Chilean Journal of Statistics
Vol. 3, No. 1, April 2012, 31–42

Nonparametric Statistics
Research Paper

On the wavelet estimation of a function in a density model with non-identically distributed observations

Christophe Chesneau[1,*] and Nargess Hosseinioun[2]

[1]LMNO, Université de Caen Basse-Normandie, Caen, France
[2]Department of Statistics, Payame Noor University, Iran

(Received: 10 March 2011 · Accepted in final form: 30 May 2011)

Abstract

A density model with possible non-identically distributed random variables is considered. We aim to estimate a common function appearing in the densities. We construct a new linear wavelet estimator and study its performance for independent and dependent data (the ρ-mixing case is explored). Then, in the independent case, we develop a new adaptive hard thresholding wavelet estimator and prove that it attains a sharp rate of convergence.

Keywords: Biased observations · Dependent data · Rates of convergence · Wavelet basis.

Mathematics Subject Classification: Primary 62G07 · Secondary 62G20.

1. Introduction

We consider the following density model. Let $(X_i)_{i\in\mathbb{Z}}$ be a random process such that, for any $i \in \mathbb{Z}$, the density of X_i is

$$g_i(x) = w_i(x)f(x), \qquad x \in \mathbb{R}, \tag{1}$$

where $(w_i(x))_{i\in\mathbb{Z}}$ is a known sequence of positive functions and f is an unknown positive function. Let $L > 0$ and $X_i(\Omega) = \{x \in \mathbb{R};\ g_i(x) \neq 0\}$. We suppose that $X_i(\Omega)$ does not depend on i, $X_1(\Omega) \subseteq [-L, L]$, there exists a constant $C_* > 0$ such that

$$\sup_{x\in\mathbb{R}} f(x) \leq C_*, \tag{2}$$

*Corresponding author. Christophe Chesneau. Department of Mathematics, LMNO, University of Caen, UFR de Sciences, F-14032, Caen, France. Email: christophe.chesneau@gmail.com

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

and there exists a sequence of real positive numbers $(v_i)_{i\in\mathbb{Z}}$ (which can depend on n) such that

$$\inf_{x\in X_1(\Omega)} w_i(x) \geq v_i. \tag{3}$$

The goal is to estimate f globally when only n random variables $X_1,\ldots,X_n$ of $(X_i)_{i\in\mathbb{Z}}$ are observed. Such an estimation problem has been recently investigated by Aubin and Leoni-Aubin (2008a,b). It can be viewed as a generalization of the standard biased density model; see e.g., Patil and Rao (1977), El Barmi and Simonoff (2000), Brunel et al. (2009) and Ramirez and Vidakovic (2010).

In this article, we investigate the estimation of f via the powerful tool of the wavelet analysis. Wavelets are attractive for nonparametric density estimation because of their spatial adaptivity, computational efficiency and asymptotic optimality properties. They enjoy excellent mean integrated squared error (MISE) properties and can achieve fast rates of convergence over a wide range of function classes (including spatially inhomogeneous function). Details on wavelet analysis in nonparametric function estimation can be found in Antoniadis (1997) and Härdle et al. (1998).

In the first part of this study, we develop a new linear wavelet estimator. We determine a sharp upper bound for the associated MISE for independent $(X_i)_{i\in\mathbb{Z}}$. Then, we extend this result for possible dependent $(X_i)_{i\in\mathbb{Z}}$ following the ρ-mixing case. In particular, we prove the upper bound obtained in the independent case is not deteriorated by our dependence condition as soon as the ρ-mixing coefficients $(\rho_m)_{m\in\mathbb{N}^*}$ of $(X_i)_{i\in\mathbb{Z}}$ (defined in Section 3) satisfy $\sum_{m=1}^{n}\rho_m \leq C$, where $C>0$ denotes a constant independent of n. The second part of the study is devoted to the adaptive estimation of f for independent $(X_i)_{i\in\mathbb{Z}}$. We construct a new hard thresholding wavelet estimator and prove that it attains a sharp upper bound, close to the one attained by the corresponding linear wavelet estimator. Let us mention that our results are proved under very mild assumptions on $w_1(x),\ldots,w_n(x)$.

Section 2 presents wavelets and the Besov balls. The linear wavelet estimation is developed in Section 3. Section 4 is devoted to our hard thresholding wavelet estimator. The proofs are postponed to Section 5.

2. Wavelets and Besov Balls

Let $L>0$, N be a positive integer, and ϕ and ψ be the Daubechies wavelets $db2N$ (which satisfy $\mathrm{supp}(\phi)=\mathrm{supp}(\psi)=[1-N,N]$). Set

$$\phi_{j,k}(x)=2^{j/2}\phi(2^jx-k), \qquad \psi_{j,k}(x)=2^{j/2}\psi(2^jx-k),$$

and

$$\Lambda_j=\{k\in\mathbb{Z};\ 1-N\leq 2^jx-k\leq N,\ x\in[-L,L]\}=\{k\in\mathbb{Z};\ -L2^j+N-1\leq k\leq L2^j-N\}.$$

Then, there exists an integer τ such that, for any integer $\ell\geq\tau$, the collection

$$\mathcal{B}=\{\phi_{\ell,k}(.),\ k\in\Lambda_\ell;\ \psi_{j,k}(.);\ j\in\mathbb{N}-\{0,\ldots,\ell-1\},\ k\in\Lambda_j\}$$

is an orthonormal basis of $\mathbb{L}^2([-L,L])=\{h:[-L,L]\to\mathbb{R};\ \int_{-L}^{L}h^2(x)dx<\infty\}$. For more details about wavelet basis, see Meyer (1992) and Cohen et al. (1993).

For any integer $\ell \geq \tau$, any $h \in \mathbb{L}^2([-L,L])$ can be expanded on $\mathcal{B}$ as

$$h(x) = \sum_{k\in\Lambda_\ell} \alpha_{\ell,k}\phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty}\sum_{k\in\Lambda_j} \beta_{j,k}\psi_{j,k}(x),$$

where $\alpha_{j,k}$ and $\beta_{j,k}$ are the wavelet coefficients of h defined by

$$\alpha_{j,k} = \int_{-L}^{L} h(x)\phi_{j,k}(x)dx, \qquad \beta_{j,k} = \int_{-L}^{L} h(x)\psi_{j,k}(x)dx. \tag{4}$$

Let $M > 0$, $s > 0$, $p \geq 1$, $r \geq 1$ and $\mathbb{L}^p([-L,L]) = \{h : [-L,L] \to \mathbb{R};\ \int_{-L}^{L} |h(x)|^p dx < \infty\}$. Set, for every measurable function h on $[-L,L]$ and $\epsilon \geq 0$, $\Delta_\epsilon(h)(x) = h(x+\epsilon) - h(x)$, $\Delta_\epsilon^2(h)(x) = \Delta_\epsilon(\Delta_\epsilon h)(x)$ and identically, for $N \in \mathbb{N}^*$, $\Delta_\epsilon^N(h)(x)$. Let

$$\rho^N(t,h,p) = \sup_{\epsilon\in[-t,t]} \left(\int_{-L}^{L} |\Delta_\epsilon^N(h)(u)|^p du\right)^{1/p}.$$

Then, for $s \in (0,N)$, we define the Besov ball $B_{p,r}^s(M)$ by

$$B_{p,r}^s(M) = \left\{h \in \mathbb{L}^p([-L,L]);\ \left[\int_{-L}^{L} \left(\frac{\rho^N(t,h,p)}{t^s}\right)^r \frac{dt}{t}\right]^{1/r} \leq M\right\}.$$

We have the following equivalence: $h \in B_{p,r}^s(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients give in Equation (4) satisfy

$$2^{\tau(1/2-1/p)}\left(\sum_{k\in\Lambda_\tau} |\alpha_{\tau,k}|^p\right)^{1/p} + \left\{\sum_{j=\tau}^{\infty}\left[2^{j(s+1/2-1/p)}\left(\sum_{k\in\Lambda_j}|\beta_{j,k}|^p\right)^{1/p}\right]^r\right\}^{1/r} \leq M^*. \tag{5}$$

In Equation (5), s is a smoothness parameter and p and r are norm parameters. The Besov balls capture a wide variety of smoothness features in a function; see e.g., Meyer (1992).

3. Linear Wavelet Estimation

For any integer $j \geq \tau$ and $k \in \Lambda_j$, we can estimate the unknown wavelet coefficient $\alpha_{j,k} = \int_{-L}^{L} f(x)\phi_{j,k}(x)dx$ by a standard empirical one given by

$$\widehat{\alpha}_{j,k}^* = \frac{1}{n}\sum_{i=1}^{n} \frac{\phi_{j,k}(X_i)}{w_i(X_i)}. \tag{6}$$

However, in this study, we consider

$$\widehat{\alpha}_{j,k} = \frac{1}{z_n}\sum_{i=1}^{n} v_i\frac{\phi_{j,k}(X_i)}{w_i(X_i)}, \qquad z_n = \sum_{i=1}^{n} v_i. \tag{7}$$

Our choice is motivated by the following upper bound results.

PROPOSITION 3.1 Suppose that $(X_i)_{i\in\mathbb{Z}}$ are independent. For any integer $j \geq \tau$ and $k \in \Lambda_j$, let $\alpha_{j,k} = \int_{-L}^{L} f(x)\phi_{j,k}(x)dx$, $\widehat{\alpha}^*_{j,k}$ be as in Equation (6) and $\widehat{\alpha}_{j,k}$ be as in Equation (7). Then, $\widehat{\alpha}^*_{j,k}$ and $\widehat{\alpha}_{j,k}$ are unbiased estimators of $\alpha_{j,k}$ and there exists a constant $C > 0$ such that

$$\mathrm{E}\left[(\widehat{\alpha}^*_{j,k} - \alpha_{j,k})^2\right] \leq C\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{v_i}, \qquad \mathrm{E}\left[(\widehat{\alpha}_{j,k} - \alpha_{j,k})^2\right] \leq C\frac{1}{z_n}.$$

These bounds are as sharp as possible and we have $1/z_n \leq (1/n^2)\sum_{i=1}^{n} 1/v_i$.

We define the linear wavelet estimator $\widehat{f}_{\mathrm{lin}}$ by

$$\widehat{f}_{\mathrm{lin}}(x) = \sum_{k\in\Lambda_{j_0}} \widehat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \tag{8}$$

where $\widehat{\alpha}_{j_0,k}$ is defined by Equation (7) and j_0 is an integer which is chosen later.

Naturally, taking $w_1(x) = \cdots = w_n(x) = 1$, Equation (1) becomes the standard density model and $\widehat{f}_{\mathrm{lin}}$ the standard linear wavelet estimator for this problem; see Härdle et al. (1998, Subsection 10.2). For a survey on wavelet linear estimators for various density models, we refer to Chaubey et al. (2010).

THEOREM 3.2 Suppose that $(X_i)_{i\in\mathbb{Z}}$ are independent and $\lim_{n\to\infty} z_n = \infty$. Suppose that $f \in B^s_{p,r}(M)$, with $s \in (0, N)$, $p \geq 2$ and $r \geq 1$. Let $\widehat{f}_{\mathrm{lin}}$ be as in Equation (8) with the integer j_0 satisfying $(1/2)z_n^{1/(2s+1)} \leq 2^{j_0} \leq z_n^{1/(2s+1)}$. Then, there exists a constant $C > 0$ such that

$$\mathrm{E}\left[\int_{-L}^{L}\left(\widehat{f}_{\mathrm{lin}}(x) - f(x)\right)^2 dx\right] \leq C z_n^{-2s/(2s+1)}.$$

Note that, when $w_1(x) = \cdots = w_n(x) = 1$, we have $z_n = n^{-2s/(2s+1)}$ and this is the optimal rate of convergence (in the minimax sense) for the standard density estimation problem; see Härdle et al. (1998, Theorem 10.1).

Let us now explore the performance of $\widehat{f}_{\mathrm{lin}}$ for a class of dependent $(X_i)_{i\in\mathbb{Z}}$.

DEFINITION 3.3 Let $(X_i)_{i\in\mathbb{Z}}$ be a random process. For any $u \in \mathbb{Z}$, let $\mathcal{F}^X_{-\infty,u}$ be the σ-algebra generated by $\ldots, X_{u-1}, X_u$ and $\mathcal{F}^X_{u,\infty}$ is the σ-algebra generated by $X_u, X_{u+1}, \ldots$ For any $m \in \mathbb{Z}$, we define the mth maximal correlation coefficient of $(X_i)_{i\in\mathbb{Z}}$ by

$$\rho_m = \sup_{\ell\in\mathbb{Z}} \sup_{(U,V)\in\mathbb{L}^2(\mathcal{F}^X_{-\infty,\ell})\times\mathbb{L}^2(\mathcal{F}^X_{m+\ell,\infty})} \frac{|\mathbb{C}(U,V)|}{\sqrt{\mathrm{V}[U]\mathrm{V}[V]}},$$

where, for any $\mathcal{A} \in \{\mathcal{F}^X_{-\infty,\ell}, \mathcal{F}^X_{m+\ell,\infty}\}$, $\mathbb{L}^2(\mathcal{A}) = \left\{U \in \mathcal{A};\ \ \mathrm{E}[U^2] < \infty\right\}$ and $\mathbb{C}(\cdot,\cdot)$ denotes the covariance function. Then, we say that $(X_i)_{i\in\mathbb{Z}}$ is ρ-mixing if and only if

$$\lim_{m\to\infty} \rho_m = 0.$$

Further details on ρ-mixing dependence can be found in, e.g., Kolmogorov and Rozanov (1960), Shao (1995) and Zhengyan and Lu (1996).

Results on wavelet estimation of a density in the ρ-mixing case can be found in Leblanc (1996) and Hosseinioun et al. (2010).

THEOREM 3.4 Suppose that $(X_i)_{i\in\mathbb{Z}}$ is ρ-mixing and there exist three constants $\upsilon > 0$, $\theta \in [0,1)$ and $\gamma \geq 0$ such that

$$\lim_{n\to\infty} \frac{1}{n^\theta [\log(n)]^\gamma} \sum_{m=1}^{n} \rho_m = \upsilon, \qquad \lim_{n\to\infty} \frac{z_n}{n^\theta [\log(n)]^\gamma} = \infty. \tag{9}$$

Suppose that $f \in B^s_{p,r}(M)$, with $s \in (0,N)$, $p \geq 2$ and $r \geq 1$. Let $\widehat{f}_{\text{lin}}$ be as in Equation (8), with the integer j_0 satisfying

$$\frac{1}{2}\left(\frac{z_n}{n^\theta[\log(n)]^\gamma}\right)^{1/(2s+1)} \leq 2^{j_0} \leq \left(\frac{z_n}{n^\theta[\log(n)]^\gamma}\right)^{1/(2s+1)}.$$

Then, there exists a constant $C > 0$ such that

$$\mathrm{E}\left[\int_{-L}^{L} \left(\widehat{f}_{\text{lin}}(x) - f(x)\right)^2 dx\right] \leq C\left(\frac{z_n}{n^\theta[\log(n)]^\gamma}\right)^{-2s/(2s+1)}.$$

The main role of the parameters θ and γ in Equation (9) is to measure the influence of the ρ-mixing dependence of $(X_i)_{i\in\mathbb{Z}}$ when $\lim_{n\to\infty}\sum_{m=1}^{n}\rho_m = \infty$ on the performance of $\widehat{f}_{\text{lin}}$. The first assumption in Equation (9) can be viewed as a generalization of the standard one, i.e., $\sum_{m=1}^{\infty}\rho_m \leq C$, which corresponds to $\theta = \gamma = 0$; see e.g., Leblanc (1996, Assumption M1). Observe that, if $\theta = \gamma = 0$, Theorem 3.4 extends the result of Theorem 3.2; the ρ-mixing dependence on $(X_i)_{i\in\mathbb{Z}}$ does not deteriorate the rate of convergence $z_n^{-2s/(2s+1)}$.

The main drawback of $\widehat{f}_{\text{lin}}$ is that it is not adaptive. It depends on the smoothness parameter s in its construction. The adaptive estimation of f for independent $(X_i)_{i\in\mathbb{Z}}$ is explored in the next section.

4. On the Adaptive Estimation of f in the Independent Case

Suppose that $(X_i)_{i\in\mathbb{Z}}$ are independent. We define the hard thresholding estimator $\widehat{f}_{\text{hard}}$ by

$$\widehat{f}_{\text{hard}}(x) = \sum_{k\in\Lambda_\tau} \widehat{\alpha}_{\tau,k}\phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1}\sum_{k\in\Lambda_j} \widehat{\beta}_{j,k} \mathbb{1}_{\left\{|\widehat{\beta}_{j,k}| \geq \kappa\theta\sqrt{\frac{\log(z_n)}{z_n}}\right\}} \psi_{j,k}(x), \tag{10}$$

where $\widehat{\alpha}_{\tau,k}$ is defined by Equation (7),

$$\widehat{\beta}_{j,k} = \frac{1}{z_n}\sum_{i=1}^{n} v_i \frac{\psi_{j,k}(X_i)}{w_i(X_i)} \mathbb{1}_{\left\{\left|v_i\frac{\psi_{j,k}(X_i)}{w_i(X_i)}\right| \leq \theta\sqrt{\frac{z_n}{\log(z_n)}}\right\}}, \tag{11}$$

for any random event $\mathcal{A}$, $\mathbb{1}_{\mathcal{A}}$ is the indicator function on $\mathcal{A}$, j_1 is the integer satisfying $(1/2)z_n < 2^{j_1} \leq z_n$, $\theta = \sqrt{C_*}$ and $\kappa \geq 8/3 + 2 + 2\sqrt{16/9+4}$.

The originality of $\widehat{f}_{\text{hard}}$ is in the definition of Equation (11). We do not estimate the unknown "mother" wavelet coefficient by the standard empirical estimator; we consider a thresholding version of it. This thresholding combined with a suitable calibration of the parameters allows us to have power MISE properties under very mild assumptions on $w_1(x), \ldots, w_n(x)$. Such a technique has been firstly introduced in a hard thresholding wavelet procedure in Delyon and Juditsky (1996) for nonparametric regression. Another application of this technique can be found in Chesneau (2011).

THEOREM 4.1 Suppose that $(X_i)_{i\in\mathbb{Z}}$ are independent and $\lim_{n\to\infty} z_n = \infty$. Let $\widehat{f}_{\mathrm{hard}}$ be as in Equation (10). 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 $s \in (1/p, N)\}$. Then, for a large enough n, there exists a constant $C > 0$ such that

$$\mathrm{E}\left[\int_{-L}^{L}\left(\widehat{f}_{\mathrm{hard}}(x) - f(x)\right)^2 dx\right] \leq C\left(\frac{\log(z_n)}{z_n}\right)^{2s/(2s+1)}.$$

Theorem 4.1 shows that $\widehat{f}_{\mathrm{hard}}$ attains a rate of convergence close to one attains by $\widehat{f}_{\mathrm{lin}}$. The only difference is the "negligible" logarithmic term $[\log(n)]^{2s/(2s+1)}$. Mention that the proof of Theorem 4.1 is based on Chesneau (2011, Theorem 2).

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 depends on ϕ or ψ.

PROOF [Proposition 3.1] We have

$$\mathrm{E}\left[\widehat{\alpha}^*_{j,k}\right] = \frac{1}{n}\sum_{i=1}^{n}\int_{-L}^{L}\frac{\phi_{j,k}(x)}{w_i(x)}g_i(x)dx = \int_{-L}^{L}\phi_{j,k}(x)f(x)dx = \alpha_{j,k}. \tag{12}$$

Using Equation (12), the independence of $X_1, \ldots, X_n$, Equations (2) and (3) and $\int_{-L}^{L}(\phi_{j,k}(x))^2 dx = 1$, we obtain

$$\begin{aligned}
\mathrm{E}\left[(\widehat{\alpha}^*_{j,k} - \alpha_{j,k})^2\right] = \mathrm{V}\left[\widehat{\alpha}^*_{j,k}\right] &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}\left[\frac{\phi_{j,k}(X_i)}{w_i(X_i)}\right] \\
&\leq \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left[\left(\frac{\phi_{j,k}(X_i)}{w_i(X_i)}\right)^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}\int_{-L}^{L}\left(\frac{1}{w_i(x)}\phi_{j,k}(x)\right)^2 g_i(x)dx \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\int_{-L}^{L}(\phi_{j,k}(x))^2\frac{f(x)}{w_i(x)}dx \leq C_*\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{v_i}.
\end{aligned}$$

We have

$$\mathrm{E}\left[\widehat{\alpha}_{j,k}\right] = \frac{1}{z_n}\sum_{i=1}^{n}v_i\int_{-L}^{L}\frac{\phi_{j,k}(x)}{w_i(x)}g_i(x)dx = \frac{1}{z_n}z_n\int_{-L}^{L}\phi_{j,k}(x)f(x)dx = \alpha_{j,k}. \tag{13}$$

Using Equation (13), again the independence of $X_1, \ldots, X_n$, Equations (2), (3) and

$\int_{-L}^{L}(\phi_{j,k}(x))^2\,dx = 1$, we obtain

$$\begin{aligned}
\mathrm{E}\left[(\widehat{\alpha}_{j,k}-\alpha_{j,k})^2\right] = \mathrm{V}\left[\widehat{\alpha}_{j,k}\right] &= \frac{1}{z_n^2}\sum_{i=1}^{n} v_i^2 \mathrm{V}\left[\frac{\phi_{j,k}(X_i)}{w_i(X_i)}\right] \\
&\leq \frac{1}{z_n^2}\sum_{i=1}^{n} v_i^2 \mathrm{E}\left[\left(\frac{\phi_{j,k}(X_i)}{w_i(X_i)}\right)^2\right] = \frac{1}{z_n^2}\sum_{i=1}^{n} v_i^2 \int_{-L}^{L}\left(\frac{\phi_{j,k}(x)}{w_i(x)}\right)^2 g_i(x)dx \\
&= \frac{1}{z_n^2}\sum_{i=1}^{n} v_i^2 \int_{-L}^{L}(\phi_{j,k}(x))^2\frac{f(x)}{w_i(x)}dx \leq C_*\frac{1}{z_n^2}z_n = C_*\frac{1}{z_n}. \qquad (14)
\end{aligned}$$

The Hölder inequality yields

$$n = \sum_{i=1}^{n}\sqrt{v_i}\frac{1}{\sqrt{v_i}} \leq z_n^{1/2}\left(\sum_{i=1}^{n}\frac{1}{v_i}\right)^{1/2}.$$

Therefore

$$\frac{1}{z_n} \leq \frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{v_i}.$$

The proof of Proposition 3.1 is complete. ■

Proof [Theorem 3.2] We expand the function f on $\mathcal{B}$ as

$$f(x) = \sum_{k\in\Lambda_{j_0}}\alpha_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j}\beta_{j,k}\psi_{j,k}(x),$$

where $\alpha_{j_0,k} = \int_{-L}^{L} f(x)\phi_{j_0,k}(x)dx$ and $\beta_{j,k} = \int_{-L}^{L} f(x)\psi_{j,k}(x)dx$.

Using the fact that $\mathcal{B}$ is an orthonormal basis of $\mathbb{L}^2([-L,L])$, Proposition 3.1 and, since $p \geq 2$, $B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$, we have

$$\begin{aligned}
\mathrm{E}\left[\int_{-L}^{L}\left(\widehat{f}_{\mathrm{lin}}(x)-f(x)\right)^2 dx\right] &= \sum_{k\in\Lambda_{j_0}}\mathrm{E}\left[(\widehat{\alpha}_{j_0,k}-\alpha_{j_0,k})^2\right] + \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j}\beta_{j,k}^2 \\
&\leq C\left(2^{j_0}\frac{1}{z_n} + 2^{-2j_0 s}\right) \leq C z_n^{-2s/(2s+1)}.
\end{aligned}$$

Theorem 3.2 is proved. ■

Proof [Theorem 3.4] First of all, let us prove the existence of a constant $C > 0$ such that

$$\mathrm{E}\left[(\widehat{\alpha}_{j_0,k}-\alpha_{j_0,k})^2\right] \leq C n^{\theta}[\log(n)]^{\gamma}\frac{1}{z_n}.$$

Since $\widehat{\alpha}_{j_0,k}$ is an unbiased estimator of $\alpha_{j_0,k}$, we have

$$\begin{aligned}
\mathrm{E}\left[(\widehat{\alpha}_{j_0,k}-\alpha_{j_0,k})^2\right] &= \mathrm{V}\left[\widehat{\alpha}_{j_0,k}\right] \\
&= \frac{1}{z_n^2}\sum_{i=1}^{n}\sum_{\ell=1}^{n} v_i v_\ell \mathbb{C}\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}, \frac{\phi_{j_0,k}(X_\ell)}{w_\ell(X_\ell)}\right) \\
&\leq \frac{1}{z_n^2}\sum_{i=1}^{n} v_i^2 \mathrm{V}\left[\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}\right] + \frac{1}{z_n^2}\sum_{i=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq i}}^{n} v_i v_\ell \left|\mathbb{C}\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}, \frac{\phi_{j_0,k}(X_\ell)}{w_\ell(X_\ell)}\right)\right|.
\end{aligned} \tag{15}$$

It follows from Equation (14) that

$$\frac{1}{z_n^2}\sum_{i=1}^{n} v_i^2 \mathrm{V}\left[\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}\right] \leq C\frac{1}{z_n}.$$

In order to bound the second term in Equation (15), we use the following result on ρ-mixing. The proof can be found in Doukhan (1994, Section 1.2.2.).

LEMMA 5.1 Let $(X_i)_{i\in\mathbb{Z}}$ be a ρ-mixing sequence. Then, for any $(i,j)\in\mathbb{Z}^2$ such that $i\neq \ell$ and any functions g and h, we have

$$|\mathbb{C}\left(h(X_i), g(X_\ell)\right)| \leq \rho_{|i-\ell|}\sqrt{\mathrm{E}\left[(h(X_i))^2\right]\mathrm{E}\left[(g(X_\ell))^2\right]},$$

whenever these quantities exist.

Using Lemma 5.1, we obtain

$$\sum_{i=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq i}}^{n} v_i v_\ell \left|\mathbb{C}\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}, \frac{\phi_{j_0,k}(X_\ell)}{w_\ell(X_\ell)}\right)\right| \leq \sum_{i=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq i}}^{n} v_i v_\ell \rho_{|i-\ell|}\sqrt{\mathrm{E}\left[\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}\right)^2\right]\mathrm{E}\left[\left(\frac{\phi_{j_0,k}(X_\ell)}{w_\ell(X_\ell)}\right)^2\right]}. \tag{16}$$

By Equations (2), (3) and $\int_{-L}^{L}\left(\phi_{j_0,k}(x)\right)^2 dx = 1$, we have

$$\mathrm{E}\left[\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}\right)^2\right] = \int_{-L}^{L}\left(\frac{\phi_{j_0,k}(x)}{w_i(x)}\right)^2 g_i(x)dx = \int_{-L}^{L}(\phi_{j_0,k}(x))^2\frac{f(x)}{w_i(x)}dx \leq C_*\frac{1}{v_i}.$$

Therefore

$$\sum_{i=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq i}}^{n} v_i v_\ell \left|\mathbb{C}\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}, \frac{\phi_{j_0,k}(X_\ell)}{w_\ell(X_\ell)}\right)\right| \leq C\sum_{i=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq i}}^{n} \sqrt{v_i}\sqrt{v_\ell}\rho_{|i-\ell|}$$
$$\leq C\sum_{i=2}^{n}\sum_{\ell=1}^{i-1} \sqrt{v_i}\sqrt{v_\ell}\rho_{i-\ell}$$
$$\leq C\sum_{i=2}^{n}\sum_{\ell=1}^{i-1} (v_i + v_\ell)\,\rho_{i-\ell}$$
$$= C\sum_{i=2}^{n}\sum_{u=1}^{i-1} (v_i + v_{i-u})\,\rho_u$$
$$= C\left(\sum_{i=2}^{n} v_i \sum_{u=1}^{i-1}\rho_u + \sum_{i=2}^{n}\sum_{u=1}^{i-1} v_{i-u}\rho_u\right).$$

Using Equation (9), we obtain

$$\sum_{i=2}^{n} v_i \sum_{u=1}^{i-1}\rho_u \leq z_n \sum_{u=1}^{n}\rho_u \leq Cn^\theta[\log(n)]^\gamma z_n,$$

and

$$\sum_{i=2}^{n}\sum_{u=1}^{i-1} v_{i-u}\rho_u = \sum_{u=1}^{n-1}\rho_u \sum_{i=u+1}^{n} v_{i-u} \leq z_n \sum_{u=1}^{n}\rho_u \leq Cn^\theta[\log(n)]^\gamma z_n.$$

Hence

$$\sum_{i=1}^{n}\sum_{\substack{\ell=1\\ \ell\neq i}}^{n} v_i v_\ell \left|\mathbb{C}\left(\frac{\phi_{j_0,k}(X_i)}{w_i(X_i)}, \frac{\phi_{j_0,k}(X_\ell)}{w_\ell(X_\ell)}\right)\right| \leq Cn^\theta[\log(n)]^\gamma z_n. \tag{17}$$

Putting Equations (15), (16) and (17) together, we obtain

$$\mathrm{E}\left[(\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2\right] \leq C\left(\frac{1}{z_n} + \frac{n^\theta[\log(n)]^\gamma}{z_n}\right) \leq C\frac{n^\theta[\log(n)]^\gamma}{z_n}.$$

Then we proceed as in Theorem 3.2. We expand the function f on $\mathcal{B}$ as

$$f(x) = \sum_{k\in\Lambda_{j_0}} \alpha_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j} \beta_{j,k}\psi_{j,k}(x),$$

where $\alpha_{j_0,k} = \int_{-L}^{L} f(x)\phi_{j_0,k}(x)dx$ and $\beta_{j,k} = \int_{-L}^{L} f(x)\psi_{j,k}(x)dx$. Using the fact that $\mathcal{B}$ is an orthonormal basis of $\mathbb{L}^2([-L,L])$ and, since $p \geq 2$, $B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$, we obtain

$$\mathrm{E}\left[\int_{-L}^{L}\left(\widehat{f}_{\mathrm{lin}}(x)-f(x)\right)^2 dx\right] = \sum_{k\in\Lambda_{j_0}} \mathrm{E}\left[(\widehat{\alpha}_{j_0,k}-\alpha_{j_0,k})^2\right] + \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j}\beta_{j,k}^2$$
$$\leq C\left(2^{j_0}\frac{n^{\theta}[\log(n)]^{\gamma}}{z_n} + 2^{-2j_0 s}\right) \leq C\left(\frac{z_n}{n^{\theta}[\log(n)]^{\gamma}}\right)^{-2s/(2s+1)}.$$

The proof of Theorem 3.4 is complete. ■

Proof [Theorem 4.1.] The result is proven using the following general result. It is a reformulation of the result given in Chesneau (2011, Theorem 2).

Theorem 5.2 (Chesneau, 2011). Let $L > 0$. We want to estimate an unknown function f with support in $[-L,L]$ from n independent random variables $U_1,\ldots,U_n$. We consider the wavelet basis $\mathcal{B}$ and the notations of Section 3. Suppose that there exist n functions $h_1,\ldots,h_n$ such that, for any $\gamma \in \{\phi,\psi\}$,

- (A1) Any integer $j \geq \tau$ and any $k \in \Lambda_j$,

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} h_i(\gamma_{j,k},U_i)\right] = \int_{-L}^{L} f(x)\gamma_{j,k}(x)dx.$$

- (A2) There exist a sequence of real numbers $(\mu_i)_{i\in\mathbb{N}^*}$ satisfying $\lim_{i\to\infty}\mu_i = \infty$ and two constants, $\theta_\gamma > 0$ and $\delta > 0$, such that, for any integer $j \geq \tau$ and any $k \in \Lambda_j$,

$$\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left[(h_i(\gamma_{j,k},U_i))^2\right] \leq \theta_\gamma^2 2^{2\delta j}\frac{1}{\mu_n}.$$

We define the hard thresholding estimator $\widehat{f}_H$ by

$$\widehat{f}_H(x) = \sum_{k\in\Lambda_\tau}\widehat{\alpha}_{\tau,k}\phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1}\sum_{k\in\Lambda_j}\widehat{\beta}_{j,k}\mathbb{1}_{\left\{|\widehat{\beta}_{j,k}|\geq\kappa\lambda_{j,n}\right\}}\psi_{j,k}(x),$$

where

$$\widehat{\alpha}_{j,k} = \frac{1}{n}\sum_{i=1}^{n} h_i(\phi_{j,k},U_i), \qquad \widehat{\beta}_{j,k} = \frac{1}{n}\sum_{i=1}^{n} h_i(\psi_{j,k},U_i)\mathbb{1}_{\{|h_i(\psi_{j,k},U_i)|\leq\eta_{j,n}\}},$$

for any random event $\mathcal{A}$, $\mathbb{1}_{\mathcal{A}}$ is the indicator function on $\mathcal{A}$,

$$\eta_{j,n} = \theta_\psi 2^{\delta j}\sqrt{\frac{\mu_n}{\log(\mu_n)}}, \qquad \lambda_{j,n} = \theta_\psi 2^{\delta j}\sqrt{\frac{\log(\mu_n)}{\mu_n}},$$

$\kappa = 8/3 + 2 + 2\sqrt{16/9+4}$ and j_1 is the integer satisfying $(1/2)\mu_n^{1/(2\delta+1)} < 2^{j_1} \leq \mu_n^{1/(2\delta+1)}$.

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

$$\mathrm{E}\left[\int_{-L}^{L} \left(\widehat{f}_H(x) - f(x)\right)^2 dx\right] \leq C \left(\frac{\log(\mu_n)}{\mu_n}\right)^{2s/(2s+2\delta+1)}.$$

Let us now investigate the assumptions (A1) and (A2) of Theorem 5.2 with, for any $i \in \{1, \ldots, n\}$, $U_i = X_i$, $\theta_\psi = \sqrt{C_*}$, $\delta = 0$, $\mu_n = z_n$ and

$$h_i(\gamma_{j,k}, y) = \frac{n}{z_n} v_i \frac{\gamma_{j,k}(y)}{w_i(y)}.$$

- On (A1). By Proposition 3.1, for any $\gamma \in \{\phi, \psi\}$, we have

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} h_i(\gamma_{j,k}, X_i)\right] = \int_{-L}^{L} f(x)\gamma_{j,k}(x)dx.$$

- On (A2). Using Equation (14), we have

$$\frac{1}{n^2}\sum_{i=1}^{n} \mathrm{E}\left[(h_i(\gamma_{j,k}, X_i))^2\right] \leq C_* \frac{1}{z_n}.$$

Let $f \in B^s_{p,r}(M)$. It follows from Theorem 5.2 that the hard thresholding estimator give in Equation (10) satisfies, for any $r \geq 1$, $\{p \geq 2$ and $s \in (0, N)\}$ or $\{p \in [1, 2)$ and $s \in (1/p, N)\}$,

$$\mathrm{E}\left[\int_{-L}^{L} \left(\widehat{f}_{\mathrm{hard}}(x) - f(x)\right)^2 dx\right] \leq C \left(\frac{\log(z_n)}{z_n}\right)^{2s/(2s+1)}.$$

The proof of Theorem 4.1 is complete. ■

Acknowledgments

The authors would like to thank three anonymous referees whose constructive comments and remarks have considerably improved the paper.

References

Antoniadis, A., 1997. Wavelets in statistics: a review (with discussion). Journal of the Italian Statistical Society Series B, 6, 97–144.

Aubin, J.-B., Leoni-Aubin, S., 2008. Adaptive Projection Density Estimation Under a m-Sample Semiparametric Model. Annales de l'I.S.U.P. Volume 5 - Numero espécial, Volume 52, Fasc. 1-2, pp. 139–156.

Aubin, J.-B., Leoni-Aubin, S., 2008. Projection density estimation under a m-sample semiparametric model. Computational Statistics and Data Analysis, 5, 2451–2468.

El Barmi, H., Simonoff, J.S., 2000. Transformation-based estimation for weighted distributions. Journal of Nonparametric Statistics, 12, 861–878.

Brunel, E., Comte, F., Guilloux, A., 2009. Nonparametric density estimation in presence of bias and censoring. TEST, 1, 166–194.

Chaubey, Y.P., Chesneau, C., Doosti, H., 2011. Wavelet linear density estimation: a review. Journal of the Indian Society of Agricultural Statistics, 65, 169-179.

Chesneau, C., 2011. Adaptive wavelet estimator for a function and its derivatives in an indirect convolution model. Journal of Statistical Theory and Practice, 2, 303–326.

Cohen, A., Daubechies, I., Jawerth, B., Vial, P., 1993. Wavelets on the interval and fast wavelet transforms. Applied and Computational Harmonic Analysis, 1, 54–81.

Delyon, B., Juditsky, A., 1996. On minimax wavelet estimators. Applied and Computational Harmonic Analysis, 3, 215–228.

Doukhan, P., 1994. Mixing: Properties and Examples, Lecture Notes in Statistics – Volume 85. Springer, New York.

Härdle, W., Kerkyacharian, G., Picard, D., Tsybakov, A., 1998. Wavelet, Approximation and Statistical Applications, Lectures Notes in Statistics. Springer Verlag, New York.

Hosseinioun, N., Doosti, H., Nirumand, H.A., 2012. Nonparametric estimation of the derivatives of a density by the method of wavelet for mixing sequences. Statistical Papers, 53, 195-203.

Kolmogorov, A.N., Rozanov, Yu.A., 1960. On strong mixing conditions for stationary Gaussian processes. Theory of Probability and Applications, 5, 204–208.

Leblanc, F., 1996. Wavelet linear density estimator for a discrete time stochastic process: L_p-losses. Statistics and Probability Letters, 27, 71–84.

Meyer, Y., 1992. Wavelets and Operators. Cambridge University Press, Cambridge.

Patil, G.P., Rao, sC.R., 1977. The weighted distributions: a survey of their applications. In Krishnaiah, P.R., (ed.). Applications of Statistics. North-Holland, Amsterdam, pp. 383–405.

Ramirez, P., Vidakovic, B., 2010. Wavelet density estimation for stratified size-biased sample. Journal of Statistical Planning and Inference, 2, 419–432.

Shao, Q.-M., 1995. Maximal inequality for partial sums of ρ-mixing sequences. Annals of Probability, 23, 948–965.

Zhengyan, L., Lu, C., 1996. Limit Theory for Mixing Dependent Random Variables. Kluwer, Dordrecht.