Journal of Mathematics Research; Vol. 12, No. 6; December 2020
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

$\mathbb{L}_p$-Adaptive Estimation Under Partially Linear Constraint in Regression Model

Kouamé Florent Kouakou[1] & Armel Fabrice Evrard Yodé[1]

[1] UFR de Mathématiques et Informatique, Université Félix Houphouët Boigny, 22 BP 582 Abidjan 22, Côte d'Ivoire

Correspondence: Armel Fabrice Evrard Yodé, UFR de Mathématiques et Informatique, Université Félix Houphouët Boigny, 22 BP 582 Abidjan 22, Côte d'Ivoire. E-mail: evrard.yode@univ-fhb.edu.ci

Received: September 16, 2020 Accepted: October 20, 2020 Online Published: November 11, 2020

doi:10.5539/jmr.v12n6p74 URL: https://doi.org/10.5539/jmr.v12n6p74

Abstract

We study the problem of multivariate estimation in the nonparametric regression model with random design. We assume that the regression function to be estimated possesses partially linear structure, where parametric and nonparametric components are both unknown. Based on Goldenshulger and Lepski methodology, we propose estimation procedure that adapts to the smoothness of the nonparametric component, by selecting from a family of specific kernel estimators. We establish a global oracle inequality (under the L_p-norm, $1 \leq p < \infty$) and examine its performance over the anisotropic Hölder space.

Keywords: adaptive estimation, minimax rate, nonparametric regression, oracle inequalities, partially linear model, structural adaptation

1. Introduction

We observe $(X_1, Y_1), \ldots, (X_n, Y_n) \in \mathbb{R}^d \times \mathbb{R}$ satisfying

$$Y_i = g(X_i) + \zeta_i, \quad i = 1, ..., n \tag{1}$$

where $d \geq 2$, g is an unknown function from $[0, 1]^d$ to $\mathbb{R}$, the design points $\{X_i\}_{i=1}^n$ are i.i.d. random variables uniformly distributed on $[0, 1]^d$ and the noise $\{\zeta_i\}_{i=1}^n$ are i.i.d. centered real random variables having symmetric distribution. The sequences $\{\zeta_i\}_{i=1}^n$ and $\{X_i\}_{i=1}^n$ are assumed to be independent. In addition, we assume that g has a partially linear structure, that is, there is an unknown parameter $\beta \in \mathbb{R}^{d_1}$ and an unknown function f defined on $[0, 1]^{d_2}$ with values in $\mathbb{R}$ such that

$$g(X_i) = \beta^{'} U_i + f(T_i) \tag{2}$$

where $X_i = (U_i, T_i)' \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with $d_1 + d_2 = d$ and the prime indicates the transposition. Moreover, we assume that d_1 and d_2 are known and the conditional covariance matrix $\mathbb{E}\Big[(U_1 - \mathbb{E}(U_1|T_1))(U_1 - \mathbb{E}(U_1|T_1))'\Big]$ is non-singular.

Partially linear models are semiparametric models since they contain both parametric and nonparametric components. They are more flexible than the standard linear models and they may be preferred to acompletely nonparametric regression because of the well-known "curse of dimensionality". Partially linear models have many applications. (Engle, Granger, Rice & Weiss, 1986) were among the first to consider this kind of models. They analyzed the relationship between temperature and electricity usage. The bibliography concerning these models is very extensive and we refer readers to (Härdle, Liang & Gao, 2000) and the references therein.

In this paper, using the ideas of (Goldenshulger & Lepski, 2012), we propose estimation procedure that adapts to the smoothness of the nonparametric component f appearing in (2), by selecting from a family of specific kernel estimators, basing on the observation $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$. For the proposed estimator, we establish global oracle inequality and show how to use it for deriving minimax adaptive results if f belongs to anisotropic Hölder space. To the best of our knowledge, Goldenshulger and Lepski method was used to have adaptive estimation in the framework of multivariate regression with random design in (Comte & Lacour, 2013), (Lepski & Serdyukova, 2014) and (Nguyen, 2014). Our work is close to these two last. However, there are major differences. Nguyen(Nguyen, 2014) considered a general regression model while we made a structural assumption about the regression function as in (Lepski & Serdyukova, 2014) who considered the case where the regression function has single-index structure. The problem of adpative estimation of multivariate function was also studied by (Goldenshulger & Lepski, 2008), (Goldenshulger & Lepski, 2009), (Goldenshulger & Lepski, 2011a), (Goldenshulger & Lepski, 2014), (Lepski, 2013a) and (Rebelles, 2015) in the white gaussian noise and density models.

To measure performance of estimators of nonparametric component f, we will use the risk function determined by the $\mathbb{L}_p$-norm $\|\cdot\|_p$, $1 \leq p < \infty$: for function f in (2) and for an arbitrary estimator $\widetilde{f_n}$ based on the observation D_n, we consider the risk:

$$\mathcal{R}_p^{(q)}[\widetilde{f_n}, f] = \Big\{\mathbb{E}_{\beta,f}^{(n)}\Big(\|\widetilde{f_n} - f\|_p^q\Big)\Big\}^{\frac{1}{q}} \quad q \geq 1$$

where $\mathbb{E}_{\beta,f}^{(n)}$ denotes the expectation with respect to the probability measure $\mathbb{P}_{\beta,f}^{(n)}$ of the observation D_n. Here, we note that

$$\|f\|_p = \Big(\int_{\mathbb{R}^{d_2}} |f(x)|^p \mu(dx)\Big)^{\frac{1}{p}}$$

with $\mu(dx) = 1_{[\delta,1-\delta]^{d_2}}(x)dx$ and $0 < \delta < \frac{1}{2}$ is a given number. We restrict the $\mathbb{L}_p$-risk from the cube $[0,1]^{d_2}$ to $[\delta, 1-\delta]^{d_2}$ in order to avoid boundary effects.

Throughout this paper, p, q and δ are supposed to be fixed, $\|.\|$ is the Euclidian norm $\mathbb{R}^{d_1}$ and the distribution of the noise satisfies the following condition.

Assumption 1 *There exists $P > 0$ and $s > \max\{2q, 2p\}$ such that :*

$$\mathbb{E}\Big(|\zeta_1|^s\Big) \leqslant P$$

This assumption represents the link between the noise and the loss function. Its can be found in (Baraud, 2002) for unidimensional case and (Nguyen, 2014) for multidimensional case.

The rest of this paper is organized as follows. In section 2, we present estimation procedure and preliminary results. Section 3 is devoted to main results. We establish oracle inequality and derive minimax adaptive properties of our estimator.

2. Preliminaries

2.1 Selection Rule

In this section we motivate and explain our procedure. Let $\mathcal{L}$ and K be two kernel functions defined respectively on $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ with values in $\mathbb{R}$ satisfying the following assumptions.

Assumption 2

(i) $\mathcal{L} \in \mathbb{C}(\mathbb{R}^{d_1})$*;* $\sup_{x\in\mathbb{R}^{d_1}} |\mathcal{L}(x)| \leqslant \bar{\mathcal{L}}$*, where $\bar{\mathcal{L}} > 0$ is a given number.*

(ii) $\int_{[0,1]^{d_1}} \mathcal{L}(u)du = 1.$

(iii) $\int_{[0,1]^{d_1}} \mathcal{L}(u)u_j du = 0, \forall j = 1, \ldots, d_1.$

Assumption 3

(i) $\int_{\mathbb{R}^{d_2}} K(x)dx = 1$ *and $supp(K) \subseteq [-\delta/2, \delta/2]^{d_2}$, where $supp(K)$ denotes the support of function K.*

(ii) $\sup_{x\in\mathbb{R}^{d_2}} |K(x)| \leqslant \kappa$; $|K(t) - K(z)| \leqslant \kappa|t - z|_\infty, \forall t, z \in \mathbb{R}^{d_2}$ *where $|\cdot|_\infty$ denotes the vector sup-norm in $\mathbb{R}^{d_2}$ and $\kappa > 0$.*

Let $\mathcal{H}_n$ be a given subset $(0,1]^{d_2}$ defined by:

$$\mathcal{H}_n = \Big\{h = (h_1, \ldots, h_{d_2}) : h_j \in H, \forall j = 1, \ldots, d_2, nV_h \geq 1\Big\}$$

where

$$H = \Big\{\exp(-i) : i = 0, \ldots, \lfloor \log n \rfloor\Big\}, \quad n \in \mathbb{N}^* \quad V_h = \prod_{j=1}^{d_2} h_j$$

and $\lfloor \log n \rfloor$ denotes the integer part of $\log n$.

Consider the family of kernel estimators :

$$\mathcal{F}(\mathcal{H}_n) = \Big\{\widehat{f_h}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{L}(U_i)K_h\Big(T_i - x\Big)Y_i;\ h \in \mathcal{H}_n;\ x \in [\delta, 1-\delta]^{d_2}\Big\}.$$

where for any $h = (h_1, ..., h_{d_2}) \in \mathcal{H}_n$, we put $K_h(\cdot) = \frac{1}{V_h}K\Big(\frac{\cdot}{h}\Big)$. Here and later, for any $u, v \in \mathbb{R}^{d_2}$, $\frac{u}{v}$ stands for the coordinate-wise division.

Remark 2.1 *The family of estimators $\mathcal{F}(\mathcal{H}_n)$ does not depend on β. Indeed, we have*

$$\begin{aligned}\mathbb{E}^{(n)}_{\beta,f}\Big(\widehat{f_h}(x)\Big) &= \sum_{j=1}^{d_1}\beta_j \int_{[0,1]^{d_1}} \mathcal{L}(u)u_j du + \int_{[0,1]^{d_2}} K_h(t-x)f(t)dt \\ &= \int_{[0,1]^{d_2}} K_h(t-x)f(t)dt\end{aligned}$$

for all $\beta \in \mathbb{R}^{d_1}$, because (U_i, T_i) are uniformly distributed on $[0,1]^d$ and according to Assumption 2 (iii). Thus, the use of the second kernel $\mathcal{L}$ makes it possible to get rid of the influence of the parameter β.

We propose the data-driven selection from the family $\mathcal{F}(\mathcal{H}_n)$ which leads to the estimator $\widehat{f_n}$. Next, for underlying function f, we find an explicit upper bound for the risk $\mathcal{R}^{(q)}_p[\widehat{f_n}, f]$. More precisely, we prove that

$$\mathcal{R}^{(q)}_p[\widehat{f_n}, f] \leq C \inf_{h\in\mathcal{H}_n} \Big\{\mathcal{R}^{(q)}_p[\widehat{f_h}, f]\Big\} + \gamma_n, \quad f \in \mathcal{F}. \tag{3}$$

Here $C \geq 1$ is an universal constant, the remainder term γ_n is independent of f and

$$\mathcal{F} = \Big\{f : [0,1]^{d_2} \longrightarrow \mathbb{R},\ \sup_{x\in[0,1]^{d_2}} |f(x)| \leq f_\infty\Big\}, \tag{4}$$

where $f_\infty > 0$ is a given number. Let

$$h^* = \arg\inf_{h\in\mathcal{H}_n} \Big\{\mathcal{R}^{(q)}_p[\widehat{f_h}, f]\Big\}.$$

Then, $\widehat{f_{h^*}}$ is called oracle and inequality (3) is called oracle inequality. It guarantees that when γ_n is negligible before $\inf_{h\in\mathcal{H}_n} \Big\{\mathcal{R}^{(q)}_p[\widehat{f_h}, f]\Big\}$, the risk of $\widehat{f_n}$ is of the same order as that of the oracle $\widehat{f_{h^*}}$.

Our selection rule uses auxiliary estimators that are constructed as follows: for $h, \eta \in \mathcal{H}_n$, define the kernel $K_h \star K_\eta$ by

$$K_{h,\eta}(u) = \Big[K_h \star K_\eta\Big](u) = \int K_h(u-w)K_\eta(w)dw.$$

Let $\widehat{f_{h,\eta}}$ denote the estimator associated with this kernel:

$$\widehat{f_{h,\eta}}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{L}(U_i)K_{h,\eta}(T_i - x)\,Y_i \quad x \in [\delta, 1-\delta]^{d_2}. \tag{5}$$

Let $\tau = (\tau_1, ..., \tau_{d_2})$ where $\tau_i = \log^{\frac{-1}{d_2}}(n)$ and consider the following notations:

$$\begin{aligned}\widehat{f_\infty} &= \|\widehat{f_\tau}\|_\infty + 5 \\ \|\widehat{f_\tau}\|_\infty &= \sup_{x\in[\frac{\delta}{2},1-\frac{\delta}{2}]^{d_2}} |\widehat{f_\tau}(x)| \\ \bar{f} &= \sup_{x\in[\frac{\delta}{2},1-\frac{\delta}{2}]^{d_2}} |f(x)| \\ e_p &= \max\Big\{(\mathbb{E}|\zeta_1|^p)^{\frac{1}{p}}, 1\Big\} \\ b(p) &= \begin{cases} 6\kappa, & p \in [1,2] \\ 10\kappa C(p), & p > 2 \end{cases}\end{aligned}$$

where $C(p) = a(p) \vee B_p$ with $a(p) = \frac{15p}{\log(p)}$ and B_p is known constant B_p given in theorem 6.36 (Folland, 1999). For every $h \in \mathcal{H}_n$, let

$$A(h) = b(p)(1+\kappa)\bar{\mathcal{L}}(e_{p\vee 2} + \widehat{f}_\infty + \|\widehat{\beta}\|d_1)\frac{1}{\sqrt{nV_h}} \tag{6}$$

$$\mathcal{R}(h) = \sup_{\eta\in\mathcal{H}_n}\Big[\|\widehat{f}_{h,\eta} - \widehat{f}_\eta\|_p - A(\eta)\Big]_+, \tag{7}$$

where

$$\widehat{\beta} = \Big(\sum_{i=1}^n (U_i - \mathbb{E}[U_i|T_i])(U_i - \mathbb{E}[U_i|T_i])'\Big)^{-1}\Big(\sum_{i=1}^n (U_i - \mathbb{E}[U_i|T_i])(Y_i - \widehat{Y}_i)\Big),$$

with

$$\widehat{Y}_i = \frac{1}{n}\sum_{\substack{j=1\\ j\neq i}}^n Y_j K_\tau\Big(T_i - T_j\Big)$$

is least squares estimator of β proposed by (Robinson, 1988). Here $\mathbb{E}[U_i|T_i]$ is known because (U_i, T_i) are uniformly distributed on $[0,1]^d$. Here, we note that $x_+ = \max(0, x)$. The selected bandwidth $\widehat{h}$ is defined by

$$\mathcal{R}(\widehat{h}) + A(\widehat{h}) \quad = \quad \inf_{h\in\mathcal{H}_n}\left[\mathcal{R}(h) + A(h)\right]. \tag{8}$$

and our final estimator is $\widehat{f}_n := \widehat{f}_{\widehat{h}}$.

2.2 Auxiliary Lemmas

According to Assumption 2, the stochastic term of estimator $\widehat{f}_h$ can be decompose as follows

$$\xi_h(x) = \widehat{f}_h(x) - \mathbb{E}^{(n)}_{\beta,f}\Big(\widehat{f}_h(x)\Big) = \xi^{(1)}_h(x) + \xi^{(2)}_h(x) + \xi^{(3)}_h(x),$$

where

$$\xi^{(1)}_h(x) = \frac{1}{n}\sum_{i=1}^n \Big[f(T_i)\mathcal{L}(U_i)K_h(T_i - x) - \mathbb{E}^{(n)}_{\beta,f}\Big[f(T_i)\mathcal{L}(U_i)K_h(T_i - x)\Big]\Big]$$
$$\xi^{(2)}_h(x) = \frac{1}{n}\sum_{i=1}^n \mathcal{L}(U_i)K_h(T_i - x)\zeta_i$$
$$\xi^{(3)}_h(x) = \frac{1}{n}\sum_{i=1}^n \Big[\sum_{j=1}^{d_1}\beta_j u_j\mathcal{L}(U_i)K_h(T_i - x)\Big].$$

Similarly for the auxiliary estimator, we have

$$\xi_{h,\eta}(x) = \xi^{(1)}_{h,\eta}(x) + \xi^{(2)}_{h,\eta}(x) + \xi^{(3)}_{h,\eta}(x)$$

where

$$\xi^{(1)}_{h,\eta}(x) = \frac{1}{n}\sum_{i=1}^n \Big[f(T_i)\mathcal{L}(U_i)K_{h,\eta}(T_i - x) - \mathbb{E}^{(n)}_{\beta,f}\Big[f(T_i)\mathcal{L}(U_i)K_{h,\eta}(T_i - x)\Big]\Big]$$
$$\xi^{(2)}_{h,\eta}(x) = \frac{1}{n}\sum_{i=1}^n \mathcal{L}(U_i)K_{h,\eta}(T_i - x)\zeta_i$$
$$\xi^{(3)}_{h,\eta}(x) = \frac{1}{n}\sum_{i=1}^n \Big[\sum_{j=1}^{d_1}\beta_j u_j\mathcal{L}(U_i)K_{h,\eta}(T_i - x)\Big].$$

Also, the biases of estimators $\widehat{f}_h(x)$ and $\widehat{f}_{h,\eta}(x)$ are denoted $B_h(x)$ and $B_{h,\eta}(x)$ respectively.

Proposition 2.2 *Suppose that Assumption 3 is fulfilled. For any $\eta \in \mathcal{H}_n$, we have*

$$\sup_{h\in\mathcal{H}_n}\|\xi^{(1)}_{h,\eta} - \xi^{(1)}_\eta\|_p \leq (1+\kappa)\|\xi^{(1)}_\eta\|_p.$$

Proof. For any $x \in \mathbb{R}^{d_2}$

$$\xi^{(1)}_{h,\eta}(x) \quad = \int_{\mathbb{R}^{d_2}} K_h(v-x)\xi^{(1)}_{\eta}(v)dv.$$

Using Young inequality and Assumption 3 we obtain:

$$\|\xi^{(1)}_{h,\eta}\|_p \leq \kappa\|\xi^{(1)}_{\eta}\|_p.$$

Thus, we have

$$\|\xi^{(1)}_{h,\eta} - \xi^{(1)}_{\eta}\|_p \leq (1+\kappa)\|\xi^{(1)}_{\eta}\|_p.$$

and we deduce the result.

Proposition 2.3 *Suppose that $f \in \mathcal{F}$ and assumptions 1, 2 and 4 are fulfilled. Then for all $n \geqslant \widetilde{N} = \max\{\tilde{n}, \bar{n}\}$ we have*

$$\Big(\mathbb{E}^{(n)}_{\beta,f} \sup_{h\in\mathcal{H}_n}\Big[\|\xi^{(1)}_h\|_p - b(p)\frac{1}{\sqrt{nV_h}}\widehat{f_\infty}\bar{\mathcal{L}}\Big]^q_+\Big)^{\frac{1}{q}} \quad \leqslant \quad \boldsymbol{D}\frac{\log^{d_2}(n)}{\sqrt{n}} + T(2q)\log^{d_2}(n)\delta_{B_\tau} + \frac{\log^{d_2+\frac{1}{q}}}{\sqrt{n}}T(2q)\widetilde{T}^{1/2q}(q)$$

where $\widetilde{T}^{1/2q}(q) = (6\exp\{4q\log(4q)\})^{1/2q} + \Big(6\kappa^2P(\bar{\mathcal{L}})^2\Big)^{1/2q}$.

To prove this Proposition we need the following Lemmas. The first is due to (Goldenshulger & Lepski, 2011b). It's an immediate consequence of the Bennett inequality for empirical processes (see Bousquet, 2002) and the standard arguments allowing to derive the Bernstein inequality from the Bennett inequality.

Lemma 2.4 *Let $p \in [1,\infty[$ and suppose that $f \in \mathcal{F}$, then for any $h \in [0,1)^{d_2}$, for any $n \geqslant 2$ et $x \geqslant 0$ we have*

$$\mathbb{P}\Big(\|\xi^{(1)}_h\|_p \geqslant U_1(h,p) + x\Big) \leqslant \exp\Big(\frac{-x^2}{A_1^2(h,p) + B_1(h,p)x}\Big)$$

where

$$B_1(h,p) = \frac{4}{3}\frac{\kappa\bar{\mathcal{L}}\bar{f}}{nV_h}V_h^{1/p}.$$

Lemma 2.5 *Suppose that $f \in \mathcal{F}$ and assumptions 2, 3 are fulfilled. Then for any $n \geqslant 2$ we have:*

$$\Big(\mathbb{E}^{(n)}_{\beta,f} \sup_{h\in\mathcal{H}_n}\Big[\|\xi^{(1)}_h\|_p - b(p)\frac{1}{\sqrt{nV_h}}\bar{f}\bar{\mathcal{L}}\Big]^q_+\Big)^{\frac{1}{q}} \leqslant \boldsymbol{D}\frac{\log^{d_2}(n)}{\sqrt{n}}$$

where

$$\boldsymbol{D} = \begin{cases} 18f_\infty\kappa\bar{\mathcal{L}}(q\Gamma(q))^{1/q}, & p \in [1,2] \\ a(p)\kappa f_\infty\bar{\mathcal{L}}\exp\{p\log(pq) + (p+2)\log(C(p))\}(q\Gamma(q))^{1/q}, & p > 2. \end{cases}$$

Proof. We can write

$$\xi^{(1)}_h(x) = \sum_{i=1}^n W(x,Z_i) \quad Z_i = (T_i, U_i, \zeta_i)$$

where

$$W(x,Z_i) = \frac{1}{nV_h}\Big[K\Big(\frac{T_i-x}{h}\Big)f(T_i)\mathcal{L}(U_i) - \mathbb{E}^{(n)}_{\beta,f}\Big[K\Big(\frac{T_i-x}{h}\Big)f(T_i)\mathcal{L}(U_i)\Big]\Big].$$

For $p \geqslant 1$, there exist a countable subset Υ of the unit ball of $L_q(\mathbb{R}^{d_2})$ with $\frac{1}{q} = 1 - \frac{1}{p}$ such that

$$\|\xi^{(1)}_h\|_p \quad = \quad \sup_{\pi\in\Upsilon}\int_{[\delta,1-\delta]^{d_2}} \pi(x)\xi^{(1)}_h(x)dx.$$

Let's pose

$$\Lambda = \Big\{\lambda(z) = \int \pi(x)W(x,z)dx; \pi \in \Upsilon\Big\}.$$

Note that for all $\lambda \in \Lambda$, we have

$$\mathbb{E}(\lambda(Z_1)) = 0.$$

We need to find the upper bound for $\mathbb{E}\Big(\|\xi_h^{(1)}\|_p\Big)$, $\sup_{\lambda\in\Lambda}\mathbb{E}(\lambda^2(Z_1))$ and $\sup_{\lambda\in\Lambda}\|\lambda\|_\infty$.

If $p \in [1,2]$, we have

$$(\mathbb{E}(\|\xi_h^{(1)}\|_p))^2 \leqslant \mathbb{E}(\|\xi_h^{(1)}\|_2^2) = n\int_{[\delta,1-\delta]^{d_2}} \mathbb{E}W^2(x,Z_1)dx.$$

Therefore, we deduce that

$$\mathbb{E}(\|\xi_h^{(1)}\|_p) \le \frac{\bar{f}\kappa\bar{\mathcal{L}}}{\sqrt{nV_h}}.$$

If $p > 2$, using Rosenthal's inequality, we have

$$\begin{aligned}(\mathbb{E}(\|\xi_h^{(1)}\|_p))^p &\leqslant \int_{[\delta,1-\delta]^{d_2}} \mathbb{E}\Big[\Big|\sum_{i=1}^n W(x,Z_i)\Big|^p\Big]dx\\ &\leqslant \Big(\frac{15p}{\log(p)}\Big)^p\int_{[\delta,1-\delta]^{d_2}}\Big[n\mathbb{E}[|W(x,Z_1)|^p] + \Big(n\mathbb{E}[|W(x,Z_1)|^2]\Big)^{p/2}\Big]dx\\ &\le \Big(\frac{30p\kappa\bar{f}\bar{\mathcal{L}}}{\log(p)}\Big)^p\Big(\frac{1}{n^{p-1}V_h^{p-1}} + \frac{1}{(nV_h)^{p/2}}\Big).\end{aligned}$$

Therefore, we obtain

$$\mathbb{E}\Big(\|\xi_h^{(1)}\|_p\Big) \le U_1(h,p) = \begin{cases}\frac{\bar{\mathcal{L}}\bar{f}\kappa}{\sqrt{nV_h}}, & \text{if } p\in[1,2]\\ \frac{60p\kappa\bar{\mathcal{L}}\bar{f}}{\log(p)\sqrt{nV_h}}, & \text{if } p>2.\end{cases}$$

Using Hölder inequality, we can write

$$|\lambda(z)| \leqslant \|W(.,z)\|_p \leqslant \frac{2}{nV_h}\sup_{(t,u)\in[0,1]^{d_2}\times[0,1]^{d_1}}\Big[\int_{[\delta,1-\delta]^{d_2}}|K(\frac{t-x}{h})f(t)\mathcal{L}(u)|^p dx\Big]^{1/p}.$$

Therefore, we obtain

$$\sup_{\lambda\in\Lambda}\|\lambda\|_\infty \le \begin{cases}\frac{2\kappa\bar{f}\bar{\mathcal{L}}}{nV_h}V_h^{1/2} & \text{if } p\in[1,2]\\ \frac{2\kappa\bar{f}\bar{\mathcal{L}}}{nV_h}V_h^{1/p} & \text{if } p>2\end{cases}$$

Applying Theorem 6.18 and Theorem 6.36 in (Folland, 1999), we obtain

$$\sup_{\lambda\in\Lambda}\mathbb{E}\Big(\lambda^2(Z_1)\Big) \le \begin{cases}\frac{(2\kappa\bar{f}\bar{\mathcal{L}})^2}{(nV_h)^2}V_h^2 & \text{if } p\in[1,2]\\ \frac{(2C(p)\kappa\bar{f}\bar{\mathcal{L}})^2}{(nV_h)^2}V_h^{1+2/p} & \text{if } p>2.\end{cases}$$

Thus, we have

$$\begin{aligned}&2n\sup_{\lambda\in\Lambda}\mathbb{E}(\lambda^2(Z_1)) + 4U_1(h,p)\sup_{\lambda\in\Lambda}\|\lambda\|_\infty \le A_1^2(h,p)\\ &= \begin{cases}\frac{8(\bar{\mathcal{L}}\bar{f}\kappa)^2}{n} + 8\frac{(\bar{\mathcal{L}}\bar{f}\kappa)^2}{nV_h\sqrt{n}}, & p\in[1,2]\\ \frac{8n(C(p)\kappa\bar{f}\bar{\mathcal{L}})^2}{(nV_h)^2}V_h^{1+\frac{2}{p}} + \frac{32(\bar{\mathcal{L}}\bar{f}\kappa)^2 a(p)}{nV_h\sqrt{nV_h}}V_h^{1/p}, & p>2\end{cases}\end{aligned}$$

Applying Lemma 2.4 we have :

$$\begin{aligned}\mathbb{E}\Big[\|\xi_h^{(1)}\|_p - 2U_1(h,p)\Big]_+^q &= q\int_0^{+\infty} x^{q-1}\mathbb{P}\Big\{\|\xi_h^{(1)}\|_p \geqslant 2U_1(h,p)+x\Big\}dx\\ &\leqslant qU_1^q(h,p)\int_0^{+\infty} t^{q-1}\exp\left\{-\frac{(1+t)^2}{\frac{A_1^2(h,p)}{U_1^2(h,p)} + \frac{B_1(h,p)(1+t)}{U_1(h,p)}}\right\}dt.\end{aligned}$$

(i) If $p \in [1,2]$, we obtain

$$\begin{aligned}\mathbb{E}\left[\|\xi_h^{(1)}\|_p - 2U_1(h,p)\right]_+^q &\leqslant q\left(\frac{\bar{\mathcal{L}}\bar{f}\kappa}{\sqrt{nV_h}}\right)^q \int_0^{+\infty} t^{q-1}\exp\Big\{-\frac{(1+t)}{8V_h+\frac{10}{\sqrt{n}}}\Big\}dt \\ &\leqslant q\Gamma(q)(18\kappa\bar{f}\bar{\mathcal{L}})^q n^{\frac{-q}{2}}\end{aligned}$$

(ii) If $p > 2$, using the inequality $\exp(-x) \leqslant \exp(m\log(m))x^{-m}$ for all $x > 0$ and $m > 0$, we obtain

$$\begin{aligned}&\mathbb{E}\left[\|\xi_h^{(1)}\|_p - 2U_1(h,p)\right]_+^q \leqslant \\ &qU_1^q(p,q)\int_0^{+\infty} t^{q-1}\exp\Big(-\frac{(1+t)^2}{\frac{1}{8}C^2(p)V_h^{2/p}+\frac{1}{2}V_h^{1/p}+\frac{1}{2}V_h^{1/p}(1+t)}\Big)dt \\ \leqslant\ & q\Gamma(q)\Big(a(p)\kappa\bar{f}\bar{\mathcal{L}}\exp\{p\log(pq)+(p+2)\log(C(p))\}\Big)^q n^{-q/2}\end{aligned}$$

Thus, we obtain

$$\left(\mathbb{E}_f^{(n)}\left[\|\xi_h^{(1)}\|_p - 2U_1(h,p)\right]_+^q\right)^{\frac{1}{q}} \leqslant \frac{\mathbf{D}}{\sqrt{n}}.$$

and

$$\left(\mathbb{E}_f^{(n)}\sup_{h\in\mathcal{H}_n}\left[\|\xi_h^{(1)}\|_p - b(p)\frac{1}{\sqrt{nV_h}}\bar{f}\bar{\mathcal{L}}\right]_+^q\right)^{\frac{1}{q}} \leqslant \mathbf{D}\frac{\log^{d_2}(n)}{\sqrt{n}}$$

Lemma 2.6 *Suppose that $f \in \mathcal{F}$ and assumptions 2, 3 are fulfilled, then for any $n \geqslant 2$ we have*

$$\left(\mathbb{E}\left[\|\xi_h^{(1)}\|_p^q\right]\right)^{\frac{1}{q}} \leqslant \frac{T(q)}{\sqrt{nV_h}}$$

where $T(q) = 4a(p\vee q)\kappa\bar{\mathcal{L}}f_\infty$

Proof. For $p \in [1,\infty)$, using Fubini's Theorem and Rosenthal's Theorem

$$\mathbb{E}[\|\xi_h^{(1)}\|_p^p] \leqslant \frac{2^{p+1}(\kappa\bar{\mathcal{L}}f_\infty a(p))^p}{(nV_h)^{p/2}}$$

If $q \leqslant p$, we have

$$\left(\mathbb{E}[\|\xi_h^{(1)}\|_p^q]\right)^{1/q} \leqslant \left(\mathbb{E}[\|\xi_h^1\|_p^p]\right)^{1/p} \leqslant \frac{4a(p)\kappa\bar{\mathcal{L}}f_\infty}{\sqrt{nV_h}}$$

If $q > p$, we have

$$\left(\mathbb{E}[\|\xi_h^{(1)}\|_p^q]\right)^{1/q} \leqslant \left(\mathbb{E}[\|\xi_h^1\|_q^q]\right)^{1/q} \leqslant \frac{4a(q)\kappa\bar{\mathcal{L}}f_\infty}{\sqrt{nV_h}}$$

We obtain

$$\left(\mathbb{E}[\|\xi_h^{(1)}\|_p^q]\right)^{1/q} \leqslant \frac{4a(q\vee p)\kappa\bar{\mathcal{L}}f_\infty}{\sqrt{nV_h}}.$$

Lemma 2.7 *Suppose that $f \in \mathcal{F}$ and the Assumptions 1, 2 and 3 are fulfilled. Then there exists integer $\widetilde{N}$ such that for any $n \geqslant \widetilde{N}$, we have*

$$\mathbb{P}\left\{\bar{f} \geqslant \hat{f}_\infty\right\} \leqslant \delta_{B_\tau} + \frac{6\exp\{4q\log(4q)\}}{n^q} + \frac{6\kappa^2 P(\bar{\mathcal{L}})^2\log^2(n)}{n^q}.$$

Proof. Recall that

$$\begin{aligned}\xi_\tau^{(1)}(x) &= \frac{1}{n}\sum_{i=1}^n\Big[\mathcal{L}(U_i)K_\tau(T_i-x)f(T_i) - \mathbb{E}_{\beta,f}^{(n)}(\mathcal{L}(U_i)K_\tau(T_i-x)f(T_i))\Big] \\ \bar{\xi}_\tau^{(2)}(x) &= \frac{1}{n}\sum_{i=1}^n\mathcal{L}(U_i)K_\tau(T_i-x)\zeta_i 1_{\{|\zeta_i|\leqslant\sqrt{n}\}} \\ \tilde{\xi}_\tau^{(2)}(x) &= \frac{1}{n}\sum_{i=1}^n\mathcal{L}(U_i)K_\tau(T_i-x)\zeta_i 1_{\{|\zeta_i|>\sqrt{n}\}} \\ \xi_\tau^{(3)}(x) &= \frac{1}{n}\sum_{i=1}^n\Big[\sum_{j=1}^{d_1}\beta_j u_j\mathcal{L}(U_i)K_\tau(T_i-x)\Big].\end{aligned}$$

We can write

$$\begin{aligned}\mathbb{P}\Big(\bar{f} \geqslant \hat{f}_\infty\Big) &\leqslant \mathbb{P}\Big(\|B_\tau\|_\infty \geqslant 1\Big) + \mathbb{P}\Big(\|\xi_\tau^{(1)}\|_\infty \geqslant 1\Big) + \mathbb{P}\Big(\|\bar{\xi}_\tau^{(2)}\|_\infty \geqslant 1\Big) + \mathbb{P}\Big(\|\tilde{\xi}_\tau^{(2)}\|_\infty \geqslant 1\Big)\\ &+ \mathbb{P}\Big(\|\xi_\tau^{(3)}\|_\infty \geqslant 1\Big).\end{aligned}$$

Using Proposition 1 in (Lepski, 2013b), with $\psi(.) = |.|$, $\epsilon = \sqrt{2} - 1$, $y = 2n^{1/4}$, $c = 1$, there exists integer $\tilde{n}$ such that for all $n > \tilde{n}$

$$\mathbb{P}\Big(\|\xi_\tau^{(1)}\|_\infty \geqslant 1\Big) \leqslant 2\exp\Big(-n^{1/4}\Big), \tag{9}$$

$$\mathbb{P}\Big(\|\bar{\xi}_\tau^{(2)}\|_\infty \geqslant 1\Big) \leqslant 2\exp\Big(-n^{1/4}\Big), \tag{10}$$

$$\mathbb{P}\Big(\|\xi_\tau^{(3)}\|_\infty \geqslant 1\Big) \leqslant 2\exp\Big(-n^{1/4}\Big). \tag{11}$$

Now, we consider $\mathbb{P}\left\{\|\tilde{\xi}_\tau^{(2)}\|_\infty \geqslant 1\right\}$. First, we remark that

$$n\mathbb{E}\Big[|\mathcal{L}(U_1)||\zeta_1|1_{\{|\zeta_1|>\sqrt{n}\}}\Big] \leqslant \bar{\mathcal{L}}Pn^{1/2},$$

$$\|\tilde{\xi}_\tau^{(2)}\|_\infty \leqslant \frac{\kappa\log(n)}{n}\sum_{i=1}^{n}|\mathcal{L}(U_i)||\zeta_i|1_{\{|\zeta_i|>\sqrt{n}\}}.$$

Putting $\bar{n} = \arg\sup_{n\in\mathbb{N}}\{n : \sqrt{n} - 2P\kappa\bar{\mathcal{L}}\log(n) \leqslant 0\}$, we obtain for all $n > \bar{n}$

$$\begin{aligned}\mathbb{P}\left\{\|\tilde{\xi}_\tau^{(2)}\|_\infty \geqslant 1\right\} &\leqslant \mathbb{P}\Big(\sum_{i=1}^{n}\Big[|\mathcal{L}(U_i)||\zeta_i|1_{\{|\zeta_i|>\sqrt{n}\}} - \mathbb{E}|\mathcal{L}(U_1)||\zeta_1|1_{\{|\zeta_1|>\sqrt{n}\}}\Big]\\ &+ n\mathbb{E}|\mathcal{L}(U_1)||\zeta_1|1_{\{|\zeta_1|>\sqrt{n}\}} \geqslant \frac{n}{\kappa\log(n)}\Big)\\ &\leqslant \mathbb{P}\left\{\sum_{i=1}^{n}\Big[|\mathcal{L}(U_i)||\zeta_i|1_{\{|\zeta_i|>\sqrt{n}\}} - \mathbb{E}|\mathcal{L}(U_1)||\zeta_1|1_{\{\zeta_1>\sqrt{n}\}}\Big] \geqslant \frac{n}{2\kappa\log(n)}\right\}\\ &\leqslant \left(\frac{2\kappa\log(n)}{n}\right)^2\mathbb{E}|\sum_{i=1}^{n}\Big(|\mathcal{L}(U_i)||\zeta_i|1_{\{|\zeta_i|>\sqrt{n}\}} - \mathbb{E}|\mathcal{L}(U_1)||\zeta_1|1_{\{|\zeta_1|>\sqrt{n}\}}\Big)|^2\\ &\leqslant 4\kappa^2\log^2(n)(\bar{\mathcal{L}})^2Pn^{-s/2}\end{aligned}$$

If $s > 2q$ then for all $n > \bar{n}$, we have

$$\mathbb{P}\left(\|\tilde{\xi}_\tau^{(2)}\|_\infty\right) \leqslant 4\kappa^2\log^2(n)(\bar{\mathcal{L}})^2Pn^{-q}.$$

Thus, We obtain for all $n > \widetilde{N} = \max\{\tilde{n}, \bar{n}\}$

$$\begin{aligned}\mathbb{P}\Big(\bar{f} \geqslant \hat{f}_\infty\Big) &\leqslant \delta_{B_\tau} + 6\exp\Big(-n^{1/4}\Big) + 4\kappa^2P\log^2(n)(\bar{\mathcal{L}})^2n^{-q}\\ &\leqslant \delta_{B_\tau} + \frac{6\exp\{4q\log(4q)\}}{n^q} + \frac{6\kappa^2P(\bar{\mathcal{L}})^2\log^2(n)}{n^q}\end{aligned}$$

Here, we have used the inequality $\exp\{-x\} \leqslant \exp\{m\log(m)\}x^{-m}$ for all $x > 0$ and $m > 0$.

Proof of Proposition 2.3. Let us define $A = \{\bar{f} \geqslant \hat{f}_\infty\}$. Using Lemma 2.5 we obtain

$$\begin{aligned}\Big(\mathbb{E}_{\beta,f}^{(n)}\sup_{h\in\mathcal{H}_n}\Big[\|\xi_h^{(1)}\|_p - b(p)\frac{1}{\sqrt{nV_h}}\hat{f}_\infty\bar{\mathcal{L}}\Big]_+^q\Big)^{\frac{1}{q}} &= \left(\mathbb{E}_{\beta,f}^{(n)}\sup_{h\in\mathcal{H}_n}\left[\|\xi_h^{(1)}\|_p - b(p)\frac{1}{\sqrt{nV_h}}\hat{f}_\infty\bar{\mathcal{L}}\right]_+^q 1_{\bar{A}}\right)^{\frac{1}{q}}\\ &+\Big(\mathbb{E}_{\beta,f}^{(n)}\sup_{h\in\mathcal{H}_n}\Big[\|\xi_h^{(1)}\|_p - b(p)\frac{1}{\sqrt{nV_h}}\hat{f}_\infty\bar{\mathcal{L}}\Big]_+^q 1_A\Big)^{\frac{1}{q}}.\\ &\leqslant \mathbf{D}\frac{\log^{d_2}(n)}{\sqrt{n}} + \sum_{h\in\mathcal{H}_n}\Big(\mathbb{E}_{\beta,f}^{(n)}[\|\xi_h^{(1)}\|_p^{2q}]\Big)^{1/2q}(\mathbb{P}_{\beta,f}^{(n)}(A))^{1/2q}\end{aligned}$$

According to Lemmas 2.6 and 2.7, we deduce that there exists integer $\widetilde{N}$ such that for all $n > \widetilde{N}$

$$\Big(\mathbb{E}^{(n)}_{\beta,f}\sup_{h\in\mathcal{H}_n}\Big[\|\xi^{(1)}_h\|_p - b(p)\frac{1}{\sqrt{nV_h}}\hat{f}_\infty\bar{\mathcal{L}}\Big]^q_+\Big)^{\frac{1}{q}} \leqslant \mathbf{D}\frac{\log^{d_2}(n)}{\sqrt{n}} + \log^{d_2}(n)T(2q)\delta_{B_\tau} + \frac{\log^{d_2+\frac{1}{q}}}{\sqrt{n}}T(2q)\widetilde{T}^{1/2q}(q).$$

Proposition 2.8 *Suppose that Assumptions 1, 2 and 3 are fulfilled. Then, there exists a constant $n_1 \geq 2$ such that for any $n \geqslant n_1$, we have*

$$\Big(\mathbb{E}^{(n)}_{\beta,f}\sup_{h\in\mathcal{H}_n}\Big[\|\xi^{(2)}_h\|_p - b(p)\bar{\mathcal{L}}e_{p\vee 2}\frac{1}{\sqrt{nV_h}}\Big]^q_+\Big)^{\frac{1}{q}} \leqslant [\boldsymbol{D}_1 + C_1(p\vee q)]\frac{\log^{d_2}(n)}{\sqrt{n}}.$$

Proof. Using Proposition 2.2, we have

$$\|\xi^{(2)}_{h,\eta} - \xi^{(2)}_\eta\|_p \leqslant (1+\kappa)\|\xi^{(2)}_\eta\|_p.$$

Put $\mathcal{M} = \mathcal{M}(\epsilon) = n^{\frac{1}{2}-\epsilon} > 0$ where

$$0 < \epsilon < \varpi := \frac{s-2l}{2s-2l}; \quad l = p\vee q.$$

The condition $s > \max\{2p, 2q\}$ given in Assumption 1 ensures that $0 < \varpi < \frac{1}{2}$.

We have the following decomposition

$$\xi^{(2)}_h(x) = \bar{\xi}^{(2)}_h(x) + \tilde{\xi}^{(2)}_h(x)$$

where

$$\bar{\xi}^{(2)}_h(x) = \frac{1}{nV_h}\sum_{i=1}^n \mathcal{L}(U_i)K\Big(\frac{T_i - x}{h}\Big)\zeta_i 1_{\{|\zeta_i|\leqslant\mathcal{M}\}}$$

$$\tilde{\xi}^{(2)}_h(x) = \frac{1}{nV_h}\sum_{i=1}^n \mathcal{L}(U_i)K\Big(\frac{T_i - x}{h}\Big)\zeta_i 1_{\{|\zeta_i|>\mathcal{M}\}}$$

Put

$$\bar{\xi}^{(2)}_h(x) = \sum_{i=1}^n W(x, Z_i) \qquad Z_i = (T_i, U_i, \zeta_i)$$

$$W(x,Z_i) = \frac{1}{nV_h}\Big[\mathcal{L}(U_i)K\Big(\frac{T_i-x}{h}\Big)\zeta_i 1_{\{|\zeta_i|\leqslant\mathcal{M}\}} - \mathbb{E}\Big[\mathcal{L}(U_i)K\Big(\frac{T_1-x}{h}\Big)\zeta_1 1_{\{|\zeta_1|\leqslant\mathcal{M}\}}\Big]\Big].$$

Since ζ_1 is centered and symmetric, we have

$$W(x,z) = \frac{1}{nV_h}\Big[\mathcal{L}(u)K\Big(\frac{t-x}{h}\Big)y 1_{\{|y|\leqslant\mathcal{M}\}}\Big]$$

For $p \geqslant 1$, we have

$$\|\bar{\xi}^{(2)}_h\|_p = \sup_{\pi\in\Upsilon}\int_{[\delta,1-\delta]^{d_2}}\pi(x)\bar{\xi}^{(2)}_h(x)dx.$$

We need to find the upper bound for $\mathbb{E}\Big(\|\bar{\xi}^{(2)}_h\|_p\Big)$, $\sup_{\lambda\in\Lambda}\mathbb{E}(\lambda^2(Z_1))$ and $\sup_{\lambda\in\Lambda}\|\lambda\|_\infty$.

If $p\in[1,2]$, we can write

$$\mathbb{E}[\|\bar{\xi}^{(2)}_h\|_p] \leqslant \frac{\bar{\mathcal{L}}\kappa e_2}{\sqrt{nV_h}}$$

If $p > 2$, using Rosenthal's inequality, we have :

$$\left(\mathbb{E}\left(\|\bar{\xi}_h^{(2)}\|_p\right)\right)^p \leqslant (a(p))^p \int_{[\delta,1-\delta]^{d_2}} \left[n\mathbb{E}|W(x,Z_1)|^p + (n\mathbb{E}|W(x,Z_1)|^2)^{p/2}\right] dx$$

As

$$\mathbb{E}[|W(x,Z_1|^p] \leq \frac{\left(\kappa\bar{\mathcal{L}}e_p\right)^p}{n^p V_h^{p-1}} \quad \text{and} \quad \left(n\mathbb{E}[|W(x,Z_1)|^2]\right)^{p/2} \leq \frac{\left(\kappa\bar{\mathcal{L}}e_p\right)^p}{(nV_h)^{p/2}}$$

then, we deduce that

$$\mathbb{E}\left(\|\bar{\xi}_h^{(2)}\|_p\right) \leqslant \frac{2a(p)\kappa\bar{\mathcal{L}}e_p}{\sqrt{nV_h}}.$$

Thus, we have

$$\mathbb{E}\left(\|\bar{\xi}_h^{(2)}\|_p\right) \leq U_2(h,p) = \begin{cases} \frac{\bar{\mathcal{L}}\kappa e_2}{\sqrt{nV_h}}, & \text{if } p \in [1,2] \\ \frac{2a(p)\kappa\bar{\mathcal{L}}e_p}{\sqrt{nV_h}}, & \text{if } p > 2. \end{cases}$$

We can write $|\lambda(z)| \leqslant \|W(.,z)\|_p$ and

$$\|W(.,z)\|_p^p \leqslant \frac{1}{(nV_h)^p} \sup_{(t,u)\in[0,1]^{d_2}\times[0,1]^{d_1}} \int_{[\delta,1-\delta]^{d_2}} |\mathcal{L}(u)K(\frac{t-x}{h})y1_{\{|y|\leqslant\mathcal{M}\}}|^p dx$$

Therefore, we obtain

$$\sup_{\lambda\in\Lambda} \|\lambda\|_\infty \leq \begin{cases} \frac{\kappa\bar{\mathcal{L}}\mathcal{M}}{nV_h} V_h^{1/2} & \text{if } p \in [1,2] \\ \frac{\kappa\bar{\mathcal{L}}\mathcal{M}}{nV_h} V_h^{1/p} & \text{if } p > 2. \end{cases}$$

Applying Theorem 6.18 and Theorem 6.36 in (Folland, 1999), we obtain

$$\sup_{\lambda\in\Lambda} \mathbb{E}\left(\lambda^2(Z_1)\right) \leq \begin{cases} \frac{(\kappa\bar{\mathcal{L}}e_2)^2}{n^2} & \text{if } p \in [1,2] \\ \frac{(C(p)\kappa\bar{\mathcal{L}}e_2)^2}{(nV_h)^2} V_h^{1+2/p} & \text{if } p > 2. \end{cases}$$

$$\begin{aligned} 2n \sup_{\lambda\in\Lambda} \mathbb{E}(\lambda^2(Z_1)) + 4U_2(h,p) \sup_{\lambda\in\Lambda} \|\lambda\|_\infty \leq A_2^2(h,p) \\ = \begin{cases} \frac{2(\bar{\mathcal{L}}\kappa e_2)^2}{n} + \frac{4(\bar{\mathcal{L}}\kappa)^2 e_2\mathcal{M}}{nV_h\sqrt{n}}, & p \in [1,2] \\ \frac{2(C(p)\kappa\bar{\mathcal{L}}e_p)^2}{nV_h} V_h^{\frac{2}{p}} + \frac{8(\bar{\mathcal{L}}\kappa)^2 a(p)e_p\mathcal{M}}{nV_h\sqrt{nV_h}} V_h^{1/p}, & p > 2 \end{cases} \end{aligned}$$

$$\frac{2}{3} \sup_{\lambda\in\Lambda} \|\lambda\|_\infty \leq B_2(h,p) = \frac{2}{3} \frac{\kappa\bar{\mathcal{L}}\mathcal{M}}{nV_h} V_h^{1/p}, p \geqslant 1.$$

Applying Lemma 2.4 with $U_2(h,p), A_2^2(h,p)$ and $B_2(h,p)$ we have

$$\begin{aligned} \mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q &\leqslant q\int_0^{+\infty} x^{q-1}\mathbb{P}\{\|\bar{\xi}_h^{(2)}\|_p \geqslant 2U_2(h,p) + x\}dx \\ &\leqslant qU_2^q(h,p)\int_0^{+\infty} t^{q-1} exp\left\{-\frac{(1+t)^2}{\frac{A_2^2(h,p)}{U_2^2(h,p)} + \frac{B_2(h,p)}{U_2(h,p)}(1+t)}\right\} dt \end{aligned}$$

(i) If $p \in [1,2]$, we have

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\left(\frac{\bar{\mathcal{L}}\kappa e_2}{\sqrt{nV_h}}\right)^q \int_0^{+\infty} t^{q-1}\exp\left\{-\frac{(1+t)^2}{2V_h + 4n^{-\epsilon} + n^{-\epsilon}(1+t)}\right\} dt$$

If $V_h \leqslant n^{-\epsilon}$, then we have

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(7\bar{\mathcal{L}}\kappa e_2)^q n^{-q\epsilon}\exp\{-\frac{n^\epsilon}{7}\}$$

There exists $n_1 = \arg\sup_{n\in\mathbb{N}}\{n^{-q\epsilon}\exp(-\frac{n^\epsilon}{7}) - n^{-q/2} \geqslant 0\}$ such that for any $n > n_1$, we have

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(7\kappa\bar{\mathcal{L}}e_2)^q n^{-q/2}.$$

If $V_h > n^{-\epsilon}$

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(7\kappa\bar{\mathcal{L}}e_2)^q n^{-q/2}.$$

If $p \in [1,2]$, we conclude that

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(7\kappa\bar{\mathcal{L}}e_2)^q n^{-q/2}, \forall n > n_1$$

(ii) If $p > 2$

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\left(\frac{2a(p)\kappa\bar{\mathcal{L}}e_p}{\sqrt{nV_h}}\right)^q \int_0^{\infty} t^{q-1}\exp\left\{-\frac{(1+t)}{6^{-1}C^2(p)[V_h^{2/p} + n^{-\epsilon} + n^{-\epsilon}]}\right\} dt$$

If $V_h^{2/p} \leqslant n^{-\epsilon}$

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(a(p)\kappa\bar{\mathcal{L}}e_p)^q n^{-\epsilon q}C^{2q}(p)\exp\left\{-\frac{n^\epsilon}{C^2(p)}\right\}$$

There exists $n_1 = \arg\sup_{n\in\mathbb{N}}\left\{C^2(p)\exp\left\{\frac{-n^\epsilon}{qC^2(p)}\right\} - n^{\epsilon-1/2} \geqslant 0\right\}$ such that for all $n > n_1$

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(a(p)\kappa\bar{\mathcal{L}}e_p)^q n^{-q/2}$$

If $V_h^{2/p} > n^{-\epsilon}$ then

$$\mathbb{E}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q \leqslant q\Gamma(q)(a(p)\kappa\bar{\mathcal{L}}e_p\mathbf{T})^q n^{-q/2}$$

where $\mathbf{T} = \mathbf{T}(p,q) = \sup_{t\in[0,1]} C^2(p)t^{-1/2+2/p}\exp\{\frac{-1}{(qt^{2/p}C^2(p))}\}$.

Recall that for all $h \in [0,1)^{d_2}$, we have $\xi_h^{(2)}(x) = \bar{\xi}_h^{(2)}(x) + \tilde{\xi}_h^{(2)}(x)$ and for all $n > n_1$,

$$\begin{aligned}\left(\mathbb{E}_f^{(n)}\sup_{h\in\mathcal{H}_n}\left[\|\xi_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q\right)^{\frac{1}{q}} &\leqslant \left(\mathbb{E}_f^{(n)}\sup_{h\in\mathcal{H}_n}\left[\|\bar{\xi}_h^{(2)}\|_p - 2U_2(h,p)\right]_+^q\right)^{\frac{1}{q}} \\ &+ \left(\mathbb{E}\sup_{h\in\mathcal{H}_n}\|\tilde{\xi}_h^{(2)}\|_p^q\right)^{1/q} \\ &\leqslant \mathbf{D}_1\frac{\log^{d_2}(n)}{\sqrt{n}} + \sum_{h\in\mathcal{H}_n}\left(\mathbb{E}\|\tilde{\xi}_h^{(2)}\|_p^q\right)^{1/q} \\ &\leqslant [\mathbf{D}_1 + C_1(p\vee q)]\frac{\log^{d_2}(n)}{\sqrt{n}}\end{aligned}$$

where

$$C_1(l) = \begin{cases} \kappa\bar{\mathcal{L}}(2P)^{1/l}, & l \in [1,2] \\ a(l)\kappa\bar{\mathcal{L}}(P + P^{l/2})^{1/l}, & l > 2. \end{cases}$$

Proposition 2.9 *Suppose that assumptions 2 and 3 are fulfilled. Then there exists integer $\bar{N}$ such that for all $n > \bar{N}$ we have*

$$\Big(\mathbb{E}^{(n)}_{\beta,f} \sup_{h\in\mathcal{H}_n} \Big[\|\xi_h^{(3)}\|_p - \frac{b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1}{\sqrt{nV_h}}\Big]_+^q\Big)^{1/q} \leqslant \frac{\log^{d_2}(n)}{\sqrt{n}}\boldsymbol{D}_2 + T_1(2q)\frac{\log^{d_2+\frac{1}{q}}(n)}{\sqrt{n}}$$

where $T_1(q) = 3a(p \vee q)\|\beta\|\kappa\bar{\mathcal{L}}d_1$

To prove Proposition 2.9, we need the following Lemma

Lemma 2.10 *Suppose that assumptions 2 and 3 are fulfilled, then for all $n \geqslant 2$:*

$$\Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_p^q\Big]\Big)^{\frac{1}{q}} \leqslant \frac{T_1(q)}{\sqrt{nV_h}}$$

where $T_1(q) = 3a(p \vee q)\kappa\bar{\mathcal{L}}\|\beta\|d_1$

Proof. For $p \in [1;\infty)$, using Fubini's theorem and Rosenthal's inequality we have:

$$\mathbb{E}\Big[\|\xi_h^{(3)}\|_p^p \leqslant \frac{2(a(p)\kappa\bar{\mathcal{L}}\|\beta\|d_1)^p}{(nV_h)^{p/2}}$$

If $q \leqslant p$, we have

$$\Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_p^q\Big]\Big)^{1/q} \leqslant \Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_p^p\Big]\Big)^{1/p} \leqslant \frac{3a(p)\kappa\bar{\mathcal{L}}\|\beta\|d_1}{\sqrt{nV_h}}$$

If $q > p$, we have

$$\Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_p^q\Big]\Big)^{1/q} \leqslant \Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_q^q\Big]\Big)^{1/q} \leqslant \frac{3a(q)\kappa\bar{\mathcal{L}}\|\beta\|d_1}{\sqrt{nV_h}}$$

Therefore, we deduce that

$$\Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_p^q\Big]\Big)^{1/q} \leqslant \frac{3a(p \vee q)\kappa\bar{\mathcal{L}}\|\beta\|d_1}{\sqrt{nV_h}}.$$

Proof of Proposition 2.9. We can write

$$\xi_h^{(3)}(x) = \sum_{i=1}^{n} W(x, Z_i)$$

where

$$W(x, Z_i) = \frac{1}{nV_h}\sum_{j=1}^{d_1}\beta_j u_j \mathcal{L}(U_i)K(\frac{T_i - x}{h})$$

For $p \geqslant 1$, we have

$$\|\xi_h^{(3)}\|_p = \sup_{\pi\in\Upsilon}\int_{[\delta,1-\delta]^{d_2}} \pi(x)\xi_h^{(3)}(x)dx.$$

We need to find the upper bound for $\mathbb{E}\Big(\|\xi_h^{(3)}\|_p\Big)$, $\sup_{\lambda\in\Lambda}\mathbb{E}(\lambda^2(Z_1))$ and $\sup_{\lambda\in\Lambda}\|\lambda\|_\infty$.

If $p \in [1, 2]$, we can write

$$\mathbb{E}\Big(\|\xi_h^{(3)}\|_p\Big) \leqslant \frac{\|\beta\|\bar{\mathcal{L}}\kappa d_1}{\sqrt{nV_h}}$$

If $p > 2$, using Rosenthal's inequality, we have :

$$\left(\mathbb{E}\left(\|\xi_h^{(3)}\|_p\right)\right)^p \leqslant (a(p))^p \int_{[\delta,1-\delta]^{d_2}} \left[n\mathbb{E}|W(x,Z_1)|^p + (n\mathbb{E}|W(x,Z_1)|^2)^{p/2}\right] dx$$

As

$$\mathbb{E}[|W(x,Z_1|^p] \leq \frac{\left(\|\beta\|\kappa\bar{\mathcal{L}}d_1\right)^p}{n^p V_h^{p-1}} \quad \text{and} \quad \left(n\mathbb{E}[|W(x,Z_1)|^2]\right)^{p/2} \leq \frac{\left(\|\beta\|\kappa\bar{\mathcal{L}}d_1\right)^p}{(nV_h)^{p/2}}$$

we deduce that

$$\mathbb{E}\left(\|\xi_h^{(3)}\|_p\right) \leqslant \frac{2a(p)\|\beta\|\kappa\bar{\mathcal{L}}d_1}{\sqrt{nV_h}}.$$

Thus, we have

$$\mathbb{E}\left(\|\xi_h^{(3)}\|_p\right) \leq U_3(h,p) = \begin{cases} \frac{\|\beta\|\bar{\mathcal{L}}\kappa d_1}{\sqrt{nV_h}}, & \text{if } p \in [1,2] \\ \frac{2a(p)\kappa\|\beta\|\bar{\mathcal{L}}d_1}{\sqrt{nV_h}}, & \text{if } p > 2 \end{cases}$$

We can write $|\lambda(z)| \leqslant \|W(.,z)\|_p$. Therefore, we obtain

$$\sup_{\lambda\in\Lambda} \|\lambda\|_\infty \leq \begin{cases} \dfrac{\kappa\bar{\mathcal{L}}\|\beta\|d_1}{nV_h} V_h^{1/2} & \text{if } p \in [1,2] \\ \dfrac{\kappa\bar{\mathcal{L}}\|\beta\|d_1}{nV_h} V_h^{1/p} & \text{if } p > 2 \end{cases}$$

Applying Theorem 6.18 and Theorem 6.36 in (Folland, 1999), we obtain

$$\sup_{\lambda\in\Lambda} \mathbb{E}\left(\lambda^2(Z_1)\right) \leq \begin{cases} \dfrac{(\kappa\bar{\mathcal{L}}d_1\|\beta\|)^2}{(nV_h)^2} V_h^2 & \text{if } p \in [1,2] \\ \dfrac{(C(p)\kappa\bar{\mathcal{L}}d_1\|\beta\|)^2}{(nV_h)^2} V_h^{1+2/p} & \text{if } p > 2. \end{cases}$$

$$\begin{aligned} &2n \sup_{\lambda\in\Lambda} \mathbb{E}(\lambda^2(Z_1)) + 4U_3(h,p) \sup_{\lambda\in\Lambda} \|\lambda\|_\infty \leq A_3^2(h,p) \\ = &\begin{cases} \frac{2(\bar{\mathcal{L}}\kappa\|\beta\|d_1)^2}{n} + \frac{4(\bar{\mathcal{L}}\kappa\|\beta\|d_1)^2}{nV_h\sqrt{n}}, & p \in [1,2] \\ \frac{2(C(p)\kappa\bar{\mathcal{L}}\|\beta\|d_1)^2}{nV_h} V_h^{2/p} + \frac{8a(p)(\bar{\mathcal{L}}\kappa\|\beta\|d_1)^2}{nV_h\sqrt{nV_h}} V_h^{1/p}, & p > 2 \end{cases} \end{aligned}$$

$$\frac{2}{3} \sup_{\lambda\in\Lambda} \|\lambda\|_\infty \leq B_3(h,p) = \frac{2}{3} \frac{\kappa\bar{\mathcal{L}}\|\beta\|d_1}{nV_h} V_h^{1/p}, p \geqslant 1$$

Applying Lemma 2.4 with $U_3(h,p), A_3^2(h,p)$ and $B_3(h,p)$ we have

(i) If $p \in [1,2]$

$$\mathbb{E}\left[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\right]_+^q \leqslant q\Big(\frac{\|\beta\|\kappa\bar{\mathcal{L}}d_1}{\sqrt{nV_h}}\Big)^q \int_0^\infty t^{q-1} \exp\Big(-\frac{(1+t)}{2V_h + \frac{5}{\sqrt{n}}}\Big) dt$$

If $V_h \leqslant \frac{1}{\sqrt{n}}$ then

$$\mathbb{E}\left[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\right]_+^q \leqslant q\Gamma(q)(7\|\beta\|\kappa\bar{\mathcal{L}}d_1)^q n^{-q/2}$$

If $V_h > \frac{1}{\sqrt{n}}$ then

$$\mathbb{E}\Big[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\Big]_+^q \leqslant q\Gamma(q)(7\|\beta\|\kappa\bar{\mathcal{L}}d_1)^q n^{-q/2}$$

If $p \in [1,2]$, we conclude that

$$\mathbb{E}\Big[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\Big]_+^q \leqslant q\Gamma(q)(7\|\beta\|\kappa\bar{\mathcal{L}}d_1)^q n^{-q/2}$$

(ii) If $p > 2$, we have

$$\mathbb{E}\Big[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\Big]_+^q \leqslant q\Gamma(q)(a(p)\|\beta\|\kappa\bar{\mathcal{L}}d_1 \exp\{p\log(pq) + (p+2)\log(C(p))\})^q n^{-q/2}.$$

Here, we have used the inequality $exp(-x) \leqslant exp(m\log(m))x^{-m}$ for all $x > 0$ and $m > 0$.

Thus, we obtain that for all $n \geqslant 2$

$$\mathbb{E}\Big[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\Big]_+^q \leqslant \frac{\mathbf{D}_2}{\sqrt{n}}$$

and finally, we deduce that

$$\Big(\mathbb{E}_f^{(n)} \sup_{h\in\mathcal{H}_n} \Big[\|\xi_h^{(3)}\|_p - 2U_3(h,p)\Big]_+^q\Big)^{\frac{1}{q}} \leqslant \mathbf{D}_2 \frac{\log^{d_2}(n)}{\sqrt{n}}$$

Let us define $\mathbf{B} = \Big[\|\beta\| \geqslant \|\hat{\beta}\|\Big]$, suppose that there exist

$$\overline{N} = \sup\Big\{n : n^{4q}\mathbb{P}\Big(\|\beta\| \geq \|\widehat{\beta}\|\Big) - \log^2(n) \geq 0\Big\}$$

$$\begin{aligned}
\Big(\mathbb{E}\sup_{h\in\mathcal{H}_n}\Big[\|\xi_h^{(3)}\|_p - \frac{b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1}{\sqrt{nV_h}}\Big]_+^q\Big)^{1/q} &= \Big(\mathbb{E}\sup_{h\in\mathcal{H}_n}\Big[\|\xi_h^{(3)}\|_p - \frac{b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1}{\sqrt{nV_h}}\Big]_+^q \mathbf{1_B}\Big)^{1/q} \\
&+ \Big(\mathbb{E}\sup_{h\in\mathcal{H}_n}\Big[\|\xi_h^{(3)}\|_p - \frac{b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1}{\sqrt{nV_h}}\Big]_+^q \mathbf{1}_{\bar{\mathbf{B}}}\Big)^{1/q} \\
&\leqslant \mathbf{D}_2\frac{\log^{d_2}(n)}{\sqrt{n}} + \sum_{h\in\mathcal{H}_n}\Big(\mathbb{E}[\|\xi_h^{(3)}\|_p^{2q}]\Big)^{1/2q}\Big(\mathbb{P}(\mathbf{B})\Big)^{1/2q} \\
&\leqslant \frac{\log^{d_2}(n)}{\sqrt{n}}\mathbf{D}_2 + T_1(2q)\frac{\log^{d_2+\frac{1}{q}}(n)}{\sqrt{n}},\ \forall n > \overline{N}.
\end{aligned}$$

Proposition 2.11 *Suppose that assumptions 1, 2 and 3 are fulfilled. Then there exists $n_3 \geq 2$ such that for any $n \geqslant n_3$.*

$$\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\widehat{f_\infty}\Big]^q\Big)^{1/q} \leqslant 69b(p)\bar{\mathcal{L}} + 3b(p)f_\infty\kappa\Big(\bar{\mathcal{L}}\Big)^2.$$

Proof. Recall that

$$\begin{aligned}
\hat{f}_\infty &= \|\hat{f}_\tau\|_\infty + 5 \\
\|\hat{f}_\tau\|_\infty + 5 &\leqslant \|\xi_\tau\|_\infty + \|\mathbb{E}[\hat{f}_\tau]\|_\infty + 5
\end{aligned}$$

Taking the account that

$$\hat{f}_\tau(x) - \mathbb{E}[\hat{f}_\tau(x)] = \xi_\tau^{(1)}(x) + \xi_\tau^{(2)}(x) + \xi_\tau^{(3)}(x)$$

We obtain

$$\begin{aligned}
\Big(\mathbb{E}[b(p)\bar{\mathcal{L}}\widehat{f_\infty}]^q\Big)^{\frac{1}{q}} &\leqslant 3b(p)\bar{\mathcal{L}}\|\mathbb{E}[\widehat{f_\tau}]\|_\infty + 9b(p)\bar{\mathcal{L}}\Big(\mathbb{E}[\|\xi_\tau^{(1)}\|_\infty^q 1_{\{\|\xi_\tau^{(1)}\|_\infty \geqslant 1\}}]\Big)^{\frac{1}{q}} \\
&+ 18b(p)\bar{\mathcal{L}}\Big(\mathbb{E}[\|\bar{\xi}_\tau^{(2)}\|_\infty^q 1_{\{\|\bar{\xi}_\tau^{(2)}\|_\infty \geqslant 1\}}]\Big)^{\frac{1}{q}} \\
&+ 18b(p)\bar{\mathcal{L}}\Big(\mathbb{E}[\|\tilde{\xi}_\tau^{(2)}\|_\infty^q]\Big)^{1/q} + 9b(p)\bar{\mathcal{L}}\Big(\mathbb{E}[\|\xi_\tau^{(3)}\|_\infty^q 1_{\{\|\xi_\tau^{(3)}\|_\infty \geqslant 1\}}]\Big)^{1/q} \\
&+ 15b(p)\bar{\mathcal{L}}.
\end{aligned}$$

and

$$\|\mathbb{E}[\hat{f}_\tau]\|_\infty \leqslant \kappa\bar{\mathcal{L}}f_\infty.$$

In view of (9), (10) and (11) we have respectively:

$$\Big(\mathbb{E}[\|\xi_\tau^{(1)}\|_\infty^q 1_{\{\|\xi_\tau^{(1)}\|_\infty \geqslant 1\}}]\Big)^{\frac{1}{q}} \leqslant 4\kappa\bar{\mathcal{L}}f_\infty \log(n)\exp\Big(-\frac{n^{1/4}}{q}\Big) \quad (12)$$

$$\Big(\mathbb{E}[\|\bar{\xi}_\tau^{(2)}\|_\infty^q 1_{\{\|\bar{\xi}_\tau^{(2)}\|_\infty \geqslant 1\}}]\Big)^{1/q} \leqslant 2\kappa\bar{\mathcal{L}}\sqrt{n}\log(n)\exp\Big(\frac{-n^{1/4}}{q}\Big) \quad (13)$$

$$\Big(\mathbb{E}[\|\xi_\tau^{(3)}\|_\infty^q 1_{\{\|\xi_\tau^{(3)}\|_\infty \geqslant 1\}}]\Big)^{1/q} \leqslant 2\bar{\mathcal{L}}\kappa\|\beta\|d_1\log(n)\exp\Big(\frac{-n^{1/4}}{q}\Big) \quad (14)$$

Using Markov inequality, we get

$$\begin{aligned}\Big(\mathbb{E}[\|\tilde{\xi}_\tau^{(2)}\|_\infty^q]\Big)^{1/q} &\leqslant \frac{\kappa\bar{\mathcal{L}}\log(n)}{n}\sum_{i=1}^n\Big(\mathbb{E}[|\zeta_i|^q 1_{\{|\zeta_i| > \sqrt{n}\}}]\Big)^{1/q} \\ &\leqslant \kappa\bar{\mathcal{L}}\log(n)Pn^{\frac{q-s}{2}}. \quad (15)\end{aligned}$$

There exists integer n_3 such that for all $n \geqslant n_3$ all right hand-sides of (12), (13), (14) and (15) are smaller than 1. From all the above, it can be deduced that

$$\Big(\mathbb{E}[b(p)\bar{\mathcal{L}}\widehat{f_\infty}]^q\Big)^{\frac{1}{q}} \leqslant 69b(p)\bar{\mathcal{L}} + 3b(p)f_\infty\kappa\Big(\bar{\mathcal{L}}\Big)^2, \forall n \geqslant n_3$$

Proposition 2.12 *There exists $n_5 \in \mathbb{N}$ such that $\forall n \geqslant n_5$ we have for any $1 \leqslant q \leqslant 2$*

$$\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1\Big]^q\Big)^{\frac{1}{q}} \leqslant 2b(p)\bar{\mathcal{L}}d_1^{\frac{5q-2}{2q}}.$$

Proof. Recall that

$$\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\|\hat{\beta}\|d_1\Big]^q\Big)^{\frac{1}{q}} = b(p)\bar{\mathcal{L}}d_1\Big(\mathbb{E}[\|\hat{\beta}\|^q]\Big)^{1/q}$$

Then, we obtain

$$\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1\Big]^q\Big)^{\frac{1}{q}} \leqslant 2b(p)\bar{\mathcal{L}}d_1^{\frac{3q-2}{2q}}\Big[\sum_{j=1}^{d_1}\Big(\mathbb{E}[|\widehat{\beta}_j - \beta_j|^q]\Big)^{1/q} + d_1\|\beta\|\Big].$$

Using the fact that for $j = 1, ..., d_1, \widehat{\beta}_j - \beta_j = O_{\mathbb{P}}(n^{-1/2})$ then we can write for all $\omega > 0$, there exist $\gamma > 0$ such that

$$\mathbb{P}\Big\{|\hat{\beta}_j - \beta_j| \geqslant \gamma n^{-1/2}\Big\} < \omega.$$

Using Hölder inequality, we have

$$\begin{aligned}\mathbb{E}\Big[|\hat{\beta}_j - \beta_j|^q\Big] &= \mathbb{E}\Big[|\hat{\beta}_j - \beta_j|^q 1_{\{|\hat{\beta}_j - \beta_j|^q \geqslant (\gamma n^{-1/2})^q\}}\Big] + \mathbb{E}\Big[|\hat{\beta}_j - \beta_j|^q 1_{\{|\hat{\beta}_j - \beta_j|^q < (\gamma n^{-1/2})^q\}}\Big] \\ &\leqslant \Big(\mathbb{E}\Big[|\hat{\beta}_j - \beta_j|^2\Big]\Big)^{q/2}\Big(\mathbb{P}\Big\{|\hat{\beta}_j - \beta_j|^q \geqslant (\gamma n^{-1/2})^q\Big\}\Big)^{1-\frac{q}{2}} + (\gamma n^{-1/2})^q\end{aligned}$$

Using the fact that for $j = 1, ..., d_1, \mathbb{E}\Big[(\hat{\beta}_j - \beta_j)^2\Big] = O(n^{-1})$, that is there exist $C > 0, N \in \mathbb{N}$ such that for any $n \geqslant N$, $\mathbb{E}\Big[|\widehat{\beta}_j - \beta_j|^2\Big] \leqslant \frac{C}{n}$, we have

$$\Big(\mathbb{E}\Big[|\hat{\beta}_j - \beta_j|^q\Big]\Big)^{1/q} \leqslant \Big[C^{1/2}\omega^{\frac{2-q}{2q}} + \gamma\Big]\frac{1}{\sqrt{n}}. \quad (16)$$

there exist n_4 such that for all $n \geqslant n_4$ the right-hand term of (16) is smaller than 1. It can be deduced that

$$\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\|\hat{\beta}\|d_1\Big]^q\Big)^{\frac{1}{q}} \leqslant 2b(p)\bar{\mathcal{L}}d_1^{\frac{5q-2}{2q}}.\ \forall n \geqslant n_5 = \max\{n_4, N\}.$$

3. Main Results

3.1 Oracle Inequality

Define $B_\tau(x) = \int_{[0,1]^{d_2}} K_\tau(u-x)f(u)du - f(x)$ and put also

$$\delta_{B_\tau} = \begin{cases} 1 & \text{if } \|B_\tau\|_\infty \geq 1 \\ 0 & \text{if } \|B_\tau\|_\infty < 1. \end{cases}$$

Note that B_τ is the bias of estimator $\widehat{f_\tau}$. Let $\widehat{f_n}$ be the estimator obtained from the selection rule (6), (7), (8).

Theorem 3.1 *Suppose that Assumptions 1, 2 and 3 are fulfilled. Then, there exists integer N_1 such that for any $n \geqslant N_1$ and f belongs to $\mathcal{F}$ defined by* (4), *we have for any $1 \leqslant q \leqslant 2$*

$$\mathcal{R}_p^{(q)}\Big[\widehat{f_n}, f\Big] \leqslant 2(1+\kappa)\inf_{h\in\mathcal{H}_n}\left\{\mathcal{R}_p^{(q)}\Big[\widehat{f_h}, f\Big] + \frac{\boldsymbol{d}}{\sqrt{nV_h}}\right\} + 2(1+\kappa)\widetilde{Q}\frac{\log^{d_2}(n)}{\sqrt{n}}$$
$$+ 2(1+\kappa)\boldsymbol{d}_1\frac{\log^{d_2+\frac{1}{q}}(n)}{\sqrt{n}} + 2(1+\kappa)T(2q)\delta_{B_\tau}\log^{d_2}(n)$$

where $\widetilde{Q}$, $\boldsymbol{d}$, $\boldsymbol{d}_1$ and $T(2q)$ are constants depending on p, q, f_∞, κ, and $\bar{\mathcal{L}}$.

Remark 3.2

(i) The oracle inequality of theorem 3.1 is the key technical tool bounding $\mathbb{L}_p$-risk of this estimator on anisotropic Hölder classes $H_{d_2}(\alpha, L)$.

(ii) The main difficulty in extending it to $q > 2$ is to use the square risk of $\widehat{\beta}_j - \beta_j$. Indeed, analyzing the proof of proposition 2.12, we remark that the following relations should be fufilled

$$\begin{aligned} \mathbb{E}\left[|\hat{\beta}_j - \beta_j|^q\right] &= \mathbb{E}\left[|\hat{\beta}_j - \beta_j|^q 1_{\{|\hat{\beta}_j-\beta_j|^q \geqslant (\gamma n^{-1/2})^q\}}\right] \\ &+ \mathbb{E}\left[|\hat{\beta}_j - \beta_j|^q 1_{\{|\hat{\beta}_j-\beta_j|^q < (\gamma n^{-1/2})^q\}}\right] \\ &\leqslant \left(\mathbb{E}\left[|\hat{\beta}_j - \beta_j|^2\right]\right)^{q/2}\left(\mathbb{P}\left\{|\hat{\beta}_j - \beta_j|^q \geqslant (\gamma n^{-1/2})^q\right\}\right)^{1-\frac{q}{2}} \\ &+ (\gamma n^{-1/2})^q \mathbb{P}\left\{|\hat{\beta}_j - \beta_j|^q < (\gamma n^{-1/2})^q\right\} \end{aligned}$$

If $1 \leqslant q \leqslant 2$, these relation hold. However, we were not able to obtain this relation in the case $q > 2$. Note nevertheless that if such relation would be found ours results could be extented to $q > 2$.

Proof. Thanks to the triangular inequality, formula (7) and the definition of $\widehat{h}$, we have for any $h \in \mathcal{H}_n$

$$\|\widehat{f_n} - f\|_p \leqslant 2[\mathcal{R}(h) + A(h)] + \|\widehat{f_h} - f\|_p.$$

Using Fubini theorem and Young inequality, for any $h, \eta \in \mathcal{H}_n$, we have

$$\begin{aligned} \|\widehat{f_{h,\eta}} - \widehat{f_\eta}\|_p - A(\eta) &\leq \|\xi_{h,\eta}^{(1)} - \xi_h^{(1)}\|_p + \|\xi_{h,\eta}^{(2)} - \xi_h^{(2)}\|_p + \|\xi_{h,\eta}^{(3)} - \xi_h^{(3)}\|_p \\ &+ \|B_{h,\eta} - B_h\|_p - A(\eta) \\ &\leq \|\xi_{h,\eta}^{(1)} - \xi_h^{(1)}\|_p + \|\xi_{h,\eta}^{(2)} - \xi_h^{(2)}\|_p + \|\xi_{h,\eta}^{(3)} - \xi_h^{(3)}\|_p \\ &+ \kappa\|B_h\|_p - A(\eta) \end{aligned}$$

According to (Goldenshluger & Lepski, 2012), Proposition 2.2 and formula (6), we obtain

$$\begin{aligned} \|\widehat{f_{h,\eta}} - \widehat{f_\eta}\|_p - A(\eta) \leq (1+\kappa)&\sup_{\eta\in\mathcal{H}_n}\left[\|\xi_\eta^{(1)}\|_p - b(p)\bar{\mathcal{L}}\widehat{f_\infty}\frac{1}{\sqrt{nV_\eta}}\right]_+ \\ + (1+\kappa)&\sup_{\eta\in\mathcal{H}_n}\left[\|\xi_\eta^{(2)}\|_p - b(p)\bar{\mathcal{L}}e_{p\vee 2}\frac{1}{\sqrt{nV_\eta}}\right]_+ \\ + (1+\kappa)&\sup_{\eta\in\mathcal{H}_n}\left[\|\xi_\eta^{(3)}\|_p - b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1\frac{1}{\sqrt{nV_\eta}}\right]_+ + \kappa\mathcal{R}_p^{(q)}[\hat{f}_h, f]. \end{aligned}$$

Thus, using Propositions 2.8, 2.9, 2.11 and 2.12, we deduce that there exists integer N_1 such that for all $n \geqslant N_1$ we get

$$\begin{aligned}
\mathcal{R}_p^{(q)}[\widehat{f_{\hat{h}}}, f] \leqslant & 2(1+\kappa)\Big(\mathbb{E}_{\beta,f}^{(n)} \sup_{\eta\in\mathcal{H}_n}\Big[\|\xi_\eta^{(1)}\|_p - b(p)\frac{1}{\sqrt{nV_\eta}}\bar{\mathcal{L}}\widehat{f}_\infty\Big]_+^q\Big)^{\frac{1}{q}} \\
& +2(1+\kappa)\Big(\mathbb{E}_{\beta,f}^{(n)} \sup_{\eta\in\mathcal{H}_n}\Big[\|\xi_\eta^{(2)}\|_p - b(p)\bar{\mathcal{L}}e_{p\vee 2}\frac{1}{\sqrt{nV_\eta}}\Big]_+^q\Big)^{\frac{1}{q}} \\
& +2(1+\kappa)\Big(\mathbb{E}_{\beta,f}^{(n)} \sup_{\eta\in\mathcal{H}_n}\Big[\|\xi_\eta^{(3)}\|_p - b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1\frac{1}{\sqrt{nV_\eta}}\Big]_+^q\Big)^{\frac{1}{q}} \\
& +2(1+\kappa)\mathcal{R}_p^{(q)}[\widehat{f_h}, f] + 2(1+\kappa)\frac{1}{\sqrt{nV_h}}b(p)\bar{\mathcal{L}}e_{p\vee 2} \\
& +2(1+\kappa)\frac{1}{\sqrt{nV_h}}\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\widehat{f}_\infty\Big]^q\Big)^{\frac{1}{q}} + 2(1+\kappa)\frac{1}{\sqrt{nV_h}}\Big(\mathbb{E}\Big[b(p)\bar{\mathcal{L}}\|\widehat{\beta}\|d_1\Big]^q\Big)^{\frac{1}{q}}
\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
\mathcal{R}_p^{(q)}[\widehat{f_{\hat{h}}}, f] \leq & 2(1+\kappa)\Big(\mathbf{D} + \mathbf{D}_1 + C_1(p\vee q) + \mathbf{D}_2\Big)\frac{\log^{d_2}(n)}{\sqrt{n}} \\
& +2(1+\kappa)\inf_{h\in\mathcal{H}_n}\Big[(\bar{\mathcal{L}}b(p)e_{p\vee 2} + 69b(p)\bar{\mathcal{L}} + 3b(p)f_\infty\kappa\bar{\mathcal{L}}^2 + 2b(p)\bar{\mathcal{L}}d_1^{\frac{5q-2}{2q}})\frac{1}{\sqrt{nV_h}} \\
& +\mathcal{R}_p^{(q)}[\widehat{f_h}, f]\Big] + 2(1+\kappa)\Big(T(2q)\widetilde{T}^{1/2q}(q) + T_1(2q)\Big)\frac{\log^{d_2+\frac{1}{q}}(n)}{\sqrt{n}} \\
& +2(1+\kappa)T(2q)\log^{d_2}(n)\delta_{B_\tau}.
\end{aligned}$$

3.2 Adaptive Estimation

In this section, we use the previously established oracle inequalities to study adaptive properties of estimators $\widehat{f_n}$ over the scale of anisotropic Hölder classes $H_{d_2}(\alpha, L)$, $\alpha \in (0, \mathbf{l}]^{d_2}$ and $L > 0$, where $\mathbf{l} > 0$ is fixed.

Definition 3.3 *Let $\alpha = (\alpha_1, \ldots, \alpha_{d_2})$ with $\alpha_i > 0$ for all $i = 1, \ldots, d_2$ and $L > 0$. We say that function f defined on $\mathbb{R}^{d_2}$ with values in $\mathbb{R}$ belongs to the anisotropic Hölder class $H_{d_2}(\alpha, L)$ if for all $i = 1, \ldots, d_2$ and $t \in \mathbb{R}^{d_2}$*

$$|D_i^{(m)} f(t)| \leqslant L, \quad \forall\ m = 1, \ldots, \lfloor\alpha_i\rfloor$$

$$\Big|D_i^{\lfloor\alpha_i\rfloor} f(t_1, \ldots, t_i + z, \ldots, t_{d_2}) - D_i^{\lfloor\alpha_i\rfloor} f(t_1, \ldots, t_i, \ldots, t_{d_2})\Big| \leqslant L|z|^{\alpha_i - \lfloor\alpha_i\rfloor}$$

where $D_i^{(m)} f$ denotes the mth-order partial derivative of f with respect to the variable t_i and $\lfloor u\rfloor$ is the largest integer strictly less than u

In this section, we assume that the kernel function K satisfies Assumption 3 and

Assumption 4

(a) $K(u) = K(u_1, \ldots, u_{d_2}) = \prod_{j=1}^{d_2} \mathcal{K}(u_j)$, where $u \in \mathbb{R}^{d_2}$ and $\mathcal{K}$ is an unidimensional kernel function.

(b) $\mathcal{K}(t) = \mathcal{K}(-t)$, for all $t \in \mathbb{R}$.

(c) $\mathcal{K}$ is compactly supported with $support(\mathcal{K}) \subseteq [-\frac{\delta}{2}, \frac{\delta}{2}]$

(d) $\|\mathcal{K}\|_\infty < +\infty$

(e) $\int \mathcal{K}(t)t^j dt = 0$, for all $0 < j < \mathbf{l}$.

Fix $\alpha = (\alpha_1, \ldots, \alpha_{d_2}) \in (1, \mathbf{l})^{d_2}$, and put $\psi_n(\alpha) = n^{-\bar{\alpha}/(2\bar{\alpha}+1)}$ where

$$1/\bar{\alpha} = 1/\alpha_1 + \cdots + 1/\alpha_{d_2}.$$

Theorem 3.4 *Suppose that Assumptions 1, 2, 3 and 4 are fulfilled. Then for all* $\alpha \in (1, \mathbf{l})^{d_2}$*, we have*

$$\limsup_{n\longrightarrow\infty} \sup_{f\in H_{d_2}(\alpha,L)} \left\{\psi_n^{-1}(\alpha)\mathcal{R}_p^{(q)}\left[\widehat{f_n}, f\right]\right\} < +\infty.$$

Proof. Using Assumptions 3 and 4, for any $\alpha \in (1, \mathbf{l}]^{d_2}$, $h \in \mathcal{H}_n$ and any $p \geq 1$, we have

$$\|B_h\|_p \leqslant L\kappa \sum_{i=1}^{d_2} h_i^{\alpha_i}, \quad f \in H_{d_2}(\alpha, L).$$

Then, we note that there exists integer $N = N(L, \kappa, \alpha)$ for all $n \geq N$, we obtain $\delta_{B_\tau} = 0$ where $\tau = (\tau_1, \ldots, \tau_{d_2})$ with $\tau_i = (\log(n))^{-1/d_2}$. Moreover, for any $h \in \mathcal{H}_n$, we have

$$\begin{aligned}
\mathcal{R}_p^{(q)}\Big(\widehat{f_h}, f\Big) &\leqslant \|B_h\|_p + \Big(\mathbb{E}\|\xi_h^{(1)}\|_p^q\Big)^{1/q} + \Big(\mathbb{E}\|\xi_h^{(2)}\|_p^q\Big)^{1/q} + \Big(\mathbb{E}\|\xi_h^{(3)}\|_p^q\Big)^{1/q} \\
&\leqslant \|B_h\|_p + \Big(\mathbb{E}\Big[\|\xi_h^{(1)}\|_p - \frac{b(p)\bar{\mathcal{L}}\widehat{f_\infty}}{\sqrt{nV_h}}\Big]_+^q\Big)^{1/q} + \Big(\mathbb{E}\Big[\|\xi_h^{(2)}\|_p - \frac{b(p)\bar{\mathcal{L}}e_{p\vee 2}}{\sqrt{nV_h}}\Big]_+^q\Big)^{1/q} \\
&+ \Big(\mathbb{E}\Big[\|\xi_h^{(3)}\|_p - \frac{b(p)\bar{\mathcal{L}}d_1\|\widehat{\beta}\|}{\sqrt{nV_h}}\Big]_+^q\Big)^{1/q} \\
&+ \frac{1}{\sqrt{nV_h}}\Big(\mathbb{E}[b(p)\bar{\mathcal{L}}\widehat{f_\infty}]^q)^{1/q} + b(p)\bar{\mathcal{L}}e_{p\vee 2} + (\mathbb{E}[b(p)\bar{\mathcal{L}}d_1\|\widehat{\beta}\|]^q)^{1/q}\Big) \\
&\leq \|B_h\|_p + C_1^*\frac{\log^{d_2}(n)}{\sqrt{n}} + C_2^*\frac{\log^{d_2+\frac{1}{q}}(n)}{\sqrt{n}} + C_3^*\frac{1}{\sqrt{nV_h}}.
\end{aligned}$$

Consider the following system of equations:

$$h_i^{\alpha_i} = h_k^{\alpha_k} = \sqrt{\frac{1}{nV_h}}, \quad i, k \in \{1, \ldots, d_2\}$$

and let h^* denotes its solution. One can check that

$$h_i^* = n^{\frac{-\bar{\alpha}}{\alpha_i(2\bar{\alpha}+1)}},\ i = 1, ..., d_2. \tag{17}$$

Here, we have used $\frac{1}{\bar{\alpha}} = \frac{1}{\alpha_1} + \ldots + \frac{1}{\alpha_{d_2}}$. Applying Theorem 3.1, we obtain

$$\mathcal{R}_p^{(q)}\left(\widehat{f_n}, f\right) \leqslant 2(1+\kappa)\left\{\frac{C_*}{\sqrt{nV_{h^*}}} + L\kappa\sum_{i=1}^{d_2}[h_i^*]^{\alpha_i}\right\} + C_1^*\frac{\log^{d_2}(n)}{\sqrt{n}} + C_2^*\frac{\log^{d_2+\frac{1}{q}}(n)}{\sqrt{n}}.$$

Thus, we deduce the result

$$n^{\frac{\bar{\alpha}}{2\bar{\alpha}+1}}\mathcal{R}_p^{(q)}\left(\widehat{f_n}, f\right) \leqslant 2(1+\kappa)(C_* + L\kappa d_2) + o(1).$$

Assumption 5 . The noise ζ_1 admits a p_ζ w.r.t. Lebesgue measure on $\mathbb{R}$ satisfying

$$\int \frac{[p_\zeta(t-u)]^2}{p_\zeta(t)}dt \leqslant 1 + p^*u^2,\ \forall\ 0 \leqslant u \leqslant v_0; \quad p^* > 0.$$

Theorem 3.5 *Suppose that assumption 5 is fulfilled. Then for all* $p, q \geq 1$ *and* $\alpha \in (1, \mathbf{l})^{d_2}$ *we have*

$$\liminf_{n\to+\infty} \inf_{\widetilde{f_n}} \sup_{f\in H_{d_2}(\alpha,L)} \{\psi_n^{-1}(\alpha)\mathcal{R}_p^{(q)}[\widetilde{f_n}, f]\} > 0.$$

Remark 3.6 *Combining Theorem 3.4 and Theorem 3.5, we can assert that* $\psi_n(\alpha)$ *is a minimax rate of convergence on* $H_{d_2}(\alpha, L)$*,* $\alpha \in (0, \mathbf{l})^{d_2}$ *and* $\widehat{f_n}$ *is minimax adaptive estimator over the collection of anisotropic Hölder classes. Thus this estimator adjusts automatically to unknown smoothness.*

Proof. The proof of Theorem 3.5 coincides with the one one of Theorem 3 in (Goldenshluger & Lepski, 2014), up minor modifications to take into account the partially linear structure. Therefore, it is omitted.

Acknowledgements

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

References

Baraud, Y. (2002). Model selection for regression on a random design. *ESAIM: Probabability and Statistics, 6*, 127-146. https://doi.org/10.1051/ps:2002007

Bousquet, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *Comptes Rendus Mathematique, 334*(6), 495-500. https://doi.org/10.1016/S1631-073X(02)02292-6

Comte, F., & Lacour, C. (2013). Anisotropic adaptive kernel deconvolution. *Annales de l'Institut Henri Poincaré-Probabilités et Statistiques, 49*(2), 569-609. https://doi.org/10.1214/11-AIHP470

Engle, R. F., Granger, C. W. J., Rice, J., & Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association, 81*(394) 310-320. https://doi.org/10.2307/2289218

Folland, G. B. (1999). *Real Analysis: Modern Techniques and their Applications* (2nd ed.). John Wiley, New York.

Goldenshluger, A., & Lepski, O. V. (2008). Universal pointwise selection rule in multivariate function estimation. *Bernouilli, 14*(4), 1150-1190. https://doi.org/10.3150/08-BEJ144

Goldenshluger, A., & Lepski, O. (2009). Structural adaptation via $\mathbb{L}_p$-norm oracle inequalities. *Probability Theory and Related Fields, 143*, 41-71. https://doi.org/10.1007/s00440-007-0119-5

Goldenshluger, A., & Lepski, O. (2011a). Bandwidth selection in kernel density estimation: oracle inequalities and adapative minimax optimality. *The Annals of Statistics, 39*(3), 1608-1632. https://doi.org/10.1214/11-AOS883

Goldenshluger, A., & Lepski, O. (2011b). Uniform bounds for norms of sums of independent random functions. *The Annals of Probability, 39*(6) 2318-2384. https://doi.org/10.1214/10.AOP595

Goldenshluger, A. V., & Lepski, O. V. (2012). General selection rule from a family of linear estimators. *Theory Probability and Its Applications, 57*(2), 209-226. https://doi.org/10.1137/S0040585X9798523

Goldenshluger, A., & Lepski, O. (2014). On adaptive minimax density estimation on $\mathbb{R}^d$. *Probability Theory and Related Fields, 159*(3), 479-543. https://doi.org/10.1007/s00440-013-0512-1

Härdle, W., Liang, H., & Gao, J. (2000). *Partially linear models*. Springer Verlag, New York. https://doi.org/10.7007/978-3-642-57700-0

Lepski, O. V. (2013a). Multivariate density estimation under sup-norm loss: oracle approach, adaptation and independance structure. *The Annals of Statistics, 41*(2), 1005-1034. https://doi.org/10.1214/13-AOS1109

Lepski, O. V. (2013b). Upper functions for positive random functionals. I. General setting and gaussian random funcions. *Mathematical Methods of Statistics, 22*(1), 1-27. https://doi.org/10.3103/S1066530713010018

Lepski, O. V., & Serdyukova, N. (2014). Adaptive estimation under single-index constraint in regression model. *The Annals of Statistics, 42*(1), 1-28. https://doi.org/10.1214/13-AOS1152

Nguyen, N. (2014). *Adaptation via des inégalités d'oracle dans le modèle de regression avec design aléatoire* (Thèse de doctorat, Aix-marseille université). Theses.fr. http://www.theses.fr/2014AIXM4716/document

Rebelles, G. (2015). *Sur l'estimation adaptative d'une densité multivariée sous l'hypothèse de la structure d'indépendance* (Thèse de doctorat, Aix-marseille université). Theses.fr. http://www.theses.fr/2015AIXM4774/document

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica, 56*(4), 931-954. https://doi.org/10.2307/1912705