



**HAL**  
open science

## A maxiset approach of a Gaussian noise model

Christophe Chesneau

► **To cite this version:**

| Christophe Chesneau. A maxiset approach of a Gaussian noise model. 2005. hal-00004331v9

**HAL Id: hal-00004331**

**<https://hal.science/hal-00004331v9>**

Preprint submitted on 7 Jun 2005

**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.

# A maxiset approach of Gaussian white noise models. (nondefinitive version)

Chesneau Christophe

June 7, 2005

Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, 175 rue de Chevaleret, F-75013 Paris, France. christophe.chesneau4@wanadoo.fr

## Abstract

We consider the problem of estimating an unknown function  $f$  in a homoscedastic Gaussian white noise setting under  $\mathbb{L}^p$  risk. The particularity of this model is that it has an intermediate function, say  $v$ , which complicates the estimate significantly. While varying the assumptions on  $v$ , we investigate the minimax rate of convergence over two balls of spaces which belong to family of Besov classes. One is defined as usual and the other called 'weighted Besov balls' used  $v$  explicitly. Adopting the maxiset approach, we develop a natural hard thresholding procedure which attained the minimax rate of convergence within a logarithmic factor over these weighted balls.

**Keywords:** Minimax, Muckenhoupt weights, Maxiset, Gaussian noise, warped wavelets, wavelet thresholding.

**AMS 1991 Subject Classification:** *Primary:* 62G07, *Secondary:* 62G20, 42B20.

## 1 Introduction

Consider the Gaussian white noise model in which we observe processes  $Y_t$  governed by

$$dY_t = H_v(f)(t)dt + \frac{1}{\sqrt{n}}dW_t, \quad n \in \mathbb{N}^*, \quad t \in [0, 1] \quad (1)$$

where the operator  $H_v : \mathcal{B}([0, 1]) \rightarrow \mathbb{L}^2([0, 1])$  is defined by

$$H_v(f)(t) = \frac{f(t)}{v(t)},$$

with

$$\mathcal{B}([0, 1]) = \{f \text{ measurable on } [0, 1], \sup_{x \in [0, 1]} |f(x)| < \infty\}$$

and

$$\mathbb{L}^2([0, 1]) = \{f \text{ measurable on } [0, 1], \|f\|_2^2 = \int_0^1 |f(x)|^2 dx < \infty\}.$$

The function  $v$  is supposed to be known and to satisfy the condition ' $\frac{1}{v}$  belongs to  $\mathbb{L}^2([0, 1])$ '. The process  $W_t$  is a standard Brownian motion on  $[0, 1]$ . The function  $f$  is the unknown function of interest. We want to reconstruct  $f$  from the observations  $\{\int_0^1 h(t)dY_t, h \in \mathbb{L}^2([0, 1])\}$ .

In the simplest case where  $v$  is constant, we observe the well known Gaussian white noise model which has been considered in several papers starting from Ibragimov and Has'minskii (1977). Under certain assumptions on the smoothness of  $f$ , the model (1) becomes an appropriate large sample limit to more general non parametric models such as probability density estimation (see Nussbaum (1996)) or nonparametric regression (see Brown and Low (1996)). Minimax properties in various risk over numerous function spaces can be found in the book of Tsybakov (2004).

In the case where  $v$  is spatially inhomogeneous, the curve estimation is significantly more complicated. For instance, consider the observation of data  $(Y_1, X_1), \dots, (Y_n, X_n)$  where

$$Y_i = f(X_i) + \sigma(X_i)\epsilon_i. \quad (2)$$

where the random variables  $X_i$  are i.i.d, independent of  $\epsilon_i$ , with density  $g$ . The  $\epsilon_i$ 's are normal i.i.d with mean zero and variance one. Brown and Low (1996) have shown that if  $\sigma$  and  $g$  satisfy some condition of boundedness and  $f$  belongs to certain Sobolev classes then the model

$$dZ_t = v(t)dY_t, \quad t \in [0, 1] \quad (3)$$

is asymptotically equivalent (in Le Cam's sense) to (2) under the calibration  $v = \frac{\sigma}{\sqrt{g}}$ . An application of this result can be found in Efromovich and Pinsker (1996). For other equivalences concerning (3), see Grama and Nussbaum (1998).

In this paper, we are focused on the model (1) for general  $v$  and we take the problem in the framework of wavelet analysis. We wish to estimate  $f$  on  $[0, 1]$  by a measurable function on  $[0, 1]$  with respect to the observations  $\{\int_0^1 h(t)dY_t, \quad h \in \mathbb{L}^2([0, 1])\}$  under the  $\mathbb{L}^p$  risk

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

We have denoted  $\mathbb{E}_f$  the expectation with respect the distribution  $\mathbb{P}_f$  of processes  $Y_t$ . Our study can be divided in two parts.

In a first part, we investigate the estimation of  $f$  over usual Besov balls  $B_{s,\pi,r}(L)$ . We show that the minimax properties obtain for the case where  $v$  is bounded from above and below can be extended (without deteriorating the rate of convergence) to more general functions  $v$ . More precisely, we show that if  $\pi \geq p \geq 1$  and  $v$  belongs to  $\mathbb{L}^{\pi'}([0, 1])$  for  $\pi' = \max(\pi, 2)$  then the minimax rate of convergence is of the form

$$n^{-\alpha_1 p} \quad \text{where} \quad \alpha_1 = \frac{s}{1 + 2s}.$$

For other values on the parameters  $(s, \pi, r)$ , we show that if  $v$  is bounded from above then the minimax rate of convergence over  $B_{s,\pi,r}(L)$  is of the form

$$\left( \frac{\ln(n)}{n} \right)^{\alpha_2 p} \quad \text{where} \quad \alpha_2 = \frac{s - \frac{1}{\pi} + \frac{1}{p}}{2(s - \frac{1}{\pi}) + 1}.$$

In the case where  $\pi \geq p > 2$ , it is natural to address the following question: can we obtain the same minimax rate over such spaces for any  $v$  which does not belong to  $\mathbb{L}^\pi([0, 1])$ ? Using an explicit example, we show that the answer is 'No'.

This result motivates us to devote a second part in which we investigate other function spaces more adapted to our model. Our choice will be made on Besov balls constructed on a wavelet basis warped by a factor depending on  $v$ . Such spaces were introduced in analysis by Qui (1982) and were recently developed in statistics by Kerkycharian and Picard (2005). These authors have established good estimation results in a regression setting with random design (i.e (2) with  $\sigma(\cdot) = 1$ ) for very general densities  $g$ . The key of the success of our study rests on the following argument : under certain conditions on the warping factor which refer to Muckenhoupt theory, the warped wavelet bases possess some interesting geometrical properties in the  $\mathbb{L}^p$  norm which allow us to consider function spaces and procedures deeply linked to the model.

Using these analytical tools, we show that if  $\pi \geq p > 1$  and if  $v$  is subject to a property of Muckenhoupt type then the minimax rate over weighted Besov balls  $B_{s,\pi,r}^G(L)$  defined starting from  $G$ , the primitive of  $\frac{1}{v^2}$ , is of the form

$$n^{-\alpha_1} \quad \text{where} \quad \alpha_1 = \frac{s}{1 + 2s}$$

for  $s$  large enough. The hypotheses made on  $v$  are more general than conditions of boundedness or other conditions of integrability depending directly of the parameter  $\pi$ .

Finally, we use this warped wavelet basis to construct a natural procedure which stay as close as possible to the standard hard thresholding algorithm. In order to measure his performance under  $\mathbb{L}^p$  risk, we isolate the associated *maxiset*. This statistical tool developed by Cohen, De Vore, Kerkycharian and Picard (2000) consists in investigating the maximal space (or maxiset) where a procedure has a given rate of convergence. One of the main advantages of this approach is to provide a functional set which is authentically connected to the procedure and the model. Thus, by choosing the rate

$$(\ln(n))^{\alpha_1} n^{-\alpha_1} \quad \text{where} \quad \alpha_1 = \frac{s}{1 + 2s}$$

and considering  $\pi \geq p > 1$ , we prove that our weighted Besov ball  $B_{s,\pi,r}^G(L)$  is included in the maxiset of our procedure. So, we conclude that it is 'near to the optimal' i.e it attains the minimax rate of convergence up to a logarithmic factor.

The paper is organized as follows.

Section 2 defines the basic tools (Muckenhoupt weights, warped wavelet basis ...), inequalities and function spaces we shall need in the study. In Section 3 we investigate the minimax rate over usual Besov balls. Section 4 do the same study but over the weighted Besov balls. Section 5 is devoted to the performance of a natural hard thresholding procedure when the unknown function of interest belongs to these weighted spaces. In Section 6, we describe another statistical model and we explain why we can have results similar to those obtained for (1). Finally, Section 7 is devoted to the proofs of technical lemmas.

## 2 Muckenhoupt condition, warped wavelet bases and function spaces

Throughout this paper, for a weight  $m$  (i.e non negative locally integrable function) on  $[0, 1]$ , we set

$$\mathbb{L}_m^p([0, 1]) = \{f \text{ measurable on } [0, 1] \mid \|f\|_{m,p}^p = \int_0^1 |f(t)|^p m(t) dt < \infty\}$$

where  $\mathbb{L}^p([0, 1]) = \mathbb{L}_1^p([0, 1])$  denotes the usual Lebesgue space i.e

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

### 2.1 Muckenhoupt weight

First recall the notion of Muckenhoupt weight.

**Definition 2.1.** Let  $1 < p < \infty$  and  $q$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ . A weight  $m$  satisfies the  $\mathcal{A}_p$ -condition (or belongs to  $\mathcal{A}_p$ ) iff there exists a constant  $C > 0$  such that for any measurable function  $h$  and any subinterval  $I$  of  $[0, 1]$  we have

$$\left(\frac{1}{|I|} \int_I |h(x)| dx\right) \leq C \left(\frac{1}{m(I)} \int_I |h(x)|^p m(x) dx\right)^{\frac{1}{p}} \quad (4)$$

where  $|I|$  denotes the Lebesgue measure of  $I$  and  $m(I) = \int_I m(x) dx$ .

If  $m$  verifies the  $\mathcal{A}_p$  condition then it is a **Muckenhoupt weight**.

**Example 2.1.** The weight  $m(x) = x^\sigma$  satisfies the  $\mathcal{A}_p$ -condition with  $p > 1$  iff  $-1 < \sigma < p - 1$ .

Let us introduce one of the most interesting property related to this notion.

**Lemma 2.1.** Let  $1 < p < \infty$ . If  $w$  satisfies the  $\mathcal{A}_p$  condition then there exists a constant  $C > 0$  such that for any subintervals  $S \subseteq B \subseteq [0, 1]$  we have

$$w(B) \left(\frac{|S|}{|B|}\right)^p \leq C w(S).$$

*Proof of Lemma 2.1.* It suffices to apply (4) with the function  $h = 1_S$  and the interval  $I = B$ . □

The previous condition has been introduced by Muckenhoupt (1972) and widely used afterwards in the context of Calderón-Zygmund theory. The  $\mathcal{A}_p$ -condition characterizes the boundedness of certain integral operators on  $\mathbb{L}_m^p$  spaces like the Hardy-Littlewood maximal operator or the Hilbert transform. For the complete theory, see the book of Stein (1993).

## 2.2 Warped wavelet bases and Muckenhoupt weights

First we introduce the warped wavelet bases. Second we set some results which will be intensively used in the sequel of this paper.

Let  $N$  be an integer strictly positive. We denote by

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

the **warped wavelet basis** adapted on the interval  $[0, 1]$  constructed starting from

- $\psi$  the wavelet associated with a multiresolution analysis on the line  $V_j = \{\phi_{j,k}, k \in \mathbb{Z}\}$  such that  $\text{Supp}(\phi) = \text{Supp}(\psi) = [-N + 1, N]$  and  $\int \psi(t)t^l dt = 0$  for  $l = 0, \dots, N - 1$ . Let us recall that on the unit interval there exists an integer  $\tau$  satisfying  $2^\tau \geq 2N$  such that one can built at each level  $j \geq \tau$  a wavelet system  $(\phi_{j,k}, \psi_{j,k})$  where

$$\phi_{j,k}(x) = 2^{\frac{j}{2}} \phi(2^j x - k), \quad k = N, N, N + 1, \dots, 2^j - N - 1$$

and

$$\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k), \quad k = N, N, N + 1, \dots, 2^j - N - 1.$$

For each functions, we add  $N$  functions on the neighborhood of 0 which have the support contained in  $[0, (2N - 1)2^{-j}]$  and  $N$  functions on the neighborhood of 1 which have the support contained in  $[1 - (2N - 1)2^{-j}, 1]$ . For simplicity, we denote by " $\tau - 1$ " the integer such that  $\psi_{\tau-1,k} = \phi_{\tau,k}$ .

- $T$  a measurable function on  $[0, 1]$  which are an known, increasing, bijective, absolutely continuous and satisfies

$$T(0) = 0 \quad \text{and} \quad T(1) = 1.$$

We associate to this function the weight

$$w(\cdot) = \frac{1}{\tilde{T}(T^{-1}(\cdot))} \quad (5)$$

where  $\tilde{T}$  denotes the derivative of  $T$  and  $T^{-1}$  its inverse function. Remark that for any measurable positive function  $z$  defined on  $[0, 1]$ ,  $w$  satisfies

$$\int_0^1 z(T(x)) dx = \int_0^1 z(x) w(x) dx.$$

Note that the warped wavelet bases can be viewed as a generalization of the regular wavelet bases. See Meyer (1990) and Daubechies (1992) for wavelet bases on the real line. See Cohen, Daubechies, Jawerth and Vial (1992) for wavelet bases on the interval.

Let  $\infty > p > 1$ . If  $w$  verifies the  $\mathcal{A}_p$  condition then, for any  $\nu \geq \tau$ , any function  $f$  of  $\mathbb{L}^p([0, 1])$  can be decomposed on  $\xi^T$  as

$$f(x) = P_\nu^T(f)(x) + \sum_{j \geq \nu} \sum_{k \in \Delta_j} \beta_{j,k}^T \psi_{j,k}(T(x)),$$

where

$$P_j^T(f)(x) = \sum_{k \in \Delta_j} \alpha_{j,k}^T \phi_{j,k}(T(x)), \quad \alpha_{j,k}^T = \int_0^1 f(T^{-1}(t)) \phi_{j,k}(t) dt$$

and

$$\beta_{j,k}^T = \int_0^1 f(T^{-1}(t)) \psi_{j,k}(t) dt.$$

Let us recall some properties linked to  $\xi^T$ .

**Property 2.1.** *Let  $\nu > 0$ . There exists a constant  $C > 0$  such that*

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

*This inequality is always true if we exchanged  $\phi$  by  $\psi$ .*

**Property 2.2.** If  $w \in \mathcal{A}_p$  then there exist two constant  $c > 0$  and  $C > 0$  such that for  $j \geq \tau$  we have

$$c2^{\frac{j\pi}{2}} \sum_{k \in \Delta_j} |\alpha_{j,k}^T|^p w(I_{j,k}) \leq \|P_j^T(f)\|_p^p \leq C2^{\frac{j\pi}{2}} \sum_{k \in \Delta_j} |\alpha_{j,k}^T|^p w(I_{j,k})$$

where we have set  $I_{j,k} = [\frac{k}{2^j}, \frac{k+1}{2^j}]$ . These inequalities are always true if we exchanged  $\phi$  by  $\psi$ .

### 2.3 Function spaces

Firstly, let us precise that weighted Besov balls  $B_{s,\pi,r}^T(L)$  has three parameters:  $s$  measures the degree of smoothness,  $\pi$  and  $r$  specify the type of norm used to measure the smoothness. Let us define these spaces with precise details below.

For any measurable function  $f$  defined on  $[0, 1]$ , denote the associated  $N$ -th order *modulus of smoothness* as

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

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

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

with the usual modification if  $r = \infty$ . These spaces can be viewed as a generalization of usual Besov balls.

If we are in the case where  $T = I_d$  then we simply denote  $\xi^T = \xi$ ,  $\alpha_{j,k}^T = \alpha_{j,k}$ ,  $\beta_{j,k}^T = \beta_{j,k}$ ,  $P_j^T(f) = P_j(f)$  and  $B_{s,\pi,r}^T(L) = B_{s,\pi,r}(L)$ .

Starting from the previous definition, we can set a list of properties which link the weighted Besov balls with the warped wavelet basis on the unit interval. See below three of these.

**Property 2.3.** If  $w \in \mathcal{A}_\pi$  we have

$$f \in B_{s,\pi,r}^T(L) \implies \left( \sum_{j \geq \tau-1} (2^{j(s+\frac{1}{2}}) \left( \sum_{k \in \Delta_j} |\beta_{j,k}^T|^\pi w(I_{j,k}) \right)^{\frac{1}{\pi}})^r \right)^{\frac{1}{r}} \leq L. \quad (6)$$

for  $\pi \geq 1$ ,  $r \geq 1$  and  $N > s > 0$ .

Moreover, we have the reciprocity for  $s$  large enough.

**Property 2.4.** If  $w \in \mathcal{A}_\pi$  we have

$$f \in B_{s,\pi,r}^T(L) \iff \left( \sum_{j \geq \tau-1} (2^{j(s+\frac{1}{2}}) \left( \sum_{k \in \Delta_j} |\beta_{j,k}^T|^\pi w(I_{j,k}) \right)^{\frac{1}{\pi}})^r \right)^{\frac{1}{r}} \leq L. \quad (7)$$

for  $\pi \geq 1$ ,  $r \geq 1$  and  $N > s \geq q(w)$  where

$$q(w) = \begin{cases} \inf_{v>1} \{w \text{ satisfies the } \mathcal{A}_v \text{ condition}\} & \text{if } w \text{ is not a constant on } [0, 1], \\ 0 & \text{if } w \text{ is constant on } [0, 1]. \end{cases} \quad (8)$$

The following property is similar to Property 2.3 but expressed in term of  $P_j(f)$ .

**Property 2.5.** If  $w \in \mathcal{A}_\pi$  we have

$$f \in B_{s,\pi,r}^T(L) \implies \left( \sum_{j \geq \tau} (2^{js} \|P_j^T(f) - f\|_\pi)^r \right)^{\frac{1}{r}} \leq L \quad (9)$$

These results are always true with the usual modification if  $r = \infty$ . For further details on this subsection, we refer the reader to the article of Kerkycharian and Picard (2005).

In the sequel, the constants  $C, C', C'', c, c', c''$  represent any constants we shall need, and can differ from one line to one other.

### 3 Minimax study over Besov balls

Let us recall that we observe the model (1) under the two following assumptions on  $v$  and  $f$ :

- $\frac{1}{v}$  belongs to  $\mathbb{L}^2([0, 1])$ ,
- $\|f\|_\infty < \infty$ .

Let us set the following notations:

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

We shall exhibit the minimax rate of convergence over usual Besov balls for several values of  $\epsilon$ . The first part of this section is devoted to the proofs of the two following theorems.

**Theorem 3.1.** *Let  $\infty > p \geq 1$  and  $\infty \geq \pi \geq p$ . Assume that  $v$  satisfies the following condition:*

$$v \in \mathbb{L}^{\pi'}([0, 1]) \tag{10}$$

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

$$\inf_{\hat{f}} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \asymp n^{-\alpha_1 p}.$$

**Theorem 3.2.** *Let  $\infty > p > 1$ . Assume that  $v$  satisfies the following condition:*

$$\|v\|_\infty < \infty. \tag{11}$$

Then for  $N > s > \frac{1}{\pi}$ ,  $\infty \geq \pi \geq 1$ ,  $\infty \geq r \geq 1$  and  $\epsilon < 0$  we have

$$\inf_{\hat{f}} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \asymp \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p}$$

For  $\epsilon = 0$ , there exist  $C > 0$  and  $c > 0$  satisfying

$$c \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} \leq \inf_{\hat{f}} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} (\ln(n))^{\left(\frac{p}{2} - \frac{\pi}{r}\right)_+}.$$

**Remark 3.1.** *For the two lower bounds, only the condition  $\frac{1}{v} \in \mathbb{L}^2([0, 1])$  is determinant.*

#### 3.1 Proof of Theorem 3.1: upper bound and lower bound

##### 3.1.1 Upper bound

Here, we use the standard method which consists in representing the unknown function  $f$  on a regular wavelet basis and in studying the upper bound attained by the associated linear wavelet procedure.

**Theorem 3.3.** *Let  $\infty > p \geq 1$  and  $\pi \geq p$ . Assume that the condition (10) holds. Consider  $\hat{f}^l$  the linear estimator defined by*

$$\hat{f}^l(x) = \sum_{k \in \Delta_{j_0}} \hat{\alpha}_{j_0, k} \phi_{j_0, k}(x) \tag{12}$$

where

$$\hat{\alpha}_{j_0, k} = \int_0^1 \phi_{j_0, k}(t) v(t) dY_t.$$

Then for  $N > s > 0$  and  $\infty \geq r \geq 1$  there exists a constant  $C > 0$  such that

$$\sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f}^l - f\|_p^p) \leq C n^{-\alpha_1 p}$$

for  $j_0$  the integer satisfying  $2^{j_0} \simeq n^{\frac{1}{1+2s}}$ .

*Proof of Theorem 3.3.* Using Hölder's inequality, for  $\pi \geq p$  we have

$$\mathbb{E}_f(\|\hat{f}^l - f\|_p^p) \leq C(\mathbb{E}_f(\|\hat{f}^l - f\|_\pi^\pi))^{\frac{p}{\pi}}. \quad (13)$$

Using Minkowski's inequality and the elementary inequality

$$(|x + y|)^\pi \leq 2^{\pi-1}(|x|^\pi + |y|^\pi), \quad x, y \in \mathbb{R},$$

the  $\mathbb{L}^\pi$  risk of  $\hat{f}$  can be decomposed as follows:

$$\begin{aligned} \mathbb{E}_f(\|\hat{f}^l - f\|_\pi^\pi) &\leq C(\mathbb{E}_f(\|\hat{f}^l - P_{j_0}(f)\|_\pi^\pi) + \|P_{j_0}(f) - f\|_\pi^\pi) \\ &= C(S_1 + S_2). \end{aligned} \quad (14)$$

Since  $f \in B_{s,\pi,r}(L) \subset B_{s,\pi,\infty}(L)$ , Property 2.5 gives us

$$S_2 \leq L2^{-j_0 s \pi}. \quad (15)$$

Using the definition of  $\hat{f}^l$  and Property 2.2, one gets

$$\begin{aligned} S_1 &= \mathbb{E}_f(\|\sum_{k \in \Delta_{j_0}} (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}) \phi_{j_0,k}(\cdot)\|_\pi^\pi) \\ &\leq C(2^{j_0(\frac{\pi}{2}-1)} \sum_{k \in \Delta_{j_0}} \mathbb{E}_f(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^\pi)) \\ &= C2^{j_0(\frac{\pi}{2}-1)} S_1^*. \end{aligned} \quad (16)$$

Let us consider  $\rho_{j,k}$  defined by

$$\rho_{j,k} = \sqrt{\int_0^1 v^2(t) \phi_{j,k}^2(t) dt}. \quad (17)$$

We have clearly

$$\hat{\alpha}_{j_0,k} - \alpha_{j_0,k} = \frac{1}{\sqrt{n}} \int_0^1 v(t) \phi_{j_0,k}(t) dW_t \sim \rho_{j_0,k} \epsilon_n$$

where  $\epsilon_n$  is random variable such that

$$\epsilon_n \sim \mathcal{N}(0, \frac{1}{n}).$$

To study  $S_1^*$ , we only need the second point of the following lemma which will be proved in Appendix. (The first point will be used later in the study.)

**Lemma 3.1.** *Let  $n \in \mathbb{N}^*$ . If  $V_n \sim \mathcal{N}(0, \frac{1}{n})$  then for  $\kappa \geq 2\sqrt{2\pi}$  there exists a constant  $C > 0$  only depending on  $p$  such that*

- $\mathbb{P}_f(|V_n| \geq \frac{\kappa}{2} \sqrt{\frac{\ln(n)}{n}}) \leq Cn^{-\frac{\pi}{2}},$
- $\mathbb{E}_f(|V_n|^\pi) \leq Cn^{-\frac{\pi}{2}}.$

Thus, one gets

$$S_1^* \leq Cn^{-\frac{\pi}{2}} \sum_{k \in \Delta_{j_0}} \rho_{j_0,k}^\pi. \quad (18)$$

First consider the case where  $2 > \pi \geq 1$ . Hölder's inequality, Property 2.1 and condition (10) yield

$$\begin{aligned} \sum_{k \in \Delta_{j_0}} \rho_{j_0,k}^\pi &\leq \left( \sum_{k \in \Delta_{j_0}} \int_0^1 v^2(t) \phi_{j_0,k}^2(t) dt \right)^{\frac{\pi}{2}} \left( \sum_{k \in \Delta_{j_0,k}} 1 \right)^{1-\frac{\pi}{2}} \\ &\leq C(2^{j_0} \int_0^1 v^2(t) dt)^{\frac{\pi}{2}} 2^{j_0(1-\frac{\pi}{2})} \\ &\leq C' 2^{j_0}. \end{aligned} \quad (19)$$



Second investigate the case where  $\infty \geq \pi \geq 2$ . Applying Hölder's inequality with the measure  $d\nu = \phi_{j_0,k}^2(t)dt$ , using Property 2.1 and condition (10), one gets

$$\begin{aligned} \sum_{k \in \Delta_{j_0}} \rho_{j_0,k}^\pi &\leq \int_0^1 v^\pi(t) \sum_{k \in \Delta_{j_0}} \phi_{j_0,k}^2(t) dt \\ &\leq C 2^{j_0} \|v\|_\pi^\pi \\ &= C' 2^{j_0}. \end{aligned} \quad (20)$$

Thus, considering (16), (19), (18) and (20) we obtain for  $\pi \geq 1$ ,

$$S_1 \leq C 2^{\frac{j_0 \pi}{2}} n^{-\frac{\pi}{2}}. \quad (21)$$

Taking in account that  $2^{j_0} \simeq n^{\frac{1}{1+2s}}$ , the inequalities (14), (15) and (21) imply that

$$\begin{aligned} \mathbb{E}_f(\|\hat{f}^l - f\|_\pi^\pi) &\leq C(2^{\frac{j_0 \pi}{2}} n^{-\frac{\pi}{2}} + 2^{-j_0 s \pi}) \\ &\leq C' n^{-\alpha_1 \pi}. \end{aligned}$$

Considering (13), we deduce that for  $\pi \geq p \geq 1$ ,

$$\sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f}^l - f\|_p^p) \leq C n^{-\alpha_1 p}.$$

This completes the proof of Theorem 3.3.  $\square$

### 3.1.2 Lower Bound

Now, introduce a key theorem which will be intensively used to exhibit the lower bounds over general Besov balls.

First of all, remark that for any  $j$  and  $N$  of  $\mathbb{N}^*$ , there exists a constant  $2N - 2 \geq C_* \geq 2$  such that

$$r_* = \frac{2^j - C_*}{2N - 1} \in \mathbb{N}^* \quad (22)$$

**Theorem 3.4.** *Let  $j$  an integer depending on  $n$ ,  $(\omega_{j,k})_{k \in R_j}$  a fixed sequence and  $\varepsilon$  a sequence such that  $\varepsilon = (\varepsilon_k)_{k \in R_j} \in \{-1, 1\}^{r_*}$  where*

$$R_j = \{2^j - l(2N - 1) + N - 2; l = 1, 2, \dots, r_*\}. \quad (23)$$

and  $r_*$  is defined by (22). Let us set the functions

$$g_\varepsilon(x) = \gamma_j \sum_{k \in R_j} \omega_{j,k} \varepsilon_k \psi_{j,k}(T(x)) \quad (24)$$

where  $\gamma_j$  is chosen in such a way that  $g_\varepsilon$  belongs to  $B_{s,\pi,r}^T(L)$ . For such  $\varepsilon$ , if we consider  $\varepsilon_k^* = (\varepsilon'_i)_{i \in R_j}$  defined by

$$\varepsilon'_i = \varepsilon_i 1_{\{i \neq k\}} - \varepsilon_i 1_{\{i=k\}},$$

then for any estimator  $\hat{f}$  we have

$$\begin{aligned} U_j &= \sup_{g_\varepsilon \in B_{s,\pi,r}^T(L)} \mathbb{E}_{g_\varepsilon}(\|\hat{f} - g_\varepsilon\|_p^p) \\ &\geq \frac{e^{-\lambda}}{2} \gamma_j^p \sum_{k \in R_j} \omega_{j,k}^p \inf_{\varepsilon_i \in \{-1, +1\}} \mathbb{P}_{g_\varepsilon}(\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}) \|\psi_{j,k}\|_{w,p}^p \end{aligned}$$

where  $\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon)$  denotes the likelihood ratio between the laws induced by  $g_{\varepsilon_k^*}$  and  $g_\varepsilon$  defined by

$$\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon) = \frac{d\mathbb{P}_{g_{\varepsilon_k^*}}}{d\mathbb{P}_{g_\varepsilon}}. \quad (25)$$

*Proof of Theorem 3.4.* Since  $T$  is increasing with  $T(0) = 0$  and  $T(1) = 1$ , for all  $k$  belonging to  $R_j$  we have the  $S_{j,k}^T$ 's defined by

$$S_{j,k}^T = \text{Supp}(\psi_{j,k}(T(\cdot))) = [T^{-1}(\frac{k-N+1}{2j}), T^{-1}(\frac{k+N}{2j})]$$

which satisfy

$$S_{j,k}^T \cap S_{j,k'}^T = \emptyset \text{ for } k \neq k', \quad k, k' \in R_j$$

and

$$\bigcup_{k \in R_j} S_{j,k}^T = [T^{-1}(\frac{C_* - 1}{2j}), T^{-1}(1 - \frac{1}{2j})] \subset [0, 1].$$

In the case where  $T = Id$ , we adopt the notation  $S_{j,k}$  instead of  $S_{j,k}^T$ . Denote by  $\mathcal{G}$  the set of all  $g_\varepsilon$  defined by (24). For any estimator  $\hat{f}$ , let

$$W_{j,k}^1 = \int_{S_{j,k}^T} |\hat{f}(x) - \gamma_j \varepsilon_k \omega_{j,k} \psi_{j,k}(T(x))|^p dx$$

and

$$W_{j,k}^2 = \int_{S_{j,k}^T} |\hat{f}(x) + \gamma_j \varepsilon_k \omega_{j,k} \psi_{j,k}(T(x))|^p dx.$$

Using the fact that the  $S_{j,k}^T$  are disjoint, for any positive sequence  $(\delta_{j,k})_{k \in R_j}$ , we have

$$\begin{aligned} U_j &\geq \frac{1}{\text{card}(\mathcal{G})} \sum_{\varepsilon} \mathbb{E}_{g_\varepsilon} (\|\hat{f} - g_\varepsilon\|_p^p) \\ &\geq \frac{1}{\text{card}(\mathcal{G})} \sum_{k \in R_j} \sum_{\varepsilon} \mathbb{E}_{g_\varepsilon} (\int_{S_{j,k}^T} |\hat{f}(x) - \gamma_j \varepsilon_k \omega_{j,k} \psi_{j,k}(T(x))|^p dx) \end{aligned} \quad (26)$$

By the definition of  $\varepsilon_k^*$  and the fact that for all  $k \in R_j$

$$\text{Card}(\mathcal{G}) = 2 \text{Card}(\varepsilon, \varepsilon_i \in \{-1, +1\}, \quad i \neq k, i, k \in R_j),$$

we obtain

$$\begin{aligned} U_j &\geq \frac{1}{\text{card}(\mathcal{G})} \sum_{k \in R_j} \sum_{\varepsilon_i \in \{-1, +1\}} \mathbb{E}_{g_\varepsilon} (W_{j,k}^1 + \wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) W_{j,k}^2) \\ &\geq \frac{1}{2} \sum_{k \in R_j} \inf_{\varepsilon_i \in \{-1, +1\}} \delta_{j,k}^p \mathbb{E}_{g_\varepsilon} (1_{\{W_{j,k}^1 \geq \delta_{j,k}^p\}} + e^{-\lambda} 1_{\{\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}\}} 1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}}) \\ &\geq \frac{e^{-\lambda}}{2} \sum_{k \in R_j} \inf_{\varepsilon_i \in \{-1, +1\}} \delta_{j,k}^p \mathbb{E}_{g_\varepsilon} (1_{\{\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}\}} (1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}} + 1_{\{W_{j,k}^1 \geq \delta_{j,k}^p\}})). \end{aligned} \quad (27)$$

Now, consider the sequence  $\delta_{j,k}$  defined by

$$\delta_{j,k} = \gamma_j \omega_{j,k} \|\psi_{j,k}\|_{w,p}.$$

Using Minkowski's inequality and the change of variable  $y = T(x)$ , we see that

$$\begin{aligned} (W_{j,k}^1)^{\frac{1}{p}} + (W_{j,k}^2)^{\frac{1}{p}} &\geq 2\gamma_j \omega_{j,k} \|\psi_{j,k}(T(\cdot))\|_p \\ &= 2\delta_{j,k}. \end{aligned}$$

Therefore

$$1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}} \geq 1_{\{W_{j,k}^1 \leq \delta_{j,k}^p\}}. \quad (28)$$

Putting (26), (27) and (28) together, we deduce that

$$U_j \geq \frac{e^{-\lambda}}{2} \gamma_j^p \sum_{k \in R_j} \omega_{j,k}^p \inf_{\varepsilon_i \in \{-1, +1\}} \mathbb{P}_{g_\varepsilon} (\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}) \|\psi_{j,k}\|_{w,p}^p.$$

□

Theorem 3.4 can be viewed as a generalization of Lemma 10.2 which appeared in the book Härdle, Kerkyacharian, Picard and Tsybakov (1998).

**Theorem 3.5.** *Let  $\infty > p \geq 1$ . There exists a constant  $c > 0$  such that for  $N > s > 0$ ,  $\infty \geq \pi \geq 1$  and  $\infty \geq r \geq 1$  we have*

$$\inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \geq cn^{-\alpha_1 p}.$$

*Proof of Theorem 3.5.* Let  $j$  an integer to be chosen below. Consider the functions  $g_\epsilon$  defined by (24) with

$$\omega_{j,k} = 1 \quad \text{and} \quad T = I_d.$$

Since the wavelet coefficients of  $g_\epsilon$  (denoted below by  $\beta_{j,k}^*$ ) are equal to  $\gamma_j \epsilon_k$  and  $\text{Card}(R_j) = r_* \leq C2^j$ , we have

$$\begin{aligned} 2^{j(s+\frac{1}{2})} \left( \sum_{k \in R_j} |\beta_{j,k}^*|^\pi 2^{-j} \right)^{\frac{1}{\pi}} &= \gamma_j 2^{j(s+\frac{1}{2})} \left( \sum_{k \in R_j} 2^{-j} \right)^{\frac{1}{\pi}} \\ &\leq C \gamma_j 2^{j(s+\frac{1}{2})}. \end{aligned}$$

Using Property 2.4 with  $q(w) = 0$  and taking  $j$  large enough, only the following constraint on  $\gamma_j$  is necessary to guarantee that  $g_\epsilon \in B_{s,\pi,r}(L)$ :

$$\gamma_j \leq C' 2^{-j(s+\frac{1}{2})}$$

where  $C'$  denotes a constant suitably chosen. Now, consider the following lemma which will be proved in Appendix.

**Lemma 3.2.** *If we chose  $\gamma_j = n^{-\frac{1}{2}}$  then there exist  $\lambda > 0$  and  $p_0 > 0$  not depending on  $n$  such that*

$$\sum_{k \in R_j} \inf_{\substack{\epsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{P}_{g_\epsilon}(\wedge_n(g_{\epsilon_k}^*, g_\epsilon) > e^{-\lambda}) \geq p_0 2^j.$$

It follows from Lemma 3.2 and Theorem 3.4 that:

$$\begin{aligned} \sup_{g_\epsilon \in B_{s,\pi,r}(L)} \mathbb{E}_{g_\epsilon}(\|\hat{f} - g_\epsilon\|_p^p) &\geq \frac{e^{-\lambda}}{2} \gamma_j^p \sum_{k \in R_j} \inf_{\substack{\epsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{P}_{g_\epsilon}(\wedge_n(g_{\epsilon_k}^*, g_\epsilon) > e^{-\lambda}) \|\psi_{j,k}\|_p^p \\ &\geq \frac{e^{-\lambda}}{2} 2^{\frac{jp}{2}} \|\psi\|_p^p \left(\frac{1}{\sqrt{n}}\right)^p p_0. \end{aligned}$$

Choosing  $j$  such that  $\gamma_j = n^{-\frac{1}{2}} \simeq 2^{-j(s+\frac{1}{2})}$  (i.e  $2^j \simeq n^{\frac{1}{1+2s}}$ ), one gets

$$\begin{aligned} \inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) &\geq \frac{e^{-\lambda}}{2} \|\psi\|_p^p \left(\frac{2^{\frac{j}{2}}}{\sqrt{n}}\right)^p p_0 \\ &\geq c'' n^{-\alpha_1 p}. \end{aligned}$$

This ends the proof of Theorem 3.5. □

Combining Theorem 3.3 and Theorem 3.5, we obtain Theorem 3.1.

**Remark 3.2.** *If  $v$  is a positive constant then we obtain the usual minimax result.*

**Example 3.1.** *Let  $\infty > \pi \geq p \geq 1$ . Consider the model (1) with the operator  $H_{v_1}$  where*

$$v_1(t) = t^{-\frac{\sigma}{2}} \quad \text{for} \quad -1 < \sigma < \frac{2}{\pi^t}.$$

*It is clear that the condition (10) holds. So we can apply Theorem 3.1.*

**Example 3.2.** *Let  $\infty > \pi \geq p \geq 1$ . Consider the model (1) with the operator  $H_{v_2}$  where*

$$v_2(t) = (1-t)^\alpha t^{-\beta} \quad \text{for} \quad 0 < \alpha < \frac{1}{2} \quad \text{and} \quad 0 < \beta < \frac{1}{\pi^t}.$$

*Remark that  $v_2$  is not bounded from above and below and that the condition (10) holds. So we can apply Theorem 3.1.*

The following subsection proposes to investigate the minimax rate over  $B_{s,\pi,r}(L)$  under  $\mathbb{L}^p$  loss for other values of the parameters  $(s, \pi, r)$  and other assumptions on  $v$ .

## 3.2 Proof of Theorem 3.2: upper bounds and lower bound

### 3.2.1 Upper bounds

Let us consider the following threshold wavelet procedure:

$$\hat{f}^{\otimes}(x) = \sum_{k \in \Delta_{j_1}} \hat{\alpha}_{j_1,k} \phi_{j_1,k}(x) + \sum_{j \in \Lambda} \sum_{k \in \Delta_j} \hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}| \geq \kappa \sqrt{\frac{j}{n}}\}} \psi_{j,k}(x) \quad (29)$$

where we have set

$$\hat{\alpha}_{j,k} = \int_0^1 \phi_{j,k}(t) v(t) dY_t, \quad \hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(t) v(t) dY_t,$$

$\Lambda = \{j \geq \tau - 1; j_1 < j < j_2\}$  for  $j_1$  and  $j_2$  the integers verifying

$$2^{j_1} \approx (n(\ln(n))^{\frac{p-\pi}{\pi}} 1_{\{\epsilon \geq 0\}})^{1-2\alpha} \quad \text{and} \quad 2^{j_2} \approx (n(\ln(n))^{-1} 1_{\{\epsilon \leq 0\}})^{\frac{\alpha}{(s-\frac{1}{\pi}+\frac{1}{p})}} \quad (30)$$

where

$$\alpha = \max(\alpha_1, \alpha_2).$$

Following step by step the proof of Theorem 3 which appeared in Donoho, Johnstone, Kerkyacharian and Picard (1996), one can show that the estimator (29) attains the upper bound describe in Theorem 3.2. This is an immediate consequence of the following lemma:

**Lemma 3.3.** *Assume that the condition (11) holds. Then for any  $\kappa$  there exist  $C > 0$  and  $\kappa'$  satisfying*

- $\mathbb{E}_f(|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}) \leq Cn^{-p}$
- $\mathbb{P}_f(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa'}{2} \sqrt{\frac{j}{n}}) \leq C2^{-\kappa j}$ .

The proof is rejected in Appendix.

### 3.2.2 Lower bound

**Theorem 3.6.** *Let  $\infty > p \geq 1$ . There exists a constant  $c > 0$  such that for  $N > s > \frac{1}{\pi}$ ,  $0 \geq \epsilon$  and  $\infty \geq r \geq 1$  we have*

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

*Proof of Theorem 3.6.* Consider the following family

$$\{g_0 = 0; g_k = \gamma_j \psi_{j,k}, k \in R_j\}$$

where  $R_j$  is defined by (23). In order to prove Theorem 3.6, let us introduce a theorem which can be view as an adapted version of Lemma 10.1 of Härdle, Kerkyacharian, Picard and Tsybakov (1998).

**Theorem 3.7.** *Assume the following conditions are fulfilled:*

- $\forall k \in R_j, \gamma_j$  is chosen such that  $g_k \in B_{s,\pi,r}(L)$ ,
- There exists a constant  $p_0 > 0$  satisfying

$$\sum_{k \in R_j} \mathbb{P}_{g_k}(\Lambda(g_0, g_k) \geq 2^{-\lambda^* j}) \geq p_0 2^j \quad (31)$$

for a fixed  $\lambda^*$  such that  $1 \geq \lambda^* > 0$ .

Then for any estimator  $\hat{f}$  we have

$$\sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \geq 2^{-p} 2^{j(\frac{p}{2}-1)} \gamma_j^p \|\psi\|_p^p \frac{p_0}{2}.$$

*Proof of Theorem 3.7.* For sake of simplicity, let us denote by  $d$  the  $\mathbb{L}^p$  metric, i.e. for any  $f$  and  $g$  which belong to  $\mathbb{L}^p([0, 1])$

$$d(f, g) = \left( \int_0^1 |f(t) - g(t)|^p dt \right)^{\frac{1}{p}}.$$

Put

$$\delta_j = \frac{\gamma_j}{2} 2^{j(\frac{1}{2} - \frac{1}{p})} \|\psi\|_p. \quad (32)$$

From Chebychev's inequality, we see that

$$\begin{aligned} \delta_j^{-p} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) &\geq \delta_j^{-p} \sup_{k \in R_j \cup \{0\}} \mathbb{E}_{g_k}(\|\hat{f} - g_k\|_p^p) \\ &\geq \sup_{k \in R_j \cup \{0\}} \mathbb{P}_{g_k}(\|\hat{f} - g_k\|_p \geq \delta_j) \\ &\geq \max\left(\frac{1}{r_*} \sum_{k \in R_j} \mathbb{P}_{g_k}(d(\hat{f}, g_k) \geq \delta_j), \mathbb{P}_{g_0}(d(\hat{f}, g_0) \geq \delta_j)\right). \end{aligned} \quad (33)$$

Since  $r_* \leq 2^j$  it suffices to prove that

$$\max(2^{-j} \sum_{k \in R_j} \mathbb{P}_{g_k}(d(\hat{f}, g_k) \geq \delta_j), \mathbb{P}_{g_0}(d(\hat{f}, g_0) \geq \delta_j)) \geq \frac{p_0}{2}. \quad (34)$$

Assume on the contrary that (34) is false. Then there exists an estimator, say  $f^*$ , such that

$$\max(2^{-j} \sum_{k \in R_j} \mathbb{P}_{g_k}(d(f^*, g_k) \geq \delta_j), \mathbb{P}_{g_0}(d(f^*, g_0) \geq \delta_j)) < \frac{p_0}{2}.$$

In particular, we have

$$\mathbb{P}_{g_0}(d(f^*, g_0) \geq \delta_j) < \frac{p_0}{2} \quad (35)$$

and since there exists  $c > 0$  such that  $r_* \geq c2^j$  (for instance,  $c = \frac{1}{N(2N-1)}$ ), we have

$$2^{-j} \sum_{k \in R_j} \mathbb{P}_{g_k}(d(f^*, g_k) < \delta_j) > c - \frac{p_0}{2}. \quad (36)$$

Putting (31) and (36) together, we obtain that for any  $k \in R_j$

$$\begin{aligned} \sum_{k \in R_j} \mathbb{P}_{g_k}(\{d(f^*, g_k) < \delta_j\} \cap \{\Lambda(g_0, g_k) \geq 2^{-\lambda^* j}\}) &\geq \sum_{k \in R_j} \mathbb{P}_{g_k}(d(f^*, g_k) < \delta_j) \\ &\quad + \sum_{k \in R_j} \mathbb{P}_{g_k}(\Lambda(g_0, g_k) \geq 2^{-\lambda^* j}) - (2N)^{-1} 2^j \\ &> (c - \frac{p_0}{2}) 2^j + p_0 2^j - c 2^j \\ &> \frac{p_0}{2} 2^j. \end{aligned} \quad (37)$$

We now use the  $\delta_j$  defined in (32). First for all  $k \in R_j$

$$d(g_k, g_0) = \gamma_j \|\psi_{j,k}\|_p = 2\delta_j$$

and the triangular inequality implies that

$$\bigcup_{k \in R_j} \{d(f^*, g_k) < \delta_j\} \subset \{d(f^*, g_0) \geq \delta_j\}.$$

Second for all  $k \neq k' \in R_j$  we have

$$\begin{aligned} d(g_k, g_{k'}) &= \gamma_j (\|\psi_{j,k}\|_p + \|\psi_{j,k'}\|_p) \\ &\geq \gamma_j \|\psi_{j,k}\|_p \\ &= 2\delta_j. \end{aligned}$$

Consequently the events  $\{d(f^*, g_k) < \delta_j\}$  are disjoint for  $k \neq k' \in R_j$ . It follows from (37) that

$$\begin{aligned}
\mathbb{P}_{g_0}(d(f^*, g_0) \geq \delta_j) &\geq \mathbb{P}_{g_0}\left(\bigcup_{k \in R_j} d(f^*, g_k) < \delta_j\right) \\
&= \sum_{k \in R_j} \mathbb{P}_{g_0}(d(f^*, g_k) < \delta_j) \\
&= \sum_{k \in R_j} \mathbb{E}_{g_k}(\Lambda(g_0, g_k) \mathbf{1}_{\{d(f^*, g_k) < \delta_j\}}) \\
&\geq 2^{-\lambda^* j} \sum_{k \in R_j} \mathbb{E}_{g_k}(\mathbf{1}_{\{d(f^*, g_k) < \delta_j\}} \mathbf{1}_{\{\Lambda(g_0, g_k) \geq 2^{-\lambda^* j}\}}) \\
&= 2^{-\lambda^* j} \sum_{k \in R_j} \mathbb{P}_{g_k}(\{d(f^*, g_k) < \delta_j\} \cap \{\Lambda(g_0, g_k) \geq 2^{-\lambda^* j}\}) \\
&> \frac{p_0}{2} 2^{(1-\lambda^*)j}.
\end{aligned}$$

Then we contradict (35). So, combining (33) and (34), we deduce that for any estimator  $\hat{f}$  we have

$$\sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \geq 2^{-p} 2^{j(\frac{p}{2}-1)} \gamma_j^p \|\psi\|_p^p \frac{p_0}{2}. \quad (38)$$

□

Since the wavelet coefficients of  $g_k$ , denoted below by  $\beta_{j,k}^*$ , are equal to  $\gamma_j$  with  $k$  fixed, we have

$$\begin{aligned}
2^{j(s+\frac{1}{2})} (|\beta_{j,k}^*|^{\pi} 2^{-j})^{\frac{1}{\pi}} &= \gamma_j 2^{j(s+\frac{1}{2})} (2^{-j})^{\frac{1}{\pi}} \\
&= \gamma_j 2^{j(s+\frac{1}{2}-\frac{1}{\pi})}.
\end{aligned}$$

So taking

$$\gamma_j \leq C' 2^{-j(s+\frac{1}{2}-\frac{1}{\pi})}$$

where  $C'$  denotes a constant suitably chosen, Property 2.4 with  $q(w) = 0$  implies that the  $g_k$  belong to  $B_{s, \pi, r}(L)$ . Now consider the following lemma:

**Lemma 3.4.** *Let  $\gamma_j = c_0 \sqrt{\frac{\ln(n)}{n}}$ . If there exists a constant  $c > 0$  such that for  $n$  large enough*

$$\ln(2^j) \geq c \ln(n) \quad (39)$$

then for a fixed  $1 \geq \lambda^* > 0$  and a  $c_0$  small enough there exists a constant  $p_0 > 0$  satisfying

$$\sum_{k \in R_j} \mathbb{P}_{g_k}(\Lambda(g_0, g_k) \geq 2^{-\lambda^* j}) \geq p_0 2^j. \quad (40)$$

Thus, choosing

$$\gamma_j = c_0 \sqrt{\frac{\ln(n)}{n}}, \quad \text{i.e. } 2^j \simeq \left(\sqrt{\frac{n}{\ln(n)}}\right)^{\frac{1}{s+\frac{1}{2}-\frac{1}{\pi}}}$$

and remarking that for  $n$  large enough we have

$$\begin{aligned}
\ln(2^j) &\geq \frac{1}{2(s+\frac{1}{2}-\frac{1}{\pi})} (\ln(n) - \ln(\ln(n))) + \ln(c) \\
&\geq \frac{1}{4(s+\frac{1}{2}-\frac{1}{\pi})} \ln(n) \\
&\geq \frac{1}{4N+2} \ln(n),
\end{aligned}$$

the condition (39) and a fortiori, the condition (31) are satisfied. So Theorem 3.7 implies that

$$\begin{aligned}
\sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) &\geq 2^{-p} 2^{j(\frac{p}{2}-1)} \gamma_j^p \|\psi\|_p^p \frac{p_0}{2} \\
&\geq c \left(\frac{\ln(n)}{n}\right)^{\frac{p}{2}} \left(\sqrt{\frac{n}{\ln(n)}}\right)^{\frac{p}{2} \frac{p-1}{s+\frac{1}{2}-\frac{1}{\pi}}} \\
&= c \left(\frac{\ln(n)}{n}\right)^{\alpha_{2p}}.
\end{aligned}$$

This ends the proof of Theorem 3.6. □

Putting Subsection 'Upper bounds' and Theorem 3.6 together, we establish Theorem 3.2.

### 3.3 When $v$ does not belong to $\mathbb{L}^{\pi'}([0, 1])$

The last part of this section is center around the following question:

**Question 3.1.** *Let  $\pi \geq p > 2$ . Can we extend the minimax properties without deteriorating the rate of convergence over  $B_{s,\pi,r}(L)$  for other  $v$  i.e functions which does not satisfy the assumption 'v belongs to  $\mathbb{L}^{\pi}([0, 1])$ ' (see Theorem 3.1)?*

The following theorem gives a beginning of answer by considering a function  $v$  which does not belong to  $\mathbb{L}^p([0, 1])$  and, a fortiori, to  $\mathbb{L}^{\pi}([0, 1])$ .

**Theorem 3.8.** *Let  $\infty > p > 2$  and  $\infty \geq \pi \geq p$ . Assume that we observe model (1) with the operator  $H_{v_*}$  where*

$$v_*(t) = t^{-\frac{\sigma}{2}} \quad \text{for } \frac{2}{p} < \sigma < 1. \quad (41)$$

Then for  $N > s > 0$  and  $\infty \geq r \geq 1$  there exist two constant  $C > 0$  and  $c > 0$  such that

$$cn^{-\tilde{\alpha}p} \leq \inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \leq Cn^{-\alpha'p}$$

where

$$\alpha' = \frac{s}{2s + 1 + \sigma - \frac{2}{\pi}}$$

and

$$\tilde{\alpha} = \begin{cases} \alpha_1 & \frac{\pi-p}{\pi(\sigma p-2)} > s > 0, \quad \pi > p, \\ \frac{s+\frac{1}{p}-\frac{1}{\pi}}{2s+1+\sigma-\frac{2}{\pi}} & N > s \geq \frac{\pi-p}{\pi(\sigma p-2)}, \quad \pi \geq p. \end{cases}$$

*Proof of Theorem 3.8.* First, introduce the following lemma which will be proved in Appendix.

**Lemma 3.5.** *Let  $\pi > 2$ . Let us consider  $\rho_{j,k}$  defined by (17) with  $v = v_*$  (see (41)) and  $\eta_{j,k}$  defined by*

$$\eta_{j,k} = \sqrt{\int_0^1 \frac{1}{v_*^2(t)} \psi_{j,k}^2(t) dt}. \quad (42)$$

Then there exist two constant  $C > 0$  and  $c > 0$  such that

$$c2^{\frac{j\sigma\pi}{2}} \leq \sum_{k \in R_j} \eta_{j,k}^{-\pi} \leq \sum_{k \in \Delta_j} \rho_{j,k}^{\pi} \leq C2^{\frac{j\sigma\pi}{2}}. \quad (43)$$

#### 3.3.1 Upper bound

Let us consider the linear estimator  $\hat{f}^l$  defined in (12) where  $v$  is defined by (41). Putting the inequalities (18) and (43) together, one gets

$$\begin{aligned} \sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f}^l - f\|_p^p) &\leq C(2^{j_0(\frac{\sigma}{2}-1)}n^{-\frac{\sigma}{2}} \sum_{k \in \Delta_{j_0}} \rho_{j,k}^{\pi} + 2^{-j_0s\pi})^{\frac{p}{\pi}} \\ &\leq C'(2^{j_0(\frac{\sigma}{2}-1+\frac{\sigma\pi}{2})}n^{-\frac{\sigma}{2}} + 2^{-j_0s\pi})^{\frac{p}{\pi}} \\ &\leq C''n^{-\alpha'p} \end{aligned}$$

for  $j_0$  the integer satisfying  $2^{j_0} \simeq n^{\frac{1}{1+2s+\sigma-\frac{2}{\pi}}}$ .

### 3.3.2 Lower bound

Consider the functions  $g_\epsilon$  introduced in (24) with

$$\omega_{j,k} = \eta_{j,k}^{-1} \quad \text{and} \quad T = I_d$$

where  $\eta_{j,k}$  is defined by (42). Since the wavelet coefficients of  $g_\epsilon$ , denoted below by  $\beta_{j,k}^*$ , are equal to  $\eta_{j,k}^{-1} \gamma_j \epsilon_k$ , the inequality (43) gives us

$$\begin{aligned} 2^{j(s+\frac{1}{2})} \left( \sum_{k \in R_j} |\beta_{j,k}^*|^\pi 2^{-j} \right)^{\frac{1}{\pi}} &= \gamma_j 2^{j(s+\frac{1}{2})} \left( \sum_{k \in R_j} \eta_{j,k}^{-\pi} 2^{-j} \right)^{\frac{1}{\pi}} \\ &\leq C \gamma_j 2^{j(s+\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{\pi})}. \end{aligned}$$

Using Property 2.4 with  $q(w) = 0$  and taking  $j$  large enough, only the following constraint on  $\gamma_j$  is necessary to guarantee that  $g_\epsilon \in B_{s,\pi,r}(L)$ :

$$\gamma_j \leq C' 2^{-j(s+\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{\pi})}$$

where  $C'$  denotes a constant suitably chosen. Now, consider the following Lemma which will be proved in Appendix.

**Lemma 3.6.** *If we chose  $\gamma_j = n^{-\frac{1}{2}}$  then there exist  $\lambda > 0$  and  $p_0 > 0$  not depending on  $n$  such that*

$$\inf_{\substack{\epsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{P}_{g_\epsilon} (\wedge_n (g_{\epsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0, \quad \forall k \in R_j.$$

Putting Lemma 3.6 and Theorem 3.4 together, we obtain

$$\begin{aligned} \sup_{g_\epsilon \in B_{s,\pi,r}(L)} \mathbb{E}_{g_\epsilon} (\|\hat{f} - g_\epsilon\|_p^p) &\geq \frac{e^{-\lambda}}{2} \gamma_j^p \sum_{k \in R_j} \inf_{\substack{\epsilon_i \in \{-1, +1\} \\ i \neq k}} \omega_{j,k}^p \mathbb{P}_{g_\epsilon} (\wedge_n (g_{\epsilon_k^*}, g_\epsilon) > e^{-\lambda}) \|\psi_{j,k}\|_p^p \\ &\geq \frac{e^{-\lambda}}{2} 2^{j(\frac{p}{2}-1)} p_0 \left( \frac{1}{\sqrt{n}} \right)^p \|\psi\|_p^p \sum_{k \in R_j} \eta_{j,k}^{-p}. \end{aligned}$$

Using the inequality (43) and the fact that  $n^{\frac{1}{2s+1+\sigma-\frac{2}{\pi}}} \simeq 2^j$ , one gets

$$\begin{aligned} \inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f (\|\hat{f} - f\|_p^p) &\geq c (n^{-\frac{1}{2}} 2^{j(\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{p})})^p \\ &\geq c' n^{-\tilde{\alpha}p}. \end{aligned}$$

This ends the proof of Theorem 3.8. □

**Remark 3.3.** *Let  $p > 2$ . Consider the function  $v_*$  defined by (41). It clear that for  $N > s > 0$ ,  $\pi = p$  and  $r \geq 1$  we have*

$$\inf_{\hat{f}} \sup_{f \in B_{s,p,r}(L)} \mathbb{E}_f (\|\hat{f} - f\|_p^p) \asymp n^{-\alpha^*p}$$

where

$$\alpha^* = \frac{s}{1 + 2s + \sigma - \frac{2}{p}}.$$

So we have prove that if  $v$  does not belong to  $\mathbb{L}^\pi([0, 1])$  for  $\pi \geq p$  then the minmax rate over usual Besov balls under  $\mathbb{L}^p$  risk can be slower than  $n^{-\alpha^*p}$ . In particular, Theorem 3.8 shows that this rate of convergence can truly depend on the nature of  $v$ .

This arises a new question:

**Question 3.2.** *Can we find function spaces over which the minimax rate under the  $\mathbb{L}^p$  risk stay 'stable' for others functions  $v$  i.e which does not belong to  $\mathbb{L}^\pi([0, 1])$  ?*

The answer is developed in the following section.



## 4 Minimax study over weighted Besov balls

Now, we analyze the minimax properties in  $\mathbb{L}^p$  loss over weighted Besov balls. This section is focused on the proof of the following theorem:

**Theorem 4.1.** *Let  $\infty > p > 1$  and  $\infty \geq \pi \geq p$ . Assume that we observe model (1). Suppose that the function  $G$  defined by*

$$G(t) = \int_0^t \frac{1}{v^2(y)} dy$$

*is bijective with  $G(1) = 1$ . Assume also that*

$$v^2(G^{-1}(\cdot)) \in \mathcal{A}_p. \quad (44)$$

*Then for  $N > s > q(w)$  (see (8)) and  $\infty \geq r \geq 1$  we have*

$$\inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}^G(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \asymp n^{-\alpha_1 p}.$$

### 4.1 Proof of Theorem 4.1: upper bound and lower bound

#### 4.1.1 Upper bound

Here we proceed as in Theorem 3.3 by taking in account the fact that we work with the warped wavelet basis  $\xi^G$ .

**Theorem 4.2.** *Let  $\infty > p > 1$  and  $\infty \geq \pi \geq p$ . Assume that the condition (44) holds. Let us consider  $\hat{f}^l$  the linear estimator defined by*

$$\hat{f}^l(x) = \sum_{k \in \Delta_{j_0}} \hat{\alpha}_{j_0,k} \phi_{j_0,k}(G(x))$$

where

$$\hat{\alpha}_{j_0,k} = \int_0^1 \phi_{j_0,k}(G(t)) \frac{1}{v(t)} dY_t.$$

*Then for  $N > s > 0$  and  $\infty \geq r \geq 1$  there exists a constant  $C > 0$  such that*

$$\sup_{f \in B_{s,\pi,r}^G(L)} \mathbb{E}_f(\|\hat{f}^l - f\|_p^p) \leq C n^{-\alpha_1 p}$$

*for  $j_0$  the integer depending on  $n$  such that  $2^{j_0} \simeq n^{\frac{1}{1+2s}}$ .*

*Proof of Theorem 4.2.* Starting from the inequality (13), it suffices to consider the  $\mathbb{L}^\pi$  risk of  $\hat{f}^l$ . Minkowski's inequality yields

$$\begin{aligned} \mathbb{E}_f(\|\hat{f}^l - f\|_\pi^\pi) &\leq C(\mathbb{E}_f(\|\hat{f}^l - P_{j_0}^G(f)\|_\pi^\pi) + \|P_{j_0}^G(f) - f\|_\pi^\pi) \\ &\leq C(Q_1 + Q_2). \end{aligned} \quad (45)$$

The condition (44) and the embedding  $\mathcal{A}_p \subseteq \mathcal{A}_\pi$  for  $\pi \geq p$  imply that Property 2.5 holds for  $T = G$  and  $w = v^2(G^{-1}(\cdot))$  (see (5)). So

$$Q_2 \leq C 2^{-j_0 s \pi}. \quad (46)$$

Using the definition of  $\hat{f}^l$  and Property 2.1, one gets

$$\begin{aligned} Q_1 &= \mathbb{E}_f(\|\sum_{k \in \Delta_{j_0}} (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}^G) \phi_{j_0,k}(G(\cdot))\|_\pi^\pi) \\ &\leq C(2^{\frac{j_0 \pi}{2}} (\sum_{k \in \Delta_{j_0}} \mathbb{E}_f(|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}^G|^\pi) w(I_{j_0,k})) \\ &= C 2^{\frac{j_0 \pi}{2}} Q_1^*. \end{aligned} \quad (47)$$

Using the change of variable  $y = G^{-1}(t)$ , we obtain

$$\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}^G = \frac{1}{\sqrt{n}} \int_0^1 \phi_{j_0,k}(G(t)) \frac{1}{v(t)} dW_t$$

so

$$\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}^G \sim \mathcal{N}\left(0, \frac{1}{n}\right).$$

Applying the second point of Lemma 3.1 we obtain

$$\begin{aligned} Q_1^* &\leq Cn^{-\frac{\pi}{2}} \sum_{k \in \Delta_{j_0}} w(I_{j_0,k}) \\ &= Cn^{-\frac{\pi}{2}}. \end{aligned} \tag{48}$$

Combining (45), (46), (47), (48) and taking in account that  $2^{j_0} \simeq n^{\frac{1}{1+2s}}$ , we deduce that

$$\begin{aligned} \mathbb{E}_f(\|\hat{f}^l - f\|_{\pi}^2) &\leq C(2^{\frac{j_0\pi}{2}} n^{-\frac{\pi}{2}} + 2^{-j_0s\pi}) \\ &\leq C'n^{-\alpha_1\pi}. \end{aligned}$$

Using the inequality (13), it comes that

$$\sup_{f \in B_{s,\pi,r}^G(L)} \mathbb{E}_f(\|\hat{f}^l - f\|_p^2) \leq Cn^{-\alpha_1 p}.$$

This completes the proof of Theorem 4.2.  $\square$

#### 4.1.2 Lower Bound

**Theorem 4.3.** *Let  $\infty > p > 1$ . Assume that the condition (44) holds. Then there exists a constant  $c > 0$  such that for  $N > s > q(w)$ ,  $\infty \geq \pi \geq 1$  and  $\infty \geq r \geq 1$  we have*

$$\inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}^G(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^2) \geq cn^{-\alpha_1 p}.$$

*Proof of Theorem 4.3.* Let  $j$  an integer to be chosen below. Consider the functions  $g_\epsilon$  defined by (24) with

$$\omega_{j,k} = 1 \quad \text{and} \quad T = G.$$

Since the warped wavelet coefficients of  $g_\epsilon$ , denoted below by  $\beta_{j,k}^{G*}$ , are equal to  $\gamma_j \epsilon_k$ , we have

$$\begin{aligned} 2^{j(s+\frac{1}{2})} \left( \sum_{k \in R_j} |\beta_{j,k}^{G*}|^\pi w(I_{j,k}) \right)^{\frac{1}{\pi}} &= \gamma_j 2^{j(s+\frac{1}{2})} \left( \sum_{k \in R_j} w(I_{j,k}) \right)^{\frac{1}{\pi}} \\ &\leq \gamma_j 2^{j(s+\frac{1}{2})}. \end{aligned}$$

Using Property 2.4, the embedding  $\mathcal{A}_p \subseteq \mathcal{A}_\pi$  (which is true for  $\pi \geq p$ ) and taking  $j$  large enough, only the following constraint on  $\gamma_j$  is necessary to guarantee that  $g_\epsilon \in B_{s,\pi,r}^G(L)$ :

$$\gamma_j \leq C' 2^{-j(s+\frac{1}{2})}$$

where  $C'$  denotes a constant suitably chosen. Now, introduce the following lemma which will be proved in Appendix.

**Lemma 4.1.** *If we chose  $\gamma_j = n^{-\frac{1}{2}}$  then there exist  $\lambda > 0$  and  $p_0$  not depending on  $n$  such that*

$$\inf_{\substack{\epsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{P}_{g_\epsilon}(\wedge_n (g_{\epsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0, \quad \forall k \in R_j, \quad n \in \mathbb{N}^*.$$

Lemma 4.1 and Theorem 3.4 yield

$$\begin{aligned} \sup_{g_\epsilon \in B_{s,\pi,r}^G(L)} \mathbb{E}_{g_\epsilon}(\|\hat{f} - g_\epsilon\|_p^2) &\geq \frac{e^{-\lambda}}{2} \gamma_j^p \sum_{k \in R_j} \inf_{\substack{\epsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{P}_{g_\epsilon}(\wedge_n (g_{\epsilon_k^*}, g_\epsilon) > e^{-\lambda}) \|\psi_{j,k}\|_{w,p}^p \\ &\geq \frac{e^{-\lambda}}{2} p_0 \left(\frac{1}{\sqrt{n}}\right)^p \sum_{k \in R_j} \|\psi_{j,k}\|_{w,p}^p. \end{aligned} \tag{49}$$

$$\tag{50}$$

Since  $w = v^2(G^{-1}(\cdot))$  satisfies the  $\mathcal{A}_p$  condition, the change of variable  $y = G(x)$  gives us

$$\begin{aligned} \|\psi_{j,k}\|_{w,p}^p &= \left( \int_{S_{j,k}} |\psi_{j,k}(x)|^p w(x) dx \right) \\ &\geq cw(S_{j,k}) \left( \frac{1}{|S_{j,k}|} \int_{S_{j,k}} |\psi_{j,k}(x)| dx \right)^p \\ &\geq c2^{\frac{jp}{2}} (2N-1)^{-p} w(S_{j,k}) \|\psi\|_1^p. \end{aligned} \quad (51)$$

Using Lemma 2.1, we obtain

$$\begin{aligned} \sum_{k \in R_j} w(S_{j,k}) &= w\left(\left[\frac{C_* - 1}{2^j}, 1 - \frac{1}{2^j}\right]\right) \\ &\geq \left(1 - \frac{C_*}{2^j}\right)^p w([0, 1]) \\ &\geq \left(1 - \frac{2N-2}{2^\tau}\right)^p w([0, 1]) \\ &= c. \end{aligned} \quad (52)$$

Putting (49), (51) and (52) together and choosing  $n^{-\frac{1}{2}} \simeq 2^{-j(s+\frac{1}{2})}$  (i.e.  $2^j \simeq n^{\frac{1}{1+2s}}$ ), we see that

$$\inf_{\hat{f}} \sup_{f \in B_{s,p,r}^G(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \geq c \left(\frac{2^{\frac{j}{2}}}{\sqrt{n}}\right)^p \quad (53)$$

$$\geq c' n^{-\alpha_1 p}. \quad (54)$$

□

Finally, by combining Theorem 4.2 and Theorem 4.3 we prove Theorem 4.1.

**Remark 4.1.** *If  $v$  is a positive constant then we obtain the usual minimax result.*

**Remark 4.2.** *If  $v$  is bounded below, then the inequality (53) become an obvious consequence of the inequality*

$$\begin{aligned} \sum_{k \in R_j} \|\psi_{j,k}\|_{w,p}^p &\geq c \sum_{k \in R_j} \|\psi_{j,k}\|_p^p \\ &\geq c' 2^{\frac{jp}{2}}. \end{aligned}$$

*In this case, one can show that Theorem 4.1 is always true if we consider the warping function*

$$G_m(x) = \int_0^x \frac{1}{v^m(x)} dx, \quad m \geq 2$$

*instead of  $G$  and if we assume that  $v^m(G_m^{-1}) \in \mathcal{A}_p$ .*

## 4.2 Other results and examples

The following lemma, proved in Kerkyacharian and Picard (2005), proposes another version of the condition (44).

**Lemma 4.2.** *Let  $p > 1$  and  $q$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ . Then  $v^2(G^{-1}(\cdot))$  satisfies the  $\mathcal{A}_p$  condition iff there exists a constant  $C > 0$  such that*

$$\left(\frac{1}{|I|} \int_I \frac{1}{v^{2q}(x)} dx\right)^{\frac{1}{q}} \leq C \left(\frac{1}{|I|} \int_I \frac{1}{v^2(x)} dx\right)$$

*for any subinterval  $I$  of  $[0, 1]$ .*

In order to illustrate our statistical results, consider some examples.

**Example 4.1.** Observe the model defined in (1) with the operator  $H_{v_1}$  where

$$v_1(t) = c_\sigma t^{-\frac{\sigma}{2}}, \quad \sigma > \frac{1}{p} - 1$$

and  $c_\sigma = (\sigma + 1)^{-\frac{1}{2}}$ . It is clear that  $\frac{1}{v_1}$  belongs to  $\mathbb{L}^2([0, 1])$ . Moreover

$$G(x) = x^{\sigma+1}, \quad G^{-1}(x) = x^{\frac{1}{\sigma+1}} \quad \text{and} \quad w(x) = v_1^2(G^{-1}(x)) = x^{\frac{-\sigma}{\sigma+1}}.$$

Then if  $\sigma > \frac{1}{p} - 1$  we have  $p - 1 > -\frac{\sigma}{\sigma+1} > -1$  so the function  $w$  satisfies the  $\mathcal{A}_p$  condition. Therefore, all conditions are satisfied to apply Theorem 4.1.

Let  $p > 2$ ,  $s > 1$  and  $\pi = p$ . The following table summarizes the results of Example 3.2, Remark 3.3 and Example 4.1.

Model (1) where	Space $A$	$\sigma$ values	$\inf_{\hat{f}} \sup_{f \in A} \mathbb{E}(\ \hat{f} - f\ _p^p)$
$v_1(t) = c_\sigma t^{-\frac{\sigma}{2}}$	$B_{s,p,r}(L)$	$0 < \sigma < \frac{2}{p}$	$\asymp n^{-\alpha_1 p}, \quad \alpha_1 = \frac{s}{1+2s}$
	$B_{s,p,r}(L)$	$\frac{2}{p} < \sigma < 1$	$\asymp n^{-\alpha^* p}, \quad \alpha^* = \frac{s}{1+2s+\sigma-\frac{2}{p}}$
	$B_{s,p,r}^G(L)$	$0 < \sigma$	$\asymp n^{-\alpha_1 p}, \quad \alpha_1 = \frac{s}{1+2s}$

For  $\frac{2}{p} < \sigma < 1$ , remark that the minimax rate over usual Besov balls is strictly slower than the minimax rate over weighted Besov balls.

**Remark 4.3.** The previous table show that if we chose  $G(t) = c_\sigma^{-2} t^{\sigma+1}$  with  $\frac{2}{p} < \sigma < 1$ , then for  $p > 2$ ,  $r \geq 1$  and  $s > 1$  the following inclusion

$$B_{s,p,r}(L) \subseteq B_{s,p,r}^G(L)$$

is impossible. Because if we assume the contrary then it should exist  $c > 0$  and  $C > 0$  such that

$$\begin{aligned} cn^{-\alpha^* p} &\leq \inf_{\hat{f}} \sup_{f \in B_{s,p,r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \\ &\leq \inf_{\hat{f}} \sup_{f \in B_{s,p,r}^G(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \\ &\leq Cn^{-\alpha_1 p}. \end{aligned}$$

This contradicts the fact that  $\alpha^* < \alpha_1$  and implies the non-embedding.

The following examples exhibit functions  $v$  which not satisfy  $v \in \mathbb{L}^\pi$  but satisfy the Muckenhoupt condition.

**Example 4.2.** Observe the model defined in (1) with the operator  $H_{v_3}$  where

$$v_3(t) = \left(\frac{\pi}{2} \alpha t^{\alpha-1} \cos\left(\frac{\pi}{2} t^\alpha\right)\right)^{-\frac{1}{2}}, \quad \frac{1}{p} < \alpha < 1.$$

It is easy to see that

$$\begin{cases} \lim_{t \rightarrow 0} v_3(t) = 0, \\ \lim_{t \rightarrow 1} v_3(t) = \infty. \end{cases}$$

Moreover, we have

$$G(x) = \int_0^x \frac{1}{v_3^2(y)} dy = \sin\left(\frac{\pi}{2} x^\alpha\right), \quad G^{-1}(x) = \left(\frac{2}{\pi} \arcsin(x)\right)^{\frac{1}{\alpha}}$$

and by definition

$$w(x) = v_3^2(G^{-1}(x)) = C \frac{(\arcsin(x))^{\frac{1-\alpha}{\alpha}}}{\sqrt{1-x^2}}.$$

Since  $w$  is continuous on  $]0, 1[$ , we only need to study  $w$  at the points 0 and 1. At the neighborhood of 0, we have

$$w(x) \sim C' x^{\frac{1-\alpha}{\alpha}}$$

and  $x^{\frac{1-\alpha}{\alpha}} \in \mathcal{A}_p$  for  $-1 < \frac{1-\alpha}{\alpha} < p-1$  i.e  $\alpha > \frac{1}{p}$ .  
At the neighborhood of 1 we have

$$w(x) \sim C''' \frac{1}{\sqrt{1-x}}$$

and  $\frac{1}{\sqrt{1-x}} \in \mathcal{A}_p$ . Thus we conclude that  $w$  belongs to  $\mathcal{A}_p$  and that we can apply Theorem 4.1. Note that there exists other functions of this type. See for instance

$$v_4(t) = \left(\frac{\pi}{2}\alpha \cos\left(\frac{\pi}{2}t\right) \left(\sin\left(\frac{\pi}{2}t\right)\right)^{\alpha-1}\right)^{-\frac{1}{2}}, \quad \frac{1}{p} < \alpha < 1$$

which gives the weight

$$w(t) = v_4^2(G^{-1}(t)) = \frac{x^{\frac{1-\alpha}{\alpha}}}{\sqrt{1-x^{\frac{2}{\alpha}}}}.$$

Thus we have shown that weighted Besov balls give us stable minimax results for certain  $v$  which does not belong to  $\mathbb{L}^\pi$  for  $\pi \geq p > 2$ . Starting from these results, we propose to investigate the performance of an adaptive procedure constructed on  $\xi^G$  over  $B_{s,\pi,r}^G(L)$  in the case where  $\pi \geq p$ ,  $r \geq 1$  and  $N > s > q(w)$ .

## 5 Hard thresholding procedure and warped wavelet bases

Among other things, we prove that the linear procedure defined by (45) are optimal over weighted Besov balls. This procedure is not adaptive, i.e achieve substantially slower rate of convergence if the smoothness of the function that we wish to estimate is misspecified.

In recent years, a variety of adaptive procedures have been proposed. Among them, let us quote the wavelet thresholding methods introduced by Donoho and Johnstone which enjoy excellent statistical results for numerous risks. See Donoho and Johnstone (1995) and Johnstone (1998).

The following section is focused on the performance of a hard thresholding procedure constructed on  $\xi^G$  over weighted Besov balls  $B_{s,\pi,r}^G(L)$ .

**Theorem 5.1.** *Let  $p > 1$ . Assume that the condition (44) holds. Let us consider the following hard thresholding estimator:*

$$\tilde{f}(x) = \sum_{j \in \Lambda^*} \sum_{k \in \Delta_j} \hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}| \geq \kappa \sqrt{\frac{\ln(n)}{n}}\}} \psi_{j,k}(G(x)) \quad (55)$$

where

$$\hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(G(t)) \frac{1}{v(t)} dY_t,$$

with  $\Lambda^* = \{j; \tau-1 \leq j \leq j_*\}$  for  $j_*$  the integer verifying

$$2^{j_*} \leq \frac{n}{\ln(n)} < 2^{j_*+1}. \quad (56)$$

We have adopted the following notation:

$$\hat{\beta}_{\tau-1,k} = \hat{\alpha}_{\tau,k} = \int_0^1 \phi_{\tau,k}(G(t)) \frac{1}{v(t)} dY_t.$$

Then for  $\kappa > 0$  a large enough constant,  $N > s > 0$ ,  $\infty \geq r \geq 1$  and  $\infty \geq \pi \geq p$ , we have

$$\sup_{f \in B_{s,\pi,r}^G(L)} \mathbb{E}(\|\tilde{f} - f\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_1 p}.$$

*Proof of Theorem 5.1.* Here we propose to exhibit the maxiset of the procedure  $\tilde{f}$  and to show that  $B_{s,\pi,r}^G(L)$  is included into this maximal space. To isolate such a maxiset, five conditions must be checked.

- Two on them concern the geometrical properties of  $\xi^G$ ,
- one concerns a weight inequality,

- two of them concern the estimator  $\hat{\beta}_{j,k}$ .

The proof rests on the article of Picard and Kerkyacharian (2000). For further details on the maxiset theory see Cohen, Picard and Kerkyacharian (2000) and Autin (2004).

The geometrical properties of the basis are concentrated in the following lemma:

**Lemma 5.1.** *The condition (44) implies that  $\xi^G$*

- *satisfies Temlyakov's property i.e there exist two positive constants  $c$  and  $C$  such that for any finite set of integer  $F \subseteq \mathbb{N} \cup \{\tau - 1\} \times \Delta_j$  we have*

$$c \sum_{j,k \in F} \|\psi_{j,k}(G(\cdot))\|_p^p \leq \left\| \left( \sum_{j,k \in F} |\psi_{j,k}(G(\cdot))|^2 \right)^{\frac{1}{2}} \right\|_p^p \leq C \sum_{j,k \in F} \|\psi_{j,k}(G(\cdot))\|_p^p,$$

- *is unconditional for the  $\mathbb{L}^p$  norm i.e there exists an absolute constant  $C$  such that if  $|u_{j,k}| \leq |v_{j,k}|$  for all  $(j,k) \in \mathbb{N} \cup \{\tau - 1\} \times \Delta_j$ , then*

$$\left\| \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} u_{j,k} \psi_{j,k}(G(\cdot)) \right\|_p^p \leq C \left\| \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} v_{j,k} \psi_{j,k}(G(\cdot)) \right\|_p^p.$$

*Proof of Lemma 5.1.* The first point was shown by Garcia-Martell (1999) and the second point was shown by Kerkyacharian and Picard (2003).  $\square$

Property 2.2 and the definition of  $j_*$  (see (56)) yield

$$\begin{aligned} \left( \frac{\ln(n)}{n} \right)^{\frac{p}{2}} \sum_{k \in \Lambda^*} \|\psi_{j,k}(G(\cdot))\|_p^p &\leq C \left( \frac{\ln(n)}{n} \right)^{\frac{p}{2}} \sum_{j \leq j_*} 2^{\frac{jp}{2}} \\ &\leq C' \left( \frac{\ln(n)}{n} \right)^{\frac{p}{2}} 2^{\frac{j_* p}{2}} \\ &\leq C''. \end{aligned}$$

Thus, the weight condition holds.

Since  $\ln(n) \geq 1$  for  $n \geq 3$ , Lemma 3.1 yields

- $\mathbb{P}_f(|\hat{\beta}_{j,k} - \beta_{j,k}^G| \geq \frac{\kappa}{2} \sqrt{\frac{\ln(n)}{n}}) \leq C \left( \frac{\ln(n)}{n} \right)^p,$
- $\mathbb{E}_f(|\hat{\beta}_{j,k} - \beta_{j,k}^G|^{2p}) \leq C \left( \frac{\ln(n)}{n} \right)^p.$

We deduce that the statistical conditions are satisfied.

Combining all these results, we can apply the maxiset theorem which said that for any  $\infty > p > 1$ ,  $1 > \tilde{\nu} > 0$  and  $\kappa$  large enough, there exists a positive constant  $C$  such that the following equivalence holds

$$\mathbb{E}_f(\|\tilde{f} - f\|_p^p) \leq C \left( \frac{\ln(n)}{n} \right)^{\frac{\tilde{\nu} p}{2}} \iff f \in \mathcal{M}(p, \tilde{\nu}, G)$$

where we have set

$$\begin{aligned} \mathcal{M}(p, \tilde{\nu}, G) &= E_1 \cap E_2, \\ E_1 &= \{f; \sup_{u>0} u^{(1-\tilde{\nu})p} \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} 1_{\{|\beta_{j,k}^G| > u\}} \|\psi_{j,k}(G(\cdot))\|_p^p < \infty\} \end{aligned}$$

and

$$E_2 = \{f; \sup_{l > \tau-1} 2^{\frac{l\tilde{\nu}p}{2}} \left\| \sum_{j \geq l} \sum_{k \in \Delta_j} \beta_{j,k}^G \psi_{j,k}(G(\cdot)) \right\|_p^p < \infty\}.$$

The following lemma allows us to conclude:

**Lemma 5.2.** *For  $N > s \geq \frac{1}{2}$  and  $\pi \geq p > 1$ , we have the following embedding*

$$B_{s,\pi,r}^G(L) \subset \mathcal{M}\left(p, \frac{2s}{2s+1}, G\right)$$

One part of the proof is given in Appendix.  $\square$

Finally, we have proved that hard thresholding procedure defined in (55) achieves the minimax rate of convergence up to a logarithmic term over weighted Besov balls  $B_{s,\pi,r}^G(L)$ .

## 6 Warped Gaussian noise model

Consider the Gaussian white noise model in which we observe Gaussian processes  $Y_t^*$  governed by the stochastic equation

$$dY_t^* = K_H(f)(t)dt + \frac{1}{\sqrt{n}}dW_t, \quad n \in \mathbb{N}^*, \quad t \in [0, 1], \quad (57)$$

where  $K_H : \mathcal{B}([0, 1]) \rightarrow \mathcal{B}([0, 1])$  denotes the warping operator defined by

$$K_H(f)(t) = f(H^{-1}(t)), \quad t \in [0, 1],$$

$G$  is a function measurable on  $[0, 1]$  which are known, increasing, bijective and absolutely continuous such that

$$H' = h, \quad H(0) = 0 \quad \text{and} \quad H(1) = 1.$$

$W_t$  is a standard Brownian motion on  $[0, 1]$ . The function  $f$  is the unknown function of interest. The basic assumptions are

- $h \in \mathbb{L}^1([0, 1])$ ,
- $\sup_{x \in [0, 1]} |f(x)| < \infty$ .

Starting from the model (57), one can set two theorems very similar to Theorem 6.1 and Theorem 6.2.

**Theorem 6.1.** *Let  $\infty > p \geq 1$  and  $\infty \geq \pi \geq p$ . Assume that  $h$  satisfies the following condition:*

$$\frac{1}{h} \in \mathbb{L}^{\pi_*}([0, 1]) \quad (58)$$

where  $\pi_* = \max(\frac{\pi}{2}, 1)$ . Then for  $N > s > 0$  and  $\infty \geq r \geq 1$  we have

$$\inf_{\hat{f}} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \asymp n^{-\alpha_1 p}.$$

**Theorem 6.2.** *Let  $\infty > p > 1$ . Assume that  $v$  satisfies the following condition:*

$$\left\| \frac{1}{h} \right\|_\infty < \infty. \quad (59)$$

Then for  $N > s > \frac{1}{\pi}$ ,  $\infty \geq \pi \geq 1$ ,  $\infty \geq r \geq 1$  and  $\epsilon < 0$  we have

$$\inf_{\hat{f}} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \asymp \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p}$$

For  $\epsilon = 0$ , there exist  $C > 0$  and  $c > 0$  satisfying

$$c \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} \leq \inf_{\hat{f}} \sup_{f \in B_{s, \pi, r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \leq C \left(\frac{\ln(n)}{n}\right)^{\alpha_2 p} (\ln(n))^{\left(\frac{p}{2} - \frac{\pi}{r}\right)_+}.$$

*Proofs of Theorem 6.1 and 6.2.* For the upper bounds, it suffices to consider the estimators

$$\hat{\phi}_{j,k} = \int_0^1 \phi_{j,k}(H^{-1}(t)) \frac{1}{h(H^{-1}(t))} dY_t^*, \quad \hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(H^{-1}(t)) \frac{1}{h(H^{-1}(t))} dY_t^*,$$

to remark that

$$\hat{\beta}_{j,k} - \beta_{j,k} \sim \mathcal{N}(0, (\eta_{j,k}^*)^2)$$

with

$$\eta_{j,k}^* = \left( \int_0^1 \psi_{j,k}^2(t) \frac{1}{h(t)} dt \right)^{\frac{1}{2}},$$

to use same procedures which appeared in Theorem 3.1, Theorem 3.2 and the same techniques of their proofs. The lower bound is an immediate consequence of Theorem 3.4 and Property 2.1.  $\square$

Moreover, one can easily show the following theorems:

**Theorem 6.3.** Let  $\infty > p > 1$  and  $\infty \geq \pi \geq p$ . Assume that

$$\frac{1}{h(H^{-1}(\cdot))} \in \mathcal{A}_p. \quad (60)$$

Then for  $N > s > q(w)$  (see (8)) and  $\infty \geq r \geq 1$  we have

$$\inf_f \sup_{f \in B_{s,\pi,r}^H(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \asymp n^{-\alpha_1 p}.$$

**Theorem 6.4.** Let  $p > 1$ . Assume that the condition (60) holds. Let us consider the following hard thresholding procedure (55) with  $G = H$  and

$$\hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(H^{-1}(t)) \frac{1}{h(H^{-1}(t))} dY_t^*.$$

Then for  $\kappa > 0$  a large enough constant,  $N > s > 0$ ,  $\infty \geq r \geq 1$  and  $\infty \geq \pi \geq p$ , we have

$$\sup_{f \in B_{s,\pi,r}^H(L)} \mathbb{E}(\|\tilde{f} - f\|_p^p) \leq C \left( \frac{\ln(n)}{n} \right)^{\alpha_1 p}.$$

**Remark 6.1.** Note that the assumption (60) authorizes the fact that  $h$  is not bounded from above and not bounded from below. For instance, consider

$$h(x) = \frac{\pi}{2} \alpha x^{\alpha-1} \cos\left(\frac{\pi}{2} x^\alpha\right), \quad \frac{1}{p} < \alpha < 1, \quad x \in [0, 1].$$

## 7 Appendix: Proofs of technical Lemmas

*Proof of Lemma 3.1.* It is well known that if  $N \sim \mathcal{N}(0, \sigma^2)$  then we have the following concentration inequality:

$$\mathbb{P}(|N| \geq x) \leq 2 \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Thus, for  $\kappa \geq 2\sqrt{2p}$  and  $n \geq 3$ , we have

$$\mathbb{P}_f \left( |V_n| \geq \frac{\kappa}{2} \sqrt{\frac{\ln(n)}{n}} \right) \leq 2 \exp\left(-\frac{\kappa^2 n \left(\frac{\ln(n)}{n}\right)}{8}\right) = 2n^{-\frac{\kappa^2}{8}} \leq 2n^{-p}.$$

Moreover, it is well known that if  $N \sim \mathcal{N}(0, \sigma^2)$  then for all  $\alpha > 1$  there exists  $C > 0$  such that  $\mathbb{E}_f(|V_n|^\alpha) \leq Cn^{-\frac{\alpha}{2}}$ .  $\square$

*Proof of Lemma 3.2.* Since  $\|f\|_\infty < \infty$  and  $\frac{1}{v}$  belongs to  $\mathbb{L}^2([0, 1])$ , one gets

$$\mathbb{E}_f \left( \exp\left(\frac{n}{2} \int_0^1 \frac{f^2(t)}{v^2(t)} dt\right) \right) \leq \exp\left(\frac{n}{2} \|f\|_\infty^2 \left\| \frac{1}{v} \right\|_2^2\right) < \infty.$$

Following Novikov's condition and Girsanov's theorem, the likelihood ratio defined by (25) can be written as follows:

$$\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon) = \exp\left(n \int_0^1 \frac{(g_{\varepsilon_k^*}(t) - g_\varepsilon(t))}{v(t)} dY_t - \frac{n}{2} \int_0^1 \frac{(g_{\varepsilon_k^*}^2(t) - g_\varepsilon^2(t))}{v^2(t)} dt\right).$$

Under  $\mathbb{P}_{g_\varepsilon}$ , we see that

$$\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon) = \exp\left(-\frac{n}{2} \int_0^1 \frac{(g_{\varepsilon_k^*}(t) - g_\varepsilon(t))^2}{v^2(t)} dt + \sqrt{n} \int_0^1 \frac{(g_{\varepsilon_k^*}(t) - g_\varepsilon(t))}{v(t)} dW_t\right).$$

Since  $g_{\varepsilon_k^*}(t) - g_\varepsilon(t) = -2\gamma_j \varepsilon_k \psi_{j,k}(t)$ , by choosing  $\gamma_j = n^{-\frac{1}{2}}$  we obtain

$$\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon) = \exp\left(-2 \int_0^1 \frac{\psi_{j,k}^2(t)}{v^2(t)} dt - 2\varepsilon_k \int_0^1 \frac{\psi_{j,k}(t)}{v(t)} dW_t\right).$$



Let

$$U_{j,k} = -2 \int_0^1 \frac{\psi_{j,k}^2(t)}{v^2(t)} dt \quad \text{and} \quad V_{j,k} = -2\epsilon_k \int_0^1 \frac{\psi_{j,k}(t)}{v(t)} dW_t.$$

We have clearly

$$\{\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) \geq e^{-\lambda}\} \supseteq \{U_{j,k} + V_{j,k} \geq -\lambda\} \supseteq \{|U_{j,k} + V_{j,k}| \leq \lambda\}.$$

Applying Chebychev's inequality and remarking that  $\mathbb{E}_f(|W|) = \sqrt{\frac{2}{\pi}}$  for  $W \sim \mathcal{N}(0, 1)$ , one obtains

$$\begin{aligned} \mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) \geq e^{-\lambda}) &\geq 1 - \mathbb{P}_{g_\epsilon}(|U_{j,k} + V_{j,k}| > \lambda) \geq 1 - \frac{1}{\lambda}(\mathbb{E}_{g_\epsilon}(|U_{j,k}|) + \mathbb{E}_{g_\epsilon}(|V_{j,k}|)) \\ &= 1 - \frac{1}{\lambda}(|U_{j,k}| + \frac{2}{\sqrt{\pi}}\sqrt{|U_{j,k}|}). \end{aligned}$$

Property 2.1 yields

$$\sum_{k \in R_j} |U_{j,k}| = 2 \int_0^1 \frac{1}{v^2(t)} \sum_{k \in R_j} \psi_{j,k}^2(t) dt \leq C \|\frac{1}{v}\|_2^2 2^j = C' 2^j \quad (61)$$

and

$$\sum_{k \in R_j} \sqrt{|U_{j,k}|} \leq \sqrt{\sum_{k \in R_j} |U_{j,k}|} \sqrt{\sum_{k \in R_j} 1} \leq C 2^j. \quad (62)$$

Thus, for  $\lambda$  large enough we have

$$\begin{aligned} \sum_{k \in R_j} \inf_{\substack{\varepsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) &\geq c 2^j - \frac{c'}{\lambda} 2^j \\ &\geq c'' 2^j. \end{aligned}$$

This ends the proof of Lemma 3.2. □

*proof of Lemma 3.3.* Since  $\|v\|_\infty < \infty$ , we have

$$\eta_{j,k} = \sqrt{\int_0^1 \psi_{j,k}^2(t) v^2(t) dt} \leq \|v\|_\infty.$$

Applying Lemma 3.1, the concentration inequality is obvious. Using the same lemma, one gets

$$\begin{aligned} \mathbb{P}_f(|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\kappa'}{2} \sqrt{\frac{j}{n}}) &\leq 2 \exp(-\frac{n(\kappa')^2 j}{8\|v\|_\infty^2 n}) \\ &\leq C 2^{-\kappa j} \end{aligned}$$

for  $\kappa' = \sqrt{8\|v\|_\infty^2 \kappa}$ . This ends the proof of Lemma 3.3. □

*Proof of Lemma 3.4.* Following Girsanov's theorem, under  $\mathbb{P}_{g_k}$  we have

$$\begin{aligned} \wedge_n(g_0, g_k) &= \exp\left(-\frac{n}{2} \int_0^1 \frac{(g_0(t) - g_k(t))^2}{v^2(t)} dt + \sqrt{n} \int_0^1 \frac{(g_0(t) - g_k(t))}{v(t)} dW_t\right) \\ &= \exp\left(-\frac{n}{2} \gamma_j^2 \int_0^1 \frac{\psi_{j,k}^2(t)}{v^2(t)} dt - \gamma_j \sqrt{n} \int_0^1 \frac{\psi_{j,k}(t)}{v(t)} dW_t\right). \end{aligned}$$

Setting

$$U'_{j,k} = -c_0^2 \frac{\ln(n)}{2} \int_0^1 \frac{\psi_{j,k}^2(t)}{v^2(t)} dt \quad \text{and} \quad V'_{j,k} = -c_0 \sqrt{\ln(n)} \int_0^1 \frac{\psi_{j,k}(t)}{v(t)} dW_t,$$

we have clearly

$$\{\wedge_n(g_0, g_k) \geq 2^{-\lambda^* j}\} \supseteq \{U'_{j,k} + V'_{j,k} \geq -\lambda^* \ln(2^j)\} \supseteq \{|U'_{j,k} + V'_{j,k}| \leq \lambda^* \ln(2^j)\}.$$

Applying Chebychev's inequality, one obtains

$$\begin{aligned} \mathbb{P}_{g_k}(\wedge_n(g_0, g_k) \geq 2^{-\lambda^* j}) &\geq 1 - \mathbb{P}_{g_k}(|U'_{j,k} + V'_{j,k}| \geq \lambda^* \ln(2^j)) \geq 1 - \frac{1}{\lambda^* \ln(2^j)} (\mathbb{E}_{g_k}(|U'_{j,k}|) + \mathbb{E}_{g_k}(|V'_{j,k}|)) \\ &= 1 - \frac{1}{\lambda^* \ln(2^j)} (|U'_{j,k}| + \frac{2}{\sqrt{\pi}} \sqrt{|U'_{j,k}|}). \end{aligned}$$

Choosing  $\gamma_j = \sqrt{\frac{\ln(n)}{n}}$  and using the inequalities (61) and (62), we show that

$$\begin{aligned} \sum_{k \in R_j} \mathbb{P}_{g_k}(\wedge_n(g_0, g_k) \geq 2^{-\lambda^* j}) &\geq \sum_{k \in R_j} (1 - \frac{1}{\lambda^* \ln(2^j)} (\mathbb{E}_{g_k}(|U'_{j,k}|) + \frac{2}{\sqrt{\pi}} \mathbb{E}_{g_k}(\sqrt{|U'_{j,k}|}))) \\ &\geq c2^j - \frac{c'2^j}{\lambda^* \ln(2^j)} (c_0^2 \ln(n) + c_0 \sqrt{\ln(n)}) \\ &\geq c2^j - \frac{c'2^j \ln(n)}{\lambda^* \ln(2^j)} (c_0^2 + c_0). \end{aligned}$$

Using assumption (39), for  $c_0$  small enough we see that

$$\begin{aligned} \sum_{k \in R_j} \mathbb{P}_{g_k}(\wedge_n(g_0, g_k) \geq 2^{-\lambda^* j}) &\geq c2^j - \frac{c''c_02^j}{\lambda^*} \\ &\geq p_02^j. \end{aligned}$$

and this ends the proof of Lemma 3.4.  $\square$

*Proof of Lemma 3.5. Upper bound:* We have the following splits:

$$\sum_{k \in \Delta_j} \rho_{j,k}^\pi = \sum_{k=0}^{N-1} \rho_{j,k}^\pi + \sum_{k=N}^{2^j-N-1} \rho_{j,k}^\pi + \sum_{k=2^j-N}^{2^j-1} \rho_{j,k}^\pi. \quad (63)$$

For the first term, we have

$$\begin{aligned} \sum_{k=0}^{N-1} \rho_{j,k}^\pi &\leq N \|\psi\|_\infty^\pi 2^{\frac{j\pi}{2}} \left( \int_0^{\frac{2N-1}{2^j}} t^{-\sigma} dt \right)^{\frac{\pi}{2}} \\ &\leq C 2^{\frac{j\pi}{2}} (2^{j(\sigma-1)\frac{\pi}{2}}) \\ &= C 2^{j\frac{\sigma\pi}{2}}. \end{aligned} \quad (64)$$

Since  $t^{-\sigma}$  is decreasing and  $\sum_{k \geq 1} k^{-\beta} < \infty$  for  $\beta > 1$ , one can bound the second term as

$$\begin{aligned} \sum_{k=N}^{2^j-N-1} \rho_{j,k}^\pi &\leq 2^{j\frac{\sigma\pi}{2}} \sum_{k=N}^{2^j-N-1} (k-N+1)^{-\frac{\sigma\pi}{2}} \\ &\leq C 2^{j\frac{\sigma\pi}{2}}. \end{aligned} \quad (65)$$

Remark that the function  $z(t) = t^{-\sigma+1}$  is concave on  $]0, 1]$  for  $\frac{2}{p} < \sigma < 1$ . So

$$\begin{aligned} \sum_{k=2^j-N}^{2^j-1} \rho_{j,k}^\pi &\leq N \|\psi\|_\infty^\pi 2^{\frac{j\pi}{2}} \left( \int_{1-\frac{2N-1}{2^j}}^1 t^{-\sigma} dt \right)^{\frac{\pi}{2}} \\ &\leq C 2^{\frac{j\pi}{2}} (1 - (1 - (2N-1)2^{-j})^{-\sigma+1})^{\frac{\pi}{2}} \\ &\leq C' 2^{j\frac{\sigma\pi}{2}}. \end{aligned} \quad (66)$$

Cauchy-Schwartz's inequality gives us

$$\begin{aligned} 1 &= \left( \int_0^1 \psi_{j,k}^2(t) dt \right)^\pi \\ &= \left( \int_0^1 \frac{v_*(t)}{v_*(t)} \psi_{j,k}(t) \psi_{j,k}(t) dt \right)^\pi \\ &\leq \rho_{j,k}^\pi \eta_{j,k}^\pi. \end{aligned} \quad (67)$$

Hence, combining (63), (64), (65), (66) and (67) we obtain

$$\sum_{k \in R_j} \eta_{j,k}^{-\pi} \leq \sum_{k \in \Delta_j} \rho_{j,k}^{\pi} \leq C 2^{\frac{j\sigma\pi}{2}}.$$

*Lower bound:* Since  $t^{-\sigma}$  is decreasing for  $\frac{2}{p} < \sigma < 1$ , we have clearly

$$c 2^{j \frac{\sigma\pi}{2}} \leq \sum_{k \in R_j} \left( \frac{k+N}{2^j} \right)^{-\frac{\sigma\pi}{2}} \leq \sum_{k \in R_j} \eta_{j,k}^{-\pi}. \quad (68)$$

This ends the proof of Lemma 3.5. □

*Proof of Lemma 3.6.* Following Girsanov's theorem, under  $\mathbb{P}_{g_\epsilon}$  we have

$$\Lambda_n(g_{\epsilon_k^*}, g_\epsilon) = \exp\left(-\frac{n}{2} \int_0^1 \frac{(g_{\epsilon_k^*}(t) - g_\epsilon(t))^2}{v_*^2(t)} dt + \sqrt{n} \int_0^1 \frac{(g_{\epsilon_k^*}(t) - g_\epsilon(t))}{v_*(t)} dW_t\right).$$

Since  $g_{\epsilon_k^*}(t) - g_\epsilon(t) = -2\eta_{j,k}^{-1} \gamma_j \epsilon_k \psi_{j,k}(t)$ , by choosing  $\gamma_j = n^{-\frac{1}{2}}$  we obtain

$$\begin{aligned} \Lambda_n(g_{\epsilon_k^*}, g_\epsilon) &= \exp\left(-2\eta_{j,k}^{-2} \int_0^1 \frac{\psi_{j,k}^2(t)}{v_*^2(t)} dt - 2\epsilon_k \eta_{j,k}^{-1} \int_0^1 \frac{\psi_{j,k}(t)}{v_*(t)} dW_t\right) \\ &= \exp\left(-2 - 2\eta_{j,k}^{-1} \epsilon_k \int_0^1 \frac{\psi_{j,k}(t)}{v_*(t)} dW_t\right). \end{aligned}$$

Since

$$-2\eta_{j,k}^{-1} \epsilon_k \int_0^1 \frac{\psi_{j,k}(t)}{v_*(t)} dW_t \sim \mathcal{N}(0, 4),$$

if we chose  $\lambda = 2$ , we have clearly

$$\mathbb{P}_{g_\epsilon}(\Lambda_n(g_{\epsilon_k^*}, g_\epsilon) > e^{-\lambda}) = \frac{1}{2}.$$

This ends the proof of Lemma 3.6. □

*Proof of Lemma 4.1.* Following Girsanov's theorem, under  $\mathbb{P}_{g_\epsilon}$  we obtain

$$\Lambda_n(g_{\epsilon_k^*}, g_\epsilon) = \exp\left(-\frac{n}{2} \int_0^1 \frac{(g_{\epsilon_k^*}(t) - g_\epsilon(t))^2}{v^2(t)} dt + \sqrt{n} \int_0^1 \frac{(g_{\epsilon_k^*}(t) - g_\epsilon(t))}{v(t)} dW_t\right).$$

Since  $g_{\epsilon_k^*}(t) - g_\epsilon(t) = -2\gamma_j \epsilon_k \psi_{j,k}(G(t))$ , by choosing  $\gamma_j = n^{-\frac{1}{2}}$  it follows that

$$\begin{aligned} \Lambda_n(g_{\epsilon_k^*}, g_\epsilon) &= \exp\left(-2 \int_0^1 \frac{\psi_{j,k}^2(G(t))}{v^2(t)} dt - 2\epsilon_k \int_0^1 \frac{\psi_{j,k}(G(t))}{v(t)} dW_t\right) \\ &= \exp\left(-2 - 2\epsilon_k \int_0^1 \frac{\psi_{j,k}(G(t))}{v(t)} dW_t\right). \end{aligned}$$

Since

$$-2\epsilon_k \int_0^1 \frac{\psi_{j,k}(G(t))}{v(t)} dW_t \sim \mathcal{N}(0, 4),$$

if we chose  $\lambda = 2$ , we have clearly

$$\mathbb{P}_{g_\epsilon}(\Lambda_n(g_{\epsilon_k^*}, g_\epsilon) > e^{-\lambda}) = \frac{1}{2}.$$

This finished the proof of Lemma 4.1. □

*Proof of Lemma 5.2.* For the following embedding

$$B_{s,\pi,r}^G(L) \subset \{f, \sup_{u>0} u^{\frac{\pi}{1+2s}} \sum_{j \geq \tau-1} \sum_{k \in \Delta_j} 1_{\{|\beta_{j,k}^G| > u\}} \|\psi_{j,k}(G(\cdot))\|_{\pi} < \infty\}$$

see Kerkyacharian and Picard (2005).

Assume that  $f$  belongs to  $B_{s,\pi,r}^G(L) \subset B_{s,\pi,\infty}^G(L)$  for all  $s \geq 0$ . Using Minkowski's inequality and Property 2.2, for any  $l \geq \tau - 1$  one gets

$$\begin{aligned} \left\| \sum_{j \geq l} \sum_{k \in \Delta_j} \beta_{j,k}^G \psi_{j,k}(G(\cdot)) \right\|_{\pi} 2^{\frac{ls}{1+2s}} &\leq \sum_{j \geq l} \left\| \sum_{k \in \Delta_j} \beta_{j,k}^G \psi_{j,k}(G(\cdot)) \right\|_{\pi} 2^{\frac{ls}{1+2s}} \\ &\leq C \sum_{j \geq l} 2^{\frac{j}{2}} \left( \sum_{k \in \Delta_j} |\beta_{j,k}^G|^{\pi} w(I_{j,k}) \right)^{\frac{1}{\pi}} 2^{\frac{ls}{1+2s}} \\ &\leq CL \sum_{j \geq l} 2^{\frac{ls}{1+2s} - js} \leq C' \sum_{j \geq l} 2^{(l-j)s} \leq C''. \end{aligned}$$

So

$$B_{s,\pi,\infty}^G(L) \subseteq \{f, \sup_{l > \tau-1} 2^{\frac{ls\pi}{1+2s}} \left\| \sum_{j \geq l} \sum_{k \in \Delta_j} \beta_{j,k}^G \psi_{j,k}(G(\cdot)) \right\|_{\pi} < \infty\}.$$

This completes the proof of Lemma 5.2. □

### Acknowledgment

The author thanks the Editor, the Associate Editor and the referees for their thorough and useful comments which have helped to improve the presentation of the paper.

### References

- Autin, F. (2004). *Point de vue maxiset en estimation non paramétrique*. Thèse de l'université Paris VII.
- Brown, L.D. and Low, M.G. (1996). *Asymptotic equivalence of nonparametric regression and white noise*. Ann. Statist. 24, 2384-2398.
- Cohen, A., Daubechies, I., Jawerth, B. and Vial, P.(1992) *Multiresolution analysis, wavelets and fast algorithms on an interval* Comptes Rendus Acad. Sc. Paris.
- Cohen, A., De Vore, R., Kerkyacharian, G., and Picard, D. (2000). *Maximal  $q$ -paces with given rate of convergence for thresholding algorithms*. Appl. Comput. Harmon. Anal., **11**, 167-191.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM, Philadelphia.
- Donoho, D.L., Johnstone, I.M. (1995). *Adaptating to unknown smoothness via wavelet shrinkage*. J. Am Stat. Assoc, **90**(432), 1200-1224.
- Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D.(1996). *Density estimation by Wavelet thresholding*. The annals of Statistics, 24(2), 508-539.
- Efromovich, S. and Pinsker, M.S. (1996). *Sharp-optimal and adaptative estimation for heteroscedastic nonparametric regression*. Statistica Sinica 6, 925-942.
- Garcia-Cuerva, J. and Martell, B. (1999). *Wavelet characterisation of weighted spaces*. Journal of Geometric Analysis.
- Grams, I. and Nussbaum, M. (1998). *Asymptotic equivalence for non parametric generalized linear models*. Probab. Theor. Related Fields 111, 493-533.
- Härdle, W., Kerkyacharian, G., Picard, D., and Tsybakov, A. (1998). *Wavelet, Approximation and Statistical Applications*. Lectures Notes in Statistics 129. Springer Verlag. New York.
- Ibragimov, I.A. and Khaminskii, R.Z. (1977). *On the estimation of an infinite dimensional parameter in Gaussian white noise*. Soviet Math. Dokl. 236, No.5, 1053-1055.

- Johnson, R. and Neugebauer, C.J. (1987). *Homeomorphisms perserving  $\mathcal{A}_p$* . Rev. Mat. Iberoamericana 3, 249-273.
- Johnstone, I.M. (1998). *Function estimation : White noise, sparsity and wavelets*. Lectures notes 130p.
- Kerkyacharian, G., Picard, D. (2000). *Thresholding algorithms, maxisets and well concentrated bases*, with discussion, Test, vol9, No2,P 283-345.
- Kerkyacharian, G., Picard, D. (2003). *Non-linear approximation and Muckenhoupt weights*. Preprint.
- Kerkyacharian, G., Picard, D. (2005). *Regression in random design and warped wavelets*. Bernoulli, 10(6), 1053-1105.
- Meyer, Y. (1990). *Ondelettes et Opérateurs*, Hermann, Paris.
- Muckenhoupt, B. (1972). *Weighted norm inequalities for the hardy maximal function*. Trans. Amer. Math. Hencec. 165, 207-226.
- Nussbaum, M. (1996). *Asymptotic equivalence of density estimation and white noise*. Annals of Statistics, 24, 2399-2430.
- Qui, B.H. (1982). *Weighted Besov and Triebel spaces: Interpolation by the real method*. Hiroshima Mathematical Journal. 581-605.
- Stein, E. (1993). *Harmonic Analysis : Real Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press.
- Tsybakov, A.B. (2004) *Introduction à l'estimation non-paramétrique*. Springer.