

BlockShrink Wavelet Density Estimator in ϕ -Mixing Framework

Mohammed Badaoui and Nouredine Rhomari

Abstract We study the integrated L_2 -risk, of a wavelet BlockShrink density estimator based on dependent observations. We prove that the BlockShrink estimator is adaptive in a class of Sobolev space with unknown regularity for uniformly mixing processes with arithmetically decreasing mixing coefficients.

1 Introduction

The functional estimation by the wavelet method has been intensively used, these last years, in various areas. The popularity of these methods comes from the ease of their implementation, their flexibility, ability to catch details, and for their high compression ratio. In the statistical literature, different types of wavelet estimators have been proposed. The performance of the first ones depended on the density's regularity. Later, adaptive procedures, as thresholding estimators, were developed to construct an estimate which does not depend on the explicit knowledge of this regularity. Those estimators are to make a fine selection of the coefficients estimators $\hat{\beta}_{jk}$ of the wavelets coefficient β_{jk} and several thresholding techniques, including local, global, and block thresholdings, have been developed.

Donoho and Johnstone developed the theory of thresholding in a general framework in the beginning of 1990s. They introduced two techniques of local thresholding: soft and hard thresholding. The word “local” means that individual coefficients independently of each other are subject to a possible thresholding.

M. Badaoui

Ecole Nationale des Sciences Appliquées (ENSA) de Khouribga, Université Hassan 1er, Settat, Morocco

LaMSD et URAC 04, Département de Mathématiques et Informatique, Faculté des Sciences, Université Mohamed 1er, Oujda, Morocco

e-mail: med_badaoui79@yahoo.fr; mohammed.badaoui@uhp.ac.ma

N. Rhomari (✉)

LaMSD et URAC 04, Département de Mathématiques et Informatique, Faculté des Sciences, Université Mohamed 1er, Oujda, Morocco

e-mail: nrhomari@yahoo.fr; n.rhomari@ump.ma

The idea of block thresholding was introduced by Efroimovich [9] as part of Fourier analysis. It has been adapted to the wavelet context analysis by Kerkycharian et al. [13]. They proposed a global thresholding. Its principle, instead of keeping or deleting individual wavelet coefficients, one can also keep or delete a whole j -level of coefficients.

The first localized block thresholding estimators have been developed, in case of iid observations, by Hall et al. [10, 11], Cai [3–6], and Chesneau [7]. This last author studied an L_p version of the BlockShrink estimator given by Cai [3] and proved by simulations that BlockShrink estimator is better than the local and global ones. All these works have been developed in the independent framework. Tribouley and Viennet [20] explored the global thresholding method in the case of β -mixing processes. But to our knowledge, the BlockShrink estimator for the density model has not been studied for dependent processes.

The aim of this work is to extend some results, of BlockShrink estimator, to dependent processes in L_2 -norm for the density estimation. The function density f is supposed to belong to the Sobolev space $\mathbf{H}^s$ with compact support. We consider the ϕ -mixing's processes [12] and we study the L_2 -error convergence for BlockShrink estimator f_n . We show that BlockShrink estimator is consistent with an optimal rate, under certain conditions on wavelet basis and mixing coefficients. Precisely we give for arithmetically decreasing ϕ -mixing processes an upper bound of the L_2 -mean error

$$\mathbb{E}\|f - f_n\|_{L_2}^2 = Cn^{-\frac{2s}{1+2s}}.$$

The rest of the paper is organized as follows: Sect. 2 describes briefly the wavelet basis and the Sobolev space. Section 3 introduces the BlockShrink estimator and presents its asymptotic properties under ϕ -mixing conditions. Section 4 contains proofs of various results.

2 Wavelets and Sobolev Space

In this section, we give a brief definition of wavelet decomposition and we refer to Meyer [16] and Mallat [14, 15] for more details. Without loss of generality, we suppose that f has $[0, 1]$ as support and consider a wavelet basis on the interval $[0, 1]$ of the form

$$\zeta = \{\varphi_{\tau k}, \tau \geq 0, k = 0, \dots, 2^j - 1; \psi_{jk}, j \geq \tau; k = 0, \dots, 2^j - 1\}.$$

In general, $\varphi_{jk}(x)$ and $\psi_{jk}(x)$ are “periodic” or “boundary adjusted” dilation and translation of a “father” wavelet φ and a “mother” wavelet ψ , respectively. This last function is supposed to be N -regular (cf. Meyer [16]).

For the sake of simplicity, we set $\varphi_{jk}(x) = 2^{j/2}\varphi(2^jx-k)$ and $\psi_{jk} = 2^{j/2}\psi(2^jx-k)$, where φ and ψ are father and mother wavelet functions, respectively. Let, now, τ be a large enough integer. For any $j_0 \geq \tau$, a function f in $L_2([0, 1])$ can be expanded in a wavelet series as

$$f = \sum_{k=0}^{2^{j_0}-1} \alpha_{j_0k} \varphi_{j_0k} + \sum_{j \geq j_0} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk}, \quad (1)$$

where the wavelet coefficients are defined by

$$\alpha_{j_0k} = \int f(x) \varphi_{j_0k}(x) dx \quad \text{et} \quad \beta_{jk} = \int f(x) \psi_{jk}(x) dx.$$

Now, let us give a definition of Sobolev space, the main function spaces used in our study. To unify notations we write $\beta_{\tau-1,k} = \alpha_{\tau k}$. We define the Sobolev space $\mathbf{H}^s$ as being the functions space f whose the associated wavelet coefficients β_{jk} , satisfy

$$\left(\sum_{j \geq \tau-1} 2^{2js} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right) \right)^{1/2} < \infty,$$

where $s \in]0, N + 1[$ with N denotes the wavelet regularity (cf. Meyer [16]).

3 BlockShrink Estimator

Let $X_1, \dots, X_n$, n observations from a strictly stationary process, with unknown density f , with respect to the Lebesgue measure on $\mathbf{R}$, which is supposed to be in $L_2([0, 1])$.

Then f admits a representation in (1) with $\alpha_{j_0k} = \mathbb{E}(\varphi_{j_0k}(X))$ and $\beta_{jk} = \mathbb{E}(\psi_{jk}(X))$, where X is a random variable having as probability density f .

We define the nonlinear BlockShrink estimator as proposed by Cai [3] by

$$f_n^b = \sum_{k=0}^{2^{j_0}-1} \hat{\alpha}_{j_0k} \varphi_{j_0k} + \sum_{j=j_0}^{j_1} \sum_{K \in \mathcal{A}_j} \sum_{k \in \mathcal{B}_{jK}} \hat{\beta}_{jk} I_{\{\hat{b}_{jK} \geq \lambda\}} \psi_{jk}, \quad (2)$$

where j_0 and j_1 are two integers to be precised below such that: j_0 is an integer chosen so that the linear variance term will not contribute to the overall error in the same spirit that the integer j_1 is chosen to make the bias term negligible in the overall error. We assume that $2^{j_1} \asymp n^{\frac{1}{2}}$ and $2^{j_0} \asymp (\log n)^\varepsilon$ with $\varepsilon > 2$. And for

$j \in \{j_0, \dots, j_1\}$, we split the set $\{0, \dots, 2^j - 1\}$ into blocks of length l_j . We define the sets:

$$\begin{aligned}\mathcal{A}_j &= \{1, \dots, 2^j l_j^{-1}\}, \\ \mathcal{B}_{jK} &= \{k \in \{0, \dots, 2^j - 1\}; (K-1)l_j \leq k \leq Kl_j - 1\},\end{aligned}$$

for $K \in \mathcal{A}_j$, where l_j is an increasing sequence in j such that $l_{j_0} \asymp (\log n)^\varepsilon$, and we get

$$\hat{b}_{jK} = \left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk}|^2 \right)^{\frac{1}{2}}$$

and λ is some threshold parameter, depending on n , to be defined below.

The choice of block size l_j and the threshold λ play a crucial role in the performance of the resulting estimator. Between the two parameters l_j and λ , the block l_j size is more important. It plays an important role similar to that of the bandwidth in the kernel estimator. To achieve the optimal adaptability, the block size must be at least of order $\log n$, see, for example, Hall et al. [10].

Before announcing our result, we recall the uniformly mixing dependance we work with. Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{F}$ and $\mathcal{G}$ be two sub σ -fields of $\mathcal{A}$. In order to estimate the dependence between $\mathcal{F}$ and $\mathcal{G}$ we use the uniform mixing (or ϕ -mixing) coefficient [12]

$$\phi(\mathcal{F}, \mathcal{G}) = \sup\{|\mathbb{P}(B) - \mathbb{P}(B/A)|; B \in \mathcal{G}, A \in \mathcal{F} \text{ and } \mathbb{P}(A) > 0\}.$$

When $\mathcal{F} = \sigma(X)$ and $\mathcal{G} = \sigma(Y)$ are sigma algebras generated, respectively, by random variables X and Y , we write $\phi(\mathcal{F}, \mathcal{G}) = \phi(X, Y)$.

Remark that, this measure of dependence is not symmetric with respect to $\mathcal{F}$ and $\mathcal{G}$. Therefore we can define an other coefficient by

$$\phi_{\text{rev}}(\mathcal{F}, \mathcal{G}) = \phi(\mathcal{G}, \mathcal{F}).$$

A stationary process $\mathbf{X} = (X_j, j \in \mathbf{Z})$ is said to be ϕ -mixing if

$$\phi(l) = \phi(\mathcal{F}_{-\infty}^0, \mathcal{F}_l^{+\infty}) \rightarrow 0 \text{ when } l \rightarrow \infty,$$

where $\mathcal{F}_i^k$ is the sigma algebra generated by the random variables $\{X_j, i \leq j \leq k\}$. In the same way, we can also define a ϕ_{rev} -mixing processes.

The sequence of the mixing coefficients $(\phi(l))_{l \geq 0}$ decrease arithmetically if there exists $\theta > 0$ and a constant $c > 0$ such that $\phi(l) \leq cl^{-(\theta+1)}$; and we say $\mathbf{X}$ is an arithmetically ϕ -mixing process.

We now give Theorem 1 providing an upper bound of the L_2 error of the BlockShrink estimator defined by (2).

Theorem 1 *Assume that f belongs to the class $\mathcal{F}(M_1, M_2, A) = \{f \in \mathbf{H}^s, \text{supp}(f) \subset [A, -A], \|f\|_{\mathbf{H}^s} \leq M_1, \|f\|_\infty \leq M_2\}$, with $1/2 < s < N + 1$, and the mixing coefficients decrease arithmetically with $\theta > 6 - \frac{4}{1+2(N+1)}$, then for*

$\lambda = 4\|\psi\|_{L_1}\|\psi\|_\infty\|f\|_\infty(\sum_{l \geq 0} \phi(l))/n^{1/2}$ there exists a positive constant C such that

$$\mathbb{E}\|f - f_n^b\|_{L_2}^2 \leq Cn^{-\frac{2s}{1+2s}}.$$

Corollary 1 *Under the same conditions of Theorem 1, the estimator f_n^b is adaptive in the class $\{\mathcal{F}(M_1, M_2, A), 1/2 < s < N + 1, A > 0, M_1 > 0, M_2 > 0\}$.*

Proofs of our main theorem are based on the following preliminary results, which are important in themselves.

Proposition 1 *Let us set $\hat{\beta}_{j_0-1k} = \hat{\alpha}_{j_0k}$, under the summability $\sum_{l \geq 0} \phi(l) \leq B < \infty$, we have for all $j \in \{j_0 - 1, \dots, j_1\}$*

$$\mathbb{E}|\hat{\beta}_{jk} - \beta_{jk}|^2 \leq 4B\|\psi\|_{L_1}\|\psi\|_\infty\|f\|_\infty \frac{1}{n}.$$

Proposition 2 *Assume that there exists $\theta > 0$ and a constant $c > 0$ such that $\phi(l) \leq cl^{-(\theta+1)}$, then there exists two positive constants C and K_2 such that, for all j*

$$\mathbb{P}\left(\left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2\right)^{1/2} \geq 2K_2/n^{1/2}\right) \leq Cn^{-\frac{\theta}{2}} \left(\frac{2^j}{l_j}\right)^{\frac{1}{2}(2+\theta)} (\log n)^{\theta+1},$$

where $K_2 = (4B\|\psi\|_{L_1}\|\psi\|_\infty\|f\|_\infty)^{\frac{1}{2}}$ and $B = \sum_{l \geq 0} \phi(l)$.

4 Derivations

In what follows, we use the symbol C as a generic positive constant, independent of n , which may take different values at different places.

Let us first prove Propositions 1 and 2. We are inspired by techniques developed in Tribouley and Viennet [20].

Proof of Proposition 1 It is easy to have

$$\begin{aligned}\mathbb{E} \left(|\hat{\beta}_{jk} - \beta_{jk}|^2 \right) &= n^{-2} \mathbb{E} \left(\left| \sum_{i=1}^n \left(\psi_{jk}(X_i) - \mathbb{E}(\psi_{jk}) \right) \right|^2 \right) \\ &\leq 2n^{-2} \sum_{i=0}^n (n-i) |\text{Cov}(\psi_{jk}(X_0), \psi_{jk}(X_i))|.\end{aligned}$$

From, Rio [17, Theorem 1.4, p. 27], we deduce

$$\begin{aligned}|\text{Cov}(\psi_{jk}(X_0), \psi_{jk}(X_i))| \\ \leq 2\phi^{\frac{1}{p}}(\psi_{jk}(X_i), \psi_{jk}(X_0))\phi^{\frac{1}{q}}(\psi_{jk}(X_0), \psi_{jk}(X_i))\mathbb{E}(\|\psi_{jk}(X_0)\|_{L_q})\mathbb{E}(\|\psi_{jk}(X_0)\|_{L_p}).\end{aligned}$$

Then, for $q = 1$ and $p = \infty$ we have

$$\mathbb{E}(\|\psi_{jk}(X_0)\|_{L_1}) \leq 2^{-\frac{i}{2}} \|\psi\|_{L_1} \|f\|_{\infty} \text{ and } \mathbb{E}(\|\psi_{jk}(X_0)\|_{\infty}) \leq 2^{\frac{i}{2}} \|\psi\|_{\infty},$$

and finally, we obtain the desired inequality

$$\begin{aligned}\mathbb{E} \left(|\hat{\beta}_{jk} - \beta_{jk}|^2 \right) &\leq 4n^{-1} \sum_{i=0}^n \phi(\psi_{jk}(X_0), \psi_{jk}(X_i)) \mathbb{E}(\|\psi_{jk}(X_0)\|_{L_1}) \|\psi_{jk}(X_0)\|_{\infty} \\ &\leq 4n^{-1} \|\psi\|_{L_1} \|\psi\|_{\infty} \|f\|_{\infty} \sum_{i=0}^n \phi(i).\end{aligned}$$

■

Proof of Proposition 2 To prove Proposition 2, we need the following theorem which gives an exponential type inequality of block of $(\hat{\beta}_{jk})$ coefficients.

Theorem 2 *Let $(X_i)_{i \geq 0}$ be a strictly stationary uniformly mixing process such that its sequence of mixing coefficients $(\phi(l))_{l \geq 0}$ satisfies the summability $\sum_{l \geq 0} \phi(l) \leq B < \infty$. We suppose that X_i has a density f with respect to the Lebesgue measure and that f is uniformly bounded. Then, there exists a positive constant K_1 depending on $\|f\|_{\infty}$ and K_2 such that for any integer $q \leq n$, $\lambda_1 > 0$ and $\lambda_2 > 0$ we have*

$$\begin{aligned}\mathbb{P} \left(\left(\sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{\frac{1}{2}} \geq \lambda_1 + K_2 l_j^{\frac{1}{2}} \frac{1}{\sqrt{n}} + \lambda_2 \right) \\ \leq \exp \left(-K_1 n \left(\lambda_1^2 l_j^{-1} \wedge \frac{\lambda_1}{q 2^{\frac{j}{2}}} \wedge \frac{\lambda_1^2 \sqrt{n}}{q l_j^{\frac{1}{2}} 2^{\frac{j}{2}}} \right) \right) + \frac{2C}{\lambda_2} 2^{\frac{j}{2}} \phi(q)\end{aligned}$$

with $K_2 = (4B \|\psi\|_{L_1} \|\psi\|_{\infty} \|f\|_{\infty})^{\frac{1}{2}}$ and $x \wedge y = \min(x, y)$.

The proof of this theorem is postponed to the end.

Let us return to the proof of Proposition 2. Applying this Theorem 2 with: $\lambda_1 = \lambda_2 = \frac{1}{2}K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}}$, we get: for all $q \leq n$

$$\begin{aligned}
& \mathbb{P} \left(\left(\sum_{k \in \mathcal{B}_{JK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{\frac{1}{2}} \geq 2K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}} \right) \\
& \leq \exp \left(-K_1n \left(\left(\frac{1}{2}K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}} \right)^2 l_j^{-1} \wedge \frac{\frac{1}{2}K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}}}{q2^{\frac{j}{2}}} \wedge \frac{\left(\frac{1}{2}K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}} \right)^2 \sqrt{n}}{ql_j^{\frac{1}{2}}2^{\frac{j}{2}}} \right) \right. \\
& \quad \left. + \frac{2C}{\frac{1}{2}K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}}} 2^{\frac{j}{2}}\phi(q) \right) \\
& \leq \exp \left(-K_1n \left(\frac{1}{n} \wedge \frac{l_j^{\frac{1}{2}}}{q\sqrt{n}2^{\frac{j}{2}}} \wedge \frac{l_j^{\frac{1}{2}}}{q\sqrt{n}2^{\frac{j}{2}}} \right) \right) + C\sqrt{nl_j}^{-\frac{1}{2}}2^{\frac{j}{2}}\phi(q) \\
& \leq \exp \left(-K_1\frac{1}{q}l_j^{\frac{1}{2}}2^{-\frac{j}{2}}\sqrt{n} \right) + C\sqrt{nl_j}^{-\frac{1}{2}}2^{\frac{j}{2}}\phi(q).
\end{aligned}$$

Now for $q = [l_j^{\frac{1}{2}}2^{-\frac{j}{2}}\sqrt{n}(\eta \log n)^{-1}]$ with η a positive constant to be precised, we have

$$\begin{aligned}
& \mathbb{P} \left(\left(\sum_{k \in \mathcal{B}_{JK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{\frac{1}{2}} \geq 2K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}} \right) \\
& \leq \exp(-K_1\eta \log n) + C\sqrt{nl_j}^{-\frac{1}{2}}2^{\frac{j}{2}}\phi \left(l_j^{\frac{1}{2}}2^{-\frac{j}{2}}\sqrt{n}(\eta \log n)^{-1} \right).
\end{aligned}$$

So for a good choice of η , the exponential term will be less than the mixing part (ie. the last term in the above inequality, thanks to the arithmetically decrease of ϕ), which gives

$$\mathbb{P} \left(\left(\sum_{k \in \mathcal{B}_{JK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{\frac{1}{2}} \geq 2K_2l_j^{\frac{1}{2}}\frac{1}{\sqrt{n}} \right) \leq C\sqrt{nl_j}^{-\frac{1}{2}}2^{\frac{j}{2}}\phi \left(l_j^{\frac{1}{2}}2^{-\frac{j}{2}}\sqrt{n}(\eta \log n)^{-1} \right).$$

As the mixing is arithmetic, $\phi(l) \leq c l^{-(\theta+1)}$, we get

$$\begin{aligned} \mathbb{P} \left(\left(\sum_{k \in \mathcal{B}_{jk}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{\frac{1}{2}} \geq 2 \frac{K_2 l_j^{\frac{1}{2}}}{n^{\frac{1}{2}}} \right) &\leq C \sqrt{n} l_j^{-\frac{1}{2}} 2^{\frac{j}{2}} \left(l_j^{\frac{1}{2}} 2^{-\frac{j}{2}} \sqrt{n} (\eta \log n)^{-1} \right)^{-(\theta+1)} \\ &\leq C n^{-\frac{\theta}{2}} \left(\frac{2^j}{l_j} \right)^{\frac{1}{2}(2+\theta)} (\log n)^{(\theta+1)}. \end{aligned}$$

■

Proof of Theorem 1 The proof uses techniques from conventional calculations. In fact, recall that

$$\begin{aligned} f &= \sum_{k=0}^{2^{j_0}-1} \alpha_{j_0 k} \varphi_{j_0 k} + \sum_{j \geq j_0} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \\ f_n &= \sum_{k=0}^{2^{j_0}-1} \hat{\alpha}_{j_0 k} \varphi_{j_0 k} + \sum_{j=j_0}^{j_1} \sum_{K \in \mathcal{A}_j} \sum_{k \in \mathcal{B}_{jK}} \hat{\beta}_{jk} I_{\{\hat{b}_{jK} \geq \lambda\}} \psi_{jk} \end{aligned}$$

with $\lambda/2 = 2K_2/n^{1/2}$. It is obvious that

$$\begin{aligned} \mathbb{E} \|f - f_n\|_2^2 &\leq 3 \left(\mathbb{E} \left\| \sum_{k=0}^{2^{j_0}-1} (\hat{\alpha}_{j_0 k} - \alpha_{j_0 k}) \varphi_{j_0 k} \right\|_{L_2}^2 \right. \\ &\quad \left. + \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} I_{\{\hat{b}_{jK} \geq \lambda\}} - \beta_{jk}) \psi_{jk} \right\|_{L_2}^2 + \left\| \sum_{j \geq j_1+1} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \right\|_{L_2}^2 \right). \end{aligned} \quad (3)$$

1. **Bias term:** Applying the Hölder's inequality and the fact that $f \in \mathcal{F}(M_1, M_2, B)$ we obtain

$$\begin{aligned} \left\| \sum_{j \geq j_1+1} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \right\|_{L_2}^2 &= \left(\sum_{j \geq j_1+1} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} \right)^2 \\ &= \left(\sum_{j \geq j_1+1} 2^{js} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} 2^{-js} \right)^2 \end{aligned}$$

$$\begin{aligned}
&\leq \left(\left(\sum_{j \geq j_1+1} 2^{2js} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right) \right)^{1/2} \left(\sum_{j \geq j_1+1} 2^{-2js} \right)^{1/2} \right)^2 \\
&\leq 2^{-2j_1 s} \leq C n^{-s} \leq C n^{-\frac{2s}{1+2s}}.
\end{aligned} \tag{4}$$

2. **Linear stochastic term:** Using the Proposition 1, we obtain

$$\begin{aligned}
\mathbb{E} \left\| \sum_{k=0}^{2^{j_0}-1} (\hat{\alpha}_{j_0 k} - \alpha_{j_0 k}) \varphi_{j_0 k} \right\|_{L_2}^2 &= C \sum_{k=0}^{2^{j_0}-1} \mathbb{E} |\hat{\alpha}_{j_0 k} - \alpha_{j_0 k}|^2 \\
&\leq C \left(\frac{2^{j_0}}{n} \right) \leq C n^{-\frac{2s}{1+2s}}.
\end{aligned} \tag{5}$$

3. **Nonlinear stochastic term:** Using techniques similar to those used in the proofs of Theorems 2 and 3 of Chesneau [8], we give

$$\begin{aligned}
&\mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \left(\hat{\beta}_{jk} I_{\{\hat{b}_{jK} \geq \lambda\}} - \beta_{jk} \right) \psi_{jk} \right\|_{L_2}^2 \\
&\leq 2 \left(\mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{\hat{b}_{jK} < \lambda\}} \psi_{jk} \right\|_{L_2}^2 + \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \left(\hat{\beta}_{jk} - \beta_{jk} \right) I_{\{\hat{b}_{jK} \geq \lambda\}} \psi_{jk} \right\|_{L_2}^2 \right) \\
&= 2(G_2 + G_3).
\end{aligned}$$

(a) The upper bound for G_2 . We have $G_2 \leq 2(G_{21} + G_{22})$, where

$$\begin{aligned}
G_{21} &= \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{\hat{b}_{jK} < \lambda\}} I_{\{b_{jK} \leq 2\lambda\}} \psi_{jk} \right\|_{L_2}^2 \\
G_{22} &= \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{\hat{b}_{jK} < \lambda\}} I_{\{b_{jK} > 2\lambda\}} \psi_{jk} \right\|_{L_2}^2.
\end{aligned}$$

Notice that the l_2 Minkowski's inequality yields

$$\begin{aligned}
I_{\{\hat{b}_{jK} < \lambda\}} I_{\{b_{jK} > 2\lambda\}} &\leq I_{\{|\hat{b}_{jK} - b_{jK}| \geq \lambda\}} \\
&\leq I_{\left\{ \left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{1/2} \geq \lambda \right\}}.
\end{aligned} \tag{6}$$

- For bounding the term G_{22} we will use the Large deviation established in Proposition 2. Using the inequality (6) and the Proposition 2 we obtain

$$\begin{aligned}
G_{22} &\leq \mathbb{E} \int \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} |\beta_{jk}|^2 I_{\{\hat{b}_{jK} < \lambda\}} I_{\{b_{jK} > 2\lambda\}} |\psi_{jk}(x)|^2 dx \\
&\leq \int \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} |\beta_{jk}|^2 |\psi_{jk}(x)|^2 \mathbb{E} \left[I_{\{\hat{b}_{jK} < \lambda\}} I_{\{b_{jK} > 2\lambda\}} \right] dx \\
&\leq \int \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} |\beta_{jk}|^2 |\psi_{jk}(x)|^2 P \left(\left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{1/2} \geq \lambda \right) dx \\
&\leq \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \left[P \left(\left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{1/2} \geq \lambda \right) \right]^{1/2} \right\|_{L_2}^2 \\
&\leq C n^{-\frac{\theta}{2}} l_{j_1}^{-\frac{1}{2}(2+\theta)} 2^{\frac{j_1}{2}(2+\theta)} (\log n)^{(\theta+1)} \left\| \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \right\|_{L_2}^2 \\
&\leq C n^{-\frac{\theta}{2}} \left(\frac{2^{j_1}}{l_{j_1}} \right)^{\frac{1}{2}(2+\theta)} (\log n)^{(\theta+1)} \|f\|_{L_2}^2.
\end{aligned}$$

This quantity is bounded by $n^{-\frac{2s}{1+2s}}$ for $2^{j_1} \asymp n^{\frac{1}{2}}$ with $\theta > 6 - \frac{4}{1+2(N+1)}$.

We then obtain for $\theta > 6 - \frac{4}{1+2(N+1)}$ that

$$G_{22} \leq C n^{-\frac{2s}{1+2s}}. \quad (7)$$

- The upper bound for G_{21} . Let j_s the integer belonging to $[j_0, j_1]$ such that j_s is the optimal level of truncation, when the regularity s of the density f is known (i.e $2^{j_s} \asymp n^{\frac{1}{1+2s}}$). So using an elementary inequality of convexity, we have the following decomposition:

$$\begin{aligned}
G_{21} &\leq C \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{b_{jK} \leq 2\lambda\}} \psi_{jk} \right\|_{L_2}^2 \\
&\leq C \left(\left\| \sum_{j=j_0}^{j_s} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{b_{jK} \leq 2\lambda\}} \psi_{jk} \right\|_{L_2}^2 + \left\| \sum_{j=j_s+1}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{b_{jK} \leq 2\lambda\}} \psi_{jk} \right\|_{L_2}^2 \right)
\end{aligned}$$

$$\begin{aligned}
&\leq C \left(\left\| \sum_{j=j_0}^{j_s} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{b_{jK} \leq 2\lambda\}} \psi_{jk} \right\|_{L_2}^2 + \left\| \sum_{j=j_s+1}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \right\|_{L_2}^2 \right) \\
&= C(S_1 + S_2).
\end{aligned} \tag{8}$$

- The upper bound for S_1 . Minkowski's inequality gives the following bound:

$$\begin{aligned}
S_1 &= \left\| \sum_{j=j_0}^{j_s} \sum_{k=0}^{2^j-1} \beta_{jk} I_{\{b_{jK} \leq 2\lambda\}} \psi_{jk} \right\|_{L_2}^2 \\
&= \left(\sum_{j=j_0}^{j_s} \left(\sum_{k=0}^{2^j-1} |\beta_{jk} I_{\{b_{jK} \leq 2\lambda\}}|^2 \right)^{1/2} \right)^2.
\end{aligned}$$

Thus

$$\begin{aligned}
S_1 &= C \left(\sum_{j=j_0}^{j_s} \left(\sum_{K \in \mathcal{A}_j} \sum_{k \in \mathcal{B}_{jK}} |\beta_{jk}|^2 I_{\{l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\beta_{jk}|^2\}^{1/2} \leq 2\lambda\}} \right)^{1/2} \right)^2 \\
&\leq C \left[\sum_{j=j_0}^{j_s} (\text{card} \mathcal{A}_j l_j \lambda^2)^{1/2} \right]^2 \\
&\leq C \left[\sum_{j=j_0}^{j_s} 2^{j/2} \lambda \right]^2 \\
&\leq C 2^{j_s} \lambda^2.
\end{aligned}$$

- The upper bound for S_2 . The Minkowski's inequality, the basis orthonormality and the inclusion $\mathbf{H}^s \subseteq \mathbf{B}_{2,\infty}^s$ imply that

$$\begin{aligned}
S_2 &= \left\| \sum_{j=j_s+1}^{j_1} \sum_{k=0}^{2^j-1} \beta_{jk} \psi_{jk} \right\|_{L_2}^2 = \left(\sum_{j=j_s+1}^{j_1} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} \right)^2 \\
&= \left(\sum_{j=j_s+1}^{j_1} 2^{js} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} 2^{-js} \right)^2.
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
 S_2 &\leq \left(\sup_{j \geq j_0} 2^{js} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} \sum_{j=j_s+1}^{j_1} 2^{-js} \right)^2 \\
 &\leq \left(\|f\|_{\mathbf{B}_{2,\infty}^s} \sum_{j=j_s+1}^{j_1} 2^{-js} \right)^2 \\
 &\leq C \left(2^{-j_s s} \sum_{j=1}^{j_1} 2^{-js} \right)^2 \\
 &\leq C 2^{-2j_s s}.
 \end{aligned}$$

Putting the upper bounds of S_1 and S_2 together in inequality (8), it yields

$$G_{21} \leq C (2^{-2j_s s} + 2^{j_s} \lambda^2).$$

Since $\lambda = 4K_2/n^{1/2}$ and $2^{j_s} \asymp n^{\frac{1}{1+2s}}$, we conclude that

$$G_{21} \leq C n^{-\frac{2s}{1+2s}}. \quad (9)$$

(b) For the term G_3 , we have

$$\begin{aligned}
 G_3 &= \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jK} \geq \lambda\}} \psi_{jk} \right\|_{L_2}^2 \\
 &\leq 2 \left(\mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jK} \geq \lambda\}} I_{\{b_{jK} \leq \lambda/2\}} \psi_{jk} \right\|_{L_2}^2 \right. \\
 &\quad \left. + \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jK} \geq \lambda\}} I_{\{b_{jK} > \lambda/2\}} \psi_{jk} \right\|_{L_2}^2 \right) \\
 &= 2 (G_{31} + G_{32}). \quad (10)
 \end{aligned}$$

- To bound the term G_{31} we will use the Large deviation established in Proposition 2. Using l_2 Minkowski's inequality, Hölder's inequality and

Jensen's inequality we obtain

$$\begin{aligned}
G_{31} &= \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jK} \geq \lambda\}} I_{\{b_{jK} < \lambda/2\}} \psi_{jk} \right\|_{L_2}^2 \\
&\leq \mathbb{E} \int \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} |\hat{\beta}_{jk} - \beta_{jk}|^2 I_{\{(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2)^{1/2} \geq \lambda/2\}} |\psi_{jk}(x)|^2 dx \\
&\leq \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \int \mathbb{E} \left(|\hat{\beta}_{jk} - \beta_{jk}|^2 I_{\{(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2)^{1/2} \geq \lambda/2\}} \right) |\psi_{jk}(x)|^2 dx \\
&\leq \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \left(\mathbb{E} |\hat{\beta}_{jk} - \beta_{jk}|^4 \right)^{\frac{1}{2}} \\
&\quad \left(\mathbb{E} \left(I_{\{(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2)^{1/2} \geq \lambda/2\}} \right) \right)^{\frac{1}{2}} \|\psi_{jk}\|_{L_2}^2 \\
&\leq C \sum_{j=j_0}^{j_1} \left(\sum_{k=0}^{2^j-1} \mathbb{E} |\hat{\beta}_{jk} - \beta_{jk}|^4 \right)^{\frac{1}{2}} \\
&\quad \left(P \left(\left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 \right)^{1/2} \geq \lambda/2 \right) \right)^{\frac{1}{2}}.
\end{aligned}$$

Using our Proposition 2 and the Proposition 4.1 from Tribouley and Viennet [20] we get

$$\begin{aligned}
G_{31} &\leq C \sum_{j=j_0}^{j_1} \frac{2^j}{n^2} \left(n^{-\frac{\theta}{2}} l_j^{-\frac{1}{2}(2+\theta)} 2^{\frac{j}{2}(2+\theta)} (\log n)^{1+\theta} \right)^{\frac{1}{2}} \\
&\leq C \frac{2^{j_1}}{n^{2+\frac{\theta}{4}}} \left(\frac{2^{j_1}}{l_{j_1}} \right)^{\frac{1}{4}(2+\theta)} (\log n)^{\frac{\theta+1}{2}} \\
&\leq C n^{-\frac{2s}{1+2s}}.
\end{aligned} \tag{11}$$

- The upper bound for G_{32} . Let j_3 an integer belonging to $[j_0, j_1]$ and M_1 such that $\|f\|_{\mathbf{H}^s} \leq M_1$. Since $\left(\sum_{j \geq \tau-1} 2^{2js} \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right) \right)^{1/2} < \infty$. We have, for

all $j \in [j_3, j_1]$

$$\left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} \leq M_1 2^{-js} \leq M_1 2^{-j_3 s}. \quad (12)$$

Recall that $\lambda/2 = 2K_2/n^{1/2}$ and with choice of $2^{j_3} = \left(\frac{M_1}{2K_2}\right)^{\frac{2}{1+2s}} n^{\frac{1}{1+2s}}$. We will prove that by absurd. Assume that for all $j \in [j_3, j_1]$ we have $b_{jK} > \lambda/2$ then

$$\begin{aligned} \left(l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\beta_{jk}|^2 \right)^{\frac{1}{2}} &> \lambda/2 \\ l_j^{-1} \sum_{k \in \mathcal{B}_{jK}} |\beta_{jk}|^2 &> (\lambda/2)^2 \\ \sum_{k=0}^{2^j-1} |\beta_{jk}|^2 &> (\lambda/2)^2 2^j \\ \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} &> 2^{\frac{j}{2}} \lambda/2 \\ \left(\sum_{k=0}^{2^j-1} |\beta_{jk}|^2 \right)^{1/2} &> 22^{\frac{j_3}{2}} \frac{K_2}{n^{\frac{1}{2}}} = M_1 2^{-j_3 s} \end{aligned}$$

which contradicts (12). We conclude therefor, that the terms corresponding to $j \in [j_3, j_1]$ are zero. Thus

$$\begin{aligned} G_{32} &= \mathbb{E} \left\| \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jK} \geq \lambda\}} I_{\{b_{jK} > \lambda/2\}} \psi_{jk} \right\|_{L_2}^2 \\ &= \mathbb{E} \left\| \sum_{j=j_0}^{j_3} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jK} \geq \lambda\}} I_{\{b_{jK} > \lambda/2\}} \psi_{jk} \right\|_{L_2}^2. \end{aligned}$$

Using the Proposition 1, we get

$$\begin{aligned}
 G_{32} &= \mathbb{E} \left\| \sum_{j=j_0}^{j_3} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) I_{\{\hat{b}_{jk} \geq \lambda\}} I_{\{b_{jk} > \lambda/2\}} \psi_{jk} \right\|_{L_2}^2 \\
 &\leq \mathbb{E} \left\| \sum_{j=j_0}^{j_3} \sum_{k=0}^{2^j-1} (\hat{\beta}_{jk} - \beta_{jk}) \psi_{jk} \right\|_{L_2}^2 \\
 &\leq C \sum_{j=j_0}^{j_3} \sum_{k=0}^{2^j-1} \mathbb{E} |\hat{\beta}_{jk} - \beta_{jk}|^2 \\
 &\leq C \frac{2^{j_3}}{n} \leq C n^{-\frac{2s}{1+2s}}. \tag{13}
 \end{aligned}$$

Combining the previous inequalities (4), (5), (8), (9), (11), and (13) in (3), we obtain the desired upper bounds. $\blacksquare$

Proof of Theorem 2 The scheme of this proof is quite classical. We take inspiration from the proof of Theorem 2.2 in Tribouley and Viennet [20]. The proof is based on a results of Talagrand [19], the Berbee's lemma [1] with the fact that the β -mixing coefficient are less than ϕ -mixing coefficient and mainly on the following observation

Remark 1 The study of the quantity $\sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2$ is related to the study of supremum of a centered empirical process v_n , where $v_n(\psi_{jk}) = \hat{\beta}_{jk} - \beta_{jk}$, over the set $F_2 = \left\{ h = \sum_{k \in \mathcal{B}_{jK}} a_k \psi_{jk} / \sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \leq 1 \right\}$. By duality arguments we have

$$\begin{aligned}
 \sum_{k \in \mathcal{B}_{jK}} |\hat{\beta}_{jk} - \beta_{jk}|^2 &= \sup_{a_k, \sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \leq 1} \left| \sum_{k \in \mathcal{B}_{jK}} a_k (\hat{\beta}_{jk} - \beta_{jk}) \right|^2 \\
 &= \sup_{h \in F_2} |v_n(h)|^2.
 \end{aligned}$$

Lemma 1 (Berbee [1]) *Let $(X_i)_{i>0}$ be a sequence of random variables taking their values in a Polish space χ . Then, there exists a sequence $(X_i^*)_{i>0}$ of independent random variables such that for any positive integer i , we have*

$$\mathbb{P}(X_i \neq X_i^*) < \beta(\sigma(X_j; j < i), \sigma(X_i)).$$

Return now to the proof of Theorem 2. For the sake of simplicity, we assume in the following that $X_i = 0$ and $h(X_i) = 0$ if $i > n$, and that $E_P(h) = 0$. Let $q = [\sqrt{n}]$ be a fixed integer, where $[\cdot]$ denotes the integer part. According to Berbee's lemma,

there exists a sequence of independent random variables $(X_i^*)_{i \geq 0}$ such that for any positive integer k , $Y_k = (X_{qk+1}, \dots, X_{q(k+1)})$ and $Y_k^* = (X_{qk+1}^*, \dots, X_{q(k+1)}^*)$ have the same distribution, and

$$\mathbb{P}(Y_k \neq Y_k^*) \leq \beta(q) \leq \phi(q).$$

Then, the random variables $(Y_{2k}^*)_{k \geq 0}$ are independent, and even $(Y_{2k+1}^*)_{k \geq 0}$.

Let v_n^* be the empirical process associated with the random variables $(X_i^*)_{i \geq 0}$. The centered empirical process $v_n(h)$ (that is $v_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i} - P_{X_1}$) is then decomposed as

$$v_n(h) = v_n(h) - v_n^*(h) + v_n^*(h).$$

Since for any $\lambda_1 > 0$ and $\lambda_2 > 0$ we have

$$\begin{aligned} & \mathbb{P} \left(\sup_{h \in F_2} |v_n(h)| \geq \lambda_1 + K_2 l_j^{\frac{1}{2}} \frac{1}{\sqrt{n}} + \lambda_2 \right) \\ & \leq \mathbb{P} \left(\sup_{h \in F_2} |v_n(h) - v_n^*(h)| \geq \lambda_2 \right) \\ & \quad + \mathbb{P} \left(\sup_{h \in F_2} |v_n^*(h)| \geq \lambda_1 + K_2 l_j^{\frac{1}{2}} \frac{1}{\sqrt{n}} \right). \end{aligned} \quad (14)$$

We just have to control these two terms to get the announced result.

For the first term in the right hand of (14), it is easy to have

$$|v_n(h) - v_n^*(h)| = \frac{1}{n} \left| \sum_{i=1}^n \delta_{X_i}(h) - \delta_{X_i^*}(h) \right| \leq \frac{2}{n} \|h\|_{\infty} \sum_{i=1}^n I_{X_i \neq X_i^*}.$$

We obtain then

$$\mathbb{E}|v_n(h) - v_n^*(h)| \leq \frac{2}{n} \|h\|_{\infty} n \phi(q).$$

Since $h \in F_2 = \left\{ \sum_{k \in \mathcal{B}_{jK}} a_k \psi_{jk} / \sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \leq 1 \right\}$ and using Hölder's inequality, we have

$$|h(x)| \leq \left(\sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k \in \mathcal{B}_{jK}} |\psi_{jk}(x)|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{k=0}^{2^j-1} |\psi_{jk}(x)|^2 \right)^{\frac{1}{2}}, x \in [0, 1].$$

According to Lemma 6 in Birgé and Massart [2], $\left\| \sum_{k=0}^{2^j-1} \psi_{jk}^2 \right\|_{\infty} \leq C 2^j$, we get that

$$\|h\|_{\infty} \leq C 2^{\frac{j}{2}}.$$

Therefore

$$\mathbb{E}|v_n(h) - v_n^*(h)| \leq 2\|h\|_{\infty}\phi(q) \leq 2C 2^{\frac{j}{2}}\phi(q).$$

By Markov inequality, we deduce that for any $\lambda_2 > 0$

$$\mathbb{P}\left(\sup_{h \in F_2} |v_n(h) - v_n^*(h)| \geq \lambda_2\right) \leq \frac{2C}{\lambda_2} 2^{\frac{j}{2}}\phi(q). \quad (15)$$

For the second term in the right hand of (14), we use the following theorem.

Theorem 3 (Talagrand [19]) *Let $X_1, \dots, X_n$, be n independent identically distributed random variables, and F_2 a family of functions that are uniformly bounded by some constant M_1 . Let V and H be defined by*

$$\frac{1}{n} \mathbb{E} \left(\sup_{h \in F_2} \left| \sum_{i=0}^n h(X_i) \right| \right) \leq H, \quad V = \mathbb{E} \left(\sup_{h \in F_2} \sum_{i=0}^n h^2(X_i) \right).$$

Then, there exists a universal constant K_1 such that for any $\lambda_1 > 0$,

$$\mathbb{P} \left(\sup_{h \in F_2} |v_n(h)| \geq \lambda_1 + H \right) \leq \exp \left[-K_1 n \left(\frac{\lambda_1^2}{V} \wedge \frac{\lambda_1}{M_1} \right) \right].$$

Moreover, according to Corollary 3.4 in Talagrand [18], the following bound holds

$$V \leq v + 8M_1\tilde{H},$$

where v and $\tilde{H}$ are defined by

$$\sup_{h \in F_2} \text{Var}(h(X)) \leq v, \quad \frac{1}{n} \mathbb{E} \left(\sup_{h \in F_2} \sum_{i=0}^n \varepsilon_i h(X_i) \right) \leq \tilde{H},$$

where $\varepsilon_1, \dots, \varepsilon_n$ are n independent Rademacher variables.

We derive then the following inequality: there exists a positive constant k_1 such that for any $\lambda_1 > 0$,

$$\mathbb{P} \left(\sup_{h \in F_2} |v_n(h)| \geq \lambda_1 + H \right) \leq \exp \left[-k_1 n \left(\frac{\lambda_1^2}{v} \wedge \frac{\lambda_1}{M_1} \wedge \frac{\lambda_1^2}{8M_1\tilde{H}} \right) \right].$$

Let $p = p(n)$ be the greatest integer such that $n = qp + r$ with $0 \leq r < q$, then

$$v_n^*(h) = \frac{p}{n} \left[\frac{1}{p} \sum_{l=0}^p \left(\sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) + \sum_{i=q(2l+1)+1}^{q(2l+2)} h(X_i^*) \right) \right].$$

The control of the second term in the right hand of (14) is made in two steps, considering odd terms and even ones. They are both treated in the same way, so we only detail the even part. Since the variables $\left(\sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) \right)_{0 \leq l \leq p}$ are independent by construction, we are allowed to apply the Talagrand's inequality to $v_p^* = \frac{1}{p} \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} h(X_i^*)$, with adequate choices of $\lambda_1, H, \tilde{H}, M_1$ and v . Let F_2 be the family of functions defined previously. We have to determine the quantities $M_1, \tilde{H}, H$ and v . Since

$$\left\| \sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) \right\|_{\infty} \leq q \|h\|_{\infty} \leq C q 2^{\frac{j}{2}},$$

we put

$$M_1 = C q 2^{\frac{j}{2}}. \quad (16)$$

Let us now determine the quantity H . Applying Hölder's inequality and Jensen's inequality, we have

$$\begin{aligned} & \mathbb{E} \left[\sup_{h \in F_2} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) \right| \right] \\ &= \mathbb{E} \left[\sup_{h \in F_2} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} \sum_{k \in \mathcal{B}_{jK}} a_k \psi_{jk}(X_i^*) \right| \right] \\ &\leq \mathbb{E} \left[\sup_{h \in F_2} \left(\sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \right)^{\frac{1}{2}} \times \left(\sum_{k \in \mathcal{B}_{jK}} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} \psi_{jk}(X_i^*) \right|^2 \right)^{\frac{1}{2}} \right] \end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{E} \left[\sum_{k \in \mathcal{B}_{jK}} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} \psi_{jk}(X_i^*) \right|^2 \right]^{\frac{1}{2}} \\
&\leq \left[E \sum_{k \in \mathcal{B}_{jK}} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} \psi_{jk}(X_i^*) \right|^2 \right]^{\frac{1}{2}}.
\end{aligned}$$

In order to bound this last term we have, under the summability condition on the mixing coefficients $B < \infty$ and the fact that the variables $\left(\sum_{i=2ql+1}^{q(2l+1)} \psi_{jk}(X_i^*) \right)_{0 \leq l \leq p}$ are independent and satisfy $\mathbb{E}(\psi_{jk}(X_i^*)) = 0$,

$$\begin{aligned}
\mathbb{E} \left[\sum_{k \in \mathcal{B}_{jK}} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} \psi_{jk}(X_i^*) \right|^2 \right] &= \sum_{k \in \mathcal{B}_{jK}} \mathbb{E} \left[\left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} \psi_{jk}(X_i^*) \right|^2 \right] \\
&\leq \sum_{k \in \mathcal{B}_{jK}} p^4 \|\psi\|_{L_1} \|\psi\|_{\infty} \|f\|_{\infty} q \sum_{i=2ql+1}^{q(2l+1)} \phi(i) \\
&\leq 4l_j \|\psi\|_{L_1} \|\psi\|_{\infty} \|f\|_{\infty} p q \sum_{i \geq 0} \phi(i) \\
&\leq 4B \|\psi\|_{L_1} \|\psi\|_{\infty} \|f\|_{\infty} l_j n.
\end{aligned}$$

We deduce then

$$\frac{1}{p} \mathbb{E} \left[\sup_{h \in F_2} \left| \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) \right| \right] \leq K_2 \frac{\sqrt{n}}{p} l_j^{\frac{1}{2}}.$$

We take, now

$$H = K_2 \sqrt{\frac{q}{p}} l_j^{\frac{1}{2}} \quad (17)$$

with

$$K_2 = (4B \|\psi\|_{L_1} \|\psi\|_{\infty} \|f\|_{\infty})^{\frac{1}{2}}.$$

We notice that

$$\frac{1}{n} \mathbb{E} \left(\sup_{h \in F_2} \sum_{i=1}^n \epsilon_i h(X_i) \right) \leq K_2 l_j^{\frac{1}{2}} \sqrt{\frac{q}{p}}.$$

and we choose

$$\tilde{H} = H. \quad (18)$$

Let us now determine the quantity v such that $\sup_{h \in F_2} \text{Var}(h(X)) \leq v$.

Let $h \in \left\{ \sum_{k \in \mathcal{B}_{jK}} a_k \psi_{jk} / \sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \leq 1 \right\}$, we get

$$\begin{aligned} \text{Var} \left(\sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) \right) &\leq qK_2^2 \int h^2 f \\ &\leq qK_2^2 \int \left(\sum_{k \in \mathcal{B}_{jK}} a_k \psi_{jk} \right)^2 f \\ &\leq qK_2^2 \int \left(\left(\sum_{k \in \mathcal{B}_{jK}} |a_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k \in \mathcal{B}_{jK}} \psi_{jk}^2 \right)^{\frac{1}{2}} \right)^2 f \\ &\leq qK_2^2 \int \sum_{k \in \mathcal{B}_{jK}} \psi_{jk}^2 f \\ &\leq K_2^2 \|f\|_{\infty} q l_j. \end{aligned}$$

Thus we set

$$v = K_2^2 \|f\|_{\infty} q l_j. \quad (19)$$

Finally, applying Talagrand's inequality with H , $\tilde{H}$, M_1 and v defined in (17), (18), (16), and (19), we obtain

$$\begin{aligned} &\mathbb{P} \left(\sup_{h \in F_2} \left| \frac{1}{p} \sum_{l=0}^p \sum_{i=2ql+1}^{q(2l+1)} h(X_i^*) \right| \geq q\lambda_1 + K_2 l_j^{\frac{1}{2}} \sqrt{\frac{q}{p}} \right) \\ &\leq \exp \left(-k_1 \mathbb{P} \left(\frac{q^2 \lambda_1^2}{4K_2^2 \|f\|_{\infty} q l_j} \wedge \frac{q\lambda_1}{C q 2^{\frac{j}{2}}} \wedge \frac{q^2 \lambda_1^2}{8C q 2^{\frac{j}{2}} K_2 l_j^{\frac{1}{2}} \sqrt{\frac{q}{p}}} \right) \right) \end{aligned}$$

$$\begin{aligned}
&\leq \exp \left(-k_1 n \left(\frac{\lambda_1^2 l_j^{-1}}{4K_2^2 \|f\|_\infty} \wedge \frac{\lambda_1}{C q 2^{\frac{j}{2}}} \wedge \frac{\lambda_1^2 \sqrt{n}}{q l_j^{\frac{1}{2}} 2^{\frac{j}{2}}} \frac{1}{8CK_2} \right) \right) \\
&\leq \exp \left(-K_1 n \left(\lambda_1^2 l_j^{-1} \wedge \frac{\lambda_1}{q 2^{\frac{j}{2}}} \wedge \frac{\lambda_1^2 \sqrt{n}}{q l_j^{\frac{1}{2}} 2^{\frac{j}{2}}} \right) \right), \tag{20}
\end{aligned}$$

where K_1 is a positive constant depending on $\|f\|_\infty$, K_2 , C and k_1 . Similarly, we find for the odd part

$$\begin{aligned}
&\mathbb{P} \left(\sup_{h \in F_2} \left| \frac{1}{p} \sum_{l=0}^p \sum_{i=q(2l+1)+1}^{q(2l+2)} h(X_i^*) \right| \geq q\lambda_1 + K_2 l_j^{\frac{1}{2}} \sqrt{\frac{q}{p}} \right) \\
&\leq \exp \left(-K_1 n \left(\lambda_1^2 l_j^{-1} \wedge \frac{\lambda_1}{q 2^{\frac{j}{2}}} \wedge \frac{\lambda_1^2 \sqrt{n}}{q l_j^{\frac{1}{2}} 2^{\frac{j}{2}}} \right) \right). \tag{21}
\end{aligned}$$

Regrouping (15), (21), and (20), the proof of Theorem 2 is complete. $\blacksquare$

References

1. Berbee, H.C.P.: Random walks with stationary increments and renewal theory. Cent. Math. Tracts, Amsterdam (1979)
2. Birgé, L., Massart, P.: Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* **4**(3), 329–375 (1998)
3. Cai, T.: Minimax wavelet estimation via block thresholding. Technical Report, pp. 96–41, Department of Statistics, Purdue University (1996)
4. Cai, T.: On adaptivity of blockshrink wavelet estimator over Besov spaces. Technical Report, pp. 97–05, Department of Statistics, Purdue University (1997)
5. Cai, T.: Adaptive wavelet estimation: a block thresholding and oracle inequality approach. *Ann. Stat.* **27**, 898–924 (1999)
6. Cai, T.: On block thresholding in wavelet regression: adaptivity, block size and threshold level. *Stat. Sin.* **12**, 1241–1273 (2002)
7. Chesneau, C.: Quelques contributions à l'estimation fonctionnelle par méthodes d'ondelettes. Ph.D. Thesis, Université Pierre et Marie Curie, Paris VI (2006)
8. Chesneau, C.: Wavelet estimation via block thresholding: a minimax study under L_p -risk. *Stat. Sin.* **18**, 1007–1024 (2008)
9. Efroimovich, S.Y.: Nonparametric estimation of a density of unknown smoothness. *Theory Probab. Appl.* **30**, 557–561 (1985)
10. Hall, P., Kerkycharian, G., Picard, D.: Block thresholding rules for curve estimation using kernel and wavelet methods. *Ann. Stat.* **26**, 942–962 (1998)
11. Hall, P., Kerkycharian, G., Picard, D.: On the minimax optimality of block thresholded wavelet estimators. *Stat. Sin.* **9**(1), 33–49 (1999)
12. Ibragimov, I.A.: Some limit theorems for stationary processus. *Theor. Probab. Appl.* **7**, 349–382 (1962)

13. Kerkyacharian, G., Picard, D., Tribouley, K.: Density estimation in Besov spaces. *Bernoulli* **2**, 229–247 (1996)
14. Mallat, S.: Une exploration des signaux en ondelettes. Ecole Polytechnique Ellipses Edition (2000)
15. Mallat, S.: A Wavelet Tour of Signal Processing, 3rd edn. Academic, New York (2008)
16. Meyer, Y.: Ondelettes et Opérateurs. Hermann, Paris (1990)
17. Rio, E.: Théorie asymptotique des processus aléatoires faiblement dépendants. *Mathématiques & Applications*, vol. 31. Springer, Berlin (2000)
18. Talagrand, M.: Sharper bounds for gaussian and empirical processes. *Ann. Probab.* **22**, 28–76 (1994)
19. Talagrand, M.: New concentration inequalities in product spaces. Technical Report, University Paris VI, Ohio State University (1995)
20. Tribouley, K., Viennet, G.: L_p -adaptive density estimation in a β -mixing framework. *Ann. Inst. Henri Poincaré* **34**, 179–208 (1998)

Functional Statistics and Applications

Selected Papers from MICPS-2013

Ould Said, E.; Ouassou, I.; Rachdi, M. (Eds.)

2015, XII, 164 p. 14 illus., 13 illus. in color., Hardcover

ISBN: 978-3-319-22475-6