

RESEARCH

Open Access



# The Berry-Esseen bounds of wavelet estimator for regression model whose errors form a linear process with a $\rho$ -mixing

Liwang Ding<sup>1\*</sup> and Yongming Li<sup>2</sup>

\*Correspondence:  
2008dingliwang@163.com  
<sup>1</sup>School of Information and Statistics, Guangxi University of Finance and Economics, Nanning, 530003, P.R. China  
Full list of author information is available at the end of the article

## Abstract

In this paper, considering the nonparametric regression model  $Y_{ni} = g(t_i) + \varepsilon_i$  ( $1 \leq i \leq n$ ), where  $\varepsilon_i = \sum_{j=-\infty}^{\infty} a_j e_{i-j}$  and  $e_{i-j}$  are identically distributed and  $\rho$ -mixing sequences. This paper obtains the Berry-Esseen bounds of the wavelet estimator of  $g(\cdot)$ , the rates of the normal approximation are shown as  $O(n^{-1/6})$  under certain conditions.

**MSC:** 60G20; 60F05

**Keywords:** wavelet estimator;  $\rho$ -mixing; Berry-Esseen bound; linear process

## 1 Introduction

The Berry-Esseen theorem of probability distribution concerns mainly research of statistics convergence to a certain distribution and the measure of the probability distributions of the statistics which determines the distribution as regards the absolute distance that can be controlled as an optimal problem.

In recent years, the Berry-Esseen bounds theorem has got extensive investigation. For instance, Xue [1] discussed the Berry-Esseen bound of an estimator for the variance in a semi-parametric regression model under some mild conditions, Liang and Li [2] studied the asymptotic normality and the Berry-Esseen type bound of the estimator with linear process error, Li *et al.* [3] derived the Berry-Esseen bounds of the wavelet estimator for a nonparametric regression model with linear process errors generated by  $\varphi$ -mixing sequences, Li *et al.* [4] investigated the Berry-Esseen bounds of the wavelet estimator in a semi-parametric regression model with linear process errors.

To investigate the estimation of the fixed design nonparametric regression model involves a regression function  $g(\cdot)$  which is defined on  $[0, 1]$ :

$$Y_i = g(t_i) + \varepsilon_i \quad (1 \leq i \leq n), \quad (1.1)$$

where  $\{t_i\}$  are known fixed design points, we assume  $\{t_i\}$  and to be ordered  $0 \leq t_1 \leq \dots \leq t_n \leq 1$ , and  $\{\varepsilon_i\}$  are random errors.

It is well known that a regression function estimation is an important method in data analysis and has a wide range of applications in filtering and prediction in communication and control systems, pattern recognition and classification, and econometrics. So the

model (1.1) has been studied widely. The model (1.1) has been applied in solving many practical problems and kinds of estimation methods have been used to obtain estimators of  $g(\cdot)$ .

For model (1.1), the following wavelet estimator of  $g(\cdot)$  will be considered:

$$g_n(t) = \sum_{i=1}^n Y_i \int_{A_i} E_m(t, s) ds. \tag{1.2}$$

The wavelet kernel  $E_m(t, s)$  can be considered as follows:  $E_m(t, s) = 2^m E_0(2^m t, 2^m s) = 2^m \sum_{k \in \mathbb{Z}} \phi(2^m t - k) \phi(2^m s - k)$ , where  $\phi(\cdot)$  is a scaling function, the smooth parameter  $m = m(n) > 0$  depending only on  $n$ , and  $A_i = [s_{i-1}, s_i]$  being a partition of interval  $[0, 1]$ ,  $s_i = (1/2)(t_i + t_{i+1})$ , and  $t_i \in A_i, 1 \leq i \leq n$ .

As we know wavelets have been used widely in many engineering and technological fields, especially in picture handling by computers. Since the 1990s, in order to meet practical demands, some authors began to consider using wavelet methods in statistics.

**Definition 1.1** Let  $\{X_i : i = 1, 2, \dots\}$  be a sequence of random variables. Write the  $\rho$ -mixing coefficient

$$\rho(n) = \sup_{k \in \mathbb{N}} \sup_{X \in L_2(\mathfrak{F}_1^k), Y \in L_2(\mathfrak{F}_{k+n}^\infty)} \frac{|E(X - EX)(Y - EY)|}{\sqrt{\text{Var } X \text{Var } Y}},$$

where  $\mathfrak{F}_a^b := \sigma\{X_i : a \leq i \leq b\}$ ,  $L_2(\mathfrak{F}_a^b)$  is the set of square integrable random variables on the condition of  $\mathfrak{F}_a^b$ .

**Definition 1.2** A sequence of random variables  $\{X_i : i = 1, 2, \dots\}$  is  $\rho$ -mixing if  $\rho(n) \rightarrow 0, n \rightarrow \infty$ .

Kolmogorov and Rozanov [5] put forward a  $\rho$ -mixing random variables sequence. For the wide use of  $\rho$ -mixing in science and technology and economics, many scholars have investigated the  $\rho$ -mixing and got fruitful meaningful results. For instance, the central limit theorem for  $\rho$ -mixing, the law of large numbers for  $\rho$ -mixing, the strong invariant principle and weak invariant principle for  $\rho$ -mixing, the complete convergence theorem of  $\rho$ -mixing, which we can see in Shao's work [6, 7]. Recently, Jiang [8] discussed the convergence rates in the law of the logarithm of the  $\rho$ -mixing random variables, obtaining a sufficient condition for the law of logarithm of  $\rho$ -mixing and the convergence rates in the law of the iterated logarithm. Chen and Liu [9] achieved the sufficient and necessary conditions of complete moment convergence for a sequence of identically distributed  $\rho$ -mixing random variables. Zhou and Lin [10] investigated the estimation problems of partially linear models for longitudinal data with  $\rho$ -mixing error structures, and they studied the strong consistency for the least squares estimator of the parametric component; the strong consistency and uniform consistency for the estimator of nonparametric function were studied under some mild conditions. Tan and Wang [11] studied the complete convergence for weighted sums of non-identically distributed  $\rho$ -mixing random variables sequence, and they gave the Marcinkiewicz-Zygmund type strong law of large numbers.

## 2 Assumptions and main results

First, we give some basic assumptions as follows:

- (A1)  $\{\varepsilon_j\}_{j \in \mathbb{Z}}$  has a linear representation  $\varepsilon_j = \sum_{k=-\infty}^{\infty} a_k e_{k-j}$ , where  $\{a_k\}$  is a sequence of real numbers with  $\sum_{k=-\infty}^{\infty} |a_k| < \infty$ ,  $\{e_j\}$  are identically distributed,  $\rho$ -mixing random variables with  $Ee_j = 0$ ,  $E|e_j|^r < \infty$  for some  $r > 2$ , and  $\rho(n) = O(n^{-\lambda})$  for  $\lambda > 2$ .
- (A2) The spectral density function  $f(\omega)$  of  $\{\varepsilon_i\}$  satisfies  $0 < c_1 \leq f(\omega) \leq c_2 < \infty$  for all  $\omega \in (-\pi, \pi]$ .
- (A3) (i)  $\phi(\cdot)$  is said to be  $\sigma$ -regular ( $\phi \in S_\sigma$ ,  $\sigma \in \mathbb{N}$ ) if for any  $\kappa \leq \sigma$  and any integer  $t$ , one has  $|d_\phi^x/dx^\kappa| \leq C_t(1 + |x|)^{-1}$ , the  $C_t$  depending only on  $t$ ;  
 (ii)  $\phi(\cdot)$  satisfies the Lipschitz condition with order 1 and  $|\hat{\phi}(\xi) - 1| = O(\xi)$  as  $\xi \rightarrow \infty$ , where  $\hat{\phi}$  is the Fourier transform of  $\phi$ .
- (A4) (i)  $g(\cdot)$  satisfies the Lipschitz condition of order 1;  
 (ii)  $g(\cdot) \in H^\mu$ ,  $\mu > 1/2$ , A function space  $H^\mu$  ( $\mu \in \mathbb{R}$ ) is said to be Sobolev space of order  $V$ , i.e., if  $h \in H^\mu$  then  $\int |\hat{h}(w)|^2(1 + w^2)^\mu dw < \infty$ , where  $\hat{h}$  is the Fourier transform of  $h$ .
- (A5)  $\max_{1 \leq i \leq n} |s_i - s_{i-1} - n^{-1}| = o(n^{-1})$ .
- (A6) Set  $p := p(n)$  and  $q := q(n)$ , write  $k := [3n/(p + q)]$  such that for  $p + q \leq 3n$ ,  $qp^{-1} \rightarrow 0$ , and  $\zeta_{in} \rightarrow 0$ ,  $i = 1, 2, 3, 4$ , where  $\zeta_{1n} = qp^{-1}2^m$ ,  $\zeta_{2n} = p^{2m}/n$ ,  $\zeta_{3n} = n(\sum_{|j|>n} |a_j|)^2$ ,  $\zeta_{4n} = k\rho(q)$ .

**Remark 2.1** (A1) are the general conditions of the  $\rho$ -mixing sequence, such as Shao [6, 7], (A1) is weaker than Li *et al.*'s [3], (A2)-(A5) are mild regularity conditions for the wavelet estimate in the recent literature, such as Li *et al.* [3, 4, 12], Liang and Qi [13]. In (A6),  $p, q, 2^m$  can be defined as increasing sequences, and  $\zeta_{in} \rightarrow 0$ ,  $i = 1, 2, 3, 4$ , are easily satisfied, if  $p, q$  and  $2^m$  are chosen reasonable. See e.g. Liang and Li [2], Li *et al.* [3, 4, 12].

In order to facilitate the discussion, write  $\sigma_n^2 := \sigma_n^2(t) = \text{Var}(\hat{g}_n(t))$ ,  $S_n := S_n(t) = \sigma_n^{-1}\{\hat{g}_n(t) - E\hat{g}_n(t)\}$ ,  $u(n) = \sum_{j=1}^{\infty} \rho(j)$ ,  $\|X\|_\beta = (E|X|^\beta)^{1/\beta}$ ,  $a \wedge b = \min\{a, b\}$ . Next, we give the main results as follows.

**Theorem 2.1** *Suppose that (A1)-(A6) hold, then for each  $t \in [0, 1]$ , we can get*

$$\sup_u |P(S_n(t) \leq u) - \Phi(u)| \leq C\{\zeta_{1n}^{1/3} + \zeta_{2n}^{1/3} + \zeta_{2n}^{\delta/2} + \zeta_{3n}^{1/3} + \zeta_{4n}^{1/4} + u(q)\}.$$

**Corollary 2.1** *Suppose that (A1)-(A6) hold, then for each  $t \in [0, 1]$ , we can get*

$$\sup_u |P(S_n(t) \leq u) - \Phi(u)| = o(1).$$

**Corollary 2.2** *Suppose that (A1)-(A6) hold, assume that*

$$\frac{2^m}{n} = O(n^{-\theta}) \quad \text{and} \quad \sup_{n \geq 1} \left( n^{\frac{6\lambda\theta + 3\lambda + 3\theta + 4}{2(6\lambda + 7)}} \right) \sum_{|j|>n} |a_j| < \infty$$

*for some  $\frac{2}{9-6\lambda} < \theta \leq 1$  and for some  $\lambda > 2$ , we can get*

$$\sup_u |P(S_n(t) \leq u) - \Phi(u)| \leq C\{\zeta_{1n}^{1/3} + \zeta_{2n}^{1/3} + \zeta_{3n}^{1/3} + \zeta_{4n}^{1/4} + u(q)\},$$

$$\sup_u |P(S_n(t) \leq u) - \Phi(u)| = O\left(n^{-\frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}}\right).$$

Observe, taking  $\theta \approx 1$  as  $\lambda \rightarrow \infty$ , that it follows that  $\sup_u |P(S_n(t) \leq u) - \Phi(u)| = O(n^{-1/6})$ .

### 3 Some lemmas

From (1.2), we can see that

$$\begin{aligned} S_n &= \sigma_n^{-1} \sum_{i=1}^n \varepsilon_i \int_{A_i} E_m(t, s) ds \\ &= \sigma_n^{-1} \sum_{i=1}^n \int_{A_i} E_m(t, s) ds \sum_{j=-n}^n a_j e_{i-j} \\ &\quad + \sigma_n^{-1} \sum_{i=1}^n \int_{A_i} E_m(t, s) ds \sum_{|j|>n} a_j e_{i-j} \\ &:= S_{1n} + S_{2n}. \end{aligned}$$

Write

$$S_{1n} = \sum_{l=1-n}^{2n} \sigma_n^{-1} \left( \sum_{i=\max\{1, l-n\}}^{\min\{n, l+n\}} a_{i-l} \int_{A_i} E_m(t, s) ds \right) e_l := \sum_{l=1-n}^{2n} W_{nl},$$

set  $S_{1n} = S'_{1n} + S''_{1n} + S'''_{1n}$ , where  $S'_{1n} = \sum_{v=1}^k y_{nv}$ ,  $S''_{1n} = \sum_{w=1}^k y'_{nw}$ ,  $S'''_{1n} = y'_{nk+1}$ ,

$$y_{nv} = \sum_{i=k_v}^{k_v+p-1} W_{ni}, \quad y'_{nw} = \sum_{i=l_w}^{l_w+q-1} W_{ni}, \quad y'_{nk+1} = \sum_{i=k(p+q)-n+1}^{2n} W_{ni},$$

$$k_v = (v-1)(p+q) + 1 - n, \quad l_w = (w-1)(p+q) + p + 1 - n, \quad w = 1, \dots, k,$$

then

$$S_n = S'_{1n} + S''_{1n} + S'''_{1n} + S_{2n}.$$

Next, we give the main lemmas as follows.

**Lemma 3.1** *Let  $\{X_i : i = 1, 2, \dots\}$  be a  $\rho$ -mixing sequence,  $p_1, p_2$  are two integers, let  $\eta_l := \sum_{\sum_{i=1}^{(l-1)(p_1+p_2)+p_1} X_i}$ , for  $1 \leq l \leq k$ . If  $r > 1, s > 1$ , and  $1/r + 1/s = 1$ , then*

$$\left| E \exp\left(it \sum_{l=1}^k \eta_l\right) - \prod_{l=1}^k E \exp(it \eta_l) \right| \leq C |t| \rho^{1/s}(p_2) \sum_{l=1}^k \|\eta_l\|_r.$$

*Proof of Lemma 3.1* We can easily see that

$$\begin{aligned} I_0 &:= \left| \varphi_{\xi_1, \dots, \xi_m}(t_1, \dots, t_m) - \varphi_{\xi_1}(t_1) \cdots \varphi_{\xi_m}(t_m) \right| \\ &= \left| \varphi_{\xi_1, \dots, \xi_m}(t_1, \dots, t_m) - \varphi_{\xi_1, \dots, \xi_{m-1}}(t_1, \dots, t_{m-1}) \varphi_{\xi_m}(t_m) \right| \\ &\quad + \left| \varphi_{\xi_1, \dots, \xi_{m-1}}(t_1, \dots, t_{m-1}) - \varphi_{\xi_1}(t_1) \cdots \varphi_{\xi_{m-1}}(t_{m-1}) \right| =: I_1 + I_2. \end{aligned} \tag{3.1}$$

As  $\exp(ix) = \cos(x) + i \sin(x)$ ,  $\sin(x + y) = \sin(x) \cos(y) + \cos(x) \sin(y)$ ,  $\cos(x + y) = \cos(x) \cos(y) - \sin(x) \sin(y)$ , we can get

$$\begin{aligned}
 I_1 &= \left| \mathbb{E} \exp\left(i \sum_{l=1}^m t_l \xi_l\right) - \mathbb{E} \exp\left(i \sum_{l=1}^{m-1} t_l \xi_l\right) \mathbb{E} \exp(it \xi_m) \right| \\
 &\leq \left| \text{Cov}\left(\cos\left(\sum_{l=1}^{m-1} t_l \xi_l\right), \cos(t_m \xi_m)\right) \right| + \left| \text{Cov}\left(\sin\left(\sum_{l=1}^{m-1} t_l \xi_l\right), \sin(t_m \xi_m)\right) \right| \\
 &\quad + \left| \text{Cov}\left(\sin\left(\sum_{l=1}^{m-1} t_l \xi_l\right), \cos(t_m \xi_m)\right) \right| + \left| \text{Cov}\left(\cos\left(\sum_{l=1}^{m-1} t_l \xi_l\right), \sin(t_m \xi_m)\right) \right| \\
 &=: I_{11} + I_{12} + I_{13} + I_{14}.
 \end{aligned} \tag{3.2}$$

It follows from Lemma 3.2 and  $|\sin(x)| \leq |x|$  that we have

$$\begin{aligned}
 I_{12} &\leq C \rho^{1/s}(1) \|\sin(t_m \xi_m)\|_r \leq C \rho^{1/s}(1) |t_m| \|\sin(\xi_m)\|_r, \\
 I_{14} &\leq C \rho^{1/s}(1) |t_m| \|\sin(\xi_m)\|_r.
 \end{aligned} \tag{3.3}$$

Notice that  $\cos(2x) = 1 - 2 \sin^2(x)$ . Then one has

$$\begin{aligned}
 I_{11} &= \left| \text{Cov}\left(\cos\left(\sum_{l=1}^{m-1} t_l \xi_l\right), 1 - 2 \sin^2(t_m \xi_m/2)\right) \right| \\
 &= 2 \left| \text{Cov}\left(\cos\left(\sum_{l=1}^{m-1} t_l \xi_l\right), \sin^2(t_m \xi_m/2)\right) \right| \leq C \rho^{1/s}(1) E^{1/s} |\sin(t_m \xi_m/2)|^{2r} \\
 &\leq C \rho^{1/s}(1) E^{1/s} |\sin(t_m \xi_m/2)|^r \leq C \rho^{1/s}(1) |t_m| \|\xi_m\|_r.
 \end{aligned} \tag{3.4}$$

Similarly,

$$I_{13} \leq C \rho^{1/s}(1) |t_m| \|\xi_m\|_r. \tag{3.5}$$

Therefore, we can obtain

$$I_1 \leq C \rho^{1/s}(1) |t_m| \|\xi_m\|_r. \tag{3.6}$$

Thus, as follows from (3.1) and (3.6), we can get

$$I_0 = |\varphi_{\xi_1, \dots, \xi_m}(t_1, \dots, t_m) - \varphi_{\xi_1}(t_1) \cdots \varphi_{\xi_m}(t_m)| \leq C \rho^{1/s}(1) |t_m| \|\xi_m\|_r + I_2. \tag{3.7}$$

For  $I_2$  in (3.7), using the same decomposition as in (3.1) above, it can be found that

$$\begin{aligned}
 I_2 &:= |\varphi_{\xi_1, \dots, \xi_{m-1}}(t_1, \dots, t_m) - \varphi_{\xi_1}(t_1) \cdots \varphi_{\xi_m}(t_m)| \\
 &= |\varphi_{\xi_1, \dots, \xi_{m-1}}(t_1, \dots, t_m) - \varphi_{\xi_1, \dots, \xi_{m-2}}(t_1, \dots, t_{m-2}) \varphi_{\xi_{m-1}}(t_{m-1})| \\
 &\quad + |\varphi_{\xi_1, \dots, \xi_{m-2}}(t_1, \dots, t_{m-2}) - \varphi_{\xi_1}(t_1) \cdots \varphi_{\xi_{m-2}}(t_{m-2})| =: I_3 + I_4,
 \end{aligned} \tag{3.8}$$

and similarly to the proof of  $I_1$ , we can get

$$I_3 \leq C\rho^{1/s}(1)\|t_{m-1}\|\|\xi_{m-1}\|_r.$$

Then, we obtain

$$I_2 \leq C\rho^{1/s}(1)\|t_{m-1}\|\|\xi_{m-1}\|_r + I_4. \tag{3.9}$$

Combining (3.7)-(3.9), it suffices to show Lemma 3.1. □

**Lemma 3.2** *Suppose that (A1)-(A5) hold, then we have*

$$\sigma_n^2(t) \geq C2^m n^{-1} \quad \text{and} \quad \sigma_n^{-2}(t) \left| \int_{A_i} E_m(t, s) ds \right| \leq C.$$

*Proof of Lemma 3.2* From (A1) and (1.2), we can get

$$\begin{aligned} \sigma_n^2 &= \sigma^2 \sum_{i=1}^n \left( \int_{A_i} E_m(t, s) ds \right)^2 + 2 \sum_{1 \leq i < j \leq n} E(\varepsilon_i, \varepsilon_j) \int_{A_i} E_m(t, s) ds \int_{A_j} E_m(t, s) ds \\ &= \sigma^2 \sum_{i=1}^n \left( \int_{A_i} E_m(t, s) ds \right)^2 + I_3. \end{aligned}$$

By Lemma A.1, we obtain

$$\begin{aligned} I_3 &\leq 10 \sum_{1 \leq i < j \leq n} \rho(j-i) \left| \int_{A_i} E_m(t, s) ds \int_{A_j} E_m(t, s) ds \right| \|\varepsilon_i\|_2 \|\varepsilon_j\|_2 \\ &= \sum_{k=1}^{n-1} \rho(k) \sum_{i=1}^{n-k} \left| \int_{A_i} E_m(t, s) ds \int_{A_{k+i}} E_m(t, s) ds \right| \|\varepsilon_i\|_2 \|\varepsilon_{k+i}\|_2 \\ &\leq 5\sigma^2 \sum_{k=1}^{n-1} \rho(k) \sum_{i=1}^{n-k} \left[ \left( \int_{A_i} E_m(t, s) ds \right)^2 + \left( \int_{A_{k+i}} E_m(t, s) ds \right)^2 \right] \\ &\leq 10\sigma^2 \sum_{k=1}^n \rho(k) \sum_{i=1}^n \left( \int_{A_i} E_m(t, s) ds \right)^2. \end{aligned}$$

Therefore, from applying (A1) and Lemma A.5, we obtain

$$\sigma_n^2 \leq \sigma^2 \left( 1 + 20 \sum_{k=1}^n \rho(k) \right) \sum_{i=1}^n \left( \int_{A_i} E_m(t, s) ds \right)^2 = C2^m n^{-1}.$$

In addition, the same result as (7) was deduced, see Liang and Qi [13], and by (A2), (A4), and (A5), we have

$$\sigma_n^2(t) \geq C2^m n^{-1} \quad \text{and} \quad \sigma_n^{-2}(t) \left| \int_{A_i} E_m(t, s) ds \right| \leq C. \tag{□}$$

**Lemma 3.3** *Assume that (A1)-(A6) hold, we can get*

- (1)  $E(S''_{1n})^2 \leq C\zeta_{1n}, E(S'''_{1n})^2 \leq C\zeta_{2n}, E(S_{2n})^2 \leq C\zeta_{3n};$
- (2)  $P(|S''_{1n}| \geq \zeta_{1n}^{1/3}) \leq C\zeta_{1n}^{1/3}, P(|S'''_{1n}| \geq \zeta_{2n}^{1/3}) \leq C\zeta_{2n}^{1/3}, P(|S_{2n}| \geq \zeta_{3n}^{1/3}) \leq C\zeta_{3n}^{1/3}.$

*Proof of Lemma 3.3* Let (A1)-(A6) be satisfied, and applying Lemmas 3.2 and A.3(i) in the Appendix, we can refer to Li *et al.*'s [3] Lemma 3.1 of the proof process. □

**Lemma 3.4** Assume that (A1)-(A6) hold, let  $s_n^2 = \sum_{\nu=1}^k \text{Var}(y_{n\nu})$ , we can get

$$|s_n^2 - 1| \leq C(\zeta_{1n}^{1/2} + \zeta_{2n}^{1/2} + \zeta_{3n}^{1/2} + u(q)).$$

Let  $\{\eta_{n\nu} : \nu = 1, \dots, k\}$  be independent random variables and  $\eta_{n\nu} \stackrel{D}{=} y_{n\nu}, \nu = 1, \dots, k$ . Set  $T_n = \sum_{\nu=1}^k \eta_{n\nu}$ . Then we get the following.

*Proof of Lemma 3.4* Let  $\Delta_n = \sum_{1 \leq i < j \leq k} \text{Cov}(y_{ni}, y_{nj})$ , then  $s_n^2 = E(S'_{1n})^2 - 2\Delta_n$ . By  $E(S_n)^2 = 1$ , Lemma 3.3(1), the  $C_r$ -inequality, and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} E(S'_{1n})^2 &= E[S_n - (S''_{1n} + S'''_{1n} + S_{2n})]^2 = 1 + E(S''_{1n} + S'''_{1n} + S_{2n})^2 - 2E[S_n(S''_{1n} + S'''_{1n} + S_{2n})], \\ E(S''_{1n} + S'''_{1n} + S_{2n})^2 &\leq 2[E(S''_{1n})^2 + E(S'''_{1n})^2 + E(S_{2n})^2] \leq C(\zeta_{1n} + \zeta_{2n} + \zeta_{3n}), \\ E[S_n(S''_{1n} + S'''_{1n} + S_{2n})] &\leq E^{1/2}(S_n^2)E^{1/2}(S''_{1n} + S'''_{1n} + S_{2n})^2 \leq C(\zeta_{1n}^{1/2} + \zeta_{2n}^{1/2} + \zeta_{3n}^{1/2}). \end{aligned}$$

It has been found that

$$\begin{aligned} |E(S'_{1n})^2 - 1| &= |E(S''_{1n} + S'''_{1n} + S_{2n})^2 - 2E\{S_n(S''_{1n} + S'''_{1n} + S_{2n})\}| \\ &\leq C(\zeta_{1n}^{1/2} + \zeta_{2n}^{1/2} + \zeta_{3n}^{1/2}). \end{aligned} \tag{3.10}$$

On the other hand, from the basic definition of  $\rho$ -mixing, Lemmas 3.2, A.5(iv), and (A1), we can prove that

$$\begin{aligned} |\Delta_n| &\leq \sum_{1 \leq i < j \leq k} \sum_{s_1=k_i}^{k_i+p-1} \sum_{t_1=k_j}^{k_j+p-1} |\text{Cov}(W_{ns_1}, W_{nt_1})| \\ &\leq \sum_{1 \leq i < j \leq k} \sum_{s_1=k_i}^{k_i+p-1} \sum_{t_1=k_j}^{k_j+p-1} \sum_{u=\max\{1, s_1-n\}}^{\min\{n, s_1+n\}} \sum_{v=\max\{1, t_1-n\}}^{\min\{n, t_1+n\}} \sigma_n^{-2} \left| \int_{A_u} E_m(t, s) ds \int_{A_v} E_m(t, s) ds \right| \\ &\quad \cdot |a_{u-s_1} a_{v-t_1}| |\text{Cov}(e_{s_1}, e_{t_1})| \\ &\leq C \sum_{1 \leq i < j \leq k} \sum_{s_1=k_i}^{k_i+p-1} \sum_{t_1=k_j}^{k_j+p-1} \sum_{u=\max\{1, s_1-n\}}^{\min\{n, s_1+n\}} \sum_{v=\max\{1, t_1-n\}}^{\min\{n, t_1+n\}} \left| \int_{A_u} E_m(t, s) ds \right| |a_{u-s_1} a_{v-t_1}| \\ &\quad \cdot \rho(t_1 - s_1) \sqrt{\text{Var}(e_{s_1}) \text{Var}(e_{t_1})} \\ &\leq C \sum_{i=1}^{k-1} \sum_{s_1=k_i}^{k_i+p-1} \sum_{u=\max\{1, s_1-n\}}^{\min\{n, s_1+n\}} \left| \int_{A_u} E_m(t, s) ds \right| |a_{u-s_1}| \sum_{j=i+1}^k \sum_{t_1=k_j}^{k_j+p-1} \rho(t_1 - s_1) \\ &\quad \cdot \sum_{v=\max\{1, t_1-n\}}^{\min\{n, t_1+n\}} |a_{v-t_1}| \end{aligned}$$

$$\begin{aligned}
 &\leq C \sum_{i=1}^{k-1} \sum_{s_1=k_i}^{k_i+p-1} \sum_{u=\max\{1, s_1-n\}}^{\min\{n, s_1+n\}} \left| \int_{A_u} E_m(t, s) ds \right| |a_{u-s_1}| \sum_{j=q}^{\infty} \rho(j) \\
 &\leq Cu(q) \sum_{i=1}^{k-1} \sum_{s_1=k_i}^{k_i+p-1} \sum_{u=1}^n \left| \int_{A_u} E_m(t, s) ds \right| |a_{u-s_1}| \\
 &\leq Cu(q) \sum_{u=1}^n \left| \int_{A_u} E_m(t, s) ds \right| \left( \sum_{i=1}^{k-1} \sum_{s_1=k_i}^{k_i+p-1} |a_{u-s_1}| \right) \leq Cu(q). \tag{3.11}
 \end{aligned}$$

Hence, combining (3.10) with (3.11), we can see that

$$|s_n^2 - 1| \leq |E(S'_{1n})^2 - 1| + 2|\Delta_n| \leq C\{\zeta_{1n}^{1/2} + \zeta_{2n}^{1/2} + \zeta_{3n}^{1/2} + u(q)\}. \quad \square$$

**Lemma 3.5** *Assume that (A1)-(A6) hold, and applying this in Lemma 3.4, we can get*

$$\sup_u |P(T_n/s_n \leq u) - \Phi(u)| \leq C\zeta_{2n}^{\delta/2}.$$

*Proof of Lemma 3.5* It follows from the Berry-Esseen inequality (Petrov [14]) that we have

$$\sup_u |P(T_n/s_n \leq u) - \Phi(u)| \leq C \frac{\sum_{w=1}^k E|y_{nw}|^r}{s_n^r} \quad \text{for } r \geq 2. \tag{3.12}$$

From Definition 1.2, hence  $\sum_{j=k_v}^{\lfloor \log p \rfloor} \rho^{2/r}(2^j) = o(\log p)$ . Further,  $\exp(C_1 \sum_{j=k_v}^{\lfloor \log p \rfloor} \rho^{2/r}(2^j)) = o(p^\iota)$  for any  $C_1 > 0$  and  $\iota > 0$  (for  $\iota$  small enough). According to Lemma A.5(i) and Lemma A.2, we can get

$$\begin{aligned}
 \sum_{v=1}^k E|y_{nv}|^r &= \sum_{v=1}^k E \left| \sum_{j=k_v}^{k_v+p-1} \sum_{i=\max\{1, j-n\}}^{\min\{n, j+n\}} \sigma_n^{-1} a_{i-j} \int_{A_i} E_m(t, s) ds e_j \right|^r \\
 &\leq C \sum_{v=1}^k p^{r/2} \exp\left( C_1 \sum_{j=k_v}^{\lfloor \log p \rfloor} \rho(2^j) \right) \\
 &\quad \cdot \max_{1 \leq j \leq p} \left( E \left| \sum_{i=\max\{1, j-n\}}^{\min\{n, j+n\}} \sigma_n^{-1} a_{i-j} \int_{A_i} E_m(t, s) ds e_j \right|^2 \right)^{r/2} \\
 &\quad + C \sum_{v=1}^k p \exp\left( C_1 \sum_{j=k_v}^{\lfloor \log p \rfloor} \rho^{2/r}(2^j) \right) \\
 &\quad \cdot \max_{1 \leq j \leq p} E \left| \sum_{i=\max\{1, j-n\}}^{\min\{n, j+n\}} \sigma_n^{-1} a_{i-j} \int_{A_i} E_m(t, s) ds e_j \right|^r \\
 &\leq C \sigma_n^{-r} \left( \sum_{j=-\infty}^{\infty} |a_j| \right)^r \sum_{v=1}^k \left( p^{r/2+\iota} \left( \frac{2^m}{n} \right)^r + p^{1+\iota} \left( \frac{2^m}{n} \right)^r \right) \\
 &\leq Ck p^{r/2+\iota} \left( \frac{2^m}{n} \right)^{r/2} \leq Cn p^{r/2-1} \left( \frac{2^m}{n} \right)^{r/2} \leq Cn \left( p \frac{2^m}{n} \right)^{r/2} = Cn \zeta_{2n}^r. \tag{3.13}
 \end{aligned}$$

Hence, by Lemma 3.4, and combining (3.12) with (3.13), we can get the result. □



**Lemma 3.6** *Assume that (A1)-(A6) hold, and applying this in Lemma 3.4, we can get*

$$\sup_u |P(S'_{1n} \leq u) - P(T_n \leq u)| \leq C\{\zeta_{2n}^{\delta/2} + \zeta_{4n}^{1/2}\}.$$

*Proof of Lemma 3.6* Suppose that  $\phi_1(t)$  and  $\psi_1(t)$  are the characteristic functions of  $S'_{1n}$  and  $T_n$ . Therefore, it follows from Lemmas 3.1, 3.2, A.5, and (A1) that we can obtain

$$\begin{aligned} |\phi_1(t) - \psi_1(t)| &= \left| E \exp\left( it \sum_{v=1}^k y_{nv} \right) - \prod_{v=1}^k E \exp(it y_{nv}) \right| \\ &\leq C|t|\rho^{1/2}(q) \sum_{v=1}^k \|y_{nv}\|_2 \\ &\leq C|t|\rho^{1/2}(q) \sum_{v=1}^k \left\{ E \left( \sum_{i=k_v}^{k_v+p-1} \sigma_n^{-1} \sum_{j=\max\{1, i-n\}}^{\min\{n, i+n\}} a_{j-i} \int_{A_j} E_m(t, s) ds |e_i| \right)^2 \right\}^{1/2} \\ &\leq C|t|\rho^{1/2}(q) \left( \sum_{l=-\infty}^{\infty} |a_l| \right) \left\{ k \sum_{v=1}^k \sum_{i=k_v}^{k_v+p-1} \left| \int_{A_j} E_m(t, s) ds \right| \right\}^{1/2} \\ &\leq C|t|(k\rho(q))^{1/2} \leq C|t|\zeta_{4n}^{1/2}, \end{aligned}$$

which implies that

$$\int_{-T}^T \left| \frac{\phi_1(t) - \psi_1(t)}{t} \right| dt \leq C\zeta_{4n}^{1/2} T. \tag{3.14}$$

Note that

$$P(T_n \leq u) = P(T_n/s_n \leq u/s_n).$$

Consequently, from Lemma 3.5, it has been found

$$\begin{aligned} &\sup_u |P(T_n \leq u + y) - P(T_n \leq u)| \\ &= \sup_u P|(T_n/s_n \leq u + y/s_n) - P(T_n/s_n \leq u/s_n)| \\ &\leq \sup_u |P(T_n/s_n \leq u + y/s_n) - \Phi(u + y/s_n)| + \sup_u |\Phi(u + y/s_n) - \Phi(u/s_n)| \\ &\quad + \sup_u |P(T_n/s_n \leq u/s_n) - \Phi(u/s_n)| \\ &\leq 2 \sup_u |P(T_n/s_n \leq u/s_n) - \Phi(u)| + \sup_u |\Phi(u + y/s_n) - \Phi(u/s_n)| \\ &\leq C\{\zeta_{2n}^{\delta/2} + |y|/s_n\} \leq C\{\zeta_{2n}^{\delta/2} + |y|\}. \end{aligned}$$

Therefore

$$T \sup_u \int_{|y| \leq c/T} |P(T_n \leq u + y) - P(T_n \leq u)| dy \leq C\{\zeta_{2n}^{\delta/2} + 1/T\}. \tag{3.15}$$

Thus, combining (3.14) with (3.15) and taking  $T = \zeta_{4n}^{-1/4}$ , it suffices to prove that

$$\begin{aligned} & \sup_u |P(S'_{1n} \leq u) - P(T_n \leq u)| \\ & \leq \int_{-T}^T \left| \frac{\phi_1(t) - \psi_1(t)}{t} \right| dt + T \sup_u \int_{|y| \leq c/T} |P(T_n \leq u + y) - P(T_n \leq u)| dy \\ & \leq C \{ \zeta_{4n}^{1/2} T + \zeta_{2n}^{\delta/2} + 1/T \} = C \{ \zeta_{2n}^{\delta/2} + \zeta_{4n}^{1/4} \}. \end{aligned} \quad \square$$

#### 4 Proofs of the main results

*Proof of Theorem 2.1*

$$\begin{aligned} & \sup_u |P(S'_{1n} \leq u) - \Phi(u)| \\ & \leq \sup_u |P(S'_{1n} \leq u) - P(T_n \leq u)| + \sup_u |P(T_n \leq u) - \Phi(u/s_n)| + \sup_u |\Phi(u/s_n) - \Phi(u)| \\ & := J_{1n} + J_{2n} + J_{3n}. \end{aligned} \quad (4.1)$$

According to Lemma 3.6, Lemma 3.5, and Lemma 3.4, it follows that

$$J_{1n} \leq C \{ \zeta_{2n}^{\delta/2} + \zeta_{4n}^{1/4} \}, \quad (4.2)$$

$$J_{2n} = \sup_u |P(T_n/s_n \leq u/s_n) - \Phi(u/s_n)| = \sup_u |P(T_n/s_n \leq u) - \Phi(u)| \leq C \zeta_{2n}^{\delta/2}, \quad (4.3)$$

$$J_{3n} \leq C |s_n^2 - 1| \leq C \{ \zeta_{1n}^{1/2} + \zeta_{2n}^{1/2} + \zeta_{3n}^{1/2} + u(q) \}. \quad (4.4)$$

Hence, by (4.2)-(4.4) and combining with (4.1), we have

$$\sup_u |P(S'_{1n} \leq u) - \Phi(u)| \leq C \{ \zeta_{1n}^{1/2} + \zeta_{2n}^{1/2} + \zeta_{2n}^{\delta/2} + \zeta_{3n}^{1/2} + \zeta_{4n}^{1/4} + u(q) \}. \quad (4.5)$$

Thus, by Lemma A.4, Lemma 3.3(2), and (4.5), it suffices to prove that

$$\begin{aligned} & \sup_u |P(S_n \leq u) - \Phi(u)| \\ & \leq C \left\{ \sup_u |P(S'_{1n} \leq u) - \Phi(u)| + \sum_{i=1}^3 \zeta_{in}^{1/3} + P(|S'_{1n}| \geq \zeta_{1n}^{1/3}) \right. \\ & \quad \left. + P(|S''_{1n}| \geq \zeta_{2n}^{1/3}) + P(|S_{2n}| \geq \zeta_{3n}^{1/3}) \right\} \\ & \leq C \{ \zeta_{1n}^{1/3} + \zeta_{2n}^{1/3} + \zeta_{2n}^{\delta/2} + \zeta_{3n}^{1/3} + \zeta_{4n}^{1/4} + u(q) \}. \end{aligned} \quad \square$$

*Proof of Corollary 2.1* By (A1), since  $\sum_{j=1}^\infty \rho(j) < \infty$ , we can easily see that  $u(q) \rightarrow 0$ , therefore Corollary 2.1 holds. □

*Proof of Corollary 2.2* Let  $p = [n^\tau]$ ,  $q = [n^{2\tau-1}]$ . Taking  $\tau = \frac{1}{2} + \frac{8\theta-1}{2(6\lambda+7)}$ ,  $\frac{2}{9-6\lambda} < \theta \leq 1$ ,  $\tau < \theta$ . Consequently,

$$\zeta_{1n}^{1/3} = \zeta_{2n}^{1/3} = O(n^{-\frac{\theta-\tau}{3}}) = O(n^{-\frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}}),$$

$$\begin{aligned} \zeta_{3n}^{1/3} &= n^{-\frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}} = \left( n^{\frac{6\lambda\theta+3\lambda+3\theta+4}{2(6\lambda+7)}} \sum_{|j|>n} |a_j| \right)^{2/3} = O\left( n^{-\frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}} \right), \\ \zeta_{4n}^{1/4} &= O\left( n^{-\frac{\tau+\lambda(2\tau-1)-1}{4}} \right) = O\left( n^{-\frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}} \right), \\ u(q) &= O\left( \sum_{i=q}^{\infty} i^{-\lambda} \right) = O(p^{-\lambda+1}) = O(n^{-(2\tau-1)(\lambda-1)}) = O\left( n^{-\frac{(8\theta-1)(\lambda-1)}{6\lambda+7}} \right). \end{aligned}$$

Finally, taking  $\frac{2}{9-6\lambda} < \theta$ , hence  $\frac{(8\theta-1)(\lambda-1)}{6\lambda+7} > \frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}$ , it has been found that  $u(q) = O\left( n^{-\frac{\lambda(2\theta-1)+(\theta-1)}{6\lambda+7}} \right)$ , therefore, the desired result is completed by Corollary 2.1 immediately.  $\square$

**Appendix**

**Lemma A.1** (Shao [7]) *Let  $\{X_i : i \geq 1\}$  be a  $\rho$ -mixing sequence,  $s, t > 1$ , and  $1/s + 1/t = 1$ . If  $X \in L_s(\mathcal{F}_1^k)$ ,  $Y \in L_t(\mathcal{F}_{k+n}^\infty)$ , then*

$$|EXY - EXEY| \leq 10\rho^{2(\frac{1}{s} \wedge \frac{1}{t})}(n)\|X\|_s\|Y\|_t.$$

**Lemma A.2** (Shao [7]) *Assume that  $EX_i = 0$  and  $\|X_i\|_q$  for some  $q \geq 2$ . Then there exists a positive constant  $K = K(q, \rho(\cdot))$  depending only on  $q$  and  $\rho(\cdot)$  such that for any  $k \geq 0, n \geq 1$ ,*

$$\begin{aligned} E \max_{1 \leq i \leq n} |S_k(i)|^q &\leq Kn^{q/2} \exp\left( K \sum_{i=0}^{[\log n]} \rho(2^i) \right) \max_{k \leq i \leq k+n} \|X_i\|_2^q \\ &\quad + nK \exp\left( K \sum_{i=0}^{[\log n]} \rho^{2/q}(2^i) \right) \max_{k \leq i \leq k+n} \|X_i\|_q^q. \end{aligned}$$

**Lemma A.3** (Yang [15]) *Let  $\{X_i : i = 1, 2, \dots\}$  be a  $\rho$ -mixing sequence, and there is a  $\lambda > 0$ , make  $\rho(n) = O(n^{-\lambda})$ , with  $EX_i = 0, E|X_i|^r < \infty$  ( $r > 1$ ), when any integer  $m \geq 1$ , there exists a positive constant  $C(m)$ , then:*

(i) *for  $1 < r \leq 2$ , we have*

$$E \left| \sum_{i=1}^n X_i \right|^r \leq C(m)n^{\beta(m)} \sum_{i=1}^n E|X_i|^r,$$

(ii) *for  $r > 2$ , we have*

$$E \left| \sum_{i=1}^n X_i \right|^r \leq C(m)n^{\beta(m)} \left\{ \sum_{i=1}^n E|X_i|^r + \left( \sum_{i=1}^n EX_i^2 \right)^{r/2} \right\}.$$

In  $\beta(m) = (r - 1)\omega^m$  and  $0 < \omega < 1$ .

**Lemma A.4** (Yang [16]) *Suppose that  $\{\zeta_n : n \geq 1\}, \{\eta_n : n \geq 1\}$ , and  $\{\xi_n : n \geq 1\}$  are three random variable sequences,  $\{\gamma_n : n \geq 1\}$  is a positive constant sequence, and  $\gamma_n \rightarrow 0$ . If  $\sup_u |F_{\gamma_n}(u) - \Phi(u)| \leq C\gamma_n$  then for any  $\varepsilon_1 > 0$ , and  $\varepsilon_2 > 0$ , then*

$$\sup_u |F_{\zeta_n + \eta_n + \gamma_n}(u) - \Phi(u)| \leq C\{\gamma_n + \varepsilon_1 + \varepsilon_2 + P(|\eta_n| \geq \varepsilon_1) + P(|\xi_n| \geq \varepsilon_2)\}.$$

**Lemma A.5** (Li *et al.* [12]) *Under assumptions (A2)-(A5), we have*

- (i)  $|\int_{A_i} E_m(t, s) ds| = O(\frac{2^m}{n}), i = 1, 2, \dots, n;$
- (ii)  $\sum_{i=1}^n (\int_{A_i} E_m(t, s) ds)^2 = O(\frac{2^m}{n});$
- (iii)  $\sup_m \int_0^1 |E_m(t, s) ds| \leq C;$
- (iv)  $\sum_{i=1}^n |\int_{A_i} E_m(t, s) ds| \leq C.$

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

The two authors contributed equally and significantly in writing this article. Both authors read and approved the final manuscript.

#### Author details

<sup>1</sup>School of Information and Statistics, Guangxi University of Finance and Economics, Nanning, 530003, P.R. China. <sup>2</sup>School of Mathematics and Computer Science, Shangrao Normal University, Shangrao, 334001, P.R. China.

#### Authors' information

Liwang Ding (1985-), male, lecturer, graduate, major in probability theory and mathematical statistics.

#### Acknowledgements

This project is supported by the National Natural Science Foundation of China (11461057).

Received: 20 October 2015 Accepted: 29 February 2016 Published online: 31 March 2016

#### References

- Xue, LG: Berry-Esseen bound of an estimate of error variance in a semiparametric regression model. *Acta Math. Sin.* **48**(1), 157-170 (2005)
- Liang, HY, Li, YY: A Berry-Esseen type bound of regression estimator based on linear process errors. *J. Korean Math. Soc.* **45**(6), 1753-1767 (2008)
- Li, YM, Wei, CD, Xin, GD: Berry-Esseen bounds of wavelet estimator in a regression with linear process errors. *Stat. Probab. Lett.* **81**(1), 103-111 (2011)
- Li, YM, Guo, JH, Yang, SC: The Berry-Esseen bounds of wavelet estimators for semiparametric regression model whose errors for a linear process with a mixing innovations. *Acta Math. Appl. Sin.* **36**(6), 1021-1036 (2013)
- Kolmogorov, AN, Rozanov, UA: On the strong mixing conditions for stationary Gaussian process. *Theory Probab. Appl.* **5**(2), 204-208 (1960)
- Shao, QM: Almost sure convergence properties of  $\rho$ -mixing sequences. *Acta Math. Sin.* **32**(3), 377-393 (1989)
- Shao, QM: Maximal inequalities for sums of  $\rho$ -mixing sequences. *Ann. Probab.* **23**, 948-965 (1995)
- Jiang, DY: On the convergence rates in the law of iterated logarithm of  $\rho$ -mixing sequences. *Math. Appl.* **15**(3), 32-37 (2002)
- Chen, PY, Liu, XD: Complete moment convergence for sequence of identically distributed  $\rho$ -mixing random variables. *Acta Math. Sin.* **51**(2), 281-290 (2008)
- Zhou, XC, Lin, JG: Strong consistency of estimators in partially linear models for longitudinal data with mixing-dependent structure. *J. Inequal. Appl.* **2011**, 112 (2011)
- Tan, XL, Wang, M: Strong convergence results for weighted sums of  $\rho$ -mixing random variables sequence. *J. Jilin Univ. Sci. Ed.* **52**(5), 927-932 (2014)
- Li, YM, Yin, CM, Wei, CD: The asymptotic normality for  $\varphi$ -mixing dependent of wavelet regression function estimator. *Acta Math. Appl. Sin.* **31**(6), 1016-1055 (2008)
- Liang, HY, Qi, YY: Asymptotic normality of wavelet estimator of regression function under NA assumptions. *Bull. Korean Math. Soc.* **4**(2), 247-257 (2007)
- Petrov, VV: *Limit Theory for Probability Theory*. Oxford University Press, New York (1995)
- Yang, SC: Moment inequality for mixing sequences and nonparametric estimation. *Acta Math. Sin.* **40**(2), 271-279 (1997)
- Yang, SC: Uniformly asymptotic normality of the regression weighted estimator for negatively associated samples. *Stat. Probab. Lett.* **62**, 101-110 (2003)