# Regression in random design and wavelet block thresholding

Christophe Chesneau

# Regression in random design and wavelet block thresholding

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

***20 May 2006***

**Abstract**

The problem of estimating an unknown regression function in a regression setting with (known) random design is concerned. By adopting the minimax point of view, we explore the asymptotic performances of an adaptive estimator based on wavelet block thresholding under the global $\mathbb{L}^p$ risk ($p > 1$). We show that the estimator achieves the optimal rates of convergence over a wide range of Besov balls.

**Key Words:** Minimax, wavelet block thresholding, Regression with random design, Besov spaces.

**AMS subject classification:** Primary : 62G07 ; Secondary : 62C20.

## 1 Motivation

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

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

where the $z_i$'s are Gaussian i.i.d with mean zero, variance one and are independent of the design $X_1, ..., X_n$. We assume that the design is known with $X_i \in [0, 1]$ and we denote by $g$ the density of $X_1$. The function $f$ is unknown with $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)| < \infty$. The objective is to estimate $f$ from observations $(X_1, Y_1), ..., (X_n, Y_n)$.

When wavelet methods are considered, this estimation problem has been investigated by many authors in various statistical setting. Let us cite Delyon and Juditsky (1996), Cai and Brown (1998, 1999), Maxim (2003), to name a few. The standard approach consists in considering constructions based on term-by-term thresholding via the VisuShrink algorithm developed by Donoho and Johnstone (1995). Traditionaly, under some assumptions on $g$, such procedures attain the minimax rate of convergence up to a logarithmic factor over a wide class of Besov balls. If we consider the $\mathbb{L}^2$ risk that means that the hard thresholding procedure $\hat{f}$ satisfies :

$$\sup_{f\in B^s_{\pi,r}(R)} \mathbb{E}^n_f \left( \int_0^1 |\hat{f}(x) - f(x)|^2 dx \right) \leq C \left( \frac{\ln(n)}{n} \right)^{\frac{2s}{1+2s}}.$$

Here $\mathbb{E}^n_f$ denotes the expectation with respect to the distribution $\mathbb{P}^n_f$ of $(X_1, Y_1), ..., (X_n, Y_n)$ and $B^s_{\pi,r}(R)$ is a Besov ball (to be defined in Section 2). Let us mention that $\hat{f}$ is adaptive in the sense where its construction does not depend on the parameter of regularity $s$.

In order to remove the logarithmic term, an adaptive wavelet construction based on the block thresholding developed by Cai (1999, 2002) has been studied by Chicken (2003) in the context of uniform design (i.e $g = 1_{[0,1]}$). A such estimator $\hat{f}$ satisfies :

$$\sup_{f\in B^s_{\infty,\infty}(R)} \mathbb{E}^n_f \left( \int_0^1 |\hat{f}(x) - f(x)|^2 dx \right) \leq C n^{-\frac{2s}{1+2s}}.$$

Here $B^s_{\infty,\infty}(R)$ is a Hölder space of function (to be defined in Section 2). In addition to attain the optimal rate of convergence under the minimax context above, good performances are observed in practice.

This present paper proposes to investigate the minimax rates of convergence of the BlockShrink procedure under the $\mathbb{L}^p$ risk ($p > 1$) over a wide class of Besov balls $B^s_{\pi,r}(R)$. We show that under some assumptions on $g$, the proposed BlockShrink procedure attains the minimax rate of convergence for numerous values of $s$, $\pi$ and $r$. It is important to mention that the proof is based on a general theorem which can be found in Chesneau (2006a). In fact, we only need to prove a technical concentration inequality to determine the desired minimax results.

The rest of the paper is organized as follows. Section 2 describes wavelet bases on the interval, Besov balls and the BlockShrink construction adapted to the $\mathbb{L}^p$ risk. Their minimax performances over Besov balls and under the $\mathbb{L}^p$ risk are investigated in Section 3.

## 2 Methodology

Throughout this paper, we set :

$$\mathbb{L}^p([0,1]) = \Big\{f \text{ measurable on } [0,1]; \quad \|f\|_p^p = \int_0^1 |f(x)|^p dx < +\infty\Big\}.$$

The notation $a \asymp b$ means : there exist two constants $C > 0$ and $c > 0$ such that $cb \leq a \leq Cb$. The notations $a \wedge b$ and $a \vee b$ mean respectively : $\min(a, b)$ and $\max(a, b)$.

### 2.1 Wavelet bases and Besov spaces

We summarize in this subsection the basics on wavelet bases on the unit interval $[0, 1]$.

Let us focus on the wavelet basis of $[0, 1]$ described by Cohen et al. (1993) : We consider $\phi$ a "father" wavelet of a multiresolution analysis on $\mathbb{R}$ and $\psi$ the associated "mother" wavelet. Assume that $Supp(\phi) = Supp(\psi) = [1 - N, N]$ and $\int_{1-N}^N \phi(x)dx = 1$, $\int_{1-N}^N x^l \psi(x)dx = 0$ for $l = 0, ..., N - 1$.

Let us set :

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

Then there exists an integer $\tau$ satisfying $2^\tau \geq 2N$ such that the collection $\zeta$ defined by :

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

with an appropriate treatments at the boundaries, is an orthonormal basis of $\mathbb{L}^2([0, 1])$.

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

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

where $\alpha_{j,k} = \int_0^1 f(x)\phi_{j,k}(x)dx$, $\beta_{j,k} = \int_0^1 f(x)\psi_{j,k}(x)dx$ and $\Delta_j = \{0, ..., 2^j - 1\}$. The following concentration inéquality is standard : Let $p \geq 1$. For any sequence $u = (u_{j,k})_{j,k}$ and any $j \geq \tau$ there exists a constant $C > 0$ such that :

$$\|\sum_{k\in\Delta_j} u_{j,k}\phi_{j,k}\|_p^p \leq C2^{j(\frac{p}{2}-1)} \sum_{k\in\Delta_j} |u_{j,k}|^p. \tag{2.1}$$

This inequality holds if we exchange $\phi$ by $\psi$.

Let us now define the main function spaces of our study. Let $N \in \mathbb{N}^*$, $R > 0$, $1 \le r \le \infty$, $1 \le \pi \le \infty$ and $\pi^{-1} < s < N$. The equivalence below shows the link which exists between the Besov balls and the wavelet basis $\zeta$. See Meyer (1990) for a detailed proof.

$$f \in B^s_{\pi,r}(R) \Longleftrightarrow \begin{cases} (\sum_{j \ge \tau-1}(2^{j(s+\frac{1}{2}-\frac{1}{\pi})}(\sum_{k\in\Delta_j}|\beta_{j,k}|^\pi)^{\frac{1}{\pi}})^r)^{\frac{1}{r}} \le R < \infty & if \quad r<\infty, \\ \sup_{j\ge \tau-1} 2^{j(s+\frac{1}{2}-\frac{1}{\pi})}(\sum_{k\in\Delta_j}|\beta_{j,k}|^\pi)^{\frac{1}{\pi}} \le R < \infty & if \quad r=\infty. \end{cases}$$

## 2.2 BlockShrink procedures

Let us now present the adaptive procedures which will be at the heart of our statistical study.

**Definition 2.1** (BlockShrink procedures)**.** *Let $1 < p < \infty$. Let $j_1$ be an integer satisfying $2^{j_1} \asymp \ln(n)^{\frac{p}{2}}$ and let $j_2$ be an integer satisfying $2^{j_2} \asymp \sqrt{\frac{n}{\ln(n)}}$. For all $j$ in $\{j_1, ..., j_2\}$, let us divide $\Delta_j$ into consecutive nonoverlapping blocks $\mathcal{B}_{j,K}$ of length $\ln(n)^{\frac{p}{2}\vee 1}$ i.e :*

$$\mathcal{B}_{j,K} = \left\{k \in \Delta_j : \quad (K-1)\ln(n)^{\frac{p}{2}\vee 1} \le k \le K\ln(n)^{\frac{p}{2}\vee 1} - 1\right\}, \quad K \in \mathcal{A}_j,$$

*where the sets $\mathcal{A}_j$ are defined by :*

$$\mathcal{A}_j = \left\{1, ..., 2^j \ln(n)^{-\frac{p}{2}}\right\}$$

*(for convenience we suppose that $n^{\frac{1}{2}}\ln(n)^{-\frac{p+1}{2}} \in \mathbb{N}^*$). We define the BlockShrink procedure $\hat{f}$ by :*

$$\hat{f}(x) = \sum_{k\in\Delta_{j_1}} \hat{\alpha}_{j_1,k}\phi_{j_1,k}(x) + \sum_{j=j_1}^{j_2}\sum_{K\in\mathcal{A}_j}\sum_{k\in\mathcal{B}_{j,K}} \hat{\beta}_{j,k}1_{\left\{\hat{b}_{j,K}(p)\ge \kappa n^{-\frac{1}{2}}\right\}}\psi_{j,k}(x), \qquad x \in [0,1], \tag{2.2}$$

*where $\kappa$ is a positive real number, $\hat{\alpha}_{j,k}$ and $\hat{\beta}_{j,k}$ are the following unbiased estimators of $\alpha_{j,k}$ and $\beta_{j,k}$ :*

$$\hat{\alpha}_{j,k} = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{g(X_i)}\phi_{j,k}(X_i), \qquad \hat{\beta}_{j,k} = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{g(X_i)}\psi_{j,k}(X_i) \tag{2.3}$$

*and $\hat{b}_{j,K}(p)$ is the normalized $l_p$-norm of estimators $(\hat{\beta}_{j,k})_{k\in\mathcal{B}_{j,K}}$ i.e :*

$$\hat{b}_{j,K}(p) = \left(\frac{1}{\ln(n)^{\frac{p}{2}\vee 1}}\sum_{k\in\mathcal{B}_{j,K}}|\hat{\beta}_{j,k}|^p\right)^{\frac{1}{p}}.$$

The wavelet construction above can be viewed as an $\mathbb{L}^p$ version of the BlockShrink procedure developed by Cai (1999) for the $\mathbb{L}^2$ risk and the white noise model. Let us mention that it appears in Picard and Tribouley (2000) in another statistical context and for the local $\mathbb{L}^p$ risk.

# 3 Minimax results over Besov balls

Theorem 3.1 below exhibits the rates of convergence attained by the BlockShrink procedure describes in Definition 2.1 over Besov balls under the $\mathbb{L}^p$ risk.

**Theorem 3.1** (Minimax results over Besov balls)**.** *Let $1 < p < \infty$, $R > 0$ and $\alpha_1$, $\alpha_2$ and $\epsilon$ such that :*

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

*Suppose that we observe the data (1.1) and that there exists two constants $C_* > 0$ and $c_* > 0$ such that :*

$$\|f\|_\infty \le C_* < \infty \quad \text{and} \quad \inf_{x \in [0,1]} |g(x)| \ge c_* > 0.$$

*Then for $s > 2^{-1} + \pi^{-1}$, $r \ge 1$, there exists a constant $C > 0$ depending on $C_*$ and $c_*$ such that the BlockShrink procedure $\hat{f}$ described in Definition 2.1 satisfies :*

$$\sup_{f \in B^s_{\pi,r}(R)} \mathbb{E}^n_f(\|\hat{f} - f\|^p_p) \le \begin{cases} Cn^{-\alpha_1 p}, & \text{if} \quad \pi \ge p, \\ C(\frac{\ln(n)}{n})^{\alpha_1 p}, & \text{if} \quad \epsilon > 0, \ p > \pi, \\ C(\frac{\ln(n)}{n})^{\alpha_2 p}, & \text{if} \quad \epsilon < 0, \\ C(\frac{\ln(n)}{n})^{\alpha_2 p} \ln(n)^{(p - \frac{\pi}{r})_+}, & \text{if} \quad \epsilon = 0, \end{cases}$$

*for $\kappa$ and $n$ large enough.*

The rates of convergence exhibit in the previous theorem are optimal in the minimax sense, except for the case $\{\epsilon > 0, \ p > \pi\}$ where a logarithmic factor appeared.

Let us mention that :

1. our block thresholding procedure has better rates of convergence than the warped wavelet thresholding procedure $\hat{f}^*$ developed by Kerkyacharian and Picard (2005). According to Kerkyacharian and Picard (2005, Proposition 2), under the assumption $g$ known with $0 < m \le g \le M < \infty$, such 'warped' procedure $\hat{f}^*$ satisfies :

$$\sup_{f(G^{-1}) \in B^s_{\pi,r}(R)} \mathbb{E}^n_f(\|\hat{f}^* - f\|^p_p) \le \begin{cases} C(\frac{\ln(n)}{n})^{\alpha_1 p}, & \text{if} \quad \epsilon > 0, \\ C(\frac{\ln(n)}{n})^{\alpha_2 p}, & \text{if} \quad \epsilon < 0, \\ C(\frac{\ln(n)}{n})^{\alpha_2 p} \ln(n)^{(p - \frac{\pi}{r})_+}, & \text{if} \quad \epsilon = 0, \end{cases}$$

Let us notice that the regularity of $f(G^{-1})$ is concerned and we have supposed that $g$ is bounded from above contrary to the assumption of Theorem 3.1.

2. Theorem 3.1 can be viewed as a generalization of Chicken (2003, Theorem 2) which proves this result in the case where $p = 2$, $\pi = \infty$, $r = \infty$ and $g = 1_{[0,1]}$ (uniform design), the BlockShrink procedure satisfies :

$$\sup_{f \in B^s_{\infty,\infty}(R)} \mathbb{E}^n_f(\|\hat{f} - f\|^2_2) \le Cn^{-2\alpha_1},$$

for $\kappa$ and $n$ large enough.

## 3.1 Proofs

**Preleminary remarks :** For the sake of legibility, we shall adopt the following notations :

$$\sum_K = \sum_{K \in \mathcal{A}_j} \quad \text{and} \quad \sum_{(K)} = \sum_{k \in \mathcal{B}_{j,K}}.$$

According to the proof of Chesneau (2006a, Theorems 4.1 and Theorem 4.2), it suffices to show that the estimators $\hat{\alpha}_{j,k}$ and $\hat{\beta}_{j,k}$ satisfy the three following assumptions :

**Assumption 1.** There exists a constant $C > 0$ such that the following moments condition holds :

$$\sup_{k \in \Delta_j} \mathbb{E}_f^n(|\hat{\alpha}_{j_1,k} - \alpha_{j_1,k}|^p) \le C n^{-\frac{p}{2}},$$

**Assumption 2.** There exists a constant $C > 0$ such that the following moments condition holds :

$$\sup_{k \in \Delta_j} \mathbb{E}_f^n(|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}) \le C n^{-p}, \qquad j \in \{j_1, ..., j_2\},$$

**Assumption 3.** There exist constants $\kappa_1$ and $C > 0$ such that the following concentration condition holds :

$$\sup_{K \in \mathcal{A}_j} \mathbb{P}_f^n((\frac{1}{\ln(n)^{\frac{p}{2} \vee 1}} \sum_{(K)} |\hat{\beta}_{j,k} - \beta_{j,k}|^p)^{\frac{1}{p}} \ge \frac{\kappa_1}{2\sqrt{n}}) \le C n^{-p}, \qquad j \in \{j_1, ..., j_2\}.$$

The assumptions 1 and 2 have been shown in the proof of Theorem 3.2 of Chesneau (2006b). That's why we focus our attention on the assumption 3.

The proofs of all major statement are based on several auxiliary lemmas. First, let us set a version of Talagrand's inequality and a version of Cirelson's inequality.

**Lemma 3.1** (Talagrand (Talagrand (1994))**.** *Let $V_1, ..., V_n$ be independent and identically distributed random variables, let $\epsilon_1, ..., \epsilon_n$ be independent Rademacher variables, also independent of $V_1, ..., V_n$, and let $\mathcal{F}$ be a class of functions uniformly bounded by $T$. If*

- $\sup_{h \in \mathcal{F}} Var(h(V_1)) \le \nu$,

- $\mathbb{E}(\sup_{h \in \mathcal{F}} \sum_{i=1}^n \epsilon_i h(V_i)) \le nH$,

*then there exist universal constants $C_1$, $C_2$ such that for $r_n(h) = n^{-1} \sum_{i=1}^n h(V_i) - \mathbb{E}(h(V_1))$ and any $t > 0$ we have :*

$$\mathbb{P}(\sup_{h \in \mathcal{F}} r_n(h) \ge t + C_2 H) \le \exp(-nC_1 \min\left(\frac{t^2}{v}, \frac{t}{T}\right)).$$

**Lemma 3.2** (Cirelson, Ibragimov, Sudakov's inequality (Cirelson et al. (1976))**.** *Let $(\eta_t,\ t \in T)$ be a centered Gaussian process. If*

- $\mathbb{E}(\sup_{t \in T} \eta_t) \le N$,

- $\sup_{t \in T} Var(\eta_t) \le W$,

*then for any $c > 0$ we have :*

$$\mathbb{P}(\sup_{t \in T} \eta_t \ge c + N) \le \exp(-\frac{c^2}{2W}). \tag{3.1}$$

We are now in position to start the proof. Here and later, $C$ denotes a constant (independent of $n$ and $f$) which may be different from one term to the other. The notation "$|\mathbb{X}$"

means conditionaly to $X_1, ..., X_n$.

**Proof ofTheorem 3.1.** Since for $1 < p < 2$ and all $\kappa > 0$, the $l_p$ norm inequality gives us :

$$\mathbb{P}_f^n((\ln(n)^{-1}\sum_{(K)}|\hat{\beta}_{j,k}-\beta_{j,k}|^p)^{\frac{1}{p}} \geq \frac{\kappa}{2\sqrt{n}}) \leq \mathbb{P}_f^n((\ln(n)^{-1}\sum_{(K)}|\hat{\beta}_{j,k}-\beta_{j,k}|^2)^{\frac{1}{2}} \geq \frac{\kappa}{2\sqrt{n}}),$$

it suffices to show the assumption 3 in the case where $p \geq 2$. We have the decomposition

$$\hat{\beta}_{j,k} - \beta_{j,k} = A_{j,k} + B_{j,k}$$

where

$$A_{j,k} = \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_1) - \mathbb{E}_f^n(\frac{f(X_1)}{g(X_1)}\psi_{j,k}(X_1))$$

and

$$B_{j,k} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{g(X_i)}\psi_{j,k}(X_i)z_i.$$

Using Minkowski's inequality, one gets :

$$\mathbb{P}_f^n(\frac{1}{\sqrt{\ln(n)}}(\sum_{(K)}|\hat{\beta}_{j,k}-\beta_{j,k}|^p)^{\frac{1}{p}} \geq \frac{\kappa}{2\sqrt{n}}) \leq \mathcal{U} + \mathcal{V}$$

where we have set :

$$\mathcal{U} = \mathbb{P}_f^n(\frac{1}{\sqrt{\ln(n)}}(\sum_{(K)}|A_{j,k}|^p)^{\frac{1}{p}} \geq \frac{\kappa}{4\sqrt{n}}) \quad \text{and} \quad \mathcal{V} = \mathbb{P}_f^n(\frac{1}{\sqrt{\ln(n)}}(\sum_{(K)}|B_{j,k}|^p)^{\frac{1}{p}} \geq \frac{\kappa}{4\sqrt{n}}).$$

Let us analyze each term $\mathcal{U}$ and $\mathcal{V}$ in turn.

• **Bound for the term $\mathcal{U}$.** Here, the aim is to apply Talagrand's inequality (Lemma 3.1). Let us consider an infinite sequence

$$\mathcal{C}_q = \{a = (a_{j,k});\ \sum_{k\in\mathbb{Z}}|a_{j,k}|^q \leq 1\}$$

where $q$ is defined by the equation $p^{-1} + q^{-1} = 1$. Let us consider a class of functions

$$\mathcal{F} = \left\{ h;\ h(x) = \frac{f(x)}{g(x)}\sum_{(K)}a_{j,k}\psi_{j,k}(x), \quad a \in \mathcal{C}_q \right\}.$$

Using an argument of duality, it is easy to notice that we have :

$$(\sum_{(K)}|\hat{\beta}_{j,k}-\beta_{j,k}|^p)^{\frac{1}{p}} = \sup_{a\in\mathcal{C}_q}\sum_{(K)}a_{j,k}(\hat{\beta}_{j,k}-\beta_{j,k})(= \sup_{h\in\mathcal{F}} r_n(h)).$$

So we only need to find the values corresponding to $v$, $H$ and $T$ in Talagrand's inequality. Let us set a preliminary remark.

**Remark 3.1.** *Since $p \geq 2$, we have clearly $q \leq 2$ and by comparison between the $l_p$ norms :*

$$\sup_{a\in\mathcal{C}_q}(\sum_{(K)}|a_{j,k}|^2)^{\frac{1}{2}} \leq \sup_{a\in\mathcal{C}_q}(\sum_{(K)}|a_{j,k}|^q)^{\frac{1}{q}} \leq 1.$$

∘ *Value of $T$.* Using Hölder's inequality, the fact that $\|f\|_\infty < \infty$, $\inf_{x\in[0,1]} |g(x)| > 0$ and the concentration inéquality (2.1), for all $h \in \mathcal{F}$ one gets :

$$\begin{aligned} |h(x)| &\leq \frac{|f(x)|}{|g(x)|}(\sum_{(K)} |\psi_{j,k}(x)|^2)^{\frac{1}{2}} \sup_{a\in\mathcal{C}_q}(\sum_{(K)} |a_{j,k}|^2)^{\frac{1}{2}} \leq C(\sum_{(K)} |\psi_{j,k}(x)|^2)^{\frac{1}{2}} \\ &\leq C2^{\frac{j}{2}}, \qquad x \in [0,1]. \end{aligned}$$

We deduce that :

$$T = C2^{\frac{j}{2}}.$$

∘ *Value of $H$.* If $\epsilon_1, ..., \epsilon_n$ are independent Rademacher variables independent of $X_1, ..., X_n$ then by Hölder's inequality and Cauchy-Schwartz's inequality we obtain :

$$\begin{aligned} \mathbb{E}_f^n(\sup_{a\in\mathcal{C}_q}\sum_{i=1}^n(\sum_{(K)} a_{j,k}\epsilon_i\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_i))) &\leq \sup_{a\in\mathcal{C}_q}(\sum_{(K)} |a_{j,k}|^q)^{\frac{1}{q}}(\sum_{(K)}\mathbb{E}_f^n(|\sum_{i=1}^n\epsilon_i\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_i)|^p))^{\frac{1}{p}} \\ &\leq (\sum_{(K)}\mathbb{E}_f^n(|\sum_{i=1}^n\epsilon_i\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_i)|^p))^{\frac{1}{p}}. \end{aligned} \tag{3.2}$$

By virtue of Khintchine's inequality, we can bound the last term by :

$$\begin{aligned} \mathbb{E}_f^n(|\sum_{i=1}^n\epsilon_i\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_i)|^p) &= \mathbb{E}_f^n(\mathbb{E}_f^n(|\sum_{i=1}^n\epsilon_i\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_i)|^p|\mathbb{X})) \\ &\leq \mathbb{E}_f^n(|\sum_{i=1}^n\frac{f^2(X_i)}{g^2(X_i)}\psi_{j,k}^2(X_i))|^{\frac{p}{2}}) = W. \end{aligned} \tag{3.3}$$

For the sake of simplicity in exposition, let us set $N_i = \frac{f^2(X_i)}{g^2(X_i)}\psi_{j,k}^2(X_i)$. An elementary inequality of convexity gives us :

$$W \leq C(\mathbb{E}_f^n(|\sum_{i=1}^n(N_i - \mathbb{E}_f^n(N_1))|^{\frac{p}{2}}) + n^{\frac{p}{2}}\mathbb{E}_f^n(N_1)^{\frac{p}{2}}) = C(W_1 + W_2). \tag{3.4}$$

Let us analyze each term $W_1$ and $W_2$ in turn.

*Bound for the term $W_1$.* Using Bernstein's inequality with the i.i.d centered variables $N_1, ..., N_n$ and Cauchy-Schwartz's inequality, it comes :

$$\begin{aligned} W_1 &\leq C(n\mathbb{E}_f^n(|N_1 - \mathbb{E}_f^n(N_1)|^{\frac{p}{2}}) + (n\mathbb{E}_f^n(|N_1 - \mathbb{E}_f^n(N_1)|^2))^{\frac{p}{4}}) \\ &\leq C(n\mathbb{E}_f^n(|N_1|^{\frac{p}{2}}) + (n\mathbb{E}_f^n(|N_1|^2))^{\frac{p}{4}}). \end{aligned}$$

Since for any $m \geq 1$, any $j \in \{j_1, ..., j_2\}$ and under the assumptions of boundedness on $f$ and $g$ we have :

$$\begin{aligned} \mathbb{E}_f^n(|N_1|^m)) &= \int_0^1 \frac{|f(x)|^{2m}}{|g(x)|^{2m-1}}|\psi_{j,k}(x)|^{2m}dx \leq C\int_0^1 |\psi_{j,k}(x)|^{2m}dx \\ &\leq C2^{j(m-1)}\|\psi\|_\infty^{2m-2}\int_0^1 |\psi_{j,k}(x)|^2dx \leq C2^{j(m-1)} \leq C2^{j_2(m-1)} \leq Cn^{m-1}, \end{aligned}$$

we deduce that :

$$W_1 \leq Cn^{\frac{p}{2}}. \tag{3.5}$$

*Bound for the term $W_2$.* Since $\mathbb{E}_f^n(N_1) \le C$, we have clearly :

$$W_2 \le C n^{\frac{p}{2}}. \tag{3.6}$$

Combining (3.2)-(3.6), we see that :

$$\mathbb{E}_f^n(\sup_{a\in\mathcal{C}_q}\sum_{i=1}^n(\sum_{(K)} a_{j,k}\epsilon_i\frac{f(X_i)}{g(X_i)}\psi_{j,k}(X_i))) \le (\sum_{(K)} W)^{\frac{1}{p}} \le C n^{\frac{1}{2}} Card(\mathcal{B}_{j,K})^{\frac{1}{p}} \le C n^{\frac{1}{2}} \ln(n)^{\frac{1}{2}}.$$

It follows that :

$$H = C\sqrt{\frac{\ln(n)}{n}}.$$

∘ *Value of $v$.* Using the fact that $\|f\|_\infty < \infty$, $\inf_{x\in[0,1]}|g(x)| > 0$ and $q \le 2$, one gets :

$$\begin{aligned}
\sup_{h\in\mathcal{F}} Var(h(X_1)) &\le \sup_{a\in\mathcal{C}_q}\mathbb{E}_f^n(\frac{|f(X_1)|^2}{|g(X_1)|^2}|\sum_{(K)} a_{j,k}\psi_{j,k}(X_1)|^2)\\
&\le C\sup_{a\in\mathcal{C}_q}\mathbb{E}_f^n(\sum_{(K)}\sum_{(K)} a_{j,k}a_{j,k'}\frac{1}{g(X_1)}\psi_{j,k}(X_1)\psi_{j,k'}(X_1))\\
&\le C\sup_{a\in\mathcal{C}_q}\sum_{(K)}\sum_{(K)} a_{j,k}a_{j,k'}\int_0^1\psi_{j,k}(x)\psi_{j,k'}(x)dx = C\sup_{a\in\mathcal{C}_q}\sum_{(K)}|a_{j,k}|^2 \le C.
\end{aligned}$$

We conclude that

$$v = C.$$

Since $n2^j \le n2^{j_2} \le Cn^{\frac{3}{2}}\ln(n)^{-\frac{1}{2}}$, if we set $t = \frac{\kappa}{8}\sqrt{\frac{\ln(n)}{n}}$ then we have :

$$\min\left(\frac{t^2}{v},\frac{t}{T}\right) = \min\left(\frac{C\kappa^2}{64}\frac{\ln(n)}{n},\frac{C}{8}\sqrt{\frac{\ln(n)}{n2^j}}\right) \ge \frac{C\kappa^2}{64}\frac{\ln(n)}{n}$$

for $n$ large enough. Thus, for a such choice of $t$, Talagrand's inequality (see Lemma 3.1) gives us :

$$\begin{aligned}
\mathcal{U} &= \mathbb{P}_f^n(\frac{1}{\sqrt{\ln(n)}}(\sum_{(K)}|A_{j,k}|^p)^{\frac{1}{p}} \ge \frac{\kappa}{4\sqrt{n}}) \le \mathbb{P}_f^n(\frac{1}{\sqrt{\ln(n)}}(\sum_{(K)}|A_{j,k}|^p)^{\frac{1}{p}} \ge \frac{C}{\sqrt{n}}+\frac{\kappa}{8\sqrt{n}})\\
&\le \exp(-nC\kappa^2\frac{\ln(n)}{n}) \le n^{-p}
\end{aligned}$$

for $\kappa$ large enough. We obtain the desired bound for the term $\mathcal{U}$.

• **Bound for the term $\mathcal{V}$.** The Cirelson's inequality (see Lemma 3.2) will be at the heart of the present proof. As in the previous subsection, let us consider the set

$$\mathcal{C}_q = \left\{a = (a_{j,k});\ \sum_{(K)}|a_{j,k}|^q \le 1\right\}$$

where $q$ is the real number satisfying $q^{-1}+p^{-1} = 1$. Let us consider :

$$\mathcal{Z}(a) = \sum_{(K)} a_{j,k}B_{j,k}$$

and let us notice that an argument of duality gives us :

$$\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a) = (\sum_{(K)} |B_{j,k}|^p)^{\frac{1}{p}}.$$

The rest of the proof is inspired by some methods developed in Kerkyacharian and Picard (2005, Proof of Proposition 3). Let us set :

$$\mathcal{B}_\kappa = \left\{ |n^{-1}\sum_{i=1}^{n} \frac{1}{g(X_i)}|\psi_{j,k}(X_i)|^2 - 1| \geq \kappa \right\}.$$

Using Holder's inequality, the fact that conditionally on $(X_1, ..., X_n)$ we have

$$\mathcal{Z}(a) \sim \mathcal{N}(0, \frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{g^2(X_i)}|\psi_{j,k}(X_i)|^2)$$

and that $\inf_{x\in[0,1]} |g(x)| > 0$, on the set $\mathcal{B}_\kappa^c$ we have :

$$\begin{aligned}
N(\mathbb{X}) &= \mathbb{E}_f^n(\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a)|\mathbb{X}) \leq (\sum_{(K)} \mathbb{E}_f^n(|B_{j,k}|^p|\mathbb{X}))^{\frac{1}{p}} \leq C(\sum_{(K)}(\frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{g(X_i)}|\psi_{j,k}(X_i)|^2)^{\frac{p}{2}})^{\frac{1}{p}} \\
&= n^{-\frac{1}{2}}(\sum_{(K)}(\frac{1}{n}\sum_{i=1}^{n} \frac{1}{g(X_i)}|\psi_{j,k}(X_i)|^2 - 1 + 1)^{\frac{p}{2}})^{\frac{1}{p}} \\
&\leq Cn^{-\frac{1}{2}}((\sum_{(K)}(\kappa+1)^{\frac{p}{2}})^{\frac{1}{p}}) \leq Cn^{-\frac{1}{2}}(\kappa+1)^{\frac{1}{2}} Card(\mathcal{B}_{j,K})^{\frac{1}{p}} \leq C(\kappa+1)^{\frac{1}{2}}\sqrt{\frac{\ln(n)}{n}},
\end{aligned}$$

we deduce that :

$$\begin{aligned}
\mathbb{P}_f^n(\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a) \geq c + C(1+\kappa)^{\frac{1}{2}}\sqrt{\frac{\ln(n)}{n}}) &= \mathbb{E}_f^n(\mathbb{P}_f^n(\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a) \geq c + C(1+\kappa)^{\frac{1}{2}}\sqrt{\frac{\ln(n)}{n}}|\mathbb{X})) \\
&\leq \mathbb{E}_f^n(\mathbb{P}_f^n(\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a) \geq c + C(1+\kappa)^{\frac{1}{2}}\sqrt{\frac{\ln(n)}{n}}|\mathbb{X})(1_{\mathcal{B}_\kappa} + 1_{\mathcal{B}_\kappa^c})) \\
&\leq \mathbb{P}_f^n(\mathcal{B}_\kappa) + \mathbb{E}_f^n(\mathbb{P}_f^n(\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a) \geq c + N(\mathbb{X})|\mathbb{X})). \qquad (3.7)
\end{aligned}$$

Since, conditionally on $(X_1, ..., X_n)$, $\mathcal{Z}(a)$ is a centered Gaussian process, we are in position to apply Cirelson's inequality which gives us :

$$\mathbb{E}_f^n(\mathbb{P}_f^n(\sup_{a\in\mathcal{C}_q} \mathcal{Z}(a) \geq c + N(\mathbb{X}))|\mathbb{X})) \leq \mathbb{E}_f^n(\exp -(\frac{c^2}{2W(\mathbb{X})})) \qquad (3.8)$$

where $W(\mathbb{X}) = \sup_{a\in\mathcal{C}_q} Var_f^n(\mathcal{Z}(a)|\mathbb{X})$.

∘ *Bound for* $W(\mathbb{X})$. Let us consider the set :

$$\mathcal{A}_\kappa = \left\{ \sup_{a\in\mathcal{C}_q}(\sum_{k\in\mathcal{B}_{j,K}}\sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}(n^{-1}\sum_{i=1}^{n} \frac{1}{g(X_i)}\psi_{j,k}(X_i)\psi_{j,k'}(X_i)) - \sum_{(K)}|a_{j,k}|^2) \geq \kappa \right\}.$$

Since $q \leq 2$, on the set $\mathcal{A}_\kappa^c$ we have :

$$\begin{aligned}
G &= \sup_{a\in\mathcal{C}_q}\sum_{k\in\mathcal{B}_{j,K}}\sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}(n^{-1}\sum_{i=1}^{n} \frac{1}{g(X_i)}\psi_{j,k}(X_i)\psi_{j,k'}(X_i)) \\
&\leq \sup_{a\in\mathcal{C}_q}(\sum_{k\in\mathcal{B}_{j,K}}\sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}(n^{-1}\sum_{i=1}^{n} \frac{1}{g(X_i)}\psi_{j,k}(X_i)\psi_{j,k'}(X_i)) - \sum_{(K)}|a_{j,k}|^2) + \sup_{a\in\mathcal{C}_q}\sum_{(K)}|a_{j,k}|^2 \\
&\leq (\kappa+1).
\end{aligned}$$

Since $\mathbb{E}^n_f(B_{j,k}) = 0$ and $\mathbb{E}^n_f(z_i z_{i'}) = 1$ if $i = i'$ and 0 otherwise, $\|f\|_\infty < \infty$ and $\inf_{x\in[0,1]} |g(x)| > 0$, it follows from the previous inequality that on $\mathcal{A}^c_\kappa$ we have :

$$\begin{aligned}
W(\mathbb{X}) &= \sup_{a\in\mathcal{C}_q} Var^n_f(\mathcal{Z}(a)|\mathbb{X}) \leq \sup_{a\in\mathcal{C}_q} \mathbb{E}^n_f(\sum_{k\in\mathcal{B}_{j,K}} \sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}B_{j,k}B_{j,k'}|\mathbb{X}) \\
&\leq C\sup_{a\in\mathcal{C}_q}(n^{-2}\sum_{k\in\mathcal{B}_{j,K}} \sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}\sum_{i=1}^n\sum_{i'=1}^n \frac{1}{g(X_i)}\psi_{j,k}(X_i)\psi_{j,k'}(X_{i'})\mathbb{E}^n_f(z_i z_{i'})) \\
&= Cn^{-1}\sup_{a\in\mathcal{C}_q}\sum_{k\in\mathcal{B}_{j,K}} \sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}(n^{-1}\sum_{i=1}^n \frac{1}{g(X_i)}\psi_{j,k}(X_i)\psi_{j,k'}(X_i))) = Cn^{-1}G \\
&\leq Cn^{-1}(\kappa+1).
\end{aligned}$$

Therefore :

$$\begin{aligned}
\mathbb{E}^n_f(\exp-(\frac{c^2}{2W_1(\mathbb{X})})) &= \mathbb{E}^n_f(\exp-(\frac{c^2}{2W_1(\mathbb{X})})1_{\mathcal{A}_\kappa}) + \mathbb{E}^n_f(\exp-(\frac{c^2}{2W_1(\mathbb{X})})1_{\mathcal{A}^c_\kappa}) \\
&\leq \mathbb{P}^n_f(\mathcal{A}_\kappa) + \exp-(\frac{nc^2}{2(\kappa+1)}). \qquad (3.9)
\end{aligned}$$

Now, let us consider the following proposition which will be proved at the end of the proof.

**Proposition 3.1.** *There exists a constant $C > 0$ such that for $n$ and $\kappa$ large enough we have :*

$$\max(\mathbb{P}^n_f(\mathcal{A}_\kappa), \mathbb{P}^n_f(\mathcal{B}_\kappa)) \leq Cn^{-p}, \qquad j \in \{j_1, ..., j_2\}. \qquad (3.10)$$

Putting (3.7)-(3.10) together and taking $c = \frac{\kappa}{8}\sqrt{\frac{\ln(n)}{n}}$, we find :

$$\begin{aligned}
\mathcal{V} &= \mathbb{P}^n_f(\sup_{a\in\mathcal{C}_q}\mathcal{Z}(a) \geq \frac{\kappa}{4}\sqrt{\frac{\ln(n)}{n}}) \leq \mathbb{P}^n_f(\sup_{a\in\mathcal{C}_q}\mathcal{Z}(a) \geq \frac{\kappa}{8}\sqrt{\frac{\ln(n)}{n}} + C(1+\kappa)^{\frac{1}{2}}\sqrt{\frac{\ln(n)}{n}}) \\
&\leq C(\mathbb{P}^n_f(\mathcal{A}_\kappa) + \mathbb{P}^n_f(\mathcal{B}_\kappa) + \exp-(\frac{\ln(n)\kappa^2}{128(\kappa+1)})) \\
&\leq C(n^{-C\kappa^2} + n^{-C\frac{\kappa^2}{1+\kappa}}) \leq Cn^{-p}
\end{aligned}$$

for $\kappa$ large enough. Combining the bounds of $\mathcal{U}$ and $\mathcal{V}$, the proof of Theorem 3.1 is complete in the case where $p \geq 2$.

***Proof of Proposition 3.1.*** Let us investigate separately the bounds of $\mathbb{P}^n_f(\mathcal{B}_\kappa)$ and $\mathbb{P}^n_f(\mathcal{A}_\kappa)$.

- *Let us show that $\mathbb{P}^n_f(\mathcal{B}_\kappa) \leq Cn^{-p}$.* Since the variables $\frac{\psi^2_{j,k}(X_i)}{g(X_i)}$ are i.i.d with $\frac{\psi^2_{j,k}(X_i)}{g(X_i)} \leq \|\frac{1}{g}\|_\infty\|\psi\|^2_\infty 2^j$ and $\mathbb{E}^n_f\left(\frac{\psi^2_{j,k}(X_1)}{g(X_1)}\right) = 1$, Hoeffding's inequality yields :

$$\mathbb{P}^n_f(\mathcal{B}_\kappa) \leq 2\exp\left(-\frac{2n^2\kappa^2}{n\|\frac{1}{g}\|^2_\infty\|\psi\|^4_\infty 2^{2j}}\right) \leq Cn^{-p}$$

for $2^j \leq 2^{j_2} \leq C\sqrt{\frac{n}{\ln(n)}}$ and $\kappa$ large enough.

- *Let us show that $\mathbb{P}^n_f(\mathcal{A}_\kappa) \leq Cn^{-p}$.* The aim is to apply the Talagrand inequality. Let us consider the set of functions $\mathcal{F}'$ defined by :

$$\mathcal{F}' = \left\{h;\ h(x) = \frac{1}{g(x)}\sum_{k\in\mathcal{B}_{j,K}}\sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}\psi_{j,k}(x)\psi_{j,k'}(x),\quad a\in\mathcal{C}_q\right\}.$$

and let us notice that :

$$\sup_{a \in \mathcal{C}_q} \Big( \sum_{k \in \mathcal{B}_{j,K}} \sum_{k' \in \mathcal{B}_{j,K}} a_{j,k} a_{j,k'} (n^{-1} \sum_{i=1}^{n} \frac{1}{g(X_i)} \psi_{j,k}(X_i)\psi_{j,k'}(X_i)) - \sum_{(K)} |a_{j,k}|^2 \Big) = \sup_{h \in \mathcal{F}'} r_n(h)$$

where $r_n$ is defined as in Lemma 3.1. Thus, it suffices to determine the values of $T$, $H$ and $v$ present in the Talagrand inequality.

∘ *Value of T.* Using the Holder inequality, the fact that $\inf_{x \in [0,1]} |g(x)| \geq c > 0$ and the concentration inequality (2.1), for any $h \in \mathcal{F}'$ it comes :

$$|h(x)| \leq C \sum_{(K)} |a_{j,k}|^2 \sum_{(K)} |\psi_{j,k}(x)|^2 \leq C 2^j.$$

Hence $T = C2^j$.

∘ *Value of H.* The $l_2$ Holder inequality combined with the Hölder inequality and the fact that $q \leq 2$ give us :

$$\begin{aligned} S &= \mathbb{E}_f^n \Big( \sup_{a \in \mathcal{C}_q} \sum_{k \in \mathcal{B}_{j,K}} \sum_{k' \in \mathcal{B}_{j,K}} a_{j,k} a_{j,k'} (\sum_{i=1}^{n} \epsilon_i \frac{1}{g(X_i)} \psi_{j,k}(X_i)\psi_{j,k'}(X_i)) \Big) \\ &= \sup_{a \in \mathcal{C}_q} \Big( \sum_{k \in \mathcal{B}_{j,K}} \sum_{k' \in \mathcal{B}_{j,K}} |a_{j,k}|^2 |a_{j,k'}|^2 \Big)^{\frac{1}{2}} \Big( \sum_{k \in \mathcal{B}_{j,K}} \sum_{k' \in \mathcal{B}_{j,K}} \mathbb{E}_f^n (|\sum_{i=1}^{n} \epsilon_i (\frac{1}{g(X_i)} \psi_{j,k}(X_i)\psi_{j,k'}(X_i))|^2) \Big)^{\frac{1}{2}} \\ &\leq C \Big( \sum_{k \in \mathcal{B}_{j,K}} \sum_{k' \in \mathcal{B}_{j,K}} \mathbb{E}_f^n (|\sum_{i=1}^{n} \epsilon_i (\frac{1}{g(X_i)} \psi_{j,k}(X_i)\psi_{j,k'}(X_i))|^2) \Big)^{\frac{1}{2}} \end{aligned} \tag{3.11}$$

Since $(\epsilon_1, ..., \epsilon_n)$ are i.i.d Rademacher variables and are independent of $(X_1, ..., X_n)$, we have :

$$\begin{aligned} \mathbb{E}_f^n (|\sum_{i=1}^{n} \epsilon_i (\frac{1}{g(X_i)} \psi_{j,k}(X_i)\psi_{j,k'}(X_i))|^2) &= \mathbb{E}_f^n (\mathbb{E}_f^n (|\sum_{i=1}^{n} \epsilon_i (\frac{1}{g(X_i)} \psi_{j,k}(X_i)\psi_{j,k'}(X_i))|^2 | \mathbb{X})) \\ &= \mathbb{E}_f^n (\sum_{i=1}^{n} |\frac{1}{g(X_i)}|^2 |\psi_{j,k}(X_i)|^2 |\psi_{j,k'}(X_i)|^2) \\ &\leq Cn \mathbb{E}_f^n (|\psi_{j,k}(X_1)|^2 |\psi_{j,k'}(X_1)|^2). \end{aligned} \tag{3.12}$$

Since the concentration inequality (2.1) says that :

$$\sum_{(K)} \mathbb{E}_f^n (|\psi_{j,k}(X_1)|^2) \leq C 2^j,$$

the inequalities (3.11) and (3.12) give us :

$$S \leq C n^{\frac{1}{2}} 2^j.$$

Hence $H = C 2^j n^{-\frac{1}{2}}$.

∘ *Value of $v$.* Taking in account the assumption done on $g$, the fact that $q \leq 2$, the Hölder inequality et the concentration inequality (2.1), one gets :

$$\begin{aligned}\sup_{h\in\mathcal{F}} Var(h(X_1)) &\leq \sup_{a\in\mathcal{C}_q} \mathbb{E}_f^n\Big(\frac{1}{|g(X_1)|^2}\Big|\sum_{k\in\mathcal{B}_{j,K}}\sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}\psi_{j,k}(X_1)\psi_{j,k'}(X_1)\Big|^2\Big)\\ &\leq C\sup_{a\in\mathcal{C}_q}\Big(\sum_{(K)}|a_{j,k}|^2\Big)^2\mathbb{E}_f^n\Big(\Big(\sum_{(K)}|\psi_{j,k}(X_1)|^2\Big)^2\Big)\\ &\leq C2^{2j}.\end{aligned}$$

Hence $v = C2^{2j}$.

If we take $t = \frac{\kappa}{2}$ then :

$$\left(\frac{t^2}{v}\wedge\frac{t}{T}\right) = \left(C\frac{\kappa^2}{2^{2j}}\wedge C\frac{\kappa}{2^j}\right) = C\frac{\kappa^2}{2^{2j}}$$

Choising a such $t$ and taking in account the fact that $2^j \leq 2^{j_2} \leq \sqrt{\frac{n}{\ln(n)}}$, the Talagrand inequality says that :

$$\mathbb{P}_f^n\Big(\sup_{a\in\mathcal{C}_q}\Big(\sum_{k\in\mathcal{B}_{j,K}}\sum_{k'\in\mathcal{B}_{j,K}} a_{j,k}a_{j,k'}\Big(n^{-1}\sum_{i=1}^n\frac{1}{g(X_i)}\psi_{j,k}(X_i)\psi_{j,k'}(X_i)\Big) - \sum_{(K)}|a_{j,k}|^2\Big|\Big) \geq \frac{\kappa}{2} + C2^j n^{-\frac{1}{2}}\Big)$$

and a fortiori $\mathbb{P}_f^n(\mathcal{A}_\kappa)$, is dominated by :

$$C\exp(-nC\kappa^2 2^{-2j_2}) \leq Cn^{-p}$$

for $\kappa$ large enough. This ends the proof of Proposition 3.1. □

## Références


Cai, T. (1999). Adaptative wavelet estimation : a block thresholding and orakle inequality approach. *Ann. Stat.*, 27 :898–924.

Cai, T. (2002). On block thresholding in wavelet regression : adaptivity, block size and threshold level. *Statist. Sinica*, 12(4) :1241–1273.

Cai, T. and Brown, L. (1998). Wavelet shrinkage for nonequispaced samples. *Ann. Stat.*, 26(5) :1783–1799.

Cai, T. and Brown, L. (1999). Wavelet estimation for samples with random uniform design. *Statistics and Probability Letters*, 42 :313–321.

Chesneau, C. (2006a). On wavelet block thresholding in nonparametric estimation : Maxiset and minimax results under the $\mathbb{L}^p$ risk. *Preprint 'Laboratoire de Probabilités et Modèles Aléatoires' Paris VI (it can be printed on http://www.chesneau-stat.com).*

Chesneau, C. (2006b). Regression with random design : a minimax study. *Preprint 'Laboratoire de Probabilités et Modèles Aléatoires' Paris VI, it can be printed on http://www.chesneau-stat.com).*

Chicken, E. (2003). Nonparametric regression on random processes and design. *Florida State University Department of Statistics, Technical Report.*

CIRELSON, B., IBRAGIMOV, I., and SUDAKOV, V. (1976). *Norm of Gaussian sample functions*. Berlin, springer verlag ed.

COHEN, A., DAUBECHIES, I., JAWERTH, B., and VIAL, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 24(1) :54–81.

DELYON, B. and JUDITSKY, A. (1996). On minimax wavelet estimators. *Applied Computational Harmonic Analysis*, 3 :215–228.

DONOHO, D. and JOHNSTONE, I. (1995). Adaptating to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432) :1200–1224.

KERKYACHARIAN, G. and PICARD, D. (2005). Regression in random design and warped wavelets. *Bernoulli*, 10(6) :1053–1105.

MAXIM, V. (2003). *Restauration de signaux bruités sur des plans d'expérience aléatoires*. PhD Thesis, Université Joseph Fourrier, Grenoble I.

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

PICARD, D. and TRIBOULEY, K. (2000). Adaptive confidence interval for pointwise curve estimation. *Ann.Statist.*, 28(1) :298–335.

TALAGRAND, M. (1994). Sharper bounds for gaussian and empirical processes. *Ann. Probab.*, 22 :28–76.