A maxiset approach of a Gaussian noise model

Christophe Chesneau

A maxiset approach of a Gaussian white noise model

Chesneau Christophe

March 18, 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 the heteroscedastic white noise setting under $\mathbb{L}^p$ risk. We show that the major connection which exists between Muckenhoupt theory and the geometrical properties of warped wavelet bases $\{\psi_{j,k}(G)\}$ allows us to consider spaces over which the minimax rate is stable for a wide class of variance functions v, contrarily to the usual wavelet approach. Adopting the maxiset point of view, we show that the hard thresholding procedure constructed on such a warped wavelet basis is close to the optimal over weighted Besov classes.

Key words and phrases: Minimax, Muckenhoupt weights, Heteroscedastic Gaussian noise, warped wavelets, wavelet thresholding.

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

1 Introduction

Consider the heteroscedastic white noise model in which we observe Gaussian processes Y_t governed by the stochastic equation:

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

The variance function v is known, positive and belongs to $\mathbb{L}^2([0,1])$. The process W_t is a standard Brownian motion on $[0,1]$. The function f is an unknown function of interest. We wish to estimate f on $[0,1]$ given a realization $y = \{Y_t,\ t \in [0,1]\}$ with small $\mathbb{L}^p$ risk:

$$\mathbb{E}(\int_0^1 |\hat{f}(t) - f(t)|^p dt).$$

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). Numerous results were established in that case under certain assumptions on the smoothness of f, model (1) is 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 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, Brown and Low (1996) have shown that model (1) with $v = \frac{\sigma}{\sqrt{g}}$ is asymptotically equivalent to the observation of data $(Y_1, X_1), ..., (Y_n, X_n)$ with:

$$Y_i = f(X_i) + \sigma(X_i)\epsilon_i \tag{2}$$

The variables X_i are i.i.d, independent of ϵ_i and with density g. The variables ϵ_i are normal i.i.d with mean zero and variance 1. Other equivalence can be found in Grama and Nussbaum (1998). Also let us quote Efromovich and Pinsker (1996) who use model (1) in order to estimate a smooth regression function under $\mathbb{L}^2$ loss for the case of model (2).

In this paper, we consider the problem in the framework of wavelet analysis. Our goal is to show that the union of Muckenhoupt theory with warped wavelet bases $\{\psi_{j,k}(G)\}$ proved themselves powerful tools which allow us to consider adaptative procedures and spaces over which they reach the minimax rate (up to a logarithmic factor) under $\mathbb{L}^p$-risk without assumption of boundedness from above and below for v. This study can be divided into three parts.

In a first part, we investigate the estimation of f from (1) over the usual Besov space $B_{s,\pi,r}(L)$ under $\mathbb{L}^p$ risk (see Definition 2.4). We show that if v and $\frac{1}{v}$ belong to $\mathbb{L}^p([0,1])$ then the minimax rate of convergence is of the form $n^{-\alpha}$ where $\alpha = \frac{sp}{1+2s}$. So the following question naturally arises : can we obtain this minimax rate over such a space for any v or $\frac{4}{v}$ which do not belong to $\mathbb{L}^p$ for $p > 2$? The answer is no. We exhibit such a function for which the minimax rate over usual Besov spaces is of the form $n^{-\beta}$ where $\beta < \alpha$.

This result motivates us to devote a second part in which we investigate another functional space more adapted to our model. Our choice will be made on Besov bodies constructed on a wavelet basis $\{\psi_{j,k}(G)\}$ warped by a factor G depending on v. Such spaces were developed by Kerkyacharian and Picard (2004) who establish good estimation results in a regression setting with random design (i.e (2) with $\sigma(.) = 1$) without assumption of boundedness from above and below for g. The key of the success of this study rests on the following argument : under certain conditions on v which refer to Muckenhoupt theory, the warped wavelet basis possesses some interesting geometrical properties in $\mathbb{L}^p$ norm which allow us to consider functional sets and procedures deeply linked to the model. Using these analytical tools, we show that if v, $\frac{1}{v}$ belong to $\mathbb{L}^2$ and if v is subject to a property of Muckenhoupt type then the minimax rate over weighted Besov spaces $B^G_{s,\pi,r}(L)$ defined starting from the primitive of $\frac{1}{v^2}$ (see Definition 2.4) is of the form $n^{-\alpha}$ where $\alpha = \frac{sp}{1+2s}$. This space has the advantage to take in account the variance functions v which do not necessarily satisfy the assumptions 'v and $\frac{1}{v}$ belong to $\mathbb{L}^p$ for $p > 2$'.

In a second part, we use such a basis to construct a natural procedure which stay as close as possible to the standard thresholding. In order to measure its performance under $\mathbb{L}^p$ risk, we adopt the maxiset point of view. This statistical tool developed by Cohen, De Vore, Kerkyacharian 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 $(\ln(n))^{\alpha}n^{-\alpha}$ where $\alpha = \frac{sp}{1+2s}$, we prove that the weighted Besov space introduced in the first part is included into the maxiset of our procedure. That allows us to conclude that it attains the minimax rate of convergence (up to a logarithmic factor) under $\mathbb{L}^p$ risk over this space.

The paper is organized as follows.
Section 2 defines the basic tools (Muckenhoupt weights, warped wavelet basis ...), inequalities and functional spaces we shall need in the study. In Section 3 we investigate the minimax rate over usual Besov spaces and we show that it truly depends on v. Section 4 investigates and discusses the minimax properties over a weighted Besov space. Section 5 is devoted to the performance of the associated hard thresholding procedure where the unknown function of interest belongs to these spaces. In Section 6, we introduce other statistical models and we explain why we can obtain results similar to those in the heteroscedastic white noise. Finally, Section 7 is devoted to the proofs of technical lemmas.

2 Muckenhoupt conditions, warped wavelet bases and functional spaces

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

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

$\mathbb{L}^p([0,1]) = \mathbb{L}^p_1([0,1])$ denotes the usual Lebesgue space.

2.1 Muckenhoupt condition

Let us first recall the following notion:

Definition 2.1 (Muckenhoupt condition). *Let $1 \le p < \infty$ and q such that $\frac{1}{p}+\frac{1}{p} = 1$. A non-negative locally integrable function m is said to verify the $\mathcal{A}_p$ condition if and only if there exists a constant $C > 0$ such that for any measurable functions h and any subintervals I of $[0,1]$ we have:*

$$(\frac{1}{|I|}\int_I |h(x)|dx) \le C(\frac{1}{m(I)}\int_I |h(x)|^p m(x)dx)^{\frac{1}{p}} \qquad (1)$$

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

Any w which verifies the $\mathcal{A}_p$ condition are called ***Muckenhoupt weights****.*

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

The previous condition has been introduced by Muckenhoupt in (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 $\mathbb{L}^p_m$ spaces like the Hardy-Littlewood maximal operator or the Hilbert transform. For the complete theory, see the book of Stein (1993).

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

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:*

$$w(B)(\frac{|S|}{|B|})^p \leq Cw(S)$$

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

2.2 Warped wavelet bases and Muckenhoupt weights

In this subsection we introduce the warped wavelet bases which can be viewed as a generalization of the regular wavelet bases. Then we state some important results that will be useful throughout the paper.

Definition 2.2 (Warped wavelet bases). *Let N be an integer of the form $2^u + 1$ where u belongs to $\{0, 1, 2, 3, 4\}$. We denote by*

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

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

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

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

with

$$S_{j,k} = Supp(\phi_{j,k}) = [\frac{k-N+1}{2^j}, \frac{k+N}{2^j}].$$

We add $N-1$ functions on the neighborhood of 0 which have the support contained in $[0, (2N-2)2^{-j}]$ and $N-1$ functions on the neighborhood of 1 which have the support contained in $[1-(2N-2)2^{-j}, 1]$. The wavelet functions are defined by

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

where

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

We add $N-1$ functions on the neighborhood of 0 which have the support contained in $[0, (2N-2)2^{-j}]$ and $N-1$ functions on the neighborhood of 1 which have the support contained in $[1-(2N-2)2^{-j}, 1]$. Let us define the integer $\tau - 1$ as $\psi_{\tau-1,k} = \phi_{\tau,k}$.

- *a known function $T : [0,1] \mapsto [0,1]$ which is bijective and absolutely continuous.*

We associate to this function the weight $w(.) = \frac{1}{\tilde{T}(T^{-1}(.))}$ where $\tilde{T}$ denotes the derivative of T and T^{-1} its inverse function. Remark that for any measurable positive function z, w satisfies:

$$\int_0^1 z(T(x))dx = \int_0^1 z(x)w(x)dx. \tag{2}$$

See Meyer (1990) and Daubechies (1992) for wavelet bases on the real line. See Cohen, Daubechies, Jawerth and Vial (1992) for further details on wavelet bases on the interval. The warped wavelet basis ξ^T was introduced by Picard and Kerkyacharian (2004).

2.3 Some properties linked to ξ^T

Let us start by setting an inequality which refers to the usual wavelet basis.

Lemma 2.2. *Let $v > 0$. There exists a constant $C > 0$ such that:*

$$\sum_{k\in\Delta_j} |\phi_{j,k}(x)|^v \leq C2^{\frac{jv}{2}}, \quad \sum_{k\in\Delta_j} |\psi_{j,k}(x)|^v \leq C2^{\frac{jv}{2}}, \quad x \in [0,1]. \tag{3}$$

The proof is given in the Appendix.

The following lemmas show that if w defined in (2) verifies a condition of Muckenhoupt type then ξ^T behaves quite similarly to a regular wavelet basis.

Lemma 2.3 (Decomposition on ξ^T). *Let $1 < p < \infty$. If w verifies the $\mathcal{A}_p$ condition then, for any $\nu \geq \tau$, a function f of $\mathbb{L}^p([0,1])$ can be decomposed on ξ^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

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

and

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

Lemma 2.4. *Under the assumptions of Lemma 2.3, there exist two constant $c > 0$ and $C > 0$ depending only on ψ, ϕ and w such that for $j \geq \tau$ we have:*

$$c2^{\frac{jp}{2}} \sum_{k\in\Delta_j} |\beta_{j,k}^T|^p w(I_{j,k}) \leq \| \sum_{k\in\Delta_j} \beta_{j,k}^T \psi_{j,k}(T(.))\|_p^p \leq C2^{\frac{jp}{2}} \sum_{k\in\Delta_j} |\beta_{j,k}^T|^p w(I_{j,k})$$

and:

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

where $I_{j,k} = [\frac{k}{2^j}, \frac{k+1}{2^j}]$ and $w(I) = \int_I w(x)dx$.

Proof of Lemmas 2.3 and 2.4. See Kerkyacharian and Picard (2004). □

These two lemmas will be in the heart of our minimax study. See Section 3 and Section 4.

2.4 Functional spaces

Let us define the Besov bodies constructed on ξ^T and weighted Besov spaces which play a determinant role in the sequel.

Definition 2.3 (weighted Besov body). *Let $L > 0$ be a constant. Under the same assumptions of Lemma 2.3, we say that a function f of $\mathbb{L}^p([0,1])$ belongs to the Besov body $b_{s,p,\infty}^T(L)$ if and only if we have:*

$$\|f\|_{s,p,\infty}^T = \sup_{j\geq\tau-1} (2^{j(s+\frac{1}{2})}(\sum_{k\in\Delta_j} |\beta_{j,k}^T|^p w(I_{j,k}))^{\frac{1}{p}}) \leq L < \infty.$$

Definition 2.4 (Weighted Besov spaces). *For any measurable function f we set:*

$$\Delta_{T,h}(f)(x) = f(T^{-1}(T(x)+h)) - f(x).$$

Recursively $\Delta_{T,h}^2(f)(x) = \Delta_{T,h}(\Delta_{T,h}(f))(x)$ and identically, for $N \in \mathbb{N}^*$,

$$\Delta_{T,h}^N(f)(x) = \Delta_{T,h}(\Delta_{T,h}^{N-1}(f))(x) = \sum_{k=0}^N \binom{N}{k}(-1)^k f(T^{-1}(T(x)+kh)).$$

We denote the associated N-th order modulus of smoothness as:

$$\rho^N(t,f,T,p) = \sup_{|h|\leq t}\left(\int_{J_{Nh}} |\Delta^N_{T,h}(f)(u)|^p du\right)^{\frac{1}{p}}$$

where $J_{Nh} = \{x\in[0,1]:\ T(x)+Nh\in[0,1]\}$. For $s>0$, $1<p,q\leq+\infty$, fix integer $N>s$. We say that a function f of $\mathbb{L}^p([0,1])$ belongs to the weighted Besov spaces $B^T_{s,p,r}(L)$ if and only if:

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

These spaces can be viewed as a generalization of the usual Besov spaces. The main interest of such spaces lies in the fact that they can model important forms of spacial inhomogeneity.

Notations 2.1. *If $T=I_d$, we simply denote $\xi^T=\xi$, $\alpha^T_{j,k}=\alpha_{j,k}$, $\beta^T_{j,k}=\beta_{j,k}$, $P^T_j(f)=P_j(f)$, $b^T_{s,p,\infty}(L)=b_{s,p,\infty}(L)$ and $B^T_{s,p,r}(L)=B_{s,p,r}(L)$.*

Lemma 2.5. *Under the assumptions of Lemma 2.3, we have the following equivalence:*

$$f\in b^T_{s,p,\infty}(L)\Longleftrightarrow f\in B^T_{s,\pi,r}(L)$$

for $N>s\geq q(w)$, $\pi\geq p$ and $r>0$ where

$$q(w)=\begin{cases}\inf_{v>1}\{\text{w satisfies the } \mathcal{A}_{\text{v}} \text{ condition}\} & \text{if w is not a constant,}\\ 0 & \text{if w is the identity.}\end{cases}$$

Proof of Lemma 2.5. See Kerkyacharian and Picard (2004) □

Remark 2.1. *In the sequel, the constants C, C', C'', c, c', c'' represent any constants we shall need, and can different from one line to one other.*

3 Minimax study over the usual Besov spaces

Let us state the main result of this section:

Theorem 3.1. *Let $2\leq p<\infty$. If v and $\frac{1}{v}$ belong to $\mathbb{L}^p([0,1])$ then for $s>0$ we have the following minimax result:*

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

where

$$\alpha=\frac{sp}{1+2s}.$$

The sequel is devoted to the proof of this theorem.

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 studying the upper bound attained by the associated linear wavelet procedure.

Theorem 3.2. *Let $2\leq p<\infty$, $s>0$. Suppose that v belongs to $\mathbb{L}^p([0,1])$. Let us consider $\hat f^l$ the linear estimator defined by:*

$$\hat f^l(x)=\sum_{k\in\Delta_{j(n)}}\hat\alpha_{j(n),k}\phi_{j(n),k}(x),\qquad \hat\alpha_{j(n),k}=\int_0^1\phi_{j(n),k}(t)dY_t.$$

Then we obtain the following upper bound:

$$\sup_{f\in b_{s,p,\infty}(L)}\mathbb{E}(\|\hat f^l-f\|_p^p)\leq Cn^{-\alpha}$$

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

Proof of Theorem 3.2. Using the Minkowski inequality and the elementary inequality

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

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

$$\begin{aligned}\mathbb{E}(\|\hat{f}^l - f\|_p^p) &\leq C(\mathbb{E}(\|\hat{f}^l - P_{j(n)}(f)\|_p^p) + \|P_{j(n)}(f) - f\|_p^p) \\ &= C(S_1+S_2). \end{aligned} \tag{2}$$

Expanding the function f on the basis ξ at level $\nu = j(n)$, assuming that $f \in b_{s,p,\infty}(L)$ and considering Lemma 2.4, one gets:

$$\begin{aligned} S_2 &= \|\sum_{j\geq j(n)}\sum_{k\in\Delta_j}\beta_{j,k}\psi_{j,k}(.)\|_p^p \leq C(\sum_{j\geq j(n)} 2^{j(\frac{1}{2}-\frac{1}{p})}(\sum_{k\in\Delta_j}|\beta_{j,k}|^p)^{\frac{1}{p}})^p \\ &\leq C(L\sum_{j\geq j(n)}2^{-js})^p = C'2^{-j(n)sp}(\sum_{j\geq 0}2^{-js})^p \leq C''2^{-j(n)sp}. \end{aligned}\tag{3}$$

Using the definition of $\hat{f}^l$ and Lemma 2.4, one gets:

$$\begin{aligned} S_1 &= \mathbb{E}(\|\sum_{k\in\Delta_{j(n)}}(\hat{\alpha}_{j(n),k}-\alpha_{j(n),k})\phi_{j(n),k}(.)\|_p^p) \\ &\leq C(2^{j(n)(\frac{p}{2}-1)}\sum_{k\in\Delta_{j(n)}}\mathbb{E}(|\hat{\alpha}_{j(n),k}-\alpha_{j(n),k}|^p)) \\ &= C2^{j(n)(\frac{p}{2}-1)}S_1^*. \end{aligned}\tag{4}$$

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

$$\rho_{j,k} = \sqrt{\int_0^1 v^2(t)\phi_{j,k}^2(t)dt}. \tag{5}$$

We have clearly

$$\hat{\alpha}_{j(n),k} - \alpha_{j(n),k} = \frac{1}{\sqrt{n}}\int_0^1 v(t)\phi_{j(n),k}(t)dW_t \sim \rho_{j(n),k}\epsilon_n \quad \text{with} \quad \epsilon_n \sim \mathcal{N}(0,\frac{1}{n}).$$

Let us now consider the following lemma:

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

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

The proof is given in the Appendix.
Using the second point of the previous lemma, one gets:

$$S_1^* \leq Cn^{-\frac{p}{2}}\sum_{k\in\Delta_{j(n)}}\rho_{j(n),k}^p \tag{6}$$

Applying the Hölder inequality with the measure $d\nu = \phi_{j(n),k}^2(t)dt$ and Lemma 2.2, one gets:

$$\begin{aligned}\sum_{k\in\Delta_{j(n)}}\rho_{j(n),k}^p &\leq Cn^{-\frac{p}{2}}\int_0^1 v^p(t)\sum_{k\in\Delta_{j(n)}}\phi_{j(n),k}^2(t)dt \\ &\leq C'n^{-\frac{p}{2}}2^{j(n)}\|v\|_p^p = C''n^{-\frac{p}{2}}2^{j(n)}. \end{aligned}\tag{7}$$

Combining (4), (6) and (7) we obtain:

$$S_1 \leq C2^{j(n)sp}n^{-\frac{p}{2}}. \tag{8}$$

Taking in account that $2^{j(n)} \simeq n^{\frac{1}{1+2s}}$, one gets:

$$\sup_{f\in b_{s,p,\infty}(L)}\mathbb{E}(\|\hat{f}^l-f\|_p^p) \leq C(2^{\frac{j(n)p}{2}}n^{-\frac{p}{2}} + 2^{-j(n)sp}) \leq C'n^{-\frac{sp}{1+2s}}.$$

This completes the proof of Theorem 3.2. □

3.1.2 Lower Bound

Theorem 3.3. *Under the assumptions of Theorem 3.1, there exists a positive constant $c > 0$ such that for $s > 0$:*

$$\inf_{\hat{f}} \sup_{f \in b_{s,p,\infty}(L)} \mathbb{E}(\|\hat{f} - f\|_p^p) \geq c n^{-\alpha}.$$

The proof of this Theorem rests on the Assouad's lemma.

Proof of Theorem 3.3. In the sequel, j denotes an integer to be chosen below and ε denotes a sequence such that $\varepsilon = (\varepsilon_k)_{k \in R_j} \in \{-1, +1\}^{\frac{2^j}{2(N-1)}}$ with $R_j = \{(2N-1)l - N;\ l = 1, 2, ..., \frac{2^j}{2(N-1)}\}$. This set is chosen in such a way that $S_{j,k}$ (see Definition 2.2) verifies $S_{j,k} \cap S_{j,k'} = \emptyset$ for $k, k' \in R_j$ with $k \neq k'$ and $\cup_{k \in R_j} S_{j,k} = [0, \frac{2N-1}{2(N-1)}] \subset [0,1]$. In the sequel, we denote $e_N = \frac{2N-1}{2(N-1)}$. Let us set:

$$g_\varepsilon(x) = \gamma_j \sum_{k \in R_j} \eta_{j,k}^{-1} \varepsilon_k \psi_{j,k}(x) \tag{9}$$

where $\eta_{j,k}$ is defined by:

$$\eta_{j,k} = \sqrt{\int_0^1 \frac{1}{v^2(t)} \psi_{j,k}^2(t) dt}. \tag{10}$$

In the sequel, we denote $\mathcal{G}$ the set of all such g_ε. Let us now consider the following proposition:

Lemma 3.2. *Let $2 \leq p < \infty$. If we assume that v and $\frac{1}{v}$ belongs to $\mathbb{L}^p([0,1])$ then there exist two constant $C > 0$ and $c > 0$ such that:*

$$c 2^j \leq \sum_{k \in R_j} \eta_{j,k}^{-p} \leq C 2^j.$$

The proof is given in the Appendix.

Using the definition of $b_{s,p,\infty}(L)$, Proposition 3.2 and the fact than $|\epsilon_i| \leq 1$, one gets:

$$\|g_\varepsilon\|_{s,p,\infty}^1 \leq \gamma_j 2^{j(s+\frac{1}{2})} (\sum_{k \in R_j} \eta_{j,k}^{-p} 2^{-j})^{\frac{1}{p}} \leq C^{\frac{1}{p}} \gamma_j 2^{j(s+\frac{1}{2})}.$$

Thus, for j large, only the following constraint on γ_j is necessary to guarantee that $g_\varepsilon \in b_{s,p,\infty}^1(L)$:

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

Let us now consider the following lemma.

Lemma 3.3. *Here and later, we denote $\mathbb{P}_f (= \mathbb{P}_f^n)$ the law of the process defined in (1) and $\mathbb{E}_f (= \mathbb{E}_f^n)$ the associated expectation. For $\varepsilon \in \{-1, +1\}^{\frac{2^j}{2(N-1)}}$, put $\varepsilon_k^* = (\varepsilon_i')_{i \in R_j}$ such that $\varepsilon_i' = \varepsilon_i 1_{\{i \neq k\}} - \varepsilon_i 1_{\{i = k\}}$. If there exist $\lambda > 0$ and $p_0 > 0$ such that*

$$\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0, \quad n \in \mathbb{N}^*$$

where we wrote:

$$\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) = \frac{d\mathbb{P}_{g_{\varepsilon_k^*}}}{d\mathbb{P}_{g_\varepsilon}}$$

then for any estimator $\hat{f}$ there exists a constant $c > 0$ depending only on ϕ, ψ and N such that :

$$U_j = max_{g_\varepsilon} \mathbb{E}_{g_\epsilon}(\|\hat{f} - g_\varepsilon\|_p^p) \geq c e^{-\lambda} \gamma_j^p 2^{\frac{jp}{2}} p_0.$$

Proof of Lemma 3.3. Let us set:

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

and

$$V_{j,k}^2 = \int_{S_{j,k}} |\hat{f}(x) + \gamma_j \varepsilon_k \eta_{j,k}^{-1} \psi_{j,k}(x)|^p dx$$

As the $S_{j,k}$ are disjoint, for any positive sequence $(\delta_{j,k})_{k\in R_j}$ we have:

$$U_j \geq \frac{1}{card(\mathcal{G})}\sum_{k\in R_j}\sum_{\substack{\varepsilon_i\in\{-1,+1\}\\ i\neq k}} \mathbb{E}_{g_\varepsilon}(\delta^p_{j,k}1_{\{V^1_{j,k}\geq\delta^p_{j,k}\}}) \tag{11}$$

$$+ \wedge_n(g_{\varepsilon^*_k},g_\epsilon)\delta^p_{j,k}1_{\{V^2_{j,k}\geq\delta^p_{j,k}\}}). \tag{12}$$

For precise mathematical statements see the book of Härdle, Picard, Kerkyacharian and Tsybakov (1998) and the book of Tsybakov (2004). Now let us consider the sequence $\delta_{j,k}$ defined as:

$$\delta_{j,k} = \gamma_j 2^{j(\frac{1}{2}-\frac{1}{p})}\eta^{-1}_{j,k}\|\psi\|_p.$$

The Minkowsky inequality yields:

$$(V^2_{j,k})^{\frac{1}{p}} + (V^1_{j,k})^{\frac{1}{p}} \geq (\int_{S_{j,k}} |2\gamma_j\eta^{-1}_{j,k}\psi_{j,k}(x)|^p dx)^{\frac{1}{p}} = 2\delta_{j,k}.$$

So:

$$1_{\{V^2_{j,k}\geq\delta^p_{j,k}\}} \geq 1_{\{V^1_{j,k}\leq\delta^p_{j,k}\}}. \tag{13}$$

Using the Markov inequality, we have:

$$min(\mathbb{E}_{g_\epsilon}(\wedge_n(g_{\varepsilon^*_k},g_\epsilon)),1) \geq e^{-\lambda}\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon^*_k},g_\epsilon)\geq e^{-\lambda}), \quad \forall\lambda>0. \tag{14}$$

Combining (11), (11) and (14) we obtain:

$$U_j \geq \frac{1}{card(\mathcal{G})}\sum_{k\in R_j}\sum_{\substack{\varepsilon_i\in\{-1,+1\}\\ i\neq k}} \delta^p_{j,k}e^{-\lambda}\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon^*_k},g_\epsilon)\geq e^{-\lambda})$$

Since Proposition 3.2 implies:

$$\sum_{k\in R_j}\delta^p_{j,k} = \gamma^p_j 2^{\frac{jp}{2}-j}\|\psi\|^p_p\sum_{k\in R_j}\eta^{-p}_{j,k} \geq c\gamma^p_j 2^{\frac{jp}{2}} \tag{15}$$

and $Card(\mathcal{G}) = 2Card(\varepsilon_i\in\{-1,+1\}, i\neq k)$, we deduce:

$$U_j \geq c\gamma^p_j\frac{e^{-\lambda}}{2}2^{\frac{jp}{2}}p_0.$$

□

Proposition 3.1. *If we chose $\gamma_j = n^{-\frac{1}{2}}$ then there exist $\lambda>0$ and $p_0>0$ such that:*

$$\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon^*_k},g_\epsilon) > e^{-\lambda}) \geq p_0, \qquad \forall\epsilon\in\{-1,1\}, \quad n\in\mathbb{N}^*.$$

The proof is given in Appendix.
Injecting the result of this proposition in Lemma 3.3 with $\gamma_j = n^{-\frac{1}{2}} \simeq 2^{-j(s+\frac{1}{2})}$ (i.e $2^j \simeq n^{\frac{1}{1+2s}}$), one gets:

$$\inf_{\hat{f}}\sup_{f\in b_{s,p,\infty}(L)} \mathbb{E}(\|\hat{f}-f\|^p_p) \geq c\gamma^p_j\frac{e^{-\lambda}}{2}2^{\frac{jp}{2}}p_0 = c'(\frac{1}{\sqrt{n}}2^{\frac{j}{2}})^p \geq c''n^{-\frac{sp}{1+2s}}.$$

This ends the proof of Theorem 3.3. □

Combining Theorem 3.2 and Theorem 3.3, we obtain Theorem 3.1.

Remark 3.1. *Thanks to Proposition 2.5, we can extend the result of Theorem 3.1 to $B_{s,\pi,r}(L)$ for $\pi\geq p$, $r>0$ and $N>s>0$.*

Remark 3.2. *If v is a positive constant then we obtain a traditional minimax result.*

Example 3.1. *Let $p\geq 2$. It is clear that $v(t) = t^{-\frac{\sigma}{2}}$ and $\frac{1}{v}$ belong to $\mathbb{L}^p([0,1])$. Thus, we can apply Theorem 3.1.*

The following subsection proposes to investigate the minimax rate over $b_{s,p,\infty}(L)$ under $\mathbb{L}^p$ loss for other assumptions on the variance function.

3.2 When v or $\frac{1}{v}$ does not belong to $\mathbb{L}^p$ for $p > 2$

This part is center around the following question:

Question 3.1. *Can we have the same minimax rate than Theorem 3.1 over $b_{s,p,\infty}(L)$ for any functions v which do not satisfy the asssumptions 'v and $\frac{1}{v}$ belong to $\mathbb{L}^p$ for $p > 2$' ?*

The answer is contained into the following theorem:

Theorem 3.4. *Let $2 < p < \infty$. Let us consider the function $v(t) = t^{-\frac{\sigma}{2}}$ for $\frac{2}{p} < \sigma < 1$. Then for $s > 0$ we have the following minimax result:*

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

where

$$\tilde{\alpha} = \frac{sp}{2s + 1 + \sigma - \frac{2}{p}}.$$

Proof of Theorem 3.4. Let us introduce the following lemma:

Lemma 3.4. *Let $p > 2$. Let us consider $\eta_{j,k}$ defined in (10) and $\rho_{j,k}$ defined in (5) where the function v is defined in Theorem 3.4. Then there exist two constant $C > 0$ and $c > 0$ such that:*

$$c 2^{\frac{j\sigma p}{2}} \leq \sum_{k \in R_j} \eta_{j,k}^{-p} \leq \sum_{k \in R_j} \rho_{j,k}^p \leq C 2^{\frac{j\sigma p}{2}}. \tag{16}$$

The proof is given in the Appendix.

3.2.1 Upper bound

Let us consider the linear estimator $\hat{f}^l$ defined in (1) where v is function defined in Theorem 3.4. Putting the inequality (16) into (7), one gets:

$$\sup_{f \in b_{s,p,\infty}(L)} \mathbb{E}(\|\hat{f}^l - f\|_p^p) \leq C(2^{j(n)(\frac{p}{2} - 1 + \frac{\sigma p}{2})} n^{-\frac{p}{2}} + 2^{-j(n)sp}) \leq C' n^{-\frac{sp}{1+2s+\sigma-\frac{2}{p}}}$$

for $j(n)$ the integer satisfying $2^{j(n)} \simeq n^{\frac{1}{1+2s+\sigma-\frac{2}{p}}}$.

3.2.2 Lower bound

Let us consider the function g_ϵ defined in (9) where $\gamma_j = n^{-\frac{1}{2}}$ and $v(t) = t^{-\frac{\sigma}{2}}$ with $\frac{2}{p} < \sigma < 1$. Using the inequality (16), we have:

$$\|g_\varepsilon\|_{s,p,\infty}^1 \leq n^{-\frac{1}{2}} 2^{j(s+\frac{1}{2})} (\sum_{k \in R_j} \eta_{j,k}^{-p} 2^{-j})^{\frac{1}{p}} \leq C n^{-\frac{1}{2}} 2^{j(s+\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{p})}.$$

So if we chose the integer j such that $n^{-\frac{1}{2}} \simeq 2^{-j(s+\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{p})}$ (i.e $n^{\frac{1}{2s+1+\sigma-\frac{2}{p}}} \simeq 2^j$) then g_ϵ belongs to $b_{s,p,\infty}(L)$. Injecting inequality (16) into (15) and using arguments similar to the proof of Theorem 3.3, we justify the existence of a constant $c > 0$ such that:

$$\inf_{\hat{f}} \sup_{f \in b_{s,p,\infty}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \geq c(n^{-\frac{1}{2}} 2^{j(\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{p})})^p \geq c' n^{-\frac{sp}{2s+1+\sigma-\frac{2}{p}}}.$$

This ends the proof of Theorem 3.4. □

So we have prove that if the variance function v does not belong to $\mathbb{L}^p$ then the mininax rate over usual Besov bodies under $\mathbb{L}^p$ risk can be slower than $n^{-\alpha}$ where $\alpha = \frac{sp}{1+2s}$. In particular, Theorem 3.4 shows that this rate of convergence can truly depends on the nature of v.

This arises a new question:

Question 3.2. *Can we find functional spaces over which the minimax rate under the $\mathbb{L}^p$ risk stay 'stable' for the functions v do not necessarily belong to $\mathbb{L}^p$ for $p > 2$?*

The answer is developed in the following section.

4 Minimax study over weighted Besov spaces

This section is focused on the proof of the following Theorem.

Theorem 4.1. *Let $2 \leq p < \infty$, $s > 0$. Assume that the functions v and $\frac{1}{v}$ belongs to $\mathbb{L}^2([0,1])$ with $\|\frac{1}{v}\|_2^2 = 1$, the primitive function $G(t) = \int_0^t \frac{1}{v^2(y)}dy$ is bijective and the function $v^2(G^{-1}(.))$ satisfies the $\mathcal{A}_p$ condition. Then we have the following minimax rate:*

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

where

$$\alpha = \frac{sp}{1+2s}.$$

As we shall see in the following proof, the $\mathcal{A}_p$ condition is truly at the heart of this result.

4.1 Proof of Theorem 4.1: upper bound and lower bound

4.1.1 Upper bound

Here we proceed as in Section 3 by taking in account that we work with the warped wavelet basis ξ^G.

Theorem 4.2. *We follow the assumptions and the notations of Theorem 4.1. Let us consider $\hat{f}^l$ the linear estimator defined by:*

$$\hat{f}^l(x) = \sum_{k \in \Delta_{j(n)}} \hat{\alpha}_{j(n),k}\phi_{j(n),k}(G(x)), \quad \hat{\alpha}_{j(n),k} = \int_0^1 \phi_{j(n),k}(G(t))\frac{1}{v^2(t)}dY_t.$$

Then for $s > 0$ we obtain the following upper bound:

$$\sup_{f \in b^G_{s,p,\infty}(L)} \mathbb{E}(\|\hat{f}^l - f\|_p^p) \leq Cn^{-\alpha}$$

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

Proof of Theorem 4.2. Using the Minkowski inequality and the inequality (1), the $\mathbb{L}^p$ risk of $\hat{f}^l$ can be decomposed as follows:

$$\begin{aligned}\mathbb{E}(\|\hat{f}^l - f\|_p^p) &\leq C(\mathbb{E}(\|\hat{f}^l - P^G_{j(n)}(f)\|_p^p) + \|P^G_{j(n)}(f) - f\|_p^p) \\ &\leq C(Q_1 + Q_2). \end{aligned} \tag{1}$$

Since conditions of Lemma 2.3 are satisfied with $T = G$ and $w = v^2(G^{-1}(.))$, we can expand f on the warped wavelet basis ξ^G at level $\nu = j(n)$. Assuming that $f \in b^G_{s,p,\infty}(L)$, Lemma 2.4 yields:

$$\begin{aligned} Q_2 &= \|\sum_{j \geq j(n)}\sum_{k \in \Delta_j} \beta^G_{j,k}\psi_{j,k}(G(.))\|_p^p \leq C(\sum_{j \geq j(n)} 2^{\frac{j}{2}}(\sum_{k \in \Delta_j} |\beta^G_{j,k}|^p w(I_{j,k}))^{\frac{1}{p}})^p \\ &\leq C(L\sum_{j \geq j(n)} 2^{-js})^p = C'2^{-j(n)sp}(\sum_{j \geq 0} 2^{-js})^p \leq C''2^{-j(n)sp}. \end{aligned} \tag{2}$$

Using the definition of $\hat{f}^l$ and Lemma 2.4, one gets:

$$\begin{aligned} Q_1 &= \mathbb{E}(\|\sum_{k \in \Delta_{j(n)}} (\hat{\alpha}_{j(n),k} - \alpha^G_{j(n),k})\phi_{j(n),k}(G(.))\|_p^p) \\ &\leq C(2^{\frac{j(n)p}{2}}(\sum_{k \in \Delta_{j(n)}} \mathbb{E}(|\hat{\alpha}_{j(n),k} - \alpha^G_{j(n),k}|^p)w(I_{j(n),k})) \\ &= C2^{\frac{j(n)p}{2}}Q_1^*. \end{aligned} \tag{3}$$

Using the change of variable $y = G^{-1}(t)$, one gets :

$$\hat{\alpha}_{j(n),k} - \alpha^G_{j(n),k} = \frac{1}{\sqrt{n}}\int_0^1 \phi_{j(n),k}(G(t))\frac{1}{v(t)}dW_t \sim \mathcal{N}(0, \frac{1}{n}).$$

Applying the second point of Lemma 3.1 and considering the fact that $\sum_{k\in\Delta_{j(n)}} w(I_{j(n),k}) = w([0,1]) = 1$, we obtain:

$$Q_1^* \leq n^{-\frac{p}{2}} \sum_{k\in\Delta_{j(n)}} w(I_{j(n),k}) = n^{-\frac{p}{2}} \tag{4}$$

Combining (1), (2), (3), (4) and taking in account that $2^{j(n)} \simeq n^{\frac{1}{1+2s}}$, one gets:

$$\sup_{f\in b^G_{s,p,\infty}(L)} \mathbb{E}(\|\hat{f}^l - f\|_p^p) \leq C(2^{\frac{j(n)p}{2}} n^{-\frac{p}{2}} + 2^{-j(n)sp}) \leq C' n^{-\frac{sp}{1+2s}}.$$

This completes the proof of Theorem 4.2. □

4.1.2 Lower Bound

Theorem 4.3. *Under the assumptions of Theorem 3.1, there exists a constant $c > 0$ such that for $s > 0$:*

$$\inf_{\hat{f}} \sup_{f\in b^G_{s,p,\infty}(L)} \mathbb{E}(\|\hat{f} - f\|_p^p) \geq cn^{-\alpha}.$$

The technique of the proof is similar to Section 3. As we shall see, the Muckenhoupt condition is determinant to proof this theorem.

Proof of Theorem 4.3. In the sequel, j denotes an integer to be chosen below and ε denotes a sequence such that $\varepsilon = (\varepsilon_k)_{k\in R_j} \in \{-1,+1\}^{\frac{2^j}{2(N-1)}}$ with $R_j = \{(2N-1)l - N;\ l = 1, 2, ..., \frac{2^j}{2(N-1)}\}$. Since G is increasing, for all k belongs to this set we have $S^G_{j,k} = Supp(\psi_{j,k}(G(.))) = [G^{-1}(\frac{k-N+1}{2^j}), G^{-1}(\frac{k+N}{2^j})]$ which verifies $S^G_{j,k} \cap S^G_{j,k'} = \emptyset$ for $k \neq k'$ and $\cup_{k\in R_j} S_{j,k} = [0, e_N] \subset [0,1]$. Let us consider

$$g_\varepsilon(x) = \gamma_j \sum_{k\in R_j} \varepsilon_k \psi_{j,k}(G(x)).$$

Let us denote by $\mathcal{G}$ the set of all such g_ε. The definition of $b^G_{s,p,\infty}(L)$ and the fact that $|\epsilon_i| \leq 1$ give us:

$$\|g_\varepsilon\|^G_{s,p,\infty} \leq \gamma_j 2^{j(s+\frac{1}{2})} (\sum_{k\in R_j} |\epsilon_k|^p w(I_{j,k}))^{\frac{1}{p}} \leq \gamma_j 2^{j(s+\frac{1}{2})} w([0,1]).$$

Thus, for j large, only the following constraint on γ_j is necessary to guarantee that $g_\varepsilon \in b^G_{s,p,\infty}(L)$:

$$\gamma_j \leq Lw([0,1])^{-1} 2^{-j(s+\frac{1}{2})}.$$

Let us now consider the following lemma.

Lemma 4.1. *For $\varepsilon \in \{-1,+1\}^{\frac{2^j}{2(N-1)}}$, put $\varepsilon_k^* = (\varepsilon_i')_{i\in R_j}$ such that $\varepsilon_i' = \varepsilon_i 1_{\{i\neq k\}} - \varepsilon_i 1_{\{i=k\}}$. If there exist $\lambda > 0$ and $p_0 > 0$ such that*

$$\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0, \quad n \in \mathbb{N}^*$$

where we wrote:

$$\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) = \frac{d\mathbb{P}_{g_{\varepsilon_k^*}}}{d\mathbb{P}_{g_\varepsilon}}$$

and if w verifies the $\mathcal{A}_p$ condition then for any estimator $\hat{f}$ there exists a constant $c > 0$ depending only on w, ϕ, ψ and N such that :

$$V_j = max_{g_\varepsilon} \mathbb{E}(\|\hat{f} - g_\varepsilon\|_p^p) \geq ce^{-\lambda} \gamma_j^p 2^{\frac{jp}{2}} p_0.$$

Proof of Lemma 4.1. Let us set:

$$W^1_{j,k} = \int_{S_{j,k}} |\hat{f}(G^{-1}(x)) - \gamma_j \varepsilon_k \psi_{j,k}(x)|^p w(x)dx$$

and

$$W_{j,k}^2 = \int_{S_{j,k}} |\hat{f}(G^{-1}(x)) + \gamma_j \varepsilon_k \psi_{j,k}(x)|^p w(x)dx.$$

As the $S_{j,k}^G$ are disjoint, for any positive sequence $(\delta_{j,k})_{k\in R_j}$ we have:

$$\begin{aligned}
V_j &\geq \frac{1}{card(\mathcal{G})} \sum_{\varepsilon} \mathbb{E}_{g_\varepsilon}(\|\hat{f} - g_\varepsilon\|_p^p) \\
&\geq \frac{1}{card(\mathcal{G})} \sum_{k\in R_j} \sum_{\varepsilon} \mathbb{E}_{g_\varepsilon}(\int_{S_{j,k}^G} |\hat{f}(x) - \gamma_j \epsilon_k \psi_{j,k}(G(x))|^p dx) \\
&= \frac{1}{card(\mathcal{G})} \sum_{k\in R_j} \sum_{\substack{\varepsilon_i\in\{-1,+1\} \\ i\neq k}} \mathbb{E}_{g_\varepsilon}(W_{j,k}^1 + \wedge_n(g_{\varepsilon_k^*}, g_\epsilon) W_{j,k}^2) \\
&\geq \frac{1}{card(\mathcal{G})} \sum_{k\in R_j} \sum_{\substack{\varepsilon_i\in\{-1,+1\} \\ i\neq k}} \mathbb{E}_{g_\varepsilon}(\delta_{j,k}^p 1_{\{W_{j,k}^1 \geq \delta_{j,k}^p\}}) \\
&+ \wedge_n(g_{\varepsilon_k^*}, g_\epsilon) \delta_{j,k}^p 1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}}). \qquad (5)
\end{aligned}$$

For further details, see the book Härdle, Picard, Kerkyacharian and Tsybakov (1998) and the book of Tsybakov (2004). Let us consider the sequence $\delta_{j,k}$ defined as:

$$\delta_{j,k} = \gamma_j (2N-1)^{-1} 2^{\frac{j}{2}} (w(S_{j,k})^{\frac{1}{p}} \|\psi\|_1.$$

Using the Minkowsky inequality in $\mathbb{L}_w^p$-norm, one gets:

$$(W_{j,k}^1)^{\frac{1}{p}} + (W_{j,k}^2)^{\frac{1}{p}} \geq 2(\int_{S_{j,k}} |\gamma_j \epsilon_k \psi_{j,k}(x)|^p w(x)dx)^{\frac{1}{p}}.$$

Since w satisfies the $\mathcal{A}_p$ condition, we have:

$$\begin{aligned}
(\int_{S_{j,k}} |\gamma_j \psi_{j,k}(x)|^p w(x)dx)^{\frac{1}{p}} &\geq \gamma_j (w(S_{j,k}))^{\frac{1}{p}} (\frac{1}{|S_{j,k}|} \int_{S_{j,k}} |\psi_{j,k}(x)|dx) \\
&\geq \gamma_j 2^{\frac{j}{2}} (2N-1)^{-1} (w(S_{j,k}))^{\frac{1}{p}} \|\psi\|_1 = \delta_{j,k}.
\end{aligned}$$

Starting from these previous inequalities, we deduce:

$$1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}} \geq 1_{\{W_{j,k}^1 \leq \delta_{j,k}^p\}}. \qquad (6)$$

Combining (5), (14) and (6) we obtain:

$$V_j \geq \frac{1}{card(\mathcal{G})} \sum_{k\in R_j} \sum_{\varepsilon_i\in\{-1,+1\}, i\neq k} \delta_{j,k}^p e^{-\lambda} \mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) \geq e^{-\lambda}).$$

Since

$$\sum_{k\in R_j} \delta_{j,k}^p = (2N-1)^{-p} \gamma_j^p 2^{\frac{jp}{2}} w([0, e_N]) \|\psi\|_1^p$$

and $Card(\mathcal{G}) = 2Card(\varepsilon_i \in \{-1,+1\}, i \neq k)$, one gets:

$$V_j \geq (2N-1)^{-p} \frac{\gamma_j^p}{2} e^{-\lambda} 2^{\frac{jp}{2}} w([0, e_N]) \|\psi\|_1^p p_0 = c e^{-\lambda} \gamma_j^p 2^{\frac{jp}{2}} p_0.$$

$\square$

Proposition 4.1. *If we choosing $\gamma_j = n^{-\frac{1}{2}}$ then there exist $\lambda > 0$ and p_0 such that:*

$$\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0, \quad \forall \epsilon, \quad n \in \mathbb{N}^*.$$

The proof is given in the Appendix.
Thus, by taking $\gamma_j = n^{-\frac{1}{2}} \simeq 2^{-j(s+\frac{1}{2})}$ (i.e $2^j \simeq n^{\frac{1}{1+2s}}$), Lemma 4.1 gives us:

$$\inf_{\hat{f}} \sup_{f \in b^G_{s,p,\infty}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) \geq c\gamma_j^p e^{-\lambda} 2^{\frac{jp}{2}} p_0 = c'(\frac{1}{\sqrt{n}} 2^{\frac{j}{2}})^p \geq c'' n^{-\frac{sp}{1+2s}}.$$

□

Finally, by combining Theorem 4.2 and Theorem 4.3 we deduce the minimax bound of Theorem 4.1.

Remark 4.1. *Thanks to the Proposition 2.5, we can extend the minimax result of Theorem 4.1 over $B^G_{s,\pi,r}(L)$ for $N > s \geq p$, $\pi \geq p$ and $r > 0$.*

Remark 4.2. *If v is a positive constant then we obtain a traditional minimax result.*

4.2 Remarks and examples

The following lemma proposes another definition of the assumption '$v^2(G^{-1}(.))$ verifies the $\mathcal{A}_p$ condition'.

Lemma 4.2 (About the Muckenhoupt property of Theorem 4.1). *Let $p > 1$ and q such that $\frac{1}{p} + \frac{1}{q} = 1$. Then $v^2(G^{-1}(.))$ verifies the $\mathcal{A}_p$ condition if and only if there exists a constant $C > 0$ such that:*

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

for any subintervals I of $[0,1]$.

Proof of Lemma 4.2. See Kerkyacharian and Picard (2004). □

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

Example 4.1. *Let us observe the model defined in (1) with $v(t) = (\sigma+1)^{-\frac{1}{2}} t^{-\frac{\sigma}{2}}$. It is clear that v and $\frac{1}{v}$ belongs to $\mathbb{L}^2([0,1])$. Moreover $G(x) = x^{\sigma+1}$, $G^{-1}(x) = x^{\frac{1}{\sigma+1}}$, $w(x) = v^2(G^{-1}(x)) = x^{\frac{-\sigma}{\sigma+1}}$ with $0 < -\frac{\sigma}{\sigma+1} < p-1$ for $1 > \sigma > \frac{1}{p} - 1$ so the function w satisfies the $\mathcal{A}_p$ condition. Thus all conditions are satisfied to apply Theorem 4.1.*

Let $p \geq 2$ and $s \geq 0$. The following table combines the results of Example 3.1 and 4.1 in the case where the previous function is not bounded from above (i.e $0 < \sigma < 1$):

Model (1) where	Space A	Hypothesis	$\inf_{\hat{f}} \sup_{f \in A} \mathbb{E}(\|\hat{f} - f\|_p^p)$
$v(t) = t^{-\frac{\sigma}{2}}$	$b^1_{s,p,\infty}(L)$	$0 < \sigma < \frac{2}{p}$	$\asymp n^{-\alpha}$, $\alpha = \frac{sp}{1+2s}$
	$b^1_{s,p,\infty}(L)$	$\frac{2}{p} < \sigma < 1$	$\asymp n^{-\beta}$, $\beta = \frac{2s}{1+2s+\sigma-\frac{2}{p}}$
	$b^G_{s,p,\infty}(L)$	$0 < \sigma < 1$	$\asymp n^{-\alpha}$, $\alpha = \frac{sp}{1+2s}$

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

Example 4.2 (Other interesting functions). *We can apply Theorem 4.1 with functions more 'complicated' such as $v(t) = (\ln(2))^{-(\sigma+1)}(1+t)^{\frac{1}{2}}(\ln(1+t))^{-\frac{\sigma}{2}}$ for $\frac{2}{p} < \sigma < 1$ or else, $v(t) = \frac{2}{\pi}(1-t^{2\sigma})^{\frac{1}{4}} t^{\frac{1-\sigma}{2}}$ with $\frac{2}{p} + 1 < \sigma < 2$. Remark that this last function is not bounded from above and below and does not belong to $\mathbb{L}^p([0,1])$ for $p > 2$.*

Thus we have shown that weighted Besov spaces give us stable minimax results for the variance functions v not belonging to $\mathbb{L}^p$ for $p > 2$. Starting from these results, we propose to investigate the performance of an adaptative procedure constructed on ξ^G over $B^G_{s,\pi,r}(L)$ when $\pi \geq p$, $r > 0$ and $s \geq 0$.

5 Hard thresholding procedure and warped wavelet bases

Among other things, we showed in the previous part that linear procedure (1) are optimal over weighted Besov spaces. This procedure is not adaptative, 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 in various risks (see Donoho and Johnstone (1995) and Johnstone (1998)). The following section focused on the performance of a hard thresholding procedure constructed on ξ^G over weighted Besov spaces.

Let us state the main results of this section:

Theorem 5.1. *Assume that assumptions of Theorem 4.1 holds. Let us consider the following hard thresholding estimator:*

$$\tilde{f}(x) = \sum_{j,k\in\Lambda_n^2} \hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|\geq\kappa\sqrt{\frac{\ln(n)}{n}}\}} \psi_{j,k}(G(x)), \quad \hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(G(t))\frac{1}{v^2(t)}dY_t,$$

where $\Lambda_n^2 = \{(j,k); j \leq j_2(n),\ k \in \Delta_j\}$ *for* $j_2(n)$ *the integer verifying*

$$2^{j_2(n)} \leq \frac{n}{\ln(n)} < 2^{j_2(n)+1}. \tag{1}$$

Then for $\kappa > 0$ *a large enough constant and* $s > 0$*, we have the following upper bound:*

$$f \in b^G_{s,p,\infty}(L) \Rightarrow \mathbb{E}(\|\tilde{f} - f\|_p^p) \leq C\left(\frac{\ln(n)}{n}\right)^\alpha$$

where

$$\alpha = \frac{sp}{1+2s}.$$

Proof of Theorem 5.1. Our strategy is the following : we exhibit the maxiset of the procedure $\tilde{f}$ and we show that $b^G_{s,p,\infty}(L)$ is included into it. To isolate such a space, we must verify five assumptions. Two on them concern the geometrical properties of ξ^T, one concerns a weight inequality and finally, two of them concern 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 a consequence of the following lemma:

Lemma 5.1. *Under the assumptions of Theorem 4.1, the basis* ξ^G

- *satisfies the Temlyakov 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:*

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

- *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:*

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

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

Proposition 2.4 and definition of $j_2(n)$ (see (1)) yield:

$$(\frac{\ln(n)}{n})^{\frac{p}{2}}\sum_{k\in\Lambda_n^2}\|\psi_{j,k}(G(.))\|_p^p \leq C(\frac{\ln(n)}{n})^{\frac{p}{2}}\sum_{j\leq j_2(n)}2^{\frac{jp}{2}} \leq C'(\frac{\ln(n)}{n})^{\frac{p}{2}}2^{\frac{j_2(n)p}{2}} \leq C''.$$

Thus, the weight condition holds.

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

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

We deduce that the statistical conditions are satisfied.

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

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

where we have set:

$$\mathcal{M}(p, \tilde{\nu}, G) = E_1 \cap E_2$$

where:

$$E_1 = \{f;\ \sup_{u>0} u^{(1-\tilde{\nu})p} \sum_{j\geq\tau-1} \sum_{k\in\Delta_j} 1_{\{|\beta^G_{j,k}|>u\}} \|\psi_{j,k}(G(.))\|_p^p < \infty\}$$

and

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

We conclude by the following lemma:

Lemma 5.2. *For $s \geq 0$ we have the following embedding:*

$$b^G_{s,p,\infty}(L) \subset \mathcal{M}(p, \frac{2s}{2s+1}, G)$$

The proof is given in Appendix.

□

Remark 5.1. *Using Proposition 2.5, this result can be extended to the weighted Besov spaces $B^G_{s,\pi,r}(L)$ with $\pi \geq p$, $r > 0$ and $s \geq 0$.*

Finally, we have proved that hard thresholding procedure defined in (1) achieves the minimax rate of convergence up to a logarithmic factor over the weighted Besov space $B^G_{s,p,\infty}(L)$.

6 Extension of the results to other statistical models

This section introduces two statistical models using linear operators over which we can reproduce the same study than previously.

6.1 Simple model extension

Consider the Gaussian white noise model in which we observe Gaussian processes $Y_t^{(1)}$ governed by the stochastic equation:

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

where v is a function submits to the assumptions of Theorem 4.1, H_v denotes the operator defined by $H_v(f)(.) = \frac{f(.)}{v(.)}$ and W_t is a standard Brownian motion on $[0,1]$. The function f is an unknown function of interest.

By remarking that $dY_t^{(1)} = \frac{1}{v(t)}dY_t$ where dY_t is defined in (1) we can establish results similar to those obtained in Theorem 4.1 and Theorem 5.1.

6.2 Warped Gaussian noise model

Consider the Gaussian white noise model in which we observe Gaussian processes $Y_t^{(2)}$ governed by the stochastic equation:

$$dY_t^{(2)} = K_G(f)(t)dt + \frac{1}{\sqrt{n}}dW_t, \quad n \in \mathbb{N}^*, \quad t \in [0,1],$$

where G is a known differentiable and bijective function with $G' = g$, K_G denotes the warped operator defined by $K_G(f)(.) = f(G^{-1}(.))$ and W_t is a standard Brownian motion on $[0,1]$. The function f is an unknown function of interest.

If we suppose that g belongs to $\mathbb{L}^1([0,1])$ with $\|g\|_1 = 1$ and $\frac{1}{g(G^{-1}(.))}$ satisfies the $\mathcal{A}_p$ condition then we can show results similar to those obtained in Theorem 4.1 and Theorem 5.1 by using the warped wavelet basis ξ^G and by condering $\hat{\beta}_{j,k}^{(2)} = \int_0^1 \psi_{j,k}(t)dY_t^{(2)}$ instead of $\hat{\beta}_{j,k}$.

Remark 6.1. *Such a function g is not necessarily bounded above or below. For instance, let us consider*

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

7 Appendix

This section contains proofs of Lemma 2.2, Lemma 3.1, Lemma 3.2, Proposition 3.1, Lemma 3.4, Proposition 4.1 and Lemma 5.2.

Proof of lemma 2.2. By construction of ξ, we have:

$$T = \sum_{k\in\Delta_j} |\phi_{j,k}(x)|^v = T_1 + T_2 + T_3$$

The compacity of ϕ gives us:

$$T_1 = \sum_{k=0}^{N-2} |\phi_{j,k}(x)|^v \leq (N-1)2^{\frac{jv}{2}}\|\phi\|_\infty^v$$

and

$$T_3 = \sum_{k=2^j-N+1}^{2^j-1} |\phi_{j,k}(x)|^v \leq (N-1)2^{\frac{jv}{2}}\|\phi\|_\infty^v.$$

For the second term, let us introduce the following sets:

$$S_{j,k,i} = [\frac{k-N+1+i}{2^j}, \frac{k-N+i+2}{2^j}] \quad \text{and} \quad \tilde{S_{j,i}} = [\frac{i}{2^j}, \frac{2^j+i-2N+2}{2^j}].$$

Thus,

$$\begin{aligned}
T_2 &= \sum_{k=N-1}^{2^j-N} |\phi_{j,k}(x)|^v = \sum_{k=N-1}^{2^j-N}\sum_{i=0}^{2N-2} |\phi_{j,k}(x)|^v 1_{S_{j,k,i}}(x) \\
&\leq 2^{\frac{jv}{2}}\|\phi\|_\infty^v \sum_{k=N-1}^{2^j-N}\sum_{i=0}^{2N-2} 1_{S_{j,k,i}}(x) = 2^{\frac{jv}{2}}\|\phi\|_\infty^v \sum_{i=0}^{2N-2} 1_{\tilde{S}_{j,i}}(x) \\
&\leq 2^{\frac{jv}{2}}\|\phi\|_\infty^v(2N-1).
\end{aligned}$$

We deduce that there exists $C > 0$ such that $T \leq C2^{\frac{jv}{2}}$. Of course, ϕ and ψ can be exchanged. This ends the proof of Lemma 2.2. □

Proof of Lemma 3.1. It is well known that if $N \sim \mathcal{N}\left(0,\sigma^2\right)$ then we have the concentration inequality $\mathbb{P}(|N| \geq x) \leq 2\exp(-\frac{x^2}{2\sigma^2})$. Therefore, for $\kappa \geq 2\sqrt{2p}$ and $n \geq 3$, we have:

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

Moreover, it is well known that if $N \sim \mathcal{N}(0,\sigma^2)$ then $\mathbb{E}(|N|^{2p}) = K\sigma^{2p}$ where $K = \frac{2^p}{\sqrt{\pi}}\int_0^{+\infty} x^{p-\frac{1}{2}}e^{-x}dx$. We deduce the existence of a constant $C > 0$ which satisfies $\mathbb{E}(|V_n|^p) \leq Cn^{-\frac{p}{2}}$. □

Proof of Lemma 3.2. Upper bound: Here we consider the $\rho_{j,k}$ defined in (5). Using Cauchy -Schwartz inequality and Hölder inequality with the measure $d\nu = \psi_{j,k}^2(t)dt$, one gets:

$$1 = (\int_0^1 \psi_{j,k}^2(t)dt)^p \leq \rho_{j,k}^p\eta_{j,k}^p \leq \int_0^1 v^p(t)\psi_{j,k}^2(t)dt\eta_{j,k}^p.$$

Using Lemma 2.2, we obtain:

$$\sum_{k\in R_j}\eta_{j,k}^{-p} \leq \int_0^1 v^p(t)\sum_{k\in\Delta_j}\psi_{j,k}^2(t)dt \leq C2^j\|v\|_p^p = C'2^j.$$

Lower bound: The Cauchy-Schwartz inequality yields:

$$c2^j = Card(R_j) = \sum_{k\in R_j}\eta_{j,k}^{\frac{p}{2}}\eta_{j,k}^{-\frac{p}{2}} \leq \sqrt{\sum_{k\in R_j}\eta_{j,k}^p}\sqrt{\sum_{k\in R_j}\eta_{j,k}^{-p}} \tag{1}$$

Using Hölder inequality with the measure $d\nu = \psi_{j,k}^2(t)dt$ and Lemma 2.2, we obtain:

$$\sum_{k\in R_j}\eta_{j,k}^p \leq \int_0^1 \frac{1}{v^p(t)}\sum_{k\in\Delta_j}\psi_{j,k}^2(t)dt \leq C2^j\|\frac{1}{v}\|_p^p = C'2^j. \tag{2}$$

Inequalities (1) and (2) yield:

$$\sum_{k\in R_j}\eta_{j,k}^{-p} \geq c2^j.$$

This completes the proof of Lemma 3.2. □

Proof of Proposition 3.1. Following Girsanov's theorem, one gets:

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

Under $\mathbb{P}_{g_\epsilon}$, we have :

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

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

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

Since $-2\eta_{j,k}^{-1}\epsilon_k\int_0^1 \frac{\psi_{j,k}(t)}{v(t)}dW_t \sim \mathcal{N}(0,4)$, this entail the existence of $\lambda > 0$ and $p_0 > 0$ such that $\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0$. □

Proof of Lemma 3.4. Upper bound: Let $v^2(S_{j,k}) = \int_{S_{j,k}} v^2(t)dt$. Since $t^{-\sigma}$ verifies the $\mathcal{A}_p$ condition for $\frac{2}{p} < \sigma < 1$, Lemma 2.1 implies the existence of a constant $C > 0$ such that $v^2(S_{j,k}) \leq Cv^2(I_{j,k})$. Thus

$$\sum_{k\in R_j}(\int_0^1 v^2(t)\psi_{j,k}^2(t)dt)^{\frac{p}{2}} \leq \|\psi\|_\infty^2 2^{\frac{jp}{2}}\sum_{k\in R_j}(v^2(S_{j,k}))^{\frac{p}{2}} \leq C'2^{\frac{jp}{2}}\sum_{k\in R_j}(v^2(I_{j,k}))^{\frac{p}{2}}.$$

Using the following elementary inequality

$$(1+z)^\alpha - 1 \leq \alpha z \text{ for } z \in [0,1],\ 0 < \alpha < 1,$$

one gets:

$$(v(I_{j,k}))^{\frac{p}{2}} = 2^{j(\sigma-1)\frac{p}{2}} k^{(-\sigma+1)\frac{p}{2}}((1+\frac{1}{k})^{-\sigma+1}-1)^{\frac{p}{2}} \leq (1-\sigma)2^{j(\sigma-1)\frac{p}{2}} k^{-\frac{\sigma p}{2}}.$$

Since $\sum_{k\geq 1} k^{-\beta} < \infty$ for $\beta > 1$, we deduce:

$$\sum_{k\in R_j} \rho_{j,k}^p \leq C' 2^{j\frac{\sigma p}{2}} \sum_{k\geq 1} k^{-\frac{\sigma p}{2}} \leq C'' 2^{j\frac{\sigma p}{2}}.$$

The Cauchy -Schwartz inequality gives us:

$$1 = (\int_0^1 \psi_{j,k}^2(t)dt)^p = (\int_0^1 \frac{v(t)}{v(t)} \psi_{j,k}(t)\psi_{j,k}(t)dt)^p \leq \rho_{j,k}^p \eta_{j,k}^p$$

so:

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

Lower bound: Since $-1 < -\sigma < -\frac{2}{p} < 0$, v^2 verifies the $\mathcal{A}_2$ condition (see Example 2.1). So:

$$\begin{aligned}
(\frac{1}{v^2(S_{j,k})} \int_{S_{j,k}} v^2(t)\psi_{j,k}^2(t)dt)^{\frac{1}{2}} &\geq c\frac{1}{|S_{j,k}|}\int_{S_{j,k}} |\psi_{j,k}(t)|dt \\
&= c(2N-1)^{-1} 2^{\frac{j}{2}} \|\psi\|_1 \\
&= c' 2^{\frac{j}{2}}.
\end{aligned}$$

Therefore:

$$\sum_{k\in R_j} \rho_{j,k}^p \geq c' 2^{\frac{jp}{2}} \sum_{k\in R_j} (v^2(S_{j,k}))^{\frac{p}{2}} \tag{3}$$

Since v^2 is decreasing, we have $\sup_{k\in R_j} v^2(S_{j,k}) = \int_0^{\frac{2N-1}{2^j}} t^{-\sigma}dt = (1-\sigma)^{-1}(2N-1)^{1-\sigma}2^{j(\sigma-1)}$. We deduce:

$$\begin{aligned}
2^{\frac{jp}{2}} \sum_{k\in R_j} (v^2(S_{j,k}))^{\frac{p}{2}} &\geq 2^{\frac{jp}{2}} (\sup_{k\in R_j} v^2(S_{j,k}))^{\frac{p}{2}} \\
&= c' 2^{\frac{jp}{2}} 2^{j(\sigma-1)\frac{p}{2}} \\
&= c' 2^{\frac{j\sigma p}{2}}. \qquad (4)
\end{aligned}$$

Using the compacity of ψ and the $\mathcal{A}_2$ condition with the function $h = v^{-2}$, one gets:

$$\eta_{j,k}^p \rho_{j,k}^p \leq \|\psi\|_\infty^{2p} (2^j \int_{S_{j,k}} v^2(t)dt)^{\frac{p}{2}} (2^j \int_{S_{j,k}} \frac{1}{v^2(t)}dt)^{\frac{p}{2}} \leq C. \tag{5}$$

Combining (3), (4) and (5), one gets:

$$\sum_{k\in R_j} \eta_{j,k}^{-p} \geq c \sum_{k\in R_j} \rho_{j,k}^p \geq c 2^{\frac{j\sigma p}{2}}.$$

This ends the proof of Lemma 3.4. □

Proof of Proposition 4.1. Following Girsanov's theorem, under $\mathbb{P}_{g_\epsilon}$ we have:

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

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

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

Since $-2\epsilon_k \int_0^1 \frac{\psi_{j,k}(G(t))}{v(t)}dW_t \sim \mathcal{N}(0,4)$, this entail the existence of $\lambda > 0$ and $p_0 > 0$ such that $\mathbb{P}_{g_\epsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq p_0$. □

Proof of Lemma 5.2. For the following embedding

$$b^G_{s,p,\infty}(L) \subset \{f\ ,\ \sup_{u>0} u^{\frac{p}{1+2s}} \sum_{j,k\in\Lambda} 1_{\{|\beta^G_{j,k}|>u\}} \|\psi_{j,k}(G(.))\|_p^p < \infty\}$$

see Picard and Kerkyacharian (2004).
Assume that f belongs to $b^G_{s,p,\infty}(L)$ for all $s \geq 0$. Using Lemma 2.4, for any $l \geq \tau - 1$ one gets:

$$\begin{aligned}
\|\sum_{j\geq l}\sum_{k\in\Delta_j} \beta^G_{j,k}\psi_{j,k}(G(.))\|_p 2^{\frac{ls}{1+2s}} &\leq \sum_{j\geq l}\|\sum_{k\in\Delta_j} \beta^G_{j,k}\psi_{j,k}(G(.))\|_p 2^{\frac{ls}{1+2s}} \\
&\leq C\sum_{j\geq l} 2^{\frac{j}{2}}(\sum_{k\in\Delta_j} |\beta^G_{j,k}|^p w(I_{j,k}))^{\frac{1}{p}} 2^{\frac{ls}{1+2s}} \\
&\leq CL\sum_{j\geq l} 2^{\frac{ls}{1+2s}-js} \leq C'\sum_{j\geq l} 2^{(l-j)s} \leq C''.
\end{aligned}$$

So:

$$b^G_{s,p,\infty}(L) \subseteq \{f\ ,\ \sup_{l>0} 2^{-\frac{lsp}{1+2s}} \|f - \sum_{j\geq l}\sum_{k\in\Delta_j} \beta^G_{j,k}\psi_{j,k}(G(.))\|_p^p < \infty\}.$$

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.

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.

Grama, 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 concentred 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. (2004). *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 ineqalities 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.

Stein, E. (1993). *Harmonic Analysis : Real Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton University Press.

Tsybakov, A.B. (2004) *Introduction à l'estimation non-parametrique.* Springer.