

Self-Normalized Weak Invariance Principle for Mixing Sequences

R.M. Balan^{*†}

University of Ottawa

R. Kulik[‡]

Wroclaw University and University of Ottawa

March 30, 2005

Abstract

In this article we give a necessary and sufficient condition for a self-normalized weak invariance principle, in the case of a strictly stationary ϕ -mixing sequence $\{X_j\}_{j \geq 1}$. This is obtained under the assumptions that the function $L(x) = EX_1^2 1_{\{|X_1| \leq x\}}$ is slowly varying at ∞ and the mixing coefficients satisfy $\sum_{n \geq 1} \phi^{1/2}(n) < \infty$.

Keywords: self-normalized; weak invariance principle; mixing sequences.

1 Introduction

Let $\{X_j\}_{j \geq 1}$ be a sequence of independent identically distributed random variables. The fact that the average \bar{X}_n of the first n observations (or their sum $S_n = \sum_{j=1}^n X_j$) has an asymptotic normal distribution lies at the foundation of the statistical inferences on the population mean $\mu = EX_1$. If $\sigma^2 = \text{Var}(X_1)$ is unknown (or is infinite), then a statistician will normally replace it by the sample variance $U_n^2 = (n-1)^{-1} \sum_{j=1}^n (X_j - \bar{X}_n)^2$ and use the t_{n-1} -distribution instead of $N(0, 1)$, provided that the distribution of X_1 is “approximately” normal. This is justified by the following fact: if the distribution of X_1 is normal,

^{*}*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:* Wroclaw University, Mathematical Institute, Pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland and Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Avenue, Ottawa, ON, K1N 6N5, Canada. *E-mail address:* rkuli438@science.uottawa.ca

then the t -statistic

$$T_n = \frac{\bar{X}_n - \mu}{U_n/\sqrt{n}}$$

has a t_{n-1} -distribution.

Suppose for simplicity that $\mu = 0$ and let $V_n^2 = \sum_{j=1}^n X_j^2$. By observing that T_n has a simple expression in terms of the self-normalized observations $\{X_j/V_n\}_{j=1,\dots,n}$, Giné, Götze and Mason showed in 1997 that $T_n \xrightarrow{d} N(0, 1)$, or equivalently $S_n/V_n \xrightarrow{d} N(0, 1)$, if and only if the following condition holds:

$$(C) \quad L(x) = EX_1^2 1_{\{|X| \leq x\}} \text{ is slowly varying at } \infty.$$

(see [7]). This proved a long-standing conjecture of [9].

Other fluctuation results for the sequence of self-normalized observations have been proved by various authors: the law of iterated logarithm was obtained in [8], the Berry-Esseen theorem can be found in [1], the large deviation principle was proved in [17] and the functional central limit theorem is due to [12]. 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 [4] showed that under (C), the behavior *in probability* of the self-normalized process $\{S_{[nt]}/V_n\}_{t \in [0,1]}$ coincides with the behavior of a standard Brownian motion $W = \{W(t)\}_{t \in [0,1]}$. More precisely, on an appropriate probability space we have

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

To prove the analogue of (1) in the sense of *almost sure* convergence is an open problem. In fact, the following weaker version of it remains a conjecture, posed by Shao in 1998 (see [18]):

$$\sup_{t \in [0,1]} \left| \frac{S_{[nt]}}{V_n} - \frac{W(nt)}{\sqrt{n}} \right| = o((\log \log n)^{1/2}) \quad \text{a.s.}$$

In the present paper we will prove that an analog to the weak invariance principle (1) holds for a strictly stationary ϕ -mixing sequence of random variables. This will be done under the assumptions that (C) holds and the mixing coefficients satisfy the condition $\sum_{n \geq 1} \phi^{1/2}(n) < \infty$.

As far as we know, this is one of the first results in the area of self-normalized limit theorems for mixing sequences without any moment assumptions. We note in passing that self-normalized limit theorems for martingale difference sequences with *finite* variance are well-known (see e.g. [5]).

We should point out that in the case of a ϕ -mixing sequence, one cannot obtain exactly (1), 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 $\sum_{j \geq 2} |EX_1 X_j| < \infty$. Then it is known that $S_n/(\sigma_0 \sqrt{n}) \xrightarrow{d} N(0, 1)$, where $\sigma_0^2 = \sigma^2 + 2 \sum_{j \geq 2} EX_1 X_j$ (see e.g. Theorem 18.5.2, [6]). On the other hand, it is not difficult to see that $V_n/(\sigma \sqrt{n}) \xrightarrow{P} 1$. Thus, $S_n/(\beta V_n) \xrightarrow{d} N(0, 1)$, where $\beta = \sigma_0/\sigma$. Note that in general $\beta \neq 1$. We conclude from here that the proper “self-normalization” for the sequence $\{X_j\}_{j=1, \dots, n}$ should be βV_n .

Throughout this paper we use the notation $a_n \sim b_n$ if $a_n/b_n \rightarrow 1$. We denote with C a generic constant that may be different in each of its appearances.

2 The main result

We begin to introduce the notation that will be used throughout this paper. A sequence $\{X_j\}_{j \geq 1}$ of random variables is called ϕ -**mixing** if $\phi(n) \rightarrow 0$ and ρ -**mixing** if $\rho(n) \rightarrow 0$, where

$$\phi(n) := \sup_{k \geq 1} \phi(\mathcal{M}_1^k, \mathcal{M}_{k+n}^\infty), \quad \rho(n) := \sup_{k \geq 1} \rho(\mathcal{M}_1^k, \mathcal{M}_{k+n}^\infty)$$

$$\phi(\mathcal{M}_1^k, \mathcal{M}_{k+n}^\infty) := \sup\{|P(B|A) - P(B)|; A \in \mathcal{M}_1^k, B \in \mathcal{M}_{k+n}^\infty\}$$

$$\rho(\mathcal{M}_1^k, \mathcal{M}_{k+n}^\infty) := \sup\{|\text{corr}(U, V)|; U \in L^2(\mathcal{M}_1^k), V \in L^2(\mathcal{M}_{k+n}^\infty)\}$$

and \mathcal{M}_a^b denotes the σ -field generated by X_a, X_{a+1}, \dots, X_b . It is known that $\rho(n) \leq 2\phi^{1/2}(n)$; hence a ϕ -mixing sequence is ρ -mixing.

There is an immense amount of literature dedicated to limit theorems for mixing sequences with finite variance; see [10] for an excellent review on this subject. However, very few results are available in the case of infinite variance. Among these, the following plays the role of the “central limit theorem”:

Theorem A. (Theorem 1, [3]) *Let $\{X_j\}_{j \geq 1}$ be a strictly stationary sequence of nondegenerate random variables such that $EX_1 = 0$ and (C) holds. Let $S_n = \sum_{j=1}^n X_j$. If $\rho(1) < 1$ and*

$$\sum_{n \geq 1} \rho(2^n) < \infty \tag{2}$$

then $S_n/\hat{A}_n \xrightarrow{d} N(0, 1)$, where $\hat{A}_n^2 = \text{Var}(\sum_{j=1}^n X_j 1_{\{|X_j| \leq \eta_n\}})$ and $\{\eta_n\}_{n \geq 1}$ is a nondecreasing sequence of positive numbers satisfying $\eta_n^2 \sim nL(\eta_n)$.

Note that the “functional” version of the previous theorem has been obtained in [15]. It is not difficult to see that under the hypothesis of Theorem A, we have $\hat{B}_n^2 := \sum_{j=1}^n \text{Var}(X_j 1_{\{|X_j| \leq \eta_n\}}) \sim \eta_n^2$. On the other hand,

$$C\eta_n^2 \leq \hat{A}_n^2 \leq C\eta_n^2 \tag{3}$$

(see (3.13) of [3]). However, it is not clear whether $\lim_{n \rightarrow \infty} \hat{A}_n^2 / \hat{B}_n^2 := \beta^2$ exists and is finite. Since one can easily prove that $V_n^2 / \hat{B}_n^2 \xrightarrow{P} 1$ (our Proposition 3.4), from Theorem A one immediately obtains the following result:

Theorem B. *Let $\{X_j\}_{j \geq 1}$ be a strictly stationary sequence of nondegenerate random variables such that $EX_1 = 0$ and (C) holds. Suppose that $\rho(1) < 1$ and (2) is satisfied. Let $S_n = \sum_{j=1}^n X_j$ and $V_n^2 = \sum_{j=1}^n X_j^2$. Then*

$$\frac{S_n}{\beta V_n} \xrightarrow{d} N(0, 1) \quad \text{if and only if} \quad \hat{A}_n^2 \sim \beta^2 \hat{B}_n^2.$$

Remark: It is possible to prove that if

$$\lim_{n \rightarrow \infty} \frac{1}{L(\eta_n)} \sum_{j=2}^n \text{Cov}(X_1 1_{\{|X_1| \leq \eta_n\}}, X_j 1_{\{|X_j| \leq \eta_n\}}) = \alpha > -1/2$$

then $\hat{A}_n^2 \sim \beta^2 \hat{B}_n^2$ with $\beta^2 = 1 + 2\alpha$. Note that a similar condition was used in [13] to obtain the central limit theorem for ϕ -mixing sequences without any moment assumptions.

In the present paper, we consider only ϕ -mixing sequences which satisfy

$$\sum_{n \geq 1} \phi^{1/2}(n) < \infty. \quad (4)$$

Moreover, it turns out that for our purposes it is better to use the truncation technique of [4], which differs slightly from that of [3], [15]. More precisely, we let

$$A_n^2 = \text{Var}\left(\sum_{j=1}^n X_j 1_{\{|X_j| \leq \eta_j\}}\right), \quad B_n^2 = \sum_{j=1}^n \text{Var}(X_j 1_{\{|X_j| \leq \eta_j\}})$$

where $\eta_n^2 \sim nL(\eta_n)$ and $\eta_n \leq \eta_{n+1}$. We are now ready to state our main result.

Theorem 2.1 *Let $\{X_j\}_{j \geq 1}$ be a strictly stationary sequence of nondegenerate random variables such that $EX_1 = 0$ and (C) holds. Suppose that $\phi(1) < 1$ and (4) is satisfied. Let $S_n = \sum_{j=1}^n X_j$ and $V_n^2 = \sum_{i=1}^n X_j^2$.*

The following statements are equivalent:

- (a) $A_n^2 \sim \beta B_n^2$.
- (b) $S_n / \beta V_n \xrightarrow{d} N(0, 1)$.
- (c) $S_{[nt]} / \beta V_n \xrightarrow{d} W(t)$ in $D[0, 1]$ endowed with the sup-norm metric.
- (d) 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 statement (d) \Rightarrow (c) is straightforward; for the sake of completeness we include its proof in Appendix A. Also, it is clear that (c) \Rightarrow (b).

To see that (b) \Rightarrow (a), we first prove that $A_n^2 \sim \hat{A}_n^2$ (see Appendix B). Using Theorem A, it follows that $S_n/A_n \xrightarrow{d} N(0, 1)$; this combined with the fact that $V_n/B_n \xrightarrow{P} 1$ (our Proposition 3.4) imply that

$$\frac{S_n}{(A_n/B_n)V_n} \xrightarrow{d} N(0, 1). \quad (5)$$

From (5) and (b) we conclude that (a) holds.

The remaining part of this article is devoted to the proof of (a) \Rightarrow (d). For this we let $\hat{X}_j = X_j 1_{\{|X_j| \leq \eta_j\}}$ and $\hat{S}_n = \sum_{j=1}^n \hat{X}_j$.

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 $\{\hat{X}_j/B_n\}_{j=1, \dots, n}$ and then proving that $\beta B_n \sim s_n$. Indeed, we have the following decomposition:

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

In the next sections we treat separately each of the four terms.

3 The first two terms

Let $\bar{X}_j = X_j 1_{\{|X_j| > \eta_j\}}$ and note that $E\bar{X}_j = -E\hat{X}_j$. We begin with some auxiliary lemmas.

Lemma 3.1 (Lemma 1, [4]) *Let X be a random variable with $EX = 0$ and $L(x) = EX^2 1_{\{|X| \leq x\}}$. The following statements are equivalent:*

- (a) $L(x)$ is a slowly varying function at ∞ ;
- (b) $x^2 P(|X| > x) = o(L(x))$;
- (c) $x E|X| 1_{\{|X| > x\}} = o(L(x))$;
- (d) $E|X|^{\alpha} 1_{\{|X| \leq x\}} = o(x^{\alpha-2} L(x))$ for $\alpha > 2$.

Lemma 3.2 *If (C) holds, then $\sum_{j=1}^n E|\bar{X}_j| = o(\eta_n)$.*

Proof: We write

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

and we note that $nE|X_1|1_{\{|X_1|>\eta_n\}} \leq C\eta_n^2 L^{-1}(\eta_n)E|X_1|1_{\{|X_1|>\eta_n\}} = o(\eta_n)$, using Lemma 3.1.(c) and $\sum_{j=1}^n E|X_1|1_{\{\eta_j < |X_1| \leq \eta_n\}} = o(\eta_n)$, using (20) of [4].
 \square

Lemma 3.3 *If (C) holds, then $B_n^2 \sim \eta_n^2$.*

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 $\sum_{j=1}^n L(\eta_j) \sim nL(\eta_n)$ using (9) of [4], and $\sum_{j=1}^n (E\bar{X}_j)^2 \leq (\sum_{j=1}^n E\bar{X}_j)^2 = o(\eta_n^2)$ by Lemma 3.2. \square

Proposition 3.4 *If (C) and (2) hold, then*

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

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 Lemmas 3.2 and 3.3 we have $B_n^{-2} \sum_{j=1}^n \bar{X}_j^2 \leq (B_n^{-1} \sum_{i=1}^n |\bar{X}_j|)^2 = o_P(1)$. It suffices to prove that

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

For this, we note that $\{\hat{X}_j^2 - E\hat{X}_j^2\}_{j \geq 1}$ is a ϕ' -mixing sequence with coefficient $\phi'(n) \leq \phi(n)$. Using Lemma 2.3, [14], Lemma 3.3 and Lemma 3.1.(d), we get

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

\square

By Lemma 3.2, Lemma 3.3 and Proposition 3.4 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 4, respectively Section 5.

4 The third term

For the third term $J_3(n)$, let H_i, I_i be the long, respectively short blocks with

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

for some $a \in (0, 1)$.

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

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

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

Proposition 4.1 *If*

$$\sum_{k \geq 1} \phi^{1/2}(e^{k^a/2}) < \infty$$

then without changing its distribution, we can redefine the sequence $\{u_i\}_{i \geq 1}$ on a larger probability space together with a sequence $\{Y_i\}_{i \geq 1}$ of independent random variables such that Y_i has the same distribution as u_i and for all $m \geq 1$

$$\left| \sum_{i=1}^m u_i - \sum_{i=1}^m Y_i \right| \leq C \quad \text{a.s.} \quad (6)$$

Proof: We apply Theorem 2, [2] with $X_k = u_k$, $\mathcal{L}_k = \sigma(u_k)$ and

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

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

$$P(|u_k - Y_k| \geq 6\phi_k) \leq 6\phi_k.$$

Since $\sum_{k \geq 1} \phi_k \leq \sum_{k \geq 1} \phi(e^{k^a/2}) < \infty$, by the Borel-Cantelli lemma we have

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

Hence $|\sum_{i=1}^m u_i - \sum_{i=1}^m Y_i| \leq \sum_{i=1}^m |u_i - Y_i| \leq C \sum_{i=1}^m \phi_i \leq C$ for all $m \geq 1$. \square

Using Sakhanenko's theorem (see Theorem B, [16]), without changing its distribution we can redefine the sequence $\{Y_i\}_{i \geq 1}$ 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 $x > 0, \delta > 0$

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

Clearly, without changing its distribution we can redefine the sequence $\{Y_i^*\}_{i \geq 1}$ 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 every m . In particular $W(s_k^2) = W(s_{m_k}^{*2}) = \sum_{i=1}^{m_k} Y_i^*$. Note that

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

We are now ready to treat the third term. For any $\varepsilon > 0$ we have

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

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

Lemma 4.2 *Suppose that (C) holds. If $\phi(1) < 1$ and (2) is satisfied, then*

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

Proof: Note that $\sum_{i=1}^{m_n} |I_i| \leq Cn^{1/2}$. Using Lemma 2.3, [14] for every $m \leq m_n$

$$\sum_{i=1}^m E v_i^2 \leq C \sum_{i=1}^m |I_i| \max_{j \in I_i} E \hat{X}_j^2 \leq CL(\eta_n) \sum_{i=1}^m |I_i| \leq CL(\eta_n) n^{1/2} \quad (8)$$

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

Note that the sequence $\{v_i\}_{i \geq 1}$ is ϕ -mixing with coefficient $\phi^{(v)}(n) \leq \phi(e^{n^a})$. Using Proposition 3.2, [11], we get

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

The result follows by Chebyshev's inequality. \square

Lemma 4.3 *Suppose that (C) holds. If $\phi(1) < 1$ and (2) is satisfied, then*

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

Proof: By Markov's inequality we get

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

Using Proposition 3.2, [11], Lemma 2.3, [14] and Lemma 3.1.(d) we have: for every $m \leq m_n$

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

Hence

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

Note that

$$\begin{aligned} \sum_{m=1}^{m_n} |H_m|^{1+\delta/2} &\leq C \sum_{m=1}^{m_n} m^{(a-1)(2+\delta)/2} e^{(2+\delta)m^a/2} = \sum_{m=1}^{m_n} o(m^{a-1} e^{(2+\delta)m^a/2}) \\ &= o(e^{(2+\delta)m_n^a/2}) = o(n^{1+\delta/2}) \end{aligned} \tag{9}$$

and hence $P'_3(n) = o(1)$. For the second part we use $\sum_{i=1}^{m_n} |H_i| \leq Cn$. Hence

$$P''_3(n) \leq \frac{C}{\eta_n^{2+\delta}} o(\eta_{N_{m_n}}^\delta L(\eta_{N_{m_n}})) \sum_{m=1}^{m_n} |H_m| = o(\eta_n^{-2} L(\eta_n)) \quad n = o(1).$$

This concludes the proof of the lemma. \square

Lemma 4.4 *Suppose that (C) holds. If (2) is satisfied, then $\sum_{i=1}^{m_n} E|u_i|^{2+\delta} = o(\eta_n^{2+\delta})$ and hence*

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

Proof: Using Lemma 2.3, [14] we obtain that

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

We treat separately the two terms. Note that for every $i \leq m_n$, we have $\max_{j \in H_i} (E\hat{X}_j^2)^{1+\delta/2} \leq L(\eta_n)^{1+\delta/2}$. Using (9) we get

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

For the second term, note that by Lemma 3.1.(d) we have for every $i \leq m_n$

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

We conclude that

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

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

5 The fourth term

We have

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

Since $\max_{k \leq n} |W(s_k^2)|/s_n = O_P(1)$, the fact that $J_4(n) = o_P(1)$ will follow from (a) and the following lemma.

Lemma 5.1 *If (4) holds, then $s_n^2 \sim A_n^2$.*

Proof: By Lemma B.1 (Appendix B) and (3), it suffices to prove that $A_n^2 - s_n^2 = o(\eta_n^2)$. This will follow from

$$A_{N_{m_n}}^2 - s_n^2 = o(\eta_n^2) \quad (10)$$

$$A_n^2 - A_{N_{m_n}}^2 = o(\eta_n^2) \quad (11)$$

To prove (10), we note that $A_{N_{m_n}}^2 = E(\hat{S}_{N_{m_n}} - E\hat{S}_{N_{m_n}})^2 = E(\sum_{i=1}^{m_n} (u_i + v_i))^2$ and hence

$$A_{N_{m_n}}^2 - s_n^2 = \sum_{i=1}^{m_n} E v_i^2 + 2 \sum_{i=1}^{m_n} E u_i v_i + 2 \sum_{i=1}^{m_n} \sum_{j=i+1}^{m_n} E(u_i + v_i)(u_j + v_j)$$

By (8), we have $\sum_{i=1}^{m_n} E v_i^2 = o(\eta_n^2)$. Using Theorem 17.2.3, [6] and (4)

$$|E u_i v_i| \leq \sum_{k \in H_i} \sum_{l \in I_i} |E \hat{X}_k \hat{X}_l - (E \hat{X}_k)(E \hat{X}_l)| \leq 2 \sum_{k \in H_i} \sum_{l \in I_i} \phi^{1/2}(l-k) (E \hat{X}_k^2)^{1/2} (E \hat{X}_l^2)^{1/2}$$

$$\leq 2L(\eta_n) \sum_{k \in H_i} |I_i| \phi(N_i + |H_i| - k) \leq CL(\eta_n) |I_i|$$

and $|\sum_{i=1}^{m_n} E u_i v_i| \leq CL(\eta_n) \sum_{i=1}^{m_n} |I_i| \leq CL(\eta_n) n^{1/2} = o(\eta_n^2)$. It remains to prove that

$$\sum_{i=1}^{m_n} \sum_{j=i+1}^{m_n} E(u_i + v_i)(u_j + v_j) = o(\eta_n^2). \quad (12)$$

For this we treat separately the four terms. Note that $E u_i^2 \leq C|H_i| \max_{j \in H_i} E \hat{X}_j^2 \leq CL(\eta_n) |H_i|$ for every $i \leq m_n$. Using Theorem 17.2.3, [6]

$$|E u_i u_j| \leq 2\phi^{1/2}(|I_i|) (E u_i^2)^{1/2} (E u_j^2)^{1/2} \leq CL(\eta_n) \phi^{1/2}(|I_i|) |H_i|^{1/2} |H_j|^{1/2}.$$

By Kronecker's lemma, under (4) we have $\sum_{i=1}^{j-1} \phi^{1/2}(|I_i|) |H_i|^{1/2} = o(|H_j|^{1/2})$. Hence

$$\begin{aligned} \left| \sum_{i=1}^{m_n} \sum_{j=i+1}^{m_n} E u_i u_j \right| &\leq CL(\eta_n) \sum_{j=1}^{m_n} |H_j|^{1/2} \sum_{i=1}^{j-1} \phi^{1/2}(|I_i|) |H_i|^{1/2} = L(\eta_n) o\left(\sum_{j=1}^{m_n} |H_j|\right) \\ &= L(\eta_n) o(n) = o(\eta_n^2). \end{aligned}$$

For any $i, j = 1, \dots, m_n$ with $i < j$ we have

$$|E v_i v_j| \leq \sum_{k \in H_i} \sum_{l \in H_j} |E \hat{X}_k \hat{X}_l - (E \hat{X}_k)(E \hat{X}_l)| \leq 2 \sum_{l \in H_j} \sum_{k \in I_i} \phi^{1/2}(l-k) (E \hat{X}_k^2)^{1/2} (E \hat{X}_l^2)^{1/2}$$

$$\leq CL(\eta_n) \sum_{l \in H_j} |I_l| \phi^{1/2}(l - N_i) \leq CL(\eta_n) |I_i|.$$

From here we get immediately $|\sum_{i=1}^{m_n} \sum_{j=i+1}^{m_n} E v_i u_j| = o(\eta_n^2)$. The other two terms in (12) are similar.

To prove (11), note that

$$A_n^2 = A_{N_{m_n}}^2 + E \left(\sum_{j=N_{m_n}+1}^n (\hat{X}_j - E\hat{X}_j) \right)^2 + 2 \sum_{j=N_{m_n}+1}^n E(\hat{X}_j - E\hat{X}_j)(\hat{S}_{N_{m_n}} - E\hat{S}_{N_{m_n}}).$$

By Lemma 2.3, [14], $E \left(\sum_{j=N_{m_n}+1}^n (\hat{X}_j - E\hat{X}_j) \right)^2 \leq C(n - N_{m_n})L(\eta_n) = o(n)L(\eta_n) = o(\eta_n^2)$. Using Theorem 17.2.3, [6] and (4)

$$\begin{aligned} \left| \sum_{j=N_{m_n}+1}^n E(\hat{X}_j - E\hat{X}_j)(\hat{S}_{N_{m_n}} - E\hat{S}_{N_{m_n}}) \right| &\leq \sum_{i=1}^{N_{m_n}} \sum_{j=N_{m_n}+1}^n |E\hat{X}_i \hat{X}_j - (E\hat{X}_i)(E\hat{X}_j)| \leq \\ &2 \sum_{i=1}^{N_{m_n}} \sum_{j=N_{m_n}+1}^n \phi^{1/2}(j-i)(E\hat{X}_i^2)^{1/2}(E\hat{X}_j^2)^{1/2} \leq CL(\eta_n)(n - N_{m_n}) = o(\eta_n^2). \end{aligned}$$

The proof of the lemma is complete. \square

A The proof of (d) \Rightarrow (c)

Since $\{W(s_{[nt]}^2)/s_n; t \in [0, 1]\} \stackrel{d}{=} \{W(s_{[nt]}^2/s_n^2); t \in [0, 1]\}$, it suffices to show that

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

which in turn will follow from

$$\max_{k \leq n} \left| \frac{s_k^2}{s_n^2} - \frac{k}{n} \right| = o(1).$$

Clearly (d) \Rightarrow (b) \Rightarrow (a). Since $s_n^2 \sim A_n^2 \sim \beta^2 B_n^2$, it is enough to show that

$$\max_{k \leq n} \left| \frac{B_k^2}{B_n^2} - \frac{k}{n} \right| = o(1)$$

To prove this, note that

$$\frac{B_k^2}{nL(\eta_n)} = \frac{k}{n} - \frac{1}{nL(\eta_n)} \sum_{j=1}^k (L(\eta_n) - L(\eta_j)) - \frac{1}{nL(\eta_n)} \sum_{j=1}^k (E\bar{X}_j)^2.$$

Hence, using (19) of [4], Lemma 3.2 and Lemma 3.3 we get

$$\max_{k \leq n} \left| \frac{B_k^2}{B_n^2} - \frac{k}{n} \right| \leq \left| \frac{B_n^2}{nL(\eta_n)} - 1 \right| + \max_{k \leq n} \left| \frac{B_k^2}{nL(\eta_n)} - \frac{k}{n} \right| = o(1).$$

This concludes the proof of (d) \Rightarrow (c).

B Equivalence of the two variances

Lemma B.1 *If (4) holds, then $\hat{A}_n^2 \sim A_n^2$.*

Proof: By (3), it is enough to show that $\hat{A}_n^2 - A_n^2 = o(\eta_n^2)$. For each $j = 1, \dots, n$ let $\hat{X}_{j,n} = X_j 1_{\{|X_j| \leq \eta_n\}}$ and $\bar{X}_{j,n} = X_j 1_{\{|X_j| > \eta_n\}}$. Similarly to Lemma 3.2, one can show that $\sum_{j=1}^n E|\bar{X}_{j,n}| = o(\eta_n)$. Note that

$$\begin{aligned} \hat{A}_n^2 - A_n^2 &= \sum_{j=1}^n (L(\eta_n) - L(\eta_j)) - \sum_{j=1}^n (E\bar{X}_{j,n})^2 + \sum_{j=1}^n (E\bar{X}_j)^2 + \\ &2 \sum_{i=1}^n \sum_{j=i+1}^n (\text{Cov}(\hat{X}_{i,n}, \hat{X}_{j,n}) - \text{Cov}(\hat{X}_i, \hat{X}_j)). \end{aligned}$$

Using (19) of [4] and Lemma 3.2, it is enough to prove that

$$\sum_{i=1}^n \sum_{j=i+1}^n (\text{Cov}(\hat{X}_{i,n}, \hat{X}_{j,n}) - \text{Cov}(\hat{X}_i, \hat{X}_j)) = o(\eta_n^2). \quad (13)$$

Note that $\hat{X}_{j,n} = \hat{X}_j + X_{j,n}^*$ where $X_{j,n}^* = X_j 1_{\{\eta_j < |X_j| \leq \eta_n\}}$. We have

$$\text{Cov}(\hat{X}_{i,n}, \hat{X}_{j,n}) - \text{Cov}(\hat{X}_i, \hat{X}_j) = \text{Cov}(X_{i,n}^*, X_{j,n}^*) + \text{Cov}(X_{i,n}^*, \hat{X}_j) + \text{Cov}(\hat{X}_i, X_{j,n}^*).$$

Using Theorem 17.2.3, [6] and (4) we get

$$\begin{aligned} \left| \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_{i,n}^*, X_{j,n}^*) \right| &\leq 2 \sum_{i=1}^n \sum_{j=i+1}^n \phi^{1/2}(j-i) (EX_{i,n}^{*2})^{1/2} (EX_{j,n}^{*2})^{1/2} \leq \\ 2L^{1/2}(\eta_n) \sum_{i=1}^n (L(\eta_n) - L(\eta_i))^{1/2} \sum_{j=i+1}^n \phi^{1/2}(j-i) &\leq CL^{1/2}(\eta_n) \sum_{i=1}^n (L(\eta_n) - L(\eta_i))^{1/2}. \end{aligned}$$

From (19) of [4] it follows that $\sum_{i=1}^n (L(\eta_n) - L(\eta_i))^{1/2} = o(nL^{1/2}(\eta_n))$. Hence $\sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_{i,n}^*, X_{j,n}^*) = o(\eta_n^2)$.

Using a similar argument one can show that $\sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_{i,n}^*, \hat{X}_j) = o(\eta_n^2)$ and $\sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(\hat{X}_i, X_{j,n}^*) = o(\eta_n^2)$. The proof of (13) is complete. \square

References

- [1] V. Bentkus and F. Götze, The Berry-Esseen bound for Student's statistic, *Ann. Probab.* 24 (1996) 491–503.
- [2] I. Berkes and W. Philipp, Approximation theorems for independent and weakly dependent random vectors, *Ann. Probab.* 7 (1979) 29–54.
- [3] R. C. Bradley, A central limit theorem for stationary ρ -mixing sequences with infinite variance, *Ann. Probab.* 16 (1988) 313–332.
- [4] M. Csörgő, B. Szyszkowicz and Q. Wang, Donsker's theorem for self-normalized partial sums processes, *Ann. Probab.* 31 (2003) 1228–1240.
- [5] P. Hall and C.C. Heyde, *Martingale Limit Theory and Its Applications*, Academic Press, New York, 1980.
- [6] I. A. Ibragimov and Y. V. Linnik, *Independent and Stationary Sequences of Random Variables*, Woltersnoordhoff, Groningen, The Netherlands, 1971.
- [7] E. Giné, F. Götze and D.M. Mason, When is the Student t -statistic asymptotically standard normal?, *Ann. Probab.* 25 (1997) 1514–1531.
- [8] P. S. Griffin and J. D. Kuelbs, Self-normalized laws of the iterated logarithm, *Ann. Probab.* 17 (1989) 1571–1601.
- [9] B. F. Logan, C. L. Mallows, S. O. Rice and L. A. Shepp, Limit distributions of self-normalized sums, *Ann. Probab.* 1 (1973) 788–809.
- [10] M. Peligrad, Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (a survey). *Dependence in probability and statistics (Oberwolfach, 1985)*, 193–223, *Progr. Probab. Statist.*, 11, Birkhäuser Boston, Boston, MA, 1986.
- [11] M. Peligrad, On Ibragimov-Iosifescu conjecture for ϕ mixing sequences, *Stochastic Process. Appl.* 35 (1990), 293–308.
- [12] A. Račkauskas and C. Suquet, Invariance principles for adaptive self-normalized partial sums processes, *Stochastic Process. Appl.* 95 (2001) 63–81.
- [13] J. D. Samur, A note on the convergence to Gaussian laws of sums of stationary ϕ -mixing triangular arrays. *Probability in Banach spaces, V (Medford, Mass., 1984)*, 387–399, *Lecture Notes in Math.*, 1153, Springer, Berlin, 1985.
- [14] Q.-M. Shao, Almost sure invariance principles for mixing sequences of random variables, *Stoch. Proc. Appl.* 48 (1993) 319–334.

- [15] Q.-M. Shao, An invariance principle for stationary ρ -mixing sequences with infinite variance, *Chinese Ann. Math. Ser. B* 14 (1993) 27–42.
- [16] Q.-M. Shao, Strong approximation theorems for independent random variables and their applications, *J. Multiv. Anal.* 52 (1995) 107–130.
- [17] Q.-M. Shao, Self-normalized large deviations, *Ann. Probab.* 25 (1997) 285–328.
- [18] Q.-M. Shao, Recent developments on self-normalized limit theorems. *Asymptotic methods in probability and statistics (Ottawa, ON, 1997)*, 467–480, North-Holland, Amsterdam, 1998.