Annals of the Institute of Statistical Mathematics (2019) 71:1143–1162
https://doi.org/10.1007/s10463-018-0677-6

The Berry–Esseen bounds of the weighted estimator in a nonparametric regression model

Xuejun Wang[1] · **Yi Wu**[1] · **Shuhe Hu**[1]

Received: 29 August 2017 / Revised: 24 April 2018 / Published online: 30 June 2018

Abstract
Consider the following nonparametric model: $Y_{ni} = g(x_{ni}) + \varepsilon_{ni}$, $1 \le i \le n$, where $x_{ni} \in \mathbb{A}$ are the nonrandom design points and $\mathbb{A}$ is a compact set of $\mathbb{R}^m$ for some $m \ge 1$, $g(\cdot)$ is a real valued function defined on $\mathbb{A}$, and $\varepsilon_{n1}, \ldots, \varepsilon_{nn}$ are ρ^--mixing random errors with zero mean and finite variance. We obtain the Berry–Esseen bounds of the weighted estimator of $g(\cdot)$. The rate can achieve nearly $O(n^{-1/4})$ when the moment condition is appropriate. Moreover, we carry out some simulations to verify the validity of our results.

Keywords Berry–Esseen bound · ρ^--mixing random errors · Nonparametric regression model · Weighted estimator

1 Introduction

Consider the following nonparametric model:

$$Y_{ni} = g(x_{ni}) + \varepsilon_{ni}, \quad 1 \le i \le n, \tag{1}$$

where $x_{ni} \in \mathbb{A}$ are the nonrandom design points and $\mathbb{A}$ is a compact set of $\mathbb{R}^m$ for some $m \ge 1$, $g(\cdot)$ is a real valued function defined on $\mathbb{A}$, and $\varepsilon_{n1}, \ldots, \varepsilon_{nn}$ are random errors with zero mean and finite variance.

It is well known that the regression models have substantial applications in practical problems. Stone (1977) first introduced a weighted regression method to get the estimator of $g(\cdot)$, and then Georgiev (1985) adapted it to the fixed design case.

Supported by the National Natural Science Foundation of China (11671012, 11501004, 11501005), the Natural Science Foundation of Anhui Province (1508085J06) and the Key Projects for Academic Talent of Anhui Province (gxbjZD2016005)..

✉ Xuejun Wang
07019@ahu.edu.cn

1 School of Mathematical Sciences, Anhui University, Hefei 230601, People's Republic of China

Since then, many authors were devoted to studying the asymptotic properties for the weighted regression estimator. One can refer to Roussas (1989), Fan (1990), Roussas et al. (1992), Tran et al. (1996), Hu et al. (2002, 2003), Liang and Jing (2005), Yang et al. (2012), Shen et al. (2015), Wang et al. (2015), and Shen (2016) among others for the details.

As we know that, the independent assumption on random errors is not always reasonable in many stochastic models and realistic applications, since the samples are usually dependent. So it is more practice to assume that the random errors satisfy dependent structures. In this work, we will study the Berry–Esseen bounds of the weighted estimator in a nonparametric regression model based on ρ^--mixing errors, the concept of which will be stated below.

Now let us recall some concepts of dependence. The first one is the concept of negatively associated (NA, for short) random variables, which was introduced by Joag-Dev and Proschan (1983) as follows.

Definition 1 A finite family of random variables $\{X_i, 1 \le i \le n\}$ is said to be NA if for every pair of disjoint subsets A and B of $\{1, 2, \ldots, n\}$ and any real coordinatewise nondecreasing functions f_1 on $\mathbb{R}^A$ and f_2 on $\mathbb{R}^B$,

$$\mathrm{Cov}(f_1(X_i, i \in A), f_2(X_j, j \in B)) \le 0,$$

whenever the covariance above exists. An infinite family of random variables is NA if every finite subfamily is NA.

Another important concept of dependent random variables is ρ^*-mixing, which was introduced by Bradley (1992) as follows.

Definition 2 A sequence $\{X_k, k \ge 1\}$ of random variables is called ρ^*-mixing if

$$\rho^*(s) = \sup\{\rho(S, T); S, T \subset \mathbb{N}, \mathrm{dist}(S, T) \ge s\} \to 0$$

as $s \to \infty$, where

$$\rho(S, T) = \sup\left\{\frac{|\mathrm{Cov}(X, Y)|}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} : X \in L_2(\sigma(X_k, k \in S)), Y \in L_2(\sigma(X_k, k \in T))\right\}.$$

The following are some examples satisfying ρ^*-mixing structure.

Example 1 Let $\{\xi_n\}$ be a sequence of i.i.d. random variables with zero mean and finite variance. Define

$$X_n = \sum_{j=0}^{l} c_j \xi_{n-j}$$

for some positive integer l and constants $c_j, j = 0, 1, \ldots, l$. Then $\{X_n\}$ is known as a moving average process with older l. It can be easily verified that $\{X_n\}$ is a ρ^*-mixing process.

Example 2 Let $\{X_n, n \geq 1\}$ be a strictly stationary, finite-state, irreducible and aperiodic Markov chain. Then it is a ρ^*-mixing process with $\rho^*(k) = o(e^{-Ck})$ for some $C > 0$. One can refer to Theorem 1.3 in Bradley (1997) for the details.

Zhang and Wang (1999) introduced the following concept of ρ^--mixing random variables.

Definition 3 A sequence $\{X_k, k \geq 1\}$ of random variables is called ρ^--mixing if

$$\rho^-(s) = \sup\{\rho^-(S, T) : S, T \subset \mathbb{N}, \text{dist}(S, T) \geq s\} \to 0$$

as $s \to \infty$, where

$$\rho^-(S, T) = 0 \vee \left\{ \frac{\text{Cov}(f_1(X_i, i \in S), f_2(X_j, j \in T))}{\sqrt{\text{Var}(f_1(X_i, i \in S))\text{Var}(f_2(X_j, j \in T))}} : f_1, f_2 \in \mathcal{C} \right\}$$

and $\mathcal{C}$ is the set of nondecreasing functions.

An array $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ of random variables is said to be rowwise ρ^--mixing if for every $n \geq 1$, $\{X_{ni}, 1 \leq i \leq n\}$ are ρ^--mixing.

It is easy to see that $\rho^-(s) \leq \rho^*(s)$ and ρ^--mixing implies NA if and only if $\rho^-(1) = 0$. Therefore, ρ^--mixing random variables include ρ^*-mixing random variables and NA random variables as special cases. Consequently, the study of the limit properties for ρ^--mixing random variables is of great interest. Since the concept of ρ^--mixing random variables was introduced by Zhang and Wang (1999), many interesting results have been established. For instance, Zhang and Wang (1999) obtained moment inequalities and the complete convergence for partial sums, Zhang (2000a, b) obtained the central limit theorems, Wang and Lu (2006) established some inequalities for the maximum of partial sums and weak convergence, Wang and Zhang (2007) obtained the law of the iterated logarithm, Liu and Liu (2009) showed moments of the maximum of normed partial sums and so on. However, as far as we know, there is no literature investigating the asymptotic properties for the estimator of the model (1) with ρ^--mixing errors.

Remark 1 We point out that ρ^*-mixing and NA are both ρ^--mixing. Hence, the sequences $\{X_n, n \geq 1\}$ in Examples 1 and 2 are both ρ^--mixing. However, the converse is not always true. The following gives an example of a ρ^--mixing sequence which is neither NA nor ρ^*-mixing.

Example 3 Let $\{\xi_n, n \geq 1\}$, $\{\eta_n, n \geq 1\}$ and $\{\tau_n, n \geq 1\}$ be three independent sequences of independent and identically distributed standard normal random variables. Let

$$X_n = \begin{cases} \xi_m, & \text{if } n = 2m - 1 \\ -\xi_m, & \text{if } n = 2m \end{cases}, \quad Y_n = \begin{cases} \eta_m, & \text{if } n = 2^{2m-1} \\ -\eta_m, & \text{if } n = 2^{2m} \\ \tau_n, & \text{otherwise} \end{cases},$$

and $Z_n = X_n^2 + Y_n$. Then $\{X_n, n \geq 1\}$ and $\{Y_n, n \geq 1\}$ are two independent sequences of identically distributed NA normal random variables. Also, $\{X_n, n \geq 1\}$ is a two-dependent sequence, so $\{X_n, n \geq 1\}$ is a ρ^*-mixing sequence with $\rho^*(2) = 0$. From Property P3 of Zhang and Wang (1999), we can see that $\{Z_n, n \geq 1\}$ is ANA with $\rho^-(2) = 0$. But $\{Z_n, n \geq 1\}$ is neither NA nor ρ^*-mixing, since

$$\text{Cov}\,(Z_{2m-1}, Z_{2m}) = \text{Cov}\left(X_{2m-1}^2, X_{2m}^2\right) = E\xi_m^4 - (E\xi_m^2)^2 = 2 > 0$$

and

$$\frac{\text{Cov}\left(Z_{2^{2m-1}}, Z_{2^{2m}}\right)}{Var(Z_{2^{2m-1}})\text{Var}(Z_{2^{2m}})} = -\frac{1}{3} \nrightarrow 0 \text{ as } \text{dist}(2^{2m-1}, 2^{2m}) = 2^{2m-1} \to \infty.$$

This example can be found in Zhang and Wang (1999).

In this paper, we will study the Berry–Esseen bounds for the weighted estimator of model (1) based on ρ^--mixing errors. The Berry–Esseen bounds of the weighted estimator can achieve $O(n^{-1/4})$ when the moment condition is appropriate. Finally we carry out some simulations to verify the validity of our theoretical results.

Throughout this paper, the symbol C represents some positive constant which may vary in different places. Let $I(B)$ be the indicator function of the set B and $\lfloor x \rfloor$ denote the integer part of x. Denote $x^+ = xI(x \geq 0)$, $x^- = -xI(x < 0)$, and $\|\xi\|_{2,1} = \int_0^\infty P^{1/2}(|\xi| \geq x)dx$. $\Phi(u)$ is the distribution function of $N(0, 1)$.

The layout of this paper is as follows: In Sect. 2, we introduce the estimators of unknown functions in model (1) and present the results. Some numerical simulations are provided in Sect. 3. Some lemmas and the proof of the main result are stated in Sect. 4.

2 Main results

In model (1), a normal estimator of $g(\cdot)$ is the following general linear smoother:

$$g_n(x) = \sum_{i=1}^{n} \omega_{ni}(x) Y_{ni}, \tag{2}$$

where the weight functions $\omega_{ni}(x)$, $i = 1, 2, \ldots, n$ depend on the fixed design points $x_{n1}, x_{n2}, \ldots, x_{nn}$ and the number of observations n.

For convenience, we need to define some notations as follows. Denote by $\omega_n = \omega_n(x) = \max_{1\leq i\leq n} \omega_{ni}(x)$, $\Delta_n^2 = \Delta_n^2(x) = \text{Var}(g_n(x))$, $S_n = S_n(x) = \Delta_n^{-1}(x)(g_n(x) - Eg_n(x))$, and $F_n(u) = P(S_n(x) < u)$. To present the results, the following assumptions are needed.

(A_1). (i) $g(\cdot)$ is a bounded real valued function defined on $\mathbb{A}$; (ii) $\{\xi_i, i \geq 1\}$ is a sequence of mean zero ρ^--mixing random variables with $\sup_{i\geq 1} E\xi_i^2 < \infty$; (iii) for each n, the joint distribution of $\{\varepsilon_{ni}, 1 \leq i \leq n\}$ is the same as that of $\{\xi_i, 1 \leq i \leq n\}$.

$(\boldsymbol{A_2})$. (i) $\omega_{ni}(x) \geq 0$ for all $1 \leq i \leq n$ and $n \geq 1$; (ii) $\sum_{i=1}^{n} \omega_{ni}(x) \leq C$ for all $n \geq 1$; (iii) $\omega_n(x) = O(\Delta_n^2(x))$ and $\Delta_n^2(x) > 0$.
$(\boldsymbol{A_3})$. There exist positive integers $p = p(n)$ and $q = q(n)$ such that for some positive constant c and all n large enough,

$$p + q \leq n, \ \ qp^{-1} \leq c, \tag{3}$$

and as $n \to \infty$,

$$nqp^{-1}\omega_n(x) \to 0, \ \ p\omega_n(x) \to 0, \ \ u(q) =: \sup_{j\geq 1} \sum_{i:|i-j|\geq q} |\mathrm{Cov}(\xi_i, \xi_j)| \to 0,$$

$$v(q) =: \sum_{i=q}^{\infty} \rho^{-}(i) \to 0.$$

Remark 2 The assumptions above are similar to those of Yang (2003). As is stated in Yang (2003), the assumption imposed upon $u(q)$ can be easily satisfied. Moreover, if $\mathrm{Var}(\xi_i) = \sigma_0^2 > 0$ for each $i \geq 1$, then the restriction on $v(q)$ can be canceled, since $v(q) = \sigma_0^{-2} \sup_{j\geq 1} \sum_{i:|i-j|\geq q} \mathrm{Cov}(\xi_i, \xi_j)^+ \leq \sigma_0^{-2} u(q)$.

Our main result on Berry–Esseen bounds for the weighted estimator (2) is presented as follows.

Theorem 1 *Suppose that Assumptions* $(\boldsymbol{A_1}) - (\boldsymbol{A_3})$ *hold. If* $\sup_{i\geq 1} E|\xi_i|^{2+\delta} < \infty$ *for some* $\delta > 0$*, then*

$$\sup_u |F_n(u) - \Phi(u)| = O\left((p\omega_n)^{\delta/2} + (u(q) + v(q))^{1/3} + (nqp^{-1}\omega_n)^{(2+\delta)/(6+2\delta)} + (p\omega_n)^{(2+\delta)/(6+2\delta)}\right).$$

If we take $p = \lfloor n^{\theta} \rfloor$ and $q = \lfloor n^{2\theta-1} \rfloor$ for some $1/2 < \theta < 1$, we can obtain the following result.

Corollary 1 *Suppose that Assumptions* $(\boldsymbol{A_1}) - (\boldsymbol{A_2})$ *hold. If* $\sup_{i\geq 1} E|\xi_i|^{2+\delta} < \infty$ *for some* $\delta \geq \sqrt{3} - 1$*,* $\omega_n = O(n^{-r})$*,* $u(n) + v(n) = O(n^{-3(r-\theta)(2+\delta)/(4\theta-2)(3+\delta)})$*, where* $1/2 < \theta < r \leq 1$*, then*

$$\sup_u |F_n(u) - \Phi(u)| = O\left(n^{-(r-\theta)(2+\delta)/(6+2\delta)}\right).$$

If we take $p = \lfloor n^{1/2} \rfloor$ and $q = \lfloor \log n \rfloor$ in Theorem 1, we can also obtain the following result.

Corollary 2 *Suppose that Assumptions* $(\boldsymbol{A_1}) - (\boldsymbol{A_2})$ *hold. If* $\sup_{i\geq 1} E|\xi_i|^{2+\delta} < \infty$ *for some* $\delta \geq \sqrt{3} - 1$*,* $\omega_n = O(n^{-r})$*,* $u(n) + v(n) = O(e^{-3(2r-1)(2+\delta)n/(12+4\delta)})$*, where* $1/2 < r \leq 1$*, then*

$$\sup_u |F_n(u) - \Phi(u)| = O\left((n^{-(r-1/2)} \log n)^{(2+\delta)/(6+2\delta)}\right).$$

Remark 3 It is easy to check that if $\delta = 1$, the Berry–Esseen bound for $g_n(\cdot)$ approximates to $O(n^{-3/16})$ for $\theta \approx 1/2$ and $r \approx 1$, and if δ can be sufficiently large, the Berry–Esseen bound is nearly $O(n^{-1/4})$. Hence, the results above generalize the corresponding ones of Yang (2003) from NA setting to ρ^--mixing setting.

Remark 4 Since ρ^*-mixing implies ρ^--mixing, our results are also available for ρ^*-mixing random errors. It is worth to mention that as far as we know, there is no literature investigating the Berry–Esseen bound for the estimator (2) under the assumption of ρ^*-mixing errors.

3 Simulation

In this section, we will carry out some simulations to study the asymptotic normality of the estimator (2). The data are generated from model (1). We will consider the following two cases.

Case 1 For $0 < \rho < 1$ and $n \geq 3$, let $(\xi_1, \xi_2, \ldots, \xi_n) \sim N(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{0}$ represents zero vector and

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1 & -\rho & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\rho & 0 \\ 0 & 0 & 0 & \cdots & -\rho & 1 & -\rho \\ 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{pmatrix}_{n\times n}.$$

By Joag-Dev and Proschan (1983), it can be seen that $(\xi_1, \xi_2, \ldots, \xi_n)$ is a NA vector, and thus a ρ^--mixing vector. In order to verify the asymptotic normality of the estimator (2), we choose $\rho = 0.1$ and, respectively, $\rho = 0.3$, $p = \lfloor n^{1/2} \rfloor$ and $q = \lfloor \log n \rfloor$. For $\omega_{ni}(x)$, we choose $\omega_{ni}(x) = 1/n$ for simplicity. One can easily check that the conditions of Corollary 2 are all satisfied. Take the sample sizes n as $n = 200, 500, 800$, respectively. We use the R software to compute $S_n = \Delta_n^{-1}(x)(g_n(x) - Eg_n(x))$ for 1000 times and obtain the Quantile–Quantile plots in Figs. 1, 2, 3, 4, 5 and 6.

Figures 1, 2 and 3 are the Quantile–Quantile plots of S_n with $\rho = 0.1$, and Figs. 4, 5 and 6 are the Quantile–Quantile plots of S_n with $\rho = 0.3$. One can see from Figs. 1, 2, 3, 4, 5 and 6 that for different values of ρ, S_n converges to standard normal distribution as n increases.

Case 2 For fixed positive integer m, let $e_i \overset{i.i.d.}{\sim} U(-\sqrt{3/(m+1)}, \sqrt{3/(m+1)})$ and $\xi_i = \sum_{k=0}^{m} e_{i+k}$ for each $i \geq 1$. It is easy to show that $\{\xi_i, i \geq 1\}$ is a sequence

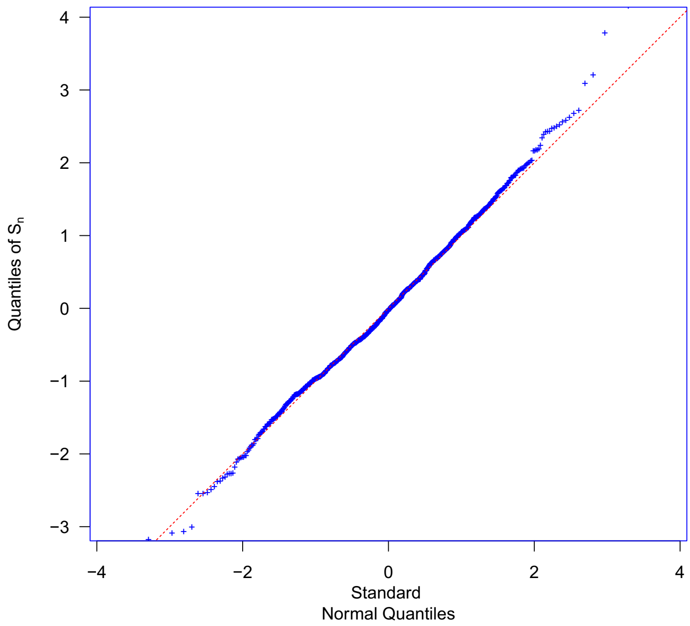

Fig. 1 200 sample with $\rho = 0.1$

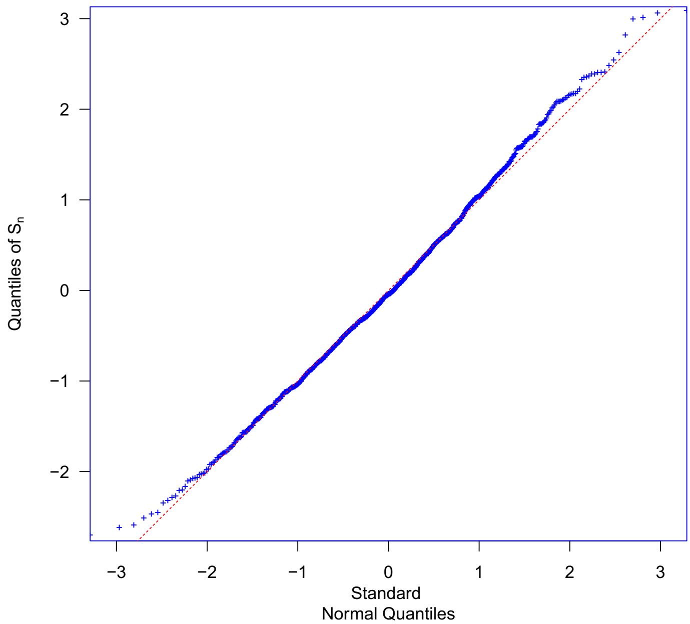

Fig. 2 500 sample with $\rho = 0.1$

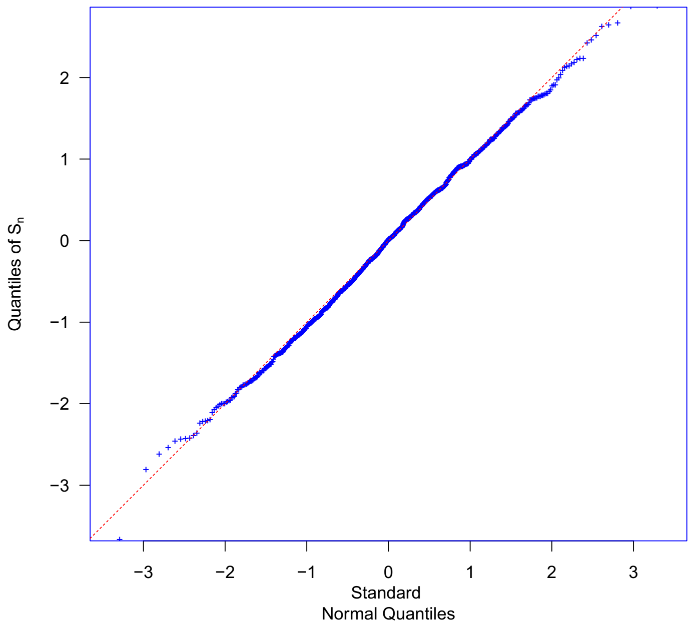

Fig. 3 800 sample with $\rho = 0.1$

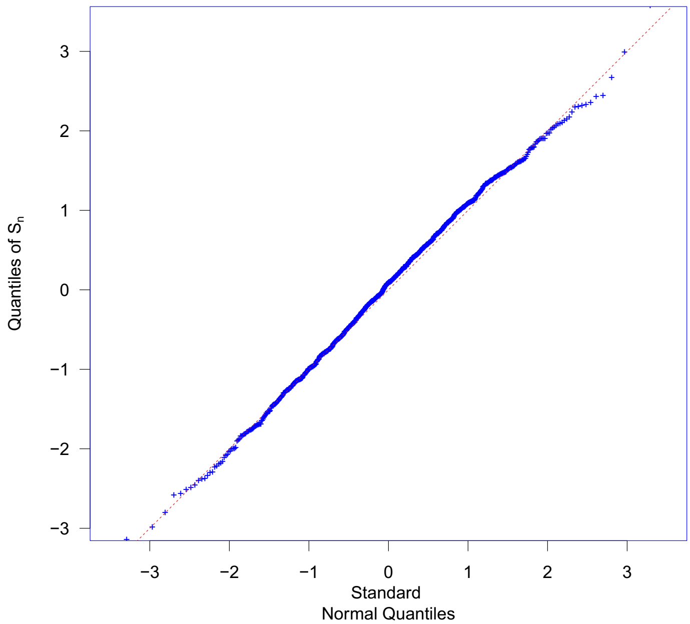

Fig. 4 200 sample with $\rho = 0.3$

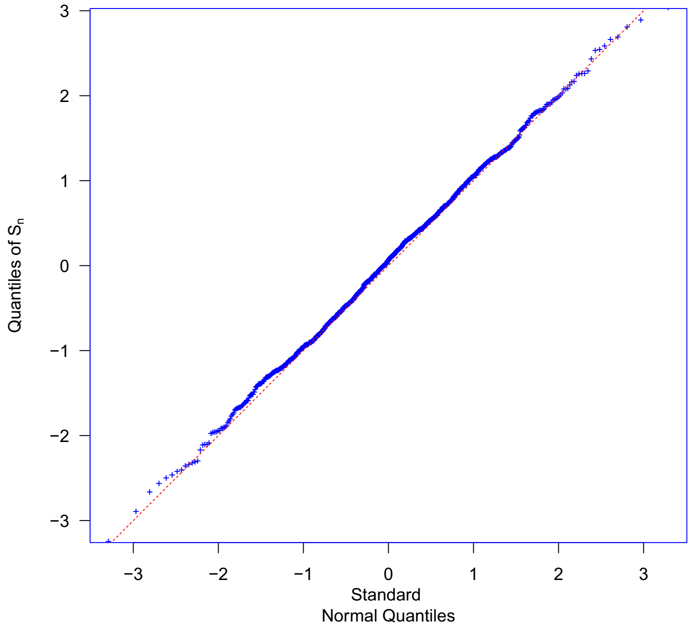

Fig. 5 500 sample with $\rho = 0.3$

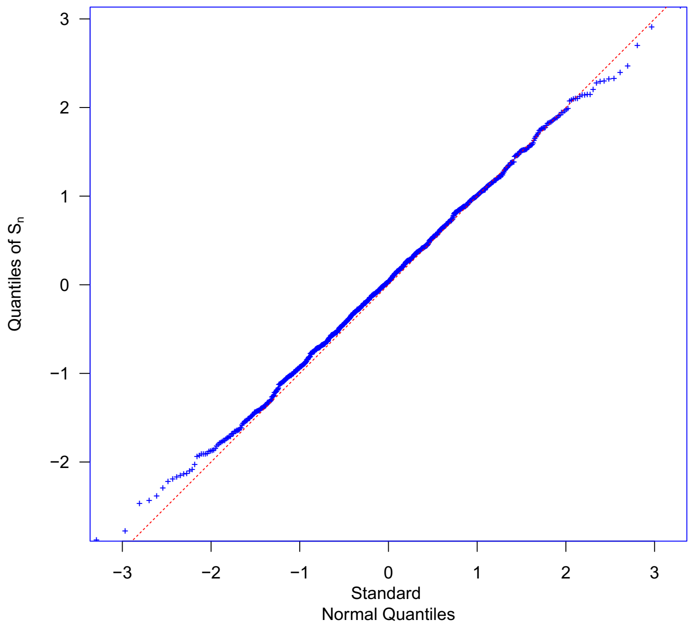

Fig. 6 800 sample with $\rho = 0.3$

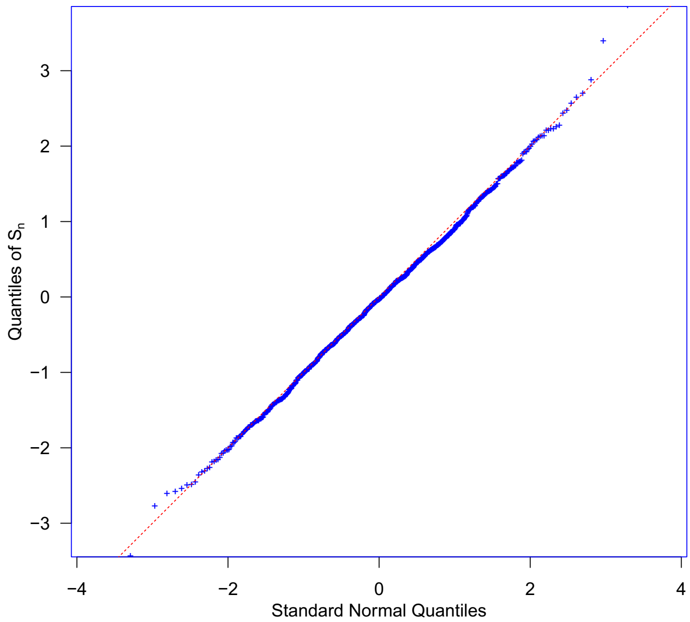

Fig. 7 200 sample with $m = 2$

of ρ^*-mixing random variables and thus a sequence of ρ^--mixing random variables with $E\xi_i = 0$ and $\text{Var}(\xi_i) = 1$. For simplicity, take $m = 2$ and $m = 3$, respectively. The other settings are the same as in Case 1. We also use the R software to obtain the Quantile–Quantile plots of S_n in Figs. 7, 8, 9, 10, 11 and 12.

From Figs. 7, 8, 9, 10, 11 and 12, we can also derive the same conclusion as that in Case 1, i.e., S_n also converges to standard normal distribution as the sample size n increases, no matter $m = 2$ or $m = 3$.

To compute the uniform Berry–Esseen bounds for the estimator (2) under Case 1 and Case 2, we first compute the empirical distribution function of S_n to estimate $F_n(u)$ and then estimate the maximum value of $|F_n(u) - \Phi(u)|$ for $u \in [-3, 3]$. The results are shown in Table 1. These simulations show a good fit of our main results established in Sect. 2.

4 Proof of main result

Before proving our main result, we first present some important lemmas as follows. The first one comes from Zhang and Wang (1999).

Lemma 1 *Increasing functions defined on disjoint subsets of a ρ^--mixing field $\{X_i, i \in \mathbb{N}^d\}$ with mixing coefficients $\rho^-(s)$ are also ρ^--mixing with mixing coefficients not greater than $\rho^-(s)$.*

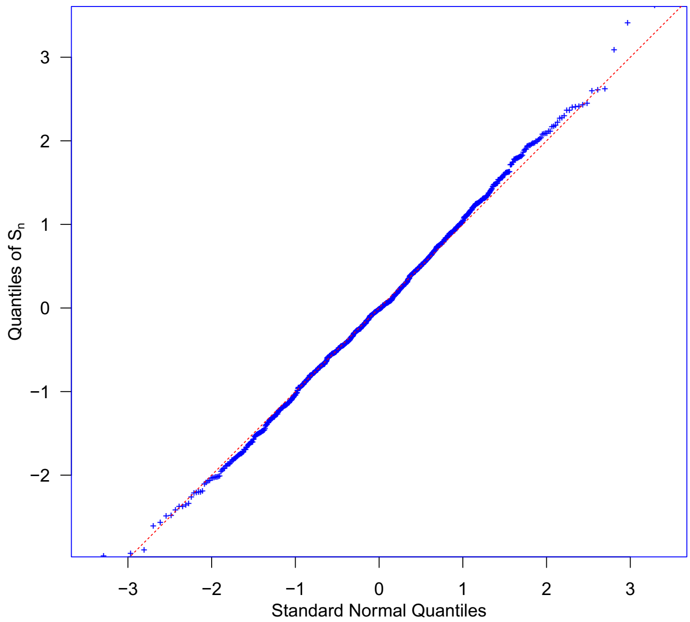

Fig. 8 500 sample with $m = 2$

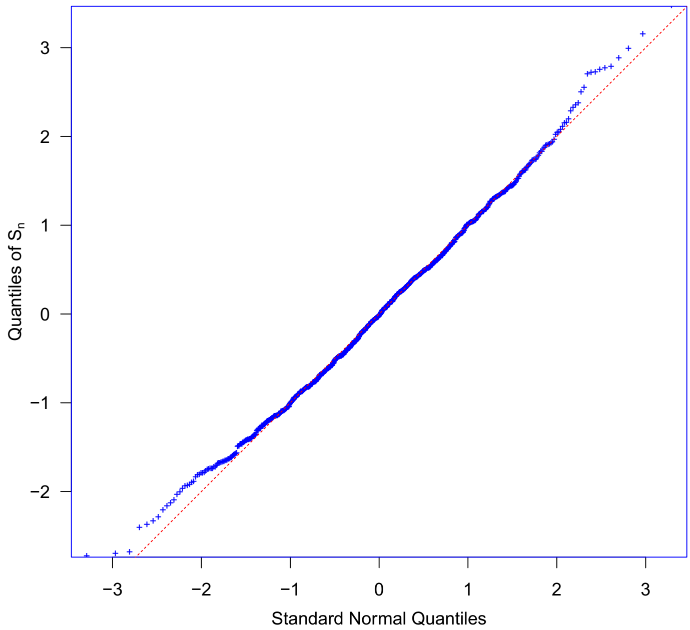

Fig. 9 800 sample with $m = 2$

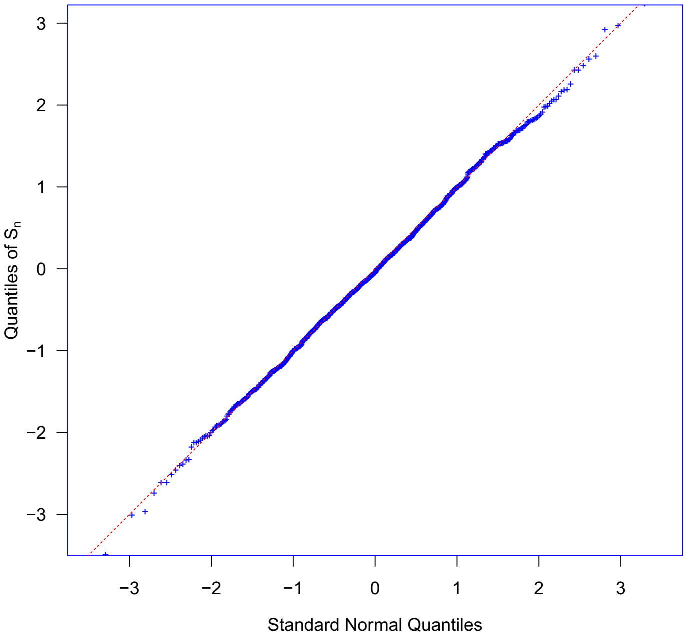

Fig. 10 200 sample with $m = 3$

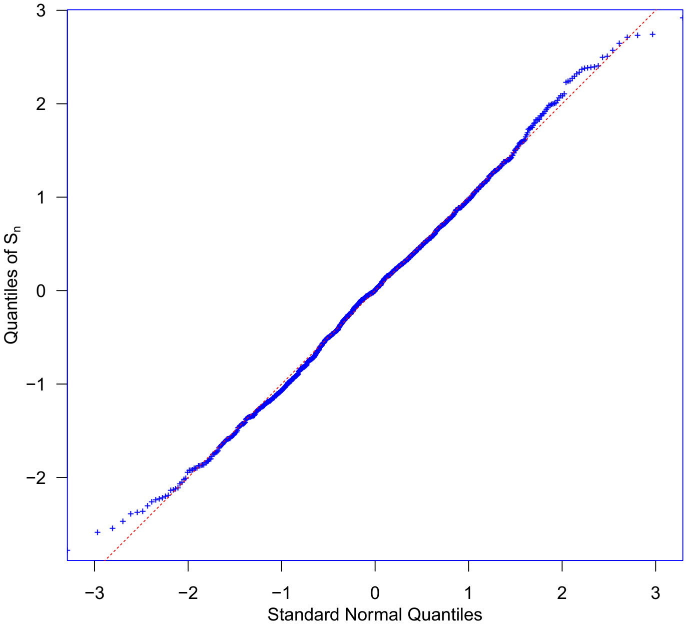

Fig. 11 500 sample with $m = 3$

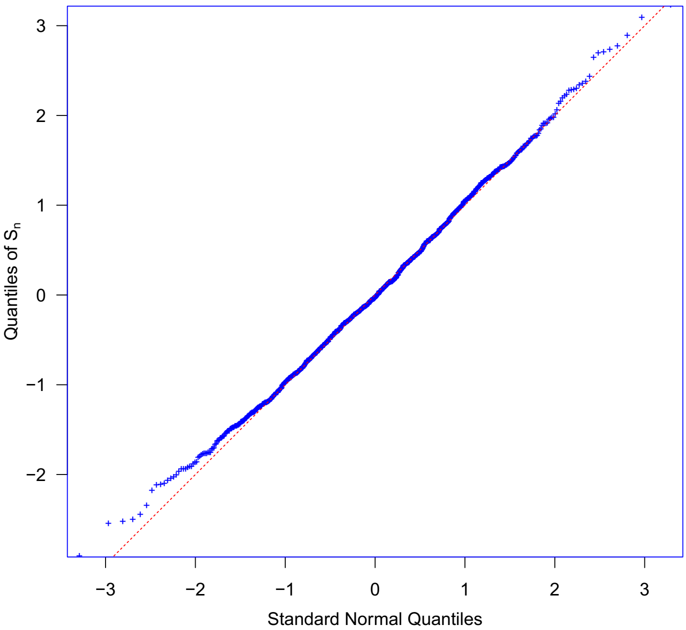

Fig. 12 800 sample with $m = 3$

Remark 5 From Lemma 1 and Definition 3, we can see that decreasing functions defined on disjoint subsets of a ρ^--mixing random variables with mixing coefficients $\rho^-(s)$ are also ρ^--mixing with mixing coefficients not greater than $\rho^-(s)$.

The following lemma can be found in Wang and Lu (2006).

Lemma 2 *Suppose that* $\{X_i, i \geq 1\}$ *is a sequence of* ρ^-*-mixing random variables with* $EX_i = 0$ *and* $E|X_i|^p < \infty$ *for some* $p \geq 2$. *Then there exists a positive constant* C_p *depending only on* p *and* $\rho^-(\cdot)$ *such that for all* $n \geq 1$,

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

Zhang (2000a) obtained the following lemma for ρ^--mixing random variables.

Lemma 3 *Suppose that* $f_1(x)$ *and* $f_2(y)$ *are real, bounded, absolutely continuous functions on* $\mathbb{R}$ *with* $|f_1^{'}(x)| \leq c_1$ *and* $|f_2^{'}(y)| \leq c_2$. *Then for any random variables* X *and* Y,

$$|Cov(f_1(x), f_2(y))| \leq c_1c_2\{-Cov(X, Y) + 8\rho^-(X, Y)\|X\|_{2,1}\|Y\|_{2,1}\}.$$

With Lemma 3 accounted for, we can obtain the following result, which plays an important role to prove the main result of the paper.

Table 1 The uniform Berry–Esseen bounds

Cases	$n = 200$	$n = 500$	$n = 800$
Case 1 with $\rho = 0.1$	0.0306	0.0280	0.0257
Case 1 with $\rho = 0.3$	0.0327	0.0277	0.0248
Case 2 with $m = 2$	0.0333	0.0271	0.0238
Case 2 with $m = 3$	0.0267	0.0226	0.0206

Lemma 4 *Let $\{X_i, i \geq 1\}$ be a sequence of ρ^--mixing random variables and $\{a_i, i \geq 1\}$ be a sequence of nonnegative (or nonpositive) constants. Denote by $Y_l = \sum_{i=(l-1)(p+q)+1}^{(l-1)(p+q)+p} a_i X_i$ for $1 \leq l \leq m$, where p and q are positive integers. Then for any $t \in \mathbb{R}$,*

$$\left|E\exp\left\{i\sum_{l=1}^{m} t_l Y_l\right\} - \prod_{l=1}^{m} E\exp\{it_l Y_l\}\right|$$
$$\leq 4\sum_{1\leq l\neq k\leq m} |t_l||t_k|\left\{-Cov(Y_l, Y_k) + 8\left\|\sum_{i=(l-1)(p+q)+1}^{(l-1)(p+q)+p} a_i X_i\right\|_{2,1}\left\|\sum_{j=(k-1)(p+q)+1}^{(k-1)(p+q)+p} a_j X_j\right\|_{2,1}\right.$$
$$\left.\rho^-(j-i)\right\}.$$

Proof Note that $\{Y_l, 1 \leq l \leq m\}$ is still a sequence of ρ^--mixing random variables by Lemma 1. By Lemma 3 we can easily prove the inequality above by adopting the method used in the proof of Theorem 3.3 in Zhang (2000a). The details are omitted. □

The following one can be found in Liang and Fan (2009) for instance.

Lemma 5 *Let $X, Y_1, \ldots, Y_m$ be random variable. For positive numbers $\omega_1, \ldots, \omega_m$, we have that*

$$\sup_u \left|P\left(X + \sum_{i=1}^{m} Y_i \leq u\right) - \Phi(u)\right| \leq \sup_u |P(X \leq u) - \Phi(u)| + \sum_{i=1}^{m} \frac{\omega_i}{\sqrt{2\pi}}$$
$$+ \sum_{i=1}^{m} P(|Y_i| > \omega_i).$$

To prove the main result, we need some notations as follows. For simplicity, we omit the argument x everywhere. Let $X_{ni} = \Delta_n^{-1}\omega_{ni}\varepsilon_{ni}$ for $i = 1, 2, \ldots, n$ and $n \geq 1$, and thus $S_n = \sum_{i=1}^{n} X_{ni}$. Let $k = \lfloor n/(p+q) \rfloor$. Then S_n can be decomposed as

$$S_n = S_n^{'} + S_n^{''} + S_n^{'''},$$

where

$$S_n^{'} = \sum_{m=1}^{k} y_{nm}, \quad S_n^{''} = \sum_{m=1}^{k} y_{nm}^{'}, \quad S_n^{'''} = \sum_{i=k(p+q)+1}^{n} X_{ni},$$

and $y_{nm}=\sum_{i=k_m}^{k_m+p-1}X_{ni}$, $y_{nm}^{'}=\sum_{i=l_m}^{l_m+q-1}X_{ni}$, $k_m=(m-1)(p+q)+1$, $l_m=(m-1)(p+q)+p+1$, $m=1,2,\ldots,k$. Denote $s_n^2=\sum_{m=1}^{k}\operatorname{Var}(y_{nm})$.

The following lemmas are the decompositions of the proof; we will state them one by one.

Lemma 6 *Suppose that Assumptions* $(\boldsymbol{A_1})-(\boldsymbol{A_2})$ *and* (3) *hold. If* $\sup_{i\geq 1}E|\xi_i|^{2+\kappa}<\infty$ *for some* $\kappa\geq 0$*, then*

$$E|S_n^{''}|^{2+\kappa}\leq C(nqp^{-1}\omega_n)^{1+\kappa/2},\quad E|S_n^{'''}|^{2+\kappa}\leq C(p\omega_n)^{1+\kappa/2}.$$

Proof It follows from $\omega_n(x)=O(\Delta_n^2(x))$ and Lemma 2 that

$$\begin{aligned}E|S_n^{''}|^{2+\kappa}&=E\left|\sum_{m=1}^{k}\sum_{i=l_m}^{l_m+q-1}\Delta_n^{-1}\omega_{ni}\varepsilon_{ni}\right|^{2+\kappa}\\&\leq C\sum_{m=1}^{k}\sum_{i=l_m}^{l_m+q-1}(\Delta_n^{-1}\omega_{ni})^{2+\kappa}E|\varepsilon_{ni}|^{2+\kappa}+C\left(\sum_{m=1}^{k}\sum_{i=l_m}^{l_m+q-1}(\Delta_n^{-1}\omega_{ni})^2E\varepsilon_{ni})^2\right)^{1+\kappa/2}\\&\leq C\sum_{m=1}^{k}\sum_{i=l_m}^{l_m+q-1}\omega_n^{1+\kappa/2}+C\left(\sum_{m=1}^{k}\sum_{i=l_m}^{l_m+q-1}\omega_n\right)^{1+\kappa/2}\leq C(nqp^{-1}\omega_n)^{1+\kappa/2},\end{aligned}$$

and similarly,

$$\begin{aligned}E|S_n^{'''}|^{2+\kappa}&=E\left|\sum_{i=k(p+q)+1}^{n}\Delta_n^{-1}\omega_{ni}\varepsilon_{ni}\right|^{2+\kappa}\\&\leq C\sum_{i=k(p+q)+1}^{n}\omega_n^{1+\kappa/2}+C\left(\sum_{i=k(p+q)+1}^{n}\omega_{ni}\right)^{1+\kappa/2}\leq C(p\omega_n)^{1+\kappa/2}.\end{aligned}$$

□

Lemma 7 *Under Assumptions* $(\boldsymbol{A_1})-(\boldsymbol{A_3})$*, we have*

$$|s_n^2-1|\leq C\left((nqp^{-1}\omega_n)^{1/2}+(p\omega_n)^{1/2}+u(q)\right).$$

Proof Let $\Psi_n=\sum_{1\leq i<j\leq n}\operatorname{Cov}(y_{ni},y_{nj})$. Thus $s_n^2=E\big(S_n^{'}\big)^2-2\Psi_n$. Noting that

$$\begin{aligned}E|S_n\big(S_n^{''}+S_n^{'''}\big)|&\leq E^{1/2}S_n^2E^{1/2}\big(S_n^{''}+S_n^{'''}\big)^2=E^{1/2}\big(S_n^{''}+S_n^{'''}\big)^2\\&\leq E^{1/2}\big(S_n^{''}\big)^2+E^{1/2}\big(S_n^{'''}\big)^2,\end{aligned}$$

and

$$E(S_n^{'})^2 = E\left(S_n - (S_n^{''} + S_n^{'''})\right)^2 = 1 + E(S_n^{''} + S_n^{'''})^2 - 2E\left(S_n(S_n^{''} + S_n^{'''})\right),$$

we have by Lemma 6 (taking $\kappa = 0$) that

$$|E(S_n^{'})^2 - 1| \leq C\left((nqp^{-1}\omega_n)^{1/2} + (p\omega_n)^{1/2}\right).$$

On the other hand, we have by $(\boldsymbol{A_2})$ and $(\boldsymbol{A_3})$ that

$$\begin{aligned} |\Psi_n| &\leq \sum_{1\leq i<j\leq n} \sum_{\mu=k_i}^{k_i+p-1} \sum_{\nu=k_j}^{k_j+p-1} \Delta_n^{-2}\omega_{n\mu}\omega_{n\nu}|\text{Cov}(\xi_\mu, \xi_\nu)| \\ &\leq C\sum_{i=1}^{k-1} \sum_{\mu=k_i}^{k_i+p-1} \omega_{n\mu} \sup_{t\geq 1} \sum_{\nu:|\nu-t|\geq q} |\text{Cov}(\xi_t, \xi_\nu)| \leq Cu(q). \end{aligned} \tag{4}$$

This completes the proof the lemma. □

Suppose that $\{\eta_{nm}, 1 \leq m \leq k\}$ are independent random variables and the distribution of η_{nm} is the same as that of y_{nm} for $m = 1, 2, \ldots, k$. Then $\sum_{m=1}^{k} \text{Var}(\eta_{nm}) = \sum_{m=1}^{k} \text{Var}(y_{nm}) = s_n^2$. Let $T_n = \sum_{m=1}^{k} \eta_{nm}$, $\tilde{F}_n(u)$, $G_n(u)$ and $\tilde{G}_n(u)$ be the distributions of $S_n^{'}$, T_n/s_n and T_n, respectively. Obviously,

$$\tilde{G}_n(u) = G_n(u/s_n).$$

Lemma 8 *Under the conditions of Theorem 1, we have*

$$\sup_u |G_n(u) - \Phi(u)| \leq C(p\omega_n)^{\delta/2}.$$

Proof It follows from Lemma 2 and Assumption $(\boldsymbol{A_2})$ that

$$\begin{aligned} \sum_{m=1}^{k} E|\eta_{nm}|^{2+\delta} &= C\sum_{m=1}^{k} E\left(\sum_{i=k_m}^{k_m+p-1} \Delta_n^{-1}\omega_{ni}\varepsilon_{ni}\right)^{2+\delta} \\ &\leq C\sum_{m=1}^{k} \sum_{i=k_m}^{k_m+p-1} (\Delta_n^{-1}\omega_{ni})^{2+\delta} E|\varepsilon_{ni}|^{2+\delta} + C\sum_{m=1}^{k}\left(\sum_{i=k_m}^{k_m+p-1} (\Delta_n^{-1}\omega_{ni})^2 E\varepsilon_{ni}^2\right)^{1+\delta/2} \\ &\leq C\omega_n^{\delta/2} \sum_{m=1}^{k} \sum_{i=k_m}^{k_m+p-1} \omega_{ni} + C(p\omega_n)^{\delta/2} \sum_{m=1}^{k} \sum_{i=k_m}^{k_m+p-1} \omega_{ni} \leq C(p\omega_n)^{\delta/2}. \end{aligned} \tag{5}$$

Moreover, by Lemma 7 we can obtain that $s_n \to 1$ as $n \to \infty$. Thus $1/s_n^{2+\delta} \sum_{m=1}^{k} E|\eta_{nm}|^{2+\delta} \leq C(p\omega_n)^{\delta/2}$, which derives the desired result directly by applying Berry–Esseen theorem. □

Lemma 9 *Under the conditions of Theorem 1, we have*

$$\sup_u |\tilde{F}_n(u) - \tilde{G}_n(u)| \le C\left((p\omega_n)^{\delta/2} + (u(q)+v(q))^{1/3}\right).$$

Proof Let $\varphi(t)$ and $\psi(t)$ be the characteristic functions of S_n' and T_n, respectively. Note that $\|X\|_{2,1} \le \frac{2+\delta}{\delta}\|X\|_{2+\delta}$ from (Ledoux and Talagrand 1991, p. 251). It follows from Lemma 2, Lemma 4 and some similar arguments as (4) and (5) that

$$\begin{aligned}
|\varphi(t) - \psi(t)| &= \left| E\exp\left\{it\sum_{m=1}^{k} y_{nm}\right\} - \prod_{m=1}^{k} E\exp\{ity_{nm}\}\right| \\
&\le 8t^2 \sum_{1\le l<m\le k} \left\{-\mathrm{Cov}(y_{nl}, y_{nm}) + 8\rho^-(q)\|y_{nl}\|_{2,1}\|y_{nm}\|_{2,1}\right\} \\
&\le 8t^2 \sum_{1\le l<m\le k} \sum_{\mu=k_l}^{k_l+p-1} \sum_{\nu=k_m}^{k_m+p-1} \Delta_n^{-2}\omega_{n\mu}\omega_{n\nu}|\mathrm{Cov}(\xi_\mu, \xi_\nu)| \\
&\quad + Ct^2\Delta_n^{-2} \sum_{1\le l<m\le k} \left\|\sum_{\mu=k_l}^{k_l+p-1} \omega_{n\mu}\varepsilon_{n\mu}\right\|_{2+\delta} \left\|\sum_{\nu=k_m}^{k_m+p-1} \omega_{n\nu}\varepsilon_{n\nu}\right\|_{2+\delta} \rho^-(\nu-\mu) \\
&\le Ct^2u(q) + Ct^2\Delta_n^{-2} \sum_{1\le l<m\le k} \sum_{\mu=k_l}^{k_l+p-1} \sum_{\nu=k_m}^{k_m+p-1} \omega_{n\mu}\omega_{n\nu}\rho^-(\nu-\mu) \\
&\le Ct^2u(q) + Ct^2 \sum_{l=1}^{m} \sum_{\mu=k_l}^{k_l+p-1} \omega_{n\mu} \sum_{i=q}^{\infty} \rho^-(i) \\
&\le Ct^2\left(u(q)+v(q)\right).
\end{aligned}$$

On the other hand, we have by Lemma 8 that

$$\begin{aligned}
\sup_u |\tilde{G}_n(u+s) - \tilde{G}_n(u)| &\le \sup_u \left| G_n\left(\frac{u+s}{s_n}\right) - \Phi\left(\frac{u+s}{s_n}\right)\right| \\
&\quad + \sup_u \left|\Phi\left(\frac{u+s}{s_n}\right) - \Phi\left(\frac{u}{s_n}\right)\right| \\
&\quad + \sup_u \left|\Phi\left(\frac{u}{s_n}\right) - G_n\left(\frac{u}{s_n}\right)\right| \\
&\le C\left((p\omega_n)^{\delta/2} + \frac{|s|}{s_n}\right) \le C\left((p\omega_n)^{\delta/2} + |s|\right).
\end{aligned}$$

Hence by Esseen inequality (see Pollard 1984), we have that for any $T > 0$ and some positive constant c,

$$\begin{aligned}
\sup_u |\tilde{F}_n(u) - \tilde{G}_n(u)| &\le \int_{-T}^{T} \left|\frac{\varphi(t)-\psi(t)}{t}\right| dt + T\sup_u \int_{|s|\le c/T} |\tilde{G}_n(u+s) - \tilde{G}_n(u)| \mathrm{d}s \\
&\le C\left((p\omega_n)^{\delta/2} + (u(q)+v(q))T^2 + \frac{1}{T}\right).
\end{aligned}$$

The desired result follows immediately by choosing $T = (u(q) + v(q))^{-1/3}$. □

Lemma 10 *Under the conditions of Theorem 1, we have*

$$P(|S_n^{''}| \geq \mu_n) \leq C\mu_n,$$
$$P(|S_n^{'''}| \geq \nu_n) \leq C\nu_n,$$

where $\mu_n = (nqp^{-1}\omega_n)^{(2+\delta)/(6+2\delta)}$ *and* $\nu_n = (p\omega_n)^{(2+\delta)/(6+2\delta)}$.

Proof It follows from Markov's inequality and Lemma 6 that

$$P(|S_n^{''}| \geq \mu_n) \leq \mu_n^{-2-\delta} E|S_n^{''}|^{2+\delta} \leq C\left(\frac{nqp^{-1}\omega_n}{(nqp^{-1}\omega_n)^{(4+2\delta)/(6+2\delta)}}\right)^{1+\delta/2} = C\mu_n,$$

and

$$P(|S_n^{'''}| \geq \nu_n) \leq \nu_n^{-2-\delta} E|S_n^{'''}|^{2+\delta} \leq C\left(\frac{p\omega_n}{(p\omega_n)^{(4+2\delta)/(6+2\delta)}}\right)^{1+\delta/2} = C\nu_n.$$

This completes the proof of the lemma. □

Now we present the proof of Theorem 1.

Proof of Theorem 1. By Lemmas 7, 8 and 9 we have that for some constant $0 < \theta < 1$,

$$\begin{aligned}
\sup_u |\tilde{F}_n(u) - \Phi(u)| &\leq \sup_u |\tilde{F}_n(u) - \tilde{G}_n(u)| + \sup_u |\tilde{G}_n(u) - \Phi(u/s_n)| \\
&\quad + \sup_u |\Phi(u/s_n) - \Phi(u)| \\
&\leq \sup_u |\tilde{F}_n(u) - \tilde{G}_n(u)| + \sup_u |\tilde{G}_n(u) - \Phi(u/s_n)| \\
&\quad + C \sup_u |u| e^{-[u+\theta(us_n-u)]^2/2} |s_n - 1| \\
&\leq \sup_u |\tilde{F}_n(u) - \tilde{G}_n(u)| + \sup_u |G_n(u) - \Phi(u)| + C|s_n^2 - 1| \\
&\leq C\left((nqp^{-1}\omega_n)^{1/2} + (p\omega_n)^{1/2} + (p\omega_n)^{\delta/2} + (u(q) + v(q))^{1/3}\right).
\end{aligned} \tag{6}$$

Thus, it follows from Lemma 5, Lemma 10 and (6) that

$$\begin{aligned}
\sup_u |F_n(u) - \Phi(u)| &\leq P(|S_n^{''}| \geq \mu_n) + P(|S_n^{'''}| \geq \nu_n) + \sup_u |\tilde{F}_n(u) - \Phi(u)| \\
&\quad + \frac{1}{\sqrt{2\pi}}(\mu_n + \nu_n) \\
&\leq C\Big\{(p\omega_n)^{\delta/2} + (u(q) + v(q))^{1/3} + (nqp^{-1}\omega_n)^{(2+\delta)/(6+2\delta)} \\
&\quad + (p\omega_n)^{(2+\delta)/(6+2\delta)}\Big\}.
\end{aligned}$$

The proof is completed. □

Acknowledgements The authors are most grateful to the Chief Editor, Associate Editor and two anonymous reviewers for carefully reading the manuscript and for valuable suggestions which helped in improving an earlier version of this paper.

References

Bradley, R. C. (1992). On the spectral density and asymptotic normality of weakly dependent random fields. *Journal of Theoretical Probability*, *5*, 355–373.

Bradley, R. C. (1997). Every "lower psi-mixing" Markov chain is "interlaced rho-mixing". *Stochastic Processes and their Applications*, *72*, 221–239.

Fan, Y. (1990). Consistent nonparametric multiple regression for dependent heterogeneous processes. *Journal of Multivariate Analysis*, *33*(1), 72–88.

Georgiev, A. A. (1985). Local properties of function fitting estimates with applications to system identification. In W. Grossmann, et al. (Eds.), *Proceedings 4th Pannonian symposium on mathematical statistics and applications, Bad Tatzmannsdorf, Austria* (Vol. B, pp. 141–151), September 4–10, 1983.

Hu, S. H., Zhu, C. H., Chen, Y. B., Wang, L. C. (2002). Fixed-design regression for linear time series. *Acta Mathematica Scientia, Series B*, *22*, 9–18.

Hu, S. H., Pan, G. M., Gao, Q. B. (2003). Estimation problems for a regression model with linear process errors. *Applied Mathematics*: *A Journal of Chinese Universities, Series A*, *18*(1), 81–90.

Joag-Dev, K., Proschan, F. (1983). Negative association of random variables with applications. *The Annals of Statistics*, *11*(1), 286–295.

Ledoux, M., Talagrand, M. (1991). *Probability in Banach space*. Now York: Springer.

Liang, H. Y., Fan, G. L. (2009). Berry–Esseen type bounds of estimators in a semiparametric model with linear process errors. *Journal of Multivariate Analysis*, *100*, 1–15.

Liang, H. Y., Jing, B. Y. (2005). Asymptotic properties for estimates of nonparametric regression models based on negatively associated sequences. *Journal of Multivariate Analysis*, *95*, 227–245.

Liu, X. D., Liu, J. X. (2009). Moments of the maximum of normed partial sums of ρ^--mixing random variables. *Applied Mathematics*: *A Journal of China Universities, Series B*, *24*(3), 355–360.

Pollard, D. (1984). *Convergence of stochastic processes*. Berlin: Springer.

Roussas, G. G. (1989). Consistent regression estimation with fixed design points under dependence conditions. *Statistics and Probability Letters*, *8*, 41–50.

Roussas, G. G., Tran, L. T., Ioannides, D. A. (1992). Fixed design regression for time series: Asymptotic normality. *Journal of Multivariate Analysis*, *40*, 262–291.

Shen, A. T. (2016). Complete convergence for weighted sums of END random variables and its application to nonparametric regression models. *Journal of Nonparametric Statistics*, *28*(4), 702–715.

Shen, A. T., Zhang, Y., Volodin, A. (2015). Applications of the Rosenthal-type inequality for negatively superadditive dependent random variables. *Metrika*, *78*, 295–311.

Stone, C. J. (1977). Consistent nonparametric regression. *The Annals of Statistics*, *5*, 595–620.

Tran, L., Roussas, G., Yakowitz, S., Van, B. T. (1996). Fixed-design regression for linear time series. *The Annals of Statistics*, *24*, 975–991.

Wang, J. F., Lu, F. B. (2006). Inequalities of maximum partial sums and weak convergence for a class of weak dependent random variables. *Acta Mathematica Sinica, English Series*, *22*(3), 693–700.

Wang, J. F., Zhang, L. X. (2007). A Berry–Esseen theorem and a law of the iterated logarithm for asymptotically negatively associated sequences. *Acta Mathematica Sinica, English Series*, *23*(1), 127–136.

Wang, X. J., Zheng, L. L., Xu, C., Hu, S. H. (2015). Complete consistency for the estimator of nonparametric regression models based on extended negatively dependent errors. *Statistics*: *A Journal of Theoretical and Applied Statistics*, *49*, 396–407.

Yang, S. C. (2003). Uniformly asymptotic normality of the regression weighted estimator for negatively associated samples. *Statistics and Probability Letters*, *62*(2), 101–110.

Yang, W. Z., Wang, X. H., Wang, X. J., Hu, S. H. (2012). The consistency for estimator of nonparametric regression model based on NOD errors. *Journal of Inequalities and Applications*, Article ID 140.

Zhang, L. X. (2000a). A functional central limit theorem for asymptotically negatively dependent random fields. *Acta Mathematica Hungarica*, *86*(3), 237–259.

Zhang, L. X. (2000b). Central limit theorems for asymptotically negatively associated random fields. *Acta Mathematica Sinica, English Series*, *16*(4), 691–710.

Zhang, L. X., Wang, X. Y. (1999). Convergence rates in the strong laws of asymptotically negatively associated random fields. *Applied Mathematics*: *A Journal of China Universities, Series B*, *14*(4), 406–416.