



**HAL**  
open science

## Regression with random design: a minimax study

Christophe Chesneau

► **To cite this version:**

| Christophe Chesneau. Regression with random design: a minimax study. 2005. hal-00014248v1

**HAL Id: hal-00014248**

**<https://hal.science/hal-00014248v1>**

Preprint submitted on 22 Nov 2005 (v1), last revised 2 Jan 2006 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Regression with random design: a minimax study

Chesneau Christophe

Laboratoire de Probabilités et Modèles Aléatoires,  
CNRS-UMR 7599, Université Paris VI, UFR de Mathématiques,  
175 rue de Chevaleret F-75013 Paris, France.  
E-mail: christophe.chesneau4@wanadoo.fr  
Stat-page: <http://www.chesneau-stat.com>

## Abstract

In this paper, minimax rates of convergence are studied for the regression model with (known) random design. Considering the framework of wavelet analysis, we establish the asymptotic minimax rate of convergence under the  $\mathbb{L}^p$  loss over Besov balls.

**Key Words:** Regression with random design, minimax rate of convergence, Besov spaces, wavelets.

**AMS subject classification:** 62G07, 62G20.

## 1 Motivation

Suppose we observe  $n$  pairs of random variable  $(X_1, Y_1), \dots, (X_n, Y_n)$  i.i.d governed by the equation

$$Y_i = f(X_i) + \xi_i, \quad i = 1, \dots, n \quad (1.1)$$

where the  $\xi_i$ 's are Gaussian i.i.d with mean zero, variance one and are independent of the design  $X_1, \dots, X_n$ . We denote by  $g$  the density of  $X_1$ . The function  $f$  is an unknown function of interest. We wish to reconstruct  $f$  from observations  $(X_1, Y_1), \dots, (X_n, Y_n)$ .

In the case where  $g$  is the density of an uniform law or satisfied some conditions of boundedness, this statistical problem was studied by many authors under various risk and over various function spaces. Quote for instance [Ibragimov and Khaminskii \(1982\)](#), [Stone \(1982\)](#) and [Delyon and Juditsky \(1996\)](#) among others. Numerous statistical results can be found in the book of [Tsybakov \(2004\)](#). In this paper, the accuracy of an estimate  $\hat{f}$  of  $f$  is measured under the global  $\mathbb{L}^p$  risk

$$R(\hat{f}, f) = \mathbb{E}_f^n \left( \int_0^1 |\hat{f}(t) - f(t)|^p dt \right), \quad p \geq 1$$

where we have denoted  $\mathbb{E}_f^n$  the expectation with respect to the distribution  $\mathbb{P}_f^n$  of  $(X_1, Y_1), \dots, (X_n, Y_n)$ . The benchmark for the performance of an estimator  $\hat{f}$  over a function class  $\mathcal{X}$  is the following minimax  $\mathbb{L}^p$  risk:

$$\mathcal{R}_n(\mathcal{X}) = \inf_{\hat{f}} \sup_{f \in \mathcal{X}} \mathbb{E}_f^n \left( \int_0^1 |\hat{f}(x) - f(x)|^p dx \right).$$

The aim of our study is to investigate the minimax rate of convergence over Besov balls  $B_{s,\pi,r}(L)$  under mild assumptions on  $g$ . First, we show that if  $g$  belongs to the following set

$$\mathcal{G} = \left\{ g; \frac{1}{g} \in \mathbb{L}^{\max(p,2)-1}([0,1]) \right\} \quad (1.2)$$

then for  $s > 0$ ,  $\pi \geq p$  and  $r \geq 1$  we have

$$\mathcal{R}_n(B_{\pi,r}^s(L)) \asymp n^{-\frac{sp}{1+2s}}.$$

Remark that if  $g$  belongs to (1.2), then it is not necessarily bounded from below. Second, we complete our minimax study by setting the minimax rate of convergence over  $B_{\pi,r}^s(L)$  for  $p > \pi$  in the simplest case where  $g$  is bounded from below. To obtain the upper bounds, we use a non adaptive procedure introduced by [Delyon and Juditsky \(1996\)](#) and some geometrical properties of the compactly supported wavelet bases under the  $\mathbb{L}^p$  norm. Let us precise that all the lower bounds are obtained via a consequence of Fano's lemma. Finally, we prove that these minimax results can be truly deteriorated for certain densities  $g$  which don't belong to the set  $\mathcal{G}$  described in (1.2).

This paper is organized as follows. Section 2 describes wavelet bases, some of their geometrical properties in  $\mathbb{L}^p$  norm and the main function spaces of the study. The minimax results over Besov balls and the associated proofs is presented in Section 3. Proofs of Propositions and technical Lemmas are given in Section 4.

## 2 Wavelet bases and function spaces

Throughout this paper we set  $\mathbb{L}^p([0, 1]) = \{f \text{ measurable on } [0, 1] \mid \|f\|_p^p = \int_0^1 |f(t)|^p dt < +\infty\}$ .

### 2.1 Wavelet bases and geometrical properties in $\mathbb{L}^p$ -norm

First, we introduce the wavelet bases on the unit interval. Second, we set some results which will be intensively used in the rest of this paper.

**Definition 2.1 (Wavelet bases on  $[0, 1]$ ).** *Let us consider  $\phi$  a father wavelet of a multiresolution analysis on  $\mathbb{R}$  and  $\psi$  the associated mother wavelet. Assume that  $\text{Supp}(\phi) = \text{Supp}(\psi) = [1 - N, N]$  and  $\int_{1-N}^N \phi(t) dt = 1$ ,  $\int_{1-N}^N t^l \psi(t) dt = 0$  for  $l = 0, \dots, N - 1$ . Let*

$$\phi_{j,k}(x) = 2^{\frac{j}{2}} \phi(2^j x - k), \quad \psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k).$$

Then there exists  $\tau$  satisfying  $2^\tau \geq 2N$  such that the collection

$$\xi = \{\phi_{\tau,k}(\cdot), k \in \Delta_\tau; \psi_{j,k}(\cdot); j \geq \tau, k \in \Delta_j\}, \quad \Delta_j = \{0, \dots, 2^j - 1\},$$

with an appropriate treatments at the boundaries, is an orthonormal basis of  $\mathbb{L}^2([0, 1])$ . See [Cohen et al. \(1993\)](#) for further details about the wavelet bases on the interval.

Let  $1 \leq p < \infty$ . Any function  $f$  of  $\mathbb{L}^p([0, 1])$  can be decomposed on  $\xi$  as

$$f(x) = \sum_{k \in \Delta_\tau} \alpha_{\tau,k} \phi_{\tau,k}(x) + \sum_{j \geq \tau} \sum_{k \in \Delta_j} \beta_{j,k} \psi_{j,k}(x),$$

where  $\alpha_{j,k} = \int_0^1 f(t) \phi_{j,k}(t) dt$  and  $\beta_{j,k} = \int_0^1 f(t) \psi_{j,k}(t) dt$ . Let us denote by  $P_\tau(f)(x)$  the first term of this decomposition. The following lemmas set some inequalities linked to the basis  $\xi$ .

**Lemma 2.1 (Concentration property).** *Let  $v > 0$ . There exists a constant  $C > 0$  such that*

$$\sum_{k \in \Delta_j} |\phi_{j,k}(x)|^v \leq C 2^{\frac{ju}{2}}, \quad x \in [0, 1].$$

**Lemma 2.2.** *Let  $p > 1$ . There exist two constants  $c > 0$  and  $C > 0$  such that for  $j \geq \tau$  we have*

$$c 2^{j(\frac{p}{2}-1)} \sum_{k \in \Delta_j} |\alpha_{j,k}|^p \leq \left\| \sum_{k \in \Delta_j} \alpha_{j,k} \phi_{j,k}(x) \right\|_p^p \leq C 2^{j(\frac{p}{2}-1)} \sum_{k \in \Delta_j} |\alpha_{j,k}|^p.$$

The Lemmas (2.1) and (2.2) are obviously true if we exchange  $\phi$  by  $\psi$ . Let us introduce two important geometrical properties of the compactly supported wavelet bases and weighted compactly supported wavelet bases.

**Lemma 2.3 (Unconditional nature-Temlyakov's property).** *Let  $p > 1$ . Let us set  $\psi_{\tau-1,k} = \phi_{\tau,k}$ . Then the basis  $\xi$  is unconditional for  $\mathbb{L}^p([0, 1])$  i.e for all sequence  $u = (u_{j,k})_{j,k}$  we have:*

$$\left\| \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} u_{j,k} \psi_{j,k} \right\|_p^p \asymp \left\| \left( \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} |u_{j,k} \psi_{j,k}|^2 \right)^{\frac{1}{2}} \right\|_p^p. \quad (2.1)$$

Let  $\sigma \in \mathbb{R}$ . Then the weighted compactly supported wavelet basis  $\xi_\sigma$  defined by

$$\xi_\sigma = \{2^{\sigma\tau} \phi_{\tau,k}(\cdot), k \in \Delta_\tau; 2^{\sigma j} \psi_{j,k}(\cdot); j \geq \tau, k \in \Delta_j\}$$

satisfies the Temlyakov's property i.e for all  $D_{j,k} \subseteq \{\tau - 1, \dots\} \times \Delta_j$  we have:

$$\left\| \left( \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} 1_{D_{j,k}} |2^{\sigma j} \psi_{j,k}|^2 \right)^{\frac{1}{2}} \right\|_p^p \asymp \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} 1_{D_{j,k}} \|2^{\sigma j} \psi_{j,k}\|_p^p. \quad (2.2)$$

The proof of the first point of Lemma 2.3 above can be viewed in [Meyer \(1990\)](#), we refer to [Johnstone et al. \(2004, Theorem 2\)](#) for the second point.

## 2.2 Function spaces

**Definition 2.2 (Besov balls).** For any measurable function  $f$  defined on  $[0, 1]$ , we denote the associated  $N$ -th order modulus of smoothness as

$$\rho^N(t, f, \pi) = \sup_{|h| \leq t} \left( \int_{J_{Nh}} \left| \sum_{k=0}^N \binom{N}{k} (-1)^k f(u + kh) \right|^\pi du \right)^{\frac{1}{\pi}}$$

where  $J_{Nh} = \{x \in [0, 1] : x + Nh \in [0, 1]\}$ . Let  $0 < s < N$ ,  $1 \leq r \leq \infty$  and  $1 \leq \pi \leq \infty$ . We say that a function  $f$  of  $\mathbb{L}^\pi([0, 1])$  belongs to the Besov balls  $B_{\pi,r}^s(L)$  if and only if

$$\left( \int_0^1 \left( \frac{\rho^N(t, f, \pi)}{t^s} \right)^r \frac{1}{t} dt \right)^{\frac{1}{r}} \leq L < \infty$$

with the usual modification if  $r = \infty$ .

Lemma 2.4 below shows the link which exist between the Besov balls and the basis  $\xi$  described in Definition 2.1.

**Lemma 2.4.** Let  $0 < s < N$ ,  $1 \leq p \leq \infty$  and  $1 \leq r \leq \infty$ . We have

$$f \in B_{\pi,r}^s(L) \iff \begin{cases} \left( \sum_{j \geq \tau-1} \left( 2^{j(s+\frac{1}{2})} \left( \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi 2^{-j} \right)^{\frac{1}{\pi}} \right)^r \right)^{\frac{1}{r}} \leq L, \\ \left( \sum_{j \geq \tau} (2^{js} \|P_j(f) - f\|_\pi)^r \right)^{\frac{1}{r}} \leq L. \end{cases} \quad (2.3)$$

## 3 Minimax study over Besov balls

Throughout this paper, we observe the model (1.1) where  $f, g$  are assumed to be compactly supported on  $[0, 1]$ . Moreover, we shall assume that  $\|f\|_\infty = \sup_{x \in [0,1]} |f(x)| < \infty$ .

Thereafter, the constants  $C$  and  $c$  represent any constants we shall need, and can be different from one line to one other. The symbol ' $\asymp$ ' is characterized by  $a \asymp b \Leftrightarrow cb \leq a \leq Cb$ . Considering the parameters  $(s, \pi, r)$  of the Besov balls  $B_{\pi,r}^s(L)$ , we adopt the following notations:

$$\alpha_1 = \frac{s}{1+2s}, \quad \alpha_2 = \frac{(s - \frac{1}{\pi}) + \frac{1}{p}}{2(s - \frac{1}{\pi}) + 1} \quad \text{and} \quad \epsilon = \pi s + \frac{\pi - p}{2}.$$

Theorem 3.1 below shows that the minimax rate of convergence over  $B_{\pi,r}^s(L)$  can be of the form  $n^{-\alpha_1 p}$  under some condition of integrability on  $g$ . Theorem 3.2 below exhibits the minimax rate of convergence for  $p > \pi$  in the case where  $g$  is bounded from below. Proposition 3.1 below completes our minimax study.

**Theorem 3.1.** Let  $1 \leq p < \infty$ . Assume that

$$\frac{1}{g} \in \mathbb{L}^{p'-1}([0, 1]) \quad (3.1)$$

where  $p' = \max(p, 2)$ . Then for  $s > 0$ ,  $1 \leq r \leq \infty$  and  $p \leq \pi \leq \infty$  we have

$$\mathcal{R}_n(B_{\pi,r}^s(L)) \asymp n^{-\alpha_1 p}.$$

**Theorem 3.2.** Let  $1 \leq p < \infty$ . Assume that  $g$  is bounded from below. Then for  $1 \leq \pi < p$ ,  $s > \frac{1}{\pi} + \frac{1}{2}$  and  $1 \leq r \leq \infty$  we have:

$$\begin{cases} \mathcal{R}_n(B_{\pi,r}^s(L)) \asymp n^{-\alpha_1 p} & s > \frac{1}{\pi} \left( \frac{p-\pi}{2} \right) \\ \mathcal{R}_n(B_{\pi,r}^s(L)) \asymp \left( \frac{\ln(n)}{n} \right)^{\alpha_2 p} & \frac{1}{\pi} \left( \frac{p-\pi}{2} \right) > s. \end{cases} \quad (3.2)$$

For the case where  $s = \frac{1}{\pi} \left( \frac{p-\pi}{2} \right)$ , there exist two constants  $C > 0$  and  $c > 0$  such that:

$$c \left( \frac{\ln(n)}{n} \right)^{\alpha_2 p} \leq \mathcal{R}_n(B_{\pi,r}^s(L)) \leq C \left( \frac{\ln(n)}{n} \right)^{\alpha_2 p} \ln(n)^{(p-\frac{\pi}{r})_+}.$$

**Proposition 3.1.** *There exist densities  $g$  such that the minimax rates of convergence obtained in Theorem 3.1 and Theorem 3.2 can not be attained.*

### 3.1 Proofs of Theorem 3.1 and Theorem 3.2

Let  $u \in \{0, 1\}$  and  $c_* \in ]0, +\infty[$ . Let us consider the following procedure:

$$\hat{f}_u(x) = \sum_{k \in \Delta_{j_0}} \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x) + u \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} \hat{\beta}_{j,k} 1_{|\hat{\beta}_{j,k}| \geq \kappa \sqrt{\frac{j-j_0}{n}}} \psi_{j,k}(x) \quad (3.3)$$

where the estimators  $\hat{\alpha}_{j,k}$  and  $\hat{\beta}_{j,k}$  are defined by:

$$\hat{\alpha}_{j,k} = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{g(X_i)} \phi_{j,k}(X_i), \quad \hat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{g(X_i)} \psi_{j,k}(X_i) \quad (3.4)$$

and the integers  $j_0$  and  $j_1$  are defined by:

$$2^{j_0} \asymp n^{\frac{1}{2s+1}}, \quad 2^{j_1} \asymp \left( \frac{n}{\ln(n)} \right)^{c_* 1_{\{\epsilon > 0\}} + \frac{1_{\{\epsilon \leq 0\}}}{2(s-\frac{1}{\pi})+1}}.$$

*Proof of Theorem 3.1. Proof of the upper bound:* Assume that the condition (3.1) holds. Using Minkowski's inequality, Lemma 2.4, Lemma 2.2 and the fact that  $f \in B_{\pi,r}^s(L) \subset B_{p,\infty}^s(L)$  (since  $\pi \geq p$ ), the  $\mathbb{L}^p$  risk of  $\hat{f}_0$  can be decomposed as follows:

$$\begin{aligned} \mathbb{E}_f^n(\|\hat{f}_0 - f\|_p^p) &\leq C \left( \mathbb{E}_f^n(\|\hat{f}_0 - P_{j_0}(f)\|_p^p) + \|P_{j_0}(f) - f\|_p^p \right) \\ &\leq C \left( 2^{j_0(\frac{p}{2}-1)} \sum_{k \in \Delta_{j_0}} \mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) + 2^{-j_0 s p} \right). \end{aligned} \quad (3.5)$$

To bound the first term, we need the following lemma:

**Lemma 3.1.** *Let  $2 \leq a < \infty$ ,  $j \geq \tau$  and  $n \in \mathbb{N}^*$ . Assume that  $\frac{1}{g} \in \mathbb{L}^{a-1}([0, 1])$ . Then there exists a constant  $C > 0$  such that:*

$$\mathbb{E}_f^n(|\hat{\alpha}_{j,k} - \alpha_{j,k}|^a) \leq C \left( n^{1-a} \int_0^1 \frac{|\phi_{j,k}(x)|^a}{g^{a-1}(x)} dx + n^{-\frac{a}{2}} \int_0^1 \frac{|\phi_{j,k}(x)|^2}{g^{\frac{a}{2}}(x)} dx \right).$$

First, study the case where  $2 > p \geq 1$ . Using Jensen's inequality and Lemma 3.1 for  $a = 2$ , we have

$$\mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) \leq \mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^2)^{\frac{p}{2}} \leq C \left( n^{-1} \int_0^1 \frac{|\phi_{j_0,k}(x)|^2}{g(x)} dx \right)^{\frac{p}{2}}.$$

Cauchy-Schwartz's inequality, the condition (3.1) and Lemma 2.1 yield:

$$\begin{aligned} \sum_{k \in \Delta_{j_0}} \mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) &\leq C n^{-\frac{p}{2}} \left( \sum_{k \in \Delta_{j_0}} \int_0^1 \frac{|\phi_{j_0,k}(x)|^2}{g(x)} dx \right)^{\frac{p}{2}} (\text{Card}(\Delta_{j_0}))^{1-\frac{p}{2}} \\ &\leq C n^{-\frac{p}{2}} \left( 2^{j_0} \int_0^1 \frac{1}{g(x)} dx \right)^{\frac{p}{2}} 2^{j_0(1-\frac{p}{2})} \leq C 2^{j_0} n^{-\frac{p}{2}}. \end{aligned} \quad (3.6)$$

Now, consider the case where  $p \geq 2$ . It follows from Lemma 3.1, the combination of Lemma 2.1 with the condition (3.1) and the choice of  $j_0$  that:

$$\begin{aligned} \sum_{k \in \Delta_{j_0}} \mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) &\leq C \left( n^{1-p} \int_0^1 \frac{\sum_{k \in \Delta_{j_0}} |\phi_{j_0,k}(x)|^p}{g^{p-1}(x)} dx + n^{-\frac{p}{2}} \int_0^1 \frac{\sum_{k \in \Delta_{j_0}} |\phi_{j_0,k}(x)|^2}{g^{\frac{p}{2}}(x)} dx \right) \\ &\leq C \left( 2^{\frac{j_0 p}{2}} n^{1-p} + 2^{j_0} n^{-\frac{p}{2}} \right) \leq C 2^{j_0} n^{-\frac{p}{2}}. \end{aligned} \quad (3.7)$$

Putting (3.5), (3.6), (3.7) together, we deduce the existence of a constant  $C > 0$  such that:

$$\sup_{f \in B_{\pi,r}^s(L)} \mathbb{E}_f^n(\|\hat{f}_0 - f\|_p^p) \leq C \left( 2^{\frac{j_0 p}{2}} n^{-\frac{p}{2}} + 2^{-j_0 s p} \right) \leq C n^{-\alpha_1 p}$$

for  $p \geq 1$ . This implies that  $\mathcal{R}_n(B_{\pi,r}^s(L)) \leq C n^{-\alpha_1 p}$ .

*Proof of the lower bound:* Before introducing a consequence of the Fano's lemma, recall that the Kullbak-Leibler divergence  $K(P, Q)$  between  $P$  and  $Q$  is defined by

$$K(P, Q) = \begin{cases} \int \ln\left(\frac{dP}{dQ}\right) dP & \text{if } P \ll Q, \\ \infty & \text{otherwise.} \end{cases}$$

**Lemma 3.2.** *Let  $A$  be a sigma algebra on the space  $\Omega$ . Let  $A_i \in A$ ,  $i \in \{0, 1, \dots, m\}$  such that  $\forall i \neq j$ ,  $A_i \cap A_j = \emptyset$ . Let  $\mathbb{P}_i$ ,  $i \in \{0, 1, \dots, m\}$  be  $m + 1$  probability measures on  $(\Omega, A)$ . Then*

$$\sup_{i=0, \dots, m} \mathbb{P}_i(A_i^c) \geq \min\left(\frac{1}{2}, \sqrt{m} \exp\left(-\frac{3}{e}\right) \exp(-\chi_m)\right)$$

where

$$\chi_m = \inf_{v \in \{0, 1, \dots, m\}} \frac{1}{m} \sum_{k \neq v} K(\mathbb{P}_k, \mathbb{P}_v). \quad (3.8)$$

For a proof, we refer the reader to DeVore et al. (2005). For further details and applications of Fano's lemma see Birge (2001) and Tsybakov (2004).

For all  $\epsilon = (\epsilon_k)_{k \in \Delta_j} \in \{0, 1\}^{2^j}$ , let us set

$$g_\epsilon(x) = \gamma_j \sum_{k \in \Delta_j} \epsilon_k \psi_{j,k}(x) \quad (3.9)$$

where  $j$  is an integer to be chosen below and  $\gamma_j \asymp 2^{-j(s+\frac{1}{2})}$ . Since the wavelet coefficients of  $g_\epsilon$  are equal to  $\gamma_j \epsilon_k$ , it is easy to see that  $g_\epsilon \in B_{\pi,r}^s(L)$ . The rest of the proof is based on the theorem of Varshamov-Gilbert (see for instance Tsybakov (2004, Lemma 2.7)). It said that there exists a subset  $E_j = \{\epsilon^{(0)}, \dots, \epsilon^{(T_j)}\}$  of  $\{0, 1\}^{2^j}$  such that  $\forall 0 \leq u < v \leq T_j$  there exist  $0 < c < 1$  and  $0 < \alpha < 1$  satisfying:

$$\sum_{k \in \Delta_j} |\epsilon_k^{(u)} - \epsilon_k^{(v)}| \geq c 2^j \quad \text{where} \quad T_j \geq e^{\alpha 2^j} \quad (3.10)$$

Considering such a  $E_j$  and using Lemma 2.2, for  $u \neq v$  and  $u, v \in \{0, \dots, T_j\}$  one gets:

$$\|g_{\epsilon^{(u)}} - g_{\epsilon^{(v)}}\|_p = c 2^{\frac{j}{2}} \gamma_j \left( 2^{-j} \sum_{k \in \Delta_j} |\epsilon_k^{(u)} - \epsilon_k^{(v)}| \right)^{\frac{1}{p}} \geq 2\delta_j$$

where  $\delta_j = c 2^{\frac{j}{2}} \gamma_j$ . Using Chebychev's inequality, for any  $\hat{f}$  we have

$$\delta_j^{-p} \sup_{f \in B_{\pi,r}^s(L)} \mathbb{E}_f^n(\|\hat{f} - f\|_p^p) \geq \sup_{u \in \{0, \dots, T_j\}} \mathbb{P}_{g_{\epsilon^{(u)}}}^n(A_u^c) = p_1$$

where the sets  $A_u$  defined by  $A_u = \{\|\hat{f} - g_{\epsilon^{(u)}}\|_p < \delta_j\}$  satisfy  $A_u \cap A_v = \emptyset$  for  $u \neq v$  and  $u, v \in \{0, \dots, T_j\}$ . Thus, Fano's lemma gives us

$$p_1 \geq \min\left(\frac{1}{2}, \sqrt{T_j} \exp\left(-\frac{3}{e}\right) \exp(-\chi_{T_j})\right). \quad (3.11)$$

Now, let us consider the following lemma which will be proved in Appendix.

**Lemma 3.3.** *For any functions  $f_1, f_2$  defined on  $[0, 1]$  and bounded from above, we have*

$$K(\mathbb{P}_{f_1}^n, \mathbb{P}_{f_2}^n) = \frac{n}{2} \int_0^1 (f_1(x) - f_2(x))^2 g(x) dx. \quad (3.12)$$

Thus, for all  $u \neq v$  and  $u, v \in \{0, \dots, T_j\}$ , Lemma 2.1 and the fact that  $\int_0^1 g(x)dx = 1$  imply that:

$$\begin{aligned} K(\mathbb{P}_{g_{\epsilon^{(u)}}}^n, \mathbb{P}_{g_{\epsilon^{(v)}}}^n) &= \frac{n}{2} \gamma_j^2 \int_0^1 \left( \sum_{k \in \Delta_j} (\epsilon_k^{(u)} - \epsilon_k^{(v)}) \psi_{j,k}(x) \right)^2 g(x) dx \leq 2n \gamma_j^2 \int_0^1 \left( \sum_{k \in \Delta_j} |\psi_{j,k}(x)| \right)^2 g(x) dx \\ &\leq Cn \gamma_j^2 2^j \int_0^1 g(x) dx = Cn \gamma_j^2 2^j. \end{aligned}$$

Hence

$$\chi_{T_j} \leq \inf_{v \in \{0, \dots, T_j\}} \sup_{u \neq v; u \in \{0, \dots, T_j\}} K(\mathbb{P}_{g_{\epsilon^{(u)}}}^n, \mathbb{P}_{g_{\epsilon^{(v)}}}^n) \leq Cn \gamma_j^2 2^j. \quad (3.13)$$

Putting (3.10), (3.18) and (3.13) together and choosing  $\gamma_j = c_0 \frac{1}{\sqrt{n}}$  where  $c_0$  is a well chosen constant, one gets:

$$\delta_j^{-p} \sup_{f \in B_{\pi, r}^s(L)} \mathbb{E}_f^n(\|\hat{f} - f\|_p^p) \geq c \exp\left(\frac{\alpha}{2} 2^j - Cc_0^2 2^j\right) \geq c$$

with  $\delta_j \asymp \frac{2^{\frac{j}{2}}}{\sqrt{n}} \asymp n^{\frac{1}{2}(\frac{1}{1+2s}-1)} \asymp n^{-\alpha_1}$ . This justifies the existence of a constant  $c > 0$  such that  $\mathcal{R}(B_{\pi, r}^s(L)) \geq cn^{-\alpha_1 p}$ . The proof of Theorem 3.1 is thus complete.  $\square$

*Proof of Theorem 3.2. Proof of the upper bound:* Proposition 3.2 below provides upper bounds over  $B_{\pi, r}^s(L)$  under  $\mathbb{L}^p$ -risk in the case where  $p > \pi$ . In particular, it proves that the minimax results obtained by Delyon and Juditsky (1996) for the procedure  $\hat{f}_1$  under the Besov norm can be extend to the  $\mathbb{L}^p$  norm under very mild assumptions on the model.

**Proposition 3.2 (General minimax results for  $p > \pi$ ).** *Let  $1 \leq p < \infty$ . Assume that we have a sequence of model  $E_n$  in which we are able to produce estimates  $\hat{\alpha}_{j,k}$  and  $\hat{\beta}_{j,k}$  of the wavelet coefficients  $\alpha_{j,k}$  and  $\beta_{j,k}$  of the unknown function  $f$ . Let us consider the procedure describes by (3.3) with  $u = 1$ . Suppose that there exists a constants  $C > 0$  such that*

$$\mathbb{E}_f^n(|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}) \leq Cn^{-p}, \quad j_0 < j < j_1 \quad (3.14)$$

for a suitably chosen  $c_*$  (with the notation  $\hat{\beta}_{j_0,k} = \hat{\alpha}_{j_0,k}$ ) and that there exist  $\kappa > 0$  and a positive function  $h$  satisfying  $\lim_{u \rightarrow \infty} h(u) = +\infty$  such that for all  $\lambda > 0$  we have:

$$\mathbb{P}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa}{2} \sqrt{\frac{\lambda}{n}} \right) \leq C2^{-h(\kappa)\lambda}. \quad (3.15)$$

Then for  $p > \pi$  and  $s > \frac{1}{\pi} - \min\left(\frac{1}{p}, \frac{1}{2}\left(1 - \frac{1}{c_*}\right)\right)$ , we have:

$$\begin{cases} \sup_{f \in B_{\pi, r}^s(L)} \mathbb{E}_f^n(\|\hat{f}_1 - f\|_p^p) \leq Cn^{-\alpha_1 p} & s > \frac{1}{\pi} \left(\frac{p-\pi}{2}\right), \\ \sup_{f \in B_{\pi, r}^s(L)} \mathbb{E}_f^n(\|\hat{f}_1 - f\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} & \frac{1}{\pi} \left(\frac{p-\pi}{2}\right) > s, \\ \sup_{f \in B_{\pi, r}^s(L)} \mathbb{E}_f^n(\|\hat{f}_1 - f\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} (\ln(n))^{(p-\frac{\pi}{r})_+} & s = \frac{1}{\pi} \left(\frac{p-\pi}{2}\right). \end{cases} \quad (3.16)$$

For the case where  $s > \frac{1}{\pi} \left(\frac{p-\pi}{2}\right)$ , let us remark that the rate of convergence of  $\hat{f}_1$  is without logarithmic factor contrary to that reached by the procedure described in Donoho et al. (1996, Section 4). Lemma 3.4 below shows that the conditions (3.14) and (3.15) of Proposition 3.2 hold for  $c_* = \frac{1}{2}$ :

**Lemma 3.4.** *Assume that we observe the regression model (1.1) where  $g$  is bounded from below. Then for  $j < j_1$  with  $2^{j_1} \leq C \sqrt{\frac{n}{\ln(n)}}$  and  $k \in \Delta_j$ , the estimators  $\hat{\beta}_{j,k}$  and  $\hat{\alpha}_{j,k}$  described by (3.4) satisfy the conditions (3.14) and (3.15).*

The proof of the upper bounds of Theorem 3.2 is thus complete. *Proof of the lower bound:* Let us introduce the following family:

$$\{g_k = \gamma_j \psi_{j,k}, k \in \Delta_j, g_{2^j} = 0\},$$

where  $j$  is an integer to be chosen below and  $\gamma_j \asymp 2^{-j(s+\frac{1}{2}-\frac{1}{p})}$ . Since the wavelet coefficients of  $g_k$  are equal to  $\gamma_j$ , it is easy to see that  $g_k \in B_{\pi,r}^s(L)$ . Moreover for  $k \neq k'$  and  $k \neq 0$ , Lemma 2.2 gives us:

$$\|g_k - g_{k'}\|_p \geq c\gamma_j 2^{j(\frac{1}{2}-\frac{1}{p})} = 2\delta_j \quad (3.17)$$

where  $\delta_j = c\gamma_j 2^{j(\frac{1}{2}-\frac{1}{p})}$ . For any  $\hat{f}$ , Chebychev's inequality gives us

$$\delta_j^{-p} \sup_{f \in B_{\pi,r}^s(L)} \mathbb{E}_f^n(\|\hat{f} - f\|_p^p) \geq \sup_{k \in \{0,1,\dots,2^j\}} \mathbb{P}_{g_k}^n(A_k^c) = p_2$$

where the sets  $A_k$  defined by  $A_k = \{\|\hat{f} - g_k\|_p < \delta_j\}$  satisfy  $A_k \cap A_{k'} = \emptyset$  for  $k \neq k'$ . It follows from Lemma 3.2 that:

$$p_2 \geq \min\left(\frac{1}{2}, \sqrt{2^j} \exp\left(-\frac{3}{e}\right) \exp(-\chi_{2^j})\right). \quad (3.18)$$

Using the equality (3.12) and Lemma 2.1 one gets:

$$\chi_{2^j} \leq 2^{-j} \sum_{k \in \Delta_j} K(\mathbb{P}_{g_k}^n, \mathbb{P}_{g_{2^j}}^n) \leq C 2^{-j} n \gamma_j^2 \int_0^1 g(x) \sum_{k \in \Delta_j} \psi_{j,k}^2(x) dx \leq C n \gamma_j^2 \int_0^1 g(x) dx = C \gamma_j^2 n.$$

Now, for a constant  $c_0 > 0$ , let us set  $\gamma_j = c_0 \sqrt{\frac{\ln(n)}{n}}$ , i.e.  $2^j \asymp \left(\sqrt{\frac{n}{\ln(n)}}\right)^{s+\frac{1}{2}-\frac{1}{p}}$ . Thus, for  $n$  large enough, we see that

$$\ln(2^j) \geq \frac{1}{2\left(s+\frac{1}{2}-\frac{1}{p}\right)} (\ln(n) - \ln(\ln(n))) \geq c\lambda_n$$

where  $\lambda_n = \ln(n)$ . Thus there exists  $c > 0$  such that  $\sqrt{2^j} \geq \exp(c\lambda_n)$ . For a good choice of  $c_0$ , it follows from Lemma 3.2 and the inequality (3.18) that

$$\delta_j^{-p} \sup_{f \in B_{\pi,r}^s(L)} \mathbb{E}_f^n(\|\hat{f} - f\|_p^p) \geq c \exp(c\lambda_n - cn\gamma_j^2) \geq c$$

where  $\delta_j \asymp \sqrt{\frac{\ln(n)}{n}} \left(\sqrt{\frac{n}{\ln(n)}}\right)^{\frac{1}{2}-\frac{1}{p}} = \left(\frac{\ln(n)}{n}\right)^{\alpha_2}$ . We conclude that there exists  $c > 0$  such that:

$$\mathcal{R}_n(B_{\pi,r}^s(L)) \geq c \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p}.$$

This ends the proof of Theorem 3.2.  $\square$

**Remark 3.1.** All the lower bounds have been obtained without extra assumption on the density  $g$ .

## 4 Appendix: proofs of Propositions and technical Lemmas

*Proof of Lemma 3.1.* Since  $\mathbb{E}_f^n\left(\frac{\phi_{j,k}(X_1)}{g(X_1)}\right) = \int_0^1 \phi_{j,k}(x) dx < \infty$ , we have

$$\mathbb{E}_f^n(\hat{\alpha}_{j,k}) = \mathbb{E}_f^n\left(\frac{f(X_1)\phi_{j,k}(X_1)}{g(X_1)}\right) = \int_0^1 \frac{f(t)\phi_{j,k}(t)g(t)}{g(t)} dt = \alpha_{j,k}.$$

Rosenthal's inequality applied with the i.i.d real variables  $W_i = \frac{Y_i \phi_{j,k}(X_i)}{g(X_i)} - \alpha_{j,k}$  for the exponent  $2 \leq a < \infty$  justifies the existence of a constant  $C > 0$  satisfying

$$\mathbb{E}_f^n(|\hat{\alpha}_{j,k} - \alpha_{j,k}|^a) \leq C \left(\frac{Z_a}{n^{a-1}} + \frac{(Z_2)^{\frac{a}{2}}}{n^{\frac{a}{2}}}\right). \quad (4.1)$$



where  $Z_a = \mathbb{E}_f^n(|W_1|^a)$ . Using an elementary inequality of convexity and the Hölder inequality, one gets:

$$\begin{aligned} Z_a &\leq 2^{a-1} \left( \mathbb{E}_f^n \left( \left| \frac{Y_i \phi_{j,k}(X_i)}{g(X_i)} \right|^a \right) + \mathbb{E}_f^n \left( \left| \frac{Y_i \phi_{j,k}(X_i)}{g(X_i)} \right|^a \right) \right) \\ &\leq C \mathbb{E}_f^n \left( \left| \frac{Y_i \phi_{j,k}(X_i)}{g(X_i)} \right|^a \right) \leq C \left( \mathbb{E}_f^n \left( \left| \frac{f(X_1) \phi_{j,k}(X_1)}{g(X_1)} \right|^a \right) + \mathbb{E}_f^n \left( \left| \frac{\phi_{j,k}(X_1)}{g(X_1)} \xi_1 \right|^a \right) \right) \\ &= C(T_a^{(1)} + T_a^{(2)}). \end{aligned} \quad (4.2)$$

The condition (3.1) gives us:

$$Z_a \leq C \max(T_a^{(1)}, T_a^{(2)}) \leq C \max(\|f\|_\infty^a, \mathbb{E}_f^n(|\xi_1|^a)) \int_0^1 \frac{|\phi_{j,k}(x)|^a}{g^{a-1}(x)} dx.$$

Using Hölder's inequality for the measure  $d\nu = \phi_{j,k}^2(x)dx$ , one gets

$$\begin{aligned} (Z_2)^{\frac{a}{2}} &\leq C \max((T_2^{(1)})^{\frac{a}{2}}, (T_2^{(2)})^{\frac{a}{2}}) \leq C \max(\|f\|_\infty^a, \mathbb{E}_f^n(|\xi_1|^a)) \left( \int_0^1 \frac{|\phi_{j,k}(x)|^2}{g(x)} dx \right)^{\frac{a}{2}} \\ &\leq C \int_0^1 \frac{|\phi_{j,k}(x)|^2}{g^{\frac{a}{2}}(x)} dx. \end{aligned} \quad (4.3)$$

Putting the inequalities (4.1), (4.2) and (4.3) together, one gets:

$$\mathbb{E}_f^n(|\hat{\alpha}_{j,k} - \alpha_{j,k}|^a) \leq C \left( n^{1-a} \int_0^1 \frac{|\phi_{j,k}(x)|^a}{g^{a-1}(x)} dx + n^{-\frac{a}{2}} \int_0^1 \frac{|\phi_{j,k}(x)|^2}{g^{\frac{a}{2}}(x)} dx \right).$$

This ends the proof of Lemma 3.1.  $\square$

*Proof of Lemma 3.12.* The law of  $(X_1, Y_1)$  is given by  $p_f^1(x, y) = \frac{1}{\sqrt{2\pi}} g(x) \exp(-\frac{1}{2}(y - f(x))^2)$ . Thus, for any functions  $f_1$  and  $f_2$  we have

$$\begin{aligned} K(\mathbb{P}_{f_1}^1, \mathbb{P}_{f_2}^1) &= \frac{1}{2\sqrt{2\pi}} \int_{\mathbb{R}} \int_0^1 ((y - f_2(x))^2 - (y - f_1(x))^2) g(x) e^{-\frac{1}{2}(y - f_1(x))^2} dx dy \\ &= \frac{1}{2\sqrt{2\pi}} \int_{\mathbb{R}} \int_0^1 (f_1(x) - f_2(x))(2z + f_1(x) - f_2(x)) e^{-\frac{1}{2}z^2} g(x) dx dz \\ &= \frac{1}{2} \int_0^1 (f_1(x) - f_2(x))^2 g(x) dx. \end{aligned}$$

Since  $K(\mathbb{P}_{f_1}^n, \mathbb{P}_{f_2}^n) = nK(\mathbb{P}_{f_1}^1, \mathbb{P}_{f_2}^1)$ , the proof is finished.  $\square$

*Proof of Proposition 3.2.* Let us denote  $\lambda_j = \kappa \sqrt{\frac{j-j_0}{n}}$ . Proceeding as in Donoho et al. (1996), we have the following decomposition:

$$\begin{aligned} \hat{f}_1(x) - f(x) &= \sum_{k \in \Delta_{j_0}} (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}) \phi_{j_0,k}(x) + \sum_{j > j_1} \sum_{k \in \Delta_j} \beta_{j,k} \psi_{j,k}(x) \\ &\quad + \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} (\hat{\beta}_{j,k} - \beta_{j,k}) \left( \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \lambda_j\}} \mathbf{1}_{\{|\hat{\beta}_{j,k}| < \frac{\lambda_j}{2}\}} + \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \lambda_j\}} \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \frac{\lambda_j}{2}\}} \right) \psi_{j,k}(x) \\ &\quad + \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} \beta_{j,k} \left( \mathbf{1}_{\{|\hat{\beta}_{j,k}| < \lambda_j\}} \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq 2\lambda_j\}} + \mathbf{1}_{\{|\hat{\beta}_{j,k}| < \lambda_j\}} \mathbf{1}_{\{|\hat{\beta}_{j,k}| < 2\lambda_j\}} \right) \psi_{j,k}(x) \\ &= e_1 + e_2 + e_{3,1} + e_{3,2} + e_{4,1} + e_{4,2}. \end{aligned}$$

By the Minkowski inequality and an elementary inequality of convexity, one gets:

$$\mathbb{E}_f^n(\|\hat{f}_1 - f\|_p^p) \leq 6^{p-1} (\mathbb{E}_f^n(\|e_1\|_p^p) + \mathbb{E}_f^n(\|e_2\|_p^p) + \mathbb{E}_f^n(\|e_{3,1}\|_p^p) + \mathbb{E}_f^n(\|e_{3,2}\|_p^p) + \mathbb{E}_f^n(\|e_{4,1}\|_p^p) + \mathbb{E}_f^n(\|e_{4,2}\|_p^p)).$$

Proceeding as in the proof of Theorem 3.1, it is clear that:

$$\mathbb{E}_f^n(\|e_1\|_p^p) \leq C \left( 2^{j_0(\frac{p}{2}-1)} \sum_{k \in \Delta_{j_0}} \mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) \right) \leq C 2^{\frac{j_0 p}{2}} n^{-\frac{p}{2}}.$$

It follows from the Minkowski inequality, Lemma 2.2 and the Sobolev embeddings  $B_{\pi,r}^s(L) \subseteq B_{p,r}^{s-\frac{1}{\pi}+\frac{1}{p}}(L) \subseteq B_{p,\infty}^{s-\frac{1}{\pi}+\frac{1}{p}}(L)$  ( $p > \pi$ ) that:

$$\mathbb{E}_f^n(\|e_2\|_p^p) \leq C \left( \sum_{j \geq j_1} 2^{j(\frac{1}{2}-\frac{1}{p})} \left( \sum_{k \in \Delta_j} |\beta_{j,k}|^p \right)^{\frac{1}{p}} \right)^p \leq C \left( \sum_{j \geq j_1} 2^{-j(s-\frac{1}{\pi}+\frac{1}{p})} \right)^p \leq C 2^{-j_1(s-\frac{1}{\pi}+\frac{1}{p})p}.$$

Since the following inclusions hold:

$$\begin{aligned} \left\{ |\hat{\beta}_{j,k}| < \lambda_j, |\beta_{j,k}| \geq 2\lambda_j \right\} \cup \left\{ |\hat{\beta}_{j,k}| \geq \lambda_j, |\beta_{j,k}| < \frac{\lambda_j}{2} \right\} &\subset \left\{ |\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2} \right\}, \\ \left\{ |\hat{\beta}_{j,k}| < \lambda_j, |\beta_{j,k}| \geq 2\lambda_j \right\} &\subset \left\{ |\beta_{j,k}| \leq 2|\hat{\beta}_{j,k} - \beta_{j,k}| \right\}, \end{aligned}$$

we can group the terms  $e_{3,1}$  and  $e_{4,1}$ . Now, remark that the Cauchy-Schwartz inequality combined with the conditions (3.14) and (3.15) gives us:

$$\begin{aligned} \mathbb{E}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}|^p \mathbf{1}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\}} \right) &\leq \mathbb{E}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}|^{2p} \right)^{\frac{1}{2}} \mathbb{P}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2} \right)^{\frac{1}{2}} \\ &\leq C n^{-\frac{p}{2}} 2^{-\frac{(j-j_0)h(\kappa)}{2}}. \end{aligned} \quad (4.4)$$

The rest of the proof uses similar arguments than Kerkycharian and Picard (2000, Subsection 5.1.1). For  $p \geq 2$ , using the unconditional nature of the basis  $\xi$ , the generalized Minkowski inequality, the inequality (4.4) and the fact that  $\xi_\sigma$  satisfies the Temlyakov's property, we see that:

$$\begin{aligned} \mathbb{E}_f^n(\|e_{3,1}\|_p^p + \|e_{4,1}\|_p^p) &\leq C \mathbb{E}_f^n \left( \int_0^1 \left( \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^2 \mathbf{1}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\}} |\psi_{j,k}(x)|^2 dx \right)^{\frac{p}{2}} \right) \\ &\leq C \int_0^1 \left( \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}|^p \mathbf{1}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\}} \right)^{\frac{2}{p}} |\psi_{j,k}(x)|^2 dx \right)^{\frac{p}{2}} \\ &\leq C n^{-\frac{p}{2}} \int_0^1 \left( \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} 2^{-\frac{(j-j_0)h(\kappa)}{p}} |\psi_{j,k}(x)|^2 dx \right)^{\frac{p}{2}} \\ &\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} 2^{-\frac{(j-j_0)h(\kappa)}{2}} \|\psi_{j,k}\|_p^p \\ &\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} 2^{\frac{jp}{2}} 2^{-\frac{(j-j_0)h(\kappa)}{2}} \leq C n^{-\frac{p}{2}} 2^{\frac{j_0 p}{2}} \end{aligned}$$

for  $\kappa$  large enough. For  $1 < p < 2$ , using the unconditional nature of the basis  $\xi$ , the comparison between the  $l_p$  norms and the inequality (4.4), one gets:

$$\begin{aligned} \mathbb{E}_f^n(\|e_{3,1}\|_p^p + \|e_{4,1}\|_p^p) &\leq C \mathbb{E}_f^n \left( \int_0^1 \left( \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^2 \mathbf{1}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\}} |\psi_{j,k}(x)|^2 dx \right)^{\frac{p}{2}} \right) \\ &\leq C \int_0^1 \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}|^p \mathbf{1}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\}} \right) |\psi_{j,k}(x)|^p dx \\ &\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} 2^{-\frac{(j-j_0)h(\kappa)}{2}} \|\psi_{j,k}\|_p^p \leq C n^{-\frac{p}{2}} 2^{\frac{j_0 p}{2}} \end{aligned}$$

for  $\kappa$  large enough. For the term  $e_{3,2}$ , in the case where  $p \geq 2$ , we use again the unconditional nature of the basis  $\xi$ , the generalized Minkowski inequality, the condition (3.14), the fact that  $\xi$  satisfies the

Temlyakov's property and the characterization of Besov spaces:

$$\begin{aligned}
\mathbb{E}_f^n(\|e_{3,2}\|_p^p) &\leq C \int_0^1 \left( \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}|^p \right)^{\frac{2}{p}} 1_{\{|\beta_{j,k}| > \frac{\lambda_j}{2}\}} |\psi_{j,k}(x)|^2 dx \right)^{\frac{p}{2}} \\
&\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} 1_{\{|\beta_{j,k}| > \frac{\lambda_j}{2}\}} \|\psi_{j,k}\|_p^p \\
&\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} 2^{j(\frac{p}{2}-1)} \lambda_j^{-\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi \\
&\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} 2^{j(\frac{p}{2}-1)} \frac{n^{\frac{\pi}{2}}}{(j-j_0)^{\frac{\pi}{2}}} 2^{-j(s+\frac{1}{2}-\frac{1}{\pi})\pi} \\
&\leq C n^{\frac{\pi-p}{2}} \sum_{j_0 < j < j_1} \frac{2^{-j\epsilon}}{(j-j_0)^{\frac{\pi}{2}}}.
\end{aligned}$$

In the case where  $1 < p < 2$ , we use again the unconditional nature of the basis  $\xi$ , the comparison between the  $l_p$  norms, the condition (3.14) and the characterization of the Besov spaces:

$$\begin{aligned}
\mathbb{E}_f^n(\|e_{3,2}\|_p^p) &\leq C \int_0^1 \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}|^p \right) 1_{\{|\beta_{j,k}| > \frac{\lambda_j}{2}\}} |\psi_{j,k}(x)|^p dx \\
&\leq C n^{-\frac{p}{2}} \sum_{j_0 < j < j_1} \sum_{k \in \Delta_j} 1_{\{|\beta_{j,k}| > \frac{\lambda_j}{2}\}} \|\psi_{j,k}\|_p^p \\
&\leq C n^{\frac{\pi-p}{2}} \sum_{j_0 < j < j_1} \frac{2^{-j\epsilon}}{(j-j_0)^{\frac{\pi}{2}}}.
\end{aligned}$$

Now, let's focus on the term  $e_{4,2}$ . Using the Minkowski inequality and Lemma 2.2, we see that:

$$\begin{aligned}
\mathbb{E}_f^n(\|e_{4,2}\|_p^p) &\leq C \left( \sum_{j_0 < j < j_1} 2^{j(\frac{1}{2}-\frac{1}{p})} \left( \sum_{k \in \Delta_j} |\beta_{j,k}|^p 1_{\{|\beta_{j,k}| < 2\lambda_j\}} \right)^{\frac{1}{p}} \right)^p \\
&\leq C \left( \sum_{j_0 < j < j_1} 2^{j(\frac{1}{2}-\frac{1}{p})} \left( (2\lambda_j)^{p-\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi \right)^{\frac{1}{p}} \right)^p \\
&\leq C n^{\frac{\pi-p}{2}} \left( \sum_{j_0 < j < j_1} 2^{j(\frac{1}{2}-\frac{1}{p})} \left( (j-j_0)^{\frac{p-\pi}{2}} 2^{-j(s+\frac{1}{2}-\frac{1}{\pi})\pi} \right)^{\frac{1}{p}} \right)^p \\
&\leq C n^{\frac{\pi-p}{2}} \left( \sum_{j_0 < j < j_1} (2^{-j\epsilon} (j-j_0)^{\frac{p-\pi}{2}})^{\frac{1}{p}} \right)^p.
\end{aligned}$$

Since  $e_{3,1}$ ,  $e_{4,1}$  and  $e_1$  are of the same order, it suffices to balance the bounds of  $e_1$ ,  $e_2$ ,  $e_{3,2}$  and  $e_{4,2}$  to obtain the optimal upper bounds.

In the case where  $s > \max\left(\frac{1}{\pi} \left(\frac{p-\pi}{2}\right), \frac{1}{\pi} - \min\left(\frac{1}{p}, \frac{1}{2}\left(1 - \frac{1}{c_*}\right)\right)\right)$  (i.e  $\epsilon > 0$ ), we have

$$\mathbb{E}_f^n(\|e_1\|_p^p) \leq C 2^{\frac{j_0 p}{2}} n^{-\frac{p}{2}} \leq C n^{-\alpha_1 p}, \quad \mathbb{E}_f^n(\|e_2\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{c_*(s-\frac{1}{\pi}+\frac{1}{p})p} \leq C n^{-\alpha_1 p},$$

$$\max(\mathbb{E}_f^n(\|e_{3,2}\|_p^p), \mathbb{E}_f^n(\|e_{4,2}\|_p^p)) \leq C n^{\frac{\pi-p}{2}} 2^{-\epsilon j_0} \left( \sum_{j>0} j^{\frac{p-\pi}{2p}} 2^{-j\frac{\epsilon}{p}} \right)^p \leq C n^{\frac{\pi-p}{2}} 2^{-\epsilon j_0} \leq C n^{-\alpha_1 p}$$

and we obtain the desired upper bound. In the case where  $\frac{1}{\pi} \left(\frac{p-\pi}{2}\right) > s > \frac{1}{\pi} - \min\left(\frac{1}{p}, \frac{1}{2}\left(1 - \frac{1}{c_*}\right)\right)$  (i.e  $\epsilon < 0$ ), we have

$$\mathbb{E}_f^n(\|e_1\|_p^p) \leq C 2^{\frac{j_0 p}{2}} n^{-\frac{p}{2}} \leq C n^{-\alpha_1 p}, \quad \mathbb{E}_f^n(\|e_2\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p},$$

$$\max(\mathbb{E}_f^n(\|e_{3,2}\|_p^p), \mathbb{E}_f^n(\|e_{4,2}\|_p^p)) \leq C j_1^{\frac{p-\pi}{2}} 2^{-\epsilon j_1} n^{\frac{\pi-p}{2}} \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_{2p}}.$$

Finally, in the special case where  $s = \frac{1}{\pi} \left(\frac{p-\pi}{2}\right)$  (i.e  $\epsilon = 0$ ), it suffices to remark that :

$$\begin{aligned} \mathbb{E}_f^n(\|e_{4,2}\|_p^p) &\leq C \left( \sum_{j_0 < j < j_1} 2^{j(\frac{1}{2} - \frac{1}{p})} \left( (2\lambda_j)^{p-\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi \right)^{\frac{1}{p}} \right)^p \\ &\leq C \left(\frac{j_1}{n}\right)^{\frac{p-\pi}{2}} \left( \sum_{j_0 < j < j_1} \left( 2^{j\pi(s + \frac{1}{2} - \frac{1}{p})} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi \right)^{\frac{1}{p}} \right)^p \\ &\leq C \left(\frac{j_1}{n}\right)^{\frac{p-\pi}{2}} j_1^{(p-\frac{\pi}{r})_+}. \end{aligned}$$

Since this last term bounds  $\mathbb{E}_f^n(\|e_{3,2}\|_p^p)$ , we obtain the desired upper bound. The proof of Proposition 3.2 is thus complete.  $\square$

*Proof of Lemma 3.4. The moment inequality:* Since  $g$  is supposed to be bounded from below, we have

$$\int_{S_{j,k}} \frac{|\psi_{j,k}(x)|^{2p}}{g^{2p-1}(x)} dx \leq \left\| \frac{1}{g} \right\|_\infty^{2p-1} \|\psi\|_\infty^{2p-2} 2^{j(p-1)}.$$

Considering Lemma 3.1 with  $\psi$  instead of  $\phi$  and using the fact that  $2^j \leq n$ , one gets:

$$\mathbb{E}_f^n(|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}) \leq C(n^{1-2p} 2^{j(p-1)} + n^{-p}) \leq Cn^{-p}.$$

*The concentration inequality:* Here, we proceed as in [Kerkycharian and Picard \(2005, subsection 9.1\)](#). Let  $\kappa > 0$ . By a simple decomposition, one gets:

$$\mathbb{P}_f^n \left( |\hat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa}{2} \sqrt{\frac{\lambda}{n}} \right) \leq S_1 + S_2$$

where

$$S_1 = \mathbb{P}_f^n \left( \left| \frac{1}{n} \sum_{i=1}^n \frac{\psi_{j,k}(X_i) f(X_i)}{g(X_i)} - \mathbb{E}_f^n \left( \frac{\psi_{j,k}(X_i) f(X_i)}{g(X_i)} \right) \right| \geq \frac{\kappa}{4} \sqrt{\frac{\lambda}{n}} \right)$$

and

$$S_2 = \mathbb{P}_f^n \left( \left| \frac{1}{n} \sum_{i=1}^n \frac{\psi_{j,k}(X_i)}{g(X_i)} \xi_i \right| \geq \frac{\kappa}{4} \sqrt{\frac{\lambda}{n}} \right).$$

First, let us focus on the term  $S_1$ . Since

$$\left\| \frac{\psi_{j,k}(X_1) f(X_1)}{g(X_1)} - \mathbb{E}_f^n \left( \frac{\psi_{j,k}(X_1) f(X_1)}{g(X_1)} \right) \right\|_\infty \leq 2 \left\| \frac{1}{g} \right\|_\infty \|f\|_\infty \|\psi\|_\infty 2^{\frac{j}{2}} \leq C 2^{\frac{j}{2}}$$

and

$$\mathbb{E}_f^n \left( \left| \frac{\psi_{j,k}(X_1) f(X_1)}{g(X_1)} - \mathbb{E}_f^n \left( \frac{\psi_{j,k}(X_1) f(X_1)}{g(X_1)} \right) \right|^2 \right) \leq C \|f\|_\infty^2,$$

it follows from Bernstein's inequality that for  $2^j \leq \sqrt{n}$  there exists  $C > 0$  satisfying

$$S_1 \leq 2 \exp \left( - \frac{n \left( \frac{\kappa}{4} \sqrt{\frac{\lambda}{n}} \right)^2}{C \left( 1 + 2^{\frac{j}{2}} \frac{\kappa}{4} \sqrt{\frac{\lambda}{n}} \right)} \right) \leq 2 \exp \left( - \frac{\kappa^2}{C} \lambda \right)$$

for  $n$  large enough. Thus there exist  $C > 0$  and a positive function  $h$  satisfying  $\lim_{u \rightarrow \infty} h(u) = \infty$  such that:

$$S_1 \leq C 2^{-h(\kappa)\lambda}. \quad (4.5)$$

Now, let us bound the term  $S_2$ . Conditionally on  $(X_1, \dots, X_n) = (x_1, \dots, x_n)$  we have

$$\frac{1}{n} \sum_{i=1}^n \frac{\psi_{j,k}(x_i)}{g(x_i)} \xi_i \sim \mathcal{N}(0, \sigma_{j,k}^2) \quad \text{with} \quad \sigma_{j,k}^2 \leq \left\| \frac{1}{g} \right\|_{\infty} \frac{1}{n^2} \sum_{i=1}^n \frac{\psi_{j,k}^2(x_i)}{g(x_i)}.$$

For  $\underline{x} = (x_1, \dots, x_n)$ , let us define

$$\mathcal{F}_{\alpha}(\underline{x}) = \left\{ \left| \frac{1}{n} \sum_{i=1}^n \frac{\psi_{j,k}^2(x_i)}{g(x_i)} - 1 \right| \geq \alpha \right\} \quad \text{and} \quad r(\underline{x}) = 2 \exp \left( - \frac{\kappa^2 \lambda}{\frac{32}{n} \left\| \frac{1}{g} \right\|_{\infty} \sum_{i=1}^n \frac{\psi_{j,k}^2(x_i)}{g(x_i)}} \right).$$

Applying the usual Gaussian concentration inequality and remarking that  $r(\underline{x}) \leq 2$ , one gets:

$$\begin{aligned} S_2 &= \mathbb{E}_f^n \left( \mathbb{P}_f^n \left( \left| \frac{1}{n} \sum_{i=1}^n \frac{\psi_{j,k}(X_i)}{g(X_i)} \xi_i \right| \geq \frac{\kappa}{4} \sqrt{\frac{\lambda}{n}} \mid X_1, \dots, X_n \right) \right) \\ &\leq \mathbb{E}_f^n (r(X_1, \dots, X_n) 1_{\mathcal{F}_{\alpha}}(X_1, \dots, X_n)) + \mathbb{E}_f^n (r(X_1, \dots, X_n) 1_{\mathcal{F}_{\alpha}^c}(X_1, \dots, X_n)) \\ &\leq 2 \mathbb{P}_f^n (\mathcal{F}_{\alpha}(X_1, \dots, X_n)) + \mathbb{E}_f^n (r(X_1, \dots, X_n) 1_{\mathcal{F}_{\alpha}^c}(X_1, \dots, X_n)). \end{aligned}$$

On the set  $\mathcal{F}_{\alpha}^c(\underline{x})$ , remark that:

$$r(\underline{x}) \leq 2 \exp \left( - \frac{\kappa^2 \lambda}{32 \left\| \frac{1}{g} \right\|_{\infty} (1 + \alpha)} \right).$$

Since  $\mathbb{P}_f^n (\mathcal{F}_{\alpha}^c(X_1, \dots, X_n)) \leq 1$ , we see that:

$$S_2 \leq 2 \left( \mathbb{P}_f^n (\mathcal{F}_{\alpha}(X_1, \dots, X_n)) + \exp \left( - \frac{\kappa^2 \lambda}{32 \left\| \frac{1}{g} \right\|_{\infty} (1 + \alpha)} \right) \right). \quad (4.6)$$

Since  $\frac{\psi_{j,k}^2(X_i)}{g(X_i)}$  are i.i.d.,  $\frac{\psi_{j,k}^2(X_i)}{g(X_i)} \leq \left\| \frac{1}{g} \right\|_{\infty} \|\psi\|_{\infty}^2 2^j$  and  $\mathbb{E}_f^n \left( \frac{\psi_{j,k}^2(X_1)}{g(X_1)} \right) = 1$ , Hoeffding's inequality gives us

$$\mathbb{P}_f^n (\mathcal{F}_{\alpha}(X_1, \dots, X_n)) \leq 2 \exp \left( - \frac{2n^2 \alpha^2}{n \left\| \frac{1}{g} \right\|_{\infty}^2 \|\psi\|_{\infty}^4 2^{2j}} \right) \leq 2 \exp \left( - \frac{2\alpha^2}{\left\| \frac{1}{g} \right\|_{\infty}^2 \|\psi\|_{\infty}^4} \right). \quad (4.7)$$

We used the fact that  $2^j \leq \sqrt{n}$ . Since the choice of  $\alpha$  is free, it follows from (4.6) and (4.7) that there exist  $C > 0$  and a positive function  $h$  satisfying  $\lim_{u \rightarrow \infty} h(u) = \infty$  such that:

$$S_2 \leq C 2^{-h(\kappa)\lambda}. \quad (4.8)$$

Combining the results (4.8) and (4.5), we obtain the desired concentration property.  $\square$

*Proof of Proposition 3.1.* The proof consists in exhibiting the lower bound of the minimax  $\mathbb{L}^p$  risk when we observe the model (1.1) for the densities of the form:

$$g(x) = (\sigma + 1)x^{\sigma}, \quad \sigma > \frac{2}{p}.$$

Let us consider the following two functions:

$$g_1(x) = \gamma_j \psi_{j,N}(x), \quad g_{-1}(x) = -\gamma_j \psi_{j,N}(x)$$

where  $\text{Supp}(\psi_{j,N}) = [\frac{1}{2^j}, \frac{2N}{2^j}]$  and  $\gamma_j \asymp 2^{-j(s+\frac{1}{2}-\frac{1}{p})}$ . Since the wavelet coefficients of  $g_1$  and  $g_{-1}$  are equal to  $\gamma_j$ , it is easy to see that  $g_1$  and  $g_{-1}$  belong to  $B_{\pi,r}^s(L)$ . We have:

$$\|g_1 - g_{-1}\|_p = 2\gamma_j 2^{j(\frac{1}{2}-\frac{1}{p})} \|\psi\|_p = 2\delta_j$$

where  $\delta_j = \gamma_j 2^{j(\frac{1}{2}-\frac{1}{p})} \|\psi\|_p$ . For any  $\hat{f}$ , Chebychev's inequality gives us

$$\delta_j^{-p} \sup_{f \in B_{\pi,r}^s(L)} \mathbb{E}_f^n (\|\hat{f} - f\|_p^p) \geq \sup_{\epsilon \in \{-1,1\}} \mathbb{P}_{g_{\epsilon}}^n (A_{\epsilon}^c) = p_3 \quad (4.9)$$

where the set  $A_\epsilon = \left\{ \|\hat{f} - g_\epsilon\|_p < \delta_j \right\}$  satisfies  $A_1 \cap A_{-1} = \emptyset$ . It follows from Lemma 3.2 that

$$p_3 \geq \min \left( \frac{1}{2}, \sqrt{2} \exp\left(-\frac{3}{e}\right) \exp(-\chi_{\{-1,1\}}) \right). \quad (4.10)$$

Using the equality (3.12) and the fact that  $x^\sigma$  is increasing for  $\sigma > 0$ , one gets

$$\chi_{\{-1,1\}} \leq \frac{1}{2} K(\mathbb{P}_{g_1}^n, \mathbb{P}_{g_{-1}}^n) \leq (\sigma + 1) n \gamma_j^2 \int_{\frac{1}{2^j}}^{\frac{2N}{2^j}} \psi_{j,N}^2(x) x^\sigma dx \leq C n \gamma_j^2 2^{-j\sigma}. \quad (4.11)$$

Putting (4.9), (4.10) and (4.11) together and choosing  $\gamma_j \asymp 2^{j\frac{\sigma}{2}} n^{-\frac{1}{2}}$  i.e  $2^j \asymp n^{\frac{1}{2s+1+\sigma-\frac{2}{\pi}}}$  we obtain:

$$\delta_j^{-p} \sup_{f \in B_{s,r}^{\frac{s}{\pi}}(L)} \mathbb{E}_f^n(\|\hat{f} - f\|_p^p) \geq c\sqrt{2} \exp(-\chi_{\{-1,1\}}) \geq c'$$

where

$$\delta_j \asymp n^{-\frac{s-\frac{1}{\pi}+\frac{1}{p}}{2s+1+\sigma-\frac{2}{\pi}}}.$$

Since the choice of  $\sigma > 0$  is free, the proof of Proposition 3.1 ends by taking  $\sigma$  large enough.  $\square$

## References

- BIRGE, L. (2001). A new look at an old result: Fano's lemma. *Technical Report, Université Paris VI*.
- COHEN, A., DAUBECHIES, I., JAWERTH, B., and VIAL, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 24(1):54–81.
- DELYON, B. and JUDITSKY, A. (1996). On minimax wavelet estimators. *Applied Computational Harmonic Analysis*, 3:215–228.
- DEVORE, R., KERKYACHARIAN, G., PICARD, D., and TEMLYAKOV, V. (2005). Mathematical methods for supervised learning. *IMI preprint*, 22:1–51.
- DONOHO, D., JOHNSTONE, I., KERKYACHARIAN, G., and PICARD, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.*, 24(2):508–539.
- IBRAGIMOV, I. and KHAMINSKII, R. (1982). Bounds for the quality of nonparametric estimation of regression. *Theory Probab. Appl.*, 27(1):81–94.
- JOHNSTONE, I., KERKYACHARIAN, G., PICARD, D., and RAIMONDO, M. (2004). Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society. Series B. Methodological*, 66:547–573.
- KERKYACHARIAN, G. and PICARD, D. (2000). Thresholding algorithms, maxisets and well concentrated bases. *Test*, 9(2):283–345.
- KERKYACHARIAN, G. and PICARD, D. (2005). Regression in random design and warped wavelets. *Bernoulli*, 10(6):1053–1105.
- MEYER, Y. (1990). *Ondelettes et Opérateurs*. Hermann, Paris.
- STONE, C. (1982). Optimal global rates of convergence for nonparametric regression. *Annals of statistics*, 10(4):1040–1053.
- TSYBAKOV, A. (2004). *Introduction à l'estimation nonparamétrique*. Springer Verlag.