Strong Approximation for Mixing Sequences with Infinite Variance

R.M. Balan*† University of Ottawa

I.M. Zamfirescu‡ Baruch College, CUNY

November 15, 2004

Abstract

In this paper we generalize Shao's (1993a) strong invariance principle for ρ-mixing sequences to the case of infinite variance, under the same assumption: $\rho(n) \leq C(\log n)^{-r}, r > 1$. The rate in this strong approximation result is $o(a_n)$, where $\{a_n\}_n$ is a nondecreasing sequence satisfying $a_n^2 \sim nL(a_n)v(a_n)$. (Here $L(x) = EX^2 1_{\{|X| \leq x\}}$ and $v(x)$ is a nondecreasing slowly varying function with $v(x) \geq C \log\log x$.) The result is proved under the assumption that the common distribution of the random variables in the sequence is *symmetric* and lies in the domain of attraction of the normal law.

Keywords: strong approximation; mixing sequences of random variables; Skorohod embedding; blocking technique; truncation.

1 Introduction

The concept of mixing is a natural generalization of independence and can be viewed as "asymptotic independence": the dependence between two random variables in a mixing sequence becomes weaker as the distance between their indices becomes larger. There is an immense amount of literature dedicated to limit theorems for mixing sequences, most of it assuming that the moments of second order (or higher) are finite. One of the most important results in this area is Shao's (1993a) strong invariance principle, from each one can easily deduce many other limit theorems.

*Corresponding author. *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.

‡*Postal address*: Department of Mathematics, Baruch College, CUNY, One Bernard Baruch Way, New York, NY 10010, USA.

In this paper we will prove that Shao's result can be generalized to the case of infinite variance, which suggests that the result may still hold for the self-normalized sequence. Self-normalized limit theorems have become increasingly popular in the past few years, but so far only the case of independent random variables was considered. Therefore, our result may contain the seeds of future research in the promising new area of self-normalized limit theorems for dependent sequences.

In order not to obscure the main point and to avoid many technicalities, we will work under the assumption that the common distribution of the random variables in our mixing sequence is symmetric. This assumption was also used by Feller (1968) and Mijnheer (1980), which constitute our main references for the independent case.

Suppose first that $\{X_n\}_{n\geq 1}$ is a sequence of **independent** identically distributed random variables with $EX=0,\ EX^2=\infty$, where X denotes a generic random variable with the same distribution as X_n. Let $S_n=\sum_{i=1}^n X_i$. In the case of infinite variance, a crucial role is played by the truncated second order moments, given by the function $L(x) := EX^2 1_{\{|X|\leq x\}}$.

If the distribution of X is *symmetric* and

$$X \in DAN$$

then it is well-known (Raikov Theorem) that the "central limit theorem" continues to hold in the form $S_n/\eta_n \to_d N(0,1)$, where $\{\eta_n\}_n$ is a nondecreasing sequence of positive numbers satisfying

$$\eta_n^2 \sim nL(\eta_n). \tag{1}$$

(The notation $X \in DAN$ means that X belongs to the domain of attraction of the normal law, which is equivalent to saying that the function L is slowly varying.) Moreover, by Theorem 1 of Feller (1968) we have

$$\limsup_{n\to\infty} \frac{S_n}{(2\eta_n^2 \log\log \eta_n)^{1/2}} = 1 \text{ or } \infty \quad \text{a.s.}$$

depending on whether the integral

$$I_{\log\log} := \int_b^\infty \frac{x^2}{L(x)\log\log x}\, dF(x)$$

converges or diverges (here $b := \inf\{x \geq 1; L(x) > 0\}$). Hence $I_{\log\log} < \infty$ is a minimum requirement for the "law of the iterated logarithm" in the case of independent random variables with infinite variance.

In the 1971 Rietz Lecture, Kesten (1972) has discussed Feller's result and raised the question of its correctness; see Kesten's Remark 9. Fortunately, he settled this problem (Theorem 7), by replacing Feller's normalizing constant

$(\eta_n^2 \log\log \eta_n)^{1/2}$ with a slightly different constant γ_n, which behaves roughly as a root of the equation

$$\gamma_n^2 = CnL(\gamma_n)\log\log\gamma_n.$$

Following these lines, Theorem 2.1 of Mijnheer (1980) proved that it is possible to obtain (on a larger probability space), the strong approximation

$$S_n - T_n = o(a_n) \quad \text{a.s.} \tag{2}$$

where $T_n = \sum_{i=1}^n Y_i$ and $\{Y_n\}_{n\geq 1}$ is a zero-mean Gaussian sequence (with $EY_n^2 = \tau_n$ for suitable constants τ_n). His rate a_n is chosen such that

$$a_n^2 \sim nL(a_n)v(a_n). \tag{3}$$

where v is a nondecreasing slowly varying function with $\lim_{x\to\infty} v(x) = \infty$.

In this paper we will prove that a strong approximation of type (2) continues to hold in the mixing case.

We begin to introduce the notation that will be used throughout this paper. A sequence $\{X_n\}_{n\geq 1}$ of random variables is called **ρ-mixing** if

$$\rho(n) := \sup_{k\geq 1} \rho(\mathcal{M}_1^k, \mathcal{M}_{k+n}^\infty) \to 0 \quad \text{as} \quad n\to\infty$$

where

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

and $\mathcal{M}_a^b$ denotes the σ-field generated by $X_a, X_{a+1}, \ldots, X_b$.

One of the few results that we found in the literature dealing with ρ-mixing sequences of random variables with infinite variance is a "functional central limit theorem" obtained by Shao (1993b) under the condition $\sum_n \rho(2^n) < \infty$. In order to obtain a strong approximation of type (2) we need to strengthen this condition to $\rho(n) \leq C(\log n)^{-r},\ r > 1$, which is the condition for the strong invariance principle in the finite variance case (see Theorem 1.1 of Shao, 1993a).

We denote $\log x := \log\{\max(x,e)\}$. Let v be a nondecreasing slowly varying function such that $v(x) \geq C\log\log x$ for x large and

$$I_{v(\cdot)} = \int_b^\infty \frac{x^2}{L(x)v(x)}\, dF(x).$$

We require that the function L satisfies the following condition: for every $\lambda > 0$ there exists $C = C_\lambda > 0$ and $K = K_\lambda > 0$ such that

$$(C) \qquad 1 - \frac{L(\lambda x)}{L(x)} \leq Cx^{-K} \text{ for } x \text{ large.}$$

We should mention here that condition (C) was used in *only one* place, namely to ensure the convergence of the sum (40) in the proof of Lemma 4.2. Unfortunately, we could not avoid it.

Here is our main result.

Theorem 1.1 *Let $\{X_n\}_{n\geq 1}$ be a ρ-mixing sequence of* symmetric *identically distributed random variables with $EX = 0$, $EX^2 = \infty$ and $X \in DAN$. Suppose that the function L satisfies (C). Let v be a nondecreasing slowly varying function such that $v(x) \geq C \log\log x$ for x large and*

$$I := I_{v(\cdot)} < \infty.$$

Let $\{a_n\}_n$ be a nondecreasing sequence of positive numbers satisfying (3). If

$$\rho(n) \leq C(\log n)^{-r} \quad \textit{for some } r > 1 \tag{4}$$

then without changing its distribution, we can redefine $\{X_n\}_{n\geq 1}$ on a larger probability space together with a standard Brownian motion $W = \{W(t)\}_{t\geq 0}$ such that for some constants s_n^2

$$S_n - W(s_n^2) = o(a_n) \quad a.s.$$

In Section 2 we give the description of the general method. The technical details are discussed in Sections 3 and 4. Appendix A contains the proofs of two results relying on the slowly varying properties of the functions v and L. Appendix B contains the proof of a martingale subsequence criterion.

Throughout this work, C denotes a generic constant that does not depend on n but may be different from place to place. We denote $A(x) = L(x)v(x)$.

2 Sketch of the Proof

Sequences $\{\eta_n\}_n$ and $\{a_n\}_n$ satisfying (1), respectively (3) can be obtained as follows:

$$\eta_n = \inf\{s \geq b+1; \frac{L(s)}{s^2} \leq \frac{1}{n}\}$$
$$a_n = \inf\{s \geq b+1; \frac{A(s)}{s^2} \leq \frac{1}{n}\}$$

(see p.1233 of Csörgő, Szyszkowicz and Wang, 2003). Hence $a_n \geq \eta_n$ and

$$a_n^2 \geq C\eta_n^2 v(\eta_n) \geq C\eta_n^2 \log\log \eta_n \tag{5}$$

Without loss of generality in what follows we will assume that

$$\eta_n^2 = nL(\eta_n) \quad \text{and} \quad a_n^2 = nA(a_n).$$

As in Lemmas 4.1 and 5.1 of Feller (1968), we consider the following truncation:

$$b_n := \frac{a_n}{v^p(a_n)} < a_n, \quad p > 1/2$$

$$\hat{X}_n = X_n I_{\{|X_n|\leq b_n\}} \quad X'_n = X_n I_{\{b_n<|X_n|\leq a_n\}}, \quad \bar{X}_n = X_n I_{\{|X_n|>a_n\}}.$$

By the *symmetry* assumption $E\hat{X}_n = EX'_n = 0$; since $EX_n = 0$, it follows that $E\bar{X}_n = 0$. We have $X_n = \hat{X}_n + X'_n + \bar{X}_n$ and hence

$$S_n = \hat{S}_n + S'_n + \bar{S}_n \tag{6}$$

where $\hat{S}_n, S'_n, \bar{S}_n$ denote the partial sums of $\hat{X}_i, X'_i$, respectively $\bar{X}_i$.

By Lemmas 3.2 and 3.3 of Feller (1968) (under the *symmetry* assumption), the condition $I < \infty$ is equivalent to: $\sum_{n\geq 1} P(|X| > \epsilon a_n) < \infty$ for all $\epsilon > 0$. By the Borel-Cantelli lemma, $\exists N$ such that $\bar{X}_n = 0$ for all $n \geq N$ a.s. Hence

$$\bar{S}_n = o(a_n) \quad \text{a.s.} \tag{7}$$

In Section 3, we show that the central part $\hat{S}_n$ gives us the approximation

$$\hat{S}_n - W(s_n^2) = o((\eta_n^2 \log\log \eta_n)^{1/2}) \quad \text{a.s.} \tag{8}$$

for some constants s_n^2. In Section 4 we show that

$$S'_n = o(a_n) \quad \text{a.s.} \tag{9}$$

Theorem 1.1 follows immediately by (5)-(9).

3 The Central Part

In this section we will show how to obtain the desired approximation (8). Throughout this work we will denote with $I(a, b]$ the measure attributed by the integral I to the interval $(a, b]$.

Let $\tau = \min(3, r + 1)$. The blocks H_i, I_i are defined exactly as in Shao (1993a), i.e.

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

where $a \in (0, 1)$ is chosen such that $(1 - a)\tau > 2$.

Let $N_m = \sum_{i=1}^m \text{card}(H_i \cup I_i) \sim \exp(m^a)$. For each n there exists a unique index m_n such that $N_{m_n} \leq n < N_{m_n+1}$. We have $N_{m_n} \sim n$ and $m_n \sim (\log n)^{1/a}$. We define

$$u_i = \sum_{j\in H_i} \hat{X}_j, \quad v_i = \sum_{j\in I_i} \hat{X}_j, \quad \xi_i = u_i - E(u_i|\mathcal{G}_{i-1})$$

where $\mathcal{G}_m = \sigma(\{u_i; i \leq m\})$. We have

$$\hat{S}_n = \sum_{i=1}^{m_n} \xi_i + \sum_{i=1}^{m_n} E(u_i|\mathcal{G}_{i-1}) + \sum_{i=1}^{m_n} v_i + \sum_{j=N_{m_n}+1}^{n} \hat{X}_j. \tag{10}$$

The first term will give us the desired approximation with rate $o((\eta_n^2 \log\log \eta_n)^{1/2})$. The other three terms will be of order $o(\eta_n)$.

The following two propositions provide us with powerful approximation tools. Their proofs are given in Appendix A.

Proposition 3.1 *Then there exists $C > 0$ such that $b_n \leq C\eta_n$ for n large, and hence*

$$nL(b_n) \leq C\eta_n^2 \quad \textit{for } n \textit{ large.} \tag{11}$$

Proposition 3.2 *For any integer $\lambda > 0$ there exists $C = C_\lambda > 0$ such that $a_{\lambda n} \leq C a_n$ and $b_{\lambda n} \leq C b_n$ for n large, and hence*

$$L(a_{\lambda n}) \leq CL(a_n) \quad \textit{for } n \textit{ large and} \tag{12}$$

$$L(b_{\lambda n}) \leq CL(b_n) \quad \textit{for } n \textit{ large.} \tag{13}$$

We begin now to prove that the last three terms in (10) are of order $o(\eta_n)$.

Lemma 3.3 *We have*

$$\sum_{i=1}^{m} v_i = o(m^2 \cdot \exp(\frac{1}{3}m^a) \cdot L^{1/2}(b_{N_m})) \quad a.s. \tag{14}$$

and hence $\sum_{i=1}^{m_n} v_i = o(\eta_n)$ a.s.

Proof: Note that $E\hat{X}_j^2 = L(b_j)$. By Lemma 2.3 of Shao (1993a), for $i \leq m$

$$Ev_i^2 \leq C\mathrm{card}(I_i) \cdot \max_{j \in I_i} E\hat{X}_j^2 \leq C i^{a-1} \exp(\frac{1}{2}i^a) \cdot L(b_{N_m})$$

and hence

$$E(\sum_{i=1}^{m} v_i)^2 \leq m \sum_{i=1}^{m} Ev_i^2 \leq CmL(b_{N_m}) \sum_{i=1}^{m} i^{a-1} \exp(\frac{1}{2}i^a) \leq CmL(b_{N_m}) \exp(\frac{2}{3}m^a).$$

By the Chebyshev's inequality and the Borel-Cantelli Lemma, we have

$$|\sum_{i=1}^{m} v_i| \leq Cm^{1+\varepsilon} \exp(\frac{1}{3}m^a) L^{1/2}(b_{N_m}) \quad \text{a.s.}$$

for any $\varepsilon > 0$. Relation (14) follows by taking $\varepsilon \in (0, 1)$. The second statement in the lemma follows by taking $m = m_n$ in (14) and using (11). $\square$

Lemma 3.4 *If $I < \infty$, then*

$$\max_{N_m < n \le N_{m+1}} \left| \sum_{j=N_m+1}^{n} \hat{X}_j \right| = o(\exp(\frac{1}{2}m^a) \cdot L^{1/2}(b_{[\exp(m^a)]})) \quad a.s. \qquad (15)$$

and hence $\max_{N_{m_n} < n \le N_{m_n+1}} \left| \sum_{j=N_{m_n}+1}^{n} \hat{X}_j \right| = o(\eta_n) \quad a.s.$

Proof: In order to prove (15), it is enough to show that for any $\varepsilon > 0$

$$\sum_{k \ge 1} P\left(\max_{N_k < n \le N_{k+1}} \left| \sum_{j=N_k+1}^{n} \hat{X}_j \right| > \varepsilon c_k^{1/2} \right) < \infty. \qquad (16)$$

where $c_k = \exp(k^a) \cdot L(b_{[\exp(k^a)]})$. We apply Lemma 2.4 of Shao (1993a) with

$$q = \tau, \quad B = k^{-a(\tau+2)/(\tau-2)} c_k^{1/2}, \quad x = \varepsilon c_k^{1/2}$$

$$n = N_{k+1} - N_k, \quad m = [k^{-a(\tau+2)/(\tau-2)} e^{k^a}].$$

For every $j = N_k + 1, \ldots, N_{k+1}$ we have

$$E\hat{X}_j^2 1_{\{|\hat{X}_j| > B\}} = EX^2 1_{\{B < |X| \le b_j\}} \le L(b_j) \le L(b_{N_{k+1}}) \le CL(b_{[\exp(k^a)]}) = C\frac{xB}{m}$$

where we used (13) for the last inequality. Relation (16) follows exactly as (2.20) of Shao (1993a), provided we show that:

$$\sum_{k \ge 1} k^{a-1} e^{-(\tau-2)k^a/2} \cdot L^{-\tau/2}(b_{[\exp(k^a)]}) \cdot E|X|^\tau 1_{\{|X| \le 2b_{[\exp(k^a)]}\}} < \infty \qquad (17)$$

To simplify the notation we let $\beta_j = b_{[\exp(j^a)]}$. We re-write the sum in (17) as

$$\sum_{k \ge 1} k^{a-1} e^{-(\tau-2)k^a/2} \cdot L^{-\tau/2}(\beta_k) \cdot (E|X|^\tau 1_{\{|X| \le 2\beta_0\}} + \sum_{j=1}^{k} E|X|^\tau 1_{\{2\beta_{j-1} < |X| \le 2\beta_j\}})$$

$$\le C + C \sum_{j \ge 1} E|X|^\tau 1_{\{2\beta_{j-1} < |X| \le 2\beta_j\}} \cdot L^{-\tau/2}(\beta_j) \cdot e^{-(\tau-2)j^a/2}$$

$$\le C + C \sum_{j \ge 1} I(\beta_{j-1}, \beta_j] \cdot \beta_j^{\tau-2} A(\beta_j) \cdot L^{-\tau/2}(\beta_j) \cdot e^{-(\tau-2)j^a/2}. \qquad (18)$$

where for the last inequality we used: $E|X|^\tau 1_{\{a < |X| \le b\}} \le I(a, b] \cdot b^{\tau-2} A(b)$. Using Potter's Theorem for the slowly varying functions v and L we get:

$$\frac{v(b_n)}{v(a_n)} \le C \left(\frac{b_n}{a_n} \right)^{-\mu} = v^{p\mu}(a_n) \qquad (19)$$

$$\frac{b_n^2}{nL(b_n)} = \frac{L(a_n)}{L(b_n)} v^{-(2p-1)}(a_n) \le C\left(\frac{a_n}{b_n}\right)^\delta v^{-(2p-1)}(a_n) = Cv^{-(2p-1-p\delta)}(a_n) \tag{20}$$

for any $\mu, \delta > 0$ and n large. Let $\alpha_j = a_{[\exp(j^a)]}$. Using (19) and (20) we get

$$\beta_j^{\tau-2} A(\beta_j) \cdot L^{-\tau/2}(\beta_j) \cdot e^{-(\tau-2)j^a/2} = v(\beta_j)\left(\frac{\beta_j^2}{\exp(j^a)\cdot L(\beta_j)}\right)^{(\tau-2)/2}$$

$$\le Cv^{1+p\mu}(\alpha_j) \cdot v^{-(\tau-2)(2p-1-p\delta)/2}(\alpha_j) = Cv^{-\gamma}(\alpha_j) \le C \tag{21}$$

where we selected μ, δ such that $\gamma := -1 - p\mu + (\tau-2)(2p-1-p\delta)/2 > 0$. Finally from (18) and (21) we conclude that the sum in (17) is smaller than

$$C + C\sum_{j\ge i} I(\beta_{j-1}, \beta_j] \le C + C \cdot I < \infty.$$

This concludes the proof of (17). The second statement in the lemma follows by taking $m = m_n$ in (15) and using (11). □

Lemma 3.5 *We have*

$$\sum_{i=1}^m E(u_i|\mathcal{G}_{i-1}) = o(m^{-(r-1/2)a} \cdot (\log m)^3 \cdot \exp(\frac{1}{2}m^a) \cdot L^{1/2}(b_{N_m})) \quad a.s. \tag{22}$$

and hence $\sum_{i=1}^{m_n} E(u_i|\mathcal{G}_{i-1}) = o(\eta_n)$ *a.s.*

Proof: We begin by noting that relationships (2.24)-(2.26) of Shao (1993a) do not rely on the assumption $EX^2 < \infty$, and therefore they hold true in our case. By Lemma 2.3 of Shao (1993a), we have for every $i = 1, \ldots, m$

$$Eu_i^2 \le C \cdot \text{card}(H_i) \cdot \max_{j\in H_i} E\hat{X}_j^2 \le Ci^{a-1}\exp(i^a) \cdot L(b_{N_m}). \tag{23}$$

Let $j_i = \text{card}(I_i)$. By (4) we have $\rho^2(j_i) \le Ci^{-2ar}$. Using (2.26) of Shao (1993a) and (23), we get:

$$E\max_{l\le m}\left(\sum_{i=1}^l E(u_i|\mathcal{G}_{i-1})\right)^2 \le C(\log m)^4 \cdot L(b_{N_m}) \cdot m^{-2ar}\exp(m^a). \tag{24}$$

Let $T_m = \sum_{i=1}^m E(u_i|\mathcal{G}_{i-1})$, $\alpha_m = m^{-(r-1/2)a} \cdot (\log m)^3 \cdot \exp(\frac{1}{2}m^a) \cdot L^{1/2}(b_{N_m})$ and $m_k = [k^{1/a}]$. Using Chebyshev's inequality and (24) we get

$$\sum_{k\ge1} P(\max_{l\le m_k}|T_l| > \varepsilon\alpha_{m_k}) \le \sum_{k\ge1}\frac{E(\max_{l\le m_k}T_l^2)}{\varepsilon^2\alpha_{m_k}^2} \le C\sum_{k\ge1}\frac{1}{m_k^a(\log m_k)^2} < \infty$$

and hence by a well-known subsequence criterion, $T_m = o(\alpha_m)$ a.s. The second statement in the lemma follows by taking $m = m_n$ in (22) and using (11). □

The next theorem gives us the desired approximation of the first term in (10) with a Brownian motion. Let

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

Theorem 3.6 *If $I < \infty$, then without changing its distribution, we can redefine the sequence $\{\xi_i\}_{i\geq 1}$ on a larger probability space together with a standard Brownian motion $W = \{W(t)\}_{t\geq 0}$ such that*

$$\sum_{i=1}^{m_n} \xi_i - W(s_n^2) = o((\eta_n^2 \log\log \eta_n)^{1/2}) \quad a.s.$$

In order to prove this theorem we need the following two lemmas. To simplify the notation we introduce the sequences

$$c_i = \exp(i^a) \cdot L(b_{[\exp(i^a)]}) \quad \text{and} \quad d_i = \eta^2_{[\exp(i^a)]}.$$

Note that by (11), $c_i \leq C d_i$ for i large.

Lemma 3.7 *If $I < \infty$, then*

$$\sum_{i\geq 1} d_i^{-\tau/2} E|\xi_i|^\tau < \infty.$$

Proof: Since $c_i \leq C d_i$ it is enough to prove the lemma with c_i instead of d_i. Note that $E|\xi_i|^\tau \leq 16 E|u_i|^\tau$. Using Lemma 2.3 of Shao (1993a), we have

$$E|u_i|^\tau \leq C\{(\mathrm{card}(H_i))^{\tau/2} \cdot \max_{j\in H_i}(E\hat{X}_j^2)^{\tau/2} + \mathrm{card}(H_i) \cdot \max_{j\in H_i} E|\hat{X}_j|^\tau\}$$

$$\leq C\left\{(i^{a-1}\exp(i^a))^{\tau/2} \cdot L^{\tau/2}(b_{[\exp(i^a)]}) + i^{a-1}\exp(i^a) \cdot E|X|^\tau 1_{\{|X|\leq 2b_{[\exp(i^a)]}\}}\right\}$$

$$= C c_i^{\tau/2}\left\{i^{-(1-a)\tau/2} + i^{a-1} e^{-(\tau-2)i^a/2} L^{-\tau/2}(b_{[\exp(i^a)]}) E|X|^\tau 1_{\{|X|\leq 2b_{[\exp(i^a)]}\}}\right\}. \tag{25}$$

The first term in the above parenthesis is summable by the choice of a; the second term is summable by (17). □

Lemma 3.8 *If $I < \infty$, then*

$$\sum_{i=1}^{m} (E(\xi_i^2|\mathcal{G}_{i-1}) - E\xi_i^2) = o(d_m) \quad a.s.$$

Proof: It is enough to prove the lemma with c_m instead of d_m. Using the inequality on top of p. 329 of Shao (1993a), the conclusion will follow from:

$$\sum_{i=1}^{m}(E(u_i^2|\mathcal{G}_{i-1}) - Eu_i^2) = o(c_m) \quad \text{a.s.} \tag{26}$$

$$\sum_{i=1}^{m}(E^2(u_i|\mathcal{G}_{i-1}) + EE^2(u_i|\mathcal{G}_{i-1})) = o(m^{-(2r-1)a} \cdot (\log m)^2 \cdot c_m) \quad \text{a.s.} \tag{27}$$

Using (2.32) of Shao (1993a), relationship (26) will follow from:

$$\sum_{i=1}^{m}[E(u_i^{**}|\mathcal{G}_{i-1}) + Eu_i^{**}] = o(c_m) \quad \text{a.s.} \tag{28}$$

$$\sum_{i=1}^{m}[E(u_i^{*}|\mathcal{G}_{i-1}) - Eu_i^{*}] = o(m^{-(r-1/2)a} \cdot (\log m)^3 \cdot c_m) \quad \text{a.s.} \tag{29}$$

where $u_i^* = u_i^2 1_{\{|u_i| \le c_i^{1/2}\}}$ and $u_i^{**} = u_i^2 1_{\{|u_i| > c_i^{1/2}\}}$.

To prove (28), note that $E|u_i|^\tau \ge E|u_i|^\tau 1_{\{|u_i| > c_i^{1/2}\}} \ge c_i^{(\tau-2)/2} Eu_i^{**}$. Relationship (28) follows by Kronecker lemma since by (25) and (17)

$$\sum_{i\ge 1}\frac{Eu_i^{**}}{c_i} \le \sum_{i\ge 1}\frac{E|u_i|^\tau}{c_i^{\tau/2}} < \infty.$$

To prove (29), note that for every $i = 1, \ldots, m$

$$Eu_i^{*2} = Eu_i^4 1_{\{|u_i| \le c_i^{1/2}\}} \le c_i \cdot Eu_i^2 \le C i^{a-1}\exp(i^a) \cdot L(b_{[\exp(m^a)]}) \cdot c_m \tag{30}$$

where we used (23) in the last inequality.

Let $U_m = \sum_{i=1}^{m}(E(u_i^*|\mathcal{G}_{i-1}) - Eu_i^*)$ and $\beta_m = m^{-(r-1/2)a}(\log m)^3 c_m$. By the first inequality in (2.34) of Shao (1993a), Corollary 4 of Moricz (1982) and (30), we get

$$E(\max_{l\le m} U_l^2) \le C(\log m)^4 \sum_{i=1}^{m}\rho^2(j_i)Eu_i^{*2} \le C(\log m)^4 m^{-2ar}e^{m^a}L(b_{[\exp(m^a)]}) \cdot c_m. \tag{31}$$

Take $m_k = [k^{1/a}]$. By Chebyshev's inequality and (31) we get

$$\sum_{k\ge 1}P(\max_{l\le m_k}|U_l| > \varepsilon\beta_{m_k}) \le \sum_{k\ge 1}\frac{E(\max_{l\le m_k} U_l^2)}{\varepsilon^2\beta_{m_k}^2} \le C\sum_{k\ge 1}\frac{1}{m_k^a(\log m_k)^2} < \infty.$$

and hence $U_m = o(\beta_m)$ a.s. Relation (29) is proved.

It remains to prove (27). By the mixing property, (23) and (4), we have $EE^2(u_i|\mathcal{G}_{i-1}) \le \rho^2(j_i)Eu_i^2 \le Ci^{-(2r-1)a-1}\exp(i^a)\cdot L(b_{[\exp_{i^a}]}) = Ci^{-(2r-1)a-1}c_i$. Hence

$$\sum_{i\ge1} \frac{EE^2(u_i|\mathcal{G}_{i-1})}{i^{-(2r-1)a}\cdot(\log i)^2\cdot c_i} \le \sum_{i\ge1}\frac{C}{i(\log i)^2} < \infty.$$

Relation (27) follows by the Kronecker lemma. □

Proof of Theorem 3.6: By Theorem 2.1 of Shao (1993a) and Lemmas 3.7, 3.8, we can redefine the sequence $\{\xi_i\}_{i\ge1}$ on a larger probability space together with a standard Brownian motion $W = \{W(t)\}_{t\ge0}$ such that

$$\sum_{i=1}^{m}\xi_i - W(s_m^{*2}) = o(\{d_m(\log\frac{s_m^{*2}}{d_m} + \log\log d_m)\}^{1/2}) \quad \text{a.s.} \tag{32}$$

Using the mixing property, (23) and (11), we have

$$s_m^{*2} = \sum_{i=1}^{m}Eu_i^2 - \sum_{i=1}^{m}E(u_iE(u_i|\mathcal{G}_{i-1})) \le C\sum_{i=1}^{m}Eu_i^2 \le C\exp(m^a)\cdot L(b_{N_m}) \le C\eta_{N_m}^2 \tag{33}$$

The result follows from (32) and (33) by taking $m = m_n$ and noting that $d_{m_n} = \eta_n^2$. □

4 Between the Two Truncations

In this section we will prove that (9) holds.

As in the previous section we define

$$u_i' = \sum_{j\in H_i}X_j', \quad v_i' = \sum_{j\in I_i}X_j', \quad \xi_i' = u_i' - E(u_i'|\mathcal{G}_{i-1}')$$

where $\mathcal{G}_m' = \sigma(\{u_i'; i\le m\})$. We have

$$S_n' = \sum_{i=1}^{m_n}\xi_i' + \sum_{i=1}^{m_n}E(u_i'|\mathcal{G}_{i-1}') + \sum_{i=1}^{m_n}v_i' + \sum_{j=N_{m_n}+1}^{n}X_j'. \tag{34}$$

We will prove that all the 4 terms in the above decomposition are of order $o(a_n)$.

We begin by treating the last three terms in (34). We will use the following facts: $EX_j'^2 = L(a_j) - L(b_j) \le L(a_j)$ and

$$nL(a_n) \le Ca_n^2. \tag{35}$$

Lemma 4.1 *We have*

$$\sum_{i=1}^{m} v_i' = o(m^2 \cdot \exp(\frac{1}{3}m^a) \cdot L^{1/2}(a_{N_m})) \quad a.s.$$

and hence $\sum_{i=1}^{m_n} v_i' = o(a_n)$ *a.s.*

Proof: Same argument as in Lemma 3.3 by replacing b_{N_m} with a_{N_m} and using (35) instead of (11). □

Lemma 4.2 *If the function L satisfies (C), then*

$$\max_{N_m < n \le N_{m+1}} \left| \sum_{j=N_m+1}^{n} X_j' \right| = o(\exp(\frac{1}{2}m^a) \cdot L^{1/2}(a_{[\exp(m^a)]})) \quad a.s. \tag{36}$$

and hence $\max_{N_{m_n} < n \le N_{m_n+1}} \left| \sum_{j=N_{m_n}+1}^{n} X_j' \right| = o(a_n) \quad a.s.$

Proof: The second statement in the lemma follows from (36) by taking $m = m_n$ and using (35). To prove (36) we employ the same argument as in Lemma 3.4, this time making use of relation (12). Hence it suffices to show that

$$\sum_{k \ge 1} k^{a-1} e^{-(\tau-2)k^a/2} \cdot L^{-\tau/2}(a_{[\exp(k^a)]}) \cdot E|X|^\tau 1_{\{|X| \le 2a_{[\exp(k^a)]}\}} < \infty. \tag{37}$$

Let $\alpha_j = a_{[\exp(j^a)]}$. Similarly to the proof of (17), we conclude that the sum in (37) is smaller than

$$C + C \sum_{j \ge 1} E|X|^\tau 1_{\{2\alpha_{j-1} < |X| \le 2\alpha_j\}} \cdot L^{-\tau/2}(\alpha_j) \cdot e^{-(\tau-2)j^a/2} \le$$

$$C + C \sum_{j \ge 1} (L(2\alpha_j) - L(2\alpha_{j-1})) \cdot \alpha_j^{\tau-2} \cdot L^{-\tau/2}(\alpha_j) \cdot e^{-(\tau-2)j^a/2} \tag{38}$$

where we used the inequality: $E|X|^\tau 1_{\{a < |X| \le b\}} \le (L(b) - L(a)) b^{\tau-2}$. Note that

$$\alpha_j^{\tau-2} \cdot L^{-\tau/2}(\alpha_j) \cdot e^{-(\tau-2)j^a/2} = L^{-1}(\alpha_j) \left(\frac{\alpha_j^2}{\exp(j^a) \cdot L(\alpha_j)} \right)^{(\tau-2)/2}$$

$$\le C L^{-1}(2\alpha_j) \cdot v^{(\tau-2)/2}(\alpha_j) \tag{39}$$

From (38) and (39) we conclude that the sum in (37) is smaller than

$$C + C \sum_{j \ge 1} \left[1 - \frac{L(2\alpha_{j-1})}{L(2\alpha_j)} \right] \cdot v^{(\tau-2)/2}(\alpha_j) \tag{40}$$

By the Representation Theorem (Theorem 1.3.1 of Bingham, Goldie and Teugels, 1987) for the slowly varying function v, we have: $\forall \delta > 0\ \exists C = C_\delta > 0$ such that

$$v(x) \le Cx^\delta \text{ for } x \text{ large.} \tag{41}$$

Using (C) and (41), the sum in (40) becomes smaller than

$$C + C\sum_{j\ge 1} \alpha_j^{-K+(\tau-2)\delta/2} \le C + C\sum_{j\ge 1} \exp(-\frac{K_0}{2} j^a) < \infty$$

where we chose δ such that $K_0 := K - (\tau-2)\delta/2 > 0$ and we used the fact that $a_n \ge Cn^{1/2}$ for n large (consequence of (35)). This concludes the proof of (37). $\square$

Lemma 4.3 *We have*

$$\sum_{i=1}^{m} E(u_i'|\mathcal{G}_{i-1}') = o(m^{-(r-1/2)a} \cdot (\log m)^3 \cdot \exp(\frac{1}{2}m^a) \cdot L^{1/2}(a_{N_m})) \quad a.s.$$

and hence $\sum_{i=1}^{m_n} E(u_i'|\mathcal{G}_{i-1}') = o(a_n)$ *a.s.*

Proof: Same argument as in Lemma 3.5 by replacing b_{N_m} with a_{N_m} and using (35). $\square$

Our last result treats the first term in the decomposition (34).

Theorem 4.4 *If* $I < \infty$*, then*

$$\sum_{i=1}^{m_n} \xi_i' = o(a_n) \quad a.s.$$

In order to prove this result, we will use the following martingale subsequence criterion, which is probably well-known. Its proof is given in Appendix B.

Lemma 4.5 *Let* $\{S_n, \mathcal{F}_n\}_{n\ge 1}$ *be a zero-mean martingale and* $\{a_n\}_{n\ge 1}$ *a nondecreasing sequence of positive numbers with* $a_n \uparrow \infty$*. If there exists a subsequence* $\{n_k\}_k$ *such that* $a_{n_{k+1}}/a_{n_k} \le C$ *for all* k *and*

$$\sum_{k\ge 1} \frac{E|S_{n_k} - S_{n_{k-1}}|^p}{a_{n_k}^p} < \infty \quad \text{for some } p \in [1,2] \tag{42}$$

then $S_n = o(a_n)$ *a.s.*

Proof of Theorem 4.4: Let $U_n := \sum_{i=1}^{m_n} \xi_i'$ and note that $\{U_n, \mathcal{G}_{m_n}'\}_{n\ge 1}$ is a zero-mean martingale. By Lemma 4.5, it is enough to prove that for a suitable subsequence $\{n_k\}_k$ we have

$$\sum_{k\ge 1} \frac{E|U_{n_k} - U_{n_{k-1}}|^2}{a_{n_k}^2} < \infty \tag{43}$$

Similarly to the proof of Lemma 2.3 of Mijnheer (1980), we take a subsequence $\{n_k\}_k$ satisfying $n_k \sim n_{k-1}(1+\phi^{-1}(k))$, where the function ϕ is chosen such that $\lim_{k\to\infty}\phi(k)=\infty$ and

$$\frac{1}{\phi(k)+1}\cdot I(b_{n_k},a_{n_k}] \le CI(a_{n_{k-1}},a_{n_k}] \tag{44}$$

Clearly $n_k \sim n_{k+1}$ and hence, using the definition of a_n and the slowly varying properties of the functions L and v, we obtain that $a_{n_k}\sim a_{n_{k+1}}$ and $b_{n_k}\sim b_{n_{k+1}}$.

We proceed now with the proof of (43). Let

$$Z_k := U_{n_k} - U_{n_{k-1}} = \sum_{m_{n_{k-1}}<i\le m_{n_k}} \xi_i'.$$

By the martingale property

$$EZ_k^2 = \sum_{m_{n_{k-1}}<i\le m_{n_k}} E\xi_i^{'2} \le (m_{n_k}-m_{n_{k-1}}) \max_{m_{n_{k-1}}<i\le m_{n_k}} E\xi_i^{'2}. \tag{45}$$

Using Lemma 2.3 of Shao (1993a) we have: for every $m_{n_{k-1}}<i\le m_{n_k}$

$$E\xi_i^{'2} \le Eu_i^{'2} \le Ci^{a-1}e^{i^a}\cdot\max_{j\in H_i} EX_j^{'2} \le C(\log n_k)^{(a-1)/a} n_k \cdot \max_{j\in H_i} EX_j^{'2}. \tag{46}$$

Now for any $j\in H_i$ and $m_{n_{k-1}}<i\le m_{n_k}$ we have

$$EX_j^{'2} \le A(a_j)I(b_j,a_j] \le A(a_{N_i})I(b_{N_{i-1}},a_{N_i}] \le CA(a_{n_k})I(b_{n_k},a_{n_k}]. \tag{47}$$

Using (47) and (46) we get: for every $m_{n_{k-1}}<i\le m_{n_k}$

$$E\xi_i^{'2} \le C(\log n_k)^{(a-1)/a} n_k\cdot A(a_{n_k})I(b_{n_k},a_{n_k}] = C(\log n_k)^{(a-1)/a}a_{n_k}^2 I(b_{n_k},a_{n_k}]. \tag{48}$$

From (45) and (48) it follows that

$$\begin{aligned}\frac{EZ_k^2}{a_{n_k}^2} &\le C[(\log n_k)^{1/a}-(\log n_{k-1})^{1/a}]\cdot(\log n_k)^{(a-1)/a}I(b_{n_k},a_{n_k}]\\ &\le C(\log n_{k-1})^{(1-a)/a}\frac{1}{n_{k-1}}(n_k-n_{k-1})\cdot(\log n_k)^{(a-1)/a}I(b_{n_k},a_{n_k}]\\ &\le C\frac{n_k-n_{k-1}}{n_{k-1}}I(b_{n_k},a_{n_k}] \le C\frac{1}{\phi(k)+1}I(b_{n_k},a_{n_k}] \le CI(a_{n_{k-1}},a_{n_k}]\end{aligned}$$

where we used the inequality $f(y)-f(x)\le f'(x)(y-x)$ for the concave function $f(x)=(\log x)^{1/a}$ for the second inequality, and the choice (44) of the function ϕ for the last inequality. Relationship (43) follows since $I<\infty$. This concludes the proof of the theorem.□

Acknowledgement. The first author would like to thank Professor Miklós Csörgő for a series of wonderful lectures given at Carleton University throughout the academic year 2003-2004, which stimulated her interest in the area of self-normalized limit theorems.

A Appendix A

Proof of Proposition 3.1: The relation $b_n \le C\eta_n$ for n large, can be written as $a_n/\eta_n \le Cv^p(a_n)$ for n large; using the definitions of a_n and η_n, this in turn is equivalent to:

$$\frac{L(a_n)}{L(\eta_n)} \le Cv^{2p-1}(a_n) \quad \text{for } n \text{ large.} \tag{49}$$

Since L is slowly varying, it follows by Potter's Theorem (Theorem 1.5.6.(i) of Bingham, Goldie and Teugels, 1987) that for any $C > 1, \delta > 0$ we have

$$\frac{L(a_n)}{L(\eta_n)} \le C\left(\frac{a_n}{\eta_n}\right)^{\delta} = C\left(\frac{L(a_n)v(a_n)}{L(\eta_n)}\right)^{\delta/2} \quad \text{for } n \text{ large}$$

and hence

$$\left(\frac{L(a_n)}{L(\eta_n)}\right)^{1-\delta/2} \le Cv^{\delta/2}(a_n) \quad \text{for } n \text{ large.}$$

This is exactly relation (49) with $\delta = 2 - 1/p$. Relationship (11) follows using the fact that L is nondecreasing and slowly varying, and the definition of η_n:

$$nL(b_n) \le nL(C\eta_n) \le CnL(\eta_n) = C\eta_n^2.$$

□

Proof of Proposition 3.2: Using the definition of a_n and Potter's theorem for the slow varying function A, we get that for any $C > 1, \delta \in (0, 2)$

$$\frac{a_{\lambda n}^2}{a_n^2} = \frac{\lambda n A(a_{\lambda n})}{nA(a_n)} \le \lambda C\left(\frac{a_{\lambda n}}{a_n}\right)^{\delta} \quad \text{for } n \text{ large}$$

and hence

$$\frac{a_{\lambda n}}{a_n} \le C\lambda^{1/(2-\delta)} \quad \text{for } n \text{ large.}$$

Using the definition of b_n and Potter's theorem for the slowly varying function v, we get that for any $C > 1, \varepsilon > 0$

$$\frac{b_{\lambda n}}{b_n} = \frac{a_{\lambda n}}{a_n} \cdot \left(\frac{v(a_n)}{v(a_{\lambda n})}\right)^p \le C\left(\frac{a_{\lambda n}}{a_n}\right)^{1+p\varepsilon} \le C\lambda^{(1+p\varepsilon)/(2-\delta)} \quad \text{for } n \text{ large.}$$

Relationships (12) and (13) follow by the slowly varying property of L. □

B Appendix B

Proof of Lemma 4.5: Note that $\{S_{n_k}, \mathcal{F}_{n_k}\}_{k \ge 1}$ is a martingale. From (42) it follows that $S_{n_k}/a_{n_k} \to 0$ a.s. (see Theorem 2.18 of Hall and Heyde, 1980). By the extended Kolmogorov inequality (see p. 65 of Loève, 1978), we have

$$\sum_{k\geq 1} P(\max_{n_{k-1}<n\leq n_k} |S_n - S_{n_k}| > \varepsilon a_{n_k}) \leq \sum_{k\geq 1} \frac{E|S_{n_k} - S_{n_{k-1}}|^p}{\varepsilon^p a_{n_k}^p} < \infty$$

for every $\varepsilon > 0$, and hence

$$T_k = \max_{n_{k-1}<n\leq n_k} \frac{|S_n - S_{n_k}|}{a_{n_k}} \to 0 \quad \text{a.s.}$$

Finally for $n_{k-1} < n \leq n_k$ we have:

$$\frac{|S_n|}{a_n} \leq \frac{|S_{n_{k-1}}|}{a_{n_{k-1}}} + \frac{|S_n - S_{n_{k-1}}|}{a_{n_{k-1}}} \leq \frac{|S_{n_{k-1}}|}{a_{n_{k-1}}} + \frac{a_{n_k}}{a_{n_{k-1}}} \cdot T_k \leq \frac{|S_{n_{k-1}}|}{a_{n_{k-1}}} + C \cdot T_k \to 0 \quad \text{a.s.}$$

□

References

[1] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*, Cambridge University Press, Cambridge.

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

[3] Feller, W. (1968). An extension of the law of the iterated logarithm to variables without variance. *J. Math. Mech.* **18**, 343-355.

[4] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Applications*, Academic Press, New York.

[5] Kesten, H. (1972). Sums of independent random variables - without moment conditions. *Ann. Math. Stat.* **43**, 701-732.

[6] Loève, M. (1978). *Probability Theory II*, 4th Edition, Springer, New York.

[7] Mijnheer, J. (1980). A strong approximation of partial sums of i.i.d. random variables with infinite variance. *Z. Wahrsch. verw. Gebiete* **52**, 1-7.

[8] Moricz, F. (1982). A general moment inequality for the maximum of partial sums of single series. *Acta Sci. Math. Szeged* **44**, 67-75.

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

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