Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2014, Article ID 403764, 12 pages
http://dx.doi.org/10.1155/2014/403764

Research Article

A General Result on the Mean Integrated Squared Error of the Hard Thresholding Wavelet Estimator under α-Mixing Dependence

Christophe Chesneau

Laboratoire de Mathéematiques Nicolas Oresme, Université de Caen, BP 5186, 14032 Caen Cedex, France

Correspondence should be addressed to Christophe Chesneau; christophe.chesneau@gmail.com

Received 10 April 2013; Accepted 22 October 2013; Published 19 January 2014

Academic Editor: Zhidong Bai

We consider the estimation of an unknown function f for weakly dependent data (α-mixing) in a general setting. Our contribution is theoretical: we prove that a hard thresholding wavelet estimator attains a sharp rate of convergence under the mean integrated squared error (MISE) over Besov balls without imposing too restrictive assumptions on the model. Applications are given for two types of inverse problems: the deconvolution density estimation and the density estimation in a GARCH-type model, both improve existing results in this dependent context. Another application concerns the regression model with random design.

1. Introduction

A general nonparametric problem is adopted: we aim to estimate an unknown function f via n random variables $V_1, \ldots, V_n$ from a strictly stationary stochastic process $(V_t)_{t\in\mathbb{Z}}$. We suppose that $(V_t)_{t\in\mathbb{Z}}$ has a weak dependence structure; the α-mixing case is considered. This kind of dependence naturally appears in numerous models as Markov chains, GARCH-type models, and discretely observed diffusions (see, e.g., [1–3]). The problems where f is the density of V_1 or a regression function have received a lot of attention. A partial list of related works includes Robinson [4], Roussas [5, 6], Truong and Stone [7], Tran [8], Masry [9, 10], Masry and Fan [11], Bosq [12], and Liebscher [13].

For an efficient estimation of f, many methods can be considered. The most popular of them are based on kernels, splines and wavelets. In this note we deal with wavelet methods that have been introduced in i.i.d.setting by Donoho and Johnstone [14, 15] and Donoho et al. [16, 17]. These methods enjoy remarkable local adaptivity against discontinuities and spatially varying degree of oscillations. Complete reviews and discussions on wavelets in statistics can be found in, for example, Antoniadis [18] and Härdle et al. [19]. In the context of α-mixing dependence, various wavelet methods have been elaborated for a wide variety of nonparametric problems. Recent developments can be found in, for example, Leblanc [20], Tribouley and Viennet [21], Masry [22], Patil and Truong [23], Doosti et al. [24], Doosti and Niroumand [25], Doosti et al. [26], Cai and Liang [27], Niu and Liang [28], Benatia and Yahia [29], Chesneau [30–32], Chaubey and Shirazi [33], and Abbaszadeh and Emadi [34].

In the general dependent setting described above, we provide a theoretical contribution to the performance of a wavelet estimator based on a hard thresholding. This nonlinear wavelet procedure has the features to be fully adaptive and efficient over a large class of functions f (see, e.g., [14–17, 35]). Following the spirit of Kerkyacharian and Picard [36], we determine necessary assumptions on $(V_t)_{t\in\mathbb{Z}}$ and the wavelet basis to ensure that the considered estimator attains a fast rate of convergence under the MISE over Besov balls. The obtained rate of convergence often corresponds to the near optimal one in the minimax sense for the standard i.i.d. case. The originality of our result is to be general and sharp; it can be applied for nonparametric models of different natures and improves some existing results. This fact is illustrated by the consideration of the density deconvolution estimation problem and the density estimation problem in a

GARCH-type model, improving ([30], Proposition 5.1) and ([31], Theorem 2), respectively. A last part is devoted to the regression model with random design. The obtained result completes the one of Patil and Truong [23].

The organization of this note is as follows. In the next section we describe the considered wavelet setting. The hard thresholding estimator and its rate of convergence under the MISE over Besov balls are presented in Section 3. Applications of our general result are given in Section 4. The proofs are carried out in Section 5.

2. Wavelets and Besov Balls

In this section we introduce some notations corresponding to wavelets and Besov balls.

2.1. Wavelet Basis. We consider the wavelet basis on $[0,1]$ constructs from the Daubechies wavelets $db2N$ with $N \geq 1$ (see, e.g., [37]). A brief description of this basis is given below. Let ϕ and ψ be the initial wavelet functions of the family $db2N$. These functions have the particularity to be compactly supported and to belong to the class $\mathscr{C}^a$ for $N > 5a$. For any $j \geq 0$, we set $\Lambda_j = \{0, \ldots, 2^j - 1\}$ and, for $k \in \Lambda_j$,

$$\phi_{j,k}(x) = 2^{j/2}\phi\left(2^j x - k\right), \qquad \psi_{j,k}(x) = 2^{j/2}\psi\left(2^j x - k\right). \tag{1}$$

With appropriated treatments at the boundaries, there exists an integer τ such that, for any integer $\ell \geq \tau$, $\mathscr{B} = \{\phi_{\ell,k}, k \in \Lambda_\ell; \psi_{j,k}; j \in \mathbb{N} - \{0, \ldots, \ell-1\}, k \in \Lambda_j\}$ is an orthonormal basis of $\mathbb{L}^2([0,1])$, where

$$\mathbb{L}^2([0,1]) = \left\{ f : [0,1] \longrightarrow \mathbb{R}; \|f\|_2 = \left(\int_0^1 |f(x)|^2 dx\right)^{1/2} < \infty \right\}. \tag{2}$$

For any integer $\ell \geq \tau$ and $f \in \mathbb{L}^2([0,1])$, we have the following wavelet expansion:

$$f(x) = \sum_{k\in\Lambda_\ell} c_{\ell,k}\phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty}\sum_{k\in\Lambda_j} d_{j,k}\psi_{j,k}(x), \quad x \in [0,1], \tag{3}$$

where $c_{j,k}$ and $d_{j,k}$ denote the wavelet coefficients of f defined by

$$c_{j,k} = \int_0^1 f(x)\phi_{j,k}(x)\,dx, \qquad d_{j,k} = \int_0^1 f(x)\psi_{j,k}(x)\,dx. \tag{4}$$

Technical details can be found in, for example, Cohen et al. [38] and Mallat [39].

In the main result of this paper, we will investigate the MISE rate of the proposed estimator by assuming that the unknown function of interest f belongs to a wide class of functions: the Besov class. Its definition in terms of wavelet coefficients is presented in the following.

2.2. Besov Balls. We say that $f \in B^s_{p,r}(M)$ with $s > 0$, $p, r \geq 1$ and $M > 0$ if and only if there exists a constant $C > 0$ such that the wavelet coefficients of f given by (4) satisfy

$$2^{\tau(1/2-1/p)}\left(\sum_{k\in\Lambda_\tau}\left|c_{\tau,k}\right|^p\right)^{1/p} + \left(\sum_{j=\tau}^{\infty}\left(2^{j(s+1/2-1/p)}\left(\sum_{k\in\Lambda_j}\left|d_{j,k}\right|^p\right)^{1/p}\right)^r\right)^{1/r} \leq C, \tag{5}$$

with the usual modifications if $p = \infty$ or $r = \infty$. Note that, for particular choices of s, p, and r, $B^s_{p,r}(M)$ contains the classical Hölder and Sobolev balls (see, e.g., [40] and [19]).

Remark 1. We have chosen a wavelet basis on $[0,1]$ to fix the notations; wavelet basis on another interval can be considered in the rest of the study without affecting the results.

3. Statistical Framework, Estimator and Result

3.1. Statistical Framework. As mentioned in Section 1, a nonparametric estimation setting as general as possible is adopted: we aim to estimate an unknown function $f \in \mathbb{L}^2([0,1])$ via n random variables (or vectors) $V_1, \ldots, V_n$ from a strictly stationary stochastic process $(V_t)_{t\in\mathbb{Z}}$ defined on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$. We suppose that $(V_t)_{t\in\mathbb{Z}}$ has a α-mixing dependence structure with exponential decay rate; that is, there exist two constants $\gamma > 0$ and $\theta > 0$ such that

$$\sup_{m\geq 1}\left(e^{\theta m}\alpha_m\right) \leq \gamma, \tag{6}$$

where $\alpha_m = \sup_{(A,B)\in\mathscr{F}^V_{-\infty,0}\times\mathscr{F}^V_{m,\infty}}|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|$, $\mathscr{F}^V_{-\infty,0}$ is the σ-algebra generated by the random variables (or vectors) $\ldots, V_{-1}, V_0$ and $\mathscr{F}^V_{m,\infty}$ is the σ-algebra generated by the random variables (or vectors) $V_m, V_{m+1}, \ldots$.

The α-mixing dependence is reasonably weak; it is satisfied by a wide variety of models including Markov chains, GARCH-type models, and discretely observed diffusions (see, for instance, [1–3, 41]).

The considered estimator for f is presented below.

3.2. Estimator. We define the hard thresholding wavelet estimator $\widehat{f}$ by

$$\widehat{f}(x) = \sum_{k\in\Lambda_{j_0}} \widehat{c}_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j} \widehat{d}_{j,k}\mathbf{1}_{\{|\widehat{d}_{j,k}|\geq\kappa\lambda_j\}}\psi_{j,k}(x), \tag{7}$$

where

$$\widehat{c}_{j,k} = \frac{1}{n}\sum_{i=1}^{n} q\left(\phi_{j,k}, V_i\right), \qquad \widehat{d}_{j,k} = \frac{1}{n}\sum_{i=1}^{n} q\left(\psi_{j,k}, V_i\right), \tag{8}$$

$\mathbf{1}$ is the indicator function, $\kappa > 0$ is a large enough constant, j_0 is the integer satisfying

$$2^{j_0} = [\tau \ln n], \tag{9}$$

where $[a]$ denotes the integer part of a and j_1 is the integer satisfying

$$2^{j_1} = \left[\left(\frac{n}{(\ln n)^3} \right)^{1/(2\rho+1)} \right], \tag{10}$$

$$\lambda_j = 2^{\rho j} \sqrt{\frac{\ln n}{n}}. \tag{11}$$

Here it is supposed that there exists a function $q : \mathbb{L}^2([0,1]) \times V_1(\Omega) \to \mathbb{C}$ such that

(H1) for $\gamma \in \{\phi, \psi\}$, any integer $j \geq j_0$ and $k \in \Lambda_j$,

$$\mathbb{E}\left(q\left(\gamma_{j,k}, V_1\right)\right) = \int_0^1 f(x)\, \gamma_{j,k}(x)\, dx, \tag{12}$$

where $\mathbb{E}$ denotes the expectation,

(H2) there exist two constants, $C > 0$ and $\rho \geq 0$, satisfying, for $\gamma \in \{\phi, \psi\}$, for any integer $j \geq j_0$ and $k \in \Lambda_j$, (i) $\sup_{x \in V_1(\Omega)} |q(\gamma_{j,k}, x)| \leq C 2^{\rho j} 2^{j/2}$, (ii) $\mathbb{E}(|q(\gamma_{j,k}, V_1)|^2) \leq C 2^{2\rho j}$, (iii) for any $m \in \{1, \ldots, n-1\} \geq 1$,

$$\left|\mathbb{C}_{ov}\left(q\left(\gamma_{j,k}, V_{m+1}\right), q\left(\gamma_{j,k}, V_1\right)\right)\right| \leq C 2^{2\rho j} 2^{-j}, \tag{13}$$

where $\mathbb{C}_{ov}$ denotes the covariance; that is, $\mathbb{C}_{ov}(X, Y) = \mathbb{E}(X\overline{Y}) - \mathbb{E}(X)\mathbb{E}(\overline{Y})$, $\overline{Y}$ denotes the complex conjugate of Y.

For well-known nonparametric models in the i.i.d. setting, hard thresholding wavelet estimators and important results can be found in, for example, Donoho and Johnstone [14, 15], Donoho et al. [16, 17], Delyon and Juditsky [35], Kerkyacharian and Picard [36], and Fan and Koo [42]. In the α-mixing context, $\widehat{f}$ defined by (7) is a general and improved version of the estimator considered in Chesneau [30, 31]. The main differences are the presence of the tuning parameter ρ and the global definition of the function q offering numerous possibilities of applications. Three of them are explored in Section 4.

Comments on the Assumptions. The assumption (H1) ensures that (8) are unbiased estimators for $c_{j,k}$ and $d_{j,k}$ given by (4), whereas (H2) is related to their good performance. See Proposition 10. These assumptions are not too restrictive. For instance, if we consider the standard density estimation problem where $(V_t)_{t \in \mathbb{Z}}$ are i.i.d. random variables with bounded density f, the function $q(\gamma, x) = \gamma(x)$ satisfies (H1) and (H2) with $\rho = 0$ (note that, thanks to the independence of $(V_t)_{t \in \mathbb{Z}}$, the covariance term in (H2)-(iii) is zero). The technical details are given in Donoho et al. [17].

Lemma 2 describes a simple situation in which assumption (H2)-(iii) is satisfied.

Lemma 2. *We make the following assumptions.*

(F1) *Let u be the density of V_1 and let $u_{(V_1, V_{m+1})}$ be the density of (V_1, V_{m+1}) for any $m \in \mathbb{Z}$. We suppose that there exists a constant $C > 0$ such that*

$$\sup_{m \in \{1, \ldots, n-1\} \mathbb{Z}} \sup_{(x,y) \in V_1(\Omega) \times V_{m+1}(\Omega)} \left| u_{(V_1, V_{m+1})}(x, y) - u(x)\, u(y) \right| \leq C. \tag{14}$$

(F2) *There exist two constants, $C > 0$ and $\rho \geq 0$, satisfying, for $\gamma \in \{\phi, \psi\}$, for any integer $j \geq j_0$ and $k \in \Lambda_j$,*

$$\int_{V_1(\Omega)} \left| q\left(\gamma_{j,k}, x\right) \right| dx \leq C 2^{\rho j} 2^{-j/2}. \tag{15}$$

Then, under (F1) and (F2), (H2)-(iii) is satisfied.

3.3. Result. Theorem 3 determines the rate of convergence attained by $\widehat{f}$ under the MISE over Besov balls.

Theorem 3. *We consider the general statistical setting described in Section 3.1. Let $\widehat{f}$ be (7) under (H1) and (H2). Suppose that $f \in B^s_{p,r}(M)$ with $r \geq 1$, $\{p \geq 2$ and $s \in (0, N)\}$, or $\{p \in [1, 2)$ and $s \in ((2\rho + 1)/p, N)\}$. Then there exists a constant $C > 0$ such that*

$$\mathbb{E}\left(\left\| \widehat{f} - f \right\|_2^2 \right) \leq C \left(\frac{\ln n}{n} \right)^{2s/(2s+2\rho+1)}. \tag{16}$$

The rate of convergence "$((\ln n)/n)^{2s/(2s+2\rho+1)}$" is often the near optimal one in the minimax sense for numerous statistical problems in a i.i.d. setting (see, e.g., [19, 43]). Moreover, note that Theorem 3 is flexible; the assumptions on $(V_t)_{t \in \mathbb{Z}}$, related to the definition of q in (H1) and (H2), are mild. In the next section, this flexibility is illustrated for three sophisticated nonparametric estimation problems: the density deconvolution estimation problem, the density estimation problem in a GARCH-type model, and the regression function estimation in the regression model with random design.

4. Applications

4.1. Density Deconvolution. Let $(V_t)_{t \in \mathbb{Z}}$ be a strictly stationary stochastic process such that

$$V_t = X_t + \epsilon_t, \quad t \in \mathbb{Z}, \tag{17}$$

where $(X_t)_{t \in \mathbb{Z}}$ is a strictly stationary stochastic process with unknown density f and $(\epsilon_t)_{t \in \mathbb{Z}}$ is a strictly stationary stochastic process with known density g. It is supposed that ϵ_t and X_t are independent for any $t \in \mathbb{Z}$ and $(V_t)_{t \in \mathbb{Z}}$ is a α-mixing process with exponential decay rate (see Section 3.1 for a precise definition). Our aim is to estimate f via $V_1, \ldots, V_n$ from $(V_t)_{t \in \mathbb{Z}}$. Some related works are Masry [44], Kulik [45], Comte et al. [46], and Van Zanten and Zareba [47].

We formulate the following assumptions.

(G1) The support of f is $[0,1]$.

(G2) There exists a constant $C>0$ such that

$$\sup_{x\in\mathbb{R}} f(x) \leq C < \infty. \tag{18}$$

(G3) Let u be the density of V_1. We suppose that there exists a constant $C>0$ such that

$$\sup_{x\in\mathbb{R}} u(x) \leq C. \tag{19}$$

(G4) For any $m \in \mathbb{Z}$, let $u_{(V_1,V_{m+1})}$ be the density of (V_1, V_{m+1}). We suppose that there exists a constant $C>0$ such that

$$\sup_{m\in\mathbb{Z}} \sup_{(x,y)\in\mathbb{R}^2} u_{(V_1,V_{m+1})}(x,y) \leq C. \tag{20}$$

(G5) For any integrable function γ, we define its Fourier transform by

$$\mathcal{F}(\gamma)(x) = \int_{-\infty}^{\infty} \gamma(y) e^{-ixy} dy, \quad x \in \mathbb{R}. \tag{21}$$

We suppose that there exist three known constants $C>0$, $c>0$, and $\delta>1$ such that, for any $x\in\mathbb{R}$,

(i) the Fourier transform of g satisfies

$$\left|\mathcal{F}(g)(x)\right| \geq \frac{c}{\left(1+x^2\right)^{\delta/2}}, \tag{22}$$

(ii) for any $\ell \in \{0,1,2\}$, the ℓth derivative of the Fourier transform of g satisfies

$$\left|\left(\mathcal{F}(g)(x)\right)^{(\ell)}\right| \leq \frac{C}{(1+|x|)^{\delta+\ell}}. \tag{23}$$

We are now in the position to present the result.

Theorem 4. *We consider the model* (17). *Suppose that (G1)–(G5) are satisfied. Let $\widehat{f}$ be defined as in* (7) *with*

$$q(\gamma,x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma)(y)}}{\mathcal{F}(g)(y)} e^{-iyx} dy, \tag{24}$$

where $\overline{\mathcal{F}(\gamma)}(y)$ denotes the complex conjugate of $\mathcal{F}(\gamma)(y)$ and $\rho=\delta$ (appearing in(G5)).

Suppose that $f \in B^s_{p,r}(M)$ with $r\geq 1$, $\{p\geq 2$, and $s \in (0,N)\}$ or $\{p\in[1,2)$ and $s\in((2\delta+1)/p, N)\}$. Then there exists a constant $C>0$ such that

$$\mathbb{E}\left(\left\|\widehat{f}-f\right\|_2^2\right) \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\delta+1)}. \tag{25}$$

Theorem 4 improves ([30], Proposition 5.1) in terms of rate of convergence; we gain a logarithmic term.

Moreover, it is established that, in the i.i.d. setting, "$((\ln n)/n)^{2s/(2s+2\delta+1)}$" is

(i) exactly the rate of convergence attained by the hard thresholding wavelet estimator,

(ii) the near optimal rate of convergence in the minimax sense.

The details can be found in Fan and Koo [42]. Thus, Theorem 4 can be viewed as an extension of this existing result to the weak dependent case.

4.2. GARCH-Type Model. We consider the strictly stationary stochastic process $(V_t)_{t\in\mathbb{Z}}$ where, for any $t\in\mathbb{Z}$,

$$V_t = \sigma_t^2 Z_t, \tag{26}$$

$(\sigma_t^2)_{t\in\mathbb{Z}}$ is a strictly stationary stochastic process with unknown density f, and $(Z_t)_{t\in\mathbb{Z}}$ is a strictly stationary stochastic process with known density g. It is supposed that σ_t^2 and Z_t are independent for any $t\in\mathbb{Z}$ and $(V_t)_{t\in\mathbb{Z}}$ is a α-mixing process with exponential decay rate (see Section 3.1 for a precise definition). Our aim is to estimate f via $V_1,\ldots,V_n$ from $(V_t)_{t\in\mathbb{Z}}$. Some related works are Comte et al. [46] and Chesneau [31].

We formulate the following assumptions.

(J1) There exists a positive integer δ such that

$$g(x) = \frac{1}{(\delta-1)!}(-\ln x)^{\delta-1}, \quad x\in[0,1]. \tag{27}$$

Let us remark that g is the density of $\prod_{i=1}^{\delta} U_i$, where $U_1,\ldots,U_\delta$ are δ i.i.d. random variables having the common distribution $\mathcal{U}([0,1])$.

(J2) The support of f is $[0,1]$ and $f\in\mathbb{L}^2([0,1])$.

(J3) Let u be the density of V_1. We suppose that there exists a constant $C>0$ such that

$$\sup_{x\in\mathbb{R}} u(x) \leq C. \tag{28}$$

(J4) For any $m\in\mathbb{Z}$, let $u_{(V_1,V_{m+1})}$ be the density of (V_1,V_{m+1}). We suppose that there exists a constant $C>0$ such that

$$\sup_{m\in\mathbb{Z}} \sup_{(x,y)\in\mathbb{R}^2} u_{(V_1,V_{m+1})}(x,y) \leq C. \tag{29}$$

We are now in the position to present the result.

Theorem 5. *We consider model* (26). *Suppose that (J1)–(J4) are satisfied. Let $\widehat{f}$ be defined as in* (7) *with*

$$q(\gamma,x) = T_\delta(\gamma)(x), \tag{30}$$

where, for any positive integer ℓ, $T(\gamma)(x) = (x\gamma(x))'$ and $T_\ell(\gamma)(x) = T(T_{\ell-1}(\gamma))(x)$ and $\rho=\delta$ (appearing in (J1)).

Suppose that $f\in B^s_{p,r}(M)$ with $r\geq 1$, $\{p\geq 2$ and $s\in(0,N)\}$, or $\{p\in[1,2)$ and $s\in((2\delta+1)/p,N)\}$. Then there exists a constant $C>0$ such that

$$\mathbb{E}\left(\left\|\widehat{f}-f\right\|_2^2\right) \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\delta+1)}. \tag{31}$$

Theorem 5 significantly improves ([31], Theorem 2) in terms of rate of convergence; we gain an exponent $1/2$.

4.3. Nonparametric Regression Model. We consider the strictly stationary stochastic process $(V_t)_{t\in\mathbb{Z}}$ where, for any $t \in \mathbb{Z}$, $V_t = (Y_t, X_t)$,

$$Y_t = f(X_t) + \xi_t, \tag{32}$$

$(X_t)_{t\in\mathbb{Z}}$ is a strictly stationary stochastic process with unknown density g, $(\xi_t)_{t\in\mathbb{Z}}$ is a strictly stationary centered stochastic process, and f is the unknown regression function. It is supposed that X_t and ξ_t are independent for any $t \in \mathbb{Z}$ and $(V_t)_{t\in\mathbb{Z}}$ is a α-mixing process with exponential decay rate (see Section 3.1 for a precise definition). Our aim is to estimate f via $V_1, \ldots, V_n$ from $(V_t)_{t\in\mathbb{Z}}$. Applications of this problem can be found in Härdle [48]. Wavelet methods can be found in Patil and Truong [23], Doosti et al. [24], Doosti et al. [26], and Doosti and Niroumand [25].

We formulate the following assumptions.

(K1) The support of f and g is $[0,1]$ and f and $g \in \mathbb{L}^2([0,1])$.

(K2) $\xi_1(\Omega)$ is bounded.

(K3) There exists a constant $C > 0$ such that

$$\sup_{x\in[0,1]} |f(x)| \le C. \tag{33}$$

(K4) There exist two constants $c_* > 0$ and $C > 0$ such that

$$c_* \le \inf_{x\in[0,1]} g(x), \qquad \sup_{x\in[0,1]} g(x) \le C. \tag{34}$$

(K5) Let u be the density of V_1. We suppose that there exists a constant $C > 0$ such that

$$\sup_{x\in\mathbb{R}\times[0,1]} u(x) \le C. \tag{35}$$

(K6) For any $m \in \mathbb{Z}$, let $u_{(V_1,V_{m+1})}$ be the density of (V_1, V_{m+1}). We suppose that there exists a constant $C > 0$ such that

$$\sup_{m\in\mathbb{Z}} \sup_{(x,y)\in(\mathbb{R}\times[0,1])\times(\mathbb{R}\times[0,1])} u_{(V_1,V_{m+1})}(x,y) \le C. \tag{36}$$

We are now in the position to present the result.

Theorem 6. *We consider the model* (32). *Suppose that (K1)–(K6) are satisfied. Let $\hat{f}$ be the truncated ratio estimator. Consider*

$$\hat{f}(x) = \frac{\hat{v}(x)}{\hat{g}(x)} \mathbf{1}_{\{|\hat{g}(x)|\ge c_*/2\}}, \tag{37}$$

where

(i) *$\hat{v}$ is defined as in* (7) *with*

$$q(\gamma, (x, x_*)) = x\gamma(x_*) \tag{38}$$

and $\rho = 0$,

(ii) *$\hat{g}$ is defined as in* (7) *with X_t instead of V_t,*

$$q(\gamma, x) = \gamma(x) \tag{39}$$

and $\rho = 0$,

(iii) *c_* is the constant defined in (K4).*

Suppose that $fg \in B^s_{p,r}(M)$ and $g \in B^s_{p,r}(M)$ with $r \ge 1$, $\{p \ge 2$ and $s \in (0,N)\}$ or $\{p \in [1,2)$, and $s \in (1/p, N)\}$. Then there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\left\|\hat{f} - f\right\|_2^2\right) \le C\left(\frac{\ln n}{n}\right)^{2s/(2s+1)}. \tag{40}$$

The estimator (37) is derived by combining the procedure of Patil and Truong [23] with the truncated approach of Vasiliev [49].

Theorem 6 completes Patil and Truong [23] in terms of rates of convergence under the MISE over Besov balls.

Remark 7. The assumption(K2) can be relaxed with another strategy to the one developed in Theorem 6. Some technical elements are given in Chesneau [32].

Conclusion. Considering the weak dependent case on the observations, we prove a general result on the rate of convergence attains by a hard wavelet thresholding estimator under the MISE over Besov balls. This result is flexible; it can be applied for a wide class of statistical models. Moreover, the obtained rate of convergence is sharp; it can correspond to the near optimal one in the minimax sense for the standard i.i.d. case. Some recent results on sophisticated statistical problems are improved. Thanks to its flexibility, the perspectives of applications of our theoretical result in other contexts are numerous.

5. Proofs

In this section, C denotes any constant that does not depend on j, k, and n. Its value may change from one term to another and may depend on ϕ or ψ.

5.1. Key Lemmas. Let us present two lemmas which will be used in the proofs.

Lemma 8 shows a sharp covariance inequality under the α-mixing condition.

Lemma 8 (see [50]). *Let $(W_t)_{t\in\mathbb{Z}}$ be a strictly stationary α-mixing process with mixing coefficient α_m, $m \ge 0$, and let h and k be two measurable functions. Let $p > 0$ and $q > 0$ satisfying $1/p + 1/q < 1$ such that $\mathbb{E}(|h(W_1)|^p)$ and $\mathbb{E}(|k(W_1)|^q)$ exist. Then there exists a constant $C > 0$ such that*

$$\begin{aligned} &\left|\mathbb{C}_{ov}(h(W_1), k(W_{m+1}))\right| \\ &\quad \le C\alpha_m^{1-1/p-1/q}\left(\mathbb{E}\left(|h(W_1)|^p\right)\right)^{1/p}\left(\mathbb{E}(|k(W_1)|^q)\right)^{1/q}. \end{aligned} \tag{41}$$

Lemma 9 below presents a concentration inequality for α-mixing processes.

Lemma 9 (see [13]). *Let $(W_t)_{t\in\mathbb{Z}}$ be a strictly stationary process with the mth strongly mixing coefficient α_m, $m \geq 0$, let n be a positive integer, let $h : \mathbb{R} \to \mathbb{C}$ be a measurable function, and, for any $t \in \mathbb{Z}$, $U_t = h(W_t)$. We assume that $\mathbb{E}(U_1) = 0$ and there exists a constant $M > 0$ satisfying $|U_1| \leq M$. Then, for any $m \in \{1, \ldots, [n/2]\}$ and $\lambda > 0$, we have*

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n} U_i\right| \geq \lambda\right) \leq 4\exp\left(-\frac{\lambda^2 n}{16\left(D_m/m + \lambda M m/3\right)}\right) + 32\frac{M}{\lambda} n\alpha_m, \tag{42}$$

where

$$D_m = \max_{l\in\{1,\ldots,2m\}} \mathbb{V}\left(\sum_{i=1}^{l} U_i\right). \tag{43}$$

5.2. Intermediary Results

Proof of Lemma 2. Using a standard expression of the covariance, and (F1) as well as(F2), we obtain

$$\begin{aligned}
&\left|\mathbb{C}_{ov}\left(q\left(\gamma_{j,k}, V_{m+1}\right), q\left(\gamma_{j,k}, V_1\right)\right)\right| \\
&= \left|\int_{V_1(\Omega)} \int_{V_1(\Omega)} q\left(\gamma_{j,k}, x\right) \overline{q\left(\gamma_{j,k}, y\right)} \right. \\
&\qquad \left. \times \left(u_{(V_1,V_{m+1})}(x,y) - u(x)u(y)\right) dx dy \right| \\
&\leq \int_{V_1(\Omega)} \int_{V_1(\Omega)} \left|q\left(\gamma_{j,k}, x\right)\right| \left|q\left(\gamma_{j,k}, y\right)\right| \\
&\qquad \times \left|u_{(V_1,V_{m+1})}(x,y) - u(x)u(y)\right| dx\, dy \\
&\leq C\left(\int_{V_1(\Omega)} \left|q\left(\gamma_{j,k}, x\right)\right| dx\right)^2 \leq C 2^{2\rho j} 2^{-j}.
\end{aligned} \tag{44}$$

This ends the proof of Lemma 2. □

Proposition 10 proves probability and moments inequalities satisfied by the estimators (8).

Proposition 10. *Let $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ be defined as in* (8) *under (H1) and (H2), let j_0 be* (9) *and let j_1 be* (10).

(a) *There exists a constant $C > 0$ such that, for any $j \in \{j_0, \ldots, j_1\}$ and $k \in \Lambda_j$,*

$$\mathbb{E}\left(\left|\widehat{c}_{j,k} - c_{j,k}\right|^2\right) \leq C 2^{2\rho j}\frac{1}{n}, \tag{45}$$

$$\mathbb{E}\left(\left|\widehat{d}_{j,k} - d_{j,k}\right|^2\right) \leq C 2^{2\rho j}\frac{1}{n}. \tag{46}$$

(b) *There exists a constant $C > 0$ such that, for any $j \in \{j_0, \ldots, j_1\}$ and $k \in \Lambda_j$,*

$$\mathbb{E}\left(\left|\widehat{d}_{j,k} - d_{j,k}\right|^4\right) \leq C 2^{4\rho j}. \tag{47}$$

(c) *Let λ_j be defined as in* (11). *There exists a constant $C > 0$ such that, for any κ large enough, $j \in \{j_0, \ldots, j_1\}$ and $k \in \Lambda_j$, we have*

$$\mathbb{P}\left(\left|\widehat{d}_{j,k} - d_{j,k}\right| \geq \frac{\kappa\lambda_j}{2}\right) \leq C\frac{1}{n^4}. \tag{48}$$

Proof of Proposition 10. (a) Using (H1) and the stationarity of $(V_t)_{t\in\mathbb{Z}}$, we obtain

$$\begin{aligned}
&\mathbb{E}\left(\left|\widehat{c}_{j,k} - c_{j,k}\right|^2\right) = \mathbb{V}\left(\widehat{c}_{j,k}\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{V}\left(q\left(\phi_{j,k}, V_i\right)\right) + \frac{2}{n^2} \\
&\quad \times \sum_{v=2}^{n}\sum_{\ell=1}^{v-1} \operatorname{Re}\left(\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_v\right), q\left(\phi_{j,k}, V_\ell\right)\right)\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{V}\left(q\left(\phi_{j,k}, V_i\right)\right) + \frac{2}{n^2} \\
&\quad \times \sum_{m=1}^{n-1} (n-m) \operatorname{Re}\left(\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_{m+1}\right), q\left(\phi_{j,k}, V_1\right)\right)\right) \\
&\leq \frac{1}{n}\left(\mathbb{E}\left(\left|q\left(\phi_{j,k}, V_1\right)\right|^2\right)\right. \\
&\qquad \left. + 2\sum_{m=1}^{n-1} \left|\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_{m+1}\right), q\left(\phi_{j,k}, V_1\right)\right)\right|\right).
\end{aligned} \tag{49}$$

By (H2)-(ii) we get

$$\mathbb{E}\left(\left|q(\phi_{j,k}, V_1)\right|^2\right) \leq C 2^{2\rho j}. \tag{50}$$

For the covariance term, note that

$$\sum_{m=1}^{n-1} \left|\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_{m+1}\right), q\left(\phi_{j,k}, V_1\right)\right)\right| = A + B, \tag{51}$$

where

$$\begin{aligned}
A &= \sum_{m=1}^{[\ln n/\theta]-1} \left|\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_{m+1}\right), q\left(\phi_{j,k}, V_1\right)\right)\right|, \\
B &= \sum_{m=[(\ln n)/\theta]}^{n-1} \left|\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_{m+1}\right), q\left(\phi_{j,k}, V_1\right)\right)\right|.
\end{aligned} \tag{52}$$

It follows from (H2)-(iii) and $2^{-j} \leq 2^{-j_0} < 2(\ln n)^{-1}$ that

$$A \leq C 2^{2\rho j} 2^{-j}\left[\frac{\ln n}{\theta}\right] \leq C 2^{2\rho j}. \tag{53}$$

The Davydov inequality described in Lemma 8 with $p = q = 4$, (H2)-(i)-(ii), and $2^j \le 2^{j_1} \le n$ give

$$\begin{aligned} B &\le C\sqrt{\mathbb{E}\left(\left|q(\phi_{j,k}, V_1)\right|^4\right)} \sum_{m=[(\ln n)/\theta]}^{n-1} \sqrt{\alpha_m} \\ &\le C2^{\rho j}2^{j/2}\sqrt{\mathbb{E}\left(\left|q(\phi_{j,k}, V_1)\right|^2\right)} \sum_{m=[(\ln n)/\theta]}^{\infty} e^{-\theta m/2} \\ &= C2^{2\rho j}\sqrt{n}e^{-(\ln n)/2} \le C2^{2\rho j}. \end{aligned} \tag{54}$$

Thus

$$\sum_{m=1}^{n-1}\left|\mathbb{C}_{ov}\left(q\left(\phi_{j,k}, V_{m+1}\right), q\left(\phi_{j,k}, V_1\right)\right)\right| \le C2^{2\rho j}. \tag{55}$$

Putting (49), (50), and (55) together, the first point in (a) is proved. The proof of the second point is identical with ψ instead of ϕ.

(b) Thanks to (H2)-(i), we have $|\hat{d}_{j,k}| \le \sup_{x\in V_1(\Omega)}|q(\psi_{j,k}, x)| \le C2^{\rho j}2^{j/2}$. It follows from the triangular inequality and $|d_{j,k}| \le \|f\|_2 \le C$ that

$$\left|\hat{d}_{j,k} - d_{j,k}\right| \le \left|\hat{d}_{j,k}\right| + \left|d_{j,k}\right| \le C2^{\rho j}2^{j/2}. \tag{56}$$

This inequality and the second result of (a) yield

$$\mathbb{E}\left(\left|\hat{d}_{j,k} - d_{j,k}\right|^4\right) \le C2^{2\rho j}2^j\mathbb{E}\left(\left|\hat{d}_{j,k} - d_{j,k}\right|^2\right) \le C2^{4\rho j}2^j\frac{1}{n}. \tag{57}$$

Using $2^j \le 2^{j_1} \le n$, the proof of (b) is completed.

(c) We will use the Liebscher inequality described in Lemma 9. Let us set

$$U_i = q\left(\psi_{j,k}, V_i\right) - \mathbb{E}\left(q\left(\psi_{j,k}, V_1\right)\right). \tag{58}$$

We have $\mathbb{E}(U_1) = 0$ and, by(H2)-(i) and $2^j \le 2^{j_1} \le n/(\ln n)^3$,

$$|U_i| \le 2\sup_{x\in V_1(\Omega)}\left|q\left(\psi_{j,k}, x\right)\right| \le C2^{\rho j}2^{j/2} \le C2^{\rho j}\sqrt{\frac{n}{(\ln n)^3}}, \tag{59}$$

(so $M = C2^{\rho j}\sqrt{n/(\ln n)^3}$).

Proceeding as for the proofs of the bounds in (a), for any integer $l \le C\ln n$, since $2^{-j} \le 2^{-j_0} \le 2(\ln n)^{-1}$, we show that

$$\begin{aligned} &\mathbb{V}\left(\sum_{i=1}^{l}U_i\right) \\ &= \mathbb{V}\left(\sum_{i=1}^{l}q\left(\psi_{j,k}, V_i\right)\right) \le C2^{2\rho j}\left(l + l^2 2^{-j}\right) \le C2^{2\rho j}l. \end{aligned} \tag{60}$$

Therefore

$$D_m = \max_{l\in\{1,\ldots,2m\}}\mathbb{V}\left(\sum_{i=1}^{l}U_i\right) \le C2^{2\rho j}m. \tag{61}$$

Owing to Lemma 9 applied with $U_1, \ldots, U_n$, $\lambda = \kappa\lambda_j/2$, $m = [\sqrt{\kappa}\ln n]$, $M = C2^{\rho j}\sqrt{n/(\ln n)^3}$, and the bound (61), we obtain

$$\begin{aligned} &\mathbb{P}\left(\left|\hat{d}_{j,k} - d_{j,k}\right| \ge \frac{\kappa\lambda_j}{2}\right) \\ &\le C\left(\exp\left(-C\frac{\kappa^2\lambda_j^2 n}{D_m/m + \kappa\lambda_j mM}\right) + \frac{M}{\lambda_j}ne^{-\theta m}\right) \\ &\le C\Bigg(\exp\Bigg(-C\Bigg(\left(\kappa^2 2^{2\rho j}\ln n\right) \\ &\qquad \times\left(2^{2\rho j} + \kappa 2^{\rho j}\sqrt{\frac{(\ln n)}{n}} \times [\sqrt{\kappa}\ln n]\, 2^{\rho j}\sqrt{\frac{n}{(\ln n)^3}}\right)^{-1}\Bigg)\Bigg) \\ &\qquad + \sqrt{\frac{n/(\ln n)^3}{(\ln n)/n}}ne^{-\theta[\sqrt{\kappa}\ln n]}\Bigg) \\ &\le C\left(n^{-C\kappa^2/(1+\kappa^{3/2})} + n^{2-\theta\sqrt{\kappa}}\right). \end{aligned} \tag{62}$$

Taking κ large enough, the last term is bounded by C/n^4. This completes the proof of(c).

This completes the proof of Proposition 10. □

Proof of Theorem 3. Theorem 3 can be proved by combining arguments of ([36], Theorem 5.1) and ([51], Theorem 4.2). It is close to ([30], Proof of Theorem 2) by taking $\theta \to \infty$. The interested reader can find the details below.

We consider the following wavelet decomposition for f:

$$f(x) = \sum_{k\in\Lambda_{j_0}} c_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j} d_{j,k}\psi_{j,k}(x), \tag{63}$$

where $c_{j_0,k} = \int_0^1 f(x)\phi_{j_0,k}(x)dx$ and $d_{j,k} = \int_0^1 f(x)\psi_{j,k}(x)dx$.

Using the orthonormality of the wavelet basis $\mathscr{B}$, the MISE of $\hat{f}$ can be expressed as

$$\mathbb{E}\left(\left\|\hat{f} - f\right\|_2^2\right) = P + Q + R, \tag{64}$$

where

$$\begin{aligned} P &= \sum_{k\in\Lambda_{j_0}}\mathbb{E}\left(\left|\hat{c}_{j_0,k} - c_{j_0,k}\right|^2\right), \\ Q &= \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(\left|\hat{d}_{j,k}\mathbf{1}_{\{|\hat{d}_{j,k}|\ge\kappa\lambda_j\}} - d_{j,k}\right|^2\right), \\ R &= \sum_{j=j_1+1}^{\infty}\sum_{k\in\Lambda_j} d_{j,k}^2. \end{aligned} \tag{65}$$

Let us now investigate sharp upper bounds for P, R and Q successively.

Upper Bound for P. The point **(a)** of Proposition 10 and $2s/(2s+2\rho+1)<1$ yield

$$P \le C\frac{2^{j_0}}{n} \le C\frac{\ln n}{n} \le C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \tag{66}$$

Upper Bound for R.

(i) For $r \ge 1$ and $p \ge 2$, we have $f \in B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$. Using $2s/(2s+2\rho+1) < 2s/(2\rho+1)$, we obtain

$$\begin{aligned} R &\le C\sum_{j=j_1+1}^{\infty} 2^{-2js} \le C\left(\frac{(\ln n)^3}{n}\right)^{2s/(2\rho+1)} \\ &\le C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \end{aligned} \tag{67}$$

(ii) For $r \ge 1$ and $p \in [1,2)$, we have $f \in B^s_{p,r}(M) \subseteq B^{s+1/2-1/p}_{2,\infty}(M)$. The condition $s > (2\rho+1)/p$ implies that $(s+1/2-1/p)/(2\rho+1) > s/(2s+2\rho+1)$. Thus

$$\begin{aligned} R &\le C\sum_{j=j_1+1}^{\infty} 2^{-2j(s+1/2-1/p)} \\ &\le C\left(\frac{(\ln n)^3}{n}\right)^{2(s+1/2-1/p)/(2\rho+1)} \\ &\le C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \end{aligned} \tag{68}$$

Hence, for $r \ge 1$, $\{p \ge 2$ and $s > 0\}$ or $\{p \in [1,2)$, and $s > (2\rho+1)/p\}$, we have

$$R \le C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \tag{69}$$

Upper Bound for Q. Adopting the notation $\widehat{D}_{j,k} = \hat{d}_{j,k} - d_{j,k}$, Q can be written as

$$Q = \sum_{i=1}^{4} Q_i, \tag{70}$$

where

$$\begin{aligned} Q_1 &= \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(\left|\widehat{D}_{j,k}\right|^2 \mathbf{1}_{\{|\hat{d}_{j,k}|\ge\kappa\lambda_j,|d_{j,k}|<\kappa\lambda_j/2\}}\right), \\ Q_2 &= \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(\left|\widehat{D}_{j,k}\right|^2 \mathbf{1}_{\{|\hat{d}_{j,k}|\ge\kappa\lambda_j,|d_{j,k}|\ge\kappa\lambda_j/2\}}\right), \\ Q_3 &= \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(d^2_{j,k} \mathbf{1}_{\{|\hat{d}_{j,k}|<\kappa\lambda_j,|d_{j,k}|\ge2\kappa\lambda_j\}}\right), \\ Q_4 &= \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(d^2_{j,k} \mathbf{1}_{\{|\hat{d}_{j,k}|<\kappa\lambda_j,|d_{j,k}|<2\kappa\lambda_j\}}\right). \end{aligned} \tag{71}$$

Upper Bound for $Q_1 + Q_3$. Owing to the inequalities $\mathbf{1}_{\{|\hat{d}_{j,k}|<\kappa\lambda_j,|d_{j,k}|\ge2\kappa\lambda_j\}} \le \mathbf{1}_{\{|\widehat{D}_{j,k}|>\kappa\lambda_j/2\}}$, $\mathbf{1}_{\{|\hat{d}_{j,k}|\ge\kappa\lambda_j,|d_{j,k}|<\kappa\lambda_j/2\}} \le \mathbf{1}_{\{|\widehat{D}_{j,k}|>\kappa\lambda_j/2\}}$ and $\mathbf{1}_{\{|\hat{d}_{j,k}|<\kappa\lambda_j,|d_{j,k}|\ge2\kappa\lambda_j\}} \le \mathbf{1}_{\{|d_{j,k}|\le2|\widehat{D}_{j,k}|\}}$, the Cauchy-Schwarz inequality, and the points(b) and (c) of Proposition 10, we have

$$\begin{aligned} Q_1 + Q_3 &\le C\sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(\left|\widehat{D}_{j,k}\right|^2\mathbf{1}_{\{|\widehat{D}_{j,k}|>\kappa\lambda_j/2\}}\right) \\ &\le C\sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\left(\mathbb{E}\left(\left|\widehat{D}_{j,k}\right|^4\right)\right)^{1/2}\left(\mathbb{P}\left(\left|\widehat{D}_{j,k}\right| > \frac{\kappa\lambda_j}{2}\right)\right)^{1/2} \\ &\le C\frac{1}{n^2}\sum_{j=j_0}^{j_1}2^{j(1+2\rho)} \le C\frac{1}{n} \le C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \end{aligned} \tag{72}$$

Upper Bound for Q_2. It follows from the point(a) of Proposition 10 that

$$\begin{aligned} Q_2 &\le \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j}\mathbb{E}\left(\left|\widehat{D}_{j,k}\right|^2\right)\mathbf{1}_{\{|d_{j,k}|\ge\kappa\lambda_j/2\}} \\ &\le C\frac{1}{n}\sum_{j=j_0}^{j_1}2^{2\rho j}\sum_{k\in\Lambda_j}\mathbf{1}_{\{|d_{j,k}|>\kappa\lambda_j/2\}}. \end{aligned} \tag{73}$$

Let us now introduce the integer j_* defined by

$$2^{j_*} = \left[\left(\frac{n}{\ln n}\right)^{1/(2s+2\rho+1)}\right]. \tag{74}$$

Note that $j_* \in \{j_0,\dots,j_1\}$ for n large enough.

Then Q_2 can be bounded as

$$Q_2 \le Q_{2,1} + Q_{2,2}, \tag{75}$$

where

$$\begin{aligned} Q_{2,1} &= C\frac{1}{n}\sum_{j=j_0}^{j_*}2^{2\rho j}\sum_{k\in\Lambda_j}\mathbf{1}_{\{|d_{j,k}|>\kappa\lambda_j/2\}}, \\ Q_{2,2} &= C\frac{1}{n}\sum_{j=j_*+1}^{j_1}2^{2\rho j}\sum_{k\in\Lambda_j}\mathbf{1}_{\{|d_{j,k}|>\kappa\lambda_j/2\}}. \end{aligned} \tag{76}$$

On the one hand we have

$$Q_{2,1} \le C\frac{\ln n}{n}\sum_{j=j_0}^{j_*}2^{j(1+2\rho)} \le C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \tag{77}$$

On the other hand, we have the following.

(i) For $r \geq 1$ and $p \geq 2$, the Markov inequality and $f \in B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$ yield

$$Q_{2,2} \leq C\frac{\ln n}{n}\sum_{j=j_*+1}^{j_1} 2^{2\rho j}\frac{1}{\lambda_j^2}\sum_{k\in\Lambda_j} d_{j,k}^2 \leq C\sum_{j=j_*+1}^{\infty}\sum_{k\in\Lambda_j} d_{j,k}^2 \leq C\sum_{j=j_*+1}^{\infty} 2^{-2js} \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \quad (78)$$

(ii) For $r \geq 1$, $p \in [1,2)$ and $s > (2\rho+1)/p$, the Markov inequality, $f \in B^s_{p,r}(M)$, and $(2s+2\rho+1)(2-p)/2 + (s+1/2-1/p+\rho-2\rho/p)p = 2s$ imply that

$$\begin{aligned} Q_{2,2} &\leq C\frac{\ln n}{n}\sum_{j=j_*+1}^{j_1} 2^{2\rho j}\frac{1}{\lambda_j^p}\sum_{k\in\Lambda_j}\left|d_{j,k}\right|^p \\ &\leq C\left(\frac{\ln n}{n}\right)^{(2-p)/2}\sum_{j=j_*+1}^{\infty} 2^{j\rho(2-p)}2^{-j(s+1/2-1/p)p} \\ &\leq C\left(\frac{\ln n}{n}\right)^{(2-p)/2} 2^{-j_*(s+1/2-1/p+\rho-2\rho/p)p} \\ &\leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \end{aligned} \quad (79)$$

Therefore, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$, and $s > (2\rho+1)/p\}$, we have

$$Q_2 \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \quad (80)$$

Upper Bound for Q_4. We have

$$Q_4 \leq \sum_{j=j_0}^{j_1}\sum_{k\in\Lambda_j} d_{j,k}^2 \mathbf{1}_{\{|d_{j,k}|<2\kappa\lambda_j\}}. \quad (81)$$

Let j_* be the integer (74). Then Q_4 can be bound as

$$Q_4 \leq Q_{4,1} + Q_{4,2}, \quad (82)$$

where

$$\begin{aligned} Q_{4,1} &= \sum_{j=j_0}^{j_*}\sum_{k\in\Lambda_j} d_{j,k}^2 \mathbf{1}_{\{|d_{j,k}|<2\kappa\lambda_j\}}, \\ Q_{4,2} &= \sum_{j=j_*+1}^{j_1}\sum_{k\in\Lambda_j} d_{j,k}^2 \mathbf{1}_{\{|d_{j,k}|<2\kappa\lambda_j\}}. \end{aligned} \quad (83)$$

On the one hand, we have

$$Q_{4,1} \leq C\sum_{j=j_0}^{j_*} 2^j\lambda_j^2 = C\frac{\ln n}{n}\sum_{j=j_0}^{j_*} 2^{j(1+2\rho)} \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \quad (84)$$

On the other hand, we have the following.

(i) For $r \geq 1$ and $p \geq 2$, since $f \in B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$, we have

$$Q_{4,2} \leq \sum_{j=j_*+1}^{\infty}\sum_{k\in\Lambda_j} d_{j,k}^2 \leq C\sum_{j=j_*+1}^{\infty} 2^{-2js} \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \quad (85)$$

(ii) For $r \geq 1$, $p \in [1,2)$ and $s > (2\rho+1)/p$, owing to the Markov inequality, $f \in B^s_{p,r}(M)$ and $(2s+2\rho+1)(2-p)/2 + (s+1/2-1/p+\rho-2\rho/p)p = 2s$, we get

$$\begin{aligned} Q_{4,2} &\leq C\sum_{j=j_*+1}^{j_1}\lambda_j^{2-p}\sum_{k\in\Lambda_j}\left|d_{j,k}\right|^p \\ &= C\left(\frac{\ln n}{n}\right)^{(2-p)/2}\sum_{j=j_*+1}^{j_1} 2^{j\rho(2-p)}\sum_{k\in\Lambda_j}\left|d_{j,k}\right|^p \\ &\leq C\left(\frac{\ln n}{n}\right)^{(2-p)/2}\sum_{j=j_*+1}^{\infty} 2^{j\rho(2-p)}2^{-j(s+1/2-1/p)p} \\ &\leq C\left(\frac{\ln n}{n}\right)^{(2-p)/2} 2^{-j_*(s+1/2-1/p+\rho-2\rho/p)p} \\ &\leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \end{aligned} \quad (86)$$

So, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$, and $s > (2\rho+1)/p\}$, we have

$$Q_4 \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \quad (87)$$

Putting (70), (72), (80), and (87) together, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$ and $s > (2\rho+1)/p\}$, we obtain

$$Q \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2\rho+1)}. \quad (88)$$

Combining (64), (66), (69), and (88), we complete the proof of Theorem 3. □

Proof of Theorem 4. The proof of Theorem 4 is a direct application of Theorem 3: under (G1)–(G5), the function q defined by (24) satisfies (H1) see ([42], equation (2)) and (H2): (i) see ([42], Lemma 6), (ii) see, ([42], equation (11)) and (iii) see ([30], Proof of Proposition 6.1), with $\rho = \delta$. □

Proof of Theorem 5. The proof of Theorem 5 is a consequence of Theorem 3: under (J1)–(J4), the function q defined by (30) satisfies (H1) and (H2): (i)-(ii) see ([31], Proposition 1) and (iii) see ([52], equation (26)), with $\rho = \delta$. □

Proof of Theorem 6. Set $v(x) = f(x)g(x)$. Following the methodology of [49], we have

$$\hat{f}(x) - f(x) = S(x) - T(x), \quad (89)$$

where

$$\begin{aligned} S(x) &= \frac{1}{\widehat{g}(x)}\left(\widehat{v}(x)-v(x)\right.\\ &\quad\left.+f(x)\left(g(x)-\widehat{g}(x)\right)\right)\mathbf{1}_{\{|\widehat{g}(x)|\geq c_*/2\}} \\ T(x) &= f(x)\,1_{\{|\widehat{g}(x)|<c_*/2\}}. \end{aligned} \tag{90}$$

Using (K3) and the indicator function, we have

$$|S(x)| \leq C\left(|\widehat{v}(x)-v(x)|+\left|\widehat{g}(x)-g(x)\right|\right). \tag{91}$$

It follows from $\{|\widehat{g}(x)| < c_*/2\} \cap \{|g(x)| > c_*\} \subseteq \{|\widehat{g}(x) - g(x)| > c_*/2\}$, (K3), (K4), and the Markov inequality that

$$|T(x)| \leq C\mathbf{1}_{\{|\widehat{g}(x)-g(x)|>c_*/2\}} \leq C\left|\widehat{g}(x)-g(x)\right|. \tag{92}$$

The triangular inequality yields

$$\left|\widehat{f}(x)-f(x)\right| \leq C\left(|\widehat{v}(x)-v(x)|+\left|\widehat{g}(x)-g(x)\right|\right). \tag{93}$$

The elementary inequality $(a+b)^2 \leq 2(a^2+b^2)$ implies that

$$\mathbb{E}\left(\left\|\widehat{f}-f\right\|_2^2\right) \leq C\left(\mathbb{E}\left(\|\widehat{v}-v\|_2^2\right)+\mathbb{E}\left(\left\|\widehat{g}-g\right\|_2^2\right)\right). \tag{94}$$

We now bound this two MISEs via Theorem 3.

Upper Bound for the MISE of $\widehat{v}$. Under (K1)–(K6), the function q defined by (38) satisfies the following.

(H1) With v instead of f: since ξ_1 and X_1 are independent with $\mathbb{E}(\xi_1) = 0$,

$$\begin{aligned} &\mathbb{E}\left(q\left(\gamma_{j,k}, V_1\right)\right) \\ &\quad= \mathbb{E}\left(Y_1\gamma_{j,k}(X_1)\right) = \mathbb{E}\left(f(X_1)\gamma_{j,k}(X_1)\right) \\ &\quad= \int_0^1 f(x)\gamma_{j,k}(x)g(x)\,dx = \int_0^1 v(x)\gamma_{j,k}(x)\,dx, \end{aligned} \tag{95}$$

(H2): (i)-(ii)-(iii) with $\rho = 0$:

(i) since $Y_1(\Omega)$ is bounded thanks to (K2) and (K3), say $|Y_1| \leq M$ with $M > 0$, we have

$$\begin{aligned} &\sup_{(x,x_*)\in V_1(\Omega)}\left|q\left(\gamma_{j,k},(x,x_*)\right)\right| \\ &\quad= \sup_{(x,x_*)\in[-M,M]\times[0,1]}\left|x\gamma_{j,k}(x_*)\right| \\ &\quad\leq M\sup_{x_*\in[0,1]}\left|\gamma_{j,k}(x_*)\right| \leq C2^{j/2} \end{aligned} \tag{96}$$

(ii) using the boundedness of $Y_1(\Omega)$, then (K4), we have

$$\begin{aligned} \mathbb{E}\left(\left|q(\gamma_{j,k},V_1)\right|^2\right) &= \mathbb{E}\left(Y_1^2\left(\gamma_{j,k}(X_1)\right)^2\right) \\ &\leq C\mathbb{E}\left(\left(\gamma_{j,k}(X_1)\right)^2\right) \\ &= C\int_0^1\left(\gamma_{j,k}(x)\right)^2 g(x)\,dx \\ &\leq C\int_0^1\left(\gamma_{j,k}(x)\right)^2 dx \leq C \end{aligned} \tag{97}$$

(iii) using the boundedness of $Y_1(\Omega)$ and making the change of variables $y = 2^j x - k$, we obtain

$$\begin{aligned} &\int_{V_1(\Omega)}\left|q\left(\gamma_{j,k},x\right)\right|dx \\ &\quad= \left(\int_{-M}^{M}|x|\,dx\right)\left(\int_0^1\left|\gamma_{j,k}(x_*)\right|dx_*\right) \\ &\quad= M^2\int_0^1\left|\gamma_{j,k}(x)\right|dx \leq C2^{-j/2}. \end{aligned} \tag{98}$$

We conclude by applying Lemma 2 with $\rho = 0$; (K5) and (K6) imply (F1), and the previous inequality implies (F2).

Therefore, assuming that $v \in B^s_{p,r}(M)$ with $r \geq 1$, $\{p \geq 2$ and $s \in (0,N)\}$ or $\{p \in [1,2)$ and $s \in (1/p,N)\}$, Theorem 3 proves the existence of a constant $C > 0$ satisfying

$$\mathbb{E}\left(\|\widehat{v}-v\|_2^2\right) \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+1)}. \tag{99}$$

Upper Bound for the MISE of $\widehat{g}$. Under (K1)–(K6), proceeding as the previous point, we show that the function q defined by (39) satisfies (H1) with g instead of f and X_t instead of V_t, and(H2): (i)-(ii)-(iii) with $\rho = 0$.

Therefore, assuming that $g \in B^s_{p,r}(M)$ with $r \geq 1$, $\{p \geq 2$ and $s \in (0,N)\}$ or $\{p \in [1,2)$, and $s \in (1/p,N)\}$, Theorem 3 proves the existance of a constant $C > 0$ satisfying

$$\mathbb{E}\left(\left\|\widehat{g}-g\right\|_2^2\right) \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+1)}. \tag{100}$$

Combining (94), (99), and (100), we end the proof of Theorem 6. □

Conflict of Interests

The author declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The author is thankful to the reviewers for their comments which have helped in improving the presentation.

References

[1] P. Doukhan, *Mixing. Properties and Examples*, vol. 85 of *Lecture Notes in Statistics*, Springer, New York, NY, USA, 1994.

[2] M. Carrasco and X. Chen, "Mixing and moment properties of various GARCH and stochastic volatility models," *Econometric Theory*, vol. 18, no. 1, pp. 17–39, 2002.

[3] R. C. Bradley, *Introduction to Strong Mixing Conditions, Volume 1, 2, 3*, Kendrick Press, 2007.

[4] P. M. Robinson, "Nonparametric estimators for time series," *Journal of Time Series Analysis*, vol. 4, pp. 185–207, 1983.

[5] G. G. Roussas, "Nonparametric estimation in mixing sequences of random variables," *Journal of Statistical Planning and Inference*, vol. 18, no. 2, pp. 135–149, 1987.

[6] G. G. Roussas, "Nonparametric regression estimation under mixing conditions," *Stochastic Processes and their Applications*, vol. 36, no. 1, pp. 107–116, 1990.

[7] Y. K. Truong and C. J. Stone, "Nonparametric function estimation involving time series," *The Annals of Statistics*, vol. 20, pp. 77–97, 1992.

[8] L. H. Tran, "Nonparametric function estimation for time series by local average estimators," *The Annals of Statistics*, vol. 21, pp. 1040–1057, 1993.

[9] E. Masry, "Multivariate local polynomial regression for time series: uniform strong consistency and rates," *Journal of Time Series Analysis*, vol. 17, no. 6, pp. 571–599, 1996.

[10] E. Masry, "Multivariate regression estimation local polynomial fitting for time series," *Stochastic Processes and their Applications*, vol. 65, no. 1, pp. 81–101, 1996.

[11] E. Masry and J. Fan, "Local polynomial estimation of regression functions for mixing processes," *Scandinavian Journal of Statistics*, vol. 24, no. 2, pp. 165–179, 1997.

[12] D. Bosq, *Nonparametric Statistics for Stochastic Processes. Estimation and Prediction*, vol. 110 of *Lecture Notes in Statistics*, Springer, New York, NY, USA, 1998.

[13] E. Liebscher, "Estimation of the density and the regression function under mixing conditions," *Statistics and Decisions*, vol. 19, no. 1, pp. 9–26, 2001.

[14] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[15] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, pp. 1200–1224, 1995.

[16] D. Donoho, I. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: asymptopia?" *Journal of the Royal Statistical Society B*, vol. 57, pp. 301–369, 1995.

[17] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Density estimation by wavelet thresholding," *Annals of Statistics*, vol. 24, no. 2, pp. 508–539, 1996.

[18] A. Antoniadis, "Wavelets in statistics: a review (with discussion)," *Journal of the Italian Statistical Society*, vol. 6, pp. 97–144, 1997.

[19] W. Hardle, G. Kerkyacharian, D. Picard, and A. Tsybakov, *Wavelet, Approximation and Statistical Applications*, vol. 129 of *Lectures Notes in Statistics*, Springer, New York, NY, USA, 1998.

[20] F. Leblanc, "Wavelet linear density estimator for a discrete-time stochastic process: L_p-losses," *Statistics and Probability Letters*, vol. 27, no. 1, pp. 71–84, 1996.

[21] K. Tribouley and G. Viennet, "L_p adaptive density estimation in a alpha-mixing framework," *Annales de l'institut Henri Poincare B*, vol. 34, no. 2, pp. 179–208, 1998.

[22] E. Masry, "Wavelet-based estimation of multivariate regression functions in besov spaces," *Journal of Nonparametric Statistics*, vol. 12, no. 2, pp. 283–308, 2000.

[23] P. N. Patil and Y. K. Truong, "Asymptotics for wavelet based estimates of piecewise smooth regression for stationary time series," *Annals of the Institute of Statistical Mathematics*, vol. 53, no. 1, pp. 159–178, 2001.

[24] H. Doosti, M. Afshari, and H. A. Niroumand, "Wavelets for nonparametric stochastic regression with mixing stochastic process," *Communications in Statistics*, vol. 37, no. 3, pp. 373–385, 2008.

[25] H. Doosti and H. A. Niroumand, "Multivariate stochastic regression estimation by wavelet methods for stationary time series," *Pakistan Journal of Statistics*, vol. 27, no. 1, pp. 37–46, 2009.

[26] H. Doosti, M. S. Islam, Y. P. Chaubey, and P. Góra, "Two-dimensional wavelets for nonlinear autoregressive models with an application in dynamical system," *Italian Journal of Pure and Applied Mathematics*, no. 27, pp. 39–62, 2010.

[27] J. Cai and H. Liang, "Nonlinear wavelet density estimation for truncated and dependent observations," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 9, no. 4, pp. 587–609, 2011.

[28] S. Niu and H. Liang, "Nonlinear wavelet estimation of conditional density under left-truncated and α-mixing assumptions," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 9, no. 6, pp. 989–1023, 2011.

[29] F. Benatia and D. Yahia, "Nonlinear wavelet regression function estimator for censored dependent data," *Journal Afrika Statistika*, vol. 7, no. 1, pp. 391–411, 2012.

[30] C. Chesneau, "On the adaptive wavelet deconvolution of a density for strong mixing sequences," *Journal of the Korean Statistical Society*, vol. 41, no. 4, pp. 423–436, 2012.

[31] C. Chesneau, "Wavelet estimation of a density in a GARCH-type model," *Communications in Statistics*, vol. 42, no. 1, pp. 98–117, 2013.

[32] C. Chesneau, "On the adaptive wavelet estimation of a multidimensional regression function under alpha-mixing dependence: beyond the standard assumptions on the noise," *Commentationes Mathematicae Universitatis Carolinae*, vol. 4, pp. 527–556, 2013.

[33] Y. P. Chaubey and E. Shirazi, "On MISE of a nonlinear wavelet estimator of the regression function based on biased data under strong mixing ," *Communications in Statistics*. In press.

[34] M. Abbaszadeh and M. Emadi, "Wavelet density estimation and statistical evidences role for a garch model in the weighted distribution," *Applied Mathematics*, vol. 4, no. 2, pp. 10–416, 2013.

[35] B. Delyon and A. Juditsky, "On minimax wavelet estimators," *Applied and Computational Harmonic Analysis*, vol. 3, no. 3, pp. 215–228, 1996.

[36] G. Kerkyacharian and D. Picard, "Thresholding algorithms, maxisets and well-concentrated bases," *Test*, vol. 9, no. 2, pp. 283–344, 2000.

[37] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[38] A. Cohen, I. Daubechies, and P. Vial, "Wavelets on the interval and fast wavelet transforms," *Applied and Computational Harmonic Analysis*, vol. 1, no. 1, pp. 54–81, 1993.

[39] S. Mallat, *A Wavelet Tour of Signal Processing*, The sparse way, with contributions from Gabriel Peyralphae, Elsevier/Academic Press, Amsterdam, The Netherlands, 3rd edition, 2009.

[40] Y. Meyer, *Ondelettes et Opalphaerateurs*, Hermann, Paris, France, 1990.

[41] V. Genon-Catalot, T. Jeantheau, and C. Larédo, "Stochastic volatility models as hidden Markov models and statistical applications," *Bernoulli*, vol. 6, no. 6, pp. 1051–1079, 2000.

[42] J. Fan and J. Koo, "Wavelet deconvolution," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 734–747, 2002.

[43] A. B. Tsybakov, *Introduction alphaa l'estimation non-paramalphaetrique*, Springer, New York, NY, USA, 2004.

[44] E. Masry, "Strong consistency and rates for deconvolution of multivariate densities of stationary processes," *Stochastic Processes and their Applications*, vol. 47, no. 1, pp. 53–74, 1993.

[45] R. Kulik, "Nonparametric deconvolution problem for dependent sequences," *Electronic Journal of Statistics*, vol. 2, pp. 722–740, 2008.

[46] F. Comte, J. Dedecker, and M. L. Taupin, "Adaptive density deconvolution with dependent inputs," *Mathematical Methods of Statistics*, vol. 17, no. 2, pp. 87–112, 2008.

[47] H. van Zanten and P. Zareba, "A note on wavelet density deconvolution for weakly dependent data," *Statistical Inference for Stochastic Processes*, vol. 11, no. 2, pp. 207–219, 2008.

[48] W. Hardle, *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, UK, 1990.

[49] V. A. Vasiliev, "One investigation method of a ratios type estimators," in *Proceedings of the 16th IFAC Symposium on System Identialphacation*, pp. 1–6, Brussels, Belgium, July 2012, (in progress).

[50] Y. Davydov, "The invariance principle for stationary processes," *Theory of Probability and Its Applications*, vol. 15, no. 3, pp. 498–509, 1970.

[51] C. Chesneau, "Wavelet estimation via block thresholding: a minimax study under L^p-risk," *Statistica Sinica*, vol. 18, no. 3, pp. 1007–1024, 2008.

[52] C. Chesneau and H. Doosti, "Wavelet linear density estimation for a GARCH model under various dependence structures," *Journal of the Iranian Statistical Society*, vol. 11, no. 1, pp. 1–21, 2012.