



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-00004331v5

HAL Id: hal-00004331

<https://hal.science/hal-00004331v5>

Preprint submitted on 15 Apr 2005 (v5), last revised 7 Jun 2005 (v9)

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 a Gaussian white noise model

Chesneau Christophe

April 15, 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 exhibit the minimax rate of convergence over usual Besov spaces and over weighted Besov spaces for a wide class of variance functions. We show via the maxiset approach that the natural hard thresholding procedure constructed on warped wavelet bases is close to the optimal over weighted Besov spaces.

Minimax, Muckenhoupt weights, Heteroscedastic Gaussian noise, warped wavelets, wavelet thresholding.

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

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

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 $[0, 1]$ with small \mathbb{L}^p risk

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

where \mathbb{E}_f is the expectation with respect the distribution \mathbb{P}_f of processes Y_t and \hat{f} denotes a measurable function on $[0, 1]$ with respect to the observations (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 , 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 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 variance function $v = \frac{\sigma}{\sqrt{g}}$ is asymptotically equivalent (in Le Cam's sense) to the observation of data $(Y_1, X_1), \dots, (Y_n, X_n)$ where

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

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). Quote also Efromovich and Pinsker (1996) who use the equivalence between the models (1) and (2) in order to estimate a smooth regression function under \mathbb{L}^2 .

In this paper, we consider the problem in the framework of wavelet analysis. Our study can be divided in two parts.

In a first part, we investigate the estimation of f from (1) over the usual Besov spaces denoted $B_{s,\pi,r}(L)$ under \mathbb{L}^p risk. We show that if $\pi \geq p$, v belongs to $\mathbb{L}^\pi([0, 1])$ and $\frac{1}{v}$ belongs to $\mathbb{L}^2([0, 1])$ then the minimax rate of convergence is of the form

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

So the following question naturally arises : can we obtain this minimax rate over such a space for any v which do not belong to \mathbb{L}^π for $\pi > 2$? 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 spaces constructed on a wavelet basis $\{\psi_{j,k}(G)\}$ warped by a factor G depending on v . Such spaces were introduced in analysis by Qui (1982) and were developed in statistics by Kerkycharian and Picard (2004). These authors establish good estimation results in a regression setting with random design (i.e (2) with $\sigma(\cdot) = 1$) 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 bases possess some interesting geometrical properties in \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, $\pi \geq p$, v and $\frac{1}{v}$ belong to \mathbb{L}^2 and if v is subject to a property of Muckenhoupt type then the minimax rate over weighted Besov spaces $B_{s,\pi,r}^G(L)$ defined starting from the primitive of $\frac{1}{v^2}$ is of the form

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

This space has the advantage to take in account the variance functions v which do not necessarily satisfy the assumption ' v belongs to $\mathbb{L}^\pi([0, 1])$ '.

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, 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

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

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 is 'near to the optimal' i.e 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 function spaces we shall need in the study. In Section 3 we investigate the minimax rate over usual Besov spaces and we show that it can truly depends on v . Section 4 investigates and discusses the minimax properties over the weighted Besov spaces. 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 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 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.

2.1 Muckenhoupt condition

Let us first recall the following notion of Muckenhoupt weight.

Definition 2.1 (Muckenhoupt condition). *Let $1 < p < \infty$ and q such that $\frac{1}{p} + \frac{1}{q} = 1$. A non-negative locally integrable function m is said to verify the \mathcal{A}_p condition (or belong to \mathcal{A}_p) 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*

$$\left(\frac{1}{|I|} \int_I |h(x)| dx\right) \leq C \left(\frac{1}{m(I)} \int_I |h(x)|^p m(x) dx\right)^{\frac{1}{p}} \quad (3)$$

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 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 (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 L_m^p 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*

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

Proof of Lemma 2.1. It suffices to apply (3) 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 set some important results that will be useful in the sequel of this paper.

Let N be an integer of the form 2^u where u denotes a 'reasonable' integer. 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

- 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 $l = 0, \dots, N - 1$. 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:

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

and

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

with

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

For each functions, 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 $[1 - (2N - 2)2^{-j}, 1]$.

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

We associate to this function the weight

$$w(\cdot) = \frac{1}{\tilde{T}(T^{-1}(\cdot))} \quad (4)$$

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

$$\int_0^1 z(T(x))dx = \int_0^1 z(x)w(x)dx. \quad (5)$$

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.

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

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

Let us recall some properties linked to ξ^T .

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

$$\sum_{k \in \Delta_j} |\phi_{j,k}(x)|^\nu \leq C 2^{\frac{j\nu}{2}}, \quad x \in [0, 1]. \quad (6)$$

This inequality are always true if we exchanged ϕ by ψ .

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*

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

These inequalities are always true if we exchanged ϕ by ψ .

For further details on this subsection see Picard and Kerkyacharian (2004).

2.3 Function spaces

For any measurable f defined on $[0, 1]$, we 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 $N > s > 0$, $1 < p, q \leq +\infty$. We say that a function f of $\mathbb{L}^\pi([0, 1])$ belongs to the weighted Besov spaces $B_{s,\pi,r}^T(L)$ if and only if

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

These spaces can be viewed as a generalization of the usual Besov spaces. The major advantage of these spaces is that they can be expressed in terms of wavelet coefficients.

Starting from the warped wavelet basis on the unit interval, if $w \in \mathcal{A}_p$ we have

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

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

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

Moreover, under the same conditions as previously, we have

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

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

Notations 2.1. If $T = I_d$, 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 $B_{s,\pi,r}^T(L) = B_{s,\pi,r}(L)$.

For further details, 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 differ 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 $\infty > p \geq 1$ and $\infty \geq \pi \geq p$. Assume we observe model (1) with v satisfying

$$(A1) \ v \in \mathbb{L}^{\max(\pi,2)}([0,1]) \quad \text{and} \quad (A2) \ \frac{1}{v} \in \mathbb{L}^2([0,1])$$

then for $N > s > 0$ we have

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

where

$$\alpha_1 = \frac{s}{1+2s}.$$

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 $\infty \geq p \geq 1$ and $\pi \geq p$. Assume (A1). Consider \hat{f}^l the linear estimator defined by

$$\hat{f}^l(x) = \sum_{k \in \Delta_{j(n)}} \hat{\alpha}_{j(n),k} \phi_{j(n),k}(x), \quad \hat{\alpha}_{j(n),k} = \int_0^1 \phi_{j(n),k}(t) dY_t. \quad (10)$$

Then for $N > s > 0$ we have

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

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

Proof of Theorem 3.2. Using the Hölder's inequality, for $\pi \geq p$ we have

$$\mathbb{E}_f(\|\hat{f}^l - f\|_p^p) \leq C(\mathbb{E}_f(\|\hat{f}^l - f\|_\pi)^\pi)^{\frac{p}{\pi}} \quad (11)$$

Using the 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}^π risk of \hat{f} can be decomposed as follows

$$\begin{aligned}\mathbb{E}_f(\|\hat{f}^l - f\|_\pi^\pi) &\leq C(\mathbb{E}_f(\|\hat{f}^l - P_{j(n)}(f)\|_\pi^\pi) + \|P_{j(n)}(f) - f\|_\pi^\pi) \\ &= C(S_1 + S_2).\end{aligned}\tag{12}$$

Since $f \in B_{s,\pi,r}(L) \subset B_{s,\pi,\infty}(L)$, the equivalence (9) gives us

$$S_2 \leq L2^{-j(n)s\pi} \leq Ln^{-\alpha_1\pi}.\tag{13}$$

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

$$\begin{aligned}S_1 &= \mathbb{E}_f(\|\sum_{k \in \Delta_{j(n)}} (\hat{\alpha}_{j(n),k} - \alpha_{j(n),k}) \phi_{j(n),k}(\cdot)\|_\pi^\pi) \\ &\leq C(2^{j(n)(\frac{\pi}{2}-1)}) \sum_{k \in \Delta_{j(n)}} \mathbb{E}_f(|\hat{\alpha}_{j(n),k} - \alpha_{j(n),k}|^\pi) \\ &= C2^{j(n)(\frac{\pi}{2}-1)} S_1^*.\end{aligned}\tag{14}$$

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

$$\rho_{j,k} = \sqrt{\int_0^1 v^2(t) \phi_{j,k}^2(t) dt}.\tag{15}$$

We have clearly

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

To study S_1^* , we need the following lemma which will be proved in the Appendix.

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*

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

Using the second point of Lemma 3.1, one gets

$$S_1^* \leq Cn^{-\frac{\pi}{2}} \sum_{k \in \Delta_{j(n)}} \rho_{j(n),k}^\pi.\tag{16}$$

In the case where $2 > \pi \geq 1$, the Hölder's inequality and Property 2.1 yield

$$\begin{aligned}\sum_{k \in \Delta_{j(n)}} \rho_{j(n),k}^\pi &\leq \left(\sum_{k \in \Delta_{j(n)}} \int_0^1 v^2(t) \phi_{j(n),k}^2(t) dt \right)^{\frac{\pi}{2}} \left(\sum_{k \in \Delta_{j(n),k}} 1 \right)^{1-\frac{\pi}{2}} \\ &\leq (2^{j(n)} \int_0^1 v^2(t) dt)^{\frac{\pi}{2}} 2^{j(n)(1-\frac{\pi}{2})} \\ &\leq C2^{j(n)}.\end{aligned}\tag{17}$$

In the case where $\infty \geq \pi \geq 2$, applying Hölder's inequality with the measure $d\nu = \phi_{j(n),k}^2(t) dt$ and using Property 2.1, one gets

$$\begin{aligned}\sum_{k \in \Delta_{j(n)}} \rho_{j(n),k}^\pi &\leq \int_0^1 v^\pi(t) \sum_{k \in \Delta_{j(n)}} \phi_{j(n),k}^2(t) dt \\ &\leq C2^{j(n)} \|v\|_\pi^\pi \\ &= C2^{j(n)}.\end{aligned}\tag{18}$$

Thus, considering (14), (17), (16) and (18) we obtain for $\infty \geq \pi \geq 1$,

$$S_1 \leq C2^{j(n)s\pi} n^{-\frac{\pi}{2}}.\tag{19}$$

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

$$\mathbb{E}_f(\|\hat{f}^l - f\|_\pi) \leq C(2^{\frac{j(n)\pi}{2}} n^{-\frac{\pi}{2}} + 2^{-j(n)s\pi}) \leq C'n^{-\frac{s\pi}{1+2s}}.$$

Considering (11), we deduce that for $\pi \geq p \geq 1$,

$$\sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f}^l - f\|_p) \leq Cn^{-\frac{sp}{1+2s}}.$$

This completes the proof of Theorem 3.2. \square

3.1.2 Lower Bound

Now, introduce a theorem which will be intensively used in the sequel.

Theorem 3.3. *Let \mathcal{F} be a function space. Let j a fixed integer, $(\omega_{j,k})_{k \in R_j}$ a fixed sequence and ε a sequence such that $\varepsilon = (\varepsilon_k)_{k \in R_j} \in \{-1, 1\}^{\frac{2^j}{2N}}$ where*

$$R_j = \{(2N-1)l - N; l = 1, 2, \dots, \frac{2^j}{2N}\}. \quad (20)$$

Consider the functions

$$g_\varepsilon(x) = \gamma_j \sum_{k \in R_j} \omega_{j,k} \varepsilon_k \psi_{j,k}(T(x)) \quad (21)$$

where γ_j is chosen in such a way that g_ε belongs to \mathcal{F} .

For such ε , put $\varepsilon_k^* = (\varepsilon'_i)_{i \in R_j}$ defined by

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

Then for any estimators \hat{f} we have

$$\begin{aligned} U_j &= \sup_{g_\varepsilon \in \mathcal{F}} \mathbb{E}_{g_\varepsilon}(\|\hat{f} - g_\varepsilon\|_p^p) \\ &\geq \frac{e^{-\lambda}}{2} \gamma_j^p \sum_{k \in R_j} \omega_{j,k}^p \inf_{\varepsilon_i \in \{-1, +1\}_{i \neq k}} \mathbb{P}_{g_\varepsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}) \|\psi_{j,k}(T(\cdot))\|_p^p \end{aligned} \quad (22)$$

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

$$\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) = \frac{d\mathbb{P}_{g_{\varepsilon_k^*}}}{d\mathbb{P}_{g_\varepsilon}}. \quad (23)$$

Proof of Theorem 3.3. Since T is increasing, for all k belonging to R_j we have the $S_{j,k}^T$'s defined by

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

which satisfy

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

and

$$\cup_{k \in R_j} S_{j,k} = [0, \frac{2N-1}{2N}] \subset [0, 1].$$

Denote by \mathcal{G} the set of all functions g_ε defined by (21). For any estimators \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 sequences $(\delta_{j,k})_{k \in R_j}$, we have

$$\begin{aligned} U_j &\geq \frac{1}{\text{card}(\mathcal{G})} \sum_{\varepsilon} \mathbb{E}_{g_{\varepsilon}} (\|\hat{f} - g_{\varepsilon}\|_p^p) \\ &\geq \frac{1}{\text{card}(\mathcal{G})} \sum_{k \in R_j} \sum_{\varepsilon} \mathbb{E}_{g_{\varepsilon}} \left(\int_{S_{j,k}^T} |\hat{f}(x) - \gamma_j \epsilon_k \omega_{j,k} \psi_{j,k}(T(x))|^p dx \right) \end{aligned} \quad (24)$$

By the definition of ϵ' and the fact that for all $k \in R_j$

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

we obtain

$$\begin{aligned} U_j &\geq \frac{1}{\text{card}(\mathcal{G})} \sum_{k \in R_j} \sum_{\substack{\varepsilon_i \in \{-1, +1\} \\ i \neq k}} \mathbb{E}_{g_{\varepsilon}} (W_{j,k}^1 + \wedge_n(g_{\varepsilon_k^*}, g_{\varepsilon}) W_{j,k}^2) \\ &\geq \frac{1}{2} \sum_{k \in R_j} \inf_{\substack{\varepsilon_i \in \{-1, +1\} \\ i \neq k}} \delta_{j,k}^p \mathbb{E}_{g_{\varepsilon}} (1_{\{W_{j,k}^1 \geq \delta_{j,k}^p\}} + e^{-\lambda} 1_{\{\wedge_n(g_{\varepsilon_k^*}, g_{\varepsilon}) > e^{-\lambda}\}} 1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}}) \\ &\geq \frac{e^{-\lambda}}{2} \sum_{k \in R_j} \inf_{\substack{\varepsilon_i \in \{-1, +1\} \\ i \neq k}} \delta_{j,k}^p \mathbb{E}_{g_{\varepsilon}} (1_{\{\wedge_n(g_{\varepsilon_k^*}, g_{\varepsilon}) > e^{-\lambda}\}} (1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}} + 1_{\{W_{j,k}^1 \geq \delta_{j,k}^p\}})). \end{aligned} \quad (25)$$

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

$$\delta_{j,k} = \gamma_j \omega_{j,k} \|\psi_{j,k}(T(\cdot))\|_p.$$

Using the Minkowsky's inequality, we see that

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

therefore

$$1_{\{W_{j,k}^2 \geq \delta_{j,k}^p\}} \geq 1_{\{W_{j,k}^1 \leq \delta_{j,k}^p\}}. \quad (26)$$

Putting (24), (25) and (26) together, we deduce that

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

□

Theorem 3.3 can be viewed as a generalization of a result which appeared in the book Härdle, Picard, Kerkyacharian and Tsybakov (1998).

Theorem 3.4. *Let $p \geq 1$. Assume (A2). Then there exists a positive constant $c > 0$ such that for $s > 0$, $\infty \geq \pi \geq 1$ 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 3.4. In the sequel, j denotes an integer to be chosen below. Consider the functions g_{ε} defined in Theorem 3.3 with

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

Using the definition of $B_{s,\pi,r}(L)$ and the fact than $|\epsilon_i| = 1$, one gets

$$\|g_{\varepsilon}\|_{s,\pi,r} \leq \gamma_j 2^{j(s+\frac{1}{2})} \left(\sum_{k \in R_j} 2^{-j} \right)^{\frac{1}{\pi}} \leq C \gamma_j 2^{j(s+\frac{1}{2})}.$$

Thus, for j large, only the following constraint on γ_j is necessary to guarantee that $g_{\varepsilon} \in B_{s,\pi,r}(L)$:

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

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 exists $\lambda > 0$ and $p_0 > 0$ not depending on n such that*

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

It follows from Lemma 3.2 and Theorem 3.3 that

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

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

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

This ends the proof of Theorem 3.4. □

Combining Theorem 3.2 and Theorem 3.4, we obtain Theorem 3.1.

Remark 3.1. *If v is a positive constant then we obtain an usual minimax result.*

Example 3.1. *Let $\infty > \pi \geq p \geq 1$. Consider $v(t) = t^{-\frac{\sigma}{2}}$ for $-1 < \sigma < \frac{2}{\pi}$. It is clear that conditions (A1) and (A2) hold. Therefore 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 assumptions on the variance function.

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

This part is center around the following question:

Question 3.1. *Can we have the same minimax rate than Theorem 3.1 over $B_{s, \pi, r}(L)$ for any functions v which do not satisfy the assumption ' v belongs to $\mathbb{L}^\pi([0, 1])$ '?*

The answer is contained into the following theorem:

Theorem 3.5. *Let $\pi \geq p > 2$. Assume that we observe model (1) with*

$$v(t) = t^{-\frac{\sigma}{2}} \quad \text{for } \frac{2}{p} < \sigma < 1. \quad (27)$$

Then for $N > s > 0$ and $\infty \geq r \geq 1$ we have

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

where

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

and

$$\tilde{\alpha} = \begin{cases} \frac{s}{2s+1} & \frac{\pi-p}{\pi(\sigma p-2)} \geq s, \\ \frac{s+\frac{1}{p}-\frac{1}{\pi}}{2s+1+\sigma-\frac{2}{\pi}} & s > \frac{\pi-p}{\pi(\sigma p-2)}. \end{cases} \quad (28)$$

Proof of Theorem 3.5. First, introduce the following lemma which will be proved in Appendix.

Lemma 3.3. *Let $\pi \geq p > 2$. Let us consider $\eta_{j,k}$ defined by*

$$\eta_{j,k} = \sqrt{\int_0^1 t^\sigma \psi_{j,k}^2(t) dt}. \quad (29)$$

and $\rho_{j,k}$ defined by (15). Then there exist two constant $C > 0$ and $c > 0$ such that

$$c 2^{\frac{j\sigma\pi}{2}} \leq \sum_{k \in R_j} \eta_{j,k}^{-\pi} \leq \sum_{k \in R_j} \rho_{j,k}^\pi \leq C 2^{\frac{j\sigma\pi}{2}}. \quad (30)$$

The proof is given in the Appendix.

3.2.1 Upper bound

Let us consider the linear estimator \hat{f}^l defined in (10) where v is defined by (27). Putting the inequality (30) into (16), one gets

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

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

3.2.2 Lower bound

Let us consider the function g_ϵ defined in Theorem 3.3 with

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

Using the inequality (30), we have

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

So if we chose the integer j such that $\gamma_j \simeq 2^{-j(s+\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{\pi})}$ (i.e $n^{\frac{1}{2s+1+\sigma-\frac{2}{\pi}}} \simeq 2^j$) then g_ϵ belongs to $B_{s,\pi,r}(L)$. Now, consider the following Lemma:

Lemma 3.4. *If we chose $\gamma_j = n^{-\frac{1}{2}}$ then there exist $\lambda > 0$ and $p_0 > 0$ independent 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 3.4 and Theorem 3.3 together, we obtain

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

Injecting inequality (30), one gets

$$\begin{aligned} \inf_{\hat{f}} \sup_{f \in B_{s,\pi,r}(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) &\geq c(n^{-\frac{1}{2}} 2^{j(\frac{1}{2}+\frac{\sigma}{2}-\frac{1}{p})})^p \\ &\geq c'n^{-\tilde{\alpha}p}. \end{aligned} \tag{31}$$

This ends the proof of Theorem 3.5. \square

Remark 3.2. *Under the assumptions of Theorem 3.5, it clear that for $\pi = p$, 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 the variance function v does not belong to \mathbb{L}^p then the minimax rate over usual Besov bodies under \mathbb{L}^p risk can be slower than $n^{-\alpha}$ for $\alpha = \frac{sp}{1+2s}$. In particular, Theorem 3.5 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 $\mathbb{L}^\pi([0,1])$?*

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 $\infty > p > 1$ and $\infty \geq \pi \geq p$. Assume that we observe model (1). Suppose that v satisfies*

$$(B1) \quad \frac{1}{v} \in \mathbb{L}^2([0, 1])$$

and that G defined by

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

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

$$(B2) \quad v^2(G^{-1}(\cdot)) \in \mathcal{A}_p.$$

Then for $N > s > q(w)$ (see (8)) we have

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

where

$$\alpha_1 = \frac{s}{1 + 2s}.$$

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. *Let $\infty > p > 1$ and $\infty \geq \pi \geq p$. Assume (B1) and (B2). 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 $N > s > q(w)$ and $\infty \geq \pi \geq p \geq 1$ we have

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

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

Proof of Theorem 4.2. Starting from the inequality (11), it suffices to consider the \mathbb{L}^π risk of \hat{f}^l . The Minkowski's inequality yield

$$\begin{aligned} \mathbb{E}_f(\|\hat{f}^l - f\|_\pi^\pi) &\leq C(\mathbb{E}_f(\|\hat{f}^l - P_{j(n)}^G(f)\|_\pi^\pi) + \|P_{j(n)}^G(f) - f\|_\pi^\pi) \\ &\leq C(Q_1 + Q_2). \end{aligned} \tag{32}$$

Note that conditions (B1) and (B2) imply that the equivalence (9) holds for $T = G$ and $w = v^2(G^{-1}(\cdot))$. By the definition of $B_{s,\pi,r}^G(L)$, we have

$$Q_2 \leq C 2^{-j(n)s\pi}. \tag{33}$$

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

$$\begin{aligned} Q_1 &= \mathbb{E}_f(\|\sum_{k \in \Delta_{j(n)}} (\hat{\alpha}_{j(n),k} - \alpha_{j(n),k}^G) \phi_{j(n),k}(G(\cdot))\|_\pi^\pi) \\ &\leq C(2^{\frac{j(n)\pi}{2}} (\sum_{k \in \Delta_{j(n)}} \mathbb{E}_f(|\hat{\alpha}_{j(n),k} - \alpha_{j(n),k}^G|^\pi) w(I_{j(n),k}))) \\ &= C 2^{\frac{j(n)\pi}{2}} Q_1^*. \end{aligned} \tag{34}$$

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

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

Applying the second point of Lemma 3.1 we obtain

$$Q_1^* \leq n^{-\frac{\pi}{2}} \sum_{k \in \Delta_{j(n)}} w(I_{j(n),k}) = n^{-\frac{\pi}{2}} w([0, 1]) = n^{-\frac{\pi}{2}}. \quad (35)$$

Combining (32), (33), (34), (35) and taking in account that $2^{j(n)} \simeq n^{\frac{1}{1+2s}}$, we deduce that

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

Using the inequality (11), it comes that

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

This completes the proof of Theorem 4.2. \square

4.1.2 Lower Bound

Theorem 4.3. *Let $p > 1$. Assume (B1) and (B2). Then there exists a constant $c > 0$ such that for $s > 0$, $\infty \geq \pi \geq 1$ and $\infty \geq r \geq 1$ we have*

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

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. Consider the functions g_ϵ with

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

The definition of $B_{s,\pi,r}^G(L)$, the fact that $|\epsilon_i| = 1$ and $w([0, 1]) = 1$ give us

$$\|g_\epsilon\|_{s,\pi,r}^G \leq \gamma_j 2^{j(s+\frac{1}{2})} \left(\sum_{k \in R_j} |\epsilon_k|^\pi w(I_{j,k}) \right)^{\frac{1}{\pi}} \leq \gamma_j 2^{j(s+\frac{1}{2})}.$$

Thus, for j large, only the following constraint on γ_j is necessary to guarantee that $g_\epsilon \in B_{s,\pi,r}^G(L)$

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

Now, introduce the following lemma which will be proved in Appendix.

Lemma 4.1. *If we choosing $\gamma_j = n^{-\frac{1}{2}}$ then there exist $\lambda > 0$ and p_0 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 3.3 give us

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

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

$$\begin{aligned} \|\psi_{j,k}(G)\|_p^p &\geq \left(\int_{S_{j,k}} |\psi_{j,k}(x)|^p w(x) dx \right) \\ &\geq c_w w(S_{j,k}) \left(\frac{1}{|S_{j,k}|} \int_{S_{j,k}} |\psi_{j,k}(x)| dx \right)^p \\ &\geq c_w 2^{\frac{jp}{2}} (2N-1)^{-p} w(S_{j,k}) \|\psi\|_1^p. \end{aligned}$$

Using the fact that $\sum_{k \in R_j} w(S_{j,k}) = w([0, e_N])$ where $e_N = \frac{2N-1}{2N}$ and that $n^{-\frac{1}{2}} \simeq 2^{-j(s+\frac{1}{2})}$ (i.e. $2^j \simeq n^{\frac{1}{1+2s}}$), we deduce that

$$\begin{aligned} \inf_{\hat{f}} \sup_{f \in B_{s,p,r}^G(L)} \mathbb{E}_f(\|\hat{f} - f\|_p^p) &\geq c_w (2N-1)^{-p} \frac{e^{-\lambda}}{2} e^{-\lambda} \left(\frac{2^{\frac{j}{2}}}{\sqrt{n}} \right)^p w([0, e_N]) \|\psi\|_1^p p_0 \\ &\geq cn^{-\alpha_1 p}. \end{aligned}$$

□

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 an usual minimax result.*

4.2 Remarks and examples

The following lemma proposes another definition of the assumption ' $v^2(G^{-1}(\cdot))$ 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 $v^2(G^{-1}(\cdot))$ verifies the \mathcal{A}_p condition if and only if 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 subintervals I of $[0, 1]$.

Proof of Lemma 4.2. See Kerkycharian 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 $p > 2$ and $s > 1$. The following table summarizes the results of Example 3.1 and 4.1.

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

Thus, 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 $\mathbb{L}^\pi([0, 1])$ for $\pi > 2$.*

Thus we have shown that weighted Besov spaces give us stable minimax results for the variance functions v not belonging to \mathbb{L}^π for $\pi > 2$. Starting from these results, we propose to investigate the performance of an adaptative procedure constructed on ξ^G over $B_{s,\pi,r}^G(L)$ 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 (32) 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 $B_{s,\pi,r}^G$.

Theorem 5.1. *Let $p > 1$. Assume (B1) and (B2). Let us consider the following hard thresholding estimator*

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

where

$$\hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(G(t)) \frac{1}{v^2(t)} dY_t,$$

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

$$2^{j_2(n)} \leq \frac{n}{\ln(n)} < 2^{j_2(n)+1}. \quad (36)$$

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

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

where

$$\alpha_1 = \frac{s}{1 + 2s}.$$

Proof of Theorem 5.1. Our strategy is the following. We exhibit the maxiset of the procedure \tilde{f} and we show that $B_{s,\pi,r}^G(L)$ is included into it for $\pi \geq p$ and $N > s > 0$. 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*

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

- *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 Picard and Kerkyacharian (2003). \square

Property 2.2 and the definition of $j_2(n)$ (see (36)) yield

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

Thus, the weight condition holds.

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

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

We deduce that the statistical conditions are satisfied.

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

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

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

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

We conclude by the following lemma

Lemma 5.2. *For $N > s \geq 0$ and $\pi \geq p$, we have the following embedding*

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

The proof is given in Appendix. □

Finally, we have proved that hard thresholding procedure defined in (36) achieves the minimax rate of convergence up to a logarithmic factor over the weighted Besov space $B_{s,\pi,r}^G(L)$.

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 $H_v(f)(\cdot) = \frac{f(\cdot)}{v(\cdot)}$ 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 $K_G(f)(\cdot) = f(G^{-1}(\cdot))$ 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 $\frac{1}{g(G^{-1}(\cdot))}$ 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 from above and below. See for instance

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

7 Appendix

This section contains proofs of Lemma 3.1, Lemma 3.2, Lemma 3.3, Lemma 3.4, Lemma 4.1 and Lemma 5.2.

Proof of Lemma 3.1. It is well known that if $N \sim \mathcal{N}(0, \sigma^2)$ 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

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

Moreover, it is well known that if $N \sim \mathcal{N}(0, \sigma^2)$ then $\mathbb{E}(|N|^{2p}) = K\sigma^{2p}$ 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 $\mathbb{E}(|V_n|^p) \leq Cn^{-\frac{p}{2}}$. \square

Proof of Lemma 3.2. Following Girsanov's theorem, one gets

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

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

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

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

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

We have the following embeddings

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

Applying Tchebitchev's inequality, one obtains

$$\mathbb{P}(\Lambda_n(g_{\varepsilon_k^*}, g_\varepsilon) \geq e^{-\lambda}) \geq 1 - \mathbb{P}(|U_{j,k} + V_{j,k}| > \lambda) \geq 1 - \frac{1}{\lambda}(\mathbb{E}(|U_{j,k}|) + \mathbb{E}(|V_{j,k}|)) \quad (37)$$

$$= 1 - \frac{1}{\lambda}(|U_{j,k}| + \frac{2}{\pi}\sqrt{|U_{j,k}|}) \quad (38)$$

The condition (A2) and Property 2.1 imply that

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

and

$$\sum_{k \in R_j} \sqrt{|U_{j,k}|} = \sqrt{\sum_{k \in R_j} |U_{j,k}|} \sqrt{\sum_{k \in R_j} 1} \leq C 2^j.$$

Therefore

$$\sum_{k \in R_j} \inf_{\varepsilon_i \in \{-1, +1\}} \mathbb{P}_{g_\epsilon}(\wedge_n (g_{\varepsilon_k^*}, g_\epsilon) > e^{-\lambda}) \geq c 2^j - \frac{c'}{\lambda} 2^j \geq c'' 2^j. \quad (39)$$

□

Proof of Lemma 3.3. Upper bound: Since $t^{-\sigma}$ is decreasing and $\sum_{k \geq 1} k^{-\beta} < \infty$ for $\beta > 1$, we have

$$\begin{aligned} \sum_{k \in R_j} \left(\int_{S_{j,k}} t^{-\sigma}(t) \psi_{j,k}^2(t) dt \right)^{\frac{\pi}{2}} &\leq \|\psi\|_\infty^\pi 2^{\frac{j\pi}{2}} \left(\int_0^{2^{N-1}} t^{-\sigma} dt \right)^{\frac{\pi}{2}} + \sum_{k \in R_j \setminus \{N-1\}} \left(\int_0^1 t^{-\sigma}(t) \psi_{j,k}^2(t) dt \right)^{\frac{\pi}{2}} \\ &\leq C (2^j)^{\frac{\sigma\pi}{2}} + 2^{j\frac{\sigma\pi}{2}} \sum_{k \in R_j \setminus \{N-1\}} (k - N + 1)^{-\frac{\sigma\pi}{2}} \\ &\leq C' 2^{j\frac{\sigma\pi}{2}}. \end{aligned}$$

Cauchy-Schwartz's inequality gives us

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

so

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

Lower bound: Since $t^{-\sigma}$ is decreasing, we have

$$c 2^{j\frac{\sigma\pi}{2}} \leq \sum_{k \in R_j} \left(\frac{k + N}{2^j} \right)^{-\frac{\sigma\pi}{2}} \leq \sum_{k \in R_j} \rho_{j,k}^\pi. \quad (40)$$

Since $-1 < -\sigma < -\frac{2}{p} < 0$, $t^{-\sigma}$ verifies the \mathcal{A}_2 condition (see Example 2.1), the inequality (3) with $h = t^\sigma$ gives us

$$\eta_{j,k}^\pi \rho_{j,k}^\pi \leq \|\psi\|_\infty^{2\pi} (2^j \int_{S_{j,k}} \frac{1}{t^\sigma} dt)^{\frac{\pi}{2}} (2^j \int_{S_{j,k}} t^\sigma dt)^{\frac{\pi}{2}} \leq C. \quad (41)$$

Combining (40) and (41), we deduce that

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

This ends the proof of Lemma 3.3. □

Proof of Lemma 3.4. Following Girsanov's theorem, one gets

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

Under \mathbb{P}_{g_ϵ} , we have

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

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

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

Since

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

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

$$\mathbb{P}_{g_\varepsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}) = \frac{1}{2}.$$

This finished the proof of Lemma 3.4. □

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

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

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

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

Since

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

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

$$\mathbb{P}_{g_\varepsilon}(\wedge_n(g_{\varepsilon_k^*}, g_\varepsilon) > e^{-\lambda}) = \frac{1}{2}.$$

This finished the proof of Lemma 4.1. □

Proof of Lemma 5.2. For the following embedding

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

see Picard and Kerkycharian (2004).

Assume that f belongs to $B_{s,p,\infty}^G(L)$ for all $s \geq 0$. Using Property 2.1 and Property 2.2, for any $l \geq \tau - 1$ one gets

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

So:

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

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.
- 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 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. (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 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.