African Diaspora Journal of Mathematics
Volume 10, Number 2, pp. 27–57 (2010)
ISSN 1539-854X
www.math-res-pub.org/adjm

MULTIDIMENSIONAL NONPARAMETRIC DENSITY ESTIMATES: MINIMAX RISK WITH RANDOM NORMALIZING FACTOR

ARMEL FABRICE YODÉ*
Laboratoire de Mathématiques Appliquées et Informatique (L.M.A.I.)
Université de Cocody-Abidjan
22 BP 582 Abidjan 22, Côte d'Ivoire.

Abstract

We consider nonparametric minimax problem of multidimensional density estimation. Using the concept of random normalizing factor, by considering the plausible hypothesis of independence, we improve the accuracy of minimax estimation $n^{-\frac{\beta}{2\beta+d}}$: with prescribed confidence level α_n, we show that the best possible attainable (random) rate is $\left\{\sqrt{\log(2/\alpha_n)}/n\right\}^{\frac{2\beta}{4\beta+d}}$. We construct an optimal estimator and an optimal random normalizing factor in the sense of Lepski.

AMS Subject Classification: 62G07; 62G10; 62G15; 62G20.

Keywords: nonparametric estimation, minimax theory, adaptive estimation, random normalizing factor, probability density model, independence hypothese testing.

1 Introduction

The nonparametric problem of minimax estimation consists of constructing asymptotically optimal estimator on a chosen space of regular functions and finding its minimax rate of convergence (MRC). This MRC is attainable and cannot be improved in the minimax sense. Therefore, the MRC can be treated as accuracy of estimation. What should one do in the situation when it is bad? How should one improve it? In this paper, we propose to discuss these issues for multidimensional probability density model using the concept of minimax risks with random normalizing factors (RNF) initiated by Lepski [7]. This concept which is a combination of adaptive estimation and hypothesis testing, introduces a new kind of risks normalized by random variable depending on the observation. A first application of this concept is given in Lepski [7] for the estimation of an unknown signal in unidimensional Gaussian white noise model. Hoffmann [3] considered the estimation of the diffusion coefficient, when one observes unidimensional diffusion process at times i/n, $i = 0, \ldots, n$

*E-mail address: yafevrard@yahoo.fr

(asymptotics are studies as $n \to \infty$). Baraud [1] solved the problem of building a non asymptotic Euclidean confidence ball with prescribed probability of coverage via model selection. Hoffmann and Lepski [4] generalized this concept with application to the problem of selecting significant variables in multidimensional Gaussian white noise model.

1.1 Statistical setting

Let the statistical experiment generated by the observation $X^n = (X_1, \ldots, X_n)$, where $X_i = \big(X_i^{(1)}, \ldots, X_i^{(d)}\big)$, $i = 1, \ldots, n$ are independent identically distributed (*i.i.d.*) random vectors with common unknown probability density f defined on d-dimensional Euclidean space $\mathbb{R}^d$, where $d \geq 2$. The asymptotics will be studied w.r.t. $n \to +\infty$.
Let Φ be the set of all density functions compactly supported on $[0,1]^d$ and $\Sigma_d(\beta, L)$, $\beta = m + \tau$, $m \in \mathbb{N}$, $\tau \in (0,1]$, $L > 0$, be the isotropic Hölder functions space. We say that the function f belongs to $\Sigma_d(\beta, L)$ if

$$\big|f(x) - P_{m,f,y}(x)\big| \quad \leq \quad L\|x-y\|^{\beta}, \tag{1.1}$$

where the Taylor's polynomial $P_{m,f,y}(x)$ of f in y of order m and the Euclidian norm $\|\cdot\|$ are defined respectively by

$$P_{m,f,y}(x) = \sum_{0 \leq i_1 + \cdots + i_d \leq m} \frac{1}{i_1! \cdots i_d!} \prod_{l=1}^{d} (x_l - y_l)^{i_l} \frac{\partial^{i_1 + \cdots + i_d} f}{\partial x_1^{i_1} \ldots \partial x_d^{i_d}}(y)$$

$$\|x\| = \left(\sum_{l=1}^{d} x_l^2\right)^{1/2}, \quad x = (x_1, \ldots, x_d),\ y = (y_1, \ldots, y_d) \in \mathbb{R}^d.$$

We suppose that the unknown density function f belongs to

$$\Sigma = \Sigma(\beta, L, S) \overset{\triangle}{=} \quad \big\{f : \mathbb{R}^d \to \mathbb{R} :\ f \in \Sigma_d(\beta, L) \cap \Phi,\ \|f\|_\infty \leq S\big\}, \tag{1.2}$$

where S is a positive constant and $\|f\|_\infty = \sup\limits_{x \in \mathbb{R}^d} |f(x)|$.
Here and in the sequel, the space Σ is known a priori, and this knowledge is used for all constructions.

1.2 Minimax approach

Let φ_n be a normalizing factor i.e. a positive sequence such that $\varphi_n \to 0$ as $n \to +\infty$. We consider the maximal risk on the set Σ normalized by φ_n:

$$R_n\big(\tilde{f}_n, \Sigma, \varphi_n\big) = \sup_{f \in \Sigma} \mathbb{E}_f^n \big(\varphi_n^{-1} \big\|\tilde{f}_n - f\big\|_2\big)^q, \tag{1.3}$$

where $\tilde{f}_n(x) = \tilde{f}_n(x, X^n)$, $x \in [0,1]^d$, is some estimator, i.e. a function defined on $[0,1]^d$ and measurable w.r.t. the observation X^n; $\mathbb{E}_f^n$ is the expectation w.r.t. the probability measure $\mathbb{P}_f^n$ associated with X^n; $q > 0$ is a fixed number and $\|f\|_2 = \left(\int_{[0,1]^d} f^2(x)dx\right)^{1/2}$.

Definition 1.1. A normalizing factor $\varphi_n(\Sigma)$ is called minimax rate of convergence (MRC) if

(i) $\liminf_{n\to+\infty} \inf_{\tilde{f}_n\in\mathcal{M}_n} R_n\left(\tilde{f}_n,\Sigma,\varphi_n(\Sigma)\right) > 0,$
where $\mathcal{M}_n$ is the set of all estimators.

(ii) There exists an estimator $\bar{f}_n \in \mathcal{M}_n$ such that

$$\limsup_{n\to+\infty} R_n\left(\bar{f}_n,\Sigma,\varphi_n(\Sigma)\right) < +\infty.$$

An estimator $\bar{f}_n$ satisfying (ii) is called asymptotically optimal estimator.

The optimality of the MRC $\varphi_n(\Sigma)$ is described by (i). According to (ii), given $\bar{f}_n$ and $\varphi_n(\Sigma)$, we can construct a confidence set for the unknown density function f as the following manner: for any $0<\gamma<1$, using Markov inequality, there is a constant $C>0$ such that

$$\inf_{f\in\Sigma}\mathbb{P}_f^n\left\{\left\|\bar{f}_n - f\right\|_2 \le \left(C/\gamma\right)^{\frac{1}{q}}\varphi_n(\Sigma)\right\} \ge 1-\gamma. \tag{1.4}$$

It means that with given probability $1-\gamma$, the unknown density f lies inside the L_2-ball with center in $\bar{f}_n$ and of the radius $\left(C/\gamma\right)^{\frac{1}{q}}\varphi_n(\Sigma)$.
The MRC on the set Σ defined as (1.2) is $\varphi_n(\Sigma) = n^{-\frac{\beta}{2\beta+d}}$ (see Ibragimov and Khasminski [5]). This rate is attained, for example, by the Parzen-Rozenblatt estimator

$$\bar{f}_n(x) = \frac{1}{n\vartheta_n^d}\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right), \quad x\in[0,1]^d, \tag{1.5}$$

where K is a kernel function satisfying the traditional conditions; the bandwidth is $\vartheta_n = C_1 n^{-\frac{1}{2\beta+d}}$, where C_1 is a positive constant depending on Σ.

Remark 1.2. The MRC $\varphi_n(\Sigma) = n^{-\frac{\beta}{2\beta+d}}$ depends on the dimension d of Euclidean space $\mathbb{R}^d$: for great value of d, $\varphi_n(\Sigma)$ tends to zero too slowly. This phenomenon is known as the "curse of dimensionality" and is discouraging for applications.

Let Φ_0 be the set of the densities of the form $f_0(x_1,\ldots,x_d) = f_1(x_1)\ldots f_d(x_d)$. Here and later we denote by f_k, $k=1,\ldots,d$ the marginal densities of a given density function f defined on $\mathbb{R}^d$. Introduce the set $\Sigma_0 = \Sigma\cap\Phi_0$. If $f\in\Sigma_0$, the random vector which has as a density f, has its components which are independent random variables. Each univariate density f_k can be estimated separately using only the corresponding observations $\left(X_1^{(k)},\ldots,X_n^{(k)}\right)$. Let

$$\bar{f}_{k,n}(x_k) = \frac{1}{nb_n}\sum_{i=1}^n K_*\left(\frac{x_k - X_i^{(k)}}{b_n}\right),\ x_k\in\mathbb{R} \tag{1.6}$$

be the Parzen-Rozenblatt estimator that attains the univariate MRC $n^{-\frac{\beta}{2\beta+1}}$; the function K_* is a kernel function and the bandwidth is $b_n = C_2 n^{-\frac{1}{2\beta+1}}$, where C_2 is a positive constant depending on β, K_* and L. Since, $\bar{f}_{k,n}$, $k=1,\dots,d$, are *i.i.d.* random variables, therefore the estimator

$$\bar{f}_n^{(0)}(x) = \prod_{k=1}^{d} \bar{f}_{k,n}(x_k), \quad x \in [0,1]^d, \tag{1.7}$$

attains the MRC $\varphi_n(\Sigma_0) = n^{-\frac{\beta}{2\beta+1}}$ of estimation for $f(x_1,\dots,x_d) = f_1(x_1)\dots f_d(x_d)$ (see Lepski [7]).

Remark 1.3. If $f \in \Sigma_0$, then the accuracy of estimation becomes $\varphi_n(\Sigma_0) = n^{-\frac{\beta}{2\beta+1}}$ and one does not pay a dimensional effects, at least asymptotically. Moreover, $\varphi_n(\Sigma_0)$ is better than $\varphi_n(\Sigma)$ i.e.

$$\lim_{n\to+\infty} \frac{\varphi_n(\Sigma_0)}{\varphi_n(\Sigma)} = 0. \tag{1.8}$$

Suppose that we have a strong guess (hypothesis)

$$H_0 : \; f \in \Sigma_0.$$

Thus, the hope for improvement of estimation accuracy is based on the hypothesis that the estimated function belongs to the set Σ_0. Later, we will show how to use in the optimal way the outcomes of independence test (Yodé [9]) to improve the accuracy of estimation of f if H_0 holds without accuracy of our estimation being degraded if hypothesis H_0 is false.

1.3 Adaptive approach

A traditional way of improvement is the very popular adaptive approach. The discussions concerning this approach and the references of the publications on this topic are available in Lepski [7], Hoffmann and Lepski [4]. We propose here a short outline . Define the adaptive rate

$$\psi_n(f) = \begin{cases} \varphi_n(\Sigma_0) & f \in \Sigma_0, \\ \varphi_n(\Sigma) & f \in \Sigma\backslash\Sigma_0. \end{cases} \tag{1.9}$$

Definition 1.4. Then, an estimator $\hat{f}_n^{(a)}$ satisfying

$$\limsup_{n\to+\infty} \sup_{f\in\Sigma} \mathbb{E}_f^n \left(\psi_n^{-1}(f) \left\| \hat{f}_n^{(a)} - f \right\|_2 \right)^q < +\infty \tag{1.10}$$

is called adaptive estimator.

The procedure $\hat{f}_n^{(a)}$ is asymptotically optimal estimator simultaneously on Σ_0 and $\Sigma\backslash\Sigma_0$. Nevertheless, the normalizing factor $\psi_n(f)$ describing its accuracy depends on the unknown function f. This is the unavoidable payment for the adaptive property. Therefore, it is impossible to provide confidence set in the sense of (1.4). However, according to (1.8), $\hat{f}_n^{(a)}$ estimates better if f actually belongs to Σ_0. Unfortunately, such information cannot be obtained from observation. It seems reasonable to test the hypothesis H_0 and then to use the outcome of the test for construction of estimators and for studying the accuracy of estimation. This has been the main justification for introduction of the random normalizing factors by Lepski [7].

1.4 Organization of the paper

The rest of the paper is organized as follows. We present in Section 2 a general mathematical framework for improving the accuracy of estimation based on the notion of minimax risk with random normalizing factors. The main results are stated in Section 3 and the proofs are delayed until Section 5. Section 4 is devoted to the preliminary results. Some sketches of proofs are given in this section. The detailed proofs of these results are available in Yodé [9].

2 Minimax risks with random normalizing factors

As mentioned before, adaptive estimator has the advantage that it estimates the parameter better but we have no idea about the order of magnitude of the distance between the adaptive estimator and the true unknown function and cannot build nontrivial confidence sets from it. The main idea of random normalizing factors approach is to replace in formula (1.10) the adaptive rate $\psi_n(f)$ by a normalizing factor which depends on the observation X^n i.e. that one can calculate. Since $\psi(f)$ depends on information whether density f belongs to the set Σ_0 or not, we test the hypothesis H_0. If H_0 is accepted then the unknown parameter f lies not far from Σ_0 where $\varphi_n(\Sigma_0)$ is better than $\varphi_n(\Sigma)$. Hopefully we can use the estimator $\bar{f}_n^{(0)}$ defined as (1.7) and ensure the existence of ϕ_n not necessarily $\varphi_n(\Sigma_0)$, which would be the accuracy of corresponding estimation. If H_0 is rejected, this means that the test result provides no new information on the unknown parameter f. Thus, we use $\bar{f}_n$ defined as (1.5) which guarantees $\varphi_n(\Sigma)$ like accuracy of estimation. Formally, we use the estimator

$$f_n^* = \bar{f}_n \mathbb{I}_{\mathcal{A}_n^c} + \bar{f}_n^{(0)} \mathbb{I}_{\mathcal{A}_n}$$

and we hope that the accuracy given by this estimator is

$$\rho_n^* = \varphi_n(\Sigma) \mathbb{I}_{\mathcal{A}_n^c} + \phi_n \mathbb{I}_{\mathcal{A}_n}$$

where the events $\mathcal{A}_n$ and $\mathcal{A}_n^c$ are treated respectively as the acceptance region and rejection region of H_0. Note that ρ_n^* is a random variable.

Introduce the family Ω_n of observable random normalizing factors defined as the class

$$\Omega_n = \{\rho_n \in (0, \varphi_n(\Sigma)] : \ \rho_n \text{ is a random variable measurable w.r.t. } X^n\}.$$

For an arbitrary $\rho_n \in \Omega_n$ and for an estimator $\tilde{f}_n$, introduce the risk

$$R_n^{(r)}\left(\tilde{f}_n, \Sigma, \rho_n\right) = \sup_{f \in \Sigma} \mathbb{E}_f^n \left(\rho_n^{-1} \left\|\tilde{f}_n - f\right\|_2\right)^q. \tag{2.1}$$

The superscript (r) is put here to emphasize the random character of the normalizing factor. Our goal is to construct an estimator f_n^* and a random normalizing factor $\rho_n^* \in \Omega_n$ such that

(i) $\limsup_{n\to+\infty} R_n^{(r)}(f_n^*, \Sigma, \rho_n^*) < +\infty$;

(ii) the event $\{\rho_n^* < \varphi_n(\Sigma)\}$ has controlled probability on Σ_0 i.e.

$$\liminf_{n\to+\infty} \inf_{f\in\Sigma_0} \mathbb{P}_f^n \{\rho_n^* < \varphi_n(\Sigma)\} > 0,$$

(iii) f_n^* is adaptive estimator.

By definition, we have $\rho_n^* \leq \varphi_n(\Sigma)$. Thus, f_n^* is asymptotically optimal estimator on Σ w.r.t. risk defined as (1.3). It means that, by considering risk (2.1), we preserve the assets of standard minimax approach. From (i), we can obtain that with given probability $1-\gamma$, where $0<\gamma<1$, the unknown density f lies inside the L_2-ball with center in f_n^* and of the radius $(C/\gamma)^{1/q}\rho_n^*$, where $C>0$. This yields a confidence set as (1.4). According to (ii), if H_0 is accepted, we obtain a value to ρ_n^* essentially better than $\varphi_n(\Sigma)$ with some probability. This ensures an improved confidence set uniformly on Σ. Moreover, from (iii), we obtain an adaptive estimator f_n^* whose accuracy is calculable in contrast to the adaptive estimator of in the previous section.

Let $0<\delta<1$ be some given number and α_n be a fixed sequence assumed to be small such that $0<\alpha_n \leq 1-\delta$ for all n. The sequence α_n is arbitrary and fixed by the statistician. We want to guarantee that if actually $f \in \Sigma_0$, then we can provide some improvement with confidence $1-\alpha_n$ uniformly on Σ_0.

Definition 2.1. For a given confidence level α_n, the characteristic of $\rho_n \in \Omega_n$ is

$$x_n(\rho_n) = \inf\left\{ x \in (0, \varphi_n(\Sigma)] : \inf_{f\in\Sigma_0} P_f^n\{\rho_n \leq x\} \geq 1-\alpha_n \right\}.$$

The characteristic of ρ_n measures the improvement rate that ρ_n provides uniformly on Σ_0 with prescribed probability $1-\alpha_n$. Therefore, the concept of characteristic is used to compare the random normalizing factors. We will say that $\rho_n^{(1)}$ is better than $\rho_n^{(2)}$ if

$$\lim_{n\to+\infty} \frac{x_n(\rho_n^{(1)})}{x_n(\rho_n^{(2)})} = 0.$$

We introduce now a criterion of optimality of random normalizing factors.

Definition 2.2. (Hoffmann and Lepski [4])
The random normalizing factor $\rho_n^* \in \Omega_n$ is α_n-optimal w.r.t. Σ_0 if

(i) for any $\rho_n \in \Omega_n$ such that

$$\lim_{n\to+\infty} \frac{x_n(\rho_n)}{x_n(\rho_n^*)} = 0,$$

we have

$$\liminf_{n\to+\infty} \inf_{\tilde{f}_n \in \mathcal{M}_n} R_n^{(r)}\left(\tilde{f}_n, \Sigma, \rho_n\right) = +\infty, \tag{2.2}$$

(ii) there exists an estimator $f_n^* \in \mathcal{M}_n$ such that

$$\limsup_{n\to+\infty} R_n^{(r)}(f_n^*, \Sigma, \rho_n^*) < +\infty. \tag{2.3}$$

f_n^* is called α_n-adaptive estimator.

The value $x_n(\rho_n^*)$ cannot be improved in order due to (2.2). This fact together with (2.3) explain why ρ_n^* is called α_n-optimal. There is no uniqueness of the α_n-optimal RNF. Indeed, two RNF with the same characteristic are considered equivalent. Let ρ_n^* be an α_n-optimal random normalizing factor w.r.t. Σ_0. According to Hoffmann and Lepski [4], the RNF defined as

$$\hat{\rho}_n = \begin{cases} x_n(\rho_n^*) & \text{if } \rho_n^* \leq x_n(\rho_n^*), \\ \varphi_n(\Sigma) & \text{if } \rho_n^* > x_n(\rho_n^*) \end{cases}$$

is an α_n-optimal RNF w.r.t. Σ_0. Moreover, we have $x_n(\hat{\rho}_n) = x_n(\rho_n^*)$. This result shows that we can only define an RNF by two values: the accuracy of estimation on Σ and another value representing the improved accuracy of estimates obtained if $f \in \Sigma_0$. We can restrict ourselves to the family of RNF ρ_n taking two values $\{\varphi_n(\Sigma)\}$ and $\{a_n\}$, where $0 < a_n < \varphi_n(\Sigma)$. In this case, events $\{\rho_n = a_n\}$ and $\{\rho_n = \varphi_n(\Sigma)\}$ are respectively considered as acceptance and rejection of the hypothesis H_0. It is clear that a_n can not be better in order than $\varphi_n(\Sigma_0)$.
Let $\varphi_n(\alpha_n)$ be the minimax rate of testing of the hypothesis

$$H_0 : f \in \Sigma_0$$

against the alternative set

$$H_n : f \in \Phi_n(C\varphi_n) = \{f \in \Sigma : \|f - f_0\|_2 \geq C\varphi_n\},$$

where $f_0(x_1,\dots,x_d) = f_1(x_1)\dots f_d(x_d)$ is the product of marginal densities of f, $\varphi_n \to 0$ when $n \to +\infty$ and $C > 0$. Let

$$\rho_n^* = \begin{cases} \max(\varphi_n(\alpha_n), \varphi_n(\Sigma_0)) & \text{if } H_0 \text{ holds}, \\ \varphi_n(\Sigma) & \text{otherwise} \end{cases} \qquad f_n^* = \begin{cases} \bar{f}_n^{(0)} & \text{if } H_0 \text{ holds}, \\ \bar{f}_n & \text{otherwise}. \end{cases}$$

If we show that $\max(\varphi_n(\alpha_n), \varphi_n(\Sigma_0))$ is the characteristic of ρ_n^*, then we can hope that ρ_n^* is α_n-optimal and that f_n^* is α_n-adaptive. The exact statements of this procedures are given in Lepski [7].

3 Main results

The kernel functions satisfy the following assumptions:

($\mathcal{C}$) K and K_* are Lipschitz-functions with compact support on $\mathbb{R}^d$ and $\mathbb{R}$ respectively and

$$\|K_{h_n} * f - f\|_\infty \leq L_0 h_n^\beta, \quad K_{h_n}(\cdot) = \frac{1}{h_n^d} K\left(\frac{\cdot}{h_n}\right), \ \forall f \in \Sigma,$$
$$\|K_{*b_n} * g - g\|_\infty \leq L_0 b_n^\beta, \quad K_{*b_n}(\cdot) = \frac{1}{b_n} K_*\left(\frac{\cdot}{b_n}\right),$$

for all marginal density g, where $L_0 > 0$.

Remark 3.1. **(Probable choice of kernel functions)**
Introduce the one-sided kernels (Lepski and Tsybakov [8]) $L_-:[0,1]\to\mathbb{R}$ and $L_+:[-1,0]\to\mathbb{R}$ defined as

$$L_-(u)=\sum_{j=0}^{m}p_j(0)p_j(u),\quad L_+(u)=L_-(-u),$$

where m is defined in (1.1), $p_0,\dots,p_m$ are the first $m+1$ orthonormal Legendre polynomials on $[0,1]$. We have

$$\operatorname{supp}(L_-)=[0,1],\quad \operatorname{supp}(L_+)=[-1,0]$$

$$\int_0^1 L_-(u)u^j du=\int_{-1}^{0}L_+(u)u^j du=0,\quad j=1,\dots,m \tag{3.1}$$

$$\int_0^1 L_-(u)du=\int_{-1}^{0}L_+(u)du=1. \tag{3.2}$$

Let $x=(x_1,\dots,x_d)$. For every $i=1,\dots,d$, put

$$W_i(x_i,u_i)=\begin{cases}L_+(x_i-u_i) & \text{si } x_i\le\frac{1}{2}\\ L_-(x_i-u_i) & \text{si } x_i\ge\frac{1}{2}.\end{cases}$$

Then, we define

$$K(x-u)\stackrel{\triangle}{=}K(x,u)=\prod_{i=1}^{d}W_i(x_i,u_i).$$

Moreover, according to (1.2), (3.1) and (3.2), we obtain conditions $(\mathcal{C})$.
Introduce the Parzen-Rozenblatt estimator

$$\hat{f}_n(x)=\frac{1}{nh_n^d}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right),\ x\in[0,1]^d, \tag{3.3}$$

and the statistic

$$T_n=\left\|\hat{f}_n-\bar{f}_n^{(0)}\right\|_2^2-\frac{1}{n^2h_n^{2d}}\sum_{i=1}^{n}\int_{[0,1]^d}K^2\left(\frac{x-X_i}{h_n}\right)dx, \tag{3.4}$$

where

$$h_n=2^{\frac{4}{4\beta+d}}\Upsilon^{\frac{1}{4\beta+d}}L_0^{\frac{-4}{4\beta+d}}\left(n^{-1}\sqrt{\log\left(\frac{2}{\alpha_n}\right)}\right)^{\frac{2}{4\beta+d}},$$

with

$$\Upsilon=S^2\int_{\mathbb{R}^{3d}}K(u)K(v)K(u+w)K(v+w)dudvdw. \tag{3.5}$$

and $\bar{f}_n^{(0)}$ is defined by (1.7). Denote

$$\varphi_n(\alpha_n)=\left(n^{-1}\sqrt{\log\frac{2}{\alpha_n}}\right)^{\frac{2\beta}{4\beta+d}}\qquad \lambda=2^{\frac{8\beta+d}{8\beta+d}}\Upsilon^{\frac{\beta}{4\beta+d}}L_0^{\frac{d}{4\beta+d}}.$$

According to Yodé [9], the sequence $\varphi_n(\alpha_n)$ is the rate of testing of test of independence where the optimal test is based on the statistic T_n. Introduce the random event

$$\mathcal{A}_n = \left\{ T_n \leq (\lambda\varphi_n(\alpha_n))^2 \right\}$$

and put

$$\rho_n^* = \begin{cases} \varphi_n(\alpha_n) & \text{if } \mathcal{A}_n \text{ holds,} \\ \varphi_n(\Sigma) & \text{if } \mathcal{A}_n^c \text{ holds,} \end{cases} \qquad f_n^* = \begin{cases} \bar{f}_n^{(0)} & \text{if } \mathcal{A}_n \text{ holds,} \\ \bar{f}_n & \text{if } \mathcal{A}_n^c \text{ holds,} \end{cases}$$

where $\mathcal{A}_n^c$ is the complement to the event $\mathcal{A}_n$.
Fix A_*, B_* and a such that

$$A_* > \frac{8S\|K_*\|_2(\beta+2)}{2\beta+1}, \quad B_* > \frac{8S\|K\|_2(4\beta+3d+4)}{2(4\beta+d)}$$
$$a < \min\left\{1; -\frac{4\beta+3d+4}{2(4\beta+d)} + \frac{B_*}{8S\|K\|_2^2}; -\frac{\beta+2}{2\beta+1} + \frac{A_*}{8S\|K_*\|_2^2}\right\}.$$

Theorem 3.2. *Suppose that conditions $(\mathcal{C})$ are satisfied. Let $q \geq 2$, $d > 2$, $\beta > \frac{d}{4}$ and α_n be a positive sequence such that $\alpha_n = n^{-a}(1+\varepsilon_n)$ with $\lim_{n\to+\infty} \varepsilon_n = 0$. Then ρ_n^* is an α_n-optimal random normalizing factor and f_n^* is an α_n-adaptive estimator w.r.t. Σ_0. In particular,*

$$\limsup_{n\to+\infty} R_n^{(r)}(f_n^*, \Sigma, \rho_n^*) \leq M_*, \tag{3.6}$$

where

$$M_* = 2^q L_0^{\frac{dq}{2\beta+d}} S^{\frac{\beta q}{2\beta+d}} \|K\|_2^{\frac{2\beta q}{2\beta+d}} + \left(\lambda\left(1+\sqrt{\frac{q}{a}}\right)^{\frac{1}{2}} + \frac{\lambda}{\sqrt{2}} + 2^{\frac{6\beta+d}{4\beta+d}} L_0^{\frac{d}{8\beta+d}} \Upsilon^{\frac{\beta}{4\beta+d}}\right)^q.$$

Remark 3.3. The constants A_* and B_* come from Lemmas 4.1, 4.2 below. Condition $\beta > \frac{d}{4}$ implies $nh_n^d \to +\infty$ as $n \to +\infty$ for the kernel estimator (3.3). The constraint $d > 2$ which comes from Yodé [9] is related to the structure of the statistic (3.4). It is due to techniques of calculations. We believe that by refining these techniques, we can include the case $d = 2$ in ours results.

According to Yodé [9], we have for n large enough

$$\sup_{f\in\Sigma_0} \mathbb{P}\{\rho_n^* = \varphi_n(\Sigma)\} \leq \alpha_n.$$

Therefore, for n large enough we guarantee that

$$\inf_{f\in\Sigma_0} \mathbb{P}\{\rho_n^* = \varphi_n(\alpha_n)\} \geq 1-\alpha_n,$$

i.e. the probability of the improvement of accuracy of estimation is controlled by α_n.
From (3.6), for all $0 < \gamma < 1$, we have

$$\sup_{f\in\Sigma} \mathbb{P}_f^n \left\{ \|f_n^* - f\|_2 \leq \left(\frac{M^*}{\gamma}\right)^{1/q} \rho_n^* \right\} \geq 1-\gamma$$

as $n \to +\infty$. Thus, if event $\{\rho_n^* = \varphi_n(\alpha_n)\}$ holds, we obtain more precise coverage of an estimated function.
There exists a relevant choice of α_n that allows f_n^* to be adaptive w.r.t. the family $\{\Sigma\backslash\Sigma_0, \Sigma_0\}$.

Theorem 3.4. *All the assumptions of Theorem 3.2 are satisfied. Hence, if* $\alpha_n = O_n\left(\varphi_n^q(\Sigma_0)\right)$, *then* f_n^* *is an adaptive estimator, i.e.,*

$$\limsup_{n\to+\infty} R_n\left(f_n^*, \Sigma, \psi_n(f)\right) < +\infty,$$

where $\psi_n(f)$ *is defined by* (1.9).

The proof of this theorem is similar to that of Lepski [7], Hoffmann and Lepski [4].

4 Preliminary results

In this section, we give several Lemmas which have been used to prove the upper bound in Theorem 3.2. Consider the Parzen-Rosenblatt estimator $\hat{f}_n$ defined as (3.3).

Lemma 4.1. *For any positive sequence* $z = o_n(1)$ *as* $n \to +\infty$, *one has*

$$\sup_{f\in\Sigma} \mathbb{P}_f^n\left\{\left\|\hat{f}_n - \mathbb{E}_f^n\hat{f}_n\right\|_\infty \geq z\right\} \leq 2\left(\frac{4Q_1}{zh_n^{d+1}} + 1\right)\exp\left\{\frac{-nh_n^d z^2}{8S\|K\|_2^2}\right\},$$

where Q_1 *is the Lipschitz constant of* K.
In particular for any $B_* > 0$, *one has*

$$\sup_{f\in\Sigma} \mathbb{P}_f^n\left\{\left\|\hat{f}_n - \mathbb{E}_f^n\hat{f}_n\right\|_\infty \geq \sqrt{\frac{B_*\log n}{nh_n^d}}\right\} \leq \frac{8Q_1}{\sqrt{B_*}} n^{\frac{4\beta+3d+4}{2(4\beta+d)} - \frac{B_*}{8S\|K\|_2^2}}\left(1 + o_n(1)\right).$$

Proof of Lemma 4.1. For any $f \in \Sigma$, $x \in [0,1]^d$, the centered random variables

$$\xi_i(x) = K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}_f^n K\left(\frac{x - X_i}{h_n}\right), \quad i = 1,\ldots,n \tag{4.1}$$

are independent identically distributed. Moreover, we obtain

$$\mathbb{E}_f^n\xi_i^2(x) \leq S\|K\|_2^2 h_n^d \tag{4.2}$$

and

$$\mathbb{E}_f^n|\xi_i(x)|^l \leq 2^{l+1}Q_*^{l+1}h_n^d \text{ for any } l \geq 3 \tag{4.3}$$

using the inequality $(a+b)^l \leq 2^l(a^l + b^l)$ with $a > 0$, $b > 0$, where $Q_* = \max\{1, S, \|K\|_\infty\}$. Let $x_j, \ldots, x_{M_n}$ be distinct points in $[0,1]^d$ and positive sequence η_n such that the family of sets $\{\mathcal{B}(x_j, \eta_n) \overset{\triangle}{=} \{x \in \mathbb{R}^d : \|x_j - x\| \leq \eta_n\}\}$ defines a partition of $[0,1]^d$. The positive sequences M_n and η_n will be chosen later. For any $z > 0$, note that

$$\begin{aligned}
\mathbb{P}_f^n\left\{\left\|\hat{f}_n - \mathbb{E}_f^n\hat{f}_n\right\|_\infty \geq z\right\} &= \mathbb{P}_f^n\left\{\max_{j=1,\ldots,M_n}\sup_{\|x-x_j\|\leq\eta_n}\left|\hat{f}_n(x) - \mathbb{E}_f^n\hat{f}_n(x)\right| \geq z\right\} \\
&\leq \mathbb{P}_f^n\left\{\max_{1\leq j\leq M_n}\left|\hat{f}_n(x_j) - \mathbb{E}_f^n\hat{f}_n(x_j)\right| + \frac{2Q_1\eta_n}{h^{d+1}} \geq z\right\} \\
&= \mathbb{P}_f^n\left\{\bigcup_{j=1}^{M_n}\left\{\left|\hat{f}_n(x_j) - \mathbb{E}_f^n\hat{f}_n(x_j)\right| \geq z - \frac{2Q_1\eta_n}{h_n^{d+1}}\right\}\right\} \\
&\leq \sum_{j=1}^{M_n}\mathbb{P}_f^n\left\{\left|\hat{f}_n(x_j) - \mathbb{E}_f^n\hat{f}_n(x_j)\right| \geq z - \frac{2Q_1\eta_n}{h_n^{d+1}}\right\},
\end{aligned}$$

Choosing $\eta_n < \dfrac{z h_n^{d+1}}{4Q_1}$ and $M_n = \left\lfloor \dfrac{4Q_1}{z h_n^{d+1}} \right\rfloor + 1$, we obtain

$$\mathbb{P}_f^n \left\{ \left\| \hat{f}_n - \mathbb{E}_f^n \hat{f}_n \right\|_\infty \geq z \right\} \leq \sum_{j=1}^{M_n} \mathbb{P}_f^n \left\{ \frac{1}{n h_n^d} \left| \sum_{i=1}^n \xi_i(x_j) \right| \geq \frac{z}{2} \right\}. \tag{4.4}$$

For any $t > 0$, Markov exponential inequality yields

$$\mathbb{P}_f^n \left\{ \frac{1}{n h_n^d} \sum_{i=1}^n \xi_i(x_j) \geq z \right\} \leq \exp\left\{ -\sqrt{n h_n^d} z t \right\} \left(\mathbb{E}_f^n \left(\exp\left\{ \frac{t \xi_1(x_j)}{\sqrt{n h_n^d}} \right\} \right) \right)^n. \tag{4.5}$$

Using Taylor formula and (4.2), (4.3) for all $f \in \Sigma$ we have the following result:

$$\begin{aligned}
\mathbb{E}_f^n \left(\exp\left\{ \frac{t \xi_1(x_j)}{\sqrt{n h_n^d}} \right\} \right) &\leq 1 + \frac{t^2}{2 n h_n^d} \mathbb{E}_f^n \xi_1^2(x_j) + \sum_{l=3}^{+\infty} \frac{1}{l!} \left(\frac{t}{\sqrt{n h_n^d}} \right)^l \mathbb{E}_f^n |\xi_1(x_j)|^l \\
&\leq 1 + \frac{S \|K\|_2^2 t^2}{2n} + \frac{16 Q_*^4 t^3}{n \sqrt{n h^d}} \sum_{l=3}^{+\infty} \frac{1}{l!} \left(\frac{2 t Q_*}{\sqrt{n h_n^d}} \right)^{l-3} \\
&\leq 1 + \frac{S \|K\|_2^2 t^2}{2n} \left(1 + \frac{32 Q_*^4}{S \|K\|_2^2} \frac{t}{\sqrt{n h^d}} \exp\left\{ \frac{2 t Q_*}{\sqrt{n h_n^d}} \right\} \right) \\
&\leq \exp\left\{ \frac{S \|K\|_2^2 t^2}{2n} (1 + o_n(1)) \right\}.
\end{aligned} \tag{4.6}$$

for any $t = o_n(\sqrt{n h_n^d})$ as $n \to +\infty$. Then, combining (4.5), (4.6) and choosing $t = \sqrt{n h_n^d} z / S \|K\|_2$ with $z = o_n(1)$, we obtain

$$\mathbb{P}_f^n \left\{ \frac{1}{n h_n^d} \sum_{i=1}^n \xi_i(x_j) \geq z \right\} \leq \exp\left\{ -\frac{n h_n^d z^2}{2 S \|K\|_2^2} \right\}.$$

for all $f \in \Sigma$. Therefore, continuing (4.4), we conclude that for $z = o_n(1)$

$$\sup_{f \in \Sigma} \mathbb{P}_f^n \left\{ \left\| \hat{f}_n - \mathbb{E}_f^n \hat{f}_n \right\|_\infty \geq z \right\} \leq 2 \left(\frac{4 Q_1}{z h_n^{d+1}} + 1 \right) \exp\left\{ -\frac{n h_n^d z^2}{8 S \|K\|_2^2} \right\}. \blacksquare$$

Consider the estimator $\bar{f}_{kn}$ defined by (1.6). The following result is a consequence of Lemma 4.1.

Lemma 4.2. *For any positive sequence $z = o_n(1)$ as $n \to +\infty$, one has*

$$\sup_{f \in \Sigma} \mathbb{P}_f^n \left\{ \sup_{k=1,\dots,d} \left\| \bar{f}_{kn} - \mathbb{E}_f^n \bar{f}_{kn} \right\|_\infty \geq z \right\} \leq 2d \left(\frac{4 Q_2}{b_n^2 z} + 1 \right) \exp\left\{ \frac{-n b_n z^2}{8 S \|K_*\|_2^2} \right\},$$

where Q_2 is a Lipschitz constant of K_*.
In particular for any $A_* > 0$, one has

$$\sup_{f \in \Sigma} \mathbb{P}_f^n \left\{ \sup_{k=1,\dots,d} \left\| \bar{f}_{kn} - \mathbb{E}_f^n \bar{f}_{kn} \right\|_\infty \geq \sqrt{\frac{A_* \log n}{n b_n}} \right\} \leq \frac{8 d Q_2}{\sqrt{A_*}} n^{\frac{\beta+2}{2\beta+1} - \frac{A_*}{8 S \|K_*\|_2^2}} \left(1 + o_n(1) \right).$$

Introduce the following degenerate U-statistic U_n defined by

$$U_n = \frac{1}{n^2}\sum_{i\neq j} H_n(X_i,X_j), \tag{4.7}$$

where $H_n(u,v) = \int_{\mathbb{R}^d} \left(K_{h_n}(x-u) - \mathbb{E}_f^n K_{h_n}(x-X_1)\right)\left(K_{h_n}(x-v) - \mathbb{E}_f^n K_{h_n}(x-X_2)\right) dx.$

Lemma 4.3. *For any positive sequence* $z = o_n\left(n^{-2}h_n^{\frac{d}{5}}\right)$ *as* $n \to +\infty$, *one has*

$$\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{\frac{1}{n^2}\sum_{i\neq j} H_n(X_i,X_j) \geq z\right\} \leq 2\exp\left\{\frac{-n^2h_n^d z^2}{16\Upsilon}\right\}, \tag{4.8}$$

where Υ *is defined by* (3.5)*.*

Proof of Lemma 4.3. Let $\mathcal{F}_l = \sigma(X_1,\ldots,X_l)$ the σ-algebra be generated by $(X_1,\ldots,X_l)$, for any $l \geq 2$. Introduce the notations

$$\xi_{l-1}(f) \stackrel{\triangle}{=} \mathbb{E}_f^n\left(\exp\left\{\frac{2t}{n^2}\sum_{j=1}^{l-1} H_n(X_l,X_j)\right\}/\mathcal{F}_{l-1}\right)$$

$$B_n = B_n(f) \stackrel{\triangle}{=} \prod_{l=2}^{n}\xi_{l-1}(f).$$

De la Peña and Giné [2] and Johnson, Schechtman and Zinn [6] imply that for centered, independent random variables $Y_1,\ldots,Y_n$ and $p \geq 2$

$$\mathbb{E}_f^n\left|\sum_{i=1}^n Y_i\right|^p \leq C(p,\mathcal{K})n^{p/2-1}\sum_{i=1}^n \mathbb{E}(|Y_i|^p), \tag{4.9}$$

where $C(p,\mathcal{K}) = 2^p(p-1)^{p/2}\mathcal{K}^p\left(\frac{p}{\ln(p)}\right)^p$ and $\mathcal{K}$ is an universal constant.

For n large enough, several calculations using properties of degenerate U-statistic U_n and (4.9) imply that (see Yodé [9])

$$\sup_{f\in\Sigma}\mathbb{E}(B_n(f)) \leq \exp\left\{\frac{2\Upsilon t^2}{n^2h_n^d}\right\}, \tag{4.10}$$

for any positive sequence $t = o\left(h_n^{\frac{6d}{5}}\right)$.

Since B_n is $\mathcal{F}_{n-1}$-measurable random variable and $\mathbb{E}(\mathbb{E}(X/\mathcal{F})) = \mathbb{E}(X)$ for any random variable X such that $\mathbb{E}(|X|) < +\infty$, we have for any integer $n \geq 2$

$$\begin{aligned}
&\mathbb{E}_f^n\left(\frac{\exp\left\{\frac{2t}{n^2}\sum_{l=2}^n\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}}{B_n}\right)\\
&= \mathbb{E}_f^n\left(\frac{1}{B_n}\mathbb{E}_f^n\left(\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n-1}\sum_{j=1}^{l-1}H_n(X_l,X_j) + \frac{2t}{n^2}\sum_{j=1}^{n-1}H_n(X_n,X_j)\right\}\right)/\mathcal{F}_{n-1}\right)\\
&= \mathbb{E}_f^n\left(\frac{1}{B_n}\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n-1}\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}\mathbb{E}_f^n\left(\exp\left\{\frac{2t}{n^2}\sum_{j=1}^{n-1}H_n(X_n,X_j)\right\}/\mathcal{F}_{n-1}\right)\right).
\end{aligned}$$

Thus, we obtain

$$
\begin{aligned}
\mathbb{E}_f^n\left(\frac{\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n}\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}}{B_n}\right) &= \mathbb{E}_f^n\left(\frac{\xi_{n-1}}{B_n}\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n-1}\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}\right)\\
&= \mathbb{E}_f^n\left(\frac{1}{B_{n-1}}\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n-1}\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}\right)\\
&= \mathbb{E}_f^n\left(\frac{1}{B_{n-2}}\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n-2}\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}\right)\\
&\dots\\
&= \mathbb{E}_f^n\left(\frac{1}{B_2}\exp\left\{\frac{2t}{n^2}H_n(X_1,X_2)\right\}\right)\\
&= \mathbb{E}_f^n\left(\mathbb{E}_f^n\left(\frac{1}{\xi_1}\exp\left\{\frac{2t}{n^2}H_n(X_1,X_2)\right\}/\mathcal{F}_1\right)\right)\\
&= \mathbb{E}_f^n\left(\frac{\xi_1}{\xi_1}\right) = 1. \qquad (4.11)
\end{aligned}
$$

Using (4.10), (4.11) and Markov inequality, for $x>0$, $t>0$, we have

$$
\begin{aligned}
\mathbb{P}_f^n\left\{\frac{1}{n^2}\sum_{i\neq j}H_n(X_i,X_j)>z\right\} &\le x\exp\{-tz\}\mathbb{E}_f^n\left\{\frac{\exp\left\{\frac{2t}{n^2}\sum_{l=2}^{n}\sum_{j=1}^{l-1}H_n(X_l,X_j)\right\}}{B_n}\right\}+\frac{\mathbb{E}_f^nB_n}{x}\\
&\le 2\exp\left\{-\frac{zn^2h_n^d}{16\Upsilon}\right\}
\end{aligned}
$$

minimizing in x, t and thanks to the choices of z. ■
Introduce the sequences

$$
\begin{aligned}
\chi_n &= \frac{1}{n^2h_n^{2d}}\sum_{i=1}^{n}\left(r_i-\mathbb{E}_f^nr_i\right) \text{ where } r_i=\int_{[0,1]^d}\left(K\left(\frac{x-X_i}{h_n}\right)-\mathbb{E}_f^nK\left(\frac{x-X_i}{h_n}\right)\right)^2dx\\
\hat{\theta}_n &= \frac{1}{n^2h_n^{2d}}\sum_{i=1}^{n}\int_{[0,1]^d}K^2\left(\frac{x-X_i}{h_n}\right)dx,\\
\zeta_n(f) &= -\frac{1}{nh_n^{2d}}\int_{[0,1]^d}\left(\mathbb{E}_f^nK\left(\frac{x-X_1}{h_n}\right)\right)^2dx
\end{aligned}
$$

and the U-statistic U_n defined by (4.7).

Lemma 4.4. *For any function $f\in\Sigma$, one has*

$$
\int_{[0,1]^d}\left(\hat{f}_n(x)-\mathbb{E}_f^n\hat{f}_n(x)\right)^2dx=\chi_n+\mathbb{E}_f^n\hat{\theta}_n+U_n+\zeta_n(f),
$$

such that

$$\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{|\chi_n|\geq\sqrt{\frac{2\Delta_1\log n}{n^3h_n^{2d}}}\right\}\leq\frac{2}{n},\tag{4.12}$$

$$\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{|\hat\theta_n-\mathbb{E}_f^n\hat\theta_n|\geq\sqrt{\frac{2\Delta_1\log n}{n^5h_n^{2d}}}\right\}\leq\frac{2}{n},\tag{4.13}$$

$$\sup_{f\in\Sigma}|\zeta_n(f)|\leq\frac{S^2}{n},\tag{4.14}$$

for n large enough, where Δ_1 is a positive constant depending only on S and K.

Proof of Lemma 4.4. For any $f\in\Sigma$, we have the following decomposition

$$\begin{aligned}\int_{[0,1]^d}\left(\hat f_n(x)-\mathbb{E}_f^n\hat f_n(x)\right)^2dx&=\frac{1}{n^2h^{2d}}\sum_{i=1}^n\int_{[0,1]^d}\left(K\left(\frac{x-X_i}{h}\right)-\mathbb{E}_f^nK\left(\frac{x-X_i}{h}\right)\right)^2+U_n\\&=\chi_n+\frac{1}{n^2h^{2d}}\mathbb{E}_f^n\sum_{i=1}^n\int_{[0,1]^d}K^2\left(\frac{x-X_i}{h}\right)dx+\zeta_n(f)+U_n\end{aligned}$$

By proceeding as in the proof of Lemma 4.1, we obtain (4.12) and (4.13) for n large enough because the random variables considered are the sums of centered independent identically distributed random variables. Moreover, we have

$$|\zeta_n(f)|\leq\frac{1}{nh_n^{2d}}\int\left(\int K\left(\frac{x-u}{h_n}\right)f(u)du\right)^2dx\leq\frac{S^2}{n}.\blacksquare$$

Lemma 4.5. *We have*

$$\sup_{f\in\Sigma}\left|\left\|\mathbb{E}_f^n\hat f_n-\prod_{k=1}^d\mathbb{E}_f^n\bar f_{kn}\right\|-\left\|f-\prod_{k=1}^d f_k\right\|\right|^2\leq L_0^2h_n^{2\beta}\left(1+o_n(1)\right),\tag{4.15}$$

Proof of Lemma 4.5: Since $|\|a\|-\|b\||\leq\|a-b\|$, therefore, we have for any $f\in\Sigma$

$$\begin{aligned}\left|\left\|\mathbb{E}_f^n\hat f_n-\prod_{k=1}^d\mathbb{E}_f^n\bar f_{kn}\right\|-\left\|f-\prod_{k=1}^d f_k\right\|\right|^2&\leq\left\|\left(\mathbb{E}_f^n\hat f_n-\prod_{k=1}^d\mathbb{E}_f^n\bar f_{kn}\right)-\left(f-\prod_{k=1}^d f_k\right)\right\|^2\\&\leq\left\|\mathbb{E}_f^n\hat f_n-f\right\|_\infty^2+o\left(h_n^{2\beta}\right)\\&\leq L_0^2h_n^{2\beta}+o\left(h_n^{2\beta}\right).\blacksquare\end{aligned}$$

Let the sequence $\varpi_n:\Sigma\times[0,1]^d\to\mathbb{R}$ such that

$$\varpi_n^*=\sup_{f\in\Sigma}\|\varpi_{n,f}\|_\infty=\sup_{f\in\Sigma}\sup_{x\in[0,1]^d}|\varpi_{n,f}(x)|<\infty.$$

Lemma 4.6. *For any positive sequence $z=o_n\left(\varpi_n^*\right)$ as $n\to+\infty$, one has*

$$\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{\left|\int_{[0,1]^d}\left(\hat f_n(x)-\mathbb{E}_f^n\hat f_n(x)\right)\varpi_{n,f}(x)dx\right|\geq z\right\}\leq 2\exp\left\{\frac{-nz^2}{2\Gamma\varpi_n^{*2}}\right\}$$

where $\Gamma=S^2(1+\Delta_2)$ with $\Delta_2=\int K(u)K(u+v)dudv$.

For any $f \in \Sigma$, put $\tau_f(x) = f(x) - \prod_{k=1}^d f_k(x_k)$ with $x = (x_1, \ldots, x_d) \in [0,1]^d$.

Lemma 4.7. *For any $f \in \Sigma$, there exists positive sequence $z_f = o_n(h^{d/2}\|\tau_f\|_2)$ as $n \to +\infty$ such that*

$$\mathbb{P}_f^n\left\{\left|\int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x)\right)\tau_f(x)dx\right| \geq z_f\right\} \leq 2\exp\left\{\frac{-nz_f^2}{2\Gamma\|\tau_f\|_2^2}\right\}.$$

where $\Gamma = S^2(1+\Delta_2)$ with $\Delta_2 = \int K(u)K(u+v)dudv$.

To prove the Lemmas 4.6 and 4.7, we use the same techniques as the proof of Lemma 4.1. Consider the estimator $\bar{f}_n$ defined by (1.5).

Lemma 4.8. *For any $\delta > 0$, one has*

$$\limsup_{n\to+\infty}\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{\left\|\bar{f}_n - f\right\|_2 \geq (C+\delta)\varphi_n(\Sigma)\right\} = 0,$$

where $C = 2L_0^{\frac{d}{2\beta+d}} S^{\frac{\beta}{2\beta+d}} \|K\|_2^{\frac{2\beta}{2\beta+d}}$.

Proof of Lemma 4.8: Using triangular inequality, we have

$$\begin{aligned}\left\|\bar{f}_n - f\right\|_2 &\leq \left\|\bar{f}_n - \mathbb{E}_f^n\bar{f}_n\right\|_2 + \left\|\mathbb{E}_f^n\bar{f}_n - f\right\|_2\\ &\leq \left\|\bar{f}_n - E_f^n\bar{f}_n\right\|_2 + L_0 h_n^\beta.\end{aligned}$$

Then, for any $\delta > 0$, we have

$$\begin{aligned}\mathbb{P}_f^n\left\{\left\|\bar{f}_n - f\right\|_2 \geq (C+\delta)\varphi_n(\Sigma)\right\} &\leq \mathbb{P}_f^n\left\{\left\|\bar{f}_n - \mathbb{E}_f^n\bar{f}_n\right\|_2 + L_0h_n^\beta \geq (C+\delta)\varphi_n(\Sigma)\right\}\\ &= \mathbb{P}_f^n\left\{\gamma_n + \mathbb{E}_f^n V_n + U_{1n} \geq \left(\frac{C}{2}+\delta\right)^2\varphi_n^2(\Sigma)\right\}\end{aligned}$$

where $\gamma_n(f) = V_n(f) - \mathbb{E}_f^n V_n(f)$, and $U_{1n} = \frac{1}{n^2}\sum_{i\neq j} H_{1n}(X_i, X_j)$ with

$$V_n(f) = \frac{1}{n^2h_n^{2d}}\sum_{i=1}^n\int\left(K\left(\frac{x-X_i}{h_n}\right) - \mathbb{E}_f^n K\left(\frac{x-X_i}{h_n}\right)\right)^2 dx,$$

and

$$H_{1n}(X_i,X_j) = \frac{1}{h_n^{2d}}\int\Big(K\Big(\frac{x-X_i}{h_n}\Big) - \mathbb{E}_f^n K\Big(\frac{x-X_i}{h_n}\Big)\Big)\Big(K\Big(\frac{x-X_j}{h_n}\Big) - \mathbb{E}_f^n K\Big(\frac{x-X_j}{h_n}\Big)\Big)dx.$$

Moreover, we can easily see that $\sup_{f\in\Sigma}\mathbb{E}_f^n V_n(f) \leq \frac{S\|K\|_2^2}{nh_n^d}$. Then, for any $f \in \Sigma$, we have

$$\begin{aligned}\mathbb{P}_f^n\left\{\left\|\bar{f}_n - f\right\|_2 \geq (C+\delta)\varphi_n(\Sigma)\right\} &\leq \mathbb{P}_f^n\left\{\gamma_n(f) + U_{1n} \geq \left(\frac{C}{2}+\delta\right)^2\varphi_n^2(\Sigma) - \frac{S\|K\|_2^2}{nh_n^d}\right\}\\ &= \mathbb{P}_f^n\left\{\gamma_n(f) + U_{1n} \geq \left(\frac{C}{2}+\delta\right)^2\varphi_n^2(\Sigma) - \frac{C^2}{4}\varphi_n^2(\Sigma)\right\}\\ &= \mathbb{P}_f^n\left\{\gamma_n + U_{1n} \geq \delta(C+\delta)\varphi_n^2(\Sigma)\right\}.\end{aligned}$$

According to (4.12) in Lemma 4.4, we have

$$\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{|\gamma_n|\geq\sqrt{\frac{2\Delta_1\log n}{n^3h_n^{2d}}}\right\}\leq\frac{2}{n}. \tag{4.16}$$

Hence, introduce the random event

$$\Delta_n=\left\{|\gamma_n|\leq\frac{\sqrt{2\Delta_3\log n}}{n^{3/2}h_n^d}\right\}.$$

Then, using Lemma 4.3 and (4.16), we have

$$\begin{aligned}\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{\left\|\tilde{f}_n-f\right\|_2\geq(C+\delta)\,\varphi_n(\Sigma)\right\}&\leq\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{U_{1n}\geq\delta\,(C+\delta)\,\varphi_n^2(\Sigma)\right\}+\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{\Delta_n^c\right\}\\&\leq 2\exp\left\{-\frac{\delta^2(C+\delta)^2}{16\Upsilon}n^{\frac{d}{2\beta+d}}\right\}+\frac{2}{n},\end{aligned}$$

for any $\delta>0$. Therefore, we have the result when $n\to+\infty$. ■

5 Proof of Theorem 3.2

5.1 Upper bound

5.1.1 Proof of (3.6)

Let us prove inequality (3.6). For any $f\in\Sigma$, put

$$\begin{aligned}R_n^{(1)}(f)&=\mathbb{E}_f^n\left(\rho_n^{*-1}\left\|f_n^*-f\right\|_2\right)^q\mathbb{I}_{\mathcal{A}_n}\\&=\mathbb{E}_f^n\left(\varphi_n^{-1}(\alpha_n)\left\|\bar{f}_n^{(0)}-f\right\|_2\right)^q\mathbb{I}_{\mathcal{A}_n}\\R_n^{(2)}(f)&=\mathbb{E}_f^n\left(\rho_n^{*-1}\left\|f_n^*-f\right\|_2\right)^q\mathbb{I}_{\mathcal{A}_n^c}\\&=\mathbb{E}_f^n\left(\varphi_n^{-1}(\Sigma)\left\|\bar{f}_n-f\right\|_2\right)^q\mathbb{I}_{\mathcal{A}_n^c}.\end{aligned}$$

Hence, we have

$$R_n^{(r)}\left(f_n^*,\Sigma,\rho_n^*\right)\leq\sup_{f\in\Sigma}R_n^{(1)}(f)+\sup_{f\in\Sigma}R_n^{(2)}(f). \tag{5.1}$$

Estimation of $\sup_{f\in\Sigma}R_n^{(2)}(f)$**.** Clearly, we get

$$\sup_{f\in\Sigma}R_n^{(2)}(f)\leq\sup_{f\in\Sigma}\mathbb{E}_f^n\left(\varphi_n^{-1}(\Sigma)\left\|\bar{f}_n-f\right\|_2\right)^q.$$

From Lemma 4.8,

$$\sup_{f\in\Sigma}\mathbb{E}_f^n\left(\varphi_n^{-1}(\Sigma)\left\|\bar{f}_n-f\right\|_2\right)^q\leq(C+\delta)^q\left(1+o_n(1)\right)$$

for any $\delta > 0$. Letting $\delta \to 0$, we obtain

$$\limsup_{n\to+\infty}\sup_{f\in\Sigma} R_n^{(2)}(f) \le 2^q L_0^{\frac{dq}{2\beta+d}} S^{\frac{\beta q}{2\beta+d}} \|K\|_2^{\frac{2\beta q}{2\beta+d}}. \tag{5.2}$$

Estimation of $\sup_{f\in\Sigma} R_n^{(1)}(f)$**.** Using triangular inequality and Lemma 4.5, one has

$$\begin{aligned}
\varphi_n^{-1}(\alpha_n)\left\|\prod_{k=1}^d \bar f_{kn} - f\right\|_2 &\le \varphi_n^{-1}(\alpha_n)\left\|\prod_{k=1}^d \bar f_{kn} - \prod_{k=1}^d \mathbb{E}_f^n \bar f_{kn}\right\|_2 \\
&\quad + \varphi_n^{-1}(\alpha_n)\left\|\mathbb{E}_f^n \hat f_n - \prod_{k=1}^d \mathbb{E}_f^n \bar f_{kn}\right\|_2 + \varphi_n^{-1}(\alpha_n)\left\|\mathbb{E}_f^n \hat f_n - f\right\|_2 \\
&\le \varphi_n^{-1}(\alpha_n)\left\|\prod_{k=1}^d \bar f_{kn} - \prod_{k=1}^d \mathbb{E}_f^n \bar f_{kn}\right\|_\infty + \varphi_n^{-1}(\alpha_n) G(f) + A_1,
\end{aligned}$$

where $A_1 = 2^{\frac{6\beta+d}{4\beta+d}} L_0^{\frac{d}{8\beta+d}} \Upsilon^{\frac{\beta}{4\beta+d}}$ and $G(f) = \left\|f - \prod_{k=1}^d f_k\right\|_2$.
For any $f\in\Sigma$, $\varepsilon > 0$, Lemma 4.2 yields

$$\begin{aligned}
R_n^{(1)}(f) &\le \mathbb{E}_f^n \left(\varphi_n^{-1}(\alpha_n) G(f) + A_1\right)^q \mathbb{1}_{\mathcal{A}_n} + \varepsilon \\
&\le \left(\varphi_n^{-1}(\alpha_n) G(f) + A_1\right)^q \mathbb{P}_f^n\{\mathcal{A}_n\} + \varepsilon \\
&\triangleq \bar R_n^{(1)}(f) + \varepsilon
\end{aligned}$$

Therefore, we have

$$\limsup_{n\to+\infty}\sup_{f\in\Sigma} R_n^{(1)}(f) \le \limsup_{n\to+\infty}\sup_{f\in\Sigma} \bar R_n^{(1)}(f).$$

Thus, it is enough to estimate $\bar R_n^{(1)}(f)$, $f\in\Sigma$. Denote

$$\Sigma_n^{(1)} = \left\{ f\in\Sigma : G(f) \le \left(\left(1+\sqrt{\frac{q\log n}{\log\frac{2}{\alpha_n}}}\right)^{\frac12} + \frac{1}{\sqrt2}\right)\lambda\varphi_n(\alpha_n)\right\}.$$

and put

$$\bar R_n^{(1,1)} \triangleq \sup_{f\in\Sigma_n^{(1)}} \bar R_n^{(1)}(f), \quad \bar R_n^{(1,2)} \triangleq \sup_{f\in\Sigma\setminus\Sigma_n^{(1)}} \bar R_n^{(1)}(f).$$

Note that

$$\sup_{f\in\Sigma} \bar R_n^{(1)}(f) = \max\left\{\bar R_n^{(1,1)}, \bar R_n^{(1,2)}\right\}.$$

Since $\lim_{n\to+\infty} \alpha_n n^a = 1$, then for n large enough we obtain

$$\begin{aligned}
\bar R_n^{(1,1)} &\le \left(\lambda\left(1+\sqrt{\frac{q\log n}{\log\frac{2}{\alpha_n}}}\right)^{\frac12} + \frac{\lambda}{\sqrt2} + A_1 + o_n(1)\right)^q \\
&= \left(\lambda\left(1+\sqrt{\frac{q\log n}{\log 2 + a\log n}}\right)^{\frac12} + \frac{\lambda}{\sqrt2} + A_1 + o_n(1)\right)^q.
\end{aligned}$$

Hence, we obtain

$$\limsup_{n\to+\infty} \bar{R}_n^{(1,1)} \leq \left(\lambda\left(1+\sqrt{\frac{q}{a}}\right)^{\frac{1}{2}} + \frac{\lambda}{\sqrt{2}} + A_1\right)^q. \tag{5.3}$$

We will prove that

$$\limsup_{n\to+\infty} \bar{R}_n^{(1,2)} = 0. \tag{5.4}$$

We have from (5.3) and (5.4)

$$\limsup_{n\to+\infty} \bar{R}_n^{(1)} \leq \left(\lambda\left(1+\sqrt{\frac{q}{a}}\right)^{\frac{1}{2}} + \frac{\lambda}{\sqrt{2}} + A_1\right)^q.$$

Last expression together with (5.1) and (5.2) imply (3.6). ■

5.1.2 **Proof of** (5.4)

Let us introduce the notations

$$\begin{aligned}
I_1(f)(x) &= \hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x), \\
I_2(f)(x) &= \mathbb{E}_f^n \hat{f}_n(x) - \prod_{k=1}^{d} \mathbb{E}_f^n \bar{f}_{kn}(x_k), \\
I_3(f)(x) &= \prod_{k=1}^{d} \mathbb{E}_f^n \bar{f}_{kn}(x_k) - \prod_{k=1}^{d} \bar{f}_{kn}(x_k).
\end{aligned}$$

We have the following decomposition

$$\begin{aligned}
T_n &= \left\| \hat{f}_n - \prod_{k=1}^{d} \bar{f}_{kn} \right\|_2^2 - \frac{1}{n^2 h_n^{2d}} \sum_{i=1}^{n} \int_{[0,1]^d} K^2\left(\frac{x-X_i}{h_n}\right) dx, \\
&= S_{1n}(f) + S_{2n}(f) + S_{3n}(f) + S_{4n}(f) + S_{5n}(f) + S_{6n}(f),
\end{aligned}$$

where

$$\begin{aligned}
S_{1n}(f) &= \int_{[0,1]^d} I_1^2(f)(x)dx - \frac{1}{n^2 h_n^{2d}} \sum_{i=1}^{n} \int_{[0,1]^d} K^2\left(\frac{x-X_i}{h_n}\right) dx, \\
S_{2n}(f) &= \int_{[0,1]^d} I_2^2(f)(x)dx, \\
S_{3n}(f) &= \int_{[0,1]^d} I_3^2(f)(x)dx, \\
S_{4n}(f) &= 2\int_{[0,1]^d} I_1(f)(x) I_2(f)(x)dx, \\
S_{5n}(f) &= 2\int_{[0,1]^d} I_1(f)(x) I_3(f)(x), \\
S_{6n}(f) &= 2\int_{[0,1]^d} I_2(f)(x) I_3(f)(x)dx.
\end{aligned}$$

According to Lemma 4.5, for $f \in \Sigma \setminus \Sigma_n^{(1)}$, for n large enough

$$\begin{aligned}\mathbb{P}_f^n \{\mathcal{A}_n\} &= \mathbb{P}_f^n \left\{ T_n \le (\lambda \varphi_n(\alpha_n))^2 \right\} \\ &\le \mathbb{P}_f^n \left\{ \sum_{l \in \{1,4,5,6\}} S_{ln}(f) + \left(G(f) - L_0 h_n^{\beta} \right)^2 \le (\lambda \varphi_n(\alpha_n))^2 \right\}.\end{aligned}$$

Let us study $S_{4n}(f)$. We have the following decomposition

$$S_{4n} = J_{1n}(f) + J_{2n}(f) + J_{3n}(f) + J_{4n}(f),$$

with

$$\begin{aligned}J_{1n}(f) &= 2 \int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x) \right) \left(\mathbb{E}_f^n \hat{f}_n(x) - f(x) \right) dx \\ J_{2n}(f) &= 2 \int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x) \right) \prod_{k=1}^{d} \left(\mathbb{E}_f^n \bar{f}_{kn} - f_k(x_k) \right) dx \\ J_{3n}(f) &= 2 \int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x) \right) \left(f(x) - \prod_{k=1}^{d} f_k(x_k) \right) dx \\ J_{4n}(f) &= 2 \int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x) \right) B_n(f)(x) dx\end{aligned}$$

and

$$B_n(f)(x) = \sum_{l=1}^{d-1} \sum_{k_1 \ne \cdots \ne k_d} \prod_{s=1}^{l} \left(\mathbb{E}_f^n \bar{f}_{k_s n}(x_{k_s}) - f_{k_s}(x_{k_s}) \right) \prod_{s=l+1}^{d} f_{ks}(x_{k_s}). \tag{5.5}$$

Using Lemma 4.6, we get

$$\begin{aligned}\lim_{n \to +\infty} \sup_{f \in \Sigma \setminus \Sigma_n^{(1)}} \mathbb{P}_f^n \left\{ |J_{1n}(f)| \ge h_n^{\beta} \sqrt{\frac{\log n}{n}} \right\} &= 0 \\ \lim_{n \to +\infty} \sup_{f \in \Sigma \setminus \Sigma_n^{(1)}} \mathbb{P}_f^n \left\{ |J_{2n}(f)| \ge b_n^{d\beta} \sqrt{\frac{\log n}{n}} \right\} &= 0, \\ \lim_{n \to +\infty} \sup_{f \in \Sigma \setminus \Sigma_n^{(1)}} \mathbb{P}_f^n \left\{ |J_{4n}(f)| \ge b_n^{\beta} \sqrt{\frac{\log n}{n}} \right\} &= 0.\end{aligned}$$

In view of Lemma 4.7, we have

$$\lim_{n \to +\infty} \mathbb{P}_f^n \left\{ |J_{3n}(f)| \ge \|\tau_f\|_2 \sqrt{\frac{\log n}{n}} \right\} = 0.$$

Let us introduce the random events for any $f \in \Sigma \setminus \Sigma_n^{(1)}$,

$$H_{1n} = \left\{ |J_{1n}(f)| \leq h_n^{\beta} \sqrt{\frac{\log n}{n}} \right\},$$

$$H_{2n} = \left\{ |J_{2n}(f)| \leq b_n^{d\beta} \sqrt{\frac{\log n}{n}} \right\},$$

$$H_{3n} = \left\{ |J_{3n}(f)| \leq \|\tau_f\|_2 \sqrt{\frac{\log n}{n}} \right\}$$

$$H_{4n} = \left\{ |J_{4n}(f)| \leq b_n^{\beta} \sqrt{\frac{\log n}{n}} \right\}.$$

Let us study S_{6n}. Denote

$$D_n = \left\{ \sup_{k=1,\ldots,d} \left\| \bar{f}_{kn} - \mathbb{E}_f^n \bar{f}_{kn} \right\|_\infty \leq \sqrt{\frac{A_* \log n}{n b_n}} \right\}.$$

We have the following decomposition:

$$S_{6n} = K_{1n}(f) + K_{2n}(f) + K_{3n}(f) + K_{4n}(f),$$

where

$$K_{1n}(f) = -2 \int_{[0,1]^d} \left(\prod_{k=1}^d \bar{f}_{kn}(x_k) - \prod_{k=1}^d \mathbb{E}_f^n \bar{f}_{kn}(x_k) \right) \left(\mathbb{E}_f^n \hat{f}_n(x) - f(x) \right) dx$$

$$K_{2n}(f) = 2 \int_{[0,1]^d} \left(\prod_{k=1}^d \bar{f}_{kn}(x_k) - \prod_{k=1}^d \mathbb{E}_f^n \bar{f}_{kn}(x_k) \right) \prod_{k=1}^d \left(\mathbb{E}_f^n \bar{f}_{kn}(x_k) - f_k(x_k) \right) dx,$$

$$K_{3n}(f) = -2 \int_{[0,1]^d} \left(\prod_{k=1}^d \bar{f}_{kn}(x_k) - \prod_{k=1}^d \mathbb{E}_f^n \bar{f}_{kn}(x_k) \right) \left(f(x) - \prod_{k=1}^d f_k(x_k) \right) dx$$

$$K_{4n}(f) = -2 \int_{[0,1]^d} \left(\prod_{k=1}^d \bar{f}_{kn}(x_k) - \prod_{k=1}^d \mathbb{E}_f^n \bar{f}_{kn}(x_k) \right) B_n(f)(x) dx.$$

Therefore, we have

$$\sup_{f \in \Sigma \setminus \Sigma_n^{(1)}} [\mathbb{I}_{D_n} |K_{1n}(f)|] \leq B_1 h_n^{\beta} \sqrt{\frac{\log n}{n b_n}}$$

$$\sup_{f \in \Sigma \setminus \Sigma_n^{(1)}} [\mathbb{I}_{D_n} |K_{2n}(f)|] \leq B_2 b_n^{d\beta} \sqrt{\frac{\log n}{n b_n}}$$

$$\sup_{f \in \Sigma \setminus \Sigma_n^{(1)}} [\mathbb{I}_{D_n} |K_{4n}(f)|] \leq B_4 b_n^{\beta} \sqrt{\frac{\log n}{n b_n}},$$

for n large enough, where B_1, B_2 and B_4 are positive constants depending on d, A_*, L_0 and M_0. Therefore, if random event D_n holds, the terms $K_{1n}(f)$, $K_{2n}(f)$ and $K_{3n}(f)$ are negligible w.r.t. $h_n^{2\beta}$ fonr any $f \in \Sigma \setminus \Sigma_n^{(1)}$. Using Cauchy-Schwarz inequality, we get for any $f \in \Sigma \setminus \Sigma_n^{(1)}$

$$|\mathrm{I\!I}_{D_n} K_{3n}(f)| \leq \left\| \prod_{k=1}^{d} \bar{f}_{kn} - \prod_{k=1}^{d} \mathbb{E}_f^n \bar{f}_{kn} \right\|_\infty \left\| f - \prod_{k=1}^{d} f_k \right\|_2 \leq B_3 \|\tau_f\|_2 \sqrt{\frac{\log n}{n b_n}},$$

where B_3 is positive constant depending on d, A_*, L_0 and M_0.
For $S_{5n}(f)$, we have

$$S_{5n}(f) = V_{1n}(f) + V_{2n}(f) + V_{3n}(f),$$

where

$$V_{1n}(f) = -2 \int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x)\right) \prod_{k=1}^{d} \left(\bar{f}_{kn}(x_k) - \mathbb{E}_f^n \bar{f}_{kn}(x_k)\right) dx,$$

$$V_{2n}(f) = -2 \sum_{l=2}^{d-1} \sum_{k_1 \neq \ldots \neq k_d} \int_{[0,1]^d} (\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x)) \prod_{s=1}^{l} \bar{f}_{k_s n}(x_{k_s}) - \mathbb{E}_f^n \bar{f}_{k_s n}(x_{k_s})) \prod_{s=l+1}^{d} \mathbb{E}_f^n \bar{f}_{k_s n}(x_{k_s}) dx,$$

$$V_{3n}(f) = -2 \sum_{k_1 \neq \ldots \neq k_d} \int_{[0,1]^d} \left(\hat{f}_n(x) - \mathbb{E}_f^n \hat{f}_n(x)\right) \left(\bar{f}_{k_1 n}(x_{k_1}) - \mathbb{E}_f^n \bar{f}_{k_1 n}(x_{k_1})\right) \prod_{s=2}^{d} \mathbb{E}_f^n \bar{f}_{k_s n}(x_{k_s}) dx.$$

Introduce the random event

$$G_n = \left\{ \left\| \hat{f}_n - \mathbb{E}_f^n \hat{f}_n \right\|_\infty \leq \sqrt{\frac{B_* \log n}{n h_n^d}} \right\}.$$

Thus, we have the inequality

$$\sup_{f \in \Sigma} |\mathrm{I\!I}_{D_n} \mathrm{I\!I}_{G_n} V_{1n}(f)| \leq \sup_{f \in \Sigma} \left[\mathrm{I\!I}_{G_n} \left\| \hat{f}_n - \mathbb{E}_f^n \hat{f}_n \right\|_\infty \left(\mathrm{I\!I}_{D_n} \sup_{k=1,\ldots,d} \left\| \bar{f}_{kn} - \mathbb{E}_f^n \bar{f}_{kn} \right\|_\infty \right)^d \right] \leq C_3 \eta_n h_n^{2\beta},$$

where $C_3 = C_3(A_*, B_*, d)$ is a positive constant and $\eta_n = \frac{(\log n)^{\frac{d+1}{2}}}{n^{\frac{d+1}{2}} h_n^{\frac{d}{2}+2\beta} b_n^{\frac{d}{2}}} \to 0$, $n \to +\infty$. Hence, we deduce that

$$\sup_{f \in \Sigma} |\mathrm{I\!I}_{D_n} \mathrm{I\!I}_{G_n} V_{1n}(f)| = o_n\left(h_n^{2\beta}\right).$$

Using the same calculations, we have

$$\sup_{f \in \Sigma} |\mathrm{I\!I}_{D_n} \mathrm{I\!I}_{G_n} V_{2n}(f)| = o_n\left(h_n^{2\beta}\right).$$

To estimate $V_{3n}(f)$, we use the following decomposition

$$V_{3n}(f) = \sum_{k_1 \neq \ldots \neq k_d} \left(V_{3n1}^{(k_1,\ldots,k_d)}(f) + V_{3n2}^{(k_1,\ldots,k_d)}(f) \right), \quad f \in \Sigma,$$

where

$$V_{3n1}^{(k_1,\dots,k_d)}(f) = \frac{2}{n^2}\sum_{i=1}^{n}\psi_n(X_i,X_i) \qquad V_{3n2}^{(k_1,\dots,k_d)}(f) = \frac{2}{n^2}\sum_{i\neq j}\psi_n(X_i,X_j),$$

with

$$\psi_n(u,v) = \int_{[0,1]^d}\left(K_{h_n}(x-u) - \mathbb{E}_f^n K_{h_n}(x-X_i)\right)\Big(K_{*b_n}(x_{k_1}-v_{k_1})$$
$$- \mathbb{E}_f^n K_{*b_n}(x_{k_1}-X_i^{(k_1)})\Big)\prod_{s=2}^{d}\mathbb{E}_f^n K_{*b_n}(x_{k_s}-X_i^{(k_s)})dx.$$

The statistic $V_{3n2}^{(k_1,\dots,k_d)}(f)$ is a degenerate U-statistic of order 2. We have large deviation result similar to Lemma 4.3. Thus, we state the following result

$$\lim_{n\to+\infty}\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{|V_{3n2}^{(k_1,\dots,k_d)}(f)| \geq \sqrt{\frac{\log n}{n^2 h_n}}\right\} = 0.$$

We have

$$V_{3n1}^{(k_1,\dots,k_d)}(f) = \frac{1}{n^2}\sum_{i=1}^{n}\left(\psi_n(X_i,X_i) - \mathbb{E}_f^n\psi_n(X_i,X_i)\right) + \frac{1}{n}\mathbb{E}_f^n\psi_n(X_1,X_1)$$
$$= Z_n(f) + \frac{1}{n}\mathbb{E}_f^n\psi_n(X_1,X_1).$$

Thus, $Z_n(f)$ is a sum of *i.i.d* random variables. We have

$$\lim_{n\to+\infty}\sup_{f\in\Sigma}\mathbb{P}_f^n\left\{|Z_n(f)| \geq \sqrt{\frac{\log n}{n^3 h_n^2}}\right\} = 0.$$

Moreover, for n large enough we obtain

$$\sup_{f\in\Sigma}\left[\frac{1}{n}\mathbb{E}_f^n\psi_n(X_1,X_1)\right] \leq \frac{S^{d+1}}{n},$$

Thus, we have

$$\sup_{f\in\Sigma}\left[\frac{1}{n}\mathbb{E}_f^n\psi_n(X_1,X_1)\right] = o_n\left(h_n^{2\beta}\right).$$

Estimation of $S_{1n}(f)$. We use Lemmas 4.3, 4.4 and the decomposition

$$S_{1n}(f) = \chi_n + U_n + \zeta_n(f) + E_f^n\hat{\theta}_n - \hat{\theta}_n.$$

Let us introduce the random events

$$A_{1,n} = \left\{ \left| V_{3n2}^{(k_1,\dots,k_d)}(f) \right| \le \sqrt{\frac{\log n}{n^2 h_n}} \right\}$$

$$A_{2,n} = \left\{ |Z_n(f)| \le \sqrt{\frac{\log n}{n^3 h_n^2}} \right\}$$

$$A_{3,n} = \left\{ |\chi_n(f)| \le \sqrt{\frac{\log n}{n^3 h_n^{2d}}} \right\}$$

$$A_{4,n} = \left\{ \left| \hat{\theta}_n - \mathbb{E}\hat{\theta}_n \right| \le \sqrt{\frac{\log n}{n^3 h_n^{2d}}} \right\}$$

Putting

$$a_n \triangleq \left(\|\tau_f\|_2 - L_0 h_n^{\beta} \right)^2 - \|\tau_f\|_2 \sqrt{\frac{\log n}{n}} - B_3 \|\tau_f\|_2 \sqrt{\frac{\log n}{n b_n}},$$

we have

$$a_n \ge (\lambda \varphi_n(\alpha_n))^2 \left(1 + \sqrt{\frac{q \log n}{\log \frac{2}{\alpha_n}}} + o_n(1) \right).$$

Then, we can state

$$\begin{aligned}
\mathbb{P}_f^n \{\mathcal{A}_n\} &\le \mathbb{P}_f^n \left\{ S_{1n} + \left(1 + \sqrt{\frac{q \log n}{\log \frac{2}{\alpha_n}}} \right) (\lambda \varphi_n(\alpha_n))^2 \le (\lambda \varphi_n(\alpha_n))^2 \right\} \\
&\quad + \sum_{i=1}^{4} \mathbb{P}_f^n \{H_{in}^c\} + \mathbb{P}_f^n\{D_n^c\} + \mathbb{P}_f^n\{G_n^c\} \\
&\le \mathbb{P}_f^n \left\{ U_n \le -\sqrt{\frac{q \log n}{\log \frac{2}{\alpha_n}}} (\lambda \varphi_n(\alpha_n))^2 \right\} + \sum_{i=1}^{4} \mathbb{P}_f^n \{A_{in}^c\} \\
&\quad + \sum_{i=1}^{4} \mathbb{P}_f^n \{H_{in}^c\} + \mathbb{P}_f^n\{D_n^c\} + \mathbb{P}_f^n\{G_n^c\}
\end{aligned}$$

Using Lemma 4.3, we obtain

$$\mathbb{P}_f^n \{\mathcal{A}_n\} \le n^{-q}(1 + o_n(1)).$$

Thus, since $\|\tau_f\|_2 < +\infty$ for any f, then

$$\bar{R}_n^{(1,2)} \le 2 \left(\varphi_n^{-1}(\alpha_n) \|\tau_f\|_2 + A_1 \right)^q n^{-q} (1 + o_n(1)).$$

Therefore, we have

$$\lim_{n \to \infty} \bar{R}_n^{(1,2)} = 0. \blacksquare$$

5.2 Lower bound

5.2.1 Construction of discrete family of functions

Fix $\sigma \leq 2^{\frac{-4}{4\beta+d}} \Upsilon^{-\frac{1}{4\beta+d}} L_0^{-\frac{4}{4\beta+d}}$ and put $\delta_n = \sigma h_n$, $M_n = \delta_n^{-1}$. Suppose that M_n is an integer. Otherwise, one can take its interger part. Denote $A_{nj} = [u_j, u_{j+1}[$, $j = 1, \ldots, M_n - 1$, $A_{nM_n} = [u_{M_n}, 1]$, where $u_j = \frac{j-1}{M_n}$ for any $j = 1, \ldots, M_n$. Then $\{A_{nj}, j = 1, \ldots, M_n\}$ is a partition of $[0,1]$. For a multi-index $s = (s_1, \ldots, s_d) \in \Xi_n \stackrel{\triangle}{=} \{1, \ldots, M_n\}^d$, define $A_{ns} = A_{ns_1} \times \ldots \times A_{ns_d}$. Then, $\{A_{ns}, s \in \Xi_n\}$ is a partition of $[0,1]^d$.
Let ψ be an infinitely differentiable function with support $[0,1]$ such that

$$\int_{\mathbb{R}} \psi(x)dx = 0 \qquad \int_{\mathbb{R}} \psi^2(x)dx = 1. \tag{5.6}$$

For any $s \in \Xi_n$, introduce the function

$$\psi_{ns}(x_1, \ldots, x_d) = \frac{1}{\delta_n^{d/2}} \prod_{r=1}^{d} \psi\left(\frac{x_r - u_{s_r}}{\delta_n}\right)$$

such that

$$\delta_n^{\beta+d/2} \sup_{x \in [0,1]^d} \left| \psi_{ns}(x) - P_{m,\psi_{ns},u}(x) \right| \leq L\|x-u\|^{\beta}. \tag{5.7}$$

The function ψ_{ns} is compactly supported in A_{ns} and using (5.6), we obtain

$$\int_{\mathbb{R}^d} \psi_{ns}(x)dx = 0, \qquad \int_{\mathbb{R}^d} \psi_{ns}^2(x)dx = 1. \tag{5.8}$$

Put $V = \{-1,1\}^{M_n^d}$. Thus, every $v \in V$ can be written as $v = (v_s)_{s \in \Xi_n}$ where $v_s \in \{-1,1\}$. Then, we introduce the class of functions $\mathcal{F}_n \stackrel{\triangle}{=} \{f_{n,v}, v \in V\}$ where the $f_{n,v}$ is defined as follows

$$f_{n,v}(x) = f_0(x) + \delta_n^{\beta+d/2} \sum_{s \in \Xi_n} v_s \psi_{ns}(x)$$

with $f_0(x) = \mathbb{I}_{[0,1]^d}(x)$. According to (5.7), we deduce that $\mathcal{F}_n \subset \Sigma$.
In the sequel, we use the following notation

$$\begin{aligned}
&\mathbb{P}_0^n = P_{f_0}^n, \quad \mathbb{P}_v^n = P_{f_v}^n, \quad \mathbb{E}_0^n = \mathbb{E}_{f_0}^n, \quad \mathbb{E}_v^n = \mathbb{E}_{f_v}^n \\
&V_s^{(1)} = \{v \in V : \ v_s = 1\}, \qquad V_s^{(-1)} = \{v \in V : \ v_s = -1\}, \\
&V_s^{(0)} = \{v = (v_l)_{l \in \Xi_n} : \forall l \neq s, \ v_l \in \{-1,1\}; \ v_s = 0\}.
\end{aligned}$$

For any $v \in V$, the vector $v^{(s)} = (v_l^{(s)})_{l \in \Xi_n}$ be defined for $l \in \Xi_n$ by

$$v_l^{(s)} = \begin{cases} v_l & \text{if } l \neq s, \\ 0 & \text{if } l = s. \end{cases}$$

5.2.2 Proof of the lower bound

This proof is inspired by these of Lepski [7] and Hoffmann and Lepski [4], but our statistical model involves some difficulties that we need to overcome.
Let ρ_n be an arbitrary random normalizing factor in Ω_n for which

$$\lim_{n\to+\infty} \frac{x_n(\rho_n)}{x_n(\rho_n^*)} = 0. \tag{5.9}$$

We need to prove that

$$\liminf_{n\to+\infty} \inf_{\tilde{f}_n\in\mathcal{M}_n} R_n^{(r)}\left(\tilde{f}_n, \Sigma, \rho_n\right) = +\infty.$$

Let $\mathcal{B}_n = \{\rho_n = x_n(\rho_n)\}$. We have

$$\begin{aligned}
R_n^{(r)}\left(\tilde{f}_n, \Sigma, \rho_n\right) &\geq \sup_{f\in\Sigma} \mathbb{E}_f^n \left\{ \left(x_n^{-1}(\rho_n)\left\|\tilde{f}_n - f\right\|_2\right)^q \mathbb{1}_{\mathcal{B}_n} \right\} \\
&\geq \sup_{v\in V} \mathbb{E}_{f_{n,v}}^n \left\{ \left(x_n^{-1}(\rho_n)\left\|\tilde{f}_n - f_{n,v}\right\|_2\right)^q \mathbb{1}_{\mathcal{B}_n} \right\} \\
&\geq \frac{1}{2^{M_n^d}} \sum_{v\in V} \mathbb{E}_v^n \left\{ \left(x_n^{-1}(\rho_n)\left\|\tilde{f}_n - f_{n,v}\right\|_2\right)^q \mathbb{1}_{\mathcal{B}_n} \right\} \\
&\geq \left(\frac{1}{2^{M_n^d}} \sum_{v\in V} \mathbb{E}_v^n \left(x_n^{-1}(\rho_n)\left\|\tilde{f}_n - f_{n,v}\right\|_2\right)^2 \mathbb{1}_{\mathcal{B}_n} \right)^{\frac{q}{2}} \\
&\triangleq \left(R_n(\tilde{f}_n)\right)^{\frac{q}{2}}.
\end{aligned}$$

We need to introduce the following likelihood ratios

$$Z_s^{(1)} \triangleq \frac{d\mathbb{P}_v^n}{d\mathbb{P}_{v^{(s)}}}(X^n),\;\; v\in V_s^{(1)} \;\text{ and }\; Z_s^{(-1)} \triangleq \frac{d\mathbb{P}_v^n}{d\mathbb{P}_{v^{(s)}}}(X^n),\;\; v\in V_s^{(-1)}.$$

Lemma 5.1. *For all* $\delta\in(0,1)$, *we have*

$$\lim_{n\to+\infty} \sup_{s\in\Xi_n} \sup_{v\in V} \mathbb{P}_{v^{(s)}}^n \left\{ Z_s^{(1)} < 1-\delta \right\} = 0 \qquad\qquad \lim_{n\to+\infty} \sup_{s\in\Xi_n} \sup_{v\in V} \mathbb{P}_{v^{(s)}}^n \left\{ Z_s^{(1)} < 1-\delta \right\} = 0.$$

Let $0<\delta<1$ and put $D_s = \left\{Z_s^{(1)} \geq 1-\delta\right\} \cap \left\{Z_s^{(-1)} \geq 1-\delta\right\}$. Then, for any $v\in V$ we have

$$\mathbb{P}_{v^{(s)}}^n \{D_s^c\} \leq \mathbb{P}_{v^{(s)}}^n \left\{\ln Z_s^{(1)} < \ln(1-\delta)\right\} + \mathbb{P}_{v^{(s)}}^n \left\{\ln Z_s^{(-1)} < \ln(1-\delta)\right\}.$$

Therefore, Lemma 5.1 implies that

$$\lim_{n\to+\infty} \mathbb{P}_{v^{(s)}}^n \{D_s^c\} = 0. \tag{5.10}$$

We have

$$R_n(\tilde{f}_n) = \frac{x_n^{-2}(\rho_n)}{2^{M_n^d}} \sum_{v\in V} \mathbb{E}_v^n \int \left(\tilde{f}_n(x) - f_{n,v}(x)\right)^2 \mathbb{I}_{\mathcal{B}_n} dx$$
$$= \frac{x_n^{-2}(\rho_n)}{2^{M_n^d}} \sum_{s\in\Xi_n} \sum_{v\in V_s^{(1)}} \mathbb{E}_v^n \int_{A_{ns}} \left(\tilde{f}_n(x) - f_{n,v}(x)\right)^2 \mathbb{I}_{\mathcal{B}_n} dx$$
$$+ \frac{x_n^{-2}(\rho_n)}{2^{M_n^d}} \sum_{s\in\Xi_n} \sum_{v\in V_s^{(-1)}} \mathbb{E}_v^n \int_{A_{ns}} \left(\tilde{f}_n(x) - f_{n,v}(x)\right)^2 \mathbb{I}_{\mathcal{B}_n} dx$$
$$= \frac{x_n^{-2}(\rho_n)}{2^{M_n^d}} \sum_{s\in\Xi_n} \sum_{v\in V} \mathbb{E}_{v^{(s)}}^n \mathbb{I}_{\mathcal{B}_n} \Big(Z_s^{(1)} \int_{A_{ns}} \left(\tilde{f}_n(x) - f_0(x) - \delta_n^{\beta+d/2}\psi_{ns}(x)\right)^2 dx$$
$$+ Z_s^{(-1)} \int_{A_{ns}} \left(\tilde{f}_n(x) - f_0(x) + \delta_n^{\beta+d/2}\psi_{n,s}(x)\right)^2 dx \Big). \tag{5.11}$$

Continuing (5.11), we obtain

$$R_n(\tilde{f}_n) \geq (1-\delta)\frac{(x_n(\rho_n))^{-2}}{2^{M_n^d}} \sum_{s\in\Xi_n} \sum_{v\in V} \mathbb{E}_{v^{(s)}}^n \mathbb{I}_{\mathcal{B}_n\cap D_s}$$
$$\times \left(\int_{A_{ns}} \left(\tilde{f}_n(x) - f_0(x) - \delta_n^{\beta+d/2}\right)^2 + \int_{A_{ns}} \left(\tilde{f}_n(x) - f_0(x) + \delta_n^{\beta+d/2}\right)^2 dx \right)$$
$$\geq (1-\delta)\frac{(x_n(\rho_n))^{-2}\delta_n^{2\beta+d}}{2^{M_n^d}} \sum_{s\in\Xi_n} \sum_{v\in V} \mathbb{P}_{v^{(s)}}^n \{\mathcal{B}_n \cap D_s\} \int_{A_{ns}} \psi_{ns}^2(x) dx$$
$$\geq (1-\delta)\left((x_n(\rho_n))^{-1}\delta_n^{\beta+d/2}\right)^2 \sum_{s\in\Xi_n} \frac{1}{2^{M_n^d-1}} \sum_{v\in V_s^{(0)}} (\mathbb{P}_v^n\{\mathcal{B}_n\} - \mathbb{P}_v^n\{D_s^c\})$$

Lemma 5.2. *There exists $p_0 > 0$ such that for any $s \in \Xi_n$*

$$\frac{1}{2^{M_n^d-1}} \sum_{v\in V_s^{(0)}} \mathbb{P}_v^n\{\mathcal{B}_n\} \geq p_0$$

for all n large enough.

According to (5.10), for all $v \in V_s^{(0)}$ and small $\delta > 0$, we have $\mathbb{P}\{D_s^c\} \leq \delta$ for n large enough. Thus, choosing $\delta < p_0/2$,, we obtain

$$R_n(\tilde{f}_n) \geq (1-\delta)\left((x_n(\rho_n))^{-1}\delta_n^{\beta+d/2}\right)^2 M_n^d \frac{p_0}{2}.$$

From the choice of M_n and δ_n, we conclude that

$$R_n(\tilde{f}_n) \geq (1-\delta)\frac{p_0}{2}(\sigma^\beta 2^{4/(4\beta+d)}\Upsilon^{1/(4\beta+d)} L_0^{-4/(4\beta+d)})^2 \left(\frac{\varphi_n(\alpha_n)}{x_n(\rho_n)}\right)^2$$
$$= (1-\delta)\frac{p_0}{2}(\sigma^\beta 2^{4/(4\beta+d)}\Upsilon^{1/(4\beta+d)} L_0^{-4/(4\beta+d)})^2 \left(\frac{x_n(\rho_n^*)}{x_n(\rho_n)}\right)^2.$$

for δ small enough. Next, using (5.9), we have

$$\liminf_{n\to\infty} \inf_{\tilde{f}_n\in\mathcal{M}_n} R_n\left(\tilde{f}_n, \Sigma, \rho_n\right) = +\infty. \blacksquare$$

5.2.3 Proofs of Lemmas

Proof of Lemma 5.1: For all $s \in \Xi_n$, we have

$$\ln(Z_s^{(1)}) = \sum_{i=1}^{n} \ln\left(\frac{f_0(X_i) + \delta_n^{\beta+d/2} \sum\limits_{l\in\Xi_n} v_l \psi_{nl}(X_i)}{f_0(X_i) + \delta_n^{\beta+d/2} \sum\limits_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)}\right)$$

$$= \sum_{i=1}^{n} \ln\left(1 + \frac{\delta_n^{\beta+d/2} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i)}{1 + \delta_n^{\beta+d/2} \sum\limits_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i)}\right)$$

$$\triangleq \sum_{i=1}^{n} B_n^{(s)}(X_i)$$

Using Taylor formula, we have

$$B_n^{(s)}(X_i) \triangleq Q_{1n}^{(i,s)} + Q_{2n}^{(i,s)} + Q_{3n}^{(i,s)},$$

where

$$Q_{1n}^{(i,s)} = \frac{\delta_n^{\beta+d/2} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i)}{1 + \delta_n^{\beta+d/2} \sum\limits_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i)}$$

$$Q_{2n}^{(i,s)} = -\frac{1}{2}\left(\frac{\delta_n^{\beta+d/2} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i)}{1 + \delta_n^{\beta+d/2} \sum\limits_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i)}\right)^2$$

$$Q_{3n}^{(i,s)} = \frac{\theta_{n,i}}{3}\left(\frac{\delta_n^{\beta+d/2} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i)}{1 + \delta_n^{\beta+d/2} \sum\limits_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i)}\right)^3,$$

$\theta_{n,i}$ is a random sequence such that $|\theta_{n,i}| \leq 1$. Applying Taylor formula again, we obtain

$$\begin{aligned} Q_{1n}^{(i,s)} &= \delta_n^{\beta+d/2} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i) - \delta_n^{2\beta+d} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i) \sum_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i) \\ &+ \delta_n^{3\beta+3d/2} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i) \left(\sum_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i)\right)^2 \\ &- \delta_n^{4\beta+2d} \theta_{n,i}^{(1)} v_s \psi_{ns}(X_i)\, \mathrm{II}_{A_{ns}}(X_i) \left(\sum_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i)\right)^3, \end{aligned}$$

where $\theta_{n,i}^{(1)}$ is a random sequence such that $\left|\theta_{n,i}^{(1)}\right| \leq 1$. For $l \neq s$, we have $A_{nl} \cap A_{ns} = \emptyset$. Then,we deduce that

$$\mathrm{II}_{A_{ns}}(X_i) \sum_{l\in\Xi_n: l\neq s} v_l \psi_{nl}(X_i)\, \mathrm{II}_{A_{nl}}(X_i) = 0.$$

Thus, we get

$$Q_{1n}^{(i,s)} = \delta_n^{\beta+d/2} v_s \psi_{ns}(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i).$$

Same calculations imply that

$$\begin{aligned} Q_{2n}^{(i,s)} &= -\frac{\delta_n^{2\beta+d}}{2}\psi_{ns}^2(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i) \\ Q_{3n}^{(i,s)} &= \frac{\theta_{n,i}\delta_n^{3\beta+3d/2}}{3} v_s \psi_{ns}^3(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i). \end{aligned}$$

Thus, we have

$$\begin{aligned} \ln(Z_s^{(1)}) = \delta_n^{\beta+d/2}\sum_{i=1}^n v_s\psi_{ns}(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i) - \frac{\delta_n^{2\beta+d}}{2}\sum_{i=1}^n \psi_{ns}^2(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i) \\ + \frac{\delta_n^{3\beta+3d/2}}{3}\sum_{i=1}^n \theta_{n,i} v_s \psi_{ns}^3(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i). \end{aligned}$$

Put

$$\delta_n^\beta = (\sigma h_n)^\beta \stackrel{\triangle}{=} c_n \frac{1}{\sqrt{n}},$$

where

$$c_n = \sigma^{\beta+\frac{d}{2}} 2^{\frac{2(2\beta+d)}{4\beta+d}} \Upsilon^{\frac{2\beta+d}{2(4\beta+d)}} L_0^{\frac{-2(2\beta+d)}{4\beta+d}} n^{\frac{-d}{2(2\beta+d)}} \left(\log\frac{2}{\alpha_n}\right)^{\frac{2\beta+d}{2(4\beta+d)}}.$$

Then, we have

$$\frac{1}{c_n}\ln(Z_s^{(1)}) = \chi_{1n}^s + \chi_{2n}^s + \chi_{3n}^s$$

where

$$\begin{aligned} \chi_{1n}^s &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_{ns}(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i) \\ \chi_{2n}^s &= -\frac{c_n}{2n}\sum_{i=1}^n \psi_{ns}^2(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i) \\ \chi_{3n}^s &= \frac{c_n^2}{n^{\frac{3}{2}}}\sum_{i=1}^n \theta_{i,n}\psi_{ns}^3(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i). \end{aligned}$$

The random variable χ_{1n}^s is a sum of independent identically distributed centered random variables $\psi_{ns}(X_i)\,\mathrm{1\!I}_{A_{ns}}(X_i)$, $i = 1,\ldots,n$. Therefore, using the Central Limit Theorem, under $P_{\nu^s}^n$,, we have the following convergence in law as $n \to +\infty$

$$\chi_{1n} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{5.12}$$

Since

$$\mathbb{E}^n_{v^s}\,|\chi^s_{2n}| = \frac{c_n}{2}\mathbb{E}^n_{v^s}\psi^2_{ns}(X_1)\,\mathrm{1\!I}_{A_{ns}}(X_1) = \frac{c_n}{2} \to 0, \quad n \to +\infty,$$

then the random variable χ^s_{2n} tends to zero in $\mathbb{P}^n_{v^s}$-probability.
We have

$$\begin{aligned}\mathbb{E}^n_{v^s}\,|\chi^s_{3n}| &= \frac{c_n^2}{\sqrt{n}}\mathbb{E}^n_{v^s}\,|\psi_{ns}(X_1)|^3\,\mathrm{1\!I}_{A_{ns}}(X_1)\\ &= \frac{c_n^2}{\sqrt{n}}\delta_n^{\frac{-d}{2}}\prod_{s=1}^{d}\int_0^1 |\psi(u_s)|^3 du_s \to 0, \quad n \to +\infty.\end{aligned}$$

Therefore, we deduce that the random variable χ^s_{3n} converges to zero in $\mathbb{P}^n_{v^s}$-probability.
The above results imply that $\dfrac{1}{c_n}\ln(Z^{(1)}_s)$ converges to the standard normal random variable $\mathcal{N}(0,1)$ in distribution. Therefore, we deduce that for n large enough

$$\mathbb{P}^n_{v^s}\left\{Z^{(1)}_s < 1-\delta\right\} = \mathbb{P}^n_{v^s}\left\{\frac{\ln(Z^{(1)}_j)}{c_n} < \frac{\ln(1-\delta)}{c_n}\right\} = \phi\left(\frac{\ln(1-\delta)}{c_n}\right),$$

where ϕ is the distribution function of $\mathcal{N}(0,1)$. Since $0<\delta<1$ and c_n tends to zero as $n \to +\infty$, we conclude that

$$\lim_{n\to+\infty}\sup_{s\in\Xi_n}\sup_{v\in V}\mathbb{P}^n_{v^{(s)}}\left\{Z^{(1)}_s < 1-\delta\right\} = 0, \qquad n \to +\infty. \tag{5.13}$$

The same techniques used above imply that

$$\lim_{n\to+\infty}\sup_{s\in\Xi_n}\sup_{v\in V}\mathbb{P}^n_{v^{(s)}}\left\{Z^{(-1)}_s < 1-\delta\right\} = 0, \qquad n \to +\infty.$$

Proof of Lemma 5.2. Put

$$Z_n = \frac{1}{2^{M_n^d-1}}\sum_{v\in V_s^{(0)}}\frac{d\mathbb{P}^n_v}{d\mathbb{P}^n_0}(X^n).$$

Since $c\alpha_n^{-1}\mathbb{P}^n_0\{\mathcal{B}_n^c\} - c \le 0$ for any $c>0$, for n large enough, we obtain

$$\begin{aligned}\frac{1}{2^{M_n^d-1}}\sum_{v\in V_s^{(0)}}\mathbb{P}^n_v\{\mathcal{B}_n\} &= \mathbb{E}^n_0(Z_n\,\mathrm{1\!I}_{\mathcal{B}_n})\\ &\ge \mathbb{E}^n_0\left(Z_n\,\mathrm{1\!I}_{\mathcal{B}_n} + c\alpha_n^{-1}\,\mathrm{1\!I}_{\mathcal{B}_n^c}\right) - c\\ &\ge \mathbb{E}^n_0\left(Z_n\,\mathrm{1\!I}_{\left\{Z_n<c\alpha_n^{-1}\right\}} + c\alpha_n^{-1}\,\mathrm{1\!I}_{\left\{Z_n\ge c\alpha_n^{-1}\right\}}\right) - c\\ &\ge \frac{1}{2^{M_n^d-1}}\sum_{v\in V_s^{(0)}}\mathbb{P}^n_v\left\{Z_n < c\alpha_n^{-1}\right\} - c\\ &\ge (1-c) - c^{-1}\alpha_n\frac{1}{2^{M_n^d-1}}\sum_{v\in V_s^{(0)}}\mathbb{E}^n_v(Z_n)\\ &\ge (1-c) - c^{-1}\alpha_n\mathbb{E}^n_0(Z_n^2)\end{aligned} \tag{5.14}$$

We have the following inequality

$$\mathbb{E}_0^n(Z_n^2) \le \exp\left(2^4 \Upsilon \sigma^{4\beta+d} L_0^{-4} \log\left(\frac{2}{\alpha_n}\right)\right), \tag{5.15}$$

Continuing (5.14), we get

$$\begin{aligned}\frac{1}{2^{M_n^d-1}} \sum_{v \in V_s^{(0)}} \mathbb{P}_v^n\{\mathcal{B}_n\} &\ge (1-c) - c^{-1}\alpha_n \exp\left(16\Upsilon\sigma^{4\beta+d}L_0^{-4}\log\left(\frac{2}{\alpha_n}\right)\right) \\ &= (1-c) - 2^{16\sigma^{4\beta+d}L_0^{-4}} c^{-1} \alpha_n^{1-16\sigma^{4\beta+d}\Upsilon L_0^4}.\end{aligned}$$

If $\sigma \le 2^{-\frac{4}{4\beta+d}} \Upsilon^{-\frac{1}{4\beta+d}} L_0^{-\frac{4}{4\beta+d}}$, $\alpha_n \to 0$ as $n \to +\infty$ then

$$\frac{1}{2^{M_n^d-1}} \sum_{v \in V_s^{(0)}} \mathbb{P}_v^n\{\mathcal{B}_n\} \ge p_0.$$

because c can be chosen arbitrary small numbers.
It remains to prove (5.15). Since the random variables X_i are independent, elementary computation and use of (5.8) imply

$$\mathbb{E}_0^n(Z_n^2) = \frac{1}{2^{2M_n^d-2}} \sum_{v \in V_s^{(0)}} \sum_{v' \in V_s^{(0)}} \left(1 + \delta_n^{2\beta+d} \sum_{l \in \Xi_n} v_l v_l'\right)^n$$

we have

$$\begin{aligned}\mathbb{E}_0^n(Z_n^2) &= \frac{1}{2^{M_n^d-1}} \sum_{v \in V_s^{(0)}} \left[\frac{1}{2^{M_n^d-1}} \sum_{v' \in V_s^{(0)}} \left(1 + \delta_n^{2\beta+d} \sum_{l \in \Xi_n} v_l v_l'\right)^n\right] \\ &= \frac{1}{2^{M_n^d-1}} \sum_{v \in V_s^{(0)}} \left(1 + \delta_n^{2\beta+d} \sum_{l \in \Xi_n} v_l\right)^n \\ &= \frac{1}{2^{M_n^d-1}} \sum_{i=0}^{M_n^d-1} \binom{M_n^d-1}{i} \left(1 + \delta_n^{2\beta+d}\left(M_n^d - 1 - 2i\right)\right)^n \\ &\le (cosh(n\delta_n^{2\beta+d}))^{M_n^d-1} \le \exp\left(n^2 \delta_n^{4\beta+2d} M_n^d\right) \le \exp\left(2^4 \Upsilon \sigma^{4\beta+d} L_0^{-4} \log\left(\frac{2}{\alpha_n}\right)\right). \blacksquare\end{aligned}$$

Acknowledgements

The author gratefully acknowledges O. V. Lepski for his help discussions and comments. Thanks to anonymous referees for careful remarks and comments.

References

[1] Y. Baraud, Confidence balls in nonparametric gaussian regression. *Ann. Statist.* **32** (2004), 528-551.

[2] V. De la Peña and E. Giné, *Decoupling: From Dependence to independence.* Springer-Verlag, New York 1999.

[3] M. Hoffmann, On estimating the diffusion coefficient: parametric versus nonparametric. *Ann. H. Poincaré.* **37** (2001), 339-372.

[4] M. Hoffmann,O. V. Lepski, Random rates in anisotropic regression. *Ann. Statist.* **30** (2002), 325-396.

[5] I. A. Ibragimov, R. Z. Khasminskii, On estimation of distribution density. *Zapiski Nauch. Seminar. LOMI.* **98** (1980), 61-86.

[6] W.B. Johnson, G. Schechtman, J. Zinn, Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.* **13** (1985), 234-253.

[7] O. V. Lepski, How to improve the accuracy of estimation. *Math. Methods. Statist.,* **8** (1999), 441-486.

[8] O. V. Lepski, A.B. Tsybakov, Asymptotically exact nonparametric hypothesis testing in sup-norm and at fixed point. *Probab. Theory Related Fields,* **117** (2000), 17-48.

[9] A. F. Yodé, Asymptotically minimax test of independence. *Math. Methods. Statist.,* **13** (2004), 201-234.