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. 2006. hal-00014248v2

HAL Id: hal-00014248
https://hal.science/hal-00014248v2

Preprint submitted on 2 Jan 2006

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
01 December 2005

Abstract

The problem of estimating a regression function based on a regression model with (known) random design is considered. By adopting the framework of wavelet analysis, we establish the asymptotic minimax rate of convergence under the $\mathbb{L}^p$ risk over Besov balls. A part of this paper is devoted to the case where the design density is vanishing.
Key Words: Regression with random design, minimax rate of convergence, Besov spaces.

AMS subject classification: 62G07, 62G20.

1 Motivation

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

$$Y_i = f(X_i) + \xi_i, \quad i = 1, ..., n, \tag{1.1}$$

where the ξ_i's are Gaussian i.i.d with mean zero, variance one and are independent of the design $X_1, ..., X_n$. 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 $(X_1, Y_1), ..., (X_n, Y_n)$.

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 $(X_1, Y_1), ..., (X_n, Y_n)$. 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:

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

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

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

Let us notice 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 $\mathbb{L}^p$ norms (unconditional nature and Temlyakov's property). 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 the $\mathbb{L}^p$ norms and the main function spaces of the study. The minimax results over Besov balls and the associated proofs are 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 $\mathbb{L}^p([0,1]) = \{f \text{ measurable on } [0,1] \mid \|f\|_p^p = \int_0^1 |f(t)|^p dt < +\infty\}$. The constants C and c represent any constants we shall need, and can be different from one line to one other. The notation $a \asymp b$ means: there exist two constants $c > 0$ and $C > 0$ such that $cb \leq a \leq Cb$. The notations $(a)_+$ and $a \wedge b$ mean respectively $max(a, 0)$ and $max(a, b)$.

2.1 Wavelet bases and geometrical properties in $\mathbb{L}^p$ norms

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

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

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

$$\xi = \{\phi_{\tau,k}(.),\ k = 0, ..., 2^\tau - 1;\ \psi_{j,k}(.);\ j \geq \tau,\ k = 0, ..., 2^j - 1\}$$

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 such wavelet bases.

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

$$f(x) = \sum_{k \in \Delta_\tau} \alpha_{\tau,k} \phi_{\tau,k}(x) + \sum_{j \geq \tau} \sum_{k \in \Delta_j} \beta_{j,k} \psi_{j,k}(x), \quad x \in [0,1],$$

where $\alpha_{j,k} = \int_0^1 f(t)\phi_{j,k}(t)dt$, $\beta_{j,k} = \int_0^1 f(t)\psi_{j,k}(t)dt$ and $\Delta_j = \{0, ..., 2^j - 1\}$. Let us denote by $P_\tau(f)(x)$ the first term of this decomposition. The following lemmas set some inequalities linked to the basis ξ.

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

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

Lemma 2.2. *Let $p > 1$. For any $j \geq \tau$ we have:*

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

The Lemmas (2.1) and (2.2) are obviously true if we exchange ϕ by ψ. Let us introduce two important geometrical properties concerning the weighted compactly supported wavelet bases under the $\mathbb{L}^p$ norms.

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

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

Let $\sigma \in \mathbb{R} - \{-2^{-1}\}$. Then the weighted compactly supported wavelet basis ξ_σ defined by

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

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

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

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 proof of the second point.

2.2 Function spaces

Definition 2.2 (Besov balls). *Let $N \in \mathbb{N}^*$, $0 < s < N$, $1 \leq r \leq \infty$ and $1 \leq \pi \leq \infty$. For any measurable function f 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 $J_{Nh} = \{x \in [0,1] : \ x + Nh \in [0,1]\}$. We say that a function f of $\mathbb{L}^{\pi}([0,1])$ belongs to the Besov balls $B^s_{\pi,r}(L)$ 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 ξ 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*

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

3 Minimax study over Besov balls

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

Considering the parameters (s,π,r) of the Besov balls $B^s_{\pi,r}(L)$, we adopt the following notations:

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

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

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

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

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

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

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

$$\begin{cases} \mathcal{R}_n(B^s_{\pi,r}(L)) \asymp n^{-\alpha_1 p} & \text{if} \quad \epsilon > 0, \\ \mathcal{R}_n(B^s_{\pi,r}(L)) \asymp \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} & \text{if} \quad 0 > \epsilon. \end{cases}$$

For the case where $\epsilon = 0$, there exist two constants $C > 0$ and $c > 0$ such that:

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

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 $c_* \in]0,1]$. Let us consider the following procedure:

$$\hat{f}_u(x) = \sum_{k\in\Delta_{j_0}} \hat{\alpha}_{j_0,k}\phi_{j_0,k}(x) + u \sum_{j_0\le j<j_1}\sum_{k\in\Delta_j} \hat{\beta}_{j,k} 1_{\left\{|\hat{\beta}_{j,k}|\ge\kappa\sqrt{\frac{j-j_0+1}{n}}\right\}}\psi_{j,k}(x), \quad x\in[0,1], \tag{3.2}$$

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

$$\hat{\alpha}_{j,k} = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{g(X_i)}\phi_{j,k}(X_i), \qquad \hat{\beta}_{j,k} = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{g(X_i)}\psi_{j,k}(X_i) \tag{3.3}$$

and the integers j_0 and j_1 are chosen such that:

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

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 $f \in B^s_{\pi,r}(L) \subset B^s_{p,\infty}(L)$ (since $\pi \ge p$), the $\mathbb{L}^p$ risk of $\hat{f}_0$ can be dominated as follows:

$$\begin{aligned}\mathbb{E}^n_f(\|\hat{f}_0 - f\|_p^p) &\le C\left(\mathbb{E}^n_f(\|\hat{f}_0 - P_{j_0}(f)\|_p^p) + \|P_{j_0}(f) - f\|_p^p\right)\\ &\le C\left(2^{j_0(\frac{p}{2}-1)}\sum_{k\in\Delta_{j_0}} \mathbb{E}^n_f(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) + 2^{-j_0 sp}\right).\end{aligned} \tag{3.4}$$

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

Lemma 3.1 (Moments inequality). *Let $2 \le a < \infty$, $j \ge \tau$ 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:*

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

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

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

The l_p-Hölder inequality, Lemma 2.1 and the condition (3.1) yield:

$$\begin{aligned}\sum_{k\in\Delta_{j_0}} \mathbb{E}^n_f(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) &\le Cn^{-\frac{p}{2}}\left(\sum_{k\in\Delta_{j_0}}\int_0^1 \frac{|\phi_{j_0,k}(x)|^2}{g(x)}dx\right)^{\frac{p}{2}} (Card(\Delta_{j_0}))^{1-\frac{p}{2}}\\ &\le Cn^{-\frac{p}{2}}\left(2^{j_0}\int_0^1 \frac{1}{g(x)}dx\right)^{\frac{p}{2}} 2^{j_0(1-\frac{p}{2})} \le C2^{j_0}n^{-\frac{p}{2}}.\end{aligned} \tag{3.5}$$

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

$$\begin{aligned}\sum_{k\in\Delta_{j_0}} \mathbb{E}^n_f(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) &\le C\left(n^{1-p}\int_0^1 \frac{\sum_{k\in\Delta_{j_0}}|\phi_{j_0,k}(x)|^p}{g^{p-1}(x)}dx + n^{-\frac{p}{2}}\int_0^1 \frac{\sum_{k\in\Delta_{j_0}}|\phi_{j_0,k}(x)|^2}{g^{\frac{p}{2}}(x)}dx\right)\\ &\le C\left(2^{\frac{j_0 p}{2}}n^{1-p} + 2^{j_0}n^{-\frac{p}{2}}\right) \le C2^{j_0}n^{-\frac{p}{2}}.\end{aligned} \tag{3.6}$$

Putting (3.4), (3.5), (3.6) together, we deduce the existence of a constant $C > 0$ such that:

$$\sup_{f\in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f}_0 - f\|_p^p) \le C\left(2^{\frac{j_0 p}{2}}n^{-\frac{p}{2}} + 2^{-j_0 sp}\right) \le Cn^{-\alpha_1 p}$$

for $p \geq 1$. This implies that $\mathcal{R}_n(B^s_{\pi,r}(L)) \leq Cn^{-\alpha_1 p}$.

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

$$K(P,Q) = \begin{cases} \int \ln(\frac{d\mathbb{P}}{d\mathbb{Q}})d\mathbb{P} & \text{if } \mathbb{P} << \mathbb{Q}, \\ \infty & \text{otherwise.} \end{cases}$$

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

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

where

$$\chi_m = \inf_{v\in\{0,1,\dots,m\}} \frac{1}{m}\sum_{k\neq v} K(\mathbb{P}_k, \mathbb{P}_v). \tag{3.7}$$

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

$$g_\varepsilon(x) = \gamma_j \sum_{k\in\Delta_j} \varepsilon_k \psi_{j,k}(x), \qquad x \in [0,1],$$

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_ϵ are equal to $\gamma_j\epsilon_k$, it follows from Lemma 2.4 that $g_\varepsilon \in B^s_{\pi,r}(L)$. 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 exist a subset $E_j = \{\epsilon^{(0)}, ..., \epsilon^{(T_j)}\}$ of $\{0,1\}^{2^j}$ and two constants $c \in]0,1[$, $\alpha \in]0,1[$ such that $\forall 0 \leq u < v \leq T_j$:

$$\sum_{k\in\Delta_j} |\epsilon_k^{(u)} - \epsilon_k^{(v)}| \geq c2^j \qquad \text{and} \qquad T_j \geq e^{\alpha 2^j}.$$

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

$$\|g_{\epsilon^{(u)}} - g_{\epsilon^{(v)}}\|_p = c2^{\frac{j}{2}}\gamma_j(2^{-j}\sum_{k\in\Delta_j}|\epsilon_k^{(u)} - \epsilon_k^{(v)}|)^{\frac{1}{p}} \geq 2\delta_j$$

where $\delta_j = c2^{\frac{j}{2}}\gamma_j$. Using Chebychev's inequality, for any $\hat{f}$ we have:

$$\delta_j^{-p} \sup_{f\in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f} - f\|_p^p) \geq \sup_{u\in\{0,\dots,T_j\}} \mathbb{P}^n_{g_{\epsilon^{(u)}}}(A_u^c) = p_1$$

where the sets A_u defined by $A_u = \left\{\|\hat{f} - g_{\epsilon^{(u)}}\|_p < \delta_j\right\}$ satisfy $A_u \cap A_v = \varnothing$ for $u \neq v$ and $u, v \in \{0, ..., T_j\}$. Thus, Lemma 3.2 gives us

$$p_1 \geq \min\left(2^{-1}, \sqrt{T_j}\exp(-3e^{-1})\exp(-\chi_{T_j})\right). \tag{3.8}$$

where $\chi_.$ is defined by (3.7). Now, let us consider the following lemma which will be proved in Appendix.

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

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

By virtue of Lemma 3.3, Lemma 2.1, the facts that $|\epsilon_k^{(u)} - \epsilon_k^{(v)}| \leq 2$ and that $\int_0^1 g(x)dx = 1$, for all $u \neq v$ and $u, v \in \{0, ..., T_j\}$ we have:

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

Hence

$$\chi_{T_j} \leq \inf_{v\in\{0,\ldots,T_j\}} \sup_{u\neq v; u\in\{0,\ldots,T_j\}} K(\mathbb{P}^n_{g_{\epsilon(u)}}, \mathbb{P}^n_{g_{\epsilon(v)}}) \leq Cn\gamma_j^2 2^j. \tag{3.9}$$

Putting (3.8) and (3.9) together and choosing $\gamma_j = c_0 \frac{1}{\sqrt{n}}$ where c_0 denotes a well chosen constant, one gets:

$$\delta_j^{-p} \sup_{f\in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f} - f\|_p^p) \geq c\exp(\frac{\alpha}{2}2^j - Cc_0^2 2^j) \geq c$$

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

Proof of Theorem 3.2. Proof of the upper bound: Proposition 3.2 below provides upper bounds over $B^s_{\pi,r}(L)$ under the $\mathbb{L}^p$ 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 risk can be extended to the $\mathbb{L}^p$ risk for $1 \leq p < \infty$ under mild assumptions on the model.

Proposition 3.2 (Upper bounds 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 (3.2) with $u = 1$. Adopting the notation $\hat{\beta}_{j_0,k} = \hat{\alpha}_{j_0,k}$, suppose that there exists a constant $C > 0$ such that:*

$$\mathbb{E}^n_f(|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}) \leq Cn^{-p}, \qquad j_0 \leq j < j_1, \tag{3.10}$$

and that there exist $\kappa > 0$ and a positive function h satisfying $\lim_{u\to\infty} h(u) = +\infty$ such that the following concentration condition holds:

$$\mathbb{P}^n_f\left(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa}{2}\sqrt{\frac{j - j_0 + 1}{n}}\right) \leq C2^{-(j-j_0)h(\kappa)}, \qquad j_0 \leq j < j_1, \tag{3.11}$$

for a suitably chosen c_*. Then for $p > \pi$ and $s > \frac{1}{\pi} - \min\left(\frac{1}{p}, \frac{1}{2}(1 - \frac{1}{c_*})\right)$, there exists a constant $C > 0$ such that:

$$\begin{cases} \sup_{f\in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f}_1 - f\|_p^p) \leq Cn^{-\alpha_1 p} & if \quad \epsilon > 0, \\ \sup_{f\in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f}_1 - f\|_p^p) \leq C\left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} & if \quad \epsilon < 0, \\ \sup_{f\in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f}_1 - f\|_p^p) \leq C\left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} (\ln(n))^{(p-\frac{\pi}{r})_+} & if \quad \epsilon = 0. \end{cases}$$

For the case where $\epsilon > 0$, 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). Let us mention that the proof of Proposition 3.10 intensively uses the geometrical properties of the basis ξ (see Lemma 2.3).

Lemma 3.4. *Assume that we observe the regression model (1.1) and that g is bounded from below. Then the estimators $\hat{\beta}_{j,k}$ and $\hat{\alpha}_{j,k}$ described by (3.3) satisfy the conditions (3.10) and (3.11) for $c_* = 2^{-1}$.*

The proofs of the upper bounds of Theorem 3.2 are thus complete.

Proof of the lower bound: Let us introduce the following family:

$$\{g_k(x) = \gamma_j \psi_{j,k}(x),\ k \in \Delta_j, \quad g_{2^j}(x) = 0\}, \quad x \in [0, 1],$$

where j is an integer to be chosen below and $\gamma_j \asymp 2^{-j\left(s+\frac{1}{2}-\frac{1}{\pi}\right)}$. Since the wavelet coefficients of g_k are equal to γ_j, it follows from Lemma 2.4 that $g_k \in B^s_{\pi,r}(L)$. Moreover for $k \neq k'$ and $k \neq 0$, Lemma 2.2 gives us:

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

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

$$\delta_j^{-p} \sup_{f \in B^s_{\pi,r}(L)} \mathbb{E}^n_f(\|\hat{f} - f\|_p^p) \geq \sup_{k \in \{0,1,\ldots,2^j\}} \mathbb{P}^n_{g_k}(A_k^c) = p_2$$

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

$$p_2 \geq \min\left(2^{-1}, \sqrt{2^j}\exp(-3e^{-1})\exp(-\chi_{2^j})\right). \tag{3.12}$$

where $\chi_{\cdot}$ is defined by (3.7). Using Lemma 3.3 and Lemma 2.1, one gets:

$$\chi_{2^j} \leq 2^{-j}\sum_{k \in \Delta_j} K(\mathbb{P}^n_{g_k}, \mathbb{P}^n_{g_{2^j}}) \leq C2^{-j}n\gamma_j^2 \int_0^1 g(x) \sum_{k \in \Delta_j} \psi^2_{j,k}(x)dx \leq Cn\gamma_j^2 \int_0^1 g(x)dx = C\gamma_j^2 n. \tag{3.13}$$

Let us set $\gamma_j = c_0\sqrt{\frac{\ln(n)}{n}}$ where c_0 denotes an arbitrary positive constant (in particular, this implies that: $2^j \asymp \left(\sqrt{\frac{n}{\ln(n)}}\right)^{\frac{1}{s+\frac{1}{2}-\frac{1}{\pi}}}$). Thus, for n large enough, we see that:

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

where $\lambda_n = \ln(n)$. For a suitable choice of c_0, it follows from (3.12) and (3.13) that:

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

where $\delta_j \asymp \sqrt{\frac{\ln(n)}{n}}\left(\sqrt{\frac{n}{\ln(n)}}\right)^{\frac{\frac{1}{2}-\frac{1}{p}}{s+\frac{1}{2}-\frac{1}{\pi}}} = \left(\frac{\ln(n)}{n}\right)^{\alpha_2}$. We conclude that there exists a constant $c > 0$ such that $\mathcal{R}_n(\mathcal{B}^s_{\pi,r}(L)) \geq c\left(\frac{\ln(n)}{n}\right)^{\alpha_2 p}$. This ends the proof of Theorem 3.2. □

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 $\mathbb{E}^n_f\left(|\frac{\phi_{j,k}(X_1)}{g(X_1)}|\right) = \int_0^1 |\phi_{j,k}(x)|dx < \infty$, it is clear that

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

Thus, 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

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

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

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

The condition (3.1) gives us:

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

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

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

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. □

Proof of Lemma 3.3. The law of (X_1, Y_1) is given by $p(x,y) = \frac{1}{\sqrt{2\pi}} g(x) \exp\left(-\frac{1}{2}(y - f(x))^2\right)$. Thus, for any function f_1 and f_2 measurable on $[0,1]$ and bounded from above 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. □

Proof of Proposition 3.2. Let us denote $\lambda_j = \kappa\sqrt{\frac{j - j_0 + 1}{n}}$. Proceeding as in Donoho et al. (1996), we obtain the following decomposition:

$$\begin{aligned} \hat{f}_1(x) - f(x) &= \sum_{k \in \Delta_{j_0}} (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}) \phi_{j_0,k}(x) + \sum_{j > j_1} \sum_{k \in \Delta_j} \beta_{j,k} \psi_{j,k}(x) \\ &+ \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} (\hat{\beta}_{j,k} - \beta_{j,k}) \left(1_{\{|\hat{\beta}_{j,k}| \geq \lambda_j\}} 1_{\left\{|\beta_{j,k}| < \frac{\lambda_j}{2}\right\}} + 1_{\{|\hat{\beta}_{j,k}| \geq \lambda_j\}} 1_{\left\{|\beta_{j,k}| \geq \frac{\lambda_j}{2}\right\}} \right) \psi_{j,k}(x) \\ &+ \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} \beta_{j,k} \left(1_{\{|\hat{\beta}_{j,k}| < \lambda_j\}} 1_{\{|\beta_{j,k}| \geq 2\lambda_j\}} + 1_{\{|\hat{\beta}_{j,k}| < \lambda_j\}} 1_{\{|\beta_{j,k}| < 2\lambda_j\}} \right) \psi_{j,k}(x) \\ &= e_1 + e_2 + e_3 + e_4 + e_5 + e_6. \end{aligned}$$

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

$$\begin{aligned} \mathbb{E}_f^n(\|\hat{f}_1 - f\|_p^p) &\leq 6^{p-1} \left(\mathbb{E}_f^n(\|e_1\|_p^p) + \mathbb{E}_f^n(\|e_2\|_p^p) + \mathbb{E}_f^n(\|e_3\|_p^p) + \mathbb{E}_f^n(\|e_4\|_p^p) + \mathbb{E}_f^n(\|e_5\|_p^p) + \mathbb{E}(\|e_6\|_p^p) \right) \\ &= C(\tilde{e}_1 + \tilde{e}_2 + \tilde{e}_3 + \tilde{e}_4 + \tilde{e}_5 + \tilde{e}_6). \end{aligned}$$

Let us analyze each term $\tilde{e}_i$, i=1,2,3,4,5,6, in turn. *The upper bound for the term $\tilde{e}_1$.* Proceeding as in the proof of Theorem 3.1, it is clear that:

$$\tilde{e}_1 = \mathbb{E}_f^n(\|e_1\|_p^p) \leq C 2^{j_0(\frac{p}{2} - 1)} \sum_{k \in \Delta_{j_0}} \mathbb{E}_f^n(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p) \leq C 2^{\frac{j_0 p}{2}} n^{-\frac{p}{2}}.$$

The upper bound for the term $\tilde{e}_2$. It follows from the Minkowski inequality, Lemma 2.2 and the inclusions $B_{\pi,r}^s(L) \subseteq B_{p,r}^{s - \frac{1}{\pi} + \frac{1}{p}}(L) \subseteq B_{p,\infty}^{s - \frac{1}{\pi} + \frac{1}{p}}(L)$ (which holds for $p > \pi$) that:

$$\tilde{e}_2 = \mathbb{E}_f^n(\|e_2\|_p^p) \leq C \left(\sum_{j \geq j_1} 2^{j(\frac{1}{2} - \frac{1}{p})} (\sum_{k \in \Delta_j} |\beta_{j,k}|^p)^{\frac{1}{p}} \right)^p \leq C (\sum_{j \geq j_1} 2^{-j(s - \frac{1}{\pi} + \frac{1}{p})})^p \leq C 2^{-j_1(s - \frac{1}{\pi} + \frac{1}{p})p}.$$

The upper bounds for the terms $\tilde{e}_3$ and $\tilde{e}_5$. Since the following inclusions hold:

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

$$\left\{|\hat{\beta}_{j,k}| < \lambda_j,\ |\beta_{j,k}| \geq 2\lambda_j\right\} \subset \left\{|\beta_{j,k}| \leq 2|\hat{\beta}_{j,k} - \beta_{j,k}|\right\},$$

we can group the terms $\tilde{e}_3$ and $\tilde{e}_5$. Now, observe that the Cauchy-Schwartz inequality combined with the conditions (3.10) and (3.11) gives us:

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

The rest of the proof uses arguments similar to Kerkyacharian and Picard (2000, Subsection 5.1.1). For $p \geq 2$, using the unconditional nature of the basis ξ (see Lemma 2.3), the generalized Minkowski inequality, the inequality (4.4) and the fact that ξ_σ satisfies the Temlyakov property (see again Lemma 2.3), one obtains:

$$\begin{aligned}
\tilde{e}_3 + \tilde{e}_5 &= \mathbb{E}_f^n(\|e_3\|_p^p + \|e_5\|_p^p) \leq C\mathbb{E}_f^n(\|(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^2 1_{\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\right\}} |\psi_{j,k}|^2)^{\frac{1}{2}}\|_p^p) \\
&\leq C\|(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n\left(|\hat{\beta}_{j,k} - \beta_{j,k}|^p 1_{\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\right\}}\right)^{\frac{2}{p}} |\psi_{j,k}|^2)^{\frac{1}{2}}\|_p^p \\
&\leq Cn^{-\frac{p}{2}}\|(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} 2^{-\frac{(j-j_0)h(\kappa)}{p}} |\psi_{j,k}|^2)^{\frac{1}{2}}\|_p^p \\
&\leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} 2^{-\frac{(j-j_0)h(\kappa)}{2}} \|\psi_{j,k}\|_p^p \leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} 2^{\frac{jp}{2}} 2^{-\frac{(j-j_0)h(\kappa)}{2}} \\
&\leq Cn^{-\frac{p}{2}} 2^{\frac{j_0 p}{2}} \sum_{j \geq 0} 2^{\frac{j}{2}(p-h(k))} \leq Cn^{-\frac{p}{2}} 2^{\frac{j_0 p}{2}}
\end{aligned}$$

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

$$\begin{aligned}
\tilde{e}_3 + \tilde{e}_5 &= \mathbb{E}_f^n(\|e_3\|_p^p + \|e_5\|_p^p) \leq C\mathbb{E}_f^n(\|(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^2 1_{\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\right\}} |\psi_{j,k}|^2)^{\frac{1}{2}}\|_p^p) \\
&\leq C\mathbb{E}_f^n(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^p 1_{\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda_j}{2}\right\}} \|\psi_{j,k}\|_p^p) \\
&\leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} 2^{-\frac{(j-j_0)h(\kappa)}{2}} \|\psi_{j,k}\|_p^p \leq Cn^{-\frac{p}{2}} 2^{\frac{j_0 p}{2}}
\end{aligned}$$

for κ large enough.

The upper bound for the term $\tilde{e}_4$. In the case where $p \geq 2$, we use again the unconditional nature of the basis ξ, the generalized Minkowski inequality, the condition (3.10), the fact that ξ satisfies the Temlyakov's property, the Makov inequality and the characterization of Besov spaces:

$$\begin{aligned}
\tilde{e}_4 &= \mathbb{E}_f^n(\|e_4\|_p^p) = \mathbb{E}_f^n(\|(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^2 1_{\left\{|\beta_{j,k}| > \frac{\lambda_j}{2}\right\}} |\psi_{j,k}|^2)^{\frac{1}{2}}\|_p^p) \\
&\leq C\|(\sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n\left(|\hat{\beta}_{j,k} - \beta_{j,k}|^p\right)^{\frac{2}{p}} 1_{\left\{|\beta_{j,k}| > \frac{\lambda_j}{2}\right\}} |\psi_{j,k}|^2)^{\frac{1}{2}}\|_p^p \\
&\leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} 1_{\left\{|\beta_{j,k}| > \frac{\lambda_j}{2}\right\}} \|\psi_{j,k}\|_p^p \leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} 2^{j(\frac{p}{2}-1)} \lambda_j^{-\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi \\
&\leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} 2^{j(\frac{p}{2}-1)} \frac{n^{\frac{\pi}{2}}}{(j-j_0+1)^{\frac{\pi}{2}}} 2^{-j(s+\frac{1}{2}-\frac{1}{\pi})\pi} \leq Cn^{\frac{\pi-p}{2}} \sum_{j_0 \leq j < j_1} \frac{2^{-j\epsilon}}{(j-j_0+1)^{\frac{\pi}{2}}}.
\end{aligned}$$

In the case where $1 \leq p < 2$, we use again the unconditional nature of the basis ξ, the comparison beetween the l_p norms, the Cauchy-Schwartz inequality, the condition (3.10) and the characterization of the Besov spaces:

$$\begin{aligned}\tilde{e}_4 &= \mathbb{E}_f^n(\|e_4\|_p^p) \leq C \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} \mathbb{E}_f^n \left(|\hat{\beta}_{j,k} - \beta_{j,k}|^p \right) 1_{\left\{ |\beta_{j,k}| > \frac{\lambda_j}{2} \right\}} \|\psi_{j,k}\|_p^p \\ &\leq Cn^{-\frac{p}{2}} \sum_{j_0 \leq j < j_1} \sum_{k \in \Delta_j} 1_{\left\{ |\beta_{j,k}| > \frac{\lambda_j}{2} \right\}} \|\psi_{j,k}\|_p^p \leq Cn^{\frac{\pi - p}{2}} \sum_{j_0 \leq j < j_1} \frac{2^{-j\epsilon}}{(j - j_0 + 1)^{\frac{\pi}{2}}}.\end{aligned}$$

The upper bound for the term $\tilde{e}_6$. By virtue of the Minkowski inequality, Lemma 2.2 and the characterization of the Besov spaces, we see that:

$$\begin{aligned}\tilde{e}_6 &= \mathbb{E}_f^n(\|e_6\|_p^p) \leq C \left(\sum_{j_0 \leq j < j_1} 2^{j(\frac{1}{2} - \frac{1}{p})} (\sum_{k \in \Delta_j} |\beta_{j,k}|^p 1_{\{|\beta_{j,k}| < 2\lambda_j\}})^{\frac{1}{p}} \right)^p \\ &\leq C \left(\sum_{j_0 \leq j < j_1} 2^{j(\frac{1}{2} - \frac{1}{p})} ((2\lambda_j)^{p-\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi)^{\frac{1}{p}} \right)^p \\ &\leq Cn^{\frac{\pi - p}{2}} \left(\sum_{j_0 \leq j < j_1} 2^{j(\frac{1}{2} - \frac{1}{p})} (j - j_0 + 1)^{\frac{p-\pi}{2p}} 2^{-j(s + \frac{1}{2} - \frac{1}{\pi})\frac{\pi}{p}} \right)^p \\ &\leq Cn^{\frac{\pi - p}{2}} \left(\sum_{j_0 \leq j < j_1} 2^{-\frac{j\epsilon}{p}} (j - j_0 + 1)^{\frac{p-\pi}{2p}} \right)^p.\end{aligned}$$

Since $\tilde{e}_3$, $\tilde{e}_5$ and $\tilde{e}_1$ are of the same order, it suffices to balance the bounds of $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_4$ and $\tilde{e}_6$ to obtain the optimal upper bounds.

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

$$\tilde{e}_1 \leq C 2^{\frac{j_0 p}{2}} n^{-\frac{p}{2}} \leq Cn^{-\alpha_1 p}, \qquad \tilde{e}_2 \leq C \left(\frac{\ln(n)}{n} \right)^{c_*(s - \frac{1}{\pi} + \frac{1}{p})p} \leq Cn^{-\alpha_1 p},$$

$$(\tilde{e}_4 \wedge \tilde{e}_6) \leq Cn^{\frac{\pi - p}{2}} 2^{-\epsilon j_0} (\sum_{j \geq 1} j^{\frac{p-\pi}{2p}} 2^{-j\frac{\epsilon}{p}})^p \leq Cn^{\frac{\pi - p}{2}} 2^{-\epsilon j_0} \leq Cn^{-\alpha_1 p}$$

and we obtain the desired upper bound.

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

$$\tilde{e}_1 \leq Cn^{-\alpha_1 p}, \ \tilde{e}_2 \leq C \left(\frac{\ln(n)}{n} \right)^{\alpha_2 p} \quad \text{and} \quad (\tilde{e}_4 \wedge \tilde{e}_6) \leq C j_1^{\frac{p-\pi}{2}} 2^{-\epsilon j_1} n^{\frac{\pi - p}{2}} \leq C \left(\frac{\ln(n)}{n} \right)^{\alpha_2 p}$$

and we establish the desired upper bound.

Finally, in the special case where $\epsilon = 0$, we have:

$$\tilde{e}_6 \leq C \left(\sum_{j_0 \leq j < j_1} 2^{j(\frac{1}{2} - \frac{1}{p})} ((2\lambda_j)^{p-\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi)^{\frac{1}{p}} \right)^p \leq Cn^{\frac{\pi - p}{2}} j_1^{\frac{p-\pi}{2}} \left(\sum_{j_0 \leq j < j_1} l_j \right)^p$$

where $l_j = (2^{j(s + \frac{1}{2} - \frac{1}{\pi})\pi} \sum_{k \in \Delta_j} |\beta_{j,k}|^\pi)^{\frac{1}{p}}$. Now, let us distinguish the case where $\pi \geq rp$ and the case where $\pi < rp$. If $\pi \geq rp$ then $B^s_{\pi,r}(L)$ is included in $B^s_{\pi,\frac{\pi}{p}}(L)$ so we have:

$$\tilde{e}_6 \leq Cn^{\frac{\pi - p}{2}} j_1^{\frac{p-\pi}{2}} \leq C \left(\frac{\ln(n)}{n} \right)^{\alpha_2 p}.$$

Let us consider the case where $\pi < rp$. Using Hölder's inequality and the fact that $f \in B^s_{\pi,r}(L) \subseteq B^s_{\pi,\infty}(L)$, it comes:

$$(\sum_{j_0 \leq j < j_1} l_j)^p \leq (\sum_{j_0 \leq j < j_1} l_j^{\frac{pr}{\pi}})^{\frac{\pi}{r}} (\sum_{j_0 \leq j < j_1} l_j^{\frac{1}{1 - \frac{\pi}{rp}}})^{p - \frac{\pi}{r}} \leq C(\sum_{j_0 \leq j < j_1} 1)^{p - \frac{\pi}{r}} \leq C j_1^{(p - \frac{\pi}{r})}$$

and a fortiori:

$$\tilde{e}_6 \leq C j_1^{(p-\frac{\pi}{r})} n^{\frac{\pi-p}{2}} j_1^{\frac{p-\pi}{2}}.$$

Since $(\tilde{e}_4 \wedge \tilde{e}_6) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} j_1^{(p-\frac{\pi}{r})_+}$, we obtain the desired upper bound. The proof of Proposition 3.2 is thus complete. □

Proof of Lemma 3.4. Proof of the moments inequality. Since g is supposed to be bounded from below, we have

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

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

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

Proof of the concentration condition. Here, we proceed as in Kerkyacharian and Picard (2005, subsection 9.1). Let $\kappa > 0$. Let us denote $u_j = j - j_0 + 1$. By a simple decomposition, one gets:

$$\begin{aligned} U_n &= \mathbb{P}_f^n \left(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa}{2} \sqrt{\frac{u_j}{n}} \right) \\ &\leq \mathbb{P}_f^n \left(|\frac{1}{n} \sum_{i=1}^n (\frac{\psi_{j,k}(X_i) f(X_i)}{g(X_i)} - \beta_{j,k})| \geq \frac{\kappa}{4} \sqrt{\frac{u_j}{n}} \right) + \mathbb{P}_f^n \left(|\frac{1}{n} \sum_{i=1}^n \frac{\psi_{j,k}(X_i)}{g(X_i)} \xi_i| \geq \frac{\kappa}{4} \sqrt{\frac{u_j}{n}} \right) \\ &= S_1 + S_2. \end{aligned}$$

Let us analyze each terms S_1 and S_2 in turn. *The upper bound for* S_1. Since $\|\frac{\psi_{j,k}(X_1) f(X_1)}{g(X_1)} - \beta_{j,k}\|_\infty \leq C 2^{\frac{j}{2}}$ and $\mathbb{E}_f^n \left(|\frac{\psi_{j,k}(X_1) f(X_1)}{g(X_1)} - \beta_{j,k}|^2 \right) \leq C$, it follows from Bernstein's inequality that for $2^j \leq \sqrt{n}$ there exists $C > 0$ satisfying

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

for n large enough. We used the fact that $\lim_{n \to +\infty} u_j^{\frac{1}{2}} n^{-\frac{1}{4}} = 0$ for all $j_0 \leq j < j_1$. Thus, there exist a constant $C > 0$ and a positive function z satisfying $\lim_{x \to \infty} z(x) = \infty$ such that:

$$S_1 \leq C 2^{-(j-j_0) z(\kappa)}. \tag{4.5}$$

The upper bound for S_2. Conditionally on $(X_1, ..., X_n) = (x_1, ..., x_n)$ we have

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

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

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

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

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

On the set $\mathcal{F}_\alpha^c(\underline{x})$, remark that $r(\underline{x}) \leq 2\exp\left(-\frac{\kappa^2 u_j}{32\|\frac{1}{g}\|_\infty(1+\alpha)}\right)$. Since $\mathbb{P}_f^n(\mathcal{F}_\alpha^c(X_1,...,X_n)) \leq 1$, we see that:

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

Since $\frac{\psi_{j,k}^2(X_i)}{g(X_i)}$ are i.i.d, $\frac{\psi_{j,k}^2(X_i)}{g(X_i)} \leq \|\frac{1}{g}\|_\infty \|\psi\|_\infty^2 2^j$, $\mathbb{E}_f^n\left(\frac{\psi_{j,k}^2(X_1)}{g(X_1)}\right) = 1$, Hoeffding's inequality and the choice $\alpha \asymp \kappa$ imply that:

$$\mathbb{P}_f^n(\mathcal{F}_\alpha(X_1,...,X_n)) \leq 2\exp\left(-\frac{2n^2\alpha^2\ln(n)}{n\|\frac{1}{g}\|_\infty^2\|\psi\|_\infty^4 2^{2j}}\right) \leq Cn^{-C\alpha^2} \leq C2^{-j_1 C\kappa^2} \leq C2^{-(j-j_0)C\kappa^2} \tag{4.7}$$

for $j_0 \leq j < j_1$ and κ large enough. We used the fact that $2^j \leq C\sqrt{\frac{n}{\ln(n)}}$. It follows from (4.6) with α such that $\alpha \asymp \kappa$ and (4.7) that there exists a constant $C > 0$ and a positive function m satisfying $\lim_{x\to\infty} m(x) = \infty$ such that:

$$S_2 \leq C2^{-C(j-j_0)m(\kappa)} \tag{4.8}$$

for κ large enough. Combining the results (4.8) and (4.5), we obtain the desired concentration property. □

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:

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

Let us consider the two following functions:

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

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

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

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

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

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

$$p_3 \geq \min\left(2^{-1}, \sqrt{2}\exp(-3e^{-1})\exp(-\chi_1)\right). \tag{4.10}$$

Using Lemma 3.3 and the fact that x^σ is increasing for $\sigma > 0$, one gets:

$$\chi_1 \leq \frac{1}{2}K(\mathbb{P}_{g_1}^n, \mathbb{P}_{g_{-1}}^n) = (\sigma+1)n\gamma_j^2\int_{\frac{1}{2^j}}^{\frac{2N}{2^j}} \psi_{j,N}^2(x)x^\sigma dx \leq Cn\gamma_j^2 2^{-j\sigma}. \tag{4.11}$$

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:

$$\delta_j^{-p} \sup_{f\in B_{\pi,r}^s(L)} \mathbb{E}_f^n(\|\hat{f}-f\|_p^p) \geq c\sqrt{2}\exp(-\chi_1) \geq c$$

where $\delta_j \asymp n^{-v(s,\pi,p,\sigma)}$ and

$$v(s,\pi,p,\sigma) = \frac{s-\frac{1}{\pi}+\frac{1}{p}}{2s+1+\sigma-\frac{2}{\pi}}.$$

The proof of Proposition 3.1 ends by taking σ large enough. □

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 concentred bases. *Test*, 9(2):283–345.

Kerkyacharian, G. and Picard, D. (2005). Regression in random design and warped wavelets. *Bernouilli*, 10(6):1053–1105.

Meyer, Y. (1990). *Ondelettes et Opérateurs*. Hermann, Paris.

Stone, C. (1982). Optimal gobal rates of convergence for nonparametric regression. *Annals of statistics*, 10(4):1040–1053.

Tsybakov, A. (2004). *Introduction à l'estimation nonparametrique*. Springer Verlag.