



# Wavelet Estimation Via Block Thresholding : A Minimax Study Under The $L^p$ Risk

Christophe Chesneau

► **To cite this version:**

Christophe Chesneau. Wavelet Estimation Via Block Thresholding : A Minimax Study Under The  $L^p$  Risk. 19 pages. 2006. <hal-00017257v4>

**HAL Id: hal-00017257**

**<https://hal.archives-ouvertes.fr/hal-00017257v4>**

Submitted on 13 Oct 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Wavelet Estimation Via Block Thresholding : A Minimax Study Under The $L^p$ Risk

Christophe Chesneau

*University of Paris VI*

*Abstract:* We investigate the asymptotic minimax properties of an adaptive wavelet block thresholding estimator under the  $L^p$  risk over Besov balls. It can be viewed as a  $\mathbb{L}^p$  version of the BlockShrink estimator developed by Cai (1996,1997,2002). Firstly, we show that it is (near) optimal for numerous statistical models, including certain inverse problems. Under this statistical context, it achieves better rates of convergence than the hard thresholding estimator introduced by Donoho and Johnstone (1995). Secondly, we apply this general result to a deconvolution problem.

*Key words and phrases* Minimax estimation;  $L^p$  risk; Besov spaces; wavelets; block thresholding; convolution in Gaussian white noise model.

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

## 1. Motivations

Wavelet shrinkage methods have been very successful in nonparametric function estimation. They provide estimators that are spatially adaptive and (near) optimal over a wide range of function classes. Standard approaches are based on the term-by-term thresholding. The well-known example is the hard thresholding estimators introduced by Donoho and Johnstone (1995).

Recent works have shown that local block thresholding methods can enjoy better theoretical (and practical) properties than conventional term-by-term thresholding methods. This is the case for the construction developed by Hall, Kerkycharian and Picard (1999), the BlockShrink algorithm proposed by Cai (1996,1997,2002) and the blockwise Stein's algorithm studied by Cavalier and Tsybakov (2002). If we adopt the minimax point of view then the resulting estimators are optimal under the  $L^2$  risk over a wide range of Besov balls for various statistical models.

In the present paper, we synthetically analyze the asymptotic performances

of a  $L^p$  version of the BlockShrink estimator. In a first part, we consider the estimation of an unknown function  $f$  in  $L^p([0, 1])$  from a general sequence of models  $\Gamma_n$ . Under very mild assumptions on  $\Gamma_n$ , we determine a simple upper bound of the  $L^p$  risk

$$R(\hat{f}, f) = E(\|\hat{f} - f\|_p^p) = E\left(\int_0^1 |\hat{f}(t) - f(t)|^p dt\right), \quad p \geq 2,$$

where  $\hat{f}$  is a  $L^p$  version of the BlockShrink estimator and  $E$  is the expectation with respect to the distribution of the observations. Then, we use this result to isolate the rates of convergence achieved by this estimator when  $f$  belongs to Besov balls. For numerous statistical models (including several inverse problems), we show that they are (near) minimax. Moreover, the estimator considered is better in the minimax sense than the hard thresholding estimator.

In a second part, we provide some applications of this general result. After a brief study of the standard Gaussian white noise model, we focus our attention on a more delicate problem : the convolution in Gaussian white noise model.

The rest of the paper is organized as follows. Section 2 describes wavelets and Besov balls. Section 3 introduces the  $L^p$  version of the BlockShrink estimator and the key assumptions. Asymptotic properties of this estimator will be presented in Section 4. In Section 5, we apply this result to the Gaussian white noise model and the convolution in Gaussian white noise model. Section 6 contains proofs of the main theorems.

## 2. Wavelets and Besov balls

We work with a wavelet basis on the interval  $[0, 1]$  of the form

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

In general,  $\phi_{j,k}(x)$  and  $\psi_{j,k}(x)$  are "periodic" or "boundary adjusted" dilation and translation of a "father" wavelet  $\phi$  and a "mother" wavelet  $\psi$ , respectively. This last function is supposed to be  $N$ -regular. The factor  $\tau$  is a large enough integer. For the sake of simplicity, we set  $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$  and  $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$ . We assume that the three following geometrical properties are satisfied.

1. *Property of concentration.* Let  $p \in ]1, \infty[$  and  $h \in \{\phi, \psi\}$ . For any  $j \in \{\tau, \dots, \infty\}$  and any sequence  $u = (u_{j,k})_{j,k}$ , there exists a constant  $C > 0$

such that

$$\left\| \sum_{k=0}^{2^j-1} u_{j,k} h_{j,k} \right\|_p^p \leq C 2^{j(p/2-1)} \sum_{k=0}^{2^j-1} |u_{j,k}|^p. \quad (2.1)$$

2. *Property of unconditionality.* Let  $p \in ]1, \infty[$ . Let us set  $\psi_{\tau-1,k} = \phi_{\tau,k}$ . For any sequence  $u = (u_{j,k})_{j,k}$ , we have

$$\left\| \sum_{j=\tau-1}^{\infty} \sum_{k=0}^{2^j-1} u_{j,k} \psi_{j,k} \right\|_p^p \asymp \left\| \left( \sum_{j=\tau-1}^{\infty} \sum_{k=0}^{2^j-1} |u_{j,k} \psi_{j,k}|^2 \right)^{1/2} \right\|_p^p. \quad (2.2)$$

(The notation  $a \asymp b$  means : there exist two constants  $C > 0$  and  $c > 0$  such that  $cb \leq a \leq Cb$ .)

3. *Temlyakov property.* Let  $\sigma \in [0, \infty[$ . Let us set  $\psi_{\tau-1,k} = \phi_{\tau,k}$ . For any subset  $A \subseteq \{\tau-1, \dots, \infty\}$  and any subset  $C \subseteq \{0, \dots, 2^j-1\}$ , we have

$$\left\| \left( \sum_{j \in A} \sum_{k \in C} |2^{\sigma j} \psi_{j,k}|^2 \right)^{1/2} \right\|_p^p \asymp \sum_{j \in A} \sum_{k \in C} 2^{\sigma j p} \|\psi_{j,k}\|_p^p. \quad (2.3)$$

The first property is standard. The others are powerful tools. See Meyer (1990) for further details about wavelets, the property of concentration and the property of unconditionality. See Johnstone, Kerkycharian, Picard and Raimondo (2004) for further details about the Temlyakov property.

For any  $l \in \{\tau, \dots, \infty\}$ , a function  $f$  in  $L^2([0, 1])$  can be expanded in a wavelet series as

$$f(x) = \sum_{k=0}^{2^l-1} \alpha_{l,k} \phi_{l,k}(x) + \sum_{j=l}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k}(x),$$

where  $\alpha_{j,k} = \int_0^1 f(t) \phi_{j,k}(t) dt$  and  $\beta_{j,k} = \int_0^1 f(t) \psi_{j,k}(t) dt$ .

A suitable choice of the wavelet basis  $\zeta$  depends on the considered statistical model. Further details are given in Section 4.

Now, let us define the main function spaces of the study. Let  $M \in ]0, \infty[$ ,  $s \in ]0, N[$  and  $\pi \in [1, \infty[$ . Let us set  $\beta_{\tau-1,k} = \alpha_{\tau,k}$ . We say that a function  $f$  belongs to the Besov balls  $B_{\pi,r}^s(M)$  if and only if the associated wavelet coefficients satisfy

$$\left[ \sum_{j=\tau-1}^{\infty} [2^{j(s+1/2-1/\pi)} \left( \sum_{k=0}^{2^j-1} |\beta_{j,k}|^\pi \right)^{1/\pi}]^r \right]^{1/r} \leq M, \quad \text{for } r \in [1, \infty[,$$

with the usual modification if  $r = \infty$ . For a particular choice of parameters  $s$ ,  $\pi$  and  $r$ , they contain the Holder and Sobolev balls. See Meyer (1990).

### 3. Estimator and assumptions

In the first part of the present paper, following the mathematical framework adopted by Picard and Kerkyacharian (2000), we consider the estimation of an unknown function  $f$  in  $L^p([0, 1])$  from a general situation. We only assume to have a sequence of models  $\Gamma_n$  in which we are able to produce estimates of the wavelet coefficients  $\alpha_{j,k}$  and  $\beta_{j,k}$  of  $f$  on the basis  $\zeta$ . The corresponding estimators will be denoted  $\hat{\alpha}_{j,k}$  and  $\hat{\beta}_{j,k}$ .

Now, let us explain the role of two factors  $\delta$  and  $\nu$  which will appear in our mathematical framework. The first is supposed to be a parameter characterizing the model. It plays a crucial role in the study of certain inverse problems. For the standard models, it is equal to zero. The second has only a technical utility. It may depend on  $\delta$ .

We are now in position to describe the main estimator of the study. It is a  $L^p$  version of the BlockShrink estimator developed by Cai (1999). It was first defined by Picard and Tribouley (2000). It is important to mention that it does not require any a priori knowledge on  $f$  in his construction.

Suppose that  $p \in [2, \infty[$ ,  $d \in ]0, \infty[$ ,  $\delta \in [0, \infty[$  and  $\nu \in ]0, (2\delta + 1)^{-1}[$ . Let  $j_1$  and  $j_2$  be integers satisfying  $2^{j_1} \asymp (\log n)^{p/2}$  and  $2^{j_2} \asymp n^\nu$  (or  $2^{j_2} \asymp (n/\log n)^\nu$ ). For any  $j \in \{j_1, \dots, j_2\}$ , let us set  $L \asymp (\log n)^{p/2}$ ,  $A_j = \{1, \dots, 2^j L^{-1}\}$  and, for any  $K \in A_j$ ,  $U_{j,K} = \{k \in \{0, \dots, 2^j - 1\}; (K-1)L \leq k \leq KL - 1\}$ . We define the ( $L^p$  version of the) BlockShrink estimator by

$$\hat{f}(x) = \sum_{k=0}^{2^{j_1}-1} \hat{\alpha}_{j_1,k} \phi_{j_1,k}(x) + \sum_{j=j_1}^{j_2} \sum_{K \in A_j} \sum_{k \in U_{j,K}} \hat{\beta}_{j,k} 1_{\{\hat{b}_{j,K} \geq d 2^{\delta j} n^{-1/2}\}} \psi_{j,k}(x), \quad (3.1)$$

where  $\hat{b}_{j,K} = (L^{-1} \sum_{k \in U_{j,K}} |\hat{\beta}_{j,k}|^p)^{1/p}$ .

For the sake of legibility, we set  $\sum_K = \sum_{K \in A_j}$  and  $\sum_{(K)} = \sum_{k \in U_{j,K}}$ . All the constants of our study are independent of  $f$  and  $n$ .

We make the following assumptions.

(H1). *Moments inequality* Let us set  $\hat{\beta}_{j_1-1,k} = \hat{\alpha}_{j_1,k}$ . There exists a constant  $C > 0$  such that, for any  $j \in \{j_1 - 1, \dots, j_2\}$ ,  $k \in \{0, \dots, 2^j - 1\}$  and  $n$  large

enough, we have

$$E(|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}) \leq C2^{2\delta j p} n^{-p}.$$

(H2). *Large deviation inequality* There exist two constants  $\mu$  and  $C > 0$  such that, for any  $j \in \{j_1, \dots, j_2\}$ ,  $K \in A_j$  and  $n$  large enough, we have

$$P\left(\left(L^{-1} \sum_{(K)} |\hat{\beta}_{j,k} - \beta_{j,k}|^p\right)^{1/p} \geq 2^{-1} \mu 2^{\delta j} n^{-1/2}\right) \leq Cn^{-p}.$$

For numerous statistical models, we can find  $\hat{\alpha}_{j,k}$ ,  $\hat{\beta}_{j,k}$ ,  $\nu$  and  $\mu$  which satisfy the assumptions (H1) and (H2). Several applications will be considered in Section 5.

#### 4. Optimality results

Theorem 4.1 below provides an upper bound of the  $L^p$  ( $p \geq 2$ ) risk of block thresholding estimator  $\hat{f}$  defined by (3.1). The function  $f$  is only supposed to belong to  $L^p([0, 1])$ .

**Theorem 4.1** *Let  $p \in [2, \infty]$ . Let us consider the general statistical framework described in Chapter 3. Suppose that the assumptions (H1) and (H2) are satisfied. Let us consider the estimator  $\hat{f}$  defined by (3.1) with the thresholding constant  $d = \mu$ . Then there exists a constant  $C > 0$  such that, for any  $\alpha \in ]0, 1[$  and  $n$  large enough, we have*

$$E(\|\hat{f} - f\|_p^p) \leq C(Q_1(f) + Q_2(f) + n^{-\alpha p/2}),$$

where

$$Q_1(f) = \sum_{m=0}^{\infty} 2^{-mp} \left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{b_{j,K} \leq 2^{-1} \mu n^{-1/2} 2^{\delta j} 2^{m+1}\}} \psi_{j,k} \right\|_p^p,$$

$$Q_2(f) = \left\| \sum_{j=j_2+1}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k} \right\|_p^p.$$

The geometrical properties of the basis  $\zeta$  under the  $L^p$  norm is at the heart of the proof. Such an inequality was proved for the hard thresholding estimator by Kerkyacharian and Picard (2000, Theorem 5.1).

Theorem 4.2 below is a consequence of Theorem 4.1. We now suppose that  $f$  belongs to Besov balls  $B_{\pi,r}^s(M)$ . We investigate the rates of convergence achieved by the block thresholding estimator  $\hat{f}$  defined by (3.1) under the  $L^p$  risk for  $p \geq 2$ .

**Theorem 4.2** *Let  $p \in [2, \infty]$ . Let us consider the general statistical framework described in Chapter 3. Suppose that the assumptions (H1) and (H2) are satisfied. Let us consider the estimator  $\hat{f}$  defined by (3.1) with the thresholding constant  $d = \mu$ . Then there exists a constant  $C > 0$  such that, for any  $\pi \in [1, \infty]$ ,  $r \in [1, \infty]$ ,  $s \in ]1/\pi - (1/2 - 1/(2\nu) + \delta)_+, N]$  and  $n$  large enough, we have*

$$\sup_{f \in B_{\pi, r}^s(M)} E(\|\hat{f} - f\|_p^2) \leq C\varphi_n,$$

where

$$\varphi_n = \begin{cases} n^{-\alpha_1 p} (\log n)^{\alpha_1 p 1_{\{p > \pi\}}}, & \text{when } \epsilon > 0, \\ (\log n/n)^{\alpha_2 p} (\log n)^{(p - \pi/r) + 1_{\{\epsilon = 0\}}}, & \text{when } \epsilon \leq 0, \end{cases}$$

with  $\alpha_1 = s/(2(s + \delta) + 1)$ ,  $\alpha_2 = (s - 1/\pi + 1/p)/(2(s - 1/\pi + \delta) + 1)$  and  $\epsilon = \pi s + (\delta + 1/2)(\pi - p)$ .

For numerous statistical models, the rates of convergence exhibited in Theorem 4.2 are minimax, except for the case  $\epsilon > 0$  with  $p > \pi$  where an additional factor logarithmic appeared. For further details about the minimax rates of convergence under the  $L^p$  risk over Besov balls, see Delyon and Juditsky (1996) and the book of Hardle, Kerkycharian, Picard and Tsybakov (1998).

Moreover, let us notice that if (H2) is satisfied then there exist two constants  $C > 0$  and  $\mu_* > 0$  such that, for any  $j \in \{j_1, \dots, j_2\}$ ,  $k \in \{0, \dots, 2^j - 1\}$  and  $n$  large enough, we have :  $P(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq 2^{-1}\mu_* 2^{\delta j} \sqrt{(\log n/n)}) \leq P((\sum_{(K)} |\hat{\beta}_{j,k} - \beta_{j,k}|^p)^{1/p} \geq 2^{-1}\mu 2^{\delta j} \sqrt{(\log n/n)}) \leq Cn^{-p}$ . So, by considering a result proved by Picard and Kerkycharian (2000, Theorem 6.1), under the assumptions (H1) and (H2), the  $\mathbb{L}^p$  version of the BlockShrink estimator achieves better rates of convergence than the hard thresholding estimator. More precisely, it removes the logarithmic term in the case  $\pi \geq p$ .

In the following section, we apply our general results to the standard Gaussian white noise model and a well-known deconvolution problem.

## 5. Applications

– *Gaussian white noise model.* We consider the random process  $\{Y(t); t \in [0, 1]\}$  defined by

$$dY(t) = f(t)dt + n^{-1/2}dW(t),$$

where  $\{W(t); t \in [0, 1]\}$  is a standard Brownian motion. We wish to estimate the unknown function  $f$  via  $\{Y(t); t \in [0, 1]\}$ .

Here, we work with the compactly supported wavelet basis on the unit interval introduced by Daubechies, Cohen et Vial (1992). It satisfies the property of concentration, the property of unconditionality and the Temlyakov property. See for instance Picard and Kerkyacharian (2000).

Picard and Tribouley (2000) have shown that assumptions (H1) and (H2) are satisfied with  $\hat{\alpha}_{j,k} = \int_0^1 \phi_{j,k}(t)dY(t)$ ,  $\hat{\beta}_{j,k} = \int_0^1 \psi_{j,k}(t)dY(t)$ ,  $\delta = 0$ ,  $\nu = 1$  and  $\mu$  large enough. Therefore, if we defined the estimator (3.1) with the previous elements, then we can apply Theorem 4.2. This theorem can be viewed as a  $L^p$  version of some results obtained by Cai (1997, Theorems 2 and 3) under the  $L^2$  risk.

– *Convolution in Gaussian white noise model.* We consider the random process  $\{Y(t); t \in [0, 1]\}$  defined by

$$dY(t) = (f \star g)(t)dt + n^{-1/2}dW(t),$$

where  $\{W(t); t \in [0, 1]\}$  is a standard Brownian motion and  $(f \star g)(t) = \int_0^1 f(t-u)g(u)du$ . The function  $f$  is unknown and the function  $g$  is known. We assume that  $f$  and  $g$  are periodic on the unit interval and that there exists a  $\delta > 2^{-1}$  satisfying

$$F(g)(l) \asymp |l|^{-\delta}, \quad l \in Z^*, \quad F(g)(0) = 1. \quad (5.1)$$

For any  $h \in L^1([0, 1])$  and real number  $l$ ,  $F(h)$  denotes the Fourier transform of  $h$  defined by  $F(h)(l) = \int_0^1 h(x)e^{-2i\pi lx}dx$ . We wish to recover the unknown function  $f$  via  $\{Y(t); t \in [0, 1]\}$ . This model has been studied in many papers. See, for instance, Cavalier and Tsybakov (2002) and Johnstone, Kerkyacharian, Picard and Raimondo (2004).

Here, we adopt the statistical framework developed by Johnstone, Kerkyacharian, Picard and Raimondo (2004). We work with a basis constructed from Meyer-type wavelet adapted to the interval  $[0, 1]$  by periodization. We denote this family by  $\zeta^M = \{\phi_{\tau,k}^M(x), k = 0, \dots, 2^\tau - 1; \psi_{j,k}^M(x); j = \tau, \dots, \infty, k = 0, \dots, 2^j - 1\}$ , where  $\tau$  denotes a large integer. The main particularity of  $\zeta^M$  is that  $F(\psi^M)$  and  $F(\phi^M)$  are compactly supported. Moreover,  $\zeta^M$  satisfies the property of concentration, the property of unconditionality and the Temlyakov property.



**Theorem 5.3** *The assumptions (H1) and (H2) are satisfied with the estimator proposed by Johnstone, Kerkyacharian, Picard and Raimondo (2004):*

$$\hat{\alpha}_{j,k} = \sum_{l \in C_j} F^*(Y)(l)F(g)(l)^{-1}F(\phi_{j,k}^M)(l), \quad \hat{\beta}_{j,k} = \sum_{l \in C_j} F^*(Y)(l)F(g)(l)^{-1}F(\psi_{j,k}^M)(l),$$

$\nu = (1 + 2\delta)^{-1}$  and  $\mu$  large enough. Here,  $C_j = \{l \in Z; F(\psi_{j,k}^M)(l) \neq 0\} = \{l \in Z; |l| \in [2\pi 3^{-1}2^j, 8\pi 3^{-1}2^j]\}$  and, for any integrable process  $\{R(t); t \in [0, 1]\}$ ,  $F^*(R)(l) = \int_0^1 e^{-2i\pi lt} dR(t)$ .

The main difficulty of the proof of Theorem 5.3 is to show the assumption (H2).

So, if we define the estimator (3.1) with the elements  $\hat{\alpha}_{j,k}$ ,  $\hat{\beta}_{j,k}$ ,  $\delta$ ,  $\nu$  and  $\mu$  of Theorem 5.3, then we can apply Theorem 4.2. In particular, under the  $L^p$  risk for  $p \geq 2$  over Besov balls, the considered estimator is better than the hard thresholding estimator developed by Johnstone, Kerkyacharian, Picard and Raimondo (2004).

## 6. Proofs

Here and latter,  $C$  represents a constant which may be different from one term to the other. We suppose that  $n$  is large enough.

**Proof of Theorem 4.1.** For the sake of simplicity in exposition, we set  $\hat{\theta}_{j,k} = \hat{\beta}_{j,k} - \beta_{j,k}$ . Applying the Minkowski inequality and an elementary inequality of convexity, we have  $E(\|\hat{f} - f\|_p^p) \leq 4^{p-1}(G_1 + G_2 + G_3 + Q_2(f))$  where

$$G_1 = E\left(\left\|\sum_{k=0}^{2^{j_1}-1} (\hat{\alpha}_{j_1,k} - \alpha_{j_1,k})\phi_{j_1,k}\right\|_p^p\right),$$

$$G_2 = E\left(\left\|\sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{\hat{b}_{j,K} < 2^{\delta j} \mu n^{-1/2}\}} \psi_{j,k}\right\|_p^p\right),$$

$$G_3 = E\left(\left\|\sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \hat{\theta}_{j,k} 1_{\{\hat{b}_{j,K} \geq 2^{\delta j} \mu n^{-1/2}\}} \psi_{j,k}\right\|_p^p\right).$$

Let us analyze each term  $G_1$ ,  $G_2$  and  $G_3$ , in turn.

• *The upper bound for  $G_1$ .* It follows from the property of concentration (2.1) and the assumption (H1) that

$$\begin{aligned} G_1 &\leq C 2^{j_1(p/2-1)} \sum_{k=0}^{2^{j_1}-1} E(|\hat{\alpha}_{j_1,k} - \alpha_{j_1,k}|^p) \leq C n^{-p/2} 2^{j_1(\delta+1/2)p} \\ &\leq C n^{-p/2} (\log n)^{(\delta/2+1/4)p^2} \leq C n^{-\alpha p/2}. \end{aligned} \tag{6.1}$$

• *The upper bound for  $G_2$ .* Applying the Minkowski inequality and an elementary inequality of convexity, we have  $G_2 \leq 2^{p-1}(G_{2,1} + G_{2,2})$ , where

$$G_{2,1} = E\left(\left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{\hat{b}_{j,K} < 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} \leq 22^{\delta j} \mu n^{-1/2}\}} \psi_{j,k} \right\|_p^p\right),$$

$$G_{2,2} = E\left(\left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{\hat{b}_{j,K} < 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} > 22^{\delta j} \mu n^{-1/2}\}} \psi_{j,k} \right\|_p^p\right).$$

– *The upper bound for  $G_{2,1}$ .* Using the property of unconditionality (2.2), we find

$$G_{2,1} \leq C \left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{b_{j,K} \leq 22^{\delta j} \mu n^{-1/2}\}} \psi_{j,k} \right\|_p^p \leq C Q_1(f).$$

– *The upper bound for  $G_{2,2}$ .* Notice that the  $l_p$  Minkowski inequality yields

$$\begin{aligned} 1_{\{b_{j,K} > 22^{\delta j} \mu n^{-1/2}\}} 1_{\{\hat{b}_{j,K} < 2^{\delta j} \mu n^{-1/2}\}} &\leq 1_{\{|\hat{b}_{j,K} - b_{j,K}| \geq 2^{\delta j} \mu n^{-1/2}\}} \\ &\leq 1_{\{(L^{-1} \sum_{(K)} |\hat{\theta}_{j,k}|^p)^{1/p} \geq 2^{\delta j} \mu n^{-1/2}\}}. \end{aligned} \quad (6.2)$$

Using the property of unconditionality (2.2), the generalized Minkowski inequality, the inequality (6.2), the assumption (H2) and again (2.2), we obtain

$$\begin{aligned} G_{2,2} &\leq CE\left(\left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} |\beta_{j,k}|^2 1_{\{b_{j,K} > 22^{\delta j} \mu n^{-1/2}\}} 1_{\{\hat{b}_{j,K} < 2^{\delta j} \mu n^{-1/2}\}} |\psi_{j,k}|^2 \right\|_p^{1/2}\right) \\ &\leq C \left\| \left( \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} |\beta_{j,k}|^2 [E(1_{\{b_{j,K} > 22^{\delta j} \mu n^{-1/2}\}} 1_{\{\hat{b}_{j,K} < 2^{\delta j} \mu n^{-1/2}\}})]^{2/p} |\psi_{j,k}|^2 \right)^{1/2} \right\|_p^p \\ &\leq C \left\| \left( \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} |\beta_{j,k}|^2 [P((L^{-1} \sum_{(K)} |\hat{\theta}_{j,k}|^p)^{1/p} \geq 2^{\delta j} \mu n^{-1/2})]^{2/p} |\psi_{j,k}|^2 \right)^{1/2} \right\|_p^p \\ &\leq C n^{-p/2} \left\| \left( \sum_{j=\tau}^{\infty} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^2 |\psi_{j,k}|^2 \right)^{1/2} \right\|_p^p \leq C \|f\|_p^p n^{-\alpha p/2} \leq C n^{-\alpha p/2}. \end{aligned}$$

It follows from the upper bounds of  $G_{2,1}$  and  $G_{2,2}$  that

$$G_2 \leq C(Q_1(f) + n^{-\alpha p/2}). \quad (6.3)$$

• *The upper bound for  $G_3$ .* By the Minkowski inequality and an elementary inequality of convexity, we have  $G_3 \leq 2^{p-1}(G_{3,1} + G_{3,2})$ , where

$$G_{3,1} = E\left(\left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \hat{\theta}_{j,k} 1_{\{\hat{b}_{j,K} \geq 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} \psi_{j,k} \right\|_p^p\right),$$

$$G_{3,2} = E\left(\left\| \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \hat{\theta}_{j,k} 1_{\{\hat{b}_{j,K} \geq 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} \geq 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} \psi_{j,k} \right\|_p^p\right).$$

– *The upper bound for  $G_{3,1}$ .* An inequality similar to (6.2), the Cauchy-Schwartz inequality and the assumptions (H1) and (H2) imply

$$\begin{aligned} & E(|\hat{\theta}_{j,k}|^p 1_{\{\hat{b}_{j,K} \geq 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2}\}}) \\ & \leq [E(|\hat{\theta}_{j,k}|^{2p})]^{1/2} [P((L^{-1} \sum_{(K)} |\hat{\theta}_{j,k}|^p)^{1/p} \geq 2^{\delta j} 2^{-1} \mu n^{-1/2})]^{1/2} \leq C 2^{\delta j p} n^{-p}. \end{aligned} \tag{6.4}$$

Using the property of unconditionality (2.2), the generalized Minkowski inequality, the inequality (6.4), the Temlyakov property (2.3) and the fact that  $\nu \in ]0, (2\delta + 1)^{-1}]$ , we have

$$\begin{aligned} G_{3,1} & \leq CE\left(\left\| \left( \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} |\hat{\theta}_{j,k}|^2 1_{\{\hat{b}_{j,K} \geq 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} |\psi_{j,k}|^2 \right)^{1/2} \right\|_p^p\right) \\ & \leq C \left\| \left( \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} [E(|\hat{\theta}_{j,k}|^p 1_{\{\hat{b}_{j,K} \geq 2^{\delta j} \mu n^{-1/2}\}} 1_{\{b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2}\}})]^{2/p} |\psi_{j,k}|^2 \right)^{1/2} \right\|_p^p \\ & \leq C n^{-p} \left\| \left( \sum_{j=\tau}^{j_2} \sum_{k=0}^{2^j-1} 2^{2\delta j} |\psi_{j,k}|^2 \right)^{1/2} \right\|_p^p \leq C n^{-p} \sum_{j=\tau}^{j_2} \sum_{k=0}^{2^j-1} 2^{\delta j p} \|\psi_{j,k}\|_p^p \\ & = C n^{-p} \sum_{j=\tau}^{j_2} 2^{j(\delta+1/2)p} \leq C n^{-p} 2^{j_2(\delta+1/2)p} \leq C n^{-p} n^{\nu p(\delta+1/2)} \leq C n^{-\alpha p/2}. \end{aligned}$$

– *The upper bound for  $G_{3,2}$ .* Using the property of unconditionality (2.2), the generalized Minkowski inequality, the assumption (H1) and the Temlyakov

property (2.3), we obtain

$$\begin{aligned}
G_{3,2} &\leq CE\left(\left\|\left(\sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} |\hat{\theta}_{j,k}|^2 1_{\{b_{j,K} \geq 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} |\psi_{j,k}|^2\right)^{1/2}\right\|_p^p\right) \\
&\leq C\left\|\left(\sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} [E(|\hat{\theta}_{j,k}|^p)]^{2/p} 1_{\{b_{j,K} \geq 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} |\psi_{j,k}|^2\right)^{1/2}\right\|_p^p \\
&\leq Cn^{-p/2}\left\|\left(\sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} 1_{\{b_{j,K} \geq 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} 2^{2\delta j} |\psi_{j,k}|^2\right)^{1/2}\right\|_p^p \\
&\leq Cn^{-p/2} \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} 1_{\{b_{j,K} \geq 2^{\delta j} 2^{-1} \mu n^{-1/2}\}} 2^{\delta j p} \|\psi_{j,k}\|_p^p.
\end{aligned}$$

By virtue of the Markov inequality and the inclusion  $B_{p,p}^0 \subseteq L^p$ , we find

$$\begin{aligned}
G_{3,2} &\leq Cn^{-p/2} \sum_{m=0}^{\infty} \sum_{j=j_1}^{j_2} \sum_K 1_{\{2^{\delta j} 2^{-1} \mu n^{-1/2} 2^m \leq b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2} 2^{m+1}\}} 2^{\delta j p} L 2^{j(p/2-1)} \\
&\leq C \sum_{m=0}^{\infty} 2^{-mp} \sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} |\beta_{j,k}|^p 1_{\{b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2} 2^{m+1}\}} 2^{j(p/2-1)} \\
&\leq C \sum_{m=0}^{\infty} 2^{-mp} \left\|\sum_{j=j_1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{b_{j,K} < 2^{\delta j} 2^{-1} \mu n^{-1/2} 2^{m+1}\}} \psi_{j,k}\right\|_p^p = CQ_1(f).
\end{aligned}$$

It follows from the upper bounds of  $G_{3,1}$  and  $G_{3,2}$  that

$$G_3 \leq C(Q_1(f) + n^{-\alpha p/2}). \quad (6.5)$$

Combining (6.1), (6.3) and (6.5), for any  $\alpha \in ]0, 1[$ , we have

$$E(\|\hat{f} - f\|_p^p) \leq C(Q_1(f) + Q_2(f) + n^{-\alpha p/2}).$$

The proof of Theorem 4.1 is complete.

**Proof of Theorem 4.2.** Let us investigate separately the case  $\pi \geq p$  and the case  $p > \pi$ .

• **If  $\pi \geq p$ .** According to Theorem 4.1, it suffices to show that, for any  $f \in B_{\pi,r}^s(M)$ , there exists a constant  $C > 0$  satisfying the inequality  $Q_1(f) \vee Q_2(f) \leq Cn^{-\alpha_1 p}$  where  $\alpha_1 = s/(2(s + \delta) + 1)$ .

• *The upper bound for  $Q_1(f)$ .* For any integer  $m$ , let  $j_3$  be an integer satisfying  $2^{j_3} \asymp 2^{-m/(2s)} n^{1/(2(s+\delta)+1)}$ . Using the Minkowski inequality, an elementary inequality of convexity and the property of unconditionality (2.2), we have  $Q_1(f) \leq 2^{p-1}(S_1 + S_2)$ , where

$$S_1 = \sum_{m=0}^{\infty} 2^{-mp} \left\| \sum_{j=j_1}^{j_3} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}\}} \psi_{j,k} \right\|_p^p,$$

$$S_2 = \sum_{m=0}^{\infty} 2^{-mp} \left\| \sum_{j=j_3+1}^{j_2} \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k} \right\|_p^p.$$

Let us analyze each term  $S_1$  and  $S_2$ , in turn.

– *The upper bound for  $S_1$ .* If  $b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}$  then we have clearly  $(\sum_{(K)} |\beta_{j,k}|^p)^{1/p} \leq \mu n^{-1/2} 2^m 2^{\delta j} L^{1/p}$ . It follows from the Minkowski inequality and the property of concentration (2.1) that

$$\begin{aligned} S_1 &\leq C \sum_{m=0}^{\infty} 2^{-mp} \left[ \sum_{j=j_1}^{j_3} 2^{j(1/2-1/p)} \left( \sum_K \sum_{(K)} |\beta_{j,k}|^p 1_{\{b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}\}} \right)^{1/p} \right]^p \\ &\leq C n^{-p/2} \sum_{m=0}^{\infty} \left[ \sum_{j=\tau}^{j_3} 2^{j(1/2-1/p)} (Card(A_j) 2^{\delta j p} L)^{1/p} \right]^p = C n^{-p/2} \sum_{m=0}^{\infty} 2^{j_3(\delta+1/2)p} \\ &\leq C n^{-sp/(2(s+\delta)+1)} \sum_{m=0}^{\infty} 2^{-mp(1+2\delta)/(4s)} \leq C n^{-\alpha_1 p}. \end{aligned}$$

– *The upper bound for  $S_2$ .* The Minkowski inequality, the property of concentration (2.1) and the inclusion  $B_{\pi,r}^s(M) \subseteq B_{p,\infty}^s(M)$  imply that

$$\begin{aligned} S_2 &\leq C \sum_{m=0}^{\infty} 2^{-mp} \left[ \sum_{j=j_3+1}^{j_2} 2^{j(1/2-1/p)} \left( \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right]^p \leq C \sum_{m=0}^{\infty} 2^{-mp} \left( \sum_{j=j_3+1}^{\infty} 2^{-js} \right)^p \\ &\leq C \sum_{m=0}^{\infty} 2^{-mp} 2^{-j_3 s p} \leq C n^{-sp/(2(s+\delta)+1)} \sum_{m=0}^{\infty} 2^{-mp/2} \leq C n^{-\alpha_1 p}. \end{aligned}$$

Putting the upper bounds of  $S_1$  and  $S_2$  together, we conclude that

$$Q_1(f) \leq C n^{-\alpha_1 p}. \quad (6.6)$$

• *The upper bound for  $Q_2(f)$ .* Using the Minkowski inequality, the property of concentration (2.1), the inclusion  $B_{\pi,r}^s(M) \subseteq B_{p,r}^s(M)$  and the fact that  $s >$

$1/\pi - \delta - 1/2 + 1/(2\nu)$ , we find

$$\begin{aligned} Q_2(f) &\leq C \left[ \sum_{j=j_2+1}^{\infty} 2^{j(1/2-1/p)} \left( \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right]^p \leq C \left( \sum_{j=j_2+1}^{\infty} 2^{-js} \right)^p \leq C 2^{-j_2 sp} \\ &\leq C (\log n/n)^{\nu sp} \leq C n^{-\alpha_1 p}. \end{aligned} \quad (6.7)$$

We obtain the desired result by combining (6.6) and (6.7) and applying Theorem 4.1 with  $\alpha = 2\alpha_1$ .

• **If  $p > \pi$ .** According to Theorem 4.1, it suffices to show that, for any  $f \in B_{\pi,r}^s(M)$ , there exists a constant  $C > 0$  satisfying the inequality  $Q_1(f) \vee Q_2(f) \leq C (\log n/n)^{\alpha_* p} (\log n)^{(p-\pi/r)+1_{\{\epsilon=0\}}}$  where  $\alpha_* = \alpha_1 1_{\{\epsilon>0\}} + \alpha_2 1_{\{\epsilon \leq 0\}}$ ,  $\alpha_1 = s/(2(s+\delta)+1)$ ,  $\alpha_2 = (s-1/\pi+1/p)/(2(s-1/\pi+\delta)+1)$  and  $\epsilon = \pi s + (\delta+1/2)(\pi-p)$ .

• *The upper bound of  $Q_1(f)$ .* Let  $j_4$  be an integer such that

$$2^{j_4} \asymp 2^{-m/(2s)} (n/\log n)^{1/(2(s+\delta)+1-(2/\pi)1_{\{\epsilon \leq 0\}})}.$$

The Minkowski inequality and an elementary of convexity give  $Q_1(f) \leq 2^{p-1}(T_1 + T_2)$ , where

$$T_1 = \sum_{m=0}^{\infty} 2^{-mp} \left\| \sum_{j=\tau}^{j_4} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}\}} \psi_{j,k} \right\|_p^p,$$

$$T_2 = \sum_{m=0}^{\infty} 2^{-mp} \left\| \sum_{j=j_4+1}^{j_2} \sum_K \sum_{(K)} \beta_{j,k} 1_{\{b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}\}} \psi_{j,k} \right\|_p^p.$$

Let us distinguish the case  $\epsilon > 0$  with  $p > \pi$  and the case  $\epsilon \leq 0$ .

• *For  $\epsilon > 0$  with  $p > \pi$ .*

– *The upper bound for  $T_1$ .* If  $b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}$  then we have clearly  $(\sum_{(K)} |\beta_{j,k}|^p)^{1/p} \leq \mu n^{-1/2} 2^m 2^{\delta j} L^{1/p}$ . The Minkowski inequality and the property

of concentration (2.1) imply that

$$\begin{aligned}
T_1 &\leq C \sum_{m=0}^{\infty} 2^{-mp} \left[ \sum_{j=\tau}^{j_4} 2^{j(1/2-1/p)} \left( \sum_K \sum_{(K)} |\beta_{j,k}|^p \mathbf{1}_{\{b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}\}} \right)^{1/p} \right]^p \\
&\leq C n^{-p/2} \sum_{m=0}^{\infty} \left( \sum_{j=\tau}^{j_4} 2^{j(1/2+\delta)} \right)^p \leq C n^{-p/2} \sum_{m=0}^{\infty} 2^{j_4(1/2+\delta)p} \\
&\leq C (\log n/n)^{sp/(2(s+\delta)+1)} \sum_{m=0}^{\infty} 2^{-mp(1+2\delta)/(4s)} \leq C (\log n/n)^{\alpha_1 p}.
\end{aligned}$$

– *The upper bound of  $T_2$ .* Since  $L \asymp (\log n)^{p/2}$ , for any  $k$  in  $U_{j,K}$ , there exists a constant  $C > 0$  such that

$$\{b_{j,K} \leq \mu 2^{m+1} n^{-1/2} 2^{\delta j}\} \subseteq \left\{ |\beta_{j,k}| \leq C \mu 2^{m+1} 2^{\delta j} \sqrt{(\log n/n)} \right\}. \quad (6.8)$$

Since  $B_{\pi,r}^s(M) \subseteq B_{p,r}^{s-1/\pi+1/p}(M)$  and  $\epsilon > 0$  with  $p > \pi$ , we have

$$\begin{aligned}
T_2 &\leq C \sum_{m=0}^{\infty} 2^{-mp} \left[ \sum_{j=j_4+1}^{j_2} 2^{j(1/2-1/p)} \left( \sum_K \sum_{(K)} |\beta_{j,k}|^p \mathbf{1}_{\{b_{j,K} \leq \mu 2^{\delta j} 2^m n^{-1/2}\}} \right)^{1/p} \right]^p \\
&\leq C (\log n)^{(p-\pi)/2} n^{(\pi-p)/2} \sum_{m=0}^{\infty} 2^{-m\pi} \left[ \sum_{j=j_4+1}^{j_2} 2^{j(1/2-1/p)} 2^{\delta j((p-\pi)/p)} \left( \sum_{k=0}^{2^j-1} |\beta_{j,k}|^{\pi} \right)^{1/p} \right]^p \\
&\leq C (\log n)^{(p-\pi)/2} n^{(\pi-p)/2} \sum_{m=0}^{\infty} 2^{-m\pi} \left( \sum_{j=j_4+1}^{j_2} 2^{-j\epsilon/p} \right)^p \\
&\leq C (\log n)^{(p-\pi)/2} n^{(\pi-p)/2} \sum_{m=0}^{\infty} 2^{-m\pi} 2^{-j_4\epsilon} \\
&\leq C (\log n)^{(p-\pi)/2} n^{(\pi-p)/2} (\log n/n)^{\epsilon/(2(s+\delta)+1)} \sum_{m=0}^{\infty} 2^{-m\pi/2+m(2\delta+1)(\pi-p)/(4s)} \\
&\leq C (\log n/n)^{\alpha_1 p}.
\end{aligned}$$

• *For  $\epsilon < 0$ .*

– *The upper bound of  $T_1$ .* Proceeding in a similar fashion to the upper bound

of  $T_2$  for  $\epsilon > 0$ , we obtain

$$\begin{aligned}
T_1 &\leq C(\log n)^{(p-\pi)/2} n^{(\pi-p)/2} \sum_{m=0}^{\infty} 2^{-m\pi} \left( \sum_{j=\tau}^{j_4} 2^{j(1/2-1/p)} 2^{\delta j((p-\pi)/p)} 2^{-j(s+1/2-1/\pi)\pi/p} \right)^p \\
&\leq C(\log n)^{(p-\pi)/2} n^{(\pi-p)/2} \sum_{m=0}^{\infty} 2^{-m\pi} \left( \sum_{j=\tau}^{j_4} 2^{-j\epsilon/p} \right)^p \\
&\leq C(\log n)^{(p-\pi)/2} n^{(\pi-p)/2} \sum_{m=0}^{\infty} 2^{-m\pi} 2^{-j_4\epsilon} \\
&\leq C(\log n/n)^{\alpha_{2p}} \sum_{m=0}^{\infty} 2^{-m\pi/2+m(2\delta+1)(\pi-p)/(4s)} \leq C(\log n/n)^{\alpha_{2p}}.
\end{aligned}$$

– *The upper bound of  $T_2$ .* Using the property of concentration (2.1) and the inclusion  $B_{\pi,r}^s(M) \subseteq B_{p,\infty}^{s-1/\pi+1/p}(M)$ , we have

$$\begin{aligned}
T_2 &\leq C \sum_{m=0}^{\infty} 2^{-mp} \left[ \sum_{j=j_4+1}^{\infty} 2^{j(1/2-1/p)} \left( \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right]^p \\
&\leq C \sum_{m=0}^{\infty} 2^{-mp} 2^{-j_4(s-1/\pi+1/p)p} \leq C(\log n/n)^{\alpha_{2p}} \sum_{m=0}^{\infty} 2^{-mp/2+(m/2s)(p/\pi-1)} \\
&\leq C(\log n/n)^{\alpha_{2p}}.
\end{aligned}$$

We deduce that

$$Q_1(f) \leq C(\log n/n)^{\alpha_{2p}}.$$

• *For  $\epsilon = 0$ .* The upper bound obtained previously for the term  $T_2$  is always valid. Thus, it suffices to analyze the upper bound of  $T_1$ . Proceeding in a similar fashion to the upper bound of  $T_1$  for  $\epsilon < 0$  and using (6.8), we find

$$T_1 \leq C n^{(\pi-p)/2} (\log n)^{(p-\pi)/2} \sum_{m=0}^{\infty} 2^{-m\pi} \left( \sum_{j=\tau}^{j_4} \Lambda_j \right)^p,$$

where  $\Lambda_j = (2^{j(s+1/2-1/\pi)\pi} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^\pi)^{1/p}$ . Let us investigate separately the case  $\pi \geq rp$  and the case  $\pi < rp$ .

– *For  $\pi \geq rp$ .* The inclusion  $B_{\pi,r}^s(M) \subseteq B_{\pi,\pi/p}^s(M)$  implies  $\sum_{j=\tau}^{\infty} \Lambda_j \leq C$  and a fortiori

$$T_1 \leq C n^{(\pi-p)/2} (\log n)^{(p-\pi)/2} \leq C(\log n/n)^{\alpha_{2p}}.$$



– For  $\pi < rp$ . Using the Holder inequality and the inclusion  $f \in B_{\pi,r}^s(M) \subseteq B_{\pi,\infty}^s(M)$ , we have  $\Lambda_j \leq L$  and  $(\sum_{j=\tau}^{\infty} \Lambda_j^{pr/\pi})^{\pi/r} \leq L$ . Therefore,

$$\begin{aligned} \sum_{m=0}^{\infty} 2^{-m\pi} \left( \sum_{j=j_1}^{j_4} \Lambda_j \right)^p &\leq \sum_{m=0}^{\infty} 2^{-m\pi} \left( \sum_{j=\tau}^{\infty} \Lambda_j^{pr/\pi} \right)^{\pi/r} \left( \sum_{j=\tau}^{j_4} \Lambda_j^{1/(1-\pi/(rp))} \right)^{p-\pi/r} \\ &\leq C \sum_{m=0}^{\infty} 2^{-m\pi} j_4^{(p-\pi/r)} \leq C(\log n)^{(p-\pi/r)}. \end{aligned}$$

Hence,

$$T_1 \leq C(\log n)^{(p-\pi/r)} n^{(\pi-p)/2} (\log n)^{(p-\pi)/2} \leq C(\log n/n)^{\alpha_2 p} (\log n)^{(p-\pi/r)}.$$

Combining the previous inequalities, we obtain the desired upper bounds.

• *The upper bound of  $Q_2(f)$ .* Using the Minkowski inequality, the property of concentration (2.1), the inclusion  $B_{\pi,r}^s(M) \subseteq B_{p,r}^{s-1/\pi+1/p}(M)$  and the fact that  $s > 1/\pi - \delta - 1/2 + 1/(2\nu)$ , we have

$$\begin{aligned} Q_2(f) &\leq C \left[ \sum_{j=j_2+1}^{\infty} 2^{j(1/2-1/p)} \left( \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right]^p \leq C \left( \sum_{j=j_2+1}^{\infty} 2^{-j(s-1/\pi+1/p)} \right)^p \\ &\leq C 2^{-j_2(s-1/\pi+1/p)p} \leq C(n^{-\alpha_1 p} \wedge (\log n/n)^{\alpha_2 p}). \end{aligned} \quad (6.9)$$

We obtain the desired upper bounds according to the sign of  $\epsilon$ .

The proof of Theorem 4.2 is complete.

**Proof of Theorem 5.3.** Let us consider the following lemma.

**Lemma 6.1 (Cirelson's inequality (1976))** *Let  $D$  be a subset of  $\mathbb{R}$  and a centered Gaussian process  $(\eta_t)_{t \in D}$ . If  $E(\sup_{t \in D} \eta_t) \leq N$  and  $\sup_{t \in D} \text{Var}(\eta_t) \leq V$  then, for all  $x > 0$ , we have*

$$P(\sup_{t \in D} \eta_t \geq x + N) \leq \exp(-x^2/(2V)). \quad (6.10)$$

For the proof of the assumption (H1), we refer the reader to Johnstone, Kerkycharian, Picard and Raimondo (2004, Theorem 1). Let us show that the assumption (H2) is satisfied. The aim is to apply the Cirelson inequality (6.10).

Set  $\hat{\theta}_{j,k} = \hat{\beta}_{j,k} - \beta_{j,k} = n^{-1/2} \sum_{l \in C_j} F^*(W)(l) F(g)(l)^{-1} F(\psi_{j,k}^M)(l)$ . Consider the set  $\Omega_q$  defined by  $\Omega_q = \{a = (a_{j,k}); \sum_{(K)} |a_{j,k}|^q \leq 1\}$  and the centered

Gaussian process  $Z(a)$  defined by  $Z(a) = \sum_{(K)} a_{j,k} \hat{\theta}_{j,k}$ . By an argument of duality, we have  $\sup_{a \in \Omega_q} Z(a) = (\sum_{(K)} |\hat{\theta}_{j,k}|^p)^{1/p}$ . Let us analyze the values of  $N$  and  $V$  which appeared in the Cirelson inequality (6.10).

– *Value of  $N$ .* The Holder inequality and the assumption (H1) imply that

$$E(\sup_{a \in \Omega_q} Z(a)) = E(|\sum_{(K)} \hat{\theta}_{j,k}|^{1/p}) \leq [\sum_{(K)} E(|\hat{\theta}_{j,k}|^p)]^{1/p} \leq Cn^{-1/2} l_j^{1/p} 2^{\delta j}.$$

Hence  $N = Cn^{-1/2} l_j^{1/p} 2^{\delta j}$ .

– *Value of  $V$ .* Notice that the assumption (5.1) yields  $|F(g)(l)|^{-2} \asymp 2^{2\delta j}$  for any  $l \in C_j$ . Using the fact that  $F^*(W)(l) \sim N(0, 1)$ , the elementary equality  $E(F^*(W)(l) \overline{F^*(W)(l')}) = \int_0^1 e^{-2i\pi(l-l')t} dt = 1_{\{l=l'\}}$  and the Plancherel inequality, we obtain

$$\begin{aligned} \sup_{a \in \Omega_q} \text{Var}(Z(a)) &= \sup_{a \in \Omega_q} [E(\sum_{k \in U_{j,K}} \sum_{k' \in U_{j,K}} a_{j,k} \hat{\theta}_{j,k} \overline{a_{j,k'} \hat{\theta}_{j,k'}})] \\ &= n^{-1} \sup_{a \in \Omega_q} [ \sum_{k \in U_{j,K}} \sum_{k' \in U_{j,K}} a_{j,k} \overline{a_{j,k'}} \sum_{l \in C_j} \sum_{l' \in C_j} F(g)(l)^{-1} F(\psi_{j,k}^M)(l) \dots \\ &\quad \overline{F(g)(l')}^{-1} \overline{F(\psi_{j,k'}^M)(l')} E(F^*(W)(l) \overline{F^*(W)(l')}) ] \\ &= n^{-1} \sup_{a \in \Omega_q} [ \sum_{k \in U_{j,K}} \sum_{k' \in U_{j,K}} a_{j,k} \overline{a_{j,k'}} \sum_{l \in C_j} |F(g)(l)|^{-2} F(\psi_{j,k}^M)(l) \overline{F(\psi_{j,k'}^M)(l)} ] \\ &\leq Cn^{-1} 2^{2\delta j} \sup_{a \in \Omega_q} [ \sum_{k \in U_{j,K}} \sum_{k' \in U_{j,K}} a_{j,k} \overline{a_{j,k'}} \sum_{l \in C_j} F(\psi_{j,k}^M)(l) \overline{F(\psi_{j,k'}^M)(l)} ] \\ &= Cn^{-1} 2^{2\delta j} \sup_{a \in \Omega_q} [ \sum_{k \in U_{j,K}} \sum_{k' \in U_{j,K}} a_{j,k} \overline{a_{j,k'}} \int_0^1 \psi_{j,k}^M(x) \overline{\psi_{j,k'}^M(x)} dx ] \\ &= Cn^{-1} 2^{2\delta j} \sup_{a \in \Omega_q} (\sum_{k \in U_{j,K}} |a_{j,k}|^2) \leq C 2^{2\delta j} n^{-1}. \end{aligned}$$

Hence  $V = C 2^{2\delta j} n^{-1}$ . By taking  $d$  large enough and  $x = 4^{-1} d n^{-1/2} L^{1/p} 2^{\delta j}$ , the Cirelson inequality (6.10) yields

$$\begin{aligned} P((L^{-1} \sum_{(K)} |\hat{\theta}_{j,k}|^p)^{1/p} \geq 2^{\delta j} 2^{-1} d n^{-1/2}) &\leq P(\sup_{a \in \Omega_q} Z(a) \geq x + N) \\ &\leq \exp(-x^2/(2Q)) \leq \exp(-Cd^2 L^{2/p}). \end{aligned}$$

Since  $L^{2/p} \asymp \log n$ , we prove the assumption (H2) by taking  $d$  large enough. The proof of Theorem 5.3 is complete.

## References

- Cai, T., (1996). Minimax wavelet estimation via block thresholding. *Technical Report, 96-41, Department of Statistics, Purdue University*
- Cai, T., (1997). On adaptivity of BlockShrink wavelet estimator over Besov spaces. *Technical Report 97-05, Department of Statistics, Purdue University.*
- Cai, T., (2002). On block thresholding in wavelet regression: adaptivity, block size and threshold level. *Statist. Sinica.* **12**, 1241-1273.
- Cavalier, L., Tsybakov, A., (2001). Penalized block-wise Stein's method, monotone oracles and sharp adaptive estimation. *Math. Meth. Statist.* **10**, (3), 247-282.
- Cavalier, L., and Tsybakov, A., (2002). Sharp adaptation for inverse problems with random noise. *Probability Theory and Related Fields.* **123**, 323-354.
- Cirelson, B.S., and Ibragimov, I.A., and Sudakov, V.N., (1976). *Norm of Gaussian sample functions*, Springer Verlag, Berlin.
- Cohen, A., and Daubechies, I., and 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.L., and Johnstone, I.M., (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, **90**, 432, 1200-1224.
- Hall, P., Kerkyacharian, G., and Picard, D., (1999). On the minimax optimality of Block thresholded wavelet estimators. *Statist. Sinica*, **9**, 1, 33-49.
- Johnstone, I.M., Kerkyacharian, G., Picard, D., and Raimondo, M., (2004). Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society. Series B. Methodological*, **66**, 547-573,

Kerkyacharian, G., and Picard, D., (2000). Thresholding algorithms, maxisets and well concentrated bases. *Test*, **9**, 2, 283-345.

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,

Laboratoire de Probabilités et Modèles Aléatoires, LPMA, Université Paris VI,  
175 rue du Chevaleret, 75013 PARIS.

E-mail : christophe.chesneau4@wanadoo.fr (or chesneau@proba.jussieu.fr)