Research Article

A Note on the Central Limit Theorems for Dependent Random Variables

Yilun Shang

Institute for Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA

Correspondence should be addressed to Yilun Shang, shylmath@hotmail.com

Received 25 June 2012; Accepted 13 September 2012

Academic Editors: D. Fiems, A. Hutt, and M. Montero

Copyright © 2012 Yilun Shang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Classical central limit theorem is considered the heart of probability and statistics theory. Our interest in this paper is central limit theorems for functions of random variables under mixing conditions. We impose mixing conditions on the differences between the joint cumulative distribution functions and the product of the marginal cumulative distribution functions. By using characteristic functions, we obtain several limit theorems extending previous results.

1. Introduction

The central limit theorem is one of the most remarkable results of the theory of probability [1], which is critical to understand inferential statistics and hypothesis testing [2, 3]. The assumption of independence for a sequence of observations $X_1, X_2, \dots$ is often a technical convenience. Real data frequently exhibit some dependence and at the least some correlation at small lags. One extensively investigated kind of dependence is the m -dependence case (see e.g., [4–8]), in which random variables are considered as independent as long as they are m -step apart. More general measures of dependence are called mixing conditions, which are derived from the estimation of the difference between characteristic functions of averages of dependent and independent random variables. These conditions have some appealing physical interpretations. Various mixing conditions have been proposed by researchers; (see e.g., [9–14]) to name just a few.

Our main interest in this note is the central limit theorem for dependent classes of random variables. Following the work [12], instead of estimating the difference between the characteristic functions of the sum of dependent random variables and those of the sum of independent random variables of the same distributions, we compute the exact value of this difference. Our results weaken the conditions imposed on the random variable sequences in

Theorem 1 and Theorem 2 of [12] and can be used to describe systems which are globally determined but locally random. It is noteworthy that the work [12] has been extended in another direction, where the sum of a random number of random variables is examined [15].

The rest of the paper is organized as follows. In Section 2, we present our central limit theorems, and in Section 3, we give the proofs.

2. Main Results

Before proceeding, we introduce some notations. For $a, b \in \mathbb{R}$, denote by $a \vee b$ and $a \wedge b$ the maximal and minimal values between them, respectively. Let $X_1, X_2, \dots$ and Z be random variables. The image of Z is denoted by $\mathbb{I}(Z) \subseteq \mathbb{R}$. We write $X_n \xrightarrow{D} Z$ meaning that X_n converges in distribution to Z as $n \rightarrow \infty$. Denote by $N(0, 1)$ the standard normal variable. Let $E(Z)$ (or simply EZ) and $\text{Var}(Z)$ represent the mean and variance of Z , respectively.

Theorem 2.1. *Let $\{X_i\}_{i \geq 1}$ be a sequence of identically distributed random variables and $\{f_i\}_{i \geq 1}$ a sequence of measurable functions such that $f_i : \mathbb{I}(X_1) \rightarrow \mathbb{R}$ satisfying, for all $x \in \mathbb{I}(X_1)$, either*

$$0 \leq \alpha_1 x^{\gamma_1} + \beta_1 \leq f_i(x) \leq \alpha_2 x^{\gamma_2} + \beta_2, \quad (2.1)$$

or

$$\alpha_1 x^{\gamma_1} + \beta_1 \leq f_i(x) \leq \alpha_2 x^{\gamma_2} + \beta_2 \leq 0, \quad (2.2)$$

for some $\gamma_i > 0$ and $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, 2$. Assume that $\alpha_1 EX_1^{\gamma_1} + \beta_1 \geq 0$, $\alpha_2 EX_1^{\gamma_2} + \beta_2 \leq 0$, and $E|X_1|^{(\gamma_1 \vee \gamma_2)(2+\varepsilon)} < +\infty$ for some $\varepsilon > 0$. Let $0 < \varepsilon_1 < \varepsilon / (2(1 + \varepsilon))$ and $s_k^2 = \sum_{i=1}^k \text{Var}(f_i(X_i))$. If for sufficiently large k

$$\begin{aligned} & \sup \left\{ \left| P \left(\bigcap_{i=1}^j \{f_{v_i}(X_{v_i}) \leq x_{v_i}\} \right) - \prod_{i=1}^j P(f_{v_i}(X_{v_i}) \leq x_{v_i}) \right| : (x_{v_1}, \dots, x_{v_j}) \in \mathbb{R}^j \right\} \\ & \leq (1 - k^{-\varepsilon_1})^{k - k^{\varepsilon_1} - j}, \end{aligned} \quad (2.3)$$

where $v_1, \dots, v_j$ is any choice of indices such that $k^{\varepsilon_1} < v_1 < \dots < v_j \leq k$, then

$$\frac{1}{s_k} \sum_{i=1}^k (f_i(X_i) - E f_i(X_i)) \xrightarrow{D} N(0, 1) \quad (2.4)$$

as $k \rightarrow \infty$.

Corollary 2.2. *Let $\{X_i\}_{i \geq 1}$ be a sequence of identically distributed random variables and $\{f_i\}_{i \geq 1}$ a sequence of measurable functions such that $f_i : \mathbb{I}(X_1) \rightarrow \mathbb{R}$ satisfying, for all $x \in \mathbb{I}(X_1)$, expressions*

(2.1) or (2.2). Assume that $E|X_1|^n < +\infty$ for $n \in \mathbb{N}$. Suppose there exists $\delta, 0 < \delta < 1$, such that for sufficiently large k the inequality

$$\begin{aligned} & \sup \left\{ \left| P \left(\bigcap_{i=1}^j \{f_{v_i}(X_{v_i}) \leq x_{v_i}\} \right) - \prod_{i=1}^j P(f_{v_i}(X_{v_i}) \leq x_{v_i}) \right| : (x_{v_1}, \dots, x_{v_j}) \in \mathbb{R}^j \right\} \\ & \leq (1 - k^{-\delta/2})^{k-k^{\delta/2}-j} \end{aligned} \quad (2.5)$$

holds, where $v_1, \dots, v_j$ is any choice of indices such that $k^{\delta/2} < v_1 < \dots < v_j \leq k$, then

$$\frac{1}{s_k} \sum_{i=1}^k (f_i(X_i) - E f_i(X_i)) \xrightarrow{D} N(0, 1) \quad (2.6)$$

as $k \rightarrow \infty$.

Generally, $\{f_i\}_{i \geq 1}$ is a sequence of nonlinearly bounded functions. If we set $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 1$ and $\beta_1 = \beta_2 = 0$, then $f_i(x) = x$ for all i , which is the special case considered in [12] (Theorem 3). It is worth noting that in the conditions of Theorem 2.1 we used the expression $X_1^{\gamma_1}$, where γ_1 is a real number. Naturally, we preclude the situation where the random variable X_1 takes negative values and γ_1 is not an integer, since such kind of power is undefined.

Fix k and let j be relatively small, compared to k , then the right-hand side of (2.3) is close to zero. If we let j be close to $k - k^{\epsilon_1}$, then the right-hand side of (2.3) is close to 1. Hence, the dependence condition (2.3) allows a stronger dependence within a larger class of random variables and demands more independence within a smaller class. This structure can be used to describe systems which are globally determined but locally random.

Theorem 2.1 and Corollary 2.2 deal with dependence conditions on an “event-to-event” basis, while the following analogous results consider dependence on an “average” matter.

Theorem 2.3. Let $\{X_i\}_{i \geq 1}$ be a sequence of identically distributed random variables and $\{f_i\}_{i \geq 1}$ a sequence of measurable functions such that $f_i : \mathbb{I}(X_1) \rightarrow \mathbb{R}$ satisfying, for all $x \in \mathbb{I}(X_1)$, either

$$0 \leq \alpha_1 x^{\gamma_1} + \beta_1 \leq f_i(x) \leq \alpha_2 x^{\gamma_2} + \beta_2, \quad (2.7)$$

or

$$\alpha_1 x^{\gamma_1} + \beta_1 \leq f_i(x) \leq \alpha_2 x^{\gamma_2} + \beta_2 \leq 0, \quad (2.8)$$

for some $\gamma_i > 0$ and $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, 2$. Assume that $\alpha_1 EX_1^{\gamma_1} + \beta_1 \geq 0$, $\alpha_2 EX_1^{\gamma_2} + \beta_2 \leq 0$, and $E|X_1|^{(\gamma_1 \vee \gamma_2)(2+\varepsilon)} < +\infty$ for some $\varepsilon > 0$. Let $0 < \varepsilon_1 < \varepsilon/(2(1+\varepsilon))$ and $s_k^2 = \sum_{i=1}^k \text{Var}(f_i(X_i))$. If for sufficiently large k

$$\begin{aligned} \int_{\mathbb{R}^j} \left| P\left(\bigcap_{i=1}^j \{f_{v_i}(X_{v_i}) \leq x_{v_i}\}\right) - \prod_{i=1}^j P(f_{v_i}(X_{v_i}) \leq x_{v_i}) \right| dx_{v_1} \cdots dx_{v_j} \\ \leq (1 - k^{-\varepsilon_1})^{k-k^{\varepsilon_1}-j}, \end{aligned} \quad (2.9)$$

where $v_1, \dots, v_j$ is any choice of indices such that $k^{\varepsilon_1} < v_1 < \dots < v_j \leq k$, then

$$\frac{1}{s_k} \sum_{i=1}^k (f_i(X_i) - Ef_i(X_i)) \xrightarrow{D} N(0, 1) \quad (2.10)$$

as $k \rightarrow \infty$.

Corollary 2.4. Let $\{X_i\}_{i \geq 1}$ be a sequence of identically distributed random variables and $\{f_i\}_{i \geq 1}$ a sequence of measurable functions such that $f_i : \mathbb{I}(X_1) \rightarrow \mathbb{R}$ satisfying, for all $x \in \mathbb{I}(X_1)$, expressions (2.7) or (2.8). Assume that $E|X_1|^n < +\infty$ for $n \in \mathbb{N}$. Suppose there exists δ , $0 < \delta < 1$, such that for sufficiently large k the inequality

$$\begin{aligned} \int_{\mathbb{R}^j} \left| P\left(\bigcap_{i=1}^j \{f_{v_i}(X_{v_i}) \leq x_{v_i}\}\right) - \prod_{i=1}^j P(f_{v_i}(X_{v_i}) \leq x_{v_i}) \right| dx_{v_1} \cdots dx_{v_j} \\ \leq (1 - k^{-\delta/2})^{k-k^{\delta/2}-j} \end{aligned} \quad (2.11)$$

holds, where $v_1, \dots, v_j$ is any choice of indices such that $k^{\delta/2} < v_1 < \dots < v_j \leq k$, then

$$\frac{1}{s_k} \sum_{i=1}^k (f_i(X_i) - Ef_i(X_i)) \xrightarrow{D} N(0, 1) \quad (2.12)$$

as $k \rightarrow \infty$.

3. Proofs

We will prove Theorems 2.1 and 2.3 in this section through several lemmas, and the proof of corollaries follows directly.

For convenience, denote by $(\prod_{i=1}^k a_i)/(a_{v_1} \cdots a_{v_j})$, $1 \leq v_1 < \dots < v_j \leq k$, the product of all a_i 's except $a_{v_1}, \dots, a_{v_j}$. The following lemma regarding the difference between characteristic functions of sums of dependent and independent random variables is stated in [12] without proof.

Lemma 3.1. Let $X_1, X_2, \dots, X_k$ be random variables satisfying $0 \leq X_i \leq M$, $1 \leq i \leq k$. Let $g_1, \dots, g_k : [0, M] \rightarrow \mathbb{C}$ be absolutely continuous integrable functions. Then one has the identity

$$\begin{aligned} & E\left(\prod_{i=1}^k g_i(X_i)\right) - \prod_{i=1}^k E g_i(X_i) \\ &= \sum_{j=2}^k \sum_{1 \leq v_1 < \dots < v_j \leq k} (-1)^j \frac{\prod_{i=1}^k g_i(M)}{g_{v_1}(M) \cdots g_{v_j}(M)} \cdot \int_0^M \cdots \int_0^M \prod_{i=1}^j g'_{v_i}(x_{v_i}) \\ &\quad \times \left(P\left(\bigcap_{i=1}^j \{X_{v_i} \leq x_{v_i}\}\right) - \prod_{i=1}^j P(X_{v_i} \leq x_{v_i}) \right) dx_{v_1} \cdots dx_{v_j}. \end{aligned} \quad (3.1)$$

Proof. Let $F_i(x) = P(X_i \leq x)$ for $i \geq 1$. Therefore, for $k = 1$, we have

$$\begin{aligned} E(g_1(X_1)) &= \int_0^M g_1(x) dF_1(x) \\ &= g_1(M) - \int_0^M F_1(x) g'_1(x) dx \end{aligned} \quad (3.2)$$

by integration by parts. Let $F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$. Then, for $k = 2$, we derive from (3.2) that

$$\begin{aligned} E(g_1(X_1)g_2(X_2)) - E g_1(X_1)E g_2(X_2) &= \int \int_0^M g'_1(x_1)g'_2(x_2)F(x_1, x_2)dx_1dx_2 \\ &\quad - \int_0^M g'_1(x_1)F_1(x_1)dx_1 \int_0^M g'_2(x_2)F_2(x_2)dx_2 \\ &= \int \int_0^M g'_1(x_1)g'_2(x_2) \\ &\quad \cdot (F(x_1, x_2) - F_1(x_1)F_2(x_2))dx_1dx_2. \end{aligned} \quad (3.3)$$

The general results then follow by induction. □

The following lemma compares with the Lindeberg condition.

Lemma 3.2. Let $\{X_i\}_{i \geq 1}$ be a sequence of identically distributed random variables and $f_i : \mathbb{I}(X_1) \rightarrow \mathbb{R}$ satisfying, for all $x \in \mathbb{I}(X_1)$, expressions (2.1) or (2.2), $i \geq 1$. Assume that $\alpha_1 EX_1^{\gamma_1} + \beta_1 \geq 0$,

$\alpha_2 EX_1^{\gamma_2} + \beta_2 \leq 0$, and $E|X_1|^{(\gamma_1 \vee \gamma_2)(2+\varepsilon)} < +\infty$ for some $\varepsilon > 0$. Let $0 < \varepsilon_1 < \varepsilon/(2(1+\varepsilon))$. Then, there exists a sequence of reals $\{\varepsilon_k\}_{k \geq 1}$ such that

$$\varepsilon_k \longrightarrow 0, \quad \varepsilon_k s_k \longrightarrow +\infty \quad \text{as } k \longrightarrow \infty, \quad (3.4)$$

$$\frac{1}{s_k} \sum_{i=1}^k \int_{\{|f_i(X_i)| > \varepsilon_k s_k\}} |f_i(X_i)| \, dP \longrightarrow 0 \quad \text{as } k \longrightarrow \infty, \quad (3.5)$$

$$\frac{1}{s_k^2} \sum_{i=1}^k \int_{\{|f_i(X_i)| > \varepsilon_k s_k\}} f_i^2(X_i) \, dP \longrightarrow 0 \quad \text{as } k \longrightarrow \infty. \quad (3.6)$$

Proof. Without loss of generality, we can assume that $f_i(X_i)$ are centered at 0, that is, $E f_i(X_i) = 0$. In fact, let $f'_i := f_i - E f_i(X_1)$, and then $E f'_i(X_i) = 0$. By the assumptions (2.1) and (2.2), we have

$$-\infty < \alpha_1 EX_1^{\gamma_1} + \beta_1 \leq E f_i(X_1) \leq \alpha_2 EX_1^{\gamma_2} + \beta_2 < +\infty. \quad (3.7)$$

Therefore, $f'_i : \mathbb{I}(X_1) \rightarrow \mathbb{R}$ satisfies

$$\alpha_1 x^{\gamma_1} + \beta_1 - (\alpha_2 EX_1^{\gamma_2} + \beta_2) \leq f'_i(x) \leq \alpha_2 x^{\gamma_2} + \beta_2 - (\alpha_1 EX_1^{\gamma_1} + \beta_1). \quad (3.8)$$

For $x \in \mathbb{I}(X_1)$, if (2.1) holds, then $\alpha_1 x^{\gamma_1} + \beta_1 - (\alpha_2 EX_1^{\gamma_2} + \beta_2) \geq 0$; if (2.2) holds, then $\alpha_2 x^{\gamma_2} + \beta_2 - (\alpha_1 EX_1^{\gamma_1} + \beta_1) \leq 0$. Furthermore, by assumptions, we have $\alpha_1 EX_1^{\gamma_1} + \beta_1 - (\alpha_2 EX_1^{\gamma_2} + \beta_2) \geq 0$ and $\alpha_2 EX_1^{\gamma_2} + \beta_2 - (\alpha_1 EX_1^{\gamma_1} + \beta_1) \leq 0$. Hence, we may assume that $E f_i(X_i) = 0$ without loss of generality.

Choose ε_2 so that $\varepsilon_1 < \varepsilon_2 < \varepsilon/(2(1+\varepsilon))$. Define $\varepsilon_k = k^{-\varepsilon_2}$ for $k \geq 3$. We will show that the constructed sequence $\{\varepsilon_k\}_{k \geq 1}$ is what we want.

Since $s_k^2 = \sum_{i=1}^k \text{Var}(f_i(X_i))$ and X_i 's are identically distributed, by virtue of the bounds in (2.1) and (2.2) we obtain

$$kC_1^2 \leq s_k^2 \leq kC_2^2, \quad (3.9)$$

where $C_1^2 = E((\alpha_1^2 X_1^{2\gamma_1} + 2\alpha_1 \beta_1 X_1^{\gamma_1} + \beta_1^2) \wedge (\alpha_2^2 X_1^{2\gamma_2} + 2\alpha_2 \beta_2 X_1^{\gamma_2} + \beta_2^2))$ and $C_2^2 = E((\alpha_1^2 X_1^{2\gamma_1} + 2\alpha_1 \beta_1 X_1^{\gamma_1} + \beta_1^2) \vee (\alpha_2^2 X_1^{2\gamma_2} + 2\alpha_2 \beta_2 X_1^{\gamma_2} + \beta_2^2))$.

By our assumptions, we have

$$\varepsilon_k s_k \geq \frac{\sqrt{k}}{k^{\varepsilon_2}} C_1 \longrightarrow +\infty \quad (3.10)$$

as $k \rightarrow \infty$, which verifies (3.4). Let $\Delta(X_1) := (|\alpha_1||X_1|^{\gamma_1} + |\beta_1|) \vee (|\alpha_2||X_1|^{\gamma_2} + |\beta_2|)$. We have

$$\begin{aligned} & \frac{1}{s_k} \sum_{i=1}^k \int_{\{|f_i(X_i)| > \varepsilon_k s_k\}} |f_i(X_i)| dP \\ & \leq \frac{k}{s_k} \int_{\{\Delta(X_1) > \varepsilon_k s_k\}} \Delta(X_1) dP \leq \frac{k}{s_k} \left(\frac{1}{\varepsilon_k s_k} \right)^{1+\varepsilon} \int_{\{\Delta(X_1) > \varepsilon_k s_k\}} \Delta(X_1)^{2+\varepsilon} dP. \end{aligned} \quad (3.11)$$

Since $\varepsilon_2 < \varepsilon/(2(1+\varepsilon))$ and $C_1\sqrt{k} \leq s_k \leq C_2\sqrt{k}$, the quantity before the integration in (3.11) tends to zero. Consequently, (3.5) readily holds since the integration in (3.11) is bounded using our assumptions.

Now we want to verify (3.6). Note that

$$\begin{aligned} & \frac{1}{s_k^2} \sum_{i=1}^k \int_{\{|f_i(X_i)| > \varepsilon_k s_k\}} f_i^2(X_i) dP \leq \frac{k}{s_k^2} \int_{\{\Delta(X_1) > \varepsilon_k s_k\}} \Delta(X_1)^2 dP \\ & = O\left(\int_{\{\Delta(X_1) > \varepsilon_k s_k\}} \Delta(X_1)^2 dP \right). \end{aligned} \quad (3.12)$$

By the Markovian inequality (see e.g., [1]),

$$P\left(\Delta(X_1)^2 > \varepsilon_k^2 s_k^2\right) \leq \frac{E\Delta(X_1)^2}{\varepsilon_k^2 s_k^2} \rightarrow 0 \quad (3.13)$$

as $k \rightarrow \infty$. We have

$$E\Delta(X_1)^2 = \int_{\{\Delta(X_1)^2 \leq \varepsilon_k^2 s_k^2\}} \Delta(X_1)^2 dP + \int_{\{\Delta(X_1)^2 > \varepsilon_k^2 s_k^2\}} \Delta(X_1)^2 dP, \quad (3.14)$$

where the first integration in (3.14) tends to $E\Delta(X_1)^2 < \infty$ involving (3.13). Therefore, the second integration in (3.14) tends to zero, which, together with (3.12), finally verifies (3.6). $\square$

Proof of Theorem 2.1. We assume that $Ef_i(X_i) = 0$. For large enough k , we have

$$k^{\varepsilon_1} < \frac{k^{\varepsilon_2}}{\ln(1 + k^{\varepsilon_2})}. \quad (3.15)$$

In what follows, we suppose that k^{ε_1} is an integer for notation convenience. In the sequel, we take a specific sequence $\{\varepsilon_k\}_{k \geq 1}$ as in Lemma 3.2, and then (3.15) amounts to $k^{\varepsilon_1} < (\varepsilon_k \ln(1 + \varepsilon_k^{-1}))^{-1}$.

We define a sequence $\{Y_{k,j}\}_{1 \leq j \leq k-k^{\varepsilon_1}}$ of random variables as follows:

$$\begin{aligned} Y_{k,j} = & f_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j})I_{\{|f_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j})| \leq \varepsilon_k s_k\}} + \varepsilon_k s_k \\ & - \varepsilon_k s_k I_{\{f_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j}) < -\varepsilon_k s_k\}} + \varepsilon_k s_k I_{\{f_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j}) > \varepsilon_k s_k\}}. \end{aligned} \quad (3.16)$$

Hence, we have $0 \leq Y_{k,j} \leq 2\varepsilon_k s_k$ and $\{Y_{k,j} \leq x\} = \{f_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j}) \leq -\varepsilon_k s_k + x\}$ for $0 \leq x \leq 2\varepsilon_k s_k$. The inequality (2.3) then implies that, for large enough k ,

$$\begin{aligned} & \sup \left\{ \left| P \left(\bigcap_{i=1}^j \{Y_{k,v_i} \leq x_{v_i}\} \right) - \prod_{i=1}^j P(Y_{k,v_i} \leq x_{v_i}) \right| : (x_{v_1}, \dots, x_{v_j}) \in [0, 2\varepsilon_k s_k]^j \right\} \\ & \leq (1 - k^{-\varepsilon_1})^{k-k^{\varepsilon_1}-j}. \end{aligned} \quad (3.17)$$

For $k \geq 1$, denote $c_k^2 = \sum_{j=1}^{k-k^{\varepsilon_1}} \text{Var}(Y_{k,j})$. Involving Lemma 3.2, it follows that $c_k^2/s_k^2 \rightarrow 1$ as $k \rightarrow \infty$.

Next, we calculate the difference between characteristic functions of sums of independent and dependent random variables. Taking $g_j(\cdot) = \exp(it(\cdot/c_k))$ and $M = 2\varepsilon_k s_k$ in Lemma 3.1, where $i = \sqrt{-1}$, we obtain

$$\begin{aligned} & \left| E \left(\exp \left(\frac{it}{c_k} \sum_{j=1}^{k-k^{\varepsilon_1}} (Y_{k,j} - EY_{k,j}) \right) \right) - \prod_{j=1}^{k-k^{\varepsilon_1}} E \left(\exp \left(\frac{it}{c_k} (Y_{k,j} - EY_{k,j}) \right) \right) \right| \\ & \leq \left| \sum_{j=2}^{k-k^{\varepsilon_1}} \sum_{1 \leq v_1 < \dots < v_j \leq k-k^{\varepsilon_1}} (-1)^j \exp \left(2\varepsilon_k \frac{s_k}{c_k} it(k - k^{\varepsilon_1} - j) \right) \right. \\ & \quad \cdot \int_{[0, 2\varepsilon_k s_k]^j} \left(\frac{it}{c_k} \right)^j \exp \left(\frac{it}{c_k} (x_{v_1} + \dots + x_{v_j}) \right) \\ & \quad \cdot \left(P \left(\bigcap_{i=1}^j \{Y_{k,v_i} \leq x_{v_i}\} \right) - \prod_{i=1}^j P(Y_{k,v_i} \leq x_{v_i}) \right) dx_{v_1} \dots dx_{v_j} \Big| \\ & \leq c \sum_{j=2}^{k-k^{\varepsilon_1}} \binom{k-k^{\varepsilon_1}}{j} (1 - k^{-\varepsilon_1})^{k-k^{\varepsilon_1}-j} \left| 2tk^{-\varepsilon_2} \left(\frac{s_k}{c_k} \right) \right|^j \\ & \leq c \left(1 - k^{-\varepsilon_1} + 2|t|k^{-\varepsilon_2} \frac{s_k}{c_k} \right)^{k-k^{\varepsilon_1}}, \end{aligned} \quad (3.18)$$

where the second inequality comes from (3.17) and c is a positive constant. Since $\varepsilon_1 < \varepsilon_2$ and $s_k/c_k \rightarrow 1$ as $k \rightarrow \infty$, (3.18) tends to 0 as $k \rightarrow \infty$.

Define a sequence $\{W_j\}_{1 \leq j \leq k-k^{\varepsilon_1}}$ of independent random variables such that W_j has the same distribution with Y_j . By our construction, for sufficiently large k ,

$$\{ |Y_{k,j} - EY_{k,j}| = \varepsilon_k c_k \} = \{ |f_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j}) - Ef_{k^{\varepsilon_1}+j}(X_{k^{\varepsilon_1}+j})| = \varepsilon_k s_k \}. \quad (3.19)$$

Involving Lemma 3.2, it yields that $Y_{k,j}$ satisfies the Lindeberg condition. Accordingly, W_j also satisfies the Lindeberg condition. Thus, the central limit theorem is true for W_j since (3.18) approaches 0, which implies that the central limit theorem holds for $Y_{k,j}$, that is,

$$\frac{1}{c_k} \sum_{j=1}^{k-k^{\varepsilon_1}} (Y_{k,j} - EY_{k,j}) \xrightarrow{D} N(0, 1) \quad (3.20)$$

as $k \rightarrow \infty$. Taking into account that $s_k/c_k \rightarrow 1$ and the definition of $Y_{k,j}$, by (3.18) we know that

$$\frac{1}{s_k} \sum_{j=k^{\varepsilon_1}+1}^k f_j(X_j) \xrightarrow{D} N(0, 1). \quad (3.21)$$

On the other hand, it is obvious that

$$\begin{aligned} \frac{1}{s_k} \sum_{i=1}^{k^{\varepsilon_1}} \int |f_i(X_i)| dP &= \frac{1}{s_k} \sum_{i=1}^{k^{\varepsilon_1}} \int_{\{|f_i(X_i)| \leq \varepsilon_k s_k\}} |f_i(X_i)| dP \\ &\quad + \frac{1}{s_k} \sum_{i=1}^{k^{\varepsilon_1}} \int_{\{|f_i(X_i)| > \varepsilon_k s_k\}} |f_i(X_i)| dP. \end{aligned} \quad (3.22)$$

The first quantity on the right-hand side of (3.22) is up-bounded by

$$\frac{1}{s_k} \varepsilon_k s_k k^{\varepsilon_1} = \varepsilon_k k^{\varepsilon_1} \leq \frac{1}{\ln(1 + \varepsilon_k^{-1})} \rightarrow 0 \quad (3.23)$$

as $k \rightarrow \infty$, using (3.15). Combining (3.5), (3.22), and (3.23), we obtain

$$\frac{1}{s_k} \sum_{i=1}^{k^{\varepsilon_1}} f_i(X_i) \xrightarrow{L_1} 0 \quad (3.24)$$

as $k \rightarrow \infty$. Therefore, (3.21) and (3.24) imply that

$$\frac{1}{s_k} \sum_{i=1}^k f_i(X_i) \xrightarrow{D} N(0, 1), \quad (3.25)$$

which concludes the proof of Theorem 2.1. $\square$

Proof of Theorem 2.3. We define $Y_{k,j}$, c_k , and ε_k as in the proof of Theorem 2.1. We obtain

$$\begin{aligned} & \left| E \left(\exp \left(\frac{it}{c_k} \sum_{j=1}^{k-k^{\varepsilon_1}} (Y_{k,j} - EY_{k,j}) \right) \right) - \prod_{j=1}^{k-k^{\varepsilon_1}} E \left(\exp \left(\frac{it}{c_k} (Y_{k,j} - EY_{k,j}) \right) \right) \right| \\ & \leq c \sum_{j=2}^{k-k^{\varepsilon_1}} \binom{k-k^{\varepsilon_1}}{j} (1-k^{-\varepsilon_1})^{k-k^{\varepsilon_1}-j} \left| \frac{t}{c_k} \right|^j \\ & \leq c \left(1-k^{-\varepsilon_1} + \left| \frac{t}{c_k} \right| \right)^{k-k^{\varepsilon_1}}, \end{aligned} \quad (3.26)$$

where c is a positive constant. Since $s_k = \Theta(\sqrt{k})$ and $s_k/c_k \rightarrow 1$ as $k \rightarrow \infty$, the expression (3.26) tends to 0 as $k \rightarrow \infty$. The remaining part of the proof follows as in the proof of Theorem 2.1. $\square$

References

- [1] P. Billingsley, *Probability and Measure*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1995.
- [2] A. DasGupta, *Asymptotic Theory of Statistics and Probability*, Springer, New York, NY, USA, 2008.
- [3] T. S. Ferguson, *A Course in Large Sample Theory*, Chapman & Hall, New York, NY, USA, 1996.
- [4] K. N. Berk, "A central limit theorem for m -dependent random variables with unbounded m ," *Annals of Probability*, vol. 1, no. 2, pp. 352–354, 1973.
- [5] T. C. Christofides and P. M. Mavrikiou, "Central limit theorem for dependent multidimensionally indexed random variables," *Statistics & Probability Letters*, vol. 63, no. 1, pp. 67–78, 2003.
- [6] P. H. Diananda, "The central limit theorem for m -dependent variables," vol. 51, pp. 92–95, 1955.
- [7] W. Hoeffding and H. Robbins, "The central limit theorem for dependent random variables," *Duke Mathematical Journal*, vol. 15, pp. 773–780, 1948.
- [8] Y. Shang, "A central limit theorem for randomly indexed m -dependent random variables," *Filomat*, vol. 26, no. 4, pp. 713–717, 2012.
- [9] R. Balan and I.-M. Zamfirescu, "Strong approximation for mixing sequences with infinite variance," *Electronic Communications in Probability*, vol. 11, pp. 11–23, 2006.
- [10] J. Beran, "Statistical methods for data with long-range dependence (with discussion)," *Statistical Science*, vol. 7, no. 4, pp. 404–427, 1992.
- [11] R. C. Bradley, "Basic properties of strong mixing conditions. A survey and some open questions," *Probability Surveys*, vol. 2, pp. 107–144, 2005.
- [12] M. Kaminski, "Central limit theorem for certain classes of dependent random variables," *Theory of Probability and its Applications*, vol. 51, no. 2, pp. 335–342, 2007.
- [13] E. Ould-Saïd and A. Tatachak, "Strong consistency rate for the kernel mode estimator under strong mixing hypothesis and left truncation," *Communications in Statistics*, vol. 38, no. 8–10, pp. 1154–1169, 2009.
- [14] M. Rosenblatt, "A central limit theorem and a strong mixing condition," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 42, no. 1, pp. 43–47, 1956.
- [15] Y. Shang, "Central limit theorem for the sum of a random number of dependent random variables," *Asian Journal of Mathematics and Statistics*, vol. 4, no. 3, pp. 168–173, 2011.