Elect. Comm. in Probab. **11** (2006), 11–23

ELECTRONIC COMMUNICATIONS in PROBABILITY

STRONG APPROXIMATION FOR MIXING SEQUENCES WITH INFINITE VARIANCE

RALUCA BALAN[1]
Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Avenue, Ottawa, ON, K1N 6N5, Canada.
email: rbalan@uottawa.ca

INGRID-MONA ZAMFIRESCU
Department of Mathematics, Baruch College, City University of New York, One Bernard Baruch Way, New York, NY 10010, USA.
email: Ingrid-Mona_Zamfirescu@baruch.cuny.edu

Submitted 20 December 2004, *accepted in final form* 14 January 2006

AMS 2000 Subject classification: Primary 60F15, 60F17; Secondary 60E05.
Keywords: mixing sequence; domain of attraction of the normal law; truncation; blocking technique; Skorohod embedding.

Abstract
In this paper we prove a strong approximation result for a mixing sequence with infinite variance and logarithmic decay rate of the mixing coefficient. The result is proved under the assumption that the distribution is symmetric and lies in the domain of attraction of the normal law. Moreover the function $L(x) = EX^2 1_{\{|X| \leq x\}}$ is supposed to be slowly varying with remainder $(\log x)^{-\alpha}(\log\log)^{-\beta}(x)$ with $\alpha, \beta > 1$.

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 (see e.g. the recent survey article [3]). One of the most important results in this area is Shao's strong invariance principle [14], from which one can easily deduce many other limit theorems.

In this paper we prove a strong approximation result for a mixing sequence of identically distributed random variables with infinite variance, whose distribution is symmetric and lies in the domain of attraction of the normal law (DAN). This suggests that it may be possible to obtain a similar result 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

[1]RESEARCH SUPPORTED BY NSERC, CANADA

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; see e.g. [1], or [12].

Suppose first that $\{X_n\}_{n\geq 1}$ is a sequence of i.i.d. random variables with $S_n = \sum_{i=1}^n X_i$ and $EX = 0,\ EX^2 = \infty$ (here X denotes a generic random variable with the same distribution as X_n). If $X \in DAN$ (or equivalently, the function $L(x) = EX^2 1_{\{|X|\leq x\}}$ is slowly varying), then 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

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

(see e.g. [6], IX.8, XVII.5). Moreover, by Theorem 1 of [5], if the distribution of X is *symmetric* then

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

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 i.i.d. random variables with infinite variance.

In the 1971 Rietz Lecture, Kesten has discussed Feller's result and raised the question of its correctness; see his Remark 9, [8]. Fortunately, he settled this problem, 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 $\gamma_n^2 = CnL(\gamma_n)\log\log\gamma_n$ (see Theorem 7). A more general form of the law of the iterated logarithm for the "trimmed" sum $S_n^{(r)}$ (i.e. the sum obtained by deleting from S_n the r-th largest terms) has been recently obtained in [9].

Following these lines, Theorem 2.1 of [11] 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 $\lim_{x\to\infty} v(x) = \infty$ and

$$I := I_{v(\cdot)} = \int_b^\infty \frac{x^2}{L(x)v(x)}\, dF(x) < \infty. \tag{4}$$

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

We recall that 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 $X_a, X_{a+1}, \ldots, X_b$.

Here is our result.

Theorem 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 $X \in DAN$, where X denotes a random variable with the same distribution as X_n. Assume that*

$$\rho(n) \leq C(\log n)^{-r} \text{ for some } r > 1. \tag{5}$$

Let v be a nondecreasing slowly varying function such that $v(x) \geq C \log\log x$ for x large; let $\tau = \min(3, r+1)$. Suppose that the function $L(x) = EX^2 1_{\{|X|\leq x\}}$ satisfies (4) and is slowly varying with remainder $(\log x)^{-\alpha} v^{-\beta}(x)$ for some $\alpha > \tau/(\tau-2)$, $\beta > \tau/2$, i.e. for any $\lambda \in (0,1)$ there exists $C > 0$ such that

$$(SR) \qquad 1 - \frac{L(\lambda x)}{L(x)} \leq C(\log x)^{-\alpha} v^{-\beta}(x) \text{ for } x \text{ large.}$$

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

$$S_n - W(s_n^2) = o(a_n) \quad a.s. \tag{6}$$

where $\{a_n\}_n$ is a nondecreasing sequence of positive numbers satisfying (3).

Condition (SR) specifies the rate of convergence of $L(\lambda x)/L(x)$ to 1, for the slowly varying function L (see p.185 of [2]). It was used in *only one* place, namely to ensure the convergence of the sum (28) in the proof of Lemma 11. Unfortunately, we could not avoid it.

We should mention here that a "functional central limit theorem" for ρ-mixing sequences with infinite variance was obtained by [15] under the condition $\sum_n \rho(2^n) < \infty$. In order to obtain the strong approximation (6) we needed to impose the logarithmic decay rate of $\rho(n)$.

The remaining part of the paper is dedicated to the proof of Theorem 1: the description of the general method is given in Section 2, while the technical details are discussed in Sections 3 and 4. Among other ingredients, the proof uses the blocking technique introduced in [13], according to which the original random variables are replaced by their sums over progressively larger blocks of integers (separated by smaller blocks, whose length is also progressively larger).

Throughout this work, C denotes a generic constant that does not depend on n but may be different from place to place. We denote by $I(a,b]$ the measure attributed by the integral I to the interval $(a,b]$. We let $A(x) = L(x)v(x)$.

2 Sketch of Proof

As in [4] we may take

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

Clearly (1) and (3) hold. We have $a_n \geq \eta_n$ and

$$a_n^2 \geq C\eta_n^2 v(\eta_n) \geq C\eta_n^2 \log\log \eta_n. \tag{7}$$

Without loss of generality we will assume that $\eta_n^2 = nL(\eta_n)$ and $a_n^2 = nA(a_n)$.

The proof is based on a double truncation technique at levels $b_n := v^{-p}(a_n)a_n$ and a_n (which is due to [5]), and a repeated application of the method of [14] on each of the "truncation" intervals $[0, b_n]$, $(b_n, a_n]$.

We assume that $p > 1/2$. Let

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

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

$$S_n = \hat{S}_n + S'_n + \bar{S}_n \tag{8}$$

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 [5] (under the *symmetry* assumption), (4) is equivalent to $\sum_{n\geq 1} P(|X| > \epsilon a_n) < \infty$ for all $\epsilon > 0$. Hence

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

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

$$\hat{S}_n - W(s_n^2) = o((\eta_n^2 \log\log \eta_n)^{1/2}) \quad \text{a.s.} \tag{10}$$

for some constants s_n^2. In Section 4 we show that between the two truncations we have

$$S'_n = o(a_n) \quad \text{a.s.} \tag{11}$$

The conclusion (6) follows immediately by (7)-(11).

3 The Central Part

The goal of this section is to prove relation (10) on a possibly larger probability space on which the sequence $\{\hat{X}_n\}_n$ is redefined (without changing its distribution). In order to do this, we introduce the blocks $H_1, I_1, H_2, I_2, \ldots$ of consecutive integers and we decompose the sum $\hat{S}_n$ into three terms containing the sums over the "small" blocks I_i, the sums over the "big" blocks H_i, and the remaining $\hat{X}_j$'s (whose sum is shown to be negligible). The idea is to construct the blocks I_i small enough to make the term depending on these blocks negligible, but large enough to give sufficient space between the blocks H_i. The sums u_i over the blocks H_i will provide us with the desired approximation (10), by applying an almost sure invariance principle (due to [14]) to the martingale differences $\xi_i = u_i - E(u_i | u_1, \ldots, u_{i-1})$, after proving that the sum of the terms $E(u_i | u_1, \ldots, u_{i-1})$ is negligible as well.

We define the blocks $H_1, I_1, H_2, I_2, \ldots$ of consecutive integers such that

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

with $a = 1/\alpha$. Note that $(1-a)\tau > 2$. Let $N_m := \sum_{i=1}^m \mathrm{card}(H_i \cup I_i) \sim \exp(m^a)$ and $N_{m_n} \le n < N_{m_n+1}$. Clearly $N_{m_n} \sim n$, $m_n \sim (\log n)^{1/a}$.

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 $\mathcal{G}_m = \sigma(\{u_i; i \le m\})$, and write

$$\hat{S}_n = \sum_{i=1}^{m_n} v_i + \sum_{j=N_{m_n}+1}^{n} \hat{X}_j + \sum_{i=1}^{m_n} E(u_i|\mathcal{G}_{i-1}) + \sum_{i=1}^{m_n} \xi_i. \tag{12}$$

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

We begin with two elementary lemmas.

Lemma 2 *There exists $C > 0$ such that $b_n \le C\eta_n$ for n large, and hence*

$$nL(b_n) \le C\eta_n^2 \quad \textit{for } n \textit{ large.} \tag{13}$$

Proof. 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:

$$\frac{L(a_n)}{L(\eta_n)} \le Cv^{2p-1}(a_n) \quad \text{for } n \text{ large.} \tag{14}$$

Since L is slowly varying, it follows by Potter's Theorem (Theorem 1.5.6.(i) of [2]) 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 (14) with $\delta = 2 - 1/p$. Relationship (13) follows using the fact that L is nondecreasing and slowly varying, and the definition of η_n: $nL(b_n) \le nL(C\eta_n) \le CnL(\eta_n) = C\eta_n^2$. □

Lemma 3 *For any integer $\lambda > 0$ there exists $C = C_\lambda > 0$ such that $a_{\lambda n} \le Ca_n$ and $b_{\lambda n} \le Cb_n$ for n large, and hence*

$$L(a_{\lambda n}) \le CL(a_n) \textit{ and } L(b_{\lambda n}) \le CL(b_n) \quad \textit{for } n \textit{ large.} \tag{15}$$

Proof. Using the definition of a_n and Potter's theorem, we get: 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 $a_{\lambda n}/a_n \le C\lambda^{1/(2-\delta)}$ for n large. By the definition of b_n and Potter's theorem, we have: for any $C > 1, \varepsilon > 0$

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

The last statement in the lemma follows since L is slowly varying. □

We are now ready to treat the first three terms in the decomposition (12).

Lemma 4 *We have* $\sum_{i=1}^m v_i = o(m^2 \exp(\frac{1}{3}m^a)L^{1/2}(b_{N_m}))$ *a.s. and hence* $\sum_{i=1}^{m_n} v_i = o(\eta_n)$ *a.s.*

Proof. We have $Ev_i^2 \le C\text{card}(I_i) \cdot \max_{j\in I_i} E\hat{X}_j^2 \le Ci^{a-1}\exp(\frac{1}{2}i^a)L(b_{N_m})$ for all $i \le m$. Hence $E(\sum_{i=1}^m v_i)^2 \le m\sum_{i=1}^m Ev_i^2 \le CmL(b_{N_m})\exp(\frac{2}{3}m^a)$. The first statement in the lemma follows by the Chebyshev's inequality and the Borel-Cantelli lemma. The second statement follows using $m_n \sim (\log n)^{1/a}$ and relation (13). □

To simplify the notation, we let $c_i = \exp(i^a)L(b_{[\exp(i^a)]})$ and $d_i = \eta^2_{[\exp(i^a)]}$. By (13), $c_i \le Cd_i$ for i large.

Lemma 5 *We have* $\max_{N_m<n\le N_{m+1}} \left|\sum_{j=N_m+1}^n \hat{X}_j\right| = o(c_m^{1/2})$ *a.s. and hence* $\max_{N_{m_n}<n\le N_{m_n+1}} \left|\sum_{j=N_{m_n}+1}^n \hat{X}_j\right| = o(\eta_n)$ *a.s.*

Proof. The second statement follows by (13). For the first part, it is enough to prove that for any $\varepsilon > 0$

$$\sum_{k\ge1} P\left(\max_{N_k<n\le N_{k+1}}\left|\sum_{j=N_k+1}^n \hat{X}_j\right| > \varepsilon c_k^{1/2}\right) < \infty. \tag{16}$$

For this we apply Lemma 2.4 of [14] 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 $j = N_k+1, \dots, N_{k+1}$ we have $E\hat{X}_j^2 1_{\{|\hat{X}_j|>B\}} = EX^2 1_{\{B<|X|\le b_j\}} \le L(b_j) \le L(b_{N_{k+1}}) \le CL(b_{[\exp(k^a)]}) = C(xB)/m$, using (15) for the last inequality. Relation (16) follows as (2.20) of [14] provided we show that:

$$\sum_{k\ge1} k^{a-1}e^{-(\tau-2)k^a/2}\ L^{-\tau/2}(b_{[\exp(k^a)]})E|X|^\tau 1_{\{|X|\le 2b_{[\exp(k^a)]}\}} < \infty. \tag{17}$$

Let $\alpha_j = a_{[\exp(j^a)]}$ and $\beta_j = b_{[\exp(j^a)]}$. The sum in (17) becomes

$$\sum_{k\ge1} k^{a-1}e^{-(\tau-2)k^a/2}L^{-\tau/2}(\beta_k)(E|X|^\tau 1_{\{|X|\le 2\beta_0\}} + \sum_{j=1}^k E|X|^\tau 1_{\{2\beta_{j-1}<|X|\le 2\beta_j\}})$$

$$\le C\sum_{j\ge1} E|X|^\tau 1_{\{2\beta_{j-1}<|X|\le 2\beta_j\}}L^{-\tau/2}(\beta_j)e^{-(\tau-2)j^a/2}$$

$$\leq C \sum_{j \geq 1} I(\beta_{j-1}, \beta_j] \cdot \beta_j^{\tau-2} A(\beta_j) L^{-\tau/2}(\beta_j) e^{-(\tau-2)j^a/2}, \tag{18}$$

where for the last inequality we used: $E|X|^\tau 1_{\{a<|X|\leq b\}} \leq I(a,b] \cdot b^{\tau-2}A(b)$. By Potter's Theorem (Theorem 1.5.6.(i) of [2]):

$$\frac{v(b_n)}{v(a_n)} \leq C \left(\frac{b_n}{a_n}\right)^{-\mu} = Cv^{p\mu}(a_n),$$

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

for any $\mu, \delta > 0$ and n large. Hence

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

$$\leq Cv^{1+p\mu}(\alpha_j) \cdot v^{-(\tau-2)(2p-1-p\delta)/2}(\alpha_j) = Cv^{-\gamma}(\alpha_j) \leq C, \tag{19}$$

where we selected μ, δ such that $\gamma := -1 - p\mu + (\tau-2)(2p-1-p\delta)/2 > 0$. From (18), (19) we see that the sum in (17) is smaller than $C \sum_{j \geq i} I(\beta_{j-1}, \beta_j] < \infty$, using (4). □

Lemma 6 *We have* $\sum_{i=1}^m E(u_i|\mathcal{G}_{i-1}) = o(m^{-(r-1/2)a} \cdot (\log m)^3 \cdot \exp(\frac{1}{2}m^a) \cdot L^{1/2}(b_{N_m}))$ *a.s. and hence* $\sum_{i=1}^{m_n} E(u_i|\mathcal{G}_{i-1}) = o(\eta_n)$ *a.s.*

Proof. Let $T_m = \sum_{i=1}^m E(u_i|\mathcal{G}_{i-1})$ and $\alpha_m = m^{-(r-1/2)a}(\log m)^3 \exp(\frac{1}{2}m^a) \cdot L^{1/2}(b_{N_m})$. For any $i \leq m$ we have

$$Eu_i^2 \leq C \cdot \mathrm{card}(H_i) \cdot \max_{j \in H_i} E\hat{X}_j^2 \leq Ci^{a-1} \exp(i^a) L(b_{N_m}). \tag{20}$$

By (2.26) of [14] and (5), we get: $E(\max_{l \leq m} T_l^2) \leq C(\log m)^4 m^{-2ar} \cdot \exp(m^a) L(b_{N_m})$. Let $m_k = [k^{1/a}]$. Using Chebyshev's inequality we get

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

From here we conclude that

$$\sum_{k \geq 1} P(T_{m_k} > \varepsilon \alpha_{m_k}) < \infty \quad \text{and} \quad \sum_{k \geq 1} P(\max_{m_{k-1} < m \leq m_k} |T_m - T_{m_k}| > \varepsilon \alpha_{m_k}) < \infty$$

for all $\varepsilon > 0$. By the Borel-Cantelli lemma, it follows that

$$\frac{T_{m_k}}{\alpha_k} \to 0 \quad \text{a.s.}, \qquad \max_{m_{k-1} < m \leq m_k} \frac{|T_m - T_{m_k}|}{\alpha_{m_k}} \to 0 \quad \text{a.s.}$$

and hence $T_m/\alpha_m \to 0$ a.s. □

Our last theorem gives us the desired approximation for the last term in (12). To prove this theorem we need two lemmas. Let $\sigma_i^{*2} = E\xi_i^2$, $s_m^{*2} = \sum_{i=1}^m \sigma_i^{*2}$, $s_n^2 = s_{m_n}^{*2}$.

Lemma 7 *We have* $\sum_{i\geq 1} d_i^{-\tau/2} E|\xi_i|^\tau < \infty$.

PROOF. It is enough to prove the lemma with c_i instead of d_i, and u_i instead of ξ_i. By Lemma 2.3 of [14] we have

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

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

$$Cc_i^{\tau/2}\left\{i^{-(1-a)\tau/2} + i^{a-1}e^{-(\tau-2)i^a/2} L^{-\tau/2}(b_{[\exp(i^a)]}) E|X|^\tau 1_{\{|X|\leq 2b_{[\exp(i^a)]}\}}\right\}. \quad (21)$$

The lemma follows by (17). □

Lemma 8 *We have* $\sum_{i=1}^m (E(\xi_i^2|\mathcal{G}_{i-1}) - E\xi_i^2) = o(d_m)$ *a.s.*

PROOF. It is enough to prove the lemma with c_m instead of d_m. Let $u_i^* = u_i^2 1_{\{|u_i|\leq c_i^{1/2}\}}$ and $u_i^{**} = u_i^2 1_{\{|u_i|> c_i^{1/2}\}}$. The conclusion will follow from:

$$\sum_{i=1}^m (E(u_i^{**}|\mathcal{G}_{i-1}) + Eu_i^{**}) = o(c_m) \quad \text{a.s.} \quad (22)$$

$$U_m := \sum_{i=1}^m (E(u_i^*|\mathcal{G}_{i-1}) - Eu_i^*) = o(m^{-(r-1/2)a}(\log m)^3 c_m) \quad \text{a.s.} \quad (23)$$

$$\sum_{i=1}^m (E^2(u_i|\mathcal{G}_{i-1}) + EE^2(u_i|\mathcal{G}_{i-1})) = o(m^{-(2r-1)a}(\log m)^2 c_m) \quad \text{a.s.} \quad (24)$$

To prove (22), note that $E|u_i|^\tau \geq E|u_i|^\tau 1_{\{|u_i|>c_i^{1/2}\}} \geq c_i^{(\tau-2)/2} Eu_i^{**}$. Relationship (22) follows by Kronecker's lemma, (21) and (17).
To prove (23), let $\beta_m = m^{-(r-1/2)a}(\log m)^3 c_m$. For any $i \leq m$

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

where we used (20) in the last inequality. By (2.34) of [14] and (5),we get: $E(\max_{l\leq m} U_l^2) \leq C(\log m)^4 m^{-2ar}\exp(m^a) L(b_{[\exp(m^a)]}) c_m$. Let $m_k = [k^{1/a}]$. Using the same argument based on a subsequence convergence criterion as in the proof of Lemma 6, we get $U_m = o(\beta_m)$ a.s.
It remains to prove (24). By the mixing property, (20) and (5), we have $EE^2(u_i|\mathcal{G}_{i-1}) \leq Ci^{-(2r-1)a-1}\exp(i^a) L(b_{[\exp_{i^a}]}) = Ci^{-(2r-1)a-1} c_i$. Relation (24) follows by the Kronecker's lemma. □

Here is the main result of this section.

Theorem 9 *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*

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

PROOF. By Theorem 2.1 of [14], Lemma 7 and Lemma 8, 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

$$\sum_{i=1}^{m}\xi_i - W(s_m^{*2}) = o(\{d_m(\log\frac{s_m^{*2}}{d_m}+\log\log d_m)\}^{1/2}) \quad \text{a.s.} \tag{25}$$

Using the mixing property, (20) and (13), we obtain:

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

The result now follows by noting that $d_{m_n}=\eta_n^2$. □

4 Between the Two Truncations

This section is dedicated to the proof of relation (11): $S_n'/a_n\to 0$ a.s. For this we consider the same blocks H_i, I_i as in Section 3 and we decompose the sum S_n' into three components, depending on the sums over the blocks I_i, the sums over the blocks H_i and the remaining terms X_j'. The sums u_i' over the blocks H_i are once again approximated by the corresponding martingale differences ξ_i' and relation (11) follows by a martingale subsequence convergence criterion.

Let $H_1, I_1, H_2, I_2, \ldots$ be the blocks introduced in Section 3. 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 $\mathcal{G}_m'=\sigma(\{u_i'; i\leq m\})$, and write

$$S_n'=\sum_{i=1}^{m_n}v_i'+\sum_{j=N_{m_n}+1}^{n}X_j'+\sum_{i=1}^{m_n}E(u_i'|\mathcal{G}_{i-1}')+\sum_{i=1}^{m_n}\xi_i'. \tag{26}$$

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

We begin by treating the first three terms. Note that $EX_j'^2=L(a_j)-L(b_j)\leq L(a_j)$ and $nL(a_n)\leq Ca_n^2$.

Lemma 10 *We have* $\sum_{i=1}^{m}v_i'=o(m^2\exp(\frac{1}{3}m^a)L^{1/2}(a_{N_m}))$ *a.s. and hence* $\sum_{i=1}^{m_n}v_i'=o(a_n)$ *a.s.*

PROOF. Same argument as in Lemma 4. □

Lemma 11 *We have* $\max_{N_m<n\leq N_{m+1}}\left|\sum_{j=N_m+1}^{n}X_j'\right|=o(\exp(\frac{1}{2}m^a)L^{1/2}(a_{[\exp(m^a)]}))$ *a.s. and hence* $\max_{N_{m_n}<n\leq N_{m_n+1}}\left|\sum_{j=N_{m_n}+1}^{n}X_j'\right|=o(a_n)$ *a.s.*

PROOF. Using the same argument as in Lemma 5, it suffices to show that

$$\sum_{k\geq 1}k^{a-1}e^{-(\tau-2)k^a/2}L^{-\tau/2}(a_{[\exp(k^a)]})E|X|^{\tau}1_{\{|X|\leq 2a_{[\exp(k^a)]}\}}<\infty. \tag{27}$$

Let $n_j = [\exp(j^a)]$ and $\alpha_j = a_{n_j}$. Note that the sum in (27) is smaller than

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

$$C \sum_{j \geq 1} (L(2\alpha_j) - L(2\alpha_{j-1})) \cdot \alpha_j^{\tau-2} L^{-\tau/2}(\alpha_j) e^{-(\tau-2)j^a/2},$$

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

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

$$\leq C L^{-1}(2\alpha_j) \cdot v^{(\tau-2)/2}(\alpha_j).$$

Since $\alpha_j \sim \alpha_{j-1}$, we have $2\alpha_{j-1} \geq \alpha_j$ for j large. We conclude that the sum in (27) is smaller than

$$C \sum_{j \geq 1} \left[1 - \frac{L(\alpha_j)}{L(2\alpha_j)} \right] v^{(\tau-2)/2}(\alpha_j). \tag{28}$$

Using (SR) and the fact that $\alpha_j \geq C n_j^{1/2}$ and $v(x) \geq C \log\log x$, we get

$$\sum_{j \geq 1} \left(1 - \frac{L(\alpha_j)}{L(2\alpha_j)} \right) v^{(\tau-2)/2}(\alpha_j) \leq C \sum_{j \geq 1} (\log \alpha_j)^{-1/a} v^{-d}(\alpha_j) \leq$$

$$C \sum_{j \geq 1} (\log n_j)^{-1/a} (\log\log n_j)^{-d} \leq C \sum_{j \geq 1} j^{-1} (\log j)^{-d} < \infty,$$

where $d := \beta - (\tau-2)/2 > 1$ (and we recall that $a = 1/\alpha$). This concludes the proof of (27). □

Lemma 12 *We have* $\sum_{i=1}^m E(u_i'|\mathcal{G}_{i-1}') = o(m^{-(r-1/2)a} \cdot (\log m)^3 \cdot \exp(\frac{1}{2}m^a) \cdot L^{1/2}(a_{N_m}))$ *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 6. □

For our last result, we will need the following martingale subsequence convergence criterion (which is probably well-known).

Lemma 13 *Let* $\{S_n, \mathcal{F}_n\}_{n\geq 1}$ *be a zero-mean martingale and* $\{a_n\}_{n \geq 1}$ *a nondecreasing sequence of positive numbers with* $\lim_n a_n = \infty$. *If there exists a subsequence* $\{n_k\}_k$ *such that* $a_{n_k}/a_{n_{k-1}} \leq C$ *for all* k, *and*

$$\sum_{k \geq 1} \frac{E|S_{n_k} - S_{n_{k-1}}|^p}{a_{n_k}^p} < \infty \quad \textit{for some } p \in [1,2], \tag{29}$$

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

PROOF. Note that $\{S_{n_k}, \mathcal{F}_{n_k}\}_{k\geq 1}$ is a martingale. From (29) it follows that $S_{n_k}/a_{n_k} \to 0$ a.s. (see Theorem 2.18 of [7]). By the extended Kolmogorov inequality (see p. 65 of [10]), 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 $T_k := a_{n_k}^{-1} \max_{n_{k-1}<n\leq n_k} |S_n - S_{n_k}| \to 0$ a.s. Finally for $n_{k-1} < n \leq n_k$ we have:

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

$\square$

Finally, we treat the last term in the decomposition (26).

Theorem 14 *We have*

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

PROOF. Let $U_n := \sum_{i=1}^{m_n} \xi_i'$ and note that $\{U_n, \mathcal{G}'_{m_n}\}_{n\geq 1}$ is a zero-mean martingale. By Lemma 13, it is enough to prove that for a suitable subsequence $\{n_k\}_k$ we have

$$\sum_{k\geq 1} \frac{E|U_{n_k} - U_{n_{k-1}}|^2}{a_{n_k}^2} < \infty. \tag{30}$$

Similarly to the proof of Lemma 2.3 of [11], 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

$$\frac{1}{\phi(k)+1} \cdot I(b_{n_k}, a_{n_k}] \leq C I(a_{n_{k-1}}, a_{n_k}]. \tag{31}$$

Clearly $n_k \sim n_{k+1}$ and hence $a_{n_k} \sim a_{n_{k+1}}$ and $b_{n_k} \sim b_{n_{k+1}}$.
We proceed now with the proof of (30). Let

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

By the martingale property

$$EZ_k^2 = \sum_{m_{n_{k-1}} < i \leq m_{n_k}} E\xi_i^{'2} \leq (m_{n_k} - m_{n_{k-1}}) \max_{m_{n_{k-1}} < i \leq m_{n_k}} E\xi_i^{'2}. \tag{32}$$

Using Lemma 2.3 of [14] we have: for every $m_{n_{k-1}} < i \leq m_{n_k}$,

$$E\xi_i^{'2} \leq Eu_i^{'2} \leq C i^{a-1} e^{i^a} \cdot \max_{j\in H_i} EX_j^{'2} \leq C(\log n_k)^{(a-1)/a} n_k \cdot \max_{j\in H_i} EX_j^{'2}. \tag{33}$$

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

$$EX_j^{'2} \leq A(a_j) I(b_j, a_j] \leq A(a_{N_i}) I(b_{N_{i-1}}, a_{N_i}] \leq C A(a_{n_k}) I(b_{n_k}, a_{n_k}]. \tag{34}$$

Using (34) and (33) we get: for every $m_{n_{k-1}} < i \le m_{n_k}$,

$$E\xi_i^{'2} \le C(\log n_k)^{(a-1)/a} n_k \cdot A(a_{n_k}) I(b_{n_k}, a_{n_k}] = C(\log n_k)^{(a-1)/a} a_{n_k}^2 I(b_{n_k}, a_{n_k}]. \tag{35}$$

From (32) and (35) and recalling that $m_n \sim (\log n)^{1/a}$, we get

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

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 (31) of the function ϕ for the last inequality. Relationship (30) follows by (4). □

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. The authors are grateful to the Editor (Professor Martin Barlow) and the two anonymous referees for their helpful suggestions.

References

[1] R.M. Balan and R. Kulik. Self-normalized weak invariance principle for mixing sequences. *Tech. Rep. Series, Lab. Research Stat. Probab., Carleton University-University of Ottawa* **417** (2005).

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

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

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

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

[6] W. Feller. *An Introduction to Probability Theory and Its Applications*, Volume II, Second Edition (1971). John Wiley, New York.

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

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

[9] H. Kesten and R.A. Maller. Some effects of trimming on the law of the iterated logarithm. *J. Appl. Probab.* **41A** (2004), 253-271.

[10] M. Loève. *Probability Theory*, Volume II, Fourth Edition (1978). Springer, New York.

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

[12] S.Y. Novak. On self-normalized sums and Student's statistic. *Theory Probab. Appl.* **49** (2005), 336-344.

[13] W. Philipp and W.F. Stout. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. AMS Vol. 2* **161** (1975).

[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. *Chin. Ann. Math.* **14B** (1993), 27-42.