Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Convergence rates of wavelet density estimation for negatively dependent sample

Junlian Xu[1*]

*Correspondence:
xujunlian83913@163.com
[1]School of Mathematics and Information Science, Baoji University of Arts and Sciences, Baoji, P.R. China

Abstract

In this paper, linear and nonlinear wavelet estimators are defined for a density in a Besov space based on a negatively dependent random sample, and their upper bounds on L^p $(1 \le p < \infty)$ risk are provided.

Keywords: Wavelet estimator; Density function; Negatively dependent; Besov space

1 Introduction

Random variables $X_1, X_2, \ldots, X_n$ are said to be negatively dependent (ND), if for any $x_1, x_2, \ldots, x_n \in \mathbb{R}$,

$$\mathbf{P}(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) \le \prod_{i=1}^{n} P(X_i \le x_i),$$

and

$$\mathbf{P}(X_1 > x_1, X_2 > x_2, \ldots, X_n > x_n) \le \prod_{i=1}^{n} P(X_i > x_i).$$

The definition was introduced by Bozorgnia [3]. Further discussion and related concepts can be found in [2, 10]. ND random variables are very useful in reliability theory and applications. Because of the wide applications, the notion of ND random variables has been receiving more and more attention recently. A series of useful results have been established (see [13–16]). Hence, we consider density estimation for ND random variables in this paper.

For density estimation, Donoho et al. [6] defined wavelet estimators and showed their convergence rates on L^p-loss, when $X_1, X_2, \ldots, X_n$ are independent. They found that the convergence rate of the nonlinear estimator is better than that of the linear one. In many cases, random variables $X_1, X_2, \ldots, X_n$ are dependent. Doosti et al. [8] proposed a linear wavelet estimator and evaluated its L^p $(1 \le p < \infty)$ risks for negatively associated random variables. Soon afterwards, the above results were extended to the case of negatively dependent sequences [7]. Chesneau [4] and Liu [12] also considered density estimation for an NA sample. Kou [11] defined linear and nonlinear wavelet estimators for mixing data and obtained their convergence rates.

Motivated by the above work, this paper will estimate the unknown density function f from a sequence of ND data $X_1, X_2, \ldots, X_n$. We shall define wavelet estimators and give their upper bounds on L^p-loss. It turns out that our results reduce to Donoho's classical theorems in [6], when the random sample is independent.

We establish our results on Besov spaces on a compact subset of the real line $\mathbb{R}$. As usual, the Sobolev spaces with integer exponents are defined as

$$W_r^n(\mathbb{R}) := \{f \in L^r(\mathbb{R}), f^{(n)} \in L^r(\mathbb{R})\}$$

with $\|f\|_{W_r^n} := \|f\|_r + \|f^{(n)}\|_r$. Then $L^r(\mathbb{R})$ can be considered as $W_r^0(\mathbb{R})$. For $1 \leq r, q \leq \infty$ and $s = n + \alpha$ with $\alpha \in (0, 1]$, a Besov space on $\mathbb{R}$ means

$$B_{r,q}^s(\mathbb{R}) := \{f \in W_r^n(\mathbb{R}), \|t^{-\alpha}\omega_r^2(f^{(n)}, t)\|_q^* < \infty\}$$

with the norm $\|f\|_{srq} := \|f\|_{W_r^n} + \|t^{-\alpha}\omega_r^2(f^{(n)}, t)\|_q^*$, where $\omega_r^2(f, t) := \sup_{|h| \leq t} \|f(\cdot + 2h) - 2f(\cdot + h) + f(\cdot)\|_r$ stands for the smoothness modulus of f and

$$\|h\|_q^* = \begin{cases} (\int_0^\infty |h(t)|^q \frac{dt}{t})^{\frac{1}{q}}, & \text{if } 1 \leq q < \infty; \\ \operatorname{ess\,sup}_t |h(t)|, & \text{if } q = \infty. \end{cases}$$

We always assume $f \in B_{r,q}^s(\mathbb{R}, L) = \{f \in B_{r,q}^s(\mathbb{R}), f \text{ is a probability density and } \|f\|_{srq} \leq L\}$ with $L > 0$. Let $\phi \in C_0^t(\mathbb{R})$ be an orthonormal scaling function with $t > \max\{s, 1\}$. Then ϕ is a function of bounded variation (BV). The corresponding wavelet function is denoted by ψ. It is well known that $\{\phi_{J,k}, \psi_{j,k}, j \geq J, k \in Z\}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$, where $\phi_{J,k}(x) := 2^{\frac{J}{2}}\psi(2^J x - k)$, $\psi_{j,k}(x) := 2^{\frac{j}{2}}\psi(2^j x - k)$ as in wavelet analysis [5]. Then for each $f \in L^2(\mathbb{R})$, $\alpha_{J,k} = \int f(x)\phi_{J,k}(x)\,dx$, and $\beta_{j,k} = \int f(x)\psi_{j,k}(x)\,dx$, we have

$$f(x) = \sum_{k \in \mathbb{Z}} \alpha_{J,k}\phi_{J,k}(x) + \sum_{j \geq J} \sum_{k \in \mathbb{Z}} \beta_{j,k}\psi_{j,k}(x).$$

Here and in what follows, $A \lesssim B$ denotes $A \leq CB$ for some constant $C > 0$; $A \gtrsim B$ means $B \lesssim A$; $A \sim B$ stands for both $A \lesssim B$ and $B \lesssim A$. The following theorems are needed in our discussion:

Theorem 1.1 (Härdle et al. [9]) *Let $f \in L^r(\mathbb{R})$ $(1 \leq r \leq \infty)$, $\alpha_{J,k} = \int f(x)\phi_{J,k}(x)\,dx$ and $\beta_{j,k} = \int f(x)\psi_{j,k}(x)\,dx$. The following assertions are equivalent*:

(i) $f \in B_{r,q}^s(\mathbb{R})$, $s > 0$, $1 \leq q \leq \infty$;

(ii) $\{2^{js}\|P_j f - f\|_r\}_{j \geq 0} \in l^q$ *with* $P_j f := \sum_{k \in \mathbb{Z}} \alpha_{j,k}\phi_{j,k}$;

(iii) $\|\alpha_{J\cdot}\|_r + \|\{2^{j(s+\frac{1}{2}-\frac{1}{r})}\|\beta_{j\cdot}\|_r\}_{j \geq 0}\|_q < +\infty$.

Moreover,

$$\|f\|_{srq} \sim \|(2^{js}\|P_j f - f\|_r)_{j \geq 0}\|_q \sim \|\alpha_{J\cdot}\|_r + \|\{2^{j(s+\frac{1}{2}-\frac{1}{r})}\|\beta_{j\cdot}\|_r\}_{j \geq 0}\|_q.$$

Theorem 1.2 (Härdle et al. [9]) *Let $\theta_\phi(x) := \sum_k |\phi(x - k)|$ and $\operatorname{ess\,sup}_x \theta_\phi(x) < \infty$. Then for $\lambda = \{\lambda_k\} \in l^r(\mathbb{Z})$ and $1 \leq r \leq \infty$,*

$$\left\|\sum_{k \in \mathbb{Z}} \lambda_k \phi_{jk}\right\|_r \sim 2^{j(\frac{1}{2}-\frac{1}{r})}\|\lambda\|_r.$$

Negatively dependent random variables possess the following property which will be used in this paper.

Theorem 1.3 (Bozorgnia et al. [3]) *Let $X_1,\dots,X_n$ be a sequence of ND random variables and let $A_1,\dots,A_m$ be some pairwise disjoint nonempty subsets of $\{1,\dots,n\}$ with $\alpha_i = \sharp(A_i)$, where $\sharp(A)$ denotes the number of elements in the set A. If $f_i : \mathbb{R}^{\alpha_i} \to \mathbb{R}$ $(i = 1,\dots,m)$ are m coordinatewise nondecreasing (nonincreasing) functions, then $f_1(X_i, i \in A_1),\dots,f_m(X_i, i \in A_m)$ are also ND. In particular, for any $t_i \geq 0 (\leq 0)$, $1 \leq i \leq m$,*

$$\mathbf{E}\left[\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right] \leq \prod_{i=1}^{n} \mathbf{E}\left[\exp(t_i X_i)\right].$$

2 Linear estimators

In this section, we shall give a linear wavelet estimator for a density function $f(x)$ in a Besov space.

The linear wavelet estimator of $f(x)$ is defined as follows:

$$\hat{f_n^{\mathrm{lin}}}(x) = \sum_{k\in K_0} \hat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \tag{1}$$

where $K_0 = \{k \in \mathbb{Z}, \operatorname{supp} f \cap \operatorname{supp} \phi_{j_0,k} \neq \emptyset\}$,

$$\hat{\alpha}_{j_0,k} = \frac{1}{n}\sum_{i=1}^{n} \phi_{j_0,k}(X_i). \tag{2}$$

The following inequalities play important roles in this paper.

Lemma 2.1 (Rosenthal's inequality, see Asadian et al. [1]) *Let $X_1,\dots,X_n$ be a sequence of ND random variables, which satisfy $\mathbf{E}X_i = 0$ and $\mathbf{E}|X_i|^p < \infty$, where $i = 1,\dots,n$. Then*

$$\mathbf{E}\left(\left|\sum_{i=1}^{n} X_i\right|^p\right) \lesssim \sum_{i=1}^{n} \mathbf{E}|X_i|^p + \left(\sum_{i=1}^{n} \mathbf{E}X_i^2\right)^{\frac{p}{2}}, \quad p \geq 2,$$

$$\mathbf{E}\left(\left|\sum_{i=1}^{n} X_i\right|^p\right) \leq \left(\sum_{i=1}^{n} \mathbf{E}X_i^2\right)^{\frac{p}{2}}, \quad 0 < p \leq 2.$$

Lemma 2.2 *Let $X_1, X_2,\dots,X_n$ be ND random variables and let the density function f be bounded and compactly supported with support length less than $H > 0$. Then for $\hat{\alpha}_{j_0,k}$ defined by (2) we have*

$$\mathbf{E}|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p \lesssim n^{-\frac{p}{2}}$$

for $1 \leq p < \infty$ and $2^{j_0} \leq n$.

Proof By the definition of $\hat{\alpha}_{j_0,k}$, one has

$$\mathbf{E}|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p = \frac{1}{n^p}\mathbf{E}\left|\sum_{i=1}^{n}\left[\phi_{j_0,k}(X_i) - \alpha_{j_0,k}\right]\right|^p. \tag{3}$$

Let $\xi_i := \phi_{j_0,k}(X_i) - \alpha_{j_0,k}$ $(i = 1, 2, \ldots, n)$. Clearly,

$$\mathbf{E}\left|\sum_{i=1}^{n}\left[\phi_{j_0,k}(X_i) - \alpha_{j_0,k}\right]\right|^p = \mathbf{E}\left|\sum_{i=1}^{n}\xi_i\right|^p. \tag{4}$$

One can choose a scaling function ϕ, which a function of bounded variation, and assume $\phi := \tilde{\phi} - \bar{\phi}$, where $\tilde{\phi}$ and $\bar{\phi}$ are bounded, nonnegative and nondecreasing functions. Define

$$\tilde{\alpha}_{j_0,k} := \int \tilde{\phi}_{j_0,k}(x)f(x)\,dx, \qquad \bar{\alpha}_{j_0,k} := \int \bar{\phi}_{j_0,k}(x)f(x)\,dx,$$

and

$$\tilde{\xi}_i := \tilde{\phi}_{j_0,k}(X_i) - \tilde{\alpha}_{j_0,k}, \qquad \bar{\xi}_i := \bar{\phi}_{j_0,k}(X_i) - \bar{\alpha}_{j_0,k}.$$

Then $\alpha_{j_0,k} = \tilde{\alpha}_{j_0,k} - \bar{\alpha}_{j_0,k}$, $\xi_i = \tilde{\xi}_i - \bar{\xi}_i$ and

$$\mathbf{E}\left|\sum_{i=1}^{n}\xi_i\right|^p = \mathbf{E}\left|\sum_{i=1}^{n}(\tilde{\xi}_i - \bar{\xi}_i)\right|^p. \tag{5}$$

It is easy to see that $\mathbf{E}\tilde{\xi}_i = 0$, the random variables $\tilde{\xi}_1, \ldots, \tilde{\xi}_n$ are ND due to the nondecreasing property $\tilde{\phi}$ and Theorem 1.3. To apply the Rosenthal's inequality, one shows an inequality

$$\mathbf{E}|\tilde{\xi}_i|^m \lesssim 2^{\frac{(m-2)j_0}{2}} \tag{6}$$

for $m \geq 2$. In fact,

$$\mathbf{E}|\tilde{\xi}_i|^m = \mathbf{E}\left|\tilde{\phi}_{j_0,k}(X_i) - \tilde{\alpha}_{j_0,k}\right|^m \lesssim \mathbf{E}\left|\tilde{\phi}_{j_0,k}(X_i)\right|^m + |\tilde{\alpha}_{j_0,k}|^m. \tag{7}$$

Note that $|\tilde{\phi}_{j_0,k}(x)| \lesssim 2^{\frac{j_0}{2}}$. Then for $m \geq 2$,

$$\begin{aligned}\mathbf{E}\left|\tilde{\phi}_{j_0,k}(X_i)\right|^m &= \mathbf{E}\left[\left|\tilde{\phi}_{j_0,k}(X_i)\right|^2\left|\tilde{\phi}_{j_0,k}(X_i)\right|^{m-2}\right] \\ &\lesssim 2^{\frac{(m-2)j_0}{2}}\mathbf{E}\left|\tilde{\phi}^2_{j_0,k}(X_i)\right|.\end{aligned} \tag{8}$$

Note that $f \in B^s_{r,q}(\mathbb{R}, L) \subseteq B^{s-\frac{1}{r}}_{\infty,q}(\mathbb{R}, L)$. Then $\|f\|_\infty \leq L$. Using $\tilde{\phi} \in L^2(\mathbb{R})$, one knows that

$$\mathbf{E}\left|\tilde{\phi}^2_{j_0,k}(X_i)\right| \lesssim \int (\tilde{\phi}_{j_0,k})^2(x)f(x)\,dx = \int \left|\tilde{\phi}(x-k)\right|^2 f\left(2^{-j}x\right)dx \lesssim 1,$$

and $|\tilde{\alpha}_{j_0,k}| = |\int f(x)\tilde{\phi}_{j_0,k}(x)\,dx| \lesssim 1$ because of supp$f$ is contained in some interval I with length $|I| \leq H$. This, together with (8) and (7), leads to (6).

By Rosenthal's inequality with $1 \leq p \leq 2$,

$$\mathbf{E}\left|\sum_{i=1}^{n}\tilde{\xi}_i\right|^p \leq \left[\sum_{i=1}^{n_m}\mathbf{E}(\tilde{\xi}_i)^2\right]^{\frac{p}{2}} \lesssim n^{\frac{p}{2}}.$$

Similarly, $\mathbf{E}|\sum_{i=1}^{n}\bar{\xi}_i|^p \lesssim n^{\frac{p}{2}}$. Combining this with (5), one has

$$\mathbf{E}\left|\sum_{i=1}^{n}\xi_i\right|^p \lesssim \mathbf{E}\left|\sum_{i=1}^{n}\tilde{\xi}_i\right|^p + \mathbf{E}\left|\sum_{i=1}^{n}\bar{\xi}_i\right|^p \lesssim n^{\frac{p}{2}}. \tag{9}$$

Substituting (9) into (4), one obtains

$$\mathbf{E}\left|\sum_{i=1}^{n}\left[\phi_{j_0,k}(X_i) - \alpha_{j_0,k}\right]\right|^p \lesssim n^{\frac{p}{2}}.$$

This with (3) shows that for $1 \le p \le 2$,

$$E|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p \lesssim \frac{1}{n^p} \times n^{\frac{p}{2}} = n^{-\frac{p}{2}}. \tag{10}$$

When $2 \le p < \infty$, Rosenthal's inequality and (6) show that

$$\mathbf{E}\left|\sum_{i=1}^{n}\tilde{\xi}_i\right|^p \lesssim \sum_{i=1}^{n}\mathbf{E}|\tilde{\xi}_i|^p + \left[\sum_{i=1}^{n}\mathbf{E}(\tilde{\xi}_i)^2\right]^{\frac{p}{2}} \lesssim n2^{\frac{(p-2)j_0}{2}} + n^{\frac{p}{2}}.$$

Similarly, $\mathbf{E}|\sum_{i=1}^{n}\bar{\xi}_i|^p \lesssim n2^{\frac{(p-2)j_0}{2}} + n^{\frac{p}{2}}$. Hence $\mathbf{E}|\sum_{i=1}^{n}\xi_i|^p \lesssim n2^{\frac{(p-2)j_0}{2}} + n^{\frac{p}{2}}$. Furthermore, it follows from (4), (3) and $2^{j_0} \le n$ that

$$E|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p \lesssim \frac{1}{n^p}\left[n2^{\frac{(p-2)j_0}{2}} + n^{\frac{p}{2}}\right] \lesssim n^{-\frac{p}{2}}.$$

Combining this with (10), one concludes the desired inequality of the lemma. □

Theorem 2.1 *Let $f(x) \in B^s_{r,q}(\mathbb{R}, L)$ $(s > \frac{1}{r}, r, q \ge 1)$ and let $\hat{f}_n^{\mathrm{lin}}$ be defined by (1). Under the conditions of Lemma 2.2, for each $1 \le p < \infty$, one has*

$$\sup_{f\in B^s_{r,q}(\mathbb{R},L)} \mathbf{E}\left\|\hat{f}_n^{\mathrm{lin}} - f\right\|_p^p \lesssim n^{-\frac{s'p}{2s'+1}},$$

where $s' := s - (\frac{1}{r} - \frac{1}{p})_+$ and $x_+ = \max(x, 0)$.

Proof Since

$$\mathbf{E}\left\|\hat{f}_n^{\mathrm{lin}} - f\right\|_p^p \lesssim \|P_{j_0}f - f\|_p^p + \mathbf{E}\left\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\right\|_p^p, \tag{11}$$

it is sufficient to estimate $\|P_{j_0}f - f\|_p^p$ and $\mathbf{E}\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\|_p^p$.

When $r \le p$, $s' = s - (\frac{1}{r} - \frac{1}{p})_+ = s - \frac{1}{r} + \frac{1}{p}$ and $B^s_{r,q}(\mathbb{R}) \subset B^{s'}_{p,q}(\mathbb{R})$, one has

$$\sup_{f\in B^s_{r,q}(\mathbb{R},L)} \|P_{j_0}f - f\|_p^p \lesssim \sup_{f\in B^{s'}_{p,q}(\mathbb{R},L)} \|P_{j_0}f - f\|_p^p.$$

By the approximation theorem in Besov spaces and from Theorem 9.4 in [9], one gets

$$\|P_{j_0}f - f\|_p^p \lesssim 2^{-j_0 s'p}. \tag{12}$$

When $r > p$, because both f and ϕ have compact supports, one can assume that $\operatorname{supp}(P_{j_0}f - f) \subseteq I$ with $|I| \leq H$. Then Hölder inequality shows

$$\|P_{j_0}f - f\|_p^p = \int_I \left|P_{j_0}f(y) - f(y)\right|^p dy \lesssim \|P_{j_0}f - f\|_r^p.$$

Since $f \in B^s_{r,q}(\mathbb{R}, L)$, one knows $\|P_{j_0}f - f\|_r \lesssim 2^{-j_0 s}$. Moreover, $\|P_{j_0}f - f\|_p^p \lesssim 2^{-j_0 sp}$. Note that $s' = s$ for $r > p$. Then $\|P_{j_0}f - f\|_p^p \lesssim 2^{-j_0 s' p}$. This, together with (12), shows that for $1 \leq p < \infty$,

$$\|P_{j_0}f - f\|_p^p \lesssim 2^{-j_0 s' p}. \tag{13}$$

Next, one estimates $\mathbf{E}\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\|_p^p$. It is easy to see that

$$\hat{f}_n^{\mathrm{lin}} - P_{j_0}f = \sum_{k \in K} (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k})\phi_{j_0,k}$$

by the definitions of $\hat{f}_n^{\mathrm{lin}}$ and $P_{j_0}f$. Furthermore,

$$\left\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\right\|_p^p \lesssim 2^{j_0 p(\frac{1}{2} - \frac{1}{p})} \sum_{k \in K} |\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p$$

due to Theorem 1.2. Let $|K_0|$ denote the number of elements in K_0. Then $|K_0| \sim 2^{j_0}$, because $K_0 := \{k \in Z, \operatorname{supp} f \cap \operatorname{supp} \phi_{j_0,k} \neq \emptyset\}$ and f, ϕ have compact supports. This, together with Lemma 2.2, leads to

$$\mathbf{E}\left\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\right\|_p^p \lesssim 2^{\frac{j_0 p}{2}} \mathbf{E}|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p \lesssim \left(\frac{2^{j_0}}{n}\right)^{\frac{p}{2}}. \tag{14}$$

Substituting (13) and (14) into (11), one obtains

$$\mathbf{E}\left\|\hat{f}_n^{\mathrm{lin}} - f\right\|_p^p \lesssim \left(\frac{2^{j_0}}{n}\right)^{\frac{p}{2}} + 2^{-j_0 s' p}.$$

Taking $2^{j_0} \sim n^{\frac{1}{2s'+1}}$, the desired conclusion follows. □

3 Nonlinear estimators

In this part, we will give a nonlinear wavelet estimator for $f(x)$, which is better than the linear one in some cases. The nonlinear (hard thresholding) wavelet estimator is defined as follows:

$$\hat{f}_n^{\mathrm{non}}(y) := \sum_{k \in K_0} \hat{\alpha}_{j_0,k}\phi_{j_0,k}(y) + \sum_{j=j_0}^{j_1} \sum_{k \in K_j} \hat{\beta}^*_{j,k}\psi_{j,k}(y). \tag{15}$$

Here $K_0 = \{k \in \mathbb{Z}, \operatorname{supp} f \cap \operatorname{supp} \phi_{j_0,k} \neq \emptyset\}$, $K_j = \{k \in \mathbb{Z}, \operatorname{supp} f \cap \operatorname{supp} \psi_{j,k} \neq \emptyset\}$,

$$\hat{\alpha}_{j_0,k} = \frac{1}{n}\sum_{i=1}^n \phi_{j_0,k}(X_i) \quad \text{and} \quad \hat{\beta}_{j,k} = \frac{1}{n}\sum_{i=1}^n \psi_{j,k}(X_i) \tag{16}$$

with $\hat{\beta}^*_{j,k} = \hat{\beta}_{j,k}\mathcal{X}\{|\hat{\beta}_{j,k}| > \lambda = c\sqrt{\frac{j}{n}}\}$ while the constant c is determined (later on) by s, r, p and L.

For the wavelet coefficients, we can get the following lemma whose proof is very similar to that of Lemma 2.2 and so we omit it.

Lemma 3.1 *Let $\hat{\beta}_{j,k}$ be defined by* (16). *Then under the assumptions of Lemma* 2.2,

$$\mathbf{E}|\hat{\beta}_{j,k} - \beta_{j,k}|^p \lesssim n^{-\frac{p}{2}}$$

for $1 \le p < \infty$ *and* $2^j \le n$.

To prove Lemma 3.3, we need an important inequality.

Lemma 3.2 (Bernstein's inequality) *Let $X_1, \ldots, X_n$ be a sequence of ND random variables such that* $\mathbf{E}(X_i) = 0$, $\mathbf{E}(X_i^2) = \sigma^2$ *and* $|X_i| \le M < \infty$ $(i = 1, \ldots, n)$. *Then for each* $v > 0$,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n X_i\right| > v\right) \le 2\exp\left(-\frac{nv^2}{2(\sigma^2 + \frac{vM}{3})}\right).$$

This above inequality is well-known, when $X_1, \ldots, X_n$ are independent; see Theorem C.1 on page 241 in [9]. We find by checking the details that the same inequality holds for ND samples: In fact, because Theorem C.1 is a direct corollary of Lemma C.1 (page 239), it suffices to prove that lemma for the ND case. Note that

$$\mathbf{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] = \prod_{i=1}^n \mathbf{E}\left[\exp(tX_i)\right]$$

for an independent sample $X_1, \ldots, X_n$, while

$$\mathbf{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] \le \prod_{i=1}^n \mathbf{E}\left[\exp(tX_i)\right]$$

for ND samples, according to Theorem 1.3. Then we only need to replace the equality

$$\exp(-\lambda t)\mathbf{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] = \exp\left\{-\left[\lambda t - \sum_{i=1}^n \log \mathbf{E}\left(e^{tX_i}\right)\right]\right\}$$

by

$$\exp(-\lambda t)\mathbf{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] \le \exp\left\{-\left[\lambda t - \sum_{i=1}^n \log \mathbf{E}\left(e^{tX_i}\right)\right]\right\}$$

on page 240 (line 8–9), in order to complete the proof of Lemma C.1, when $X_1, \ldots, X_n$ are ND.

Lemma 3.3 *Let $\hat{\beta}_{j,k}$ be given by* (16). *Under the assumptions of Lemma* 2.2 *and if $j2^j \leq n$, then for each $\omega > 0$, there exists $c > 0$ such that*

$$\mathbb{P}\left(|\hat{\beta}_{j,k} - \beta_{j,k}| > \lambda = c\sqrt{\frac{j}{n}}\right) \lesssim 2^{-\omega j}.$$

Proof It is easy to see that

$$|\hat{\beta}_{j,k} - \beta_{j,k}| = \frac{1}{n}\left|\sum_{i=1}^{n}\left[\psi_{j,k}(X_i) - \beta_{j,k}\right]\right|.$$

Hence,

$$I := \mathbb{P}\left(|\hat{\beta}_{j,k} - \beta_{j,k}| > \lambda\right) = \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\left[\psi_{j,k}(X_i) - \beta_{j,k}\right]\right| > \lambda\right).$$

In order to estimate I, denote $\eta_i := \psi_{j,k}(X_i) - \beta_{j,k}$ $(i = 1, 2, \ldots, n)$. Then

$$I = \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\eta_i\right| > \lambda\right).$$

Since ψ is a function of BV, $\psi := \tilde{\psi} - \bar{\psi}$, where $\tilde{\psi}$ and $\bar{\psi}$ are bounded, nonnegative and nondecreasing functions. Denote

$$\tilde{\beta}_{j,k} := \int \tilde{\psi}_{j,k}(x)f(x)\,dx, \qquad \bar{\beta}_{j,k} := \int \bar{\psi}_{j,k}(x)f(x)\,dx,$$

and

$$\tilde{\eta}_i := \tilde{\psi}_{j,k}(X_i) - \tilde{\beta}_{j,k}, \qquad \bar{\eta}_i := \bar{\psi}_{j,k}(X_i) - \bar{\beta}_{j,k}.$$

Then $\beta_{j,k} = \tilde{\beta}_{j,k} - \bar{\beta}_{j,k}$, $\eta_i = \tilde{\eta}_i - \bar{\eta}_i$ and

$$\begin{aligned} I &= \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}(\tilde{\eta}_i - \bar{\eta}_i)\right| > \lambda\right) \\ &\leq \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\tilde{\eta}_i\right| > \frac{\lambda}{2}\right) + \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\bar{\eta}_i\right| > \frac{\lambda}{2}\right). \end{aligned} \tag{17}$$

Note that $\tilde{\eta}_1, \ldots, \tilde{\eta}_n$ are ND thanks to the monotonicity of $\tilde{\psi}$ and Theorem 1.3. On the other hand, $\mathbf{E}\tilde{\eta}_i = 0$, $\mathbf{E}(\tilde{\eta}_i)^2 \lesssim 1$ and $|\tilde{\eta}_i| \lesssim 2^{\frac{j}{2}}$. Using Bernstein's inequality, one obtains that

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\tilde{\eta}_i\right| > \frac{\lambda}{2} = \frac{c}{2}\sqrt{\frac{j}{n}}\right) \leq 2\exp\left(-\frac{c^2 j}{C(1 + c\sqrt{\frac{j2^j}{n}})}\right)$$

for some fixed constant $C > 0$. Due to $j2^j \leq n$, one can take $c > 0$ such that $\frac{c^2}{C(1+c)} \geq \omega$ and

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\tilde{\eta}_i\right| > \frac{\lambda}{2}\right) \lesssim 2^{-\omega j}. \tag{18}$$

Similarly, $\mathbb{P}(\frac{1}{n}|\sum_{i=1}^{n}\bar{\eta}_i| > \frac{\lambda}{2}) \lesssim 2^{-\omega j}$. This, with (18) and (17), leads to

$$I = \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\eta_i\right| > \lambda\right) \lesssim 2^{-\omega j}.$$

The desired conclusion follows. □

Theorem 3.1 *Let $f(x) \in B^s_{r,q}(\mathbb{R},L)$ $(s > \frac{1}{r}, r,q \geq 1)$, and let $\hat{f}_n^{\mathrm{non}}$ be defined by (15). Under the assumptions of Lemma 2.2, for each $1 \leq p < \infty$, $s' := s - (\frac{1}{r} - \frac{1}{p})_+$ and $x_+ = \max(x,0)$, there exist $\theta_i \in \mathbb{R}$ $(i = 1,2,3)$ such that*

$$\sup_{f\in B^s_{r,q}(\mathbb{R},L)} \mathbf{E}\left\|\hat{f}_n^{\mathrm{non}} - f\right\|_p^p \lesssim \begin{cases} (\ln n)^{\theta_1} n^{-\frac{sp}{2s+1}}, & \frac{p}{2s+1} < r < p, \\ (\ln n)^{\theta_2} (\frac{\ln n}{n})^{\frac{s'p}{2(s-1/r)+1}}, & r = \frac{p}{2s+1}, \\ (\ln n)^{\theta_3} (\frac{\ln n}{n})^{\frac{s'p}{2(s-1/r)+1}}, & r < \frac{p}{2s+1}. \end{cases} \tag{19}$$

Proof Clearly,

$$\hat{f}_n^{\mathrm{non}} - f = \left(\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\right) + (P_{j_1+1}f - f) + \sum_{j=j_0}^{j_1}\sum_{k\in K_j}\left(\hat{\beta}^*_{jk} - \beta_{jk}\right)\psi_{jk}.$$

Then

$$\mathbf{E}\left\|\hat{f}_N^{\mathrm{non}} - f^X\right\|_p^p \lesssim T_1 + T_2 + T_3, \tag{20}$$

where $T_1 := \mathbf{E}\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f\|_p^p$, $T_2 := \|P_{j_1+1}f - f\|_p^p$ and $T_3 := \mathbf{E}\|\sum_{j=j_0}^{j_1}\sum_{k\in K_j}(\hat{\beta}^*_{jk} - \beta_{jk})\psi_{jk}\|_p^p$. By (13) and (14),

$$T_1 \lesssim \left(\frac{2^{j_0}}{n}\right)^{\frac{p}{2}} \quad \text{and} \quad T_2 \lesssim 2^{-j_1 s' p}. \tag{21}$$

For estimating T_3, one uses Minkowski and Jensen's inequalities to get

$$\left\|\sum_{j=j_0}^{j_1}\sum_{k\in K_j}\left(\hat{\beta}^*_{j,k} - \beta_{j,k}\right)\psi_{j,k}\right\|_p^p \leq (j_1 - j_0 + 1)^{p-1}\sum_{j=j_0}^{j_1}\left\|\sum_{k\in K_j}\left(\hat{\beta}^*_{j,k} - \beta_{j,k}\right)\psi_{j,k}\right\|_p^p.$$

This, together with Theorem 1.2, leads to

$$T_3 \leq (j_1 - j_0 + 1)^{p-1}\mathbf{E}\sum_{j=j_0}^{j_1}2^{j(\frac{p}{2}-1)}\left(\sum_{k\in K_j}\left|\hat{\beta}^*_{j,k} - \beta_{j,k}\right|^p\right).$$

Since $\hat{\beta}^*_{j,k} = \delta^H(\hat{\beta}_{j,k},\lambda)$,

$$\begin{aligned}\left|\hat{\beta}^*_{j,k} - \beta_{j,k}\right|^p &= |\hat{\beta}_{j,k} - \beta_{j,k}|^p[\mathcal{X}_{\{|\hat{\beta}_{j,k}|>\lambda,|\beta_{j,k}|<\frac{\lambda}{2}\}} + \mathcal{X}_{\{|\hat{\beta}_{j,k}|>\lambda,|\beta_{j,k}|\geq\frac{\lambda}{2}\}}] \\ &\quad + |\beta_{j,k}|^p[\mathcal{X}_{\{|\hat{\beta}_{j,k}|\leq\lambda,|\beta_{j,k}|>2\lambda\}} + \mathcal{X}_{\{|\hat{\beta}_{j,k}|\leq\lambda,|\beta_{j,k}|\leq2\lambda\}}].\end{aligned}$$

Therefore,

$$T_3 \lesssim (j_1 - j_0 + 1)^{p-1} \Bigg\{ \mathbf{E} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^p [\mathcal{X}_{\{|\hat{\beta}_{j,k}| > \lambda, |\beta_{j,k}| < \frac{\lambda}{2}\}} + \mathcal{X}_{\{|\hat{\beta}_{j,k}| > \lambda, |\beta_{j,k}| \geq \frac{\lambda}{2}\}}] + \mathbf{E} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} |\beta_{j,k}|^p [\mathcal{X}_{\{|\hat{\beta}_{j,k}| \leq \lambda, |\beta_{j,k}| > 2\lambda\}} + \mathcal{X}_{\{|\hat{\beta}_{j,k}| \leq \lambda, |\beta_{j,k}| \leq 2\lambda\}}] \Bigg\}. \tag{22}$$

When $|\hat{\beta}_{jk}| > \lambda$ and $|\beta_{jk}| < \frac{\lambda}{2}$, $|\hat{\beta}_{jk} - \beta_{jk}| \geq |\hat{\beta}_{jk}| - |\beta_{jk}| > \frac{\lambda}{2}$, one has

$$I_{\{|\hat{\beta}_{jk}| > \lambda, |\beta_{jk}| < \frac{\lambda}{2}\}} \leq I_{\{|\hat{\beta}_{jk} - \beta_{jk}| > \frac{\lambda}{2}\}}.$$

Similarly, when $|\hat{\beta}_{jk}| \leq \lambda$ and $|\beta_{jk}| > 2\lambda$, $|\hat{\beta}_{jk}| \leq \lambda < \frac{|\beta_{jk}|}{2}$. Hence,

$$|\hat{\beta}_{jk} - \beta_{jk}| \geq |\beta_{jk}| - |\hat{\beta}_{jk}| > \frac{|\beta_{jk}|}{2} > \lambda \quad \text{and} \quad |\beta_{jk}| < 2|\hat{\beta}_{jk} - \beta_{jk}|.$$

Furthermore,

$$|\beta_{jk}|^p I_{\{|\hat{\beta}_{jk}| \leq \lambda, |\beta_{jk}| > 2\lambda\}} \lesssim |\hat{\beta}_{jk} - \beta_{jk}|^p I_{\{|\hat{\beta}_{jk} - \beta_{jk}| > \frac{\lambda}{2}\}}.$$

Then (22) reduces to

$$T_3 \lesssim T_{31} + T_{32} + T_{33},$$

where

$$T_{31} := (j_1 - j_0 + 1)^{p-1} \mathbf{E} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^p \mathcal{X}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda}{2}\}},$$

$$T_{32} := (j_1 - j_0 + 1)^{p-1} \mathbf{E} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} |\hat{\beta}_{j,k} - \beta_{j,k}|^p \mathcal{X}_{\{|\beta_{j,k}| \geq \frac{\lambda}{2}\}}$$

and $T_{33} := (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} |\beta_{j,k}|^p \mathcal{X}_{\{|\beta_{j,k}| \leq 2\lambda\}}$.

In order to estimate T_{31}, first one assumes $\frac{1}{q} + \frac{1}{q'} = 1$. Then Jensen's inequality shows that

$$\begin{aligned} T_{31} &\leq (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} \big[\mathbf{E}|\hat{\beta}_{j,k} - \beta_{j,k}|^{qp}\big]^{\frac{1}{q}} \big[\mathbf{E}(\mathcal{X}_{\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda}{2}\}})^{q'}\big]^{\frac{1}{q'}} \\ &\leq (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k \in K_j} \big(\mathbf{E}|\hat{\beta}_{j,k} - \beta_{j,k}|^{qp}\big)^{\frac{1}{q}} \left[P\left(|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{\lambda}{2}\right)\right]^{\frac{1}{q'}}. \end{aligned}$$

This, together with Lemmas 3.1 and 3.3, leads to

$$\begin{aligned} T_{31} &\lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)}2^j n^{-\frac{p}{2}}2^{-\frac{\omega j}{q'}} = (j_1-j_0+1)^{p-1}n^{-\frac{p}{2}}\sum_{j=j_0}^{j_1}2^{j(\frac{p}{2}-\frac{\omega}{q'})} \\ &\lesssim (j_1-j_0+1)^{p-1}n^{-\frac{p}{2}}2^{j_0(\frac{p}{2}-\frac{\omega}{q'})} \le (j_1-j_0+1)^{p-1}n^{-\frac{p}{2}}2^{\frac{j_0p}{2}} \end{aligned} \tag{23}$$

by choosing ω such that $\frac{p}{2} < \frac{\omega}{q'}$.

It is easy to see that $\|\beta_{j\cdot}\|_r \lesssim 2^{-j(s+\frac{1}{2}-\frac{1}{r})}$ thanks to Theorem 1.1. Combining this with Lemma 3.1 and $\mathcal{X}_{\{|\beta_{j,k}|\ge\frac{\lambda}{2}\}} \le (\frac{|\beta_{j,k}|}{\frac{\lambda}{2}})^r$, one has

$$\begin{aligned} T_{32} &\lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0}^{j_1}2^{j(\frac{p}{2}-1)}\sum_{k\in K_j}n^{-\frac{p}{2}}\left|\frac{\beta_{j,k}}{\frac{\lambda}{2}}\right|^r \\ &\lesssim (j_1-j_0+1)^{p-1}n^{-\frac{p}{2}}\sum_{j=j_0}^{j_1}\lambda^{-r}2^{j(\frac{p-r}{2}-rs)}. \end{aligned} \tag{24}$$

Similarly, it can be shown that

$$\begin{aligned} T_{33} &\le (j_1-j_0+1)^{p-1}\sum_{j=j_0}^{j_1}2^{j(\frac{p}{2}-1)}\sum_{k\in K_j}|\beta_{j,k}|^p\left(\frac{2\lambda}{|\beta_{j,k}|}\right)^{p-r} \\ &\lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0}^{j_1}\lambda^{p-r}2^{j(\frac{p-r}{2}-rs)} \end{aligned} \tag{25}$$

due to $r<p$ and $\mathcal{X}_{\{|\beta_{j,k}|\le 2\lambda\}} \le (\frac{2\lambda}{|\beta_{j,k}|})^{p-r}$.

Take

$$2^{j_0} \sim \begin{cases} [(\ln n)^{\frac{p-r}{r}}n]^{\frac{1}{2s+1}}, & r > \frac{p}{2s+1}, \\ n^{\frac{1-2/p}{2(s-1/r)+1}}, & r \le \frac{p}{2s+1}, \end{cases} \quad \text{and} \quad 2^{j_1} \sim \begin{cases} n^{\frac{s}{s'(2s+1)}}, & r > \frac{p}{2s+1}, \\ (n/\ln n)^{\frac{1}{2(s-1/r)+1}}, & r \le \frac{p}{2s+1}. \end{cases} \tag{26}$$

Then $j_0 < j_1$, $j_1 - j_0 \sim \ln n$ and for $j_0 \le j \le j_1$, $\lambda := c\sqrt{\frac{j}{n}} \sim c\sqrt{\frac{\ln n}{n}}$. Moreover, (24) and (25) reduce to

$$T_{32} \lesssim (j_1-j_0+1)^{p-1}n^{\frac{r-p}{2}}(\ln n)^{-\frac{r}{2}}\left[2^{j_0\xi}\mathcal{X}_{\{\xi<0\}} + (j_1-j_0+1)\mathcal{X}_{\{\xi=0\}} + 2^{j_1\xi}\mathcal{X}_{\{\xi>0\}}\right] \tag{27}$$

and

$$T_{33} \lesssim (j_1-j_0+1)^{p-1}\left(\frac{\ln n}{n}\right)^{\frac{p-r}{2}}\left[2^{j_0\xi}\mathcal{X}_{\{\xi<0\}} + (j_1-j_0+1)\mathcal{X}_{\{\xi=0\}} + 2^{j_1\xi}\mathcal{X}_{\{\xi>0\}}\right], \tag{28}$$

where $\xi = \frac{p-r}{2} - rs$.

Note that $\xi \geq 0$ holds if and only if $r \leq \frac{p}{2s+1}$. Then substituting (26) into (23), (27) and (28), one obtains

$$T_3 \lesssim T_{31} + T_{32} + T_{33} \lesssim \begin{cases} (\ln n)^{\theta_1} n^{-\frac{sp}{2s+1}}, & \frac{p}{2s+1} < r < p, \\ (\ln n)^{\theta_2} (\frac{\ln n}{n})^{\frac{s'p}{2(s-1/r)+1}}, & r = \frac{p}{2s+1}, \\ (\ln n)^{\theta_3} (\frac{\ln n}{n})^{\frac{s'p}{2(s-1/r)+1}}, & r < \frac{p}{2s+1}. \end{cases} \tag{29}$$

Similarly, it is easy to check that

$$T_1 + T_2 \lesssim \begin{cases} n^{-\frac{sp}{2s+1}}, & \frac{p}{2s+1} < r < p, \\ (\frac{\ln n}{n})^{\frac{s'p}{2(s-1/r)+1}}, & r \leq \frac{p}{2s+1} \end{cases} \tag{30}$$

by (26) and (21). Finally, the desired conclusion (19) follows from (20), (29), and (30). □

Remark 3.1 From Theorems 2.1 and 3.1, we easily found that our results are consistent with those in [6] for independent samples.

Remark 3.2 In [7], Doosti and Chaubey provided a convergence rate of $n^{-\frac{s'p}{2s'+1}}$ for ND samples, which is a little weaker than $n^{-\frac{sp}{2s+1}}$ in Theorem 3.1 for $r < p$ (note that $s < s'$ when $r < p$).

Acknowledgements
The author would like to thank three referees and the editor for their important comments and suggestions, which substantially improved the manuscript.

Funding
This work is supported by the National Natural Science Foundation of China (No. 11626034, 61602010), and the Projection of Baoji University of Arts and Sciences (No. ZK16051).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
JX completed this work. The author read and approved the final manuscript.

Publisher's Note

Received: 7 December 2018 Accepted: 5 June 2019 Published online: 14 June 2019

References

1. Asadian, N., Fakoor, V., Bozorgnia, A.: Rosenthal's type inequalities for negatively orthant dependent random variables. J. Iran. Stat. Soc. **5**, 69–75 (2006)
2. Block, H.W., Savits, T.H., Shaked, M.: Some concepts of negative dependence. Ann. Probab. **10**, 765–772 (1982)
3. Bozorgnia, A., Patterson, R.F., Taylor, R.L.: Limit Theorems for ND. University of Georgia, Athens (1993)
4. Chesneau, C., Dewan, I., Doosti, H.: Wavelet linear density estimation for associated stratified size-biased sample. J. Nonparametr. Stat. **2**, 429–445 (2012)
5. Daubechies, I.: Ten Lectures on Wavelets. SIAM, Philadelphia (1992)
6. Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D.: Density estimation by wavelet thresholding. Ann. Stat. **2**, 508–539 (1996)
7. Doosti, H., Chaubey, Y.P.: Wavelet linear density estimation for negatively dependent random variables. Curr. Dev. Theory Appl. Wavelets **1**(1), 57–64 (2007)
8. Doosti, H., Fakoor, V., Chaubey, Y.P.: Wavelet linear density estimation for negative associated sequences. J. Indian Stat. Assoc. **44**, 127–136 (2006)
9. Härdle, W., Kerkyacharian, G., Picard, D., Tsybakov, A.: Wavelets, Approximations, and Statistical Applications. Lecture Notes in Statistics. Springer, Berlin (1998)
10. Joag-dev, K., Proschan, F.: Negative association of random variables with application. Ann. Stat. **11**, 286–295 (1983)

11. Kou, J.K., Guo, H.J.: Wavelet density estimation for mixing and size-biased data. J. Inequal. Appl. **2018**, 189 (2018)
12. Liu, Y.M., Xu, J.L.: Wavelet density estimation for negatively associated stratified size-biased sample. J. Nonparametr. Stat. **26**, 537–554 (2014)
13. Newman, C.M.: Asymptotic independence and limit theorems for positively and negatively dependent random variables. Inequa. Stat. Probab. **5**, 127–140 (1984)
14. Sung, S.H.: A note on the complete convergence for weighted sums of negatively dependent random variables. J. Inequal. Appl. **2012**, 158 (2012)
15. Wu, Q.Y.: Complete convergence for negatively dependent sequences of random variables. J. Inequal. Appl. **2010**, Article ID 507293 (2010)
16. Zhang, L.X.: Rosenthal's inequalities for independent and negatively dependent random variables under sub-linear expectations with applications. Sci. China Math. **59**(4), 751–768 (2016)