tsybakov [30] in the Gaussian white noise model but for orthogonal series estimators of the signal function. We introduce here kernel estimators which are more suitable for density estimation and are largely used in practice. Moreover, we provide optimal kernel (in the sense of exact adaptivity) and bandwidth expression of the adaptive estimator.

For a simulation study of the behavior of this estimator we refer to Butucea [6]. We have implemented the described adaptive estimator and tested it over ten different densities. We have considered a set $B = \{1, \dots, 6\}$ and $L = 10$. As a preliminary estimator we used a Gaussian kernel estimator, with a rather large bandwidth.

We came to the conclusion that the adaptive estimator is particularly robust with respect to this preliminary estimator. Also, the choice of L is not of crucial importance. Choosing L adaptively reduces to taking a finer grid on β . The local sharp adaptive estimator behaves uniformly well for many different densities in the described class. Also, this study cleared out the behaviour of this procedure for more general analytic functions, as the classical Gaussian density. This function belongs to each Sobolev class $W(\beta, L)$ (where L gets larger with β) and $\hat{\beta}$ chooses β_{N_n} (for large enough L). The adaptive estimator achieves the rate $n^{-(\beta_{N_n}-1/2)/\beta_{N_n}}$ which is close to $n^{-1/2}$ for large β_{N_n} .

Finally, a major difference with respect to the nonparametric regression or Gaussian white noise model, as considered in Lepski and Spokoiny [28], Tsybakov [30], is the fact that the density model is heteroscedastic. This means more precisely that the variance of the kernel estimator is proportional to $f(x_0)$, the value of the unknown density at the estimation point. In consequence, this value appears into our exact normalization and optimal estimators and hence the use of a preliminary estimator of $f(x_0)$, which has to be free of unknown values $f(x_0)$ and β .

2.2. Statement of results

Let $c = c(\beta, L, q, f(x_0))$ be a positive constant defined by

$$c(\beta, L, q, f(x_0)) = b_\beta L^{\frac{1}{2\beta}} (2\beta)^{\frac{\beta+1/2}{2\beta}} \left(\frac{qf(x_0)}{2\beta-1} \right)^{\frac{\beta-1/2}{2\beta}}. \quad (2.14)$$

Theorem 2.2. *The estimator $f_n^*(x_0)$ defined by (2.13) is rate adaptive estimator and the adaptive rate $\psi_{n,\beta}$ associated to the constant $c(\beta, L, q, f(x_0))$ in (2.14) is such that*

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} R_{n,\beta}(f_n^*, \psi_{n,\beta}) \leq 1 \quad (2.15)$$

and

$$\liminf_{n \rightarrow \infty} \inf_{f_n} \sup_{\beta \in B_-} R_{n,\beta}(f_n, \psi_{n,\beta}) \geq 1. \quad (2.16)$$

Moreover,

$$\limsup_{n \rightarrow \infty} R_{n,\beta_{N_n}}(f_n^*, \psi_{n,\beta_{N_n}}) < \infty. \quad (2.17)$$

This exact constant is obviously similar to the case of Gaussian white noise model in Tsybakov [30]. Nevertheless, as we stressed in our introduction very few results of this type are known in the density estimation model.

The pointwise estimation allows more flexibility than the global adaptation and, in particular, the bandwidth and the kernel can be adjusted locally. Moreover, the adaptive setup in the pointwise estimation is more interesting because the rates are significantly different (slower by a $\log n$ factor) than the nonadaptive rates, except at the last point of the set B_{N_n} . This is usually not happening in global \mathbb{L}_p estimation.

Our theorem states that $\psi_{n,\beta}$ and f_n^* satisfy the exact adaptive problem on B_- :

$$\lim_{n \rightarrow \infty} \inf_{f_n} \sup_{\beta \in B_-} R_{n,\beta}(f_n, \psi_{n,\beta}) = \lim_{n \rightarrow \infty} \sup_{\beta \in B_-} R_{n,\beta}(f_n^*, \psi_{n,\beta}) = 1$$

split in two inequalities (2.15) – also called upper bound and (2.16) – known as the lower bound. The same procedure attains the optimal rate of convergence at the last point of the set, β_{N_n} .

The proof of the upper bounds (2.15) and (2.17) is given in Section 5. The different tools and techniques that are used here are described in Sections 3 and 4. In Section 3, we study the preliminary estimator and give

exact bounds for the bias and the variance of the optimal kernel estimator $f_{n,\beta}$. Section 4 contains the main tools in our proof, stated as theorems.

On one hand, in order to bound uniform risks of the adaptive estimator both exponential and uniform exponential inequalities are needed (Bernstein's and, respectively, van de Geer's inequalities). An original technique is used in Theorem 4.6, where we split the integration domain of the risk and treat each case differently. A global entropy reasoning or chaining would not provide us the right constants. On the other hand, it is necessary to study the estimator of β and to quantify exactly the probability that $\hat{\beta}$ is strictly less than the true β . We see that this event happens with exponentially small probability.

Section 6 gives a constructive proof of the lower bounds. Namely, the subexperiments that are difficult enough and lead to the exact lower bounds are given.

3. AUXILIARY RESULTS

From now on $o(1)$ denotes any sequence that tends to 0 when $n \rightarrow \infty$, $O(1)$ and d_i , $i = 0, 1, 2, \dots$ are positive constants, depending eventually on fixed, positive β_1 , q and L , while c_i , $i = 1, 2, \dots$ are absolute positive constants.

Proof of (2.5) to (2.7). We apply (2.3) and (2.4) and we pass to limits:

$$\frac{\beta_{N_n}}{\log n} = \frac{\beta_{N_n}^2 \log \log n}{\Delta_n \log n} \frac{\Delta_n}{\beta_{N_n} \log \log n} \leq o(1) \limsup_{n \rightarrow \infty} \Delta_n = o(1).$$

Next, for n large enough, $\beta_{N_n} \log^2 \beta_{N_n} \leq \beta_{N_n}^2 \log \beta_{N_n} \leq \beta_{N_n}^2 \log \log n$ and then

$$\left(\frac{\beta_{N_n}}{\log n} \right)^{1/2} \log \beta_{N_n} \leq \left(\frac{\beta_{N_n}^2 \log \log n}{\log n} \right)^{1/2} \leq \left(\Delta_n \frac{\beta_{N_n}^2 \log \log n}{\Delta_n \log n} \right)^{1/2} \xrightarrow{n \rightarrow \infty} 0.$$

Finally, we remark that for n large enough, $\frac{1}{\Delta_n} \leq \log n$ and then $\log \frac{1}{\Delta_n} \leq \log \log n$. We also have $\log \frac{1}{\Delta_n} \leq \frac{1}{\Delta_n}$ and we write $\log \frac{1}{\Delta_n} \leq \left(\frac{\log \log n}{\Delta_n} \right)^{1/2}$. Then

$$\left(\frac{\beta_{N_n}}{\log n} \right)^{1/2} \log \frac{1}{\Delta_n} \leq \left(\frac{1}{\beta_{N_n}} \frac{\beta_{N_n}^2 \log \log n}{\Delta_n \log n} \right)^{1/2} \xrightarrow{n \rightarrow \infty} 0.$$

□

Let us remark that the kernel K_β in (2.12) has the Fourier transform

$$\mathcal{F}(K_\beta)(u) = \frac{1}{1 + |u|^{2\beta}}$$

and by Plancherel formula $\|K_\beta\|_2^2 = \frac{1}{2\pi} \|\mathcal{F}(K_\beta)\|_2^2 = \nu_\beta^2$.

Lemma 3.1. *There exist positive constants K_{\max} , k_{\max} , k_{\min} , ν_{\max} , ν_{\min} and v_{\max} depending only on fixed β_1 , q and L such that:*

1. $\|K_\beta\|_\infty \leq K_{\max}$, $k_{\min} \leq k_\beta \leq k_{\max}$ and $\nu_{\min} \leq \nu_\beta \leq \nu_{\max}$, for any β in B_n ;
2. $\max_{\beta \in B} \sup_{f \in W_n(\beta, L)} \|f\|_\infty \leq v_{\max}$.

Proof. 1. We shall prove only the first statement, the rest being an easy consequence of the definition of k_β and ν_β . We have for $2\beta_1 > 1$:

$$\|K_\beta\|_\infty \leq \frac{1}{2\pi} \left(1 + \int_{|u|>1} \frac{du}{1+|u|^{2\beta_1}} \right) = K_{\max}(\beta_1) < +\infty.$$

2. For $\beta > 1/2$, f in $W_n(\beta, L)$ is a continuous function and

$$|f(x)| = \left| \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}(f)(y) e^{ixy} dy \right| \leq \frac{\nu_\beta}{\sqrt{2\pi}} \left(\int_{\mathbb{R}} |\mathcal{F}(f)(y)|^2 (1+|y|^\beta)^2 dy \right)^{1/2}.$$

We finish the proof by writing for $\mathcal{F}(f)$ continuous that

$$\int_{\mathbb{R}} |\mathcal{F}(f)(y)|^2 (1+|y|^\beta)^2 dy \leq 4 \left(\int_{|y|\leq 1} |\mathcal{F}(f)(y)|^2 dy + \int_{|y|>1} |\mathcal{F}(f)(y)|^2 |y|^{2\beta} dy \right)$$

which is finite. \square

From now on, we suppose that β is fixed, $\beta \in B$, f belongs to the class $W_n(\beta, L)$ and γ is also in B such that $1/2 < \gamma \leq \beta$. Let us define the nonrandom sequences

$$h_{n,\gamma} = k_\gamma \left(\frac{f(x_0) \log n}{n} \right)^{\frac{1}{2\gamma}} \quad \text{and} \quad \eta_{n,\gamma} = \nu_\gamma \sqrt{\frac{q}{2\gamma k_\gamma}} \left(\frac{f(x_0) \log n}{n} \right)^{\frac{\gamma-1/2}{2\gamma}}.$$

Let $\delta_n = 1/\log n$ and let $1/2 < \beta_0 \leq 1$. Then (2.1) and (2.11) entail:

$$\frac{n(h_n \delta_n \rho_n)^2}{\log n} \rightarrow \infty \quad \text{and} \quad \frac{h_n^{\beta_0-1/2}}{\delta_n \rho_n} \rightarrow 0. \quad (3.1)$$

We put for any γ in B

$$\tilde{\gamma} = \begin{cases} \beta, & \text{when } \gamma = \beta \text{ or } \gamma < \beta < 2\gamma \\ \gamma + \frac{\beta_1+1/2}{2}, & \text{when } 2\gamma \leq \beta, \end{cases} \quad (3.2)$$

which satisfies $\tilde{\gamma} < 2\gamma$ and $\gamma < \tilde{\gamma} \leq \beta$. We define the random event

$$A_{n,\gamma} = \left\{ \left| \left(\frac{\hat{h}_{n,\gamma}}{h_{n,\gamma}} \right)^{\tilde{\gamma}-1/2} - 1 \right| \leq \delta_n \right\}$$

and the corresponding nonrandom set

$$\mathcal{H}_{n,\gamma} = \left\{ h : \left| \left(\frac{h}{h_{n,\gamma}} \right)^{\tilde{\gamma}-1/2} - 1 \right| \leq \delta_n \right\}.$$

The following result concerns the preliminary kernel estimator $\hat{f}_n(x_0)$ in (2.10), having the bandwidth h_n that satisfies (2.11) and a bounded, kernel K , such that $\int |u| |K(u)| du < \infty$.

Lemma 3.2. *If $\overline{A}_{n,\gamma}$ denotes the complementary set of $A_{n,\gamma}$, then for sufficiently small, fixed $\alpha > 0$:*

$$P_f [\overline{A}_{n,\gamma}] \leq 2 \exp \left\{ -\frac{n((1-\alpha)h_n\delta_n\rho_n)^2}{2\|K\|_\infty^2} \right\} = o(1).$$

Proof. We use the facts that $|x^\alpha - 1| \leq |x - 1|$ for any $0 < \alpha < 1$, $x > 0$ and $|\max\{x, \rho_n\} - \max\{y, \rho_n\}| \leq |x - y|$ for any fixed n and $x, y > 0$. For $\alpha = \frac{\tilde{\gamma}-1/2}{2\gamma} \leq 1 - \frac{1}{4\beta_1}$ and $f(x_0) \geq \rho_n$, we have

$$\begin{aligned} P_f [\overline{A}_{n,\gamma}] &= P_f \left[\left| \left(\frac{\hat{\rho}_n(x_0)}{f(x_0)} \right)^{\frac{\tilde{\gamma}-1/2}{2\gamma}} - 1 \right| \geq \delta_n \right] \leq P_f \left[\left| \hat{f}_n(x_0) - f(x_0) \right| \geq \delta_n \rho_n \right] \\ &\leq P_f \left[\left| \hat{f}_n(x_0) - E_f \hat{f}_n(x_0) \right| + \left| E_f \hat{f}_n(x_0) - f(x_0) \right| \geq \delta_n \rho_n \right]. \end{aligned}$$

By embedding Theorem 15.1 in Besov *et al.* [2] we deduce that f which belongs to $W_n(\beta, L)$ belongs also to $W_n(\beta_0, L)$ with $\beta_0 = \min\{\beta_1, 1\}$, which is included in the Hölder class $H(\beta_0 - 1/2, L_0)$, $L_0 > 0$. This means f verifies:

$$|f(x) - f(y)| \leq L_0 |x - y|^{\beta_0 - 1/2}.$$

Let us consider a kernel K such that $\int |u| |K(u)| du < \infty$. Then, we have:

$$\begin{aligned} \left| E_f \hat{f}_n(x_0) - f(x_0) \right| &\leq \int |K(x)| |f(x_0 + h_n x) - f(x_0)| dx \\ &\leq L_0 h_n^{\beta_0 - 1/2} \left(\int_{|x| \leq 1} |K(x)| |x|^{\beta_0 - 1/2} dx + \int_{|x| > 1} |K(x)| |x| dx \right) \\ &\leq c(L_0, \beta_0) h_n^{\beta_0 - 1/2}, \end{aligned}$$

with a constant $c(L_0, \beta_0) > 0$. The last term of the inequality above is equal to $o(\delta_n \rho_n)$ by (3.1). We obtain, for fixed, small $\alpha > 0$

$$P_f [\overline{A}_{n,\gamma}] \leq P_f \left[\left| \hat{f}_n(x_0) - E_f \hat{f}_n(x_0) \right| \geq (1 - \alpha) \delta_n \rho_n \right].$$

We apply Hoeffding's inequality for the i.i.d. variables $K((X_i - x_0)/h_n)$, bounded by $\|K\|_\infty < \infty$:

$$P_f \left[\left| \frac{1}{n} \sum_{i=1}^n \left(K \left(\frac{X_i - x_0}{h_n} \right) - E_f K \left(\frac{X_i - x_0}{h_n} \right) \right) \right| \geq z \right] \leq 2 \exp \left\{ -\frac{nz^2}{2\|K\|_\infty^2} \right\}$$

for any $z \geq 0$, $f \in W_n(\beta, L)$. It suffices to take $z = (1 - \alpha) h_n \delta_n \rho_n$ in order to get the stated result. Moreover, $n(h_n \delta_n \rho_n)^2 \rightarrow \infty$ with n , by (3.1), then the right-hand side in the lemma is $o(1)$. \square

Denote $f_{n,\gamma}(x_0, h) = \frac{1}{nh} \sum_{i=1}^n K_\gamma \left(\frac{X_i - x_0}{h} \right)$ the kernel estimator of $f(x_0)$, having kernel K_γ and bandwidth h in the set $\mathcal{H}_{n,\gamma}$. In particular, we denote $f_{n,\gamma}(x_0) = f_{n,\gamma}(x_0, h_{n,\gamma})$.

We study in the following this kernel estimator using the classical decomposition

$$|f_{n,\gamma}(x_0, h) - f(x_0)| \leq B_{n,\gamma}(h) + Z_{n,\gamma}(h),$$

where $B_{n,\gamma}(h) = |E_f f_{n,\gamma}(x_0, h) - f(x_0)|$ is the bias of the estimator $f_{n,\gamma}(x_0, h)$ of $f(x_0)$, for a bandwidth $h \in \mathcal{H}_{n,\gamma}$ and $Z_{n,\gamma}(h) = |f_{n,\gamma}(x_0, h) - E_f f_{n,\gamma}(x_0, h)|$ its stochastic term.

Lemma 3.3. For any h in $\mathcal{H}_{n,\gamma}$, $f \in W_n(\beta, L)$ and γ, β in B_{N_n} such that $1/2 < \gamma \leq \beta$, let $\tilde{L} > 0$ be some constant, $b_{n,\gamma}(h) = \begin{cases} Lb_\beta h^{\beta-\frac{1}{2}}, & \gamma = \beta \\ \tilde{L}\nu_{\max} h^{\tilde{\gamma}-\frac{1}{2}}, & \gamma < \beta \end{cases}$, where $\tilde{\gamma}$ is defined by (3.2) and $s_{n,\gamma}^2(h) = f(x_0)\nu_\gamma^2/(nh)$.

(In particular, we denote $b_{n,\gamma} = b_{n,\gamma}(h_{n,\gamma})$ and $s_{n,\gamma} = s_{n,\gamma}(h_{n,\gamma})$). We have:

1. $B_{n,\gamma}(h) \leq b_{n,\gamma}(h)$;
2. $E_f [Z_{n,\gamma}^2(h)] \leq s_{n,\gamma}^2(h)(1 + o(1))$, where $o(1) \xrightarrow{n \rightarrow \infty} 0$ uniformly in γ and β .

Proof. 1) We have $\mathcal{F}\left(\frac{1}{h}K_\gamma\left(\frac{\cdot-x_0}{h}\right)\right)(x) = e^{ixx_0}\mathcal{F}(K_\gamma)(hx)$. Then

$$\begin{aligned} B_{n,\gamma}(h) &= \frac{1}{2\pi} \left| \int_{\mathbb{R}} \mathcal{F}(f)(x) e^{ixx_0} [\mathcal{F}(K_\gamma)(hx) - 1] dx \right| \leq \frac{1}{2\pi} \int_{\mathbb{R}} |\mathcal{F}(f)(x)| \frac{|hx|^{2\gamma}}{1 + |hx|^{2\gamma}} dx \\ &\leq \frac{1}{\sqrt{2\pi}} \left(\int_{\mathbb{R}} \frac{1}{2\pi} |\mathcal{F}(f)(x)|^2 |x|^{2\tilde{\gamma}} dx \right)^{1/2} \left(\int_{\mathbb{R}} \frac{h^{2\tilde{\gamma}} |hx|^{2(2\gamma-\tilde{\gamma})}}{(1 + |hx|^{2\gamma})^2} dx \right)^{1/2}. \end{aligned}$$

by the Cauchy–Schwarz inequality.

- a) When $\gamma = \beta$ we have $\tilde{\gamma} = \beta$ and

$$B_{n,\beta}(h) \leq \frac{Lh^{\beta-\frac{1}{2}}}{\sqrt{2\pi}} \left(\int_{\mathbb{R}} \frac{|u|^{2\beta}}{(1 + |u|^{2\beta})^2} du \right)^{1/2} = b_{n,\beta}(h).$$

- b) When $\gamma < \beta$, we have $\tilde{\gamma}$ such that $\gamma < \tilde{\gamma} < 2\gamma$ and $\tilde{\gamma} \leq \beta$. Then $f \in W_n(\beta, L) \subseteq W_n(\tilde{\gamma}, \tilde{L})$, for some fixed constant $\tilde{L} > 0$, by embedding theorem of Sobolev spaces. We get

$$B_{n,\gamma}(h) \leq \frac{\tilde{L}h^{\tilde{\gamma}-\frac{1}{2}}}{\sqrt{2\pi}} \left(\int_{\mathbb{R}} \frac{|u|^{2(2\gamma-\tilde{\gamma})}}{(1 + |u|^{2\gamma})^2} du \right)^{1/2} \leq b_{n,\gamma}(h).$$

- 2) We have $E_f [Z_{n,\gamma}^2(h)] \leq \frac{1}{nh} \int \frac{1}{h} K_\gamma^2\left(\frac{x-x_0}{h}\right) f(x) dx$ and we conclude by Bochner's lemma applied to continuous f and kernel K_γ , that

$$\int \frac{1}{h} K_\gamma^2\left(\frac{x-x_0}{h}\right) f(x) dx \leq f(x_0) \|K_\gamma\|_2^2 (1 + o(1)).$$

□

We give rough non-random upper bounds concerning the kernel estimator $\hat{f}_{n,\gamma}(x_0)$ with bandwidth $\hat{h}_{n,\gamma}$ depending on random data X_1, \dots, X_n .

Lemma 3.4. For $\gamma < \beta$, there exist constants d_0 and d_1 , such that, a.s.

$$\sup_{\beta \in B} \sup_{f \in W_n(\beta, L)} \psi_{n,\beta}^{-q} \left(\hat{\eta}_{n,\beta} + \left| \hat{f}_{n,\beta}(x_0) - f(x_0) \right| \right)^q \leq 2d_0 n^{d_1}$$

and

$$\sup_{\beta \in B} \sup_{f \in W_n(\beta, L)} \psi_{n,\beta}^{-q} \left| \hat{f}_{n,\gamma}(x_0) - f(x_0) \right|^q \leq d_0 n^{d_1}.$$

Proof. Let us bound at first the preliminary estimator $|\widehat{f}_n(x_0)| \leq \|K_0\|_\infty / h_n$. Then, a.s.

$$|\widehat{f}_{n,\gamma}(x_0)| \leq \frac{\|K_\gamma\|_\infty}{\widehat{h}_{n,\gamma}} \leq \frac{K_{\max}}{k_{\min}} \left(\frac{n}{\widehat{\rho}_n(x_0) \log n} \right)^{\frac{1}{2\gamma}} \leq O(1) n^{\frac{1}{2\beta_1}}.$$

At last, for β in B_-

$$\psi_{n,\beta}^{-1} \leq \frac{\sqrt{2\beta-1}}{2\beta L \nu_\beta k_\beta^\beta} \sqrt{k_\beta} \left(\frac{n}{f(x_0) \log n} \right)^{\frac{1}{2}-\frac{1}{4\beta}} \leq O(1) \sqrt{\beta_{N_n}} n^{\frac{1}{2}-\frac{1}{4\beta_1}}$$

and the same holds for $\beta = \beta_{N_n}$, $\psi_{n,\beta_{N_n}}^{-1} \leq O(1) n^{1/2-1/(4\beta_1)}$.

We may conclude that a.s., for any $1/2 < \gamma \leq \beta$:

$$\psi_{n,\beta}^{-1} |\widehat{f}_{n,\gamma}(x_0) - f(x_0)| \leq O(1) \sqrt{\beta_{N_n}} \left(n^{\frac{1}{2}+\frac{1}{4\beta_1}} + n^{\frac{1}{2}-\frac{1}{4\beta_1}} \right) \leq d_0 n^{d_1},$$

since $\beta_{N_n} = o(1)$. Similarly,

$$\frac{\widehat{\eta}_{n,\beta}}{\psi_{n,\beta}} = \frac{\eta_{n,\beta}}{\psi_{n,\beta}} \cdot \left(\frac{\widehat{\rho}_n(x_0)}{f(x_0)} \right)^{\frac{(\beta-1/2)}{2\beta}} \leq \left(1 - \frac{1}{2\beta} \right) \left(\frac{\widehat{\rho}_n}{\rho_n} \right)^{\frac{1}{2}-\frac{1}{4\beta}} \leq d_0 n^{d_1}.$$

□

Let us choose a convenient sequence which shall appear in the proof of the upper bounds:

$$\tau_{n,\gamma} = s_{n,\gamma} \left[\left[\frac{q}{2} \left(\frac{1}{\gamma} - \frac{1}{\beta} \right) \log n \right]^{\frac{1}{2}} + \left(\frac{\log n}{\beta_{N_n}} \right)^{\frac{1}{4}} \right], \text{ for any } \gamma \leq \beta.$$

Remark that $\tau_{n,\beta}/s_{n,\beta} = (\log n/\beta_{N_n})^{1/4} \rightarrow \infty$, when $n \rightarrow \infty$.

Lemma 3.5. 1. The following inequalities hold for any $\beta \in B_n$ and $f \in W_n(\beta, L)$

$$\begin{aligned} \frac{s_{n,\beta}}{\psi_{n,\beta}} &\leq \sqrt{\frac{2}{q}} \left(\frac{\beta_{N_n}}{\log n} \right)^{1/2}, \quad \frac{\tau_{n,\beta}}{\psi_{n,\beta}} \leq \sqrt{\frac{2}{q}} \left(\frac{\beta_{N_n}}{\log n} \right)^{\frac{1}{4}}, \\ \left(\frac{\log n}{n} \right)^{\frac{\Delta_n}{\beta_{N_n}^2}} &\leq \exp \left\{ -\frac{\Delta_n \log n}{\beta_{N_n}^2 \log \log n} \right\} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

2. If $\gamma < \beta$ are in B_n we have

$$\sup_{f \in W_n(\beta, L)} \frac{b_{n,\gamma}^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \leq O(1) (\log n)^{q/2} n^{q(\frac{1}{4\gamma} - \frac{1}{4\beta})}.$$

Proof. 1. It suffices to note that $\eta_{n,\beta}^2 = \left(\frac{2\beta-1}{2\beta} \right)^2 \psi_{n,\beta}^2 = \frac{q}{2\beta} s_{n,\beta}^2 \log n$ and use (2.4) for the limit.

2. First, if β is in B_- , we have

$$\psi_{n,\beta}^{-1} \leq O(1) \sqrt{\beta_N} \left(\frac{n}{f(x_0) \log n} \right)^{\frac{1}{2}-\frac{1}{4\beta}} \leq O(1) \sqrt{\log n} \left(\frac{n}{f(x_0) \log n} \right)^{\frac{1}{2}-\frac{1}{4\beta}}.$$

We see as well that

$$\tau_{n,\gamma} \leq s_{n,\gamma} \sqrt{\frac{q \log n}{2\gamma}} (1 + o(1)) \leq O(1) \left(\frac{f(x_0) \log n}{n} \right)^{\frac{1}{2} - \frac{1}{4\gamma}}$$

and also

$$b_{n,\gamma} \leq c_b h_{n,\gamma}^{\tilde{\gamma} - \frac{1}{2}} \leq O(1) \left(\frac{f(x_0) \log n}{n} \right)^{\frac{1}{2} - \frac{1}{4\gamma}}$$

as $\gamma \leq \tilde{\gamma} < 2\gamma$ and $k_{\min} \leq k_\gamma \leq k_{\max}$. Moreover, $f(x_0) \log n \geq \rho_n \log n > 0$, then

$$\sup_{f \in W_n(\beta, L)} \frac{b_{n,\gamma}^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \leq O(1) (\log n)^{q/2} n^{q(\frac{1}{4\gamma} - \frac{1}{4\beta})}.$$

If $\beta = \beta_{N_n}$, the result is immediate. \square

4. EXPONENTIAL INEQUALITIES AND APPLICATIONS

The key lemmas in the proof of the upper bounds are given here and the results as we apply them in the next section are stated as theorems.

4.1. Exponential bounds

The next proposition recalls two inequalities on empirical processes. Denote the empirical distribution associated to the i.i.d. observations X_1, \dots, X_n of common law P by $\mathbb{P}_n = 1/n \sum_{i=1}^n \delta_{X_i}$.

Proposition 4.1. *Let us consider a class of functions $\mathcal{K} = \{f_h(\cdot) / h \in H\}$ satisfying*

$$\sup_{h \in H} \|f_h\|_\infty \leq K \text{ and } \sup_{h \in H} \|f_h\|_{\mathbb{L}_2(P)} \leq M$$

for some positive K and M .

Then for all $u > 0$ and for fixed $h \in H$

$$P_f \left[\left| \int f_h d(\mathbb{P}_n - P) \right| \geq u \right] \leq 2 \exp \left\{ -\frac{nu^2}{2(M^2 + 2uK)} \right\}$$

(Bernstein's Inequality see Pollard [29]);

Moreover, if

$$u \leq \min \left\{ 8M, \frac{2C_1 M^2}{K} \right\} \text{ and } u \geq \frac{C_0}{\sqrt{n}} \int_{\frac{u}{2^6}}^M \max \left\{ H_B^{1/2}(\mathcal{K}, x), 1 \right\} dx,$$

where H_B denotes the entropy with bracketing with respect to the \mathbb{L}_2 -norm of a class of functions, then there exists some $a < 1/2$ such that

$$P_f \left[\sup_{h \in H} \left| \int f_h d(\mathbb{P}_n - P) \right| \geq u \right] \leq 8 \exp \left\{ -a \frac{nu^2}{M^2} \right\}$$

(Van de Geer [31]).

One can prove without difficulty that there exists a positive constant $d_2 = 1/(\beta_1 - 1/2)$, such that:

$$\mathcal{H}_{n,\gamma} \subseteq \left\{ h : \left| \frac{h}{h_{n,\gamma}} - 1 \right| \leq d_2 \delta_n \right\} = H_{n,\gamma}. \quad (4.1)$$

We shall apply the previous results, successively, to two classes of measurable functions, for observations having distribution P_f , corresponding to the probability density $f \in W_n(\beta, L)$ and empirical distribution \mathbb{P}_n . Let us take on one hand

$$\mathcal{K} = \left\{ K_{\gamma,h}(\cdot) = \frac{1}{h} K_\gamma \left(\frac{\cdot - x_0}{h} \right) : h \in H_{n,\gamma} \right\}.$$

Then

$$\sup_{h \in H_{n,\gamma}} \|K_{\gamma,h}\|_\infty \leq \frac{\|K_\gamma\|_\infty}{h_{n,\gamma}} \leq \frac{K_{\max}}{h_{n,\gamma}}$$

and by Lemma 3.3

$$\sup_{h \in H_{n,\gamma}} \|K_{\gamma,h}\|_{\mathbb{L}_2(P_f)} \leq \frac{\nu_\gamma \sqrt{f(x_0)}}{\sqrt{h_{n,\gamma}}} = s_{n,\gamma} \sqrt{n}.$$

For different h_1, h_2 in $H_{n,\gamma}$ we have

$$\begin{aligned} \|K_{\gamma,h_1} - K_{\gamma,h_2}\|_{\mathbb{L}_2(P_f)} &\leq \frac{1}{\sqrt{2\pi}} \|\mathcal{F}(K_\gamma)(h_1 \cdot) - \mathcal{F}(K_\gamma)(h_2 \cdot)\|_2 \\ &\leq \frac{1}{\sqrt{2\pi}} \left(\int \frac{|h_1^{2\gamma} - h_2^{2\gamma}|^2 |y|^{4\gamma} dy}{(1 + |h_1 y|^{2\gamma})^2 (1 + |h_2 y|^{2\gamma})^2} \right)^{1/2} \\ &\leq \frac{O(1)}{\sqrt{h_{n,\gamma}}} \left| 1 - \left(\frac{h_1}{h_2} \right)^{2\gamma} \right|. \end{aligned}$$

Then, we may write that the ε -entropy with bracketing of the class \mathcal{K} with respect to the $\mathbb{L}_2(P_f)$ norm is: $H_B(\mathcal{K}, \varepsilon) \leq O(\log \frac{1}{\varepsilon} + \log n)$. Thus

$$\frac{C_0}{\sqrt{n}} \int_0^M \max(H_B^{1/2}(\mathcal{K}, x), 1) dx \leq M \sqrt{\frac{\log n}{n}}. \quad (4.2)$$

On the other hand, consider the following class of measurable functions

$$\mathcal{K}(h_{n,\gamma}) = \{K_{\gamma,h}(\cdot) - K_{\gamma,h_{n,\gamma}}(\cdot) : h \in H_{n,\gamma}\}$$

for which

$$\begin{aligned}
\sup_{h \in H_{n,\gamma}} \|K_{\gamma,h} - K_{\gamma,h_{n,\gamma}}\|_{\infty} &\leq \sup_{h \in H_{n,\gamma}} \sup_x \frac{1}{\sqrt{2\pi}} \left| \int \mathcal{F}(K_{\gamma,h} - K_{\gamma,h_{n,\gamma}})(y) e^{ixy} dy \right| \\
&\leq \sup_{h \in H_{n,\gamma}} \frac{1}{\sqrt{2\pi}} \int \frac{|h^{2\gamma} - h_{n,\gamma}^{2\gamma}| |y|^{2\gamma} dy}{(1 + |hy|^{2\gamma})(1 + |h_{n,\gamma}y|^{2\gamma})} \\
&\leq \frac{O(1)}{h_{n,\gamma}} \left| 1 - \left(\frac{h}{h_{n,\gamma}} \right)^{2\gamma} \right| \leq O(1) \frac{\beta_N \delta_n}{h_{n,\gamma}}
\end{aligned}$$

and

$$\sup_{h \in H_{n,\gamma}} \|K_{\gamma,h} - K_{\gamma,h_{n,\gamma}}\|_{L_2(P_f)} \leq \sup_{h \in H_{n,\gamma}} \frac{O(1)}{\sqrt{h_{n,\gamma}}} \left| 1 - \left(\frac{h}{h_{n,\gamma}} \right)^{2\gamma} \right| \leq O(1) \frac{\beta_N \delta_n}{\sqrt{h_{n,\gamma}}}. \quad (4.3)$$

It is easy to see, that (4.2) still holds for this class.

Lemma 4.2. *For the kernel estimator $f_{n,\gamma}(x_0, h)$, for any fixed x_0 and for $u \leq c_1 s_{n,\gamma} \sqrt{\log n}$ ($c_1 > 0$, an absolute constant), there exists $\delta_{n0} > 0$ small enough, independent of γ , such that:*

$$P_f [|Z_{n,\gamma}(h_{n,\gamma})| \geq u] \leq 2 \exp \left\{ -\frac{u^2}{2s_{n,\gamma}^2} (1 - \delta_{n0}) \right\}.$$

Proof. Write $Z_{n,\gamma}(h) = \int K_{\gamma,h} d(\mathbb{P}_n - P_f)$ and apply Bernstein's inequality of Proposition 4.1 for $K = K_{\max}/h_{n,\gamma}$ and $M = s_{n,\gamma} \sqrt{n}$ to get

$$P_f [|Z_{n,\gamma}(h_{n,\gamma})| \geq u] \leq 2 \exp \left\{ -\frac{u^2}{2s_{n,\gamma}^2 (1 + 2uK_{\max}/\nu_\gamma^2 f(x_0))} \right\}.$$

We remark that for $u \leq c_1 s_{n,\gamma} \sqrt{\log n}$ ($c_1 > 0$) we have

$$\frac{2uK_{\max}}{\nu_\gamma^2 f(x_0)} \leq \frac{2c_1 K_{\max} \sqrt{\log n}}{\nu_\gamma \sqrt{f(x_0)} \sqrt{n} h_{n,\gamma}} \leq O(1) \left(\frac{\rho_n \log n}{n} \right)^{\frac{1}{2} - \frac{1}{4\gamma}} = o(1)$$

and thus we can find a $\delta_{n0} > 0$ small enough, such that

$$\left(1 + \frac{2uK_{\max}}{\nu_\gamma^2 f(x_0)} \right) \leq \frac{1}{1 - \delta_{n0}}$$

which we can replace in the exponential inequality above in order to get the stated lemma. \square

Lemma 4.3. *For u , such that $\tau_{n,\gamma} \leq u \leq c_1 s_{n,\gamma} \sqrt{\log n}$, there exist a constant $d_3 > 0$ and $\delta_{n1} > 0$ sufficiently small, independent of γ , such that*

$$P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] \leq d_3 \exp \left\{ -\frac{u^2 (1 - \delta_{n1})^2}{2s_{n,\gamma}^2} (1 - \delta_{n0}) \right\},$$

for n large enough, where δ_{n0} is the sequence in Lemma 4.2.

Proof. Let us consider the sequence $\delta_{n1} = \beta_{N_n} \delta_n \sqrt{\beta_{N_n} \log n} \xrightarrow{n \rightarrow \infty} 0$. We have

$$P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] \leq P_f [|Z_{n,\gamma}(h_{n,\gamma})| \geq u(1 - \delta_{n1})] \\ + P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h) - Z_{n,\gamma}(h_{n,\gamma})| \geq u\delta_{n1} \right].$$

We apply Lemma 4.2 to the first term on the right-hand side of the previous inequality. We see that $u(1 - \delta_{n1}) \leq u \leq c_1 s_{n,\gamma} \sqrt{\log n}$ and thus there exist a sufficiently small $\delta_{n0} > 0$ such that:

$$P_f [|Z_{n,\gamma}(h_{n,\gamma})| \geq u(1 - \delta_{n1})] \leq 2 \exp \left\{ -\frac{u^2 (1 - \delta_{n1})^2}{2s_{n,\gamma}^2} (1 - \delta_{n0}) \right\}.$$

For the second term we first see that

$$Z_{n,\gamma}(h) - Z_{n,\gamma}(h_{n,\gamma}) = \int (K_{\gamma,h} - K_{\gamma,h_{n,\gamma}}) d(\mathbb{P}_n - P_f).$$

Then we apply Van de Geer's inequality from Proposition 4.1, where $K = \beta_{N_n} \delta_n / h_{n,\gamma}$ and $M = \beta_{N_n} \delta_n / \sqrt{h_{n,\gamma}}$

$$P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h) - Z_{n,\gamma}(h_{n,\gamma})| \geq u\delta_{n1} \right] \leq 8 \exp \left\{ -anh_{n,\gamma} \frac{(u\delta_{n1})^2}{(\beta_{N_n} \delta_n)^2} \right\}$$

for all u in $[\tau_{n,\gamma}, c_1 s_{n,\gamma} \sqrt{\log n}]$. Indeed, $\min \left\{ 8M, \frac{2C_1 M^2}{K} \right\} = O(1) \beta_{N_n} \delta_n$,

$$\frac{C_0}{\sqrt{n}} \int_0^M \max \left\{ H_B^{1/2}(\mathcal{K}, x), 1 \right\} dx \leq O(1) \frac{\beta_{N_n} \delta_n}{\sqrt{nh_{n,\gamma}}} \sqrt{\log n}$$

and

$$[\tau_{n,\gamma} \delta_{n1}, c_1 \delta_{n1} s_{n,\gamma} \sqrt{\log n}] \subseteq \left[\frac{\beta_{N_n} \delta_n}{\sqrt{nh_{n,\gamma}}} \sqrt{\log n}, \beta_{N_n} \delta_n \right],$$

for n large enough. By summing the two terms again, we get

$$P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] \leq 2 \exp \left\{ -\frac{u^2 (1 - \delta_{n1})^2}{2s_{n,\gamma}^2} (1 - \delta_{n0}) \right\} + 8 \exp \left\{ -anh_{n,\gamma} \frac{(u\delta_{n1})^2}{(\beta_{N_n} \delta_n)^2} \right\} \\ \leq 8(1 + o(1)) \exp \left\{ -\frac{u^2 (1 - \delta_{n1})^2}{2s_{n,\gamma}^2} (1 - \delta_{n0}) \right\}. \quad \square$$

Theorem 4.4. *We have*

$$\sup_{\beta \in B_n} \sup_{f \in W_n(\beta, L)} P_f \left[\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \geq \tau_{n,\beta} \right] \xrightarrow{n \rightarrow \infty} 0.$$

Proof. Let us apply the Lemma 4.3, for $\gamma = \beta$, $u = \tau_{n,\beta}$ and $\delta_{n2} \xrightarrow{n \rightarrow \infty} 0$ such that $(1 - \delta_{n1})^2 (1 - \delta_{n0}) = 1 - \delta_{n2}$

$$P_f \left[\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \geq \tau_{n,\beta} \right] \leq d_3 \exp \left\{ -\frac{\tau_{n,\beta}^2}{2s_{n,\beta}^2} (1 - \delta_{n2}) \right\} \leq d_3 \exp \left\{ -\frac{1}{2} \left(\frac{\log n}{\beta_{N_n}} \right)^{\frac{1}{2}} (1 - \delta_{n2}) \right\}$$

and this sequence tends to 0, by Lemma 3.5, uniformly in $f \in W_n(\beta, L)$ and $\beta \in B_n$. \square

Lemma 4.5. *Let $c_2 > 0$ be a constant such that $c_2 \geq c_1 s_{n,\gamma} \sqrt{\log n}$. For n large enough and $c_1 s_{n,\gamma} \sqrt{\log n} \leq u \leq c_2$, we have that*

$$P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] \leq 8 \exp \left\{ -a \frac{u^2}{s_{n,\gamma}^2} \right\}.$$

Proof. Let us note that $Z_{n,\gamma}(h) = \int K_{\gamma,h} d(\mathbb{P}_n - P_f)$ and apply Van de Geer's inequality of Proposition 4.1, for $K = K_{\max}/h_{n,\gamma}$ and $M = s_{n,\gamma} \sqrt{\log n}$. Indeed,

$$\min \left\{ 8M, \frac{2C_1 M^2}{K} \right\} = \frac{2C_1 M^2}{K} \leq c_2$$

for c_2 large enough, and

$$\frac{C_0}{\sqrt{n}} \int_0^M \max \left\{ H_B^{1/2}(\mathcal{K}, x), 1 \right\} dx \leq C_0 \frac{M \sqrt{\log n}}{\sqrt{n}} \leq c_1 s_{n,\gamma} \sqrt{\log n}.$$

The hypothesis are verified, then

$$P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] \leq 8 \exp \left\{ -a \frac{nu^2}{M^2} \right\} \leq 8 \exp \left\{ -a \frac{u^2}{s_{n,\gamma}^2} \right\}.$$

\square

Theorem 4.6. *We have*

$$\sup_{\substack{\gamma \in B_n \\ \gamma < \beta}} \frac{\beta_{N_n}}{\Delta_n} \sup_{f \in W_n(\beta, L)} \psi_{n,\beta}^{-q} E_f \left[\left(\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \right)^q I \left(\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq \tau_{n,\gamma} \right) \right] \xrightarrow{n \rightarrow \infty} 0,$$

uniformly in $\beta \in B_n$. In particular, when $\gamma = \beta$, we have

$$\sup_{f \in W_n(\beta, L)} \psi_{n,\beta}^{-q} E_f \left[\left(\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \right)^q I \left(\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \geq \tau_{n,\beta} \right) \right] \xrightarrow{n \rightarrow \infty} 0,$$

uniformly in $\beta \in B_n$.

Proof. Let us write the expectation as follows:

$$E_f \left[\left(\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \right)^q I \left(\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq \tau_{n,\gamma} \right) \right] = \int_{\tau_{n,\gamma}}^{c_3/h_{n,\gamma}} P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] du^q.$$

Indeed, we can bound $\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \leq c_3/h_{n,\gamma}$ for some $c_3 > 0$ and it is useless to integrate further on.

We split the integration domain into three intervals: $I = [\tau_{n,\gamma}, c_1 s_{n,\gamma} \sqrt{\log n}]$, $II = [c_1 s_{n,\gamma} \sqrt{\log n}, c_2]$ and $III = [c_2, c_3/h_{n,\gamma}]$, which are non-empty for n large enough. Let us apply Lemma 4.3 on the interval I and get

$$\begin{aligned} \int_I P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] du^q &\leq d_3 \int_I \exp \left\{ -\frac{u^2}{2s_{n,\gamma}^2} (1 - \delta_{n2}) \right\} du^q \\ &\leq O(1) s_{n,\gamma}^q \exp \left\{ -\frac{\tau_{n,\gamma}^2}{2s_{n,\gamma}^2} (1 - \delta_{n2}) \right\}. \end{aligned}$$

Now

$$\begin{aligned} \psi_{n,\beta}^{-q} \int_I &\leq O(1) \frac{s_{n,\gamma}^q}{\psi_{n,\beta}^q} \exp \left\{ -\left(\frac{q}{4} \left(\frac{1}{\gamma} - \frac{1}{\beta} \right) \log n + \frac{1}{2} \sqrt{\frac{\log n}{\beta N_n}} \right) (1 - \delta_{n2}) \right\} \\ &\leq O(1) \exp \left\{ -\frac{1}{2} \sqrt{\frac{\log n}{\beta N_n}} (1 - \delta_{n2}) \right\} = o(1). \end{aligned}$$

On the second interval II , we write

$$\begin{aligned} \int_{II} P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] du^q &\leq \int_{II} P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq c_1 s_{n,\gamma} \sqrt{\log n} \right] du^q \\ &\leq c_2^q P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq c_1 s_{n,\gamma} \sqrt{\log n} \right] \\ &\leq 8c_2^q \exp \{ -ac_1^2 \log n \} \end{aligned}$$

by Lemma 4.5 and

$$\begin{aligned} \psi_{n,\beta}^{-q} \int_{II} &\leq O(1) n^{\frac{q}{2} - \frac{q}{4\beta}} \exp \{ -ac_1^2 \log n \} \\ &\leq O(1) \exp \left\{ -\log n \left(ac_1^2 - \frac{q}{2} + \frac{q}{4\beta} \right) \right\} \\ &\leq O(1) \exp \left\{ -q \frac{\log n}{4\beta N_n} - \log n \left(ac_1^2 - \frac{q}{2} \right) \right\} = o(1) \end{aligned}$$

by choosing c_1 sufficiently large. Similarly,

$$\begin{aligned} \int_{III} P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq u \right] du^q &\leq \int_{III} P_f \left[\sup_{h \in H_{n,\gamma}} |Z_{n,\gamma}(h)| \geq c_2 \right] du^q \\ &\leq \frac{K_{\max}^q}{s_{n,\gamma}^q} 8 \exp \left\{ -a \frac{c_2^2}{s_{n,\gamma}^2} \right\} \\ &\leq \frac{O(1)}{s_{n,\gamma}^q} \exp \left\{ -O(1) n^{1 - \frac{1}{2\beta_1}} \right\}. \end{aligned}$$

Now,

$$\begin{aligned} \psi_{n,\beta}^{-q} \int_{III} &\leq O(1) n^{q - \frac{q}{4\beta} - \frac{q}{4\gamma}} \exp \left\{ -O(1) n^{1 - \frac{1}{2\beta_1}} \right\} \\ &\leq O(1) \exp \left\{ -O(1) n^{1 - \frac{1}{2\beta_1}} + q \left(1 - \frac{1}{2\beta_{N_n}} \right) \log n \right\} = o(1). \end{aligned}$$

To finish the proof it suffices to remember relations (2.5) to (2.7) and see that the three sequences do not depend on f in $W_n(\beta, L)$, nor on γ and β in B_n . \square

4.2. Probability of undershooting

We want to bound here the probability that the estimated value $\hat{\beta}$ is strictly less than the true value β i.e. $P_f \left[\hat{\beta} = \gamma \right]$ for γ in $(1/2, \beta)$ for the worst density in $W_n(\beta, L)$ and uniformly in β over B_n . We show in fact that $\hat{\beta}$ typically overestimates the true value β and that the undershooting is a rare event.

Lemma 4.7. *Let us choose the sequence $\delta_{n3} = \exp \left\{ -\frac{1}{8} \cdot \frac{\Delta_n \log n}{\beta_N^2} \right\}$ which tends to 0, by condition (2.4). Then for any γ_0, γ two elements of the set B_- such that $\gamma_0 \leq \gamma < \beta$, we have for $\gamma_1 = \min \{ \alpha \in B_{N_n} : \alpha > \gamma \}$, $\gamma < \gamma_1 \leq \beta$ and*

$$\frac{b_{n,\gamma_0}}{\delta_{n3}\eta_{n,\gamma_0}} \xrightarrow{n \rightarrow \infty} 0, \quad \frac{b_{n,\gamma_1}}{\delta_{n3}\eta_{n,\gamma_0}} \xrightarrow{n \rightarrow \infty} 0, \quad \frac{\delta_{n3}s_{n,\gamma_0}}{s_{n,\gamma_1}} \xrightarrow{n \rightarrow \infty} \infty.$$

Proof. We see that for $\gamma_0 < \gamma_1$ in B_- : $h_{n,\gamma_0}/h_{n,\gamma_1} \leq O(1) (\log n/n)^{\frac{\Delta_n}{2\beta_N}} \leq o(1)$ by Lemma (3.5). We write

$$\begin{aligned} \frac{b_{n,\gamma_0}}{\delta_{n3}\eta_{n,\gamma_0}} &= \sqrt{\frac{2\gamma_0}{q \log n}} \frac{c_b h_{n,\gamma_0}^{\tilde{\gamma}_0 - \frac{1}{2}}}{\delta_{n3}s_{n,\gamma_0}} \leq O(1) \sqrt{\beta_N} \frac{h_{n,\gamma_0}^{\tilde{\gamma}_0 - \gamma_0}}{\delta_{n3}} \\ &\leq O(1) \sqrt{\beta_N} \exp \left\{ \frac{1}{8} \frac{\Delta_n \log n}{\beta_N^2} - \frac{(\tilde{\gamma}_0 - \gamma_0) \log n}{2\beta_N} \right\}. \end{aligned}$$

By definition of $\tilde{\gamma}_0$, see (3.2), $\tilde{\gamma}_0 - \gamma_0 > \min \{ \beta - \gamma, \frac{1}{2} \} \geq \min \{ \Delta_n, \frac{1}{2} \} > 0$. Then

$$\frac{b_{n,\gamma_0}}{\delta_{n3}\eta_{n,\gamma_0}} \leq O(1) \exp \left\{ -\frac{\log n}{2\beta_N} \left(\min \left\{ \Delta_n, \frac{1}{2} \right\} - \frac{\Delta_n}{4\beta_N} - \frac{\log \beta_N}{2} \frac{\beta_N}{\log n} \right) \right\}$$

which is an $o(1)$. It is easy to see now, that

$$\frac{b_{n,\gamma_1}}{\delta_{n3}\eta_{n,\gamma_0}} = \sqrt{\frac{2\gamma_0}{q \log n}} \frac{c_b h_{n,\gamma_1}^{\tilde{\gamma}_1 - \frac{1}{2}}}{\delta_{n3}s_{n,\gamma_0}} \leq O(1) \frac{\sqrt{\beta_N}}{\delta_{n3}} \sqrt{\frac{h_{n,\gamma_0}}{h_{n,\gamma_1}}} \leq O(1) \exp \left\{ -\frac{1}{4} \frac{\Delta_n \log n}{\beta_N^2} \left(1 - \frac{4 \log \beta_N}{\Delta_n \log n} \beta_N^2 \right) \right\}$$

which tends to 0, for $\tilde{\gamma}_1$ defined as in (3.2).

We conclude with the third limit, using also Lemma (3.5)

$$\begin{aligned} \frac{\delta_{n3} s_{n,\gamma_0}}{s_{n,\gamma_1}} &= \delta_{n3} \frac{\nu_{\gamma_0}}{\nu_{\gamma_1}} \sqrt{\frac{h_{n,\gamma_1}}{h_{n,\gamma_0}}} \geq O(\delta_{n3}) n^{\frac{1}{4}(\frac{1}{\gamma_0} - \frac{1}{\gamma_1})} \\ &\geq O(1) \exp\left\{\frac{1}{8} \frac{\Delta_n \log n}{\beta_N^2}\right\} \xrightarrow{n \rightarrow \infty} \infty. \end{aligned}$$

□

Lemma 4.8. *Let γ, β in B_{N_n} , $\gamma < \beta$. Then, for the estimator $\hat{\beta}$ of β , there exists a constant $d_4 > 0$ such that:*

$$\sup_{\beta \in B_{N_n}} \sup_{f \in W_n(\beta, L)} P_f \left[\hat{\beta} = \gamma \right] \leq d_4 \frac{\beta_N}{\Delta_n} n^{-\frac{q}{4\gamma}}.$$

Proof. By the definition of $\hat{\beta}$, when the event $\{\hat{\beta} = \gamma\}$ occurs, then for $\gamma_1 = \min\{\alpha \in B_{N_n} : \alpha > \gamma\}$, there exists at least one γ_0 , $\gamma_0 < \gamma_1 \leq \beta$, for which

$$\left| \hat{f}_{n,\gamma_1}(x_0) - \hat{f}_{n,\gamma_0}(x_0) \right| > \hat{\eta}_{n,\gamma_0}.$$

Then

$$\begin{aligned} P_f \left[\hat{\beta} = \gamma \right] &\leq \sum_{\substack{\gamma_0 \leq \gamma \\ \gamma_0 \in B_-}} P_f \left[\left| \hat{f}_{n,\gamma_1}(x_0) - \hat{f}_{n,\gamma_0}(x_0) \right| > \hat{\eta}_{n,\gamma_0} \right] \\ &\leq \text{card} B_- \cdot \sup_{\substack{\gamma_0 \leq \gamma \\ \gamma_0 \in B_-}} P_f \left[\left| \hat{f}_{n,\gamma_1}(x_0) - \hat{f}_{n,\gamma_0}(x_0) \right| > \hat{\eta}_{n,\gamma_0} \right] \\ &\leq \frac{\beta_N}{\Delta_n} \cdot \sup_{\substack{\gamma_0 \leq \gamma \\ \gamma_0 \in B_-}} (p_1 + p_2), \end{aligned}$$

where

$$\begin{aligned} p_1 &= P_f \left[\left\{ \left| \hat{f}_{n,\gamma_1}(x_0) - \hat{f}_{n,\gamma_0}(x_0) \right| > \hat{\eta}_{n,\gamma_0} \right\} \cap \{A_{n,\gamma_0} \cap A_{n,\gamma_1}\} \right], \\ p_2 &= P_f \left[\left\{ \left| \hat{f}_{n,\gamma_1}(x_0) - \hat{f}_{n,\gamma_0}(x_0) \right| > \hat{\eta}_{n,\gamma_0} \right\} \cap \{\bar{A}_{n,\gamma_0} \cup \bar{A}_{n,\gamma_1}\} \right]. \end{aligned}$$

In the case where the event A_{n,γ_0} holds, we have that

$$\hat{\eta}_{n,\gamma_0} = \eta_{n,\gamma_0} \left(\frac{\hat{h}_{n,\gamma_0}}{h_{n,\gamma_0}} \right)^{\tilde{\gamma}_0 - 1/2} \geq \eta_{n,\gamma_0} (1 - \delta_n)$$

and we may replace:

$$p_1 \leq P_f \left[\sup_{h \in \mathcal{H}_{n,\gamma_1}} |f_{n,\gamma_1}(h) - f| + \sup_{h \in \mathcal{H}_{n,\gamma_0}} |f_{n,\gamma_0}(h) - f| > \eta_{n,\gamma_0} (1 - \delta_n) \right].$$

Let us write that

$$\begin{aligned}
& \sup_{h \in \mathcal{H}_{n,\gamma_1}} |f_{n,\gamma_1}(h) - f| + \sup_{h \in \mathcal{H}_{n,\gamma_0}} |f_{n,\gamma_0}(h) - f| \\
& \leq \sup_{h \in \mathcal{H}_{n,\gamma_1}} (B_{n,\gamma_1}(h) + |Z_{n,\gamma_1}(h)|) + \sup_{h \in \mathcal{H}_{n,\gamma_0}} (B_{n,\gamma_0}(h) + |Z_{n,\gamma_0}(h)|) \\
& \leq (b_{n,\gamma_1} + b_{n,\gamma_0})(1 + \delta_n) + \sup_{h \in \mathcal{H}_{n,\gamma_1}} |Z_{n,\gamma_1}(h)| + \sup_{h \in \mathcal{H}_{n,\gamma_0}} |Z_{n,\gamma_0}(h)|.
\end{aligned}$$

By Lemma 4.7, $(b_{n,\gamma_1} + b_{n,\gamma_0})(1 + \delta_n) \leq d_5 \delta_{n3} \eta_{n,\gamma_0} (1 - \delta_n)$, for some $d_5 > 0$ and for n large enough. Then, we replace

$$\begin{aligned}
p_1 & \leq P_f \left[\sup_{h \in \mathcal{H}_{n,\gamma_1}} |Z_{n,\gamma_1}(h)| + \sup_{h \in \mathcal{H}_{n,\gamma_0}} |Z_{n,\gamma_0}(h)| \geq \eta_{n,\gamma_0} (1 - \delta_n) (1 - d_5 \delta_{n3}) \right] \\
& \leq P_f \left[\sup_{h \in \mathcal{H}_{n,\gamma_0}} |Z_{n,\gamma_0}(h)| \geq \eta_{n,\gamma_0} (1 - \delta_n) (1 - 2d_5 \delta_{n3}) \right] \\
& \quad + P_f \left[\sup_{h \in \mathcal{H}_{n,\gamma_1}} |Z_{n,\gamma_1}(h)| \geq d_5 \delta_{n3} \eta_{n,\gamma_0} (1 - \delta_n) \right].
\end{aligned}$$

For the first term on the right-hand side we apply Lemma 4.3, for

$$\tau_{n,\gamma_0} \leq \eta_{n,\gamma_0} (1 - \delta_n) (1 - 2d_5 \delta_{n3}) \leq c_1 s_{n,\gamma_0} \sqrt{\log n}$$

(that holds for c_1 large enough) and by the choice of δ_{n3} we get

$$P_f \left[\sup_{h \in \mathcal{H}_{n,\gamma_0}} |Z_{n,\gamma_0}(h)| \geq \eta_{n,\gamma_0} (1 - \delta_n) (1 - 2d_5 \delta_{n3}) \right] \leq d_3 \exp \left\{ -\frac{q \log n}{4\gamma_0} (1 - O(1) \delta_{n3}) \right\} \leq d_3 n^{-\frac{q}{4\gamma}}.$$

For the second term, $\tau_{n,\gamma_1} \leq \sqrt{\frac{q}{2\gamma_1}} s_{n,\gamma_1} \sqrt{\log n} \leq d_5 \delta_{n3} \eta_{n,\gamma_0} (1 - \delta_n) \leq c_2$ and we apply Lemma 4.5 for $\gamma_0 \leq \gamma$ as follows

$$P_f \left[\sup_{h \in \mathcal{H}_{n,\gamma_1}} |Z_{n,\gamma_1}(h)| \geq d_5 \delta_{n3} \eta_{n,\gamma_0} (1 - \delta_n) \right] \leq 8 \exp \left\{ -a \frac{\log n}{\gamma} \left(\frac{\delta_{n3} s_{n,\gamma_0}}{s_{n,\gamma_1}} \right)^2 (1 - \delta_n)^2 \right\}$$

and by the third statement of Lemma 4.7 this exponential is infinitely small with respect to $n^{-\frac{q}{4\gamma}}$. That allows us to conclude that

$$p_1 \leq O(1) n^{-\frac{q}{4\gamma}}.$$

We finish the proof of the lemma by using Lemma 3.2

$$p_2 \leq P_f [\overline{A}_{n,\gamma_0}] + P_f [\overline{A}_{n,\gamma_1}] \leq 4 \exp \left\{ -\frac{n((1-\alpha)h_n \delta_n \rho_n)^2}{2 \|K_0\|_\infty^2} \right\},$$

which is $o(1) n^{-\frac{q}{4\gamma}}$ by (3.1). □

Theorem 4.9. *We have the following convergence*

$$\sup_{\beta \in B_{N_n}} \text{card} B_- \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} \frac{(1 + \delta_n)^q b_{n,\gamma}^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \sup_{f \in W_n(\beta, L)} P_f \left[\hat{\beta} = \gamma \right] \xrightarrow{n \rightarrow \infty} 0.$$

Proof. Let us use Lemma 4.8 and, respectively, Lemma 3.5, for $\gamma < \beta$

$$\begin{aligned} \frac{(1 + \delta_n)^q b_{n,\gamma}^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \sup_{f \in W_n(\beta, L)} P_f \left[\hat{\beta} = \gamma \right] &\leq d_4 \frac{(1 + \delta_n)^q b_{n,\gamma}^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \frac{\beta_N}{\Delta_n} n^{-\frac{q}{4\gamma}} \\ &\leq O(1) (\log n)^{q/2} n^{q(\frac{1}{4\gamma} - \frac{1}{4\beta})} \frac{\beta_N}{\Delta_n} n^{-\frac{q}{4\gamma}}. \end{aligned}$$

We also have $\text{card} B_- \leq \frac{\beta_N}{\Delta_n}$ and then

$$\text{card} B_- \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} \frac{(1 + \delta_n)^q b_{n,\gamma}^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \sup_{f \in W_n(\beta, L)} P_f \left[\hat{\beta} = \gamma \right] \leq O(1) \frac{\beta_N^2 (\log n)^{q/2}}{\Delta_n^2} n^{-\frac{q}{4\beta_N}}$$

which is a $o(1)$, by means of relations (2.5) to (2.7). This expression is also independent of β in B and then the convergence is uniform. \square

5. UPPER BOUND

Let us prove (2.15) and (2.17) by successive decomposition of the risk and using the tools developed above. Denote $R_{n,\beta} = R_{n,\beta}(f_n^*, \psi_{n,\beta})$. In order to bound from above this expression, we consider the two possibilities when the estimated value $\hat{\beta}$ is greater or less than the true, unknown value of β . So we write:

$$R_{n,\beta} \leq R_{n,\beta}^+ + R_{n,\beta}^-,$$

where

$$\begin{aligned} R_{n,\beta}^+ &= \sup_{f \in W_n(\beta, L)} \psi_{n,\beta}^{-q} E_f \left[|f_n^*(x_0) - f(x_0)|^q I(\hat{\beta} \geq \beta) \right], \\ R_{n,\beta}^- &= \sup_{f \in W_n(\beta, L)} \psi_{n,\beta}^{-q} E_f \left[|f_n^*(x_0) - f(x_0)|^q I(\hat{\beta} < \beta) \right]. \end{aligned}$$

Thus, the proof of (2.15) consists of the following parts:

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} R_{n,\beta}^+ \leq 1 \tag{5.1}$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} R_{n,\beta}^- = 0. \tag{5.2}$$

5.1. Proof of (5.1)

Under the assumption $\hat{\beta} \geq \beta$, we have $|\hat{f}_{n,\hat{\beta}}(x_0) - \hat{f}_{n,\beta}(x_0)| \leq \hat{\eta}_{n,\beta}$, by the definition of the estimator $\hat{\beta}$. Therefore

$$R_{n,\beta}^+ \leq \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\hat{\eta}_{n,\beta} + \left| \hat{f}_{n,\beta}(x_0) - f(x_0) \right| \right]^q,$$

where $\hat{f}_{n,\beta}(x_0)$ is the kernel estimator with kernel K_β and random bandwidth $\hat{h}_{n,\beta}$.

We shall split again this upper bound, by integrating separately over $A_{n,\beta}$ and $\bar{A}_{n,\beta}$ (the complementary event of $A_{n,\beta}$), respectively. Then

$$R_{n,\beta}^+ \leq R_{n,\beta}^+(A_{n,\beta}) + R_{n,\beta}^+(\bar{A}_{n,\beta}),$$

where

$$R_{n,\beta}^+(A_{n,\beta}) = \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left(\hat{\eta}_{n,\beta} + \left| \hat{f}_{n,\beta}(x_0) - f(x_0) \right| \right)^q I(A_{n,\beta}) \right]$$

$$R_{n,\beta}^+(\bar{A}_{n,\beta}) = \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left(\hat{\eta}_{n,\beta} + \left| \hat{f}_{n,\beta}(x_0) - f(x_0) \right| \right)^q I(\bar{A}_{n,\beta}) \right].$$

We shall prove that

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} R_{n,\beta}^+(A_{n,\beta}) \leq 1 \tag{5.3}$$

and that

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} R_{n,\beta}^+(\bar{A}_{n,\beta}) = 0. \tag{5.4}$$

Proof of (5.3). We are in the case where the event $A_{n,\beta}$ holds. Then

$$\begin{aligned} \hat{\eta}_{n,\beta} + \left| \hat{f}_{n,\beta}(x_0) - f(x_0) \right| &\leq \sup_{h \in \mathcal{H}_{n,\beta}} \left[\eta_{n,\beta} \left(\frac{h}{h_{n,\beta}} \right)^{\beta-1/2} + |f_n(x_0, h, K_\beta) - f(x_0)| \right] \\ &\leq \sup_{h \in \mathcal{H}_{n,\beta}} \left[\eta_{n,\beta} \left(\frac{h}{h_{n,\beta}} \right)^{\beta-1/2} + b_{n,\beta}(h) + |Z_{n,\beta}(h)| \right] \\ &\leq \psi_{n,\beta} (1 + \delta_n) + \sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \end{aligned}$$

because $H_{n,\beta} \supseteq \mathcal{H}_{n,\beta}$ by relation (4.1) and $\eta_{n,\beta} + b_{n,\beta} = \psi_{n,\beta}$. Then

$$\begin{aligned} R_{n,\beta}^+(A_{n,\beta}) &\leq \sup_{f \in W_n(\beta,L)} E_f \left[(1 + \delta_n) + \psi_{n,\beta}^{-1} \sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \right]^q \\ &\leq \sup_{f \in W_n(\beta,L)} \left(1 + \delta_n + \psi_{n,\beta}^{-1} \tau_{n,\beta} \right)^q \\ &\quad + 2^{q-1} (1 + \delta_n)^q \sup_{f \in W_n(\beta,L)} P_f \left[\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \geq \tau_{n,\beta} \right] \\ &\quad + 2^{q-1} \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left(\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \right)^q I \left(\sup_{h \in H_{n,\beta}} |Z_{n,\beta}(h)| \geq \tau_{n,\beta} \right) \right]. \end{aligned}$$

For the first term on the right-hand side we apply Lemma 3.5 and this term is $1 + o(1)$, uniformly in β over B_- . The last two terms on the right-hand side tend to zero uniformly in β over B_- , as we see from Theorem 4.4 and the particular case of Theorem 4.6. This proves statement (5.3).

Proof of (5.4). We give direct upper bounds by Lemma 3.4 and then apply Lemma 3.2

$$\begin{aligned} R_{n,\beta}^+(\bar{A}_{n,\beta}) &\leq \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} \left(\hat{\eta}_{n,\beta} + \left| \hat{f}_{n,\beta}(x_0) - f(x_0) \right| \right)^q \sup_{f \in W_n(\beta,L)} P_f [\bar{A}_{n,\beta}] \\ &\leq 4d_0 n^{d_1} \exp \left\{ -\frac{n(h_n \delta_n \rho_n)^2}{2 \|K\|_\infty^2} \right\}. \end{aligned}$$

This bound is exponentially small uniformly in β over B_- by (3.1), then (5.4) holds.

5.2. Proof of (5.2)

As $\{\hat{\beta} < \beta\} \subset \bigcup_{\substack{\gamma \in B_- \\ \gamma < \beta}} \{\hat{\beta} = \gamma\}$, we may write:

$$\begin{aligned} R_{n,\beta}^- &\leq \sum_{\substack{\gamma \in B_- \\ \gamma < \beta}} \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left| \hat{f}_{n,\gamma}(x_0) - f(x_0) \right|^q I(\hat{\beta} = \gamma) \right] \\ &\leq \text{card} B_{N_n} \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left| \hat{f}_{n,\gamma}(x_0) - f(x_0) \right|^q I(\hat{\beta} = \gamma) \right] \\ &\leq \frac{\beta_{N_n}}{\Delta_n} \cdot \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} R_{n,\beta}^\gamma, \end{aligned}$$

where

$$R_{n,\beta}^\gamma = \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left| \hat{f}_{n,\gamma}(x_0) - f(x_0) \right|^q I(\hat{\beta} = \gamma) \right].$$

Similarly to the proof of (5.3) we integrate over $A_{n,\gamma}$ and $\bar{A}_{n,\gamma}$, respectively. We obtain

$$R_{n,\beta}^\gamma \leq R_{n,\beta}^\gamma(A_{n,\gamma}) + R_{n,\beta}^\gamma(\bar{A}_{n,\gamma}),$$

with

$$\begin{aligned} R_{n,\beta}^\gamma(A_{n,\gamma}) &= \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left| \widehat{f}_{n,\gamma}(x_0) - f(x_0) \right|^q I(\widehat{\beta} = \gamma) I(A_{n,\gamma}) \right] \\ R_{n,\beta}^\gamma(\overline{A}_{n,\gamma}) &= \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left| \widehat{f}_{n,\gamma}(x_0) - f(x_0) \right|^q I(\widehat{\beta} = \gamma) I(\overline{A}_{n,\gamma}) \right]. \end{aligned}$$

Therefore, $R_{n,\beta}^- \leq \frac{\beta_{N_n}}{\Delta_n} \cdot \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} \left(R_{n,\beta}^\gamma(A_{n,\gamma}) + R_{n,\beta}^\gamma(\overline{A}_{n,\gamma}) \right)$. We shall prove that

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} \left(\frac{\beta_{N_n}}{\Delta_n} \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} R_{n,\beta}^\gamma(A_{n,\gamma}) \right) = 0 \quad (5.5)$$

and that

$$\limsup_{n \rightarrow \infty} \sup_{\beta \in B_-} \left(\frac{\beta_{N_n}}{\Delta_n} \sup_{\substack{\gamma \in B_- \\ \gamma < \beta}} R_{n,\beta}^\gamma(\overline{A}_{n,\gamma}) \right) = 0. \quad (5.6)$$

Proof of (5.5). The event $A_{n,\gamma}$ is supposed to hold and we write directly that

$$\left| \widehat{f}_{n,\gamma}(x_0) - f(x_0) \right| \leq \sup_{h \in \mathcal{H}_{n,\gamma}} |f_n(x_0, h, K_\gamma) - f(x_0)| \leq \sup_{h \in \mathcal{H}_{n,\gamma}} (b_{n,\gamma}(h) + |Z_{n,\gamma}(h)|).$$

We use this bound in the expression of $A_{n,\gamma}$ and again relation (4.1)

$$\begin{aligned} R_{n,\beta}^\gamma(A_{n,\gamma}) &\leq \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left(b_{n,\gamma}(1 + \delta_n) + \sup_{h \in \mathcal{H}_{n,\gamma}} |Z_{n,\gamma}(h)| \right)^q I(\widehat{\beta} = \gamma) \right] \\ &\leq 2^{q-1} (1 + \delta_n)^q \sup_{f \in W_n(\beta,L)} \left(\frac{b_{n,\gamma}^q}{\psi_{n,\beta}^q} \cdot P_f[\widehat{\beta} = \gamma] \right) \\ &\quad + 2^{q-1} \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left(\sup_{h \in \mathcal{H}_{n,\gamma}} |Z_{n,\gamma}(h)| \right)^q I(\widehat{\beta} = \gamma) \right] \\ &\leq 2^{q-1} \sup_{f \in W_n(\beta,L)} \frac{b_{n,\gamma}^q (1 + \delta_n)^q + \tau_{n,\gamma}^q}{\psi_{n,\beta}^q} \sup_{f \in W_n(\beta,L)} P_f[\widehat{\beta} = \gamma] \\ &\quad + 2^{q-1} \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} E_f \left[\left(\sup_{h \in \mathcal{H}_{n,\gamma}} |Z_{n,\gamma}(h)| \right)^q I \left(\sup_{h \in \mathcal{H}_{n,\gamma}} |Z_{n,\gamma}(h)| \geq \tau_{n,\gamma} \right) \right]. \end{aligned}$$

Then, we have to apply Theorems 4.9 and 4.6. Thus, both terms on the right-hand side tend to 0, when we multiply by $\text{card } B_n$ and we take their supremum over $\gamma < \beta$, uniformly over $\beta \in B_-$. This finishes the proof of (5.5).

Proof of (5.6). We apply again Lemma 3.4 together with Lemma 3.2 which allow us to conclude that

$$\begin{aligned} R_{n,\beta}^\gamma(\overline{A}_{n,\gamma}) &\leq \sup_{f \in W_n(\beta,L)} \psi_{n,\beta}^{-q} |\widehat{f}_{n,\gamma}(x_0) - f(x_0)|^q \sup_{f \in W_n(\beta,L)} P_f[\overline{A}_{n,\gamma}] \\ &\leq 4d_0 n^{d_1} \exp\left\{-\frac{n(h_n \delta_n \rho_n)^2}{2\|K\|_\infty^2}\right\}. \end{aligned}$$

Since those bounds are exponentially small by (3.1) and depend neither on γ , nor on β , we have

$$\sup_{\beta \in B} \left(\frac{\beta_{N_n}}{\Delta_n} \cdot \sup_{\substack{\gamma \in B \\ \gamma < \beta}} R_{n,\beta}^\gamma(\overline{A}_{n,\gamma}) \right) = o(1).$$

5.3. Proof of (2.17)

For the case $\beta = \beta_{N_n} = \beta_N$, we have very similarly to (5.1):

$$\begin{aligned} R_{n,\beta_N}^+(f_n^*, \psi_{n,\beta_N}) &= \sup_{f \in W_n(\beta,L)} \psi_{n,\beta_N}^{-q} E_f \left[|f_n^*(x_0) - f(x_0)|^q I(\widehat{\beta} = \beta_N) \right] \\ &= \sup_{f \in W_n(\beta,L)} \psi_{n,\beta_N}^{-q} E_f \left[|f_{n,\beta_N}(x_0) - f(x_0)|^q \right] \\ &\leq \sup_{f \in W_n(\beta,L)} \frac{b_{n,\beta_N}^q + c(q) s_{n,\beta_N}^q}{\psi_{n,\beta_N}^q} < \infty, \end{aligned}$$

by Lemma 3.3. Similarly to (5.2),

$$R_{n,\beta_N}^-(f_n, \psi_{n,\beta_N}) = \sup_{f \in W_n(\beta,L)} \psi_{n,\beta_N}^{-q} E_f \left[|f_n^*(x_0) - f(x_0)|^q I(\widehat{\beta} < \beta_N) \right] \leq \frac{\beta_N}{\Delta_n} \sup_{\gamma \in B_-} R_{n,\beta_N}^\gamma \leq o(1).$$

Indeed, the proof goes exactly like in the case $\beta \in B_-$, for $\beta = \beta_N$ and $\gamma \in B_-$.

The proof of the upper bounds is now completed. \square

6. LOWER BOUND

The proof of the inequality (2.16) will now be given. In order to bound from below the minimax risk $R_{n,\beta}(f_n, \psi_{n,\beta})$ in (2.2), it suffices to bound it over a suitably chosen subset of probability densities in $W_n(\beta, L)$.

6.1. Introduction

Let f be a positive probability density on \mathbb{R} , infinitely continuously differentiable. For β in B_{N_n} , we denote $\|f^{(\beta)}\|_2^2 = \frac{1}{2\pi} \int_{\mathbb{R}} (\mathcal{F}(f)(x))^2 |x|^{2\beta} dx$. Let us consider the extreme values: β_1 and $\beta_{N-1} = \beta_{N_n-1}$ (the last element of the set B_-) and let f be such that both $\|f^{(\beta_1)}\|_2 \leq L$ and $\|f^{(\beta_{N-1})}\|_2 \leq L$ hold, for each n . Let $\delta \in (0, 1/2)$ be a small constant and define

$$f_{n,0}(x) = \left(\frac{\delta}{2}\right)^{\frac{1}{\beta_1+1/2}} f\left(x \left(\frac{\delta}{2}\right)^{\frac{1}{\beta_1+1/2}}\right),$$

such that $f_{n,0}(x_0) \geq \rho_n$.

Lemma 6.1. Define $K_\beta^*(x) = \nu_\beta^{-1} (2\beta - 1)^{\frac{1}{4\beta}} K_\beta \left((2\beta - 1)^{\frac{1}{2\beta}} x \right)$, where K_β is given by (2.12). Then $\|K_\beta^*\|_2 = 1$, $\|K_\beta^{*(\beta)}\|_2 = 1$ and

$$K_\beta^*(0) = b_\beta (2\beta) (2\beta - 1)^{-\frac{2\beta-1}{4\beta}}.$$

Indeed, it suffices to recall that $\|K_\beta\|_2 = \nu_\beta$ and that $\nu_\beta^2 = (2\beta - 1) b_\beta^2$. We can modify the function K_β^* in order to get the compact supported function \tilde{K}_β having the following properties.

Lemma 6.2. (Tsybakov [30]) For any fixed $\delta \in (0, \frac{1}{2})$ and $\beta > \frac{1}{2}$ there exists a number $D = D(\beta, \delta) > 0$ and an integrable function \tilde{K}_β supported on $(-D, D)$ such that $\|\tilde{K}_\beta\|_2 \leq 1 - \delta/2$, $\|\tilde{K}_\beta^{(\beta)}\|_2 \leq 1 - \delta/2$ and $(1 - \delta) K_\beta^*(0) \leq \tilde{K}_\beta(0) \leq K_\beta^*(0)$.

By this lemma, we get an integrable kernel, $\tilde{K}_{\beta_1} \in W(\beta_1, 1)$, satisfying the same conditions. Let

$$g_{n, \beta_1}(x) = L \tilde{h}_{n, \beta_1}^{\beta_1 - \frac{1}{2}} \cdot \tilde{K}_{\beta_1} \left(\frac{x - x_0}{\tilde{h}_{n, \beta_1}} \right),$$

where $\tilde{h}_{n, \beta_1} = \tilde{k}_{\beta_1} \left(\frac{f_{n,0}(x_0) \cdot \log n}{n} \right)^{\frac{1}{2\beta_1}}$ and $\tilde{k}_{\beta_1} = \left(\frac{q}{L^2 \cdot 2\beta_1} \right)^{\frac{1}{2\beta_1}}$.

Denote $\varepsilon_{n, \beta_1} = \int_{\mathbb{R}} g_{n, \beta_1}(x) dx$ which is finite by the previous lemma. Consider

$$f_{n,1}(x) = f_{n,0}(x)(1 - \varepsilon_{n, \beta_1}) + g_{n, \beta_1}(x).$$

Lemma 6.3. Let $f_{n,0}$, g_{n, β_1} and ε_{n, β_1} be defined as above. Then, $f_{n,0} \in W_n(\beta_{N-1}, L)$ and $f_{n,1} \in W_n(\beta_1, L)$, $\varepsilon_{n, \beta_1} = O\left(\tilde{h}_{n, \beta_1}^{\beta_1 + \frac{1}{2}}\right)$ and $(1 - \delta)\psi_{n, \beta_1} \leq g_{n, \beta_1}(x_0) \leq \psi_{n, \beta_1}$.

Proof. We can easily prove that $f_{n,0}$ is also a positive density function, infinitely continuously differentiable on \mathbb{R} . Moreover,

$$\|f_{n,0}^{(\beta_1)}\|_2 = \frac{\delta}{2} \|f^{(\beta_1)}\|_2 \leq \frac{L\delta}{2}.$$

Consequently, $f_{n,0} \in W_n(\beta_1, L \frac{\delta}{2})$ and obviously $f_{n,0} \in W_n(\beta_{N-1}, L)$, too.

We can see that

$$\varepsilon_{n, \beta_1} = L \tilde{h}_{n, \beta_1}^{\beta_1 + \frac{1}{2}} \int_{\mathbb{R}} \tilde{K}_{\beta_1}(x) dx = O\left(\tilde{h}_{n, \beta_1}^{\beta_1 + \frac{1}{2}}\right)$$

and that $g_{n, \beta_1}(x_0) = L \tilde{h}_{n, \beta_1}^{\beta_1 - \frac{1}{2}} \tilde{K}_{\beta_1}(0)$. Remark that $\psi_{n, \beta_1} = L \tilde{h}_{n, \beta_1}^{\beta_1 - \frac{1}{2}} K_{\beta_1}(0)$ and by Lemma 6.3

$$(1 - \delta)\psi_{n, \beta_1} \leq g_{n, \beta_1}(x_0) \leq \psi_{n, \beta_1}.$$

Remark also that $\|g_{n, \beta_1}^{(\beta_1)}\|_2 = L \|\tilde{K}_{\beta_1}^{(\beta_1)}\|_2 \leq L(1 - \frac{\delta}{2})$.

It is easy to prove that $f_{n,1}$ is a density function, positive for n large enough and that

$$f_{n,1}(x_0) \geq \rho_n (1 - \varepsilon_{n, \beta_1}) + g_{n, \beta_1}(x_0) \geq \rho_n$$

since $g_{n,\beta_1}(x_0) \geq \rho_n \varepsilon_{n,\beta_1}$, for n large enough. Moreover,

$$\left\| f_{n,1}^{(\beta_1)} \right\|_2 \leq (1 - \varepsilon_{n,\beta_1}) \left\| f_{n,0}^{(\beta_1)} \right\|_2 + \left\| g_{n,\beta_1}^{(\beta_1)} \right\|_2 \leq L \frac{\delta}{2} + L \left(1 - \frac{\delta}{2} \right) = L.$$

Then $f_{n,1} \in W_n(\beta_1, L)$. \square

6.2. Proof of the lower bound

For chosen $f_{n,0} \in W_n(\beta_{N-1}, L)$ and $f_{n,1} \in W_n(\beta_1, L)$ as in the previous section, we have

$$\begin{aligned} \inf_{\hat{f}_n} \sup_{\beta \in B} R_{n,\beta}(\hat{f}_n, \psi_{n,\beta}) &\geq \inf_{\hat{f}_n} \max_{\beta \in \{\beta_1, \beta_{N-1}\}} R_{n,\beta}(\hat{f}_n, \psi_{n,\beta}) \\ &\geq \inf_{\hat{f}_n} \max \left\{ E_{f_{n,0}}[\psi_{n,\beta_{N-1}}^{-q} \cdot |\hat{f}_n(x_0) - f_{n,0}(x_0)|^q], \right. \\ &\quad \left. E_{f_{n,1}}[\psi_{n,\beta_1}^{-q} \cdot |\hat{f}_n(x_0) - f_{n,1}(x_0)|^q] \right\}. \end{aligned}$$

Let us denote:

$$T_n = \psi_{n,\beta_1}^{-1} \cdot |\hat{f}_n(x_0) - f_{n,0}(x_0)| \quad \text{and} \quad q_n = \frac{\psi_{n,\beta_1}}{\psi_{n,\beta_{N-1}}}.$$

Remark that $q_n \geq a \left(\frac{n}{\log n} \right)^{\frac{\beta_{N-1} - \beta_1}{4\beta_1\beta_{N-1}}} \rightarrow +\infty$, when $n \rightarrow \infty$. Then, we may write

$$\begin{aligned} \inf_{\hat{f}_n} \sup_{\beta \in B} R_{n,\beta}(\hat{f}_n, \psi_{n,\beta}) &\geq \inf_{T_n} \max \left\{ E_0 |q_n T_n|^q, E_1 \left| T_n - \frac{f_{n,1}(x_0) - f_{n,0}(x_0)}{\psi_{n,\beta_1}} \right|^q \right\} \\ &\geq \inf_{T_n} \max \{ E_0 |q_n T_n|^q, E_1 |T_n - \theta_1|^q \}, \end{aligned} \quad (6.1)$$

with the notation $E_0 = E_{f_{n,0}}$, $E_1 = E_{f_{n,1}}$ (with the associated probability laws P_0 and P_1) and $\theta_1 = (g_{n,\beta_1}(x_0) - \varepsilon_{n,\beta_1} \cdot f_{n,0}(x_0)) / \psi_{n,\beta_1}$. Let us denote $R_n(T_n, \theta_1)$ the right-hand side term in (6.1), where \inf_{T_n} denotes the infimum over all random functions T_n .

Following the proof in Theorem 6, Tsybakov [30] we state here a similar result.

Lemma 6.4. *Let the numbers $q_n, q > 0, \tau > 0, 0 < \delta < \frac{1}{2}$ be fixed, let θ_1 be a real number such that $|\theta_1| \geq 1 - 2\delta$ and let P_0, P_1 be such that $P_1 \left[\frac{dP_0}{dP_1} \geq \tau \right] \geq 1 - \delta$. Then*

$$R_n(T_n, \theta_1) \geq \frac{(1 - \delta) \tau q_n^q (2\delta)^q (1 - 2\delta)^{2q}}{(1 - 2\delta)^q + \tau q_n^q (2\delta)^q}. \quad (6.2)$$

Proof. Remark that

$$|\theta_1| = \left| \frac{g_{n,\beta_1}(x_0)}{\psi_{n,\beta_1}} - \frac{\varepsilon_{n,\beta_1} \cdot f_{n,0}(x_0)}{\psi_{n,\beta_1}} \right| \geq \left| \left| \frac{g_{n,\beta_1}(x_0)}{\psi_{n,\beta_1}} \right| - \left| \frac{\varepsilon_{n,\beta_1} \cdot f_{n,0}(x_0)}{\psi_{n,\beta_1}} \right| \right|.$$

By Lemma 6.3, $1 - \delta < \left| \frac{g_{n,\beta_1}(x_0)}{\psi_{n,\beta_1}} \right| < 1$. By the definition of ε_{n,β_1} , $\frac{\varepsilon_{n,\beta_1} \cdot f_{n,0}(x_0)}{\psi_{n,\beta_1}} \leq O(\tilde{h}_{n,\beta_1}) = o(1)$ which allows us to consider $\left| \frac{\varepsilon_{n,\beta_1} \cdot f_{n,0}(x_0)}{\psi_{n,\beta_1}} \right| < \delta$, for sufficiently large n . Then $|\theta_1| \geq 1 - 2\delta$.

Let $B = \{|T_n| \geq 2\delta(1-2\delta)\}$. For $\delta < \frac{1}{2}$, the event $\{|T_n - \theta_1| \geq (1-2\delta)^2\}$ contains \overline{B} and we deduce that

$$P_1 \left[|T_n - \theta_1| \geq (1-2\delta)^2 \right] \geq P_1 [\overline{B}].$$

Then:

$$R_n(T_n, \theta_1) \geq \inf_{T_n} \max \left\{ q_n^q (2\delta)^q (1-2\delta)^q \cdot P_0[B], (1-2\delta)^{2q} \cdot P_1[\overline{B}] \right\}. \quad (6.3)$$

Moreover, let $A = \left\{ \frac{dP_0}{dP_1} \geq \tau \right\}$ and suppose for the moment that

$$P_1[A] \geq 1 - \delta \quad (6.4)$$

for arbitrary small $\delta \in (0, \frac{1}{2})$, and $\tau > 0$. Then

$$P_0[B] = E_1 \left[\frac{dP_0}{dP_1} I(B) \right] \geq \tau P_1[A \cap B] \geq \tau (P_1[B] - \delta)$$

and we use this to bound from below the right hand expression in (6.3) to get

$$\begin{aligned} R_n(T_n, \theta_1) &\geq \inf_{T_n} \max \left\{ (q_n 2\delta (1-\delta))^q \tau (P_1[B] - \delta), (1-2\delta)^{2q} P_1[\overline{B}] \right\} \\ &\geq \inf_{0 \leq t \leq 1} \max \left\{ (q_n 2\delta (1-\delta))^q \tau (t - \delta), (1-2\delta)^{2q} (1-t) \right\} \\ &\geq \frac{(1-\delta) \tau q_n^q (2\delta)^q (1-2\delta)^{2q}}{(1-2\delta)^q + \tau q_n^q (2\delta)^q}. \end{aligned}$$

□

Let us prove that (6.4) holds for n large enough, $\tau > 0$. We write

$$P_1[A] = P_1 \left[\prod_{i=1}^n \frac{f_{n,0}(X_i)}{f_{n,1}(X_i)} \geq \tau \right] = P_1 \left[\sum_{i=1}^n \frac{1}{\sqrt{\log n}} \log \frac{f_{n,0}(X_i)}{f_{n,1}(X_i)} \geq \frac{\log \tau}{\sqrt{\log n}} \right].$$

Consider the random variables $Z_{n,i} = \frac{1}{\sqrt{\log n}} \log \frac{f_{n,0}}{f_{n,1}}(X_i)$, for $i = 1, \dots, n$, $n \geq 1$, which are independent and identically distributed variables within each series. Let $V_1[Z_{n,i}]$ be the variance with respect to the distribution of X_1, \dots, X_n when the underlying probability density is $f_{n,1}$ and let $U_{n,i} = (Z_{n,i} - E_1[Z_{n,i}]) / \sigma_n$.

Lemma 6.5. *We have, for arbitrary small $\delta \in (0, \frac{1}{2})$:*

1. $\sum_{i=1}^n E_1[Z_{n,i}] \geq -\sqrt{\log n} \frac{q}{4\beta_1} \left(1 - \frac{\delta^2}{4}\right)^2$;
2. $\sigma_n^2 = \sum_{i=1}^n V_1[Z_{n,i}] \leq \frac{q}{2\beta_1} \left(1 - \frac{\delta^2}{4}\right)^2$;
3. $E_1[|U_{n,i}|^3] < +\infty$ and $\lim_{n \rightarrow \infty} \sum_{i=1}^n E_1[|U_{n,i}|^3] = 0$.

Proof. 1. Let us recall that there exists a fixed $a_0(\delta) \in (0, 1)$ such that

$$-x - \frac{x^2}{2} \left(1 + \frac{\delta}{2}\right) \leq \log(1-x) \leq -x, \quad (6.5)$$

for any $0 < x < a_0(\delta)$. Then

$$\sum_{i=1}^n E_1 [Z_{n,i}] = \frac{n}{\sqrt{\log n}} \left[-\log(1 - \varepsilon_{n,\beta_1}) + \int \log \left(1 - \frac{g_{n,\beta_1}}{f_{n,1}}(x) \right) f_{n,1}(x) dx \right].$$

As $\varepsilon_{n,\beta_1} \rightarrow 0$ and $\sup_x g_{n,\beta_1}/f_{n,1}(x) \rightarrow 0$ when $n \rightarrow \infty$, we may suppose without loss of generality that they are smaller than $a_0(\delta)$ and apply the inequality (6.5). Then, for n large enough,

$$\begin{aligned} \sum_{i=1}^n E_1 [Z_{n,i}] &\geq -\frac{1}{2} \left(1 + \frac{\delta}{2} \right) \frac{n}{\sqrt{\log n}} \int \frac{g_{n,\beta_1}^2}{f_{n,1}}(x) dx \\ &\geq -\frac{1}{2} \left(1 + \frac{\delta}{2} \right) \frac{q f_{n,0}(x_0)}{2\beta_1} \frac{\sqrt{\log n}}{\tilde{h}_{n,\beta_1}} \int \tilde{K}_{\beta_1}^2 \left(\frac{x - x_0}{\tilde{h}_{n,\beta_1}} \right) \frac{dx}{f_{n,1}(x)} \\ &\geq -\sqrt{\log n} \left(1 + \frac{\delta}{2} \right)^2 \frac{q}{4\beta_1} \|\tilde{K}_{\beta_1}\|_2^2 (1 + o(1)) \\ &\geq -\sqrt{\log n} \left(1 - \frac{\delta^2}{4} \right)^2 \frac{q}{4\beta_1} (1 + o(1)), \end{aligned}$$

where we used Lemma 6.2, Bochner's lemma and the fact that $f_{n,0}/f_{n,1}(x_0) = o(1)$ and becomes less than $1 + \delta/2$ when n is large enough.

2. For the variance, we have similarly

$$\begin{aligned} \sigma_n^2 &\leq \frac{n}{\log n} E_1 \left[\log^2 \left(1 - \frac{g_{n,\beta_1}}{f_{n,1}}(X_1) \right) \right] \leq \frac{n}{\log n} \left(\int \frac{g_{n,\beta_1}^2}{f_{n,1}}(x) dx + \frac{1}{2} \left(1 + \frac{\delta}{2} \right) \int \frac{g_{n,\beta_1}^4}{f_{n,1}}(x) dx \right) \\ &\leq \left(1 + \frac{\delta}{2} \right)^2 \frac{q}{2\beta_1} \|\tilde{K}_{\beta_1}\|_2^2 \leq \left(1 - \frac{\delta^2}{4} \right)^2 \frac{q}{2\beta_1} \end{aligned}$$

for n large enough.

3. By similar considerations, we get:

$$E_1 \left[\log^2 \left(1 - \frac{g_{n,\beta_1}}{f_{n,1}}(X_1) \right) \right] \geq \int \frac{g_{n,\beta_1}^2}{f_{n,1}}(x) dx \geq \frac{\log n}{n} \left(1 - \frac{\delta}{2} \right) \frac{q}{2\beta_1} \|\tilde{K}_{\beta_1}\|_2^2$$

and

$$\left[E_1 \log \left(1 - \frac{g_{n,\beta_1}}{f_{n,1}}(X_1) \right) \right]^2 \leq \left(-\varepsilon_{n,\beta_1} - \frac{1}{2} \left(1 + \frac{\delta}{2} \right) \int \frac{g_{n,\beta_1}^2}{f_{n,1}}(x) dx \right)^2 = o(1) \frac{\log n}{n}.$$

From those two last inequalities, we conclude that $\sigma_n^2 \geq (1 - \frac{\delta}{2})^2 \frac{q}{2\beta_1} \|\tilde{K}_{\beta_1}\|_2^2$ for n large enough. Then there

exists $\liminf_{n \rightarrow \infty} \sum_{i=1}^n V_1 [Z_{n,i}] = \sigma^2 > 0$, for $\delta \in (0, 1/2)$. Then

$$E_1 \left[|U_{i,n}|^3 \right] \leq \frac{O(1)}{(\log n)^{3/2}} E_1 \left| \log \left(1 - \frac{g_{n,\beta_1}}{f_{n,1}}(X_i) \right) \right|^3.$$

Let us note that

$$\begin{aligned} E_1 \left[\left| \log \left(1 - \frac{g_{n,\beta_1}}{f_{n,1}}(X_i) \right) \right|^3 \right] &\leq O(1) \int \frac{|g_{n,\beta_1}|^3}{f_{n,1}^2}(x) dx \\ &\leq O(1) \tilde{h}_{n,\beta_1}^{3\beta_1-1/2} \frac{1}{\tilde{h}_{n,\beta_1}} \int \left| \tilde{K}_{\beta_1} \left(\frac{x-x_0}{\tilde{h}_{n,\beta_1}} \right) \right|^3 \frac{dx}{f_{n,1}^2(x)} \\ &\leq O(1) \left(\frac{\log n}{n} \right)^{\frac{3\beta_1-1/2}{2\beta_1}} = o(1). \end{aligned}$$

Thus

$$\sum_{i=1}^n E_1 \left[|U_{i,n}|^3 \right] \leq O(1) \left(\frac{1}{\log n} \right)^{\frac{1}{4\beta_1}} \left(\frac{1}{n} \right)^{\frac{\beta_1-1/2}{2\beta_1}} = o(1).$$

□

So, the variables $(U_{n,i})_{1 \leq i \leq n, n \geq 1}$ are centered, $\sum_{i=1}^n V_1[U_{n,i}] = 1$ and they satisfy the property 3 of Lemma 6.5. Then we may apply Lyapounov's central limit theorem and get the following convergence in law to the standard Gaussian distribution:

$$U_n = \sum_{i=1}^n U_{n,i} \xrightarrow{D} N(0, 1), \text{ as } n \rightarrow \infty.$$

We take $m_n = \frac{\log \tau - n E_1 Z_{n,1}}{\sqrt{n V_1 Z_{n,1}}}$ and see that

$$P_1[A] = P_1 \left[U_n \geq \frac{\log \tau - n E_1 Z_{n,1}}{\sqrt{n V_1 Z_{n,1}}} \right] = P_1[U_n \geq m_n].$$

Let us choose $\tau = n^{-\xi}$, $\xi = -\frac{q}{4\beta_1} \left(1 - \frac{\delta^2}{4} \right)$ and use Lemma 6.5

$$m_n \leq -\sqrt{\log n} \frac{1}{2} \sqrt{\frac{q}{2\beta_1}} \frac{\delta^2}{4}.$$

Then, it is easy to see that $m_n \rightarrow -\infty$, for arbitrarily small $\delta \in (0, 1/2)$ which implies that $P_1[A] \xrightarrow{n \rightarrow \infty} 1$. This proves that (6.4) holds for n large enough and $\delta \in (0, 1/2)$ arbitrarily small.

By (6.1) and (6.2), we conclude that:

$$\inf_{\hat{f}_n} \sup_{\beta \in B} R_{n,\beta}(\hat{f}_n, \psi_{n,\beta}) \geq \frac{(1-\delta) \tau q_n^q (2\delta)^q (1-2\delta)^{2q}}{(1-2\delta)^q + \tau q_n^q (2\delta)^q}. \quad (6.6)$$

As, furthermore:

$$\begin{aligned} \liminf_{n \rightarrow \infty} \tau q_n^q &\geq \liminf_{n \rightarrow \infty} n^{-\frac{q}{4\beta_1} \left(1 - \frac{\delta^2}{4} \right)} \left(\frac{\psi_{n,\beta_1}}{\psi_{n,\beta_{N-1}}} \right)^q \\ &\geq \liminf_{n \rightarrow \infty} \exp \left[\frac{q}{4} \log n \left(\frac{1}{\beta_1} \frac{\delta^2}{4} - \frac{1}{\beta_{N-1}} - \frac{\log \log n}{\log n} \left(\frac{1}{\beta_1} - \frac{1}{\beta_{N-1}} \right) \right) \right] \end{aligned}$$

which is ∞ , for arbitrarily small $\delta \in (0, 1/2)$. Then by (6.6) we get the stated result (2.16). \square

The author is very grateful to Alexandre Tsybakov for his generous help.

REFERENCES

- [1] A. Barron, L. Birge and P. Massart, Risk bounds for model selection *via* penalization. *Probab. Theory Related Fields* **113** (1995) 301-413.
- [2] O.V. Besov, V.L. Il'in and S.M. Nikol'skii, *Integral representations of functions and imbedding theorems*. J. Wiley, New York (1978).
- [3] L. Birge and P. Massart, *From model selection to adaptive estimation*, Festschrift fur Lucien Le Cam. Springer (1997) 55-87.
- [4] L.D. Brown and M.G. Low, A constrained risk inequality with application to nonparametric functional estimation. *Ann. Statist.* **24** (1996) 2524-2535.
- [5] C. Butucea, The adaptive rates of convergence in a problem of pointwise density estimation. *Statist. Probab. Lett.* **47** (2000) 85-90.
- [6] C. Butucea, Numerical results concerning a sharp adaptive density estimator. *Comput. Statist.* **1** (2001).
- [7] L. Devroye and G. Lugosi, A universally acceptable smoothing factor for kernel density estimates. *Ann. Statist.* **24** (1996) 2499-2512.
- [8] D.L. Donoho, I. Johnstone, G. Kerkycharian and D. Picard, Wavelet shrinkage: Asymptopia? *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57** (1995) 301-369.
- [9] D.L. Donoho, I. Johnstone, G. Kerkycharian and D. Picard, Density estimation by wavelet thresholding. *Ann. Statist.* **24** (1996) 508-539.
- [10] D.L. Donoho and M.G. Low, Renormalization exponents and optimal pointwise rates of convergence. *Ann. Statist.* **20** (1992) 944-970.
- [11] S.Yu. Efromovich, Nonparametric estimation of a density with unknown smoothness. *Theory Probab. Appl.* **30** (1985) 557-568.
- [12] S.Yu. Efromovich and M.S. Pinsker, An adaptive algorithm of nonparametric filtering. *Automat. Remote Control* **11** (1984) 1434-1440.
- [13] A. Goldenshluger and A. Nemirovski, On spatially adaptive estimation of nonparametric regression. *Math. Methods Statist.* **6** (1997) 135-170.
- [14] G.K. Golubev, Adaptive asymptotically minimax estimates of smooth signals. *Problems Inform. Transmission* **23** (1987) 57-67.
- [15] G.K. Golubev, Quasilinear estimates for signals in \mathbb{L}_2 . *Problems Inform. Transmission* **26** (1990) 15-20.
- [16] G.K. Golubev, Nonparametric estimation of smooth probability densities in \mathbb{L}_2 . *Problems Inform. Transmission* **28** (1992) 44-54.
- [17] G.K. Golubev and M. Nussbaum, Adaptive spline estimates in a nonparametric regression model. *Theory Probab. Appl.* **37** (1992) 521-529.
- [18] I.A. Ibragimov and R.Z. Hasminskii, *Statistical estimation: Asymptotic theory*. Springer-Verlag, New York (1981).
- [19] A. Juditsky, Wavelet estimators: Adapting to unknown smoothness. *Math. Methods Statist.* **6** (1997) 1-25.
- [20] G. Kerkycharian and D. Picard, Density estimation by kernel and wavelet method, optimality in Besov space. *Statist. Probab. Lett.* **18** (1993) 327-336.
- [21] G. Kerkycharian, D. Picard and K. Tribouley, \mathbb{L}_p adaptive density estimation. *Bernoulli* **2** (1996) 229-247.
- [22] J. Klemelä and A.B. Tsybakov, *Sharp adaptive estimation of linear functionals*, Prépublication 540. LPMA Paris 6 (1999).
- [23] O.V. Lepskii, On a problem of adaptive estimation in Gaussian white noise. *Theory Probab. Appl.* **35** (1990) 454-466.
- [24] O.V. Lepskii, Asymptotically minimax adaptive estimation I: Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36** (1991) 682-697.
- [25] O.V. Lepskii, On problems of adaptive estimation in white Gaussian noise. *Advances in Soviet Math. Amer. Math. Soc.* **12** (1992b) 87-106.
- [26] O.V. Lepski and B.Y. Levit, Adaptive minimax estimation of infinitely differentiable functions. *Math. Methods Statist.* **7** (1998) 123-156.
- [27] O.V. Lepski, E. Mammen and V.G. Spokoiny, Optimal spatial adaptation to inhomogeneous smoothness: An approach based on kernel estimates with variable bandwidth selectors. *Ann. Statist.* **25** (1997) 929-947.
- [28] O.V. Lepski and V.G. Spokoiny, Optimal pointwise adaptive methods in nonparametric estimation. *Ann. Statist.* **25** (1997) 2512-2546.
- [29] D. Pollard, *Convergence of Stochastic Processes*. Springer-Verlag, New York (1984).
- [30] A.B. Tsybakov, Pointwise and sup-norm sharp adaptive estimation of functions on the Sobolev classes. *Ann. Statist.* **26** (1998) 2420-2469.
- [31] S. Van de Geer, *A maximal inequality for empirical processes*, Technical Report TW 9505. University of Leiden, Leiden (1995).