Weak invariance principle for mixing sequences in the domain of attraction of normal law

R.M. Balan* R. Kulik †

July 16, 2007

Abstract

In this article we prove a weak invariance principle for a strictly stationary ϕ-mixing sequence $\{X_j\}_{j\geq 1}$, whose truncated variance function $L(x) := EX_1^2 1_{\{|X_1|\leq x\}}$ is slowly varying at ∞ and mixing coefficients satisfy the logarithmic growth condition: $\sum_{n\geq 1}\phi^{1/2}(2^n)<\infty$. This will be done under the condition that $\lim_n \mathrm{Var}(\sum_{j=1}^n \hat{X}_j)/[\sum_{j=1}^n \mathrm{Var}(\hat{X}_j)] = \beta^2$ exists in $(0,\infty)$, where $\hat{X}_j = X_j I_{\{|X_j|\leq \eta_j\}}$ and $\eta_n^2 \sim nL(\eta_n)$.

Keywords: weak invariance principle, mixing sequence, truncation

AMS Subject Classification Numbers: 60F05, 60F17

1 Introduction

Let $\{X_j\}_{j\geq 1}$ be a sequence of independent identically distributed random variables with $E(X_1)=0$, and let $S_n=\sum_{j=1}^n X_j$, $V_n^2=\sum_{j=1}^n X_j^2$. A long-standing conjecture of [10], which was recently proved in [9], states that the self-normalized central limit theorem:

$$\frac{S_n}{V_n} \xrightarrow{d} N(0,1) \tag{1}$$

holds true if and only if the distribution of X_1 lies in the domain of attraction of the normal law, i.e.

$$\text{(DAN)} \qquad L(x)=EX_1^2 1_{\{|X_1|\leq x\}} \text{ is slowly varying at } \infty$$

Other self-normalized fluctuation results for sequences of independent observations have been proved by various authors; we refer to [9] and the references

* *Postal address*: Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Avenue, Ottawa, ON, K1N 6N5, Canada. *E-mail address*: rbalan@uottawa.ca. Research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

† Corresponding author. *Postal address*: School of Mathematics and Statistics, University of Sydney, Australia, NSW, 2006. *E-mail address*: rkuli@maths.usyd.edu.au

therein. A common feature of all these results is that the distributional assumptions under which a self-normalized limit theorem would hold are in general milder than the assumptions of the corresponding classical limit theorem; in particular, most of these results do not require that the variance be finite.

The recent paramount result of [8] showed that if (DAN) holds, then it is possible to redefine the sequence $\{X_j\}_{j\geq 1}$ on a larger probability space together with a standard Brownian motion $W = \{W(t)\}_{t\in[0,\infty)}$ such that

$$\sup_{t\in[0,1]} \left| \frac{S_{[nt]}}{V_n} - \frac{W(nt)}{\sqrt{n}} \right| = o_P(1). \tag{2}$$

Recall that a sequence $\{X_j\}_{j\geq 1}$ of random variables is called ϕ-*mixing* if $\phi(n) := \sup_{k\geq 1} \phi(M_1^k, M_{k+n}^\infty) \to 0$ as $n \to \infty$, where M_a^b denotes the σ-field generated by $X_a, X_{a+1}, \ldots, X_b$ and

$$\phi(M_1^k, M_{k+n}^\infty) := \sup\{|P(B|A) - P(B)|; A \in M_1^k, B \in M_{k+n}^\infty\}.$$

In the present paper we prove that a weak invariance principle similar to (2) continues to hold under (DAN), in the case of a strictly stationary ϕ-mixing sequence $\{X_j\}_{j\geq 1}$ whose mixing coefficients $\phi(n)$ satisfy the "logarithmic rate" condition:

$$\text{(L)} \qquad \sum_{n\geq 1} \phi^{1/2}(2^n) < \infty.$$

To our knowledge, the only results that are available in the literature in this general framework are the central limit theorem of [6], its "functional" version found in [19], and a recent strong approximation result due to [1].

We note in passing that there is an immense amount of literature dedicated to (both classical and self-normalized) limit theorems for mixing sequences with *finite* variance; see [7] for an excellent review of the classical results, and [3], [14], [15], [16] for some self-normalized results.

We begin by noting that in the case of a ϕ-mixing sequence, one cannot obtain exactly (2), even if the variance is finite. To see this, assume that $\{X_j\}_{j\geq 1}$ is a strictly stationary ϕ-mixing sequence such that $EX_1 = 0$, $EX_1^2 = \sigma^2 < \infty$ and (L) holds. Let $S_n = \sum_{j=1}^n X_j$ and $\sigma_n^2 = E(S_n^2)$. Assume $\sigma_n^2 \to \infty$. The following facts are well-known:

$$\text{(i)}\ \frac{S_n}{\sigma_n} \xrightarrow{d} N(0,1) \quad \text{and} \quad \text{(ii)}\ \frac{V_n}{\sigma\sqrt{n}} \xrightarrow{P} 1$$

((i) was proved in [11], while (ii) can be proved using standard methods). Therefore the fact that

$$\lim_{n\to\infty} \frac{\sigma_n}{\sigma\sqrt{n}} := \beta \quad \text{exists in } (0,\infty)$$

(which was proved in [5] using a procedure specific to the finite variance case), allows us to conclude that the central limit theorem for a ϕ-mixing sequence with finite variance should be of the form

$$\frac{S_n}{\beta V_n} \xrightarrow{d} N(0,1). \tag{3}$$

In the present paper, we adapt this method of identifying a suitable normalizer β for which (3) holds, to the case when the variance may not exist, but (DAN) holds. Our method relies on the truncation technique of [8] and [1], which differs slightly from that of [6] and [19]. More precisely, let

$$\eta_n = \inf \left\{ s \geq 1; \frac{L(s)}{s^2} \leq \frac{1}{n} \right\} \tag{4}$$

and note that $\eta_n^2 \sim nL(\eta_n)$. We define

$$\hat{X}_j = X_j I_{\{|X_j| \leq \eta_j\}}, \quad \hat{S}_n = \sum_{j=1}^n \hat{X}_j, \quad A_n^2 = \mathrm{Var}(\hat{S}_n), \quad B_n^2 = \sum_{j=1}^n \mathrm{Var}(\hat{X}_j).$$

The following result lies at the origin of our developments: its part (a) is an immediate consequence of the central limit theorem of [6], mentioned above (details are given in the Appendix); part (b) is our Proposition 2.3 (Section 2).

Theorem 1.1 *Let $\{X_j\}_{j\geq 1}$ be a strictly stationary sequence of nondegenerate random variables such that $EX_1 = 0$ and (DAN) holds. Let $S_n = \sum_{i=1}^n X_i$. Suppose that $\phi(1) < 1/4$ and the mixing coefficients satisfy (L). Then*

$$(a)\ \frac{S_n}{A_n} \xrightarrow{d} N(0,1) \quad \textit{and} \quad (b)\ \frac{V_n}{B_n} \xrightarrow{P} 1.$$

An immediate consequence of the previous theorem is that if

$$\text{(C)} \qquad \lim_{n\to\infty} \frac{A_n}{B_n} := \beta \quad \text{exists in } (0,\infty),$$

then relation (3) holds. (It is also clear that if (3) holds for a certain constant $\beta > 0$, then $\lim_n A_n/B_n$ exists in $(0,\infty)$ and has to be equal to β.)

Using standard techniques, it can be proved that condition (C) holds true if the mixing coefficients satisfy the following "polynomial rate" condition:

$$\text{(P)} \qquad \sum_{n\geq 1} \phi^{1/2}(n) < \infty$$

(see [2] for details). Proving that condition (C) holds true in the case of a ϕ-mixing sequence whose coefficients have infinite variance and satisfy (L) remains an open problem, which we do not attempt to solve in the present paper.

The next example shows how to construct a sequence $\{X_j\}_{j\geq 1}$ with infinite variance which satisfies the conditions of Theorem 1.1, and for which (C) holds, but (P) fails.

Example. Let $\{Y_j\}_{j\geq 1}$ be a sequence which satisfies the conditions of Theorem 1.1, $\{\varepsilon_j\}_{j\geq 1}$ be an independent sequence of i.i.d. random variables with $P(\varepsilon_1 = 1) = P(\varepsilon_1 = -1) = 1/2$, and $X_j = \varepsilon_j Y_j, j \geq 1$. Then the sequence $\{X_j\}_{j\geq 1}$ also satisfies the conditions of Theorem 1.1 (see Theorem 5.2.(d), [7]).

With $\{\eta_j\}_{j\geq 1}$ given by (4), let $\hat{X}_j = X_j I_{\{|X_j|\leq \eta_j\}}$ and $\hat{Y}_j = Y_j I_{\{|Y_j|\leq \eta_j\}}$. Then $\hat{X}_j = \varepsilon_j \hat{Y}_j$, and hence $E\hat{X}_j = E(\varepsilon_j \hat{Y}_j) = (E\varepsilon_j)E\hat{Y}_j = 0$ for all j, and

$$E(\hat{X}_i \hat{X}_j) = E(\varepsilon_i \varepsilon_j \hat{Y}_i \hat{Y}_j) = (E\varepsilon_i)(E\varepsilon_j)E(\hat{Y}_i \hat{Y}_j) = 0 \quad \text{for all} \quad i \neq j.$$

Therefore $A_n^2 = E(\sum_{j=1}^n \hat{X}_j)^2 = \sum_{j=1}^n E\hat{X}_j^2 = B_n^2$, and (C) holds (with $\beta = 1$).

We are now ready to state our main result.

Theorem 1.2 *Let $\{X_j\}_{j\geq 1}$ be a sequence of random variables satisfying the conditions of Theorem 1.1. If (C) holds, then without changing its distribution, we can redefine the sequence $\{X_j\}_{j\geq 1}$ on a larger probability space together with a standard Brownian motion $W = \{W(t)\}_{t\geq 0}$ such that for some suitable constants s_k^2 we have*

$$\sup_{t\in[0,1]} \left| \frac{S_{[nt]}}{\beta V_n} - \frac{W(s_{[nt]}^2)}{s_n} \right| = o_P(1).$$

The remaining part of the paper is dedicated to the proof of Theorem 1.2.

The argument is based on the idea of replacing the original sequence with the truncated sequence $\{\hat{X}_j\}_{j\geq 1}$, approximating the random variable V_n^2 by B_n^2, establishing the weak invariance principle for the sequence $\{\hat{X}_j/B_n\}_{j=1,\ldots,n}$ and then proving that $\beta B_n \sim s_n$. Indeed, we have the following decomposition:

$$\max_{k\leq n} \left| \frac{S_k}{\beta V_n} - \frac{W(s_k^2)}{s_n} \right| \leq \max_{k\leq n} \left| \frac{S_k}{\beta V_n} - \frac{\hat{S}_k - E\hat{S}_k}{\beta V_n} \right| + \max_{k\leq n} \left| \frac{\hat{S}_k - E\hat{S}_k}{\beta V_n} - \frac{\hat{S}_k - E\hat{S}_k}{\beta B_n} \right| +$$

$$\max_{k\leq n} \left| \frac{\hat{S}_k - E\hat{S}_k}{\beta B_n} - \frac{W(s_k^2)}{\beta B_n} \right| + \max_{k\leq n} \left| \frac{W(s_k^2)}{\beta B_n} - \frac{W(s_k^2)}{s_n} \right| := J_1(n) + J_2(n) + J_3(n) + J_4(n).$$

In the next three sections we treat separately each of the four terms. We should point out that condition (C) is used *only* for proving that the last term is negligible.

Remark on Notation: Throughout this article, we write $a_n \sim b_n$ if $\lim_n a_n/b_n = 1$. We denote by $|S|$ the cardinality of the set S, and by $[x]$ the integer part of the real number x. We denote by C a generic constant that may be different in each of its appearances. We let $\| X \| = (EX^2)^{1/2}$, for a random variable X.

2 The first two terms

In this section we treat the terms $J_1(n)$ and $J_2(n)$, using some properties of the "tail" variables:

$$\bar{X}_j = X_j I_{\{|X_j|>\eta_j\}}, \quad \bar{S}_n = \sum_{j=1}^n \bar{X}_j.$$

Lemma 2.1 *Under the conditions of Theorem 1.1, $\sum_{j=1}^n E|\bar{X}_j| = o(\eta_n)$.*

Proof: We write

$$\sum_{j=1}^{n} E|\bar{X}_j| = nE|X_1|1_{\{|X_1|>\eta_n\}} + \sum_{j=1}^{n} E|X_1|1_{\{\eta_j<|X_j|\leq\eta_n\}}.$$

We have $nE|X_1|1_{\{|X_1|>\eta_n\}} \leq C\eta_n^2 L^{-1}(\eta_n)E|X_1|1_{\{|X_1|>\eta_n\}} = o(\eta_n)$, by Lemma 1.(c), [8], and $\sum_{j=1}^{n} E|X_1|1_{\{\eta_j<|X_1|\leq\eta_n\}} = o(\eta_n)$, by relation (20) of [8]. □

Lemma 2.2 *Under the conditions of Theorem 1.1, $B_n^2 \sim \eta_n^2$.*

Proof: We write

$$B_n^2 = \sum_{j=1}^{n}(E\hat{X}_j^2 - (E\bar{X}_j)^2) = \sum_{j=1}^{n} L(\eta_j) - \sum_{j=1}^{n}(E\bar{X}_j)^2.$$

We have $\sum_{j=1}^{n} L(\eta_j) \sim nL(\eta_n)$ by relation (9) of [8], and $\sum_{j=1}^{n}(E\bar{X}_j)^2 \leq (\sum_{j=1}^{n} |E\bar{X}_j|)^2 \leq (\sum_{j=1}^{n} E|\bar{X}_j|)^2 = o(\eta_n^2)$ by Lemma 2.1. □

Proposition 2.3 *Under the conditions of Theorem 1.1, we have*

$$\frac{V_n^2}{B_n^2} \xrightarrow{P} 1.$$

Proof: Note that

$$\frac{V_n^2}{B_n^2} - 1 = \frac{1}{B_n^2}\sum_{j=1}^{n}(\hat{X}_j^2 - E\hat{X}_j^2) + \frac{1}{B_n^2}\sum_{j=1}^{n}\bar{X}_j^2 + \frac{1}{B_n^2}\sum_{j=1}^{n}(E\bar{X}_j)^2.$$

By Lemma 2.1 and Lemma 2.2, it suffices to prove that

$$\frac{1}{\eta_n^2}\sum_{j=1}^{n}(\hat{X}_j^2 - E\hat{X}_j^2) \xrightarrow{L^2} 0.$$

For this, we note that $\{\hat{X}_j^2 - E\hat{X}_j^2\}_{j\geq 1}$ is a ϕ-mixing sequence with mixing coefficient $\hat{\phi}(n) \leq \phi(n)$. Using Lemma 2.3, [18] and Lemma 1.(d), [8], we get

$$\frac{1}{\eta_n^4}E\left(\sum_{j=1}^{n}(\hat{X}_j^2 - E\hat{X}_j^2)\right)^2 \leq C\frac{n}{\eta_n^4}\max_{j\leq n} E(\hat{X}_j^2 - E\hat{X}_j^2)^2 \leq C\frac{n}{\eta_n^4}E\hat{X}_n^4 = o(1)$$

□

By Lemma 2.1, Lemma 2.2 and Proposition 2.3 we get

$$J_1(n) = \frac{1}{\beta V_n}\max_{k\leq n}|\bar{S}_k - E\bar{S}_k| \leq \frac{B_n}{V_n}\cdot\frac{1}{\beta B_n}\sum_{j=1}^{n}(|\bar{X}_j| - E|\bar{X}_j|) = o_P(1).$$

We have

$$J_2(n) = \left|\frac{B_n}{V_n} - 1\right|\max_{k\leq n}\frac{|\hat{S}_k - E\hat{S}_k|}{\beta B_n} \leq \left|\frac{B_n}{V_n} - 1\right|\left(J_3(n) + J_4(n) + \max_{k\leq n}\frac{|W(s_k^2)|}{s_n}\right)$$

and hence $J_2(n) = o_P(1)$, provided that $J_3(n) = o_p(1)$ and $J_4(n) = o_P(1)$. This will be proved in Section 3, respectively Section 4.

3 The third term

To treat the term $J_3(n)$, we use the blocking technique of [18]. More precisely, let $a \in (1/2, 1)$ be fixed and $H_1, I_1, H_2, I_2, \dots$ be consecutive blocks of integers such that

$$|H_i| = [ai^{a-1} \exp(i^a)] \quad \text{and} \quad |I_i| = [ai^{a-1} \exp(i^a/2)].$$

We denote with u_i and v_i the sums of the (centered) truncated variables over the block H_i, respectively I_i, i.e.

$$u_i = \sum_{j \in H_i} (\hat{X}_j - E\hat{X}_j), \quad v_i = \sum_{j \in I_i} (\hat{X}_j - E\hat{X}_j).$$

Let $N_m = \sum_{i=1}^m |H_i \cup I_i| \sim \exp(m^a)$. Clearly, for each n there exists a unique m_n such that $N_{m_n} \le n < N_{m_n+1}$; we have $m_n \sim (\log n)^a$. Note that

$$\hat{S}_k - E\hat{S}_k = \sum_{i=1}^{m_k} u_i + \sum_{i=1}^{m_k} v_i + \sum_{j=N_{m_k}+1}^{k} (\hat{X}_j - E\hat{X}_j). \tag{5}$$

Let

$$\sigma_i^{*2} = Eu_i^2, \quad s_m^{*2} = \sum_{i=1}^m \sigma_i^{*2}, \quad s_n^2 = s_{m_n}^{*2}.$$

The idea is to approximate the sequence $\{u_i\}_{i\ge1}$ by a sequence $\{Y_i^*\}_{i\ge1}$ of independent Gaussian random variables (with the same variance), and then to prove that the remaining terms are negligible (in probability).

The desired approximation will be achieved in two steps, by using a classical result of Berkes and Philipp [4], combined with a more recent result of Sakhanenko [17].

Proposition 3.1 *Under the hypothesis of Theorem 1.1, without changing its distribution, we can redefine the sequence $\{u_i\}_{i\ge1}$ on a larger probability space together with a sequence $\{Y_i\}_{i\ge1}$ of independent random variables such that Y_i has the same distribution as u_i and for all $m \ge 1$*

$$|\sum_{i=1}^m u_i - \sum_{i=1}^m Y_i| \le C \quad a.s. \tag{6}$$

Proof: We apply Theorem 2, [4] with $X_k = u_k$, $L_k = \sigma(u_k)$ and

$$\phi_k := \sup_{A\in\sigma(u_1,\dots u_{k-1}); B\in\sigma(u_k)} |P(B|A) - P(B)| \le \phi(|I_k|) \le \phi(e^{k^a/2}).$$

We conclude that without changing its distribution, we can redefine the sequence $\{u_k\}_{k\ge1}$ together with a sequence $\{Y_k\}_{k\ge1}$ of independent random variables such that Y_k has the same distribution as u_k and for all $k \ge 1$

$$P(|u_k - Y_k| \ge 6\phi_k) \le 6\phi_k.$$

Since (L) holds, $n\phi^{1/2}(2^n) = o(1)$, and in particular $\phi(2^n) \leq Cn^{-2}$. Hence

$$\sum_{k\geq 1} \phi_k \leq \sum_{k\geq 1} \phi(e^{k^a/2}) \leq \sum_{k\geq 1} \phi(2^{k^a/2}) \leq C \sum_{k\geq 1} k^{-2a} < \infty,$$

since $a > 1/2$. By the Borel-Cantelli lemma, we conclude that

$$|u_k - Y_k| \leq C\phi_k \quad \text{a.s.},$$

and hence $|\sum_{i=1}^m u_i - \sum_{i=1}^m Y_i| \leq \sum_{i=1}^m |u_i - Y_i| \leq C\sum_{i=1}^m \phi_i \leq C$ a.s. $\square$

Theorem 3.2 (Theorem B, [20]) *Without changing its distribution we can redefine the sequence $\{Y_i\}_{i\geq 1}$ on a larger probability space together with a sequence $\{Y_i^*\}_{i\geq 1}$ of independent normal random variables with $EY_i^* = 0, EY_i^{*2} = \sigma_i^{*2}$ such that for every M and for every $x > 0, \delta > 0$*

$$P\left(\max_{m\leq M} |\sum_{i=1}^m Y_i - \sum_{i=1}^m Y_i^*| > x\right) \leq C\frac{1}{x^{2+\delta}} \sum_{i=1}^m E|Y_i|^{2+\delta}. \tag{7}$$

It is not difficult to see that, without changing its distribution we can redefine the sequence $\{Y_i^*\}_{i\geq 1}$ on a larger probability space together with a standard Brownian motion $W = \{W(t)\}_{t\geq 0}$ such that $W(s_m^{*2}) = \sum_{i=1}^m Y_i^*$ for any m. In particular

$$W(s_k^2) = W(s_{m_k}^{*2}) = \sum_{i=1}^{m_k} Y_i^*. \tag{8}$$

We are now ready to treat the term $J_3(n)$. Let $\varepsilon > 0$ be arbitrary. Using the decomposition (5) and relation (8), we have

$$\begin{aligned}
P\left(J_3(n) > \varepsilon/\beta\right) &= P\left(\max_{k\leq n} |\hat{S}_k - E\hat{S}_k - W(s_k^2)| > \varepsilon B_n\right) \\
&\leq P\left(\max_{m\leq m_n} \max_{N_m\leq k<N_m+1} |\sum_{i=1}^m u_i + \sum_{i=1}^m v_i + \sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j) - \sum_{i=1}^m Y_i^*| > \varepsilon B_n\right) \\
&\leq P\left(\max_{m\leq m_n} |\sum_{i=1}^m u_i - \sum_{i=1}^m Y_i| > \frac{\varepsilon B_n}{4}\right) + P\left(\max_{m\leq m_n} |\sum_{i=1}^m v_i| > \frac{\varepsilon B_n}{4}\right) + \\
&\quad P\left(\max_{m\leq m_n} \max_{N_m\leq k<N_{m+1}} |\sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j)| > \frac{\varepsilon B_n}{4}\right) + \\
&\quad P\left(\max_{m\leq m_n} |\sum_{i=1}^m Y_i - \sum_{i=1}^m Y_i^*| > \frac{\varepsilon B_n}{4}\right) := P_1(n) + P_2(n) + P_3(n) + P_4(n).
\end{aligned}$$

Using (6), we have $P_1(n) = 0$ for n large. The following results will show that $\lim_{n\to\infty} P_i(n) = 0$ for $i = 2, 3, 4$. This will conclude the proof of $J_3(n) = o_P(1)$.

Lemma 3.3 *Under the conditions of Theorem 1.1,*

$$P_2(n) := P\left(\max_{m\leq m_n} |\sum_{i=1}^m v_i| > \frac{\varepsilon B_n}{4}\right) \to 0 \quad \text{as} \quad n \to \infty.$$

Proof: Note that $\sum_{i=1}^{m_n} |I_i| \le C n^{1/2}$. Using Lemma 2.3, [18], for every $m \le m_n$,

$$\sum_{i=1}^{m} E v_i^2 \le C \sum_{i=1}^{m} |I_i| \max_{j \in I_i} E\hat{X}_j^2 \le C L(\eta_n) \sum_{i=1}^{m} |I_i| \le C L(\eta_n) n^{1/2} \tag{9}$$

and $E\left(\sum_{i=1}^m v_i\right)^2 \le m \sum_{i=1}^m E v_i^2 \le C m_n L(\eta_n) n^{1/2}$.

Note that the sequence $\{v_i\}_{i \ge 1}$ is ϕ-mixing with coefficient $\phi^{(v)}(n) \le \phi(e^{n^a})$. Using Proposition 3.2 [12], we get

$$E \max_{m \le m_n} \left(\sum_{i=1}^m v_i \right)^2 \le C \max_{m \le m_n} E(\sum_{i=1}^m v_i)^2 \le C m_n L(\eta_n) n^{1/2} = o(\eta_n^2). \tag{10}$$

The result follows by the Chebyshev's inequality. □

Lemma 3.4 *Under the conditions of Theorem 1.1,*

$$P_3(n) := P\left(\max_{m \le m_n} \max_{N_m < k \le N_{m+1}} |\sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j)| > \frac{\varepsilon B_n}{4} \right) \longrightarrow 0 \quad \text{as } n \to \infty.$$

Proof: By Markov's inequality we get: for any $\delta > 0$

$$\begin{aligned} P_3(n) &\le \sum_{m=1}^{m_n} P\left(\max_{N_m < k \le N_{m+1}} |\sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j)| > \frac{\varepsilon B_n}{4} \right) \\ &\le \frac{C}{\eta_n^{2+\delta}} \sum_{m=1}^{m_n} E\left(\max_{N_m < k \le N_{m+1}} |\sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j)| \right)^{2+\delta}. \end{aligned}$$

By Proposition 3.2, [12], Lemma 2.3, [18] and Lemma 1.(d), [8], for every $m \le m_n$

$$\begin{aligned} &E\left(\max_{N_m < k \le N_{m+1}} |\sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j)| \right)^{2+\delta} \le C \max_{N_m < k \le N_{m+1}} E\left| \sum_{j=N_m+1}^{k} (\hat{X}_j - E\hat{X}_j) \right|^{2+\delta} \\ &\le C\{(N_{m+1} - N_m)^{1+\delta/2} L(\eta_{N_m})^{1+\delta/2} + (N_{m+1} - N_m) o(\eta_{N_m}^{\delta} L(\eta_{N_m}))\} \\ &\le C\{2|H_m|^{1+\delta/2} L(\eta_{N_m})^{1+\delta/2} + 2|H_m| o(\eta_{N_m}^{\delta} L(\eta_{N_m}))\}. \end{aligned}$$

Hence

$$P_3(n) \le \frac{C L(\eta_n)^{1+\delta/2}}{\eta_n^{2+\delta}} \sum_{m=1}^{m_n} |H_m|^{1+\delta/2} + \frac{C}{\eta_n^{2+\delta}} \sum_{m=1}^{m_n} |H_m| o(\eta_{N_m}^{\delta} L(\eta_{N_m})) := P_3'(n) + P_3''(n).$$

Note that

$$\begin{aligned} \sum_{m=1}^{m_n} |H_m|^{1+\delta/2} &\le C \sum_{m=1}^{m_n} m^{(a-1)(2+\delta)/2} e^{(2+\delta)m^a/2} \\ &= \sum_{m=1}^{m_n} o(m^{a-1} e^{(2+\delta)m^a/2}) = o(e^{(2+\delta)m_n^a/2}) = o(n^{1+\delta/2}) \end{aligned}$$

and hence $P_3'(n) = o(1)$. For the second part we use $\sum_{i=1}^{m_n} |H_i| \le Cn$. Hence

$$P_3''(n) \le \frac{C}{\eta_n^{2+\delta}} o(\eta_{N_{m_n}}^{\delta} L(\eta_{N_{m_n}})) \sum_{m=1}^{m_n} |H_m| = o(\eta_n^{-2} L(\eta_n))\ n = o(1).$$

This concludes the proof of the lemma. □

Lemma 3.5 *Under the conditions of Theorem 1.1,* $\sum_{i=1}^{m_n} E|u_i|^{2+\delta} = o(\eta_n^{2+\delta})$ *for any* $\delta > 0$ *and hence*

$$P_4(n) := P\left(\max_{m \le m_n} |\sum_{i=1}^{m} Y_i - \sum_{i=1}^{m} Y_i^*| > \frac{\varepsilon B_n}{4} \right) \longrightarrow 0 \quad as \quad n \to \infty.$$

Proof: Using Lemma 2.3, [18] we obtain that

$$\sum_{i=1}^{m_n} E|u_i|^{2+\delta} \le C \sum_{i=1}^{m_n} |H_i|^{1+\delta/2} \max_{j \in H_i} (E\hat{X}_j^2)^{1+\delta/2} + C \sum_{i=1}^{m_n} |H_i| \max_{j \in H_i} E|\hat{X}_j|^{2+\delta}$$
$$:= T_1(n) + T_2(n).$$

We treat separately the two terms. Note that for every $i \le m_n$, we have $\max_{j \in H_i} (E\hat{X}_j^2)^{1+\delta/2} \le L(\eta_n)^{1+\delta/2}$. Using (11) we get

$$T_1(n) \le CL(\eta_n)^{1+\delta/2} \sum_{i=1}^{m_n} |H_i|^{1+\delta/2} \le CL(\eta_n)^{1+\delta/2} o(n^{1+\delta/2}) = o(\eta_n^{2+\delta}).$$

For the second term, note that by Lemma 1.(d), [8], for every $i \le m_n$

$$\max_{j \in H_i} E|\hat{X}_j|^{2+\delta} \le E|\hat{X}_{N_{m_n}}|^{2+\delta} = o(\eta_{N_{m_n}}^{\delta} L(\eta_{N_{m_n}})) = o(\eta_n^{\delta} L(\eta_n)).$$

We conclude that

$$T_2(n) \le o(\eta_n^{\delta} L(\eta_n)) \sum_{i=1}^{m_n} |H_i| = o(\eta_n^{\delta} L(\eta_n) n) = o(\eta_n^{2+\delta}).$$

The final statement of the lemma follows by (7). □

4 The fourth term

Note that

$$J_4(n) = \left| \frac{s_n}{\beta B_n} - 1 \right| \max_{k \le n} \frac{|W(s_k^2)|}{s_n}.$$

Since $\max_{k \le n} |W(s_k^2)|/s_n = O_P(1)$, the fact that $J_4(n) = o_P(1)$ will follow from the following lemma, which uses condition (C) in an essential way.

Lemma 4.1 *Under the conditions of Theorem 1.2, we have* $s_n^2 \sim \beta^2 B_n^2$.

Proof: Note that $B_n^2 = B_{N_{m_n}}^2 + \sum_{j=N_{m_n}+1}^{n} E(\hat{X}_j - E\hat{X}_j)^2 = B_{N_{m_n}}^2 + o(\eta_n^2)$, since $\sum_{j=N_{m_n}+1}^{n} E(\hat{X}_j - E\hat{X}_j)^2 \le (n - N_{m_n})L(\eta_n) = o(n)L(\eta_n) = o(\eta_n^2)$. By Lemma 2.2, it follows that $B_n^2 \sim B_{N_{m_n}}^2$. Hence, using condition (C), we get

$$A_{N_{m_n}}^2 \sim \beta^2 B_{N_{m_n}}^2 \sim \beta^2 B_n^2 \sim A_n^2. \tag{11}$$

By the Minkowski inequality and (10), we get

$$\left| A_{N_{m_n}} - \| \sum_{i=1}^{m_n} u_i \| \right| = \left| \| \sum_{i=1}^{m_n} (u_i + v_i) \| - \| \sum_{i=1}^{m_n} u_i \| \right| \le \| \sum_{i=1}^{m_n} v_i \| = o(\eta_n) = o(A_{N_{m_n}}).$$

(Note that for the last equality we used (11) and Lemma 2.2.) Therefore

$$A_{N_{m_n}}^2 \sim \| \sum_{i=1}^{m_n} u_i \|^2. \tag{12}$$

We note that a consequence of (12) is the fact that $\| \sum_{i=1}^{m_n} u_i \|^2 \to \infty$ as $n \to \infty$. Using the results of Section 3, we have

$$s_n^2 = \sum_{i=1}^{m_n} E u_i^2 = \sum_{i=1}^{m_n} E Y_i^2 = || \sum_{i=1}^{m_n} Y_i ||^2$$

where $\{Y_i\}_{i \ge 1}$ is a sequence of independent random variables. Finally, by Minkowski's inequality and (6), we get

$$\left| || \sum_{i=1}^{m_n} u_i || - s_n \right| = \left| || \sum_{i=1}^{m_n} u_i || - || \sum_{i=1}^{m_n} Y_i || \right| \le || \sum_{i=1}^{m_n} (u_i - Y_i) || \le C.$$

Therefore, $|| \sum_{i=1}^{m_n} u_i || - s_n = o(a_n)$ for any sequence $\{a_n\}_n$ of positive numbers with $a_n \to \infty$. In particular, this happens for $a_n = s_n$ and hence

$$|| \sum_{i=1}^{m_n} u_i ||^2 \sim s_n^2 \tag{13}$$

The conclusion of the lemma follows from (11), (12) and (13). $\square$

Appendix

Theorem 1 of [6] is proved using a slightly different truncation technique. More precisely, this theorem states that under the conditions of Theorem 1.1,

$$\frac{S_n}{\hat{A}_n} \xrightarrow{d} N(0,1), \quad \text{where} \quad \hat{A}_n = \text{Var}(\sum_{j=1}^{n} X_j I_{\{|X_j| \le \eta_n\}}).$$

The next lemma shows that the two truncations are essentially the same.

Lemma 4.2 *Under the conditions of Theorem 1.1, we have $\hat{A}_n^2 \sim A_n^2$.*

Proof: We denote $S_k(n) = \sum_{j=1}^{k} X_j I_{\{|X_j| \leq \eta_n\}}$. Without loss of generality we will assume that X_1 is symmetric. Hence $\hat{A}_n = \| \hat{S}_n(n) \|$ and $A_n = \| \hat{S}_n \|$. Using the Minkowski's inequality and a well-known property of a mixing sequence (see e.g. Proposition 3.1, [13]), we get

$$|\hat{A}_n - A_n|^2 = \Big| \| \hat{S}_n(n) \| - \| \hat{S}_n \| \Big|^2 \leq \| \hat{S}_n(n) - \hat{S}_n \|^2 = E\left(\sum_{j=1}^{n} X_j I_{\{\eta_j < |X_j| \leq \eta_n\}} \right)^2$$

$$\leq C \sum_{j=1}^{n} E(X_j^2 I_{\{\eta_j < |X_j| \leq \eta_n\}}) = C \sum_{j=1}^{n} (L(\eta_n) - L(\eta_j)) = o(\eta_n^2)$$

where the last equality follows by (20) of [8]. Hence $\hat{A}_n - A_n = o(\eta_n)$.

Finally, by relation (3.15) of [19], there exists some constants $C_0, D_0 > 0$ such that $C_0 \eta_n^2 \leq \hat{A}_n^2 \leq D_0 \eta_n^2$. This concludes the proof of the lemma. □

References

[1] R. M. Balan, I. M. Zamfirescu. Strong Approximation for Mixing Sequences with Infinite Variance, *Electr. Comm. Probab.* **11** (2006), 11-23.

[2] R. M. Balan and R. Kulik. Self-normalized weak invariance principle for mixing sequences. *Lab. Reas. Probab. Stat., Univ. Ottawa-Carleton Univ., Tech. Rep. Ser.* **417** (2005).

[3] V. Bentkus, F. Götze, A. Tikhomirov. Berry-Esseen bounds for statistics of weakly dependent samples, *Bernoulli* **3** (1997), 329–349.

[4] I. Berkes, W. Philipp. Approximation theorems for independent and weakly dependent random vectors, *Ann. Probab.* **7** (1979), 29-54.

[5] R. C. Bradley. A Sufficient Condition for Linear Growth of Variances in a Stationary Random Sequence, *Proc. Amer. Math. Soc.* **83** (1981), 586–589.

[6] R. C. Bradley. A central limit theorem for stationary ρ-mixing sequences with infinite variance, *Ann. Probab.* **16** (1988), 313–332.

[7] R. C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions, *Probab. Surveys* **2** (2005), 107–144.

[8] M. Csörgő, B. Szyszkowicz, Q. Wang. Donsker's theorem for self-normalized partial sums processes, *Ann. Probab.* **31** (2003), 1228–1240.

[9] E. Giné, F. Götze, D. M. Mason. When is the Student t-statistic asymptotically standard normal?, *Ann. Probab.* **25** (1997), 1514–1531.

[10] B. F. Logan, C. L. Mallows, S. O. Rice, L. A. Shepp. Limit distributions of self-normalized sums, *Ann. Probab.* **1** (1973), 788–809.

[11] M. Peligrad. Invariance principles for mixing sequences of random variables, *Ann. Probab.* **10** (1982), 968–981.

[12] M. Peligrad. On Ibragimov-Iosifescu conjecture for ϕ mixing sequences, *Stoch. Proc. Appl.* **35** (1990), 293-308.

[13] M. Peligrad. Maximum of partial sums and an invariance principle for a class of weak dependent random variables, *Proc. Amer. Math. Soc.* **126** (1998), 1181–1189.

[14] M. Peligrad, Q.-M. Shao. Self-normalized central limit theorem for sums of weakly dependent random variables, *J. Theor. Probab.* **7** (1994), 309–338.

[15] M. Peligrad, Q.-M. Shao. Estimation of the variance of partial sums for ρ-mixing random variables, *J. Multiv. Anal.* **52** (1995), 140–157.

[16] M. Peligrad, Q.-M. Shao. A note on estimation of variance for ρ-mixing sequences, *Stat. Probab. Letters* **26** (1996), 141–145.

[17] Sakhanenko, A. I. (1984). On estimates of the rate of convergence in the invariance principle. *Advances in Probability Theory: Limit Theorems and Related Problems.* Ed. A.A. Borovkov, 124–135.

[18] Q.-M. Shao. Almost sure invariance principles for mixing sequences of random variables, *Stoch. Proc. Appl.* **48** (1993), 319-334.

[19] Q.-M. Shao. An invariance principle for stationary ρ-mixing sequences with infinite variance, *Chinese Ann. Math. Ser. B* **14** (1993), 27–42.

[20] Q.-M. Shao. Strong approximation theorems for independent random variables and their applications, *J. Multiv. Anal.* **52** (1995), 107–130.