SPA 1047
pp: 1-19 (col.fig.: nil)
PROD. TYPE: COM
ED: Mariamma
PAGN: Saritha – SCAN: nil

ELSEVIER

Stochastic Processes and their Applications 000 (2001) 000–000

stochastic processes and their applications

www.elsevier.com/locate/spa

Invariance principles for adaptive self-normalized partial sums processes ☆

Alfredas Račkauskas[a], Charles Suquet[b,*]

[a] *Department of Mathematics, Vilnius University and Institute of Mathematics and Informatics, Naugarduko 24, Lt-2006 Vilnius, Lithuania*

[b] *Laboratorie de Statistique et Probabilités F.R.E. CNRS 2222, Université des Sciences et Technologies de Lille, Bât. M2, U.F.R. de Mathématiques, F-59655 Villeneuve d'Ascq Cedex, France*

Received 27 February 2000; received in revised form 15 February 2001; accepted 26 March 2001

Abstract

Let ζ_n^{se} be the *adaptive* polygonal process of *self-normalized* partial sums $S_k = \sum_{1 \leqslant i \leqslant k} X_i$ of i.i.d. random variables defined by linear interpolation between the points $(V_k^2/V_n^2, S_k/V_n)$, $k \leqslant n$, where $V_k^2 = \sum_{i \leqslant k} X_i^2$. We investigate the weak Hölder convergence of ζ_n^{se} to the Brownian motion W. We prove particularly that when X_1 is symmetric, ζ_n^{se} converges to W in each Hölder space supporting W *if and only if* X_1 belongs to the *domain of attraction of the normal distribution*. This contrasts strongly with Lamperti's FCLT where a moment of X_1 of order $p > 2$ is requested for some Hölder weak convergence of the classical partial sums process. We also present some partial extension to the nonsymmetric case. © 2001 Published by Elsevier Science B.V.

MSC: 60F05; 60B05; 60G17; 60E10

Keywords: Functional central limit theorem; Domain of attraction; Hölder space; Randomization

1. Introduction and results

Various partial sums processes can be built from the sums $S_n = X_1 + \cdots + X_n$ of independent identically distributed mean zero random variables. In this paper we focus attention on what we call the *adaptive self-normalized* partial sums process, denoted ζ_n^{se}. We investigate its weak convergence to the Brownian motion, trying to obtain it under the mildest integrability assumptions on X_1 and in the strongest topological framework. We basically show that in both respects, ζ_n^{se} behaves better than the classical Donsker–Prohorov partial sum processes ξ_n^{sr}. *Self-normalized* means here that the classical normalization by $\sqrt{n}$ is replaced by

$$V_n = (X_1^2 + \cdots + X_n^2)^{1/2}.$$

Adaptive means that the vertices of the corresponding random polygonal line have their abscissas at the random points V_k^2/V_n^2 $(0 \leqslant k \leqslant n)$ instead of the deterministic

☆ Preprint. Research supported by a cooperation agreement CNRS/LITHUANIA (4714).

* Corresponding author. Fax: +33-320436774.
E-mail address: charles.suquet@univ.lille1.fr (C. Suquet).

0304-4149/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S0304-4149(01)00096-5

UNCORRECTED PROOF

equispaced points k/n. By this construction the slope of each line adapts itself to the value of the corresponding random variable.

As a lot of different partial sums processes will appear throughout the paper, we need to explain our typographical conventions and fix notations.

By ζ_n (respectively ξ_n) we denote the random polygonal partial sums process defined on $[0,1]$ by linear interpolation between the vertices $(V_k^2/V_n^2,\ S_k)$, $k=0,1,\dots,n$ (respectively $(k/n, S_k)$, $k=0,1,\dots,n$), where

$$S_k = X_1 + \cdots + X_k, \quad V_k^2 = X_1^2 + \cdots + X_k^2.$$

For the special case $k=0$, we put $S_0=0$, $V_0=0$.

The upper scripts $^{\mathrm{sr}}$ or $^{\mathrm{se}}$ mean, respectively, normalization by square root of n or self-normalization. Hence,

$$\xi_n^{\mathrm{sr}} = \frac{\xi_n}{\sqrt{n}}, \quad \xi_n^{\mathrm{se}} = \frac{\xi_n}{V_n}, \quad \zeta_n^{\mathrm{sr}} = \frac{\zeta_n}{\sqrt{n}}, \quad \zeta_n^{\mathrm{se}} = \frac{\zeta_n}{V_n}.$$

By convention the random functions ξ_n^{se} and ζ_n^{se} are defined to be the null function on the event $\{V_n = 0\}$. Finally, the step partial sums processes Ξ_n, Z_n, Ξ_n^{se}, etc., are the piecewise constant random càdlàg functions whose jump points are vertices for the polygonal process denoted by the corresponding lowercase Greek letter.

Classical Donsker–Prohorov invariance principle states, that if $EX_1^2=1$, then

$$\xi_n^{\mathrm{sr}} \xrightarrow{\mathscr{D}} W, \tag{1}$$

in $C[0,1]$, where $(W(t), t\in[0,1])$ is a standard Wiener process and $\xrightarrow{\mathscr{D}}$ denotes convergence in distribution. Since (1) yields the central limit theorem, the finiteness of the second moment of X_1 therefore is necessary.

Lamperti (1962) considered the convergence (1) with respect to a stronger topology. He proved that if $E|X_1|^p<\infty$, where $p>2$, then (1) takes place in the Hölder space $H_\alpha[0,1]$, where $0<\alpha<1/2-1/p$. This result was derived again by Kerkyacharian and Roynette (1991) by another method using Ciesielski (1960) analysis of Hölder spaces by triangular functions. Further generalizations were given by Erickson (1981), Hamadouche (1998), Račkauskas and Suquet (1999c).

Considering a symmetric random variable X_1 such that $P(X_1 \geqslant u) = 1/(2u^p)$, $u\geqslant 1$, Lamperti (1962) noticed that the corresponding sequence (ξ_n^{sr}) is not tight in $H_\alpha[0,1]$ for $\alpha = 1/2-1/p$. It is then hopeless in general to look for an invariance principle in $H_\alpha[0,1]$ without some moment assumption beyond the square integrability of X_1. Recently, Račkauskas and Suquet (1999c) proved more precisely that if (ξ_n^{sr}) satisfies the invariance principle in $H_\alpha[0,1]$ for some $0<\alpha<1/2$, then necessarily

$$\sup_{t>0} t^p P(|X_1|>t) < \infty \tag{2}$$

for any $p<1/(1/2-\alpha)$.

Let us see now, how self-normalization and adaptiveness help to improve this situation. Recall that "X_1 belongs to the domain of attraction of the normal distribution"

UNCORRECTED PROOF

(denoted by $X_1 \in DAN$) means that there exists a sequence $b_n \uparrow \infty$ such that

$$b_n^{-1} S_n \xrightarrow{\mathscr{D}} N(0,1). \tag{3}$$

According to O'Brien's (1980) result: $X_1 \in DAN$ if and only if

$$V_n^{-1} \max_{1 \leqslant k \leqslant n} |X_k| \xrightarrow{\mathrm{P}} 0, \tag{4}$$

where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability. In the classical framework of $C[0,1]$, we obtain the following improvements of the Donsker–Prohorov theorem.

Theorem 1. *The convergence*

$$\xi_n^{\mathrm{se}} \xrightarrow{\mathscr{D}} W \tag{5}$$

holds in the space $C[0,1]$ *if and only if* $X_1 \in DAN$.

Theorem 2. *The convergence*

$$\zeta_n^{\mathrm{se}} \xrightarrow{\mathscr{D}} W \tag{6}$$

holds in the space $C[0,1]$ *if and only if* $X_1 \in DAN$.

Let us remark that the necessity of $X_1 \in DAN$ in both Theorems 1 and 2 follows from Giné, et al. (1997). Let us notice also that (5) or (6) both exclude the degenerated case $P(X_1 = 0) = 1$, so that almost surely $V_n > 0$ for large enough n. We have similar results (Račkauskas and Suquet, 2000) for the step processes Ξ_n^{se} and $\mathrm{Z}_n^{\mathrm{se}}$ within the Skorohod space $D(0,1)$.

For a modulus of continuity $\rho : [0,1] \to \mathbf{R}$, denote by $H_\rho[0,1]$ the set of continuous functions $x : [0,1] \to \mathbf{R}$ such that $\omega_\rho(x,1) < \infty$, where

$$\omega_\rho(x,\delta) := \sup_{\substack{t,s \in [0,1], \\ 0 < |t-s| < \delta}} \frac{|x(t) - x(s)|}{\rho(|s-t|)}.$$

The set $H_\rho[0,1]$ is a Banach space when endowed with the norm

$$\|x\|_\rho := |x(0)| + \omega_\rho(x,1).$$

Define

$$H_\rho^o[0,1] = \{x \in H_\rho[0,1] \colon \lim_{\delta \to 0} \omega_\rho(x,\delta) = 0\}.$$

Then $H_\rho^o[0,1]$ is a closed separable subspace of $H_\rho[0,1]$. In what follows we assume that the function ρ satisfies technical conditions (12) to (16) (see Section 2). These assumptions are fulfilled particularly when $\rho = \rho_{\alpha,\beta}$, $0 < \alpha < 1$, $\beta \in \mathbf{R}$, defined by

$$\rho_{\alpha,\beta}(h) := h^\alpha \ln^\beta (c/h), \quad 0 < h \leqslant 1$$

for a suitable constant c. We write $H_{\alpha,\beta}$ and $H_{\alpha,\beta}^o$ for $H_\rho[0,1]$ and $H_\rho^o[0,1]$, respectively, when $\rho = \rho_{\alpha,\beta}$ and we abbreviate $H_{\alpha,0}$ in H_α.

With respect to this Hölder scale $H_{\alpha,\beta}$, we obtain an optimal result when X_1 is symmetric.

UNCORRECTED PROOF

Theorem 3. *Assume that ρ satisfies conditions* (12)–(16) *and*

$$\lim_{j\to\infty} \frac{2^j \rho^2(2^{-j})}{j} = \infty. \tag{7}$$

If X_1 is symmetric and $X_1 \in DAN$ then

$$\zeta_n^{\mathrm{se}} \xrightarrow{\mathscr{D}} W, \tag{8}$$

in $H_\rho^o[0,1]$.

Corollary 4. *If X_1 is symmetric and $X_1 \in DAN$ then* (8) *holds in the space $H^o_{1/2,\beta}$, for any $\beta > 1/2$.*

It is well known that the Wiener process has a version in the space $H_{1/2,1/2}$ but none in $H^o_{1/2,1/2}$. Hence Corollary 4 gives the best result possible in the scale of the separable Hölder spaces $H_{\alpha,\beta}$. In Račkauskas and Suquet (1999c) it is proved that if the classical partial sums process ξ_n^{sr} converges in $H^o_{1/2,\beta}$ for some $\beta > 1/2$, then $\|X_1\|_{\psi_\gamma} < \infty$, where $\|X_1\|_{\psi_\gamma}$ is the Orlicz norm related to the Young function $\psi_\gamma(r) = \exp(r^\gamma) - 1$ with $\gamma = 1/\beta$. This shows the striking improvement of weak Hölder convergence due to self-normalization and adaptation.

It seems worth noticing here, that without adaptive construction of the polygonal process, the existence of moments of order bigger than 2 is necessary for Hölder weak convergence. Indeed, if $\xi_n^{\mathrm{se}} \xrightarrow{\mathscr{D}} W$ in H_α, then one can prove that $\mathbf{E}X_1^2 < \infty$. Therefore $\xi_n^{\mathrm{sr}} \xrightarrow{\mathscr{D}} W$ in H_α and the moment restriction (2) is necessary.

Naturally it is very desirable to remove the symmetry assumption in Corollary 4. Although the problem remains open, we can propose the following partial results in this direction.

Theorem 5. *Let $\beta > 1/2$ and suppose that we have*

$$P\left(\max_{1\leqslant k\leqslant n} \frac{X_k^2}{V_n^2} \geqslant \delta_n\right) \xrightarrow[n\to\infty]{} 0 \tag{9}$$

and

$$P\left(\max_{1\leqslant k\leqslant n} \left|\frac{V_k^2}{V_n^2} - \frac{k}{n}\right| \geqslant \delta_n\right) \xrightarrow[n\to\infty]{} 0, \tag{10}$$

with

$$\delta_n = c\frac{2^{-(\log n)^\gamma}}{\log n} \quad \textit{for some } \frac{1}{2\beta} < \gamma < 1 \quad \textit{and some } c > 0. \tag{11}$$

Then

$$\zeta_n^{\mathrm{se}} \xrightarrow{\mathscr{D}} W \quad \textit{in } H^o_{1/2,\beta}.$$

Observe that $n^{-\varepsilon} = o(\delta_n)$ for any $\varepsilon > 0$. This mild convergence rate δ_n may be obtained as soon as $\mathbf{E}|X_1|^{2+m\epsilon}$ is finite.

Corollary 6. *If for some $\varepsilon > 0$, $\mathbf{E}|X_1|^{2+\epsilon} < \infty$, then for any $\beta > 1/2$, ζ_n^{se} converges weakly to W in the space $H^o_{1/2,\beta}$.*

UNCORRECTED PROOF

This result contrasts strongly with the extension of Lamperti's invariance principle in the same functional framework (Račkauskas and Suquet, 1999c)

The present contribution is a new illustration of the now well established fact, that in general, self-normalization improves the asymptotic properties of sums of independent random variables.

A rich literature is devoted to limit theorems for self-normalized sums. Logan et al. (1973) investigate the various possible limit distributions of self-normalized sums. Giné et al. (1997) prove that S_n/V_n converges to the Gaussian standard distribution if and only if X_1 is in the domain of attraction of the normal distribution (the symmetric case was previously treated in Griffin and Mason (1991)). Egorov (1997) investigates the non identically distributed case. Bentkus and Götze (1996) obtain the rate of convergence of S_n/V_n when $X_1 \in DAN$. Griffin and Kuelbs (1989) prove the LIL for self-normalized sums when $X_1 \in DAN$. Moderate deviations (of Linnik's type) are studied in Shao (1999) and Christiakov and Götze (1999). Large deviations (of Cramér–Chernoff type) are investigated in Shao (1997) without moment conditions. Chuprunov (1997) gives invariance principles for various partial sums processes under self-normalization in $C[0,1]$ or $D[0,1]$. Our Theorems 1 and 2 improve on Chuprunov's results in the i.i.d. case.

2. Preliminaries

2.1. Analytical background

In this section we collect some facts about the Hölder spaces $H_\rho[0,1]$ including the tightness criterion for distributions in these spaces. All these facts may be found e.g. in Račkauskas and Suquet (1999b).

In what follows, we assume that the modulus of smoothness ρ satisfies the following technical conditions where c_1, c_2 and c_3 are positive constants:

$$\rho(0) = 0, \ \rho(\delta) > 0, \quad 0 < \delta \leqslant 1, \tag{12}$$

$$\rho \text{ is nondecreasing on } [0,1], \tag{13}$$

$$\rho(2\delta) \leqslant c_1 \rho(\delta), \quad 0 \leqslant \delta \leqslant 1/2, \tag{14}$$

$$\int_0^\delta \frac{\rho(u)}{u}\,\mathrm{d}u \leqslant c_2 \rho(\delta), \quad 0 < \delta \leqslant 1, \tag{15}$$

$$\delta \int_\delta^1 \frac{\rho(u)}{u^2}\,\mathrm{d}u \leqslant c_3 \rho(\delta), \quad 0 < \delta \leqslant 1. \tag{16}$$

For instance, elementary computations show that the functions

$$\rho(\delta) := \delta^\alpha \ln^\beta\left(\frac{c}{\delta}\right), \quad 0 < \alpha < 1, \ \beta \in \mathbf{R},$$

satisfy conditions (12)–(16), for a suitable choice of the constant c, namely $c \geqslant \exp(\beta/\alpha)$ if $\beta > 0$ and $c > \exp(-\beta/(1-\alpha))$ if $\beta < 0$.

UNCORRECTED PROOF

Write D_j for the set of dyadic numbers of level j in $[0,1]$, *i.e.* $D_0 = \{0,1\}$ and for $j \geqslant 1$,

$$D_j = \{(2k+1)2^{-j};\ 0 \leqslant k < 2^{j-1}\}.$$

For any continuous function $x:[0,1] \to \mathbf{R}$, define

$$\lambda_{0,t}(x) := x(t), \quad t \in D_0$$

and for $j \geqslant 1$,

$$\lambda_{j,t}(x) := x(t) - \tfrac{1}{2}(x(t+2^{-j}) + x(t-2^{-j})), \quad t \in D_j.$$

The $\lambda_{j,t}(x)$ are the coefficients of the expansion of x in a series of triangular functions. The jth partial sum $E_j x$ of this series is exactly the polygonal line interpolating x between the dyadic points $k2^{-j}(0 \leqslant k \leqslant 2^j)$. Under (12)–(16), the norm $\|x\|_\rho$ is equivalent to the sequence norm

$$||x||_\rho^{\mathrm{seq}} := \sup_{j \geqslant 0} \frac{1}{\rho(2^{-j})} \max_{t \in D_j} |\lambda_{j,t}(x)|.$$

In particular, both norms are finite if and only if x belongs to H_ρ. It is easy to check that

$$\|x - E_j x\|_\rho^{\mathrm{seq}} = \sup_{i > j} \frac{1}{\rho(2^{-i})} \max_{t \in D_i} |\lambda_{i,t}(x)|.$$

Proposition 7. *The sequence* (Y_n) *of random elements in* H_ρ^o *is tight if and only if the following two conditions are satisfied*:

(i) *For each* $t \in [0,1]$, *the sequence* $(Y_n(t))_{n \geqslant 1}$ *is tight on* $\mathbf{R}$.

(ii) *For each* $\varepsilon > 0$,

$$\lim_{j \to \infty} \sup_{n \geqslant 1} P(\|Y_n - E_j Y_n\|_\rho^{\mathrm{seq}} > \varepsilon) = 0.$$

Remark 8. Condition (ii) in Proposition 7 may be replaced by

$$\lim_{j \to \infty} \limsup_{n \to \infty} P(\|Y_n - E_j Y_n\|_\rho^{\mathrm{seq}} > \varepsilon) = 0. \tag{17}$$

2.2. Adaptive time and DAN

We establish here the technical results on the adaptive time when $X_1 \in DAN$ which will be used throughout the paper. These results rely on the common assumption that X_1 is in the domain of normal attraction. This provides the following properties on the distribution of X_1. Since $X_1 \in DAN$, there exists a sequence $b_n \uparrow \infty$ such that $b_n^{-1} S_n$ converges weakly to $N(0,1)$. Then Raikov's theorem yields

$$b_n^{-2} V_n^2 \xrightarrow{\mathrm{P}} 1. \tag{18}$$

UNCORRECTED PROOF

We have moreover for each $\tau > 0$, putting $b_n = n^{-1/2}\ell_n$,

$$nP(|X_1| > \tau\ell_n\sqrt{n}) \to 0, \tag{19}$$

$$\ell_n^{-2}\mathbf{E}(X_1^2; |X_1| \leqslant \tau\ell_n\sqrt{n}) \to 1, \tag{20}$$

$$n\mathbf{E}(X_1; |X_1| \leqslant \tau\ell_n\sqrt{n}) \to 0, \tag{21}$$

see for instance Araujo and Giné (1980, Chapter 2, Corollaries 4.8(a) and 6.18(b) and Theorem 6.17(i)). Here and in all the paper $(X; E)$ means the product of the random variable X by the indicator function of the event E.

Lemma 9. *If $X_1 \in DAN$, then*

$$\sup_{0\leqslant t\leqslant 1}\left|\frac{V_{[nt]}^2}{V_n^2} - t\right| \xrightarrow{\mathrm{P}} 0. \tag{22}$$

Proof. Consider the truncated random variables

$$X_{n,i} := b_n^{-1}(X_i; X_i^2 \leqslant b_n^2), \quad i = 1, \dots, n.$$

Define $V_{n,0} := 0$ and $V_{n,k}^2 = X_{n,1}^2 + \cdots + X_{n,k}^2$ for $k = 1, \dots, n$. Set

$$v_n = \sup_{0\leqslant t\leqslant 1}\left|\frac{V_{[nt]}^2}{V_n^2} - t\right| \quad \text{and} \quad \tilde{v}_n = \sup_{0\leqslant t\leqslant 1}\left|\frac{V_{n,[nt]}^2}{V_{n,n}^2} - t\right|.$$

Then we have for $\lambda > 0$,

$$P(v_n > \lambda) \leqslant P(\tilde{v}_n > \lambda) + nP(X_1^2 > b_n^2).$$

Due to (19) the proof of (22) reduces to the proof of

$$\tilde{v}_n \xrightarrow{\mathrm{P}} 0. \tag{23}$$

Since $V_{n,k}^2 \leqslant V_{n,n}^2$ for $k = 0, \dots, n$, the elementary estimate

$$\left|\frac{V_{n,k}^2}{V_{n,n}^2} - \frac{k}{n}\right| \leqslant \frac{V_{n,k}^2}{V_{n,n}^2}|1 - V_{n,n}^2| + \left|V_{n,k}^2 - \frac{k}{n}\right|$$

leads to

$$\tilde{v}_n \leqslant \max_{0\leqslant k\leqslant n}\left|V_{n,k}^2 - \frac{k}{n}\right| + |1 - V_{n,n}^2| + \frac{1}{n}. \tag{24}$$

Noting that $V_{n,n}^2 = b_n^{-2}V_n^2 R_n$ with

$$R_n := \frac{1}{V_n^2}\sum_{i=1}^{n}(X_i^2; X_i^2 \leqslant b_n^2),$$

we clearly have $R_n \leqslant 1$ a.s. and

$$P(R_n < 1) = P\left(\max_{1\leqslant i\leqslant n}|X_i| > b_n\right) \leqslant nP(|X_1| > b_n),$$

UNCORRECTED PROOF

which goes to zero by (19). This together with (18) gives

$$V_{n,n}^2 \xrightarrow{\mathrm{P}} 1. \tag{25}$$

Hence the proof of (23) reduces to

$$\max_{0\leqslant k\leqslant n} \left| V_{n,k}^2 - \frac{k}{n} \right| \xrightarrow{\mathrm{P}} 0. \tag{26}$$

For this convergence we have

$$\max_{0\leqslant k\leqslant n} |V_{n,k}^2 - k/n| \leqslant \max_{0\leqslant k\leqslant n} |V_{n,k}^2 - \mathbf{E}V_{n,k}^2| + \max_{0\leqslant k\leqslant n} |\mathbf{E}V_{n,k}^2 - k/n|.$$

Noting that

$$\mathbf{E}V_{n,k}^2 - \frac{k}{n} = \frac{k}{n}(nb_n^{-2}\mathbf{E}(X_1^2; X_1^2 \leqslant b_n^2) - 1)$$

gives

$$\max_{0\leqslant k\leqslant n} \left| \mathbf{E}V_{n,k}^2 - \frac{k}{n} \right| \leqslant |nb_n^{-2}\mathbf{E}(X_1^2; X_1^2 \leqslant b_n^2) - 1|,$$

which goes to zero by (20). Hence it remains to prove

$$\max_{0\leqslant k\leqslant n} |V_{n,k}^2 - \mathbf{E}V_{n,k}^2| \xrightarrow{\mathrm{P}} 0. \tag{27}$$

Putting $T_{n,k} := V_{n,k}^2 - \mathbf{E}V_{n,k}^2$, we have by Ottaviani inequality

$$P\left(\max_{1\leqslant k\leqslant n} |T_{n,k}| > 2\lambda \right) \leqslant \frac{P(|T_{n,n}| > \lambda)}{1 - \max_{1\leqslant k\leqslant n} P(|T_{n,n} - T_{n,k}| > \lambda)}. \tag{28}$$

Due to (25), we are left with the control of $I := \max_{1\leqslant k\leqslant n} P(|T_{n,k}| > \lambda)$. By Chebyshev's inequality

$$I \leqslant \lambda^{-2} \max_{1\leqslant k\leqslant n} \mathbf{E}T_{n,k}^2 \leqslant \lambda^{-2} nEX_{n,1}^4$$

and we have to consider $I_1 = n\mathbf{E}X_{n,1}^4 = nb_n^{-4}\mathbf{E}(X_1^4; |X_1| \leqslant b_n)$. For any $0 < \tau < 1$,

$$\begin{aligned}\mathbf{E}(X_1^4; |X_1| \leqslant b_n) &\leqslant \mathbf{E}(X_1^4; |X_1| \leqslant \tau b_n) + \mathbf{E}(X_1^4; \tau b_n \leqslant |X_1| \leqslant b_n)\\ &\leqslant \tau^2 b_n^2 \mathbf{E}(X_1^2; |X_1| \leqslant \tau b_n) + b_n^4 P(|X_1| \geqslant \tau b_n).\end{aligned}$$

So

$$I_1 \leqslant \tau^2 nb_n^{-2}\mathbf{E}(X_1^2; |X_1| \leqslant \tau b_n) + nP(|X_1| \geqslant \tau b_n).$$

Choosing $\tau = \lambda/2$ in (19) and (20), we can achieve $I \leqslant 1/2$ for n large enough and the proof is complete. □

Remark 10. If $X_1 \in DAN$, we also have

$$\sup_{0\leqslant t\leqslant 1} \left| \frac{V_{[nt]+1}^2}{V_n^2} - t \right| \xrightarrow{\mathrm{P}} 0. \tag{29}$$

UNCORRECTED PROOF

Indeed, recalling (4), it suffices to write

$$\frac{V_{[nt]+1}^2 - V_{[nt]}^2}{V_n^2} = \frac{X_{[nt]+1}^2}{V_n^2} \leqslant \left(\frac{1}{V_{n+1}^2} \max_{1\leqslant k\leqslant n+1} X_k^2 \right) \frac{V_{n+1}^2}{V_n^2},$$

and observe that V_{n+1}^2/V_n^2 converges to 1 in probability since by Lemma 9,

$$\left| \frac{V_n^2}{V_{n+1}^2} - \frac{n}{n+1} \right| \leqslant \sup_{0\leqslant t\leqslant 1} \left| \frac{V_{[(n+1)t]}^2}{V_{n+1}^2} - t \right| \xrightarrow{\mathrm{P}} 0.$$

Remark 11. For each $t \in [0,1]$,

$$\frac{b_{[nt]}^2}{b_n^2} \to t. \tag{30}$$

This is a simple by-product of Lemma 9, writing

$$\frac{b_{[nt]}^2}{b_n^2} = \frac{V_n^2}{b_n^2} \times \frac{b_{[nt]}^2}{V_{[nt]}^2} \times \frac{V_{[nt]}^2}{V_n^2}$$

and noting that for fixed $t > 0$ and $n \geqslant n_0$ large enough $[nt] < [(n+1)t]$ so the sequence $(b_{[nt]}^2/V_{[nt]}^2)_{n\geqslant n_0}$ is a subsequence of $(b_n^2/V_n^2)_{n\geqslant n_0}$ which converges in probability to 1 by (18).

Define the random variables

$$\tau_n(t) = \max\{k = 0,\dots,n;\ V_k^2 \leqslant tV_n^2\}, \quad t \in [0,1], \tag{31}$$

so that we have $\tau_n(1) = n$ and for $0 \leqslant t < 1$,

$$\frac{V_{\tau_n(t)}^2}{V_n^2} \leqslant t < \frac{V_{\tau_n(t)+1}^2}{V_n^2}. \tag{32}$$

Lemma 12. *If $X_1 \in DAN$ then*

$$\sup_{t\in[0,1]} |n^{-1}\tau_n(t) - t| \xrightarrow{\mathrm{P}} 0. \tag{33}$$

Proof. The result will follow from Remark 10, if we check the inclusion of events

$$\left\{ \sup_{t\in[0,1]} |n^{-1}\tau_n(t) - t| > \varepsilon \right\} \subset \left\{ \sup_{u\in[0,1]} \left| \frac{V_{[nu]+1}^2}{V_n^2} - u \right| \geqslant \varepsilon \right\}. \tag{34}$$

The occurrence of the left-hand side in (34) is equivalent to the existence of one $s \in [0,1]$ such that $|n^{-1}\tau_n(s) - s| > \varepsilon$, *i.e.* such that

$$\tau_n(s) > n(s+\varepsilon) \tag{35}$$

or

$$\tau_n(s) < n(s-\varepsilon). \tag{36}$$

Observe that under (35), $s + \varepsilon < 1$, while under (36), $s - \varepsilon > 0$. From the definition of τ_n, (35) gives an integer $k > n(s+\varepsilon)$ such that $V_k^2/V_n^2 \leqslant s$, whence

$$\frac{V_{[n(s+\varepsilon)]+1}^2}{V_n^2} \leqslant s. \tag{37}$$

UNCORRECTED PROOF

On the other hand, under (36), we have $V_k^2/V_n^2 > s$ for every $k \geqslant n(s-\varepsilon)$ and in particular

$$\frac{V^2_{[n(s-\varepsilon)]+1}}{V_n^2} > s. \tag{38}$$

Recasting (37) and (38) under the form

$$\frac{V^2_{[n(s+\varepsilon)]+1}}{V_n^2} - (s+\varepsilon) \leqslant -\varepsilon$$
$$\frac{V^2_{[n(s-\varepsilon)]+1}}{V_n^2} - (s-\varepsilon) > \varepsilon,$$

shows that both (35) and (36) imply the occurrence of the event in the right-hand side of (34). □

3. Proofs

Proof of Theorem 1. First we prove the convergence of finite dimensional distributions (f.d.d.) of the process ξ_n^{se} to the corresponding f.d.d. of the Wiener process W.

To this aim, consider the process $\Xi_n = (S_{[nt]}, t \in [0,1])$. By (4) applied to the obvious bound

$$\sup_{0\leqslant t\leqslant 1} V_n^{-1}|\xi_n(t) - \Xi_n(t)| \leqslant V_n^{-1} \max_{1\leqslant k\leqslant n} |X_k|,$$

the convergence of f.d.d. of ξ_n^{se} follows from those of the process Ξ_n^{se}.

Let $0 \leqslant t_1 < t_2 < \cdots < t_d \leqslant 1$. From (3), independence of the X_i's and Remark 11, we get

$$b_n^{-1}(S_{[nt_1]}, S_{[nt_2]} - S_{[nt_1]}, \ldots, S_{[nt_d]} - S_{[nt_{d-1}]})$$
$$\xrightarrow{\mathscr{D}} (W(t_1), W(t_2) - W(t_1), \ldots, W(t_d) - W(t_{d-1})).$$

Now (18) and the continuity of the map

$$(x_1, x_2, \ldots, x_d) \mapsto (x_1, x_2 + x_1, \ldots, x_d + \cdots + x_1)$$

yields the convergence of f.d.d. of Ξ_n^{se}. The convergence of finite dimensional distributions of the process ξ_n^{se} is thus established.

To prove the tightness we shall use Theorem 8.3 from Billingsley (1968). Since $\xi_n^{\mathrm{se}}(0) = 0$, the proof reduces in showing that for all ε, $\eta > 0$ there exist $n_0 \geqslant 1$ and δ, $0 < \delta < 1$, such that

$$\frac{1}{\delta} P\left\{ \sup_{1\leqslant i\leqslant n\delta} V_n^{-1}|S_{k+i} - S_k| \geqslant \varepsilon \right\} \leqslant \eta, \quad n \geqslant n_0 \tag{39}$$

for all $1 \leqslant k \leqslant n$.

Let us introduce the truncated variables

$$Y_i := \ell_n^{-1}(X_i;\ X_i^2 \leqslant \tau^2 b_n^2), \quad i = 1, \ldots, n$$

UNCORRECTED PROOF

with $\ell_n = n^{-1/2} b_n$ as above and τ to be chosen later. Denote by $\tilde{S}_k$ and $\tilde{V}_k$ the corresponding partial sums with their self-normalizing random variables:

$$\tilde{S}_k = Y_1 + \cdots + Y_k, \quad \tilde{V}_k = (Y_1^2 + \cdots + Y_n^2)^{1/2}, \quad k = 1, \ldots, n.$$

Then we have

$$P\left\{ \sup_{1 \leqslant i \leqslant n\delta} V_n^{-1} |S_{k+i} - S_k| \geqslant \varepsilon \right\} \leqslant A + B + C, \tag{40}$$

where

$$A := P\left\{ \sup_{1 \leqslant i \leqslant n\delta} |\tilde{S}_{k+i} - \tilde{S}_k| \geqslant \varepsilon\sqrt{n/2} \right\},$$

$$B := P\{\tilde{V}_n < \sqrt{n}/2\},$$

$$C := nP\{|X_1| \geqslant \tau \ell_n \sqrt{n}\}.$$

Due to (21) we can choose n_1 such that $\sqrt{n}|\mathbf{E}Y_1| \leqslant 1/4$ for $n \geqslant n_1$. Then with $n \geqslant n_1$ and $\delta \leqslant \varepsilon$ we have

$$A \leqslant P\left\{ \max_{1 \leqslant i \leqslant n\delta} \left| \sum_{j=k+1}^{k+i} (Y_j - \mathbf{E}Y_j) \right| + n\delta |\mathbf{E}Y_1| \geqslant \sqrt{n}\varepsilon/2 \right\}$$

$$\leqslant P\left\{ \max_{1 \leqslant i \leqslant n\delta} \left| \sum_{j=k+1}^{k+i} (Y_j - \mathbf{E}Y_j) \right| \geqslant \sqrt{n}\varepsilon/4 \right\}.$$

By Chebyshev's inequality and Rosenthal inequality with $p > 2$, we have for each $1 \leqslant k \leqslant n$

$$P\left\{ n^{-1/2} \left| \sum_{j=k+1}^{k+n\delta} (Y_j - \mathbf{E}Y_j) \right| \geqslant \frac{\varepsilon}{8} \right\} \leqslant \frac{8^p}{\varepsilon^p n^{p/2}} \mathbf{E} \left| \sum_{j=k+1}^{k+n\delta} (Y_j - \mathbf{E}Y_j) \right|^p$$

$$\leqslant \frac{8^p}{\varepsilon^p n^{p/2}} [(n\delta)^{p/2} (\mathbf{E}Y_1^2)^{p/2} + n\delta \mathbf{E}|Y_1|^p].$$

By (20) we can choose n_2 such that

$$3/4 \leqslant \mathbf{E}Y_1^2 \leqslant 3/2 \quad \text{for} \quad n \geqslant n_2. \tag{41}$$

Then we have $\mathbf{E}|Y_1|^p \leqslant 2n^{(p-2)/2} \tau^{p-2}$ and then assuming that $\tau \leqslant \delta^{1/2}$ we obtain

$$P\left\{ n^{-1/2} \left| \sum_{j=k+1}^{k+n\delta} (Y_j - \mathbf{E}Y_j) \right| \geqslant \frac{\varepsilon}{8} \right\} \leqslant \frac{8^p}{\varepsilon^p n^{p/2}} [2^{p/2} (n\delta)^{p/2} + \delta n^{p/2} \tau^{p-2}]$$

$$\leqslant \frac{2 \cdot 16^p \delta^{p/2}}{\varepsilon^p}.$$

Now by Ottaviani inequality we find

$$A \leqslant \frac{\delta\eta}{3}, \tag{42}$$

provided $\delta^{p/2} \leqslant \varepsilon^p / (4 \cdot 16^p)$ and $\delta^{(p-2)/2} \leqslant \eta\varepsilon^p / (6 \cdot 16^p)$.

UNCORRECTED PROOF

Next we consider B. Since $n^{-1}\mathbf{E}\tilde{V}_n^2 = \mathbf{E}Y_1^2$ we have by (41) $n^{-1}\mathbf{E}\tilde{V}_n^2 \geqslant 3/4$, for $n \geqslant n_2$. Furthermore,

$$B \leqslant P\{n^{-1}|\tilde{V}_n^2 - \mathbf{E}\tilde{V}_n^2| \geqslant 1/2\} \leqslant 4n^{-1}\mathbf{E}Y_1^4 \leqslant 4\tau^2\mathbf{E}Y_1^2 \leqslant \delta\eta/3, \tag{43}$$

provided $n \geqslant n_2$ and $\tau^2 \leqslant \delta\eta/18$.

Finally choose n_3 such that $C \leqslant \eta\delta/3$ when $n \geqslant n_3$ and join to that estimates (42) and (43) to conclude (39). The proof is complete. □

Proof of Theorem 2. Due to Theorem 1, it suffices to check that $\|V_n^{-1}(\xi_n - \zeta_n)\|_\infty$ goes to zero in probability, where $\|f\|_\infty := \sup_{0\leqslant t\leqslant 1}|f(t)|$. To this end let us introduce the random change of time θ_n defined as follows. When $V_n > 0$, θ_n is the map from $[0,1]$ *onto* $[0,1]$ which interpolates linearly between the points $(k/n, V_k^2/V_n^2)$, $k = 0, 1, \ldots, n$. When $V_n = 0$, we simply take $\theta_n = I$, the identity on $[0,1]$. With the usual convention $S_k/V_n := 0$ for $V_n = 0$, we always have

$$\zeta_n^{\mathrm{se}}(\theta_n(t)) = \xi_n^{\mathrm{se}}(t), \quad 0 \leqslant t \leqslant 1. \tag{44}$$

Clearly for each $t \in [0,1]$,

$$\left| \frac{V_{[nt]}^2}{V_n^2} - \theta_n(t) \right| \leqslant \max_{1\leqslant k\leqslant n} \frac{X_k^2}{V_n^2}.$$

It follows by (4) that

$$\sup_{0\leqslant t\leqslant 1} \left| \frac{V_{[nt]}^2}{V_n^2} - \theta_n(t) \right| \xrightarrow{\mathrm{P}} 0$$

and this together with Lemma 9 gives

$$\|\theta_n - I\|_\infty \xrightarrow{\mathrm{P}} 0. \tag{45}$$

Let $\omega(f;\delta) := \sup\{|f(t) - f(s)|; |t - s \leqslant \delta\}$ denote the modulus of continuity of $f \in C[0,1]$. Then recalling (44) we have

$$\|\xi_n^{\mathrm{se}} - \zeta_n^{\mathrm{se}}\|_\infty = \sup_{0\leqslant t\leqslant 1} |\xi_n^{\mathrm{se}}(\theta_n(t)) - \zeta_n^{\mathrm{se}}(\theta_n(t))| \leqslant \omega(\xi_n^{\mathrm{se}}; \|\theta_n - I\|_\infty).$$

It follows that for any $\lambda > 0$ and $0 < \delta \leqslant 1$,

$$P(\|\xi_n^{\mathrm{se}} - \zeta_n^{\mathrm{se}}\|_\infty \geqslant \lambda) \leqslant P(\|\theta_n - I\|_\infty > \delta) + P(\omega(\xi_n^{\mathrm{se}};\delta) \geqslant \lambda). \tag{46}$$

Now since the Brownian motion has a version in $C[0,1]$, we can find for each positive ε, some $\delta \in (0,1]$ such that $P(\omega(W;\delta) \geqslant \lambda) < \varepsilon$. As the functional ω is continuous on $C[0,1]$, it follows from Theorem 1 that

$$\limsup_{n\to\infty} P(\omega(\xi_n^{\mathrm{se}};\delta) \geqslant \lambda) \leqslant P(\omega(W;\delta) \geqslant \lambda).$$

Hence for $n \geqslant n_1$ we have $P(\omega(\xi_n^{\mathrm{se}};\delta) \geqslant \lambda) < 2\varepsilon$. Having in mind (45) and (46) we see that the proof is complete. □

Proof of Theorem 3. The convergence of finite dimensional distributions is already established in the proof of Theorem 2.

UNCORRECTED PROOF

It remains to prove tightness of $\zeta_n^{\rm se}$ in the space $H_\rho[0,1]$. To this aim, we have to check the second condition of Proposition 7 only.

Let $\varepsilon_1,\dots,\varepsilon_n,\dots$ be an independent Rademacher sequence which is independent on (X_i). By symmetry of X_1, both sequences (X_i) and $(\varepsilon_i X_i)$ have the same distribution. Noting also that $\varepsilon_i^2=1$ a.s., we have that $\zeta_n^{\rm se}$ has the same distribution as the random process $\tilde{\zeta}_n^{\rm se}$ which is defined linearly between the points

$$\left(\frac{V_k^2}{V_n^2},\frac{U_k}{V_n}\right),$$

where $U_0=0$ and $U_k=\sum_{i=1}^k \varepsilon_i X_i$, for $k\geqslant 1$. Hence, it suffices to prove that

$$\lim_{J\to\infty}\sup_n\sum_{j>J}2^j\max_{0\leqslant k<2^j}P(|\tilde{\zeta}_n^{\rm se}|((k+1)2^{-j})-\tilde{\zeta}_n^{\rm se}(k2^{-j})|>\varepsilon\rho(2^{-j}))=0. \tag{47}$$

To this aim we shall estimate

$$\delta(t,h,r):=P(|\tilde{\zeta}_n^{\rm se}(t+h)-\tilde{\zeta}_n^{\rm se}(t)|>r),$$

uniformly in n. First consider the case, where

$$0\leqslant\frac{V_{k-1}^2}{V_n^2}\leqslant t<t+h\leqslant\frac{V_k^2}{V_n^2},$$

so

$$0\leqslant h\leqslant\frac{V_k^2}{V_n^2}-\frac{V_{k-1}^2}{V_n^2}=\frac{X_k^2}{V_n^2}.$$

We have then by linear interpolation

$$\begin{aligned}|\tilde{\zeta}_n^{\rm se}(t+h)-\tilde{\zeta}_n^{\rm se}(t)|&=\frac{|\varepsilon_k X_k|}{V_n}\frac{V_n^2}{X_k^2}h\\&=\left(\frac{V_n}{|X_k|}\sqrt{h}\right)\sqrt{h}\leqslant\sqrt{h}.\end{aligned} \tag{48}$$

Next consider the following configuration:

$$0\leqslant\frac{V_{k-1}^2}{V_n^2}\leqslant t<\frac{V_k^2}{V_n^2}\leqslant\frac{V_l^2}{V_n^2}\leqslant t+h<\frac{V_{l+1}^2}{V_n^2}.$$

Then we have

$$|\tilde{\zeta}_n^{\rm se}(t+h)-\tilde{\zeta}_n^{\rm se}(t)|\leqslant\delta_1+\delta_2+\delta_3,$$

where

$$\delta_1:=|\tilde{\zeta}_n^{\rm se}(t+h)-\tilde{\zeta}_n^{\rm se}(V_l^2/V_n^2)|\leqslant\sqrt{t+h-V_l^2/V_n^2}\leqslant\sqrt{h},$$

$$\delta_2:=|\tilde{\zeta}_n^{\rm se}(V_l^2/V_n^2)-\tilde{\zeta}_n^{\rm se}(V_k^2/V_n^2)|=V_n^{-1}|U_l-U_k|\leqslant\frac{|U_l-U_k|}{\sqrt{V_l^2-V_k^2}}\sqrt{h},$$

$$\delta_3:=|\tilde{\zeta}_n^{\rm se}(V_k^2/V_n^2)-\zeta_n^{\rm se}(t)|\leqslant\sqrt{V_k^2/V_n^2-t}\leqslant\sqrt{h}.$$

Hence, for any configuration we obtain

$$|\tilde{\zeta}_n^{\rm se}(t+h)-\tilde{\zeta}_n^{\rm se}(t)|\leqslant\frac{|U_l-U_k|}{\sqrt{V_l^2-V_k^2}}\sqrt{h}+2\sqrt{h}, \tag{49}$$

UNCORRECTED PROOF

if we agree that $|U_l - U_k|(V_l^2 - V_k^2)^{-1/2} := 0$ when $k = l$. Therefore,

$$\delta(t,h,r) \leqslant P(|U_l - U_k|/\sqrt{V_l^2 - V_k^2} > r/(2\sqrt{h})), \tag{50}$$

provided $r > 4\sqrt{h}$. Observe that in this formula the indexes l and k are random variables depending on t, h and the sequence (X_i), but independent of the sequence (ε_i). Thus conditioning on $X_1,\dots,X_n$ and applying the well known Hoeffding's inequality we obtain

$$\delta(t,h,r) \leqslant c\exp\{-r^2/(8h)\}. \tag{51}$$

Now (47) clearly follows if for every $\varepsilon > 0$,

$$\sum_{j=1}^{\infty} 2^j \exp\{-\varepsilon 2^j \rho^2(2^{-j})\} < \infty, \tag{52}$$

which is easily seen to be equivalent to our hypothesis (7). The proof is completed. □

Proof of Theorem 5. From (9) and the characterization (4) of *DAN*, X_1 is clearly in the domain of normal attraction. So the convergence of finite dimensional distributions is already given by Theorem 2.

To establish the tightness we have to prove that

$$\lim_{J\to\infty} \limsup_{n\to\infty} P(\|\zeta_n^{\mathrm{se}} - E_J\zeta_n^{\mathrm{se}}\|_\rho^{\mathrm{seq}} > 4\varepsilon) = 0. \tag{53}$$

To this end, it suffices to prove that with some sequence $J_n \uparrow \infty$ to be precised later,

$$\limsup_{n\to\infty} P\left(\sup_{j>J_n} \max_{0\leqslant k<2^j} \frac{1}{\rho(2^{-j})} |\lambda'_{j,k}(\zeta_n^{\mathrm{se}})| > \varepsilon\right) = 0 \tag{54}$$

and

$$\lim_{J\to\infty} \limsup_{n\to\infty} P\left(\sup_{J\leqslant j\leqslant J_n} \max_{0\leqslant k<2^j} \frac{1}{\rho(2^{-j})} |\lambda'_{j,k}(\zeta_n^{\mathrm{se}})| > 3\varepsilon\right) = 0, \tag{55}$$

where

$$\lambda'_{j,k}(\zeta_n^{\mathrm{se}}) := \zeta_n^{\mathrm{se}}((k+1)2^{-j}) - \zeta_n^{\mathrm{se}}(k2^{-j}), \quad 0 \leqslant k < 2^j.$$

To start with (54), following the same steps which led to (49) we obtain with k, l such that

$$\frac{V_{k-1}^2}{V_n^2} < t \leqslant \frac{V_k^2}{V_n^2}, \quad \frac{V_{l-1}^2}{V_n^2} < t+h \leqslant \frac{V_l^2}{V_n^2},$$

the upper bound

$$|\zeta_n^{\mathrm{se}}(t+h) - \zeta_n^{\mathrm{se}}(t)| \leqslant \left(2 + \frac{|S_{(l,k]}|}{V_{(l,k]}}\right)\sqrt{h},$$

where we use the notations

$$S_{(i,j]} := \sum_{i<k\leqslant j} X_k, \quad V_{(i,j]} := \left(\sum_{i<k\leqslant j} X_k^2\right)^{1/2}$$

UNCORRECTED PROOF

with the usual convention of null value for a sum indexed by the empty set. Writing $T_{k,l} := 2 + |S_{(l,k]}|/V_{(l,k]}$, this gives

$$|\zeta_n^{\mathrm{se}}(t+h) - \zeta_n^{\mathrm{se}}(t)| \leqslant \sqrt{h} \max_{1 \leqslant k \leqslant l \leqslant n} T_{k,l}. \tag{56}$$

By Giné et al. (1997, Theorem 2.5), the $T_{k,l}$ are *uniformly* subgaussian. It is worth recalling here and for further use, that if the random variables Y_i $(1 \leqslant i \leqslant N)$ are subgaussian, then so is $\max_{1 \leqslant i \leqslant N} |Y_i|$, which more precisely satisfies

$$\left\| \max_{1 \leqslant i \leqslant N} |Y_i| \right\|_{\phi_2} \leqslant a(\log N)^{1/2} \max_{1 \leqslant i \leqslant N} \|Y_i\|_{\phi_2}, \tag{57}$$

where a is an absolute constant and $\| \, \|_{\phi_2}$ denotes the Orlicz norm associated to the Young function $\phi_2(t) := \exp(t^2) - 1$. Applying (57) to the n^2 random variables $T_{k,l}$, we obtain (with constants c, C whose value may vary at each occurence)

$$\begin{aligned} P\left(\sup_{j > J_n} \max_{0 \leqslant k < 2^j} \frac{1}{\rho(2^{-j})} |\lambda'_{j,k}(\zeta_n^{\mathrm{se}})| > \varepsilon \right) &\leqslant \sum_{j > J_n} P\left(\max_{1 \leqslant k \leqslant l \leqslant n} T_{k,l} > c\varepsilon j^{\beta} \right) \\ &\leqslant \sum_{j > J_n} C \exp\left(\frac{-cj^{2\beta}}{\log n} \right). \end{aligned} \tag{58}$$

Now choose $J_n = (\log n)^{\gamma}$ with $1 > \gamma > (2\beta)^{-1}$. Then $2\beta - 1/\gamma$ is strictly positive and using

$$j^{2\beta} = j^{1/\gamma} j^{2\beta - 1/\gamma} > J_n^{1/\gamma} j^{2\beta - 1/\gamma} = j^{2\beta - 1/\gamma} \log n,$$

we see that the right-hand side in (58) is bounded by $\sum_{j > J_n} C \exp(-cj^{2\beta - 1/\gamma})$, whence (54) follows.

To prove (55), we start with

$$P\left(\max_{J \leqslant j \leqslant J_n} \max_{0 \leqslant k < 2^j} \frac{1}{\rho(2^{-j})} |\lambda'_{j,k}(\zeta_n^{\mathrm{se}})| > 3\varepsilon \right) \leqslant P_1 + P_2 + P_3 \tag{59}$$

with P_1, P_2 and P_3 defined below. First introduce the event

$$A_n = \left\{ \sup_{t \in [0,1]} \left| \frac{V^2_{\tau_n(t)}}{V_n^2} - \frac{V^2_{[nt]}}{V_n^2} \right| \leqslant \delta_n \right\} \cap \left\{ \sup_{t \in [0,1]} \left| \frac{V^2_{[nt]}}{V_n^2} - t \right| \leqslant \delta_n \right\}.$$

where δ_n is chosen as in (11), keeping the freedom of choice of the constant c.

Now we define

$$\begin{aligned} P_1 &:= P(A_n^c), \\ P_2 &:= P\left(A_n \cap \left\{ \max_{J \leqslant j \leqslant J_n} \max_{0 \leqslant k < 2^j} \frac{1}{\rho(2^{-j})} \frac{|S_{[(k+1)2^{-j}n]} - S_{[k2^{-j}n]}|}{V_n} > \varepsilon \right\} \right), \\ P_3 &:= P\left(A_n \cap \left\{ \max_{J \leqslant j \leqslant J_n} \max_{0 \leqslant k < 2^j} \frac{1}{\rho(2^{-j})} \max_{|l - [k2^{-j}n]| \leqslant n\delta_n} \right. \right. \\ &\qquad \left. \left. \left[\frac{|S_l - S_{[k2^{-j}n]}|}{V_n} + \frac{2}{2^{j/2}} \right] > 2\varepsilon \right\} \right). \end{aligned}$$

UNCORRECTED PROOF

The following easy estimates

$$\sup_{t\in[0,1]}\left|\frac{V_{[nt]}^2}{V_n^2}-t\right| \leqslant \max_{1\leqslant k\leqslant n}\left|\frac{V_k^2}{V_n^2}-\frac{k}{n}\right|+\frac{1}{n},$$

$$\sup_{t\in[0,1]}\left|\frac{V_{\tau_n(t)}^2}{V_n^2}-\frac{V_{[nt]}^2}{V_n^2}\right| \leqslant \max_{1\leqslant k\leqslant n}\frac{X_k^2}{V_n^2}+\max_{1\leqslant k\leqslant n}\left|\frac{V_k^2}{V_n^2}-\frac{k}{n}\right|+\frac{1}{n},$$

lead by (9) and (10) to

$$P(A_n^c)\to 0. \tag{60}$$

So P_1 will be killed by taking the lim sup in n.

To control P_2, first write with self-explanatory notations

$$\frac{|S_{[(k+1)2^{-j}n]}-S_{[k2^{-j}n]}|}{V_n}=\frac{|S_{[(k+1)2^{-j}n]}-S_{[k2^{-j}n]}|}{V_{([k2^{-j}n],[(k+1)2^{-j}n]]}}\times\frac{V_{([k2^{-j}n],[(k+1)2^{-j}n]]}}{V_n}.$$

Observing that on the event A_n, we have

$$\frac{V_{([k2^{-j}n],[(k+1)2^{-j}n]]}}{V_n}\leqslant\sqrt{2^{-j}+\delta_n}$$

and assuming that

$$\delta_n\leqslant 2^{-J_n}, \tag{61}$$

we get

$$P_2\leqslant\sum_{J\leqslant j\leqslant J_n}P\left(\max_{0\leqslant k<2^j}\frac{1}{\rho(2^{-j})}\frac{|S_{[(k+1)2^{-j}n]}-S_{[k2^{-j}n]}|}{V_{([k2^{-j}n],[(k+1)2^{-j}n]]}}>\sqrt{2}\varepsilon 2^{j/2}\right).$$

Since we are dealing now with the maximum of 2^j uniformly subgaussian random variables (their φ_2 norms are bounded by a constant which depends only on the distribution of X_1), this leads to

$$P_2\leqslant\sum_{J\leqslant j\leqslant J_n}C\exp(-cj^{2\beta-1})\leqslant\sum_{j=J}^{\infty}C\exp(-cj^{2\beta-1}). \tag{62}$$

To control P_3, we first get rid of the residual term by noting that

$$\frac{2}{\rho(2^{-j})2^{j/2}}=\frac{c}{j^\beta}<\varepsilon\quad\text{for } j\geqslant J\geqslant J(\varepsilon),$$

uniformly in n. So for $J\geqslant J(\varepsilon)$,

$$P_3\leqslant P\left(A_n\cap\left\{\max_{J\leqslant j\leqslant J_n}\max_{0\leqslant k<2^j}\frac{1}{\rho(2^{-j})}\max_{|l-[k2^{-j}n]|\leqslant n\delta_n}\frac{|S_l-S_{[k2^{-j}n]}|}{V_n}>\varepsilon\right\}\right).$$

On the event A_n we have for any l such that $|l-[k2^{-j}n]|\leqslant n\delta_n$,

$$\frac{|V_{[k2^{-j}n]}^2-V_l^2|}{V_n^2}\leqslant 2\delta_n.$$

It follows that

$$P_3\leqslant P\left(\max_{J\leqslant j\leqslant J_n}\max_{0\leqslant k<2^j}\max_{|l-[k2^{-j}n]|\leqslant n\delta_n}\frac{|S_l-S_{[k2^{-j}n]}|}{|V_{[k2^{-j}n]}^2-V_l^2|^{1/2}}>\frac{\varepsilon\rho(2^{-j})}{\sqrt{2\delta_n}}\right).$$

UNCORRECTED PROOF

Using the invariance of distributions under translations on k, we get

$$P_3 \leqslant \sum_{J \leqslant j \leqslant J_n} 2^j P\left(\max_{0<l\leqslant[2n\delta_n]} \frac{|S_l|}{V_l} > \frac{\varepsilon\rho(2^{-j})}{\sqrt{2\delta_n}}\right)$$

$$\leqslant \sum_{J \leqslant j \leqslant J_n} 2^j C \exp\left(-\frac{c2^{-j}j^{2\beta}}{\delta_n \log n}\right)$$

$$\leqslant C \sum_{J \leqslant j \leqslant J_n} 2^j \exp\left(-\frac{c2^{-J_n}}{\delta_n \log n} j^{2\beta}\right).$$

Now we see that the following convergence rate (stronger than (61))

$$\delta_n = \frac{1}{2^{J_n} \log n} = \frac{2^{-(\log n)^\gamma}}{\log n}, \quad \text{with} \quad \frac{1}{2\beta} < \gamma < 1,$$

is sufficient to obtain (55). The proof is complete. □

Proof of Corollary 6. As is X_1 is square integrable, X_1 is in *DAN*. The convergence rates (9) and (10) required by Theorem 5 are provided by the two following lemmas, recalling that with our choice (11) of δ_n, we have $n^{-\varepsilon} = o(\delta_n)$ for any $\varepsilon > 0$. □

Lemma 13. *If* $\mathbf{E}|X_1|^{2+\delta} < \infty$ *for some* $\delta > 0$, *then almost surely*

$$n^{-c} \max_{1\leqslant k\leqslant n} \left|\frac{V_k^2}{V_n^2} - \frac{k}{n}\right| \to 0, \tag{63}$$

where $c = \delta/(2+2\delta)$.

Proof. By Marcinkiewicz SLLN, if the i.i.d. sequence (Y_k) satisfies $\mathbf{E}|Y_1|^p < \infty$ for some $1 \leqslant p < 2$, then $n^{-1/p}(\sum_{k\leqslant n} Y_k - n\mathbf{E}Y_1)$ goes to 0 almost surely. Applying this to $Y_1 = X_1^2$ and $p = 1 + \delta/2$ gives

$$\frac{V_n^2}{n} = 1 + n^{1/p-1}\varepsilon_n, \quad n \geqslant 1,$$

where the random sequence (ε_n) goes to zero almost surely. Since we assume $P(X_1 = 0) < 1$, we have $P(\forall n \geqslant 1, V_n = 0) = 0$. On each event $\{V_n^2 > 0\}$, we may write with $a = 1 - 1/p$,

$$\frac{V_k^2}{V_n^2} - \frac{k}{n} = \frac{k}{n}\left(\frac{V_k^2}{k}\frac{n}{V_n^2} - 1\right) = \frac{k}{n} \times \frac{k^{-a}\varepsilon_k - n^{-a}\varepsilon_n}{1 + n^{-a}\varepsilon_n}.$$

For each $n \geqslant n_0 = n_0(\omega)$ large enough, $n^{-a}\varepsilon_n > -1/2$. Now for an exponent $0 < b < 1$ to be precised later, we have

$$\left|\frac{V_k^2}{V_n^2} - \frac{k}{n}\right| \leqslant 4n^{b-1} \sup_{i\geqslant 1}|\varepsilon_i| \quad \text{for} \quad n \geqslant n_0,\ 1 \leqslant k \leqslant n^b$$

and

$$\left|\frac{V_k^2}{V_n^2} - \frac{k}{n}\right| \leqslant 4n^{-ab} \sup_{i\geqslant n^b}|\varepsilon_i| \quad \text{for} \quad n \geqslant n_0,\ n^b < k \leqslant n.$$

The optimal choice of b given by $1 - b = ab$ leads to the announced conclusion with $c = a/(a+1) = \delta/(2+2\delta)$. □

UNCORRECTED PROOF

SPA 1047

Lemma 14. *If* $\mathbf{E}|X_1|^{2+\delta} < \infty$ *for some* $\delta > 0$, *then almost surely*

$$n^d \max_{1\leqslant k\leqslant n} \frac{X_k^2}{V_n^2} \to 0 \tag{64}$$

for any $d < \delta/(2+\delta)$.

Proof. We use the same trick as in O'Brien (1980, p. 542). For any positive ε we have (noting the key role of i.o. in the following inequalities)

$$P\left(\max_{1\leqslant k\leqslant n} \frac{X_k^2}{V_n^2} > \varepsilon n^{-d}, \text{ i.o.}\right) \leqslant P\left(V_n^2 < \frac{n}{2}, \text{ i.o.}\right) + P\left(\max_{1\leqslant k\leqslant n} X_k^2 > \frac{n}{2}\varepsilon n^{-d}, \text{ i.o.}\right)$$
$$= 0 + P\left(X_n^2 > \frac{n}{2}\varepsilon n^{-d}, \text{ i.o.}\right).$$

Now observe that

$$\sum_{n=1}^{\infty} P\left(X_n^2 > \frac{n}{2}\varepsilon n^{-d}\right) \leqslant \left(\frac{2}{\varepsilon}\right)^{1+\delta/2} \mathbf{E}|X_1|^{2+\delta} \sum_{n=1}^{\infty} \frac{1}{n^{(1-d)(1+\delta/2)}}.$$

For any d such that $(1-d)(1+\delta/2) > 1$, Borel–Cantelli's Lemma leads to

$$P\left(\max_{1\leqslant k\leqslant n} \frac{X_k^2}{V_n^2} > \varepsilon n^{-d}, \text{ i.o.}\right) = 0.$$

As ε is arbitrary, the result is proved. □

Uncited Reference

Račkauskas and Suquet (1999a)

Acknowledgements

The first author would like to thank Vidmantas Bentkus for a number of stimulating discussions on the invariance principle for self-normalized sums.

References

[1] Araujo, A., Giné, E., 1980. The Central Limit Theorem for Real and Banach Valued Random Variables. Wiley, New York.
[2] Bentkus, V., Götze, F., 1996. The Berry–Esseen bound for student's statistic. Ann. Probab. 24, 491–503.
[3] Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.
[4] Christiakov, G.P., Götze, F., 1999. Moderate deviations for self-normalized sums. Preprint 99-048, SFB 343, University of Bielefeld.
[5] Chuprunov, A.N., 1997. On convergence of random polygonal lines under Student-type normalizations. Theory Probab. Appl. 41, 756–761.
[6] Ciesielski, Z., 1960. On the isomorphisms of the spaces H_α and m. Bull. Acad. Pol. Sci. Ser. Sci. Math. Phys. 8, 217–222.
[7] Egorov, V.A., 1997. On the asymptotic behavior of self-normalized sums of random variables. Theory Probab. Appl. 41, 542–548.

UNCORRECTED PROOF

[8] Erickson, R.V., 1981. Lipschitz smoothness and convergence with applications to the central limit theorem for summation processes. Ann. Probab. 9, 831–851.

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

[10] Griffin, P.S., Kuelbs, J., 1989. Self-normalized laws of the iterated logarithm. Ann. Probab. 17, 1571–1601.

[11] Griffin, P.S., Mason, D.M., 1991. On the asymptotic normality of self-normalized sums. Proc. Cambridge Philos. Soc. 109, 597–610.

[12] Hamadouche, D., 1998. Invariance principles in Hölder spaces. Portugal. Math. 57 (2000) 127–151.

[13] Kerkyacharian, G., Roynette, B., 1991. Une démonstration simple des théorèmes de Kolmogorov, Donsker et Ito-Nisio. C. R. Acad. Sci. Paris Sér. I 312, 877–882.

[14] Lamperti, J., 1962. On convergence of stochastic processes. Trans. Amer. Math. Soc. 104, 430–435.

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

[16] O'Brien, G.L., 1980. A limit theorem for sample maxima and heavy branches in Galton–Watson trees. J. Appl. Probab. 17, 539–545.

[17] Račkauskas, A., Suquet, Ch., 1999a. Central limit theorem in Hölder spaces. Probab. Math. Statist. 19, 155–174.

[18] Račkauskas, A., Suquet, Ch., 1999b. Random fields and central limit theorem in some generalized Hölder spaces. In: Grigelionis, B., et al. (Eds.), Probability Theory and Mathematical Statistics, Proceedings of the seventh Vilnius Conference, 1998. TEV, Vilnius and VSP, Utrecht, pp. 599–616.

[19] Račkauskas, A., Suquet, Ch., 1999c. On the Hölderian functional central limit theorem for i.i.d. random elements in Banach space, Pub. IRMA Lille 50-III, Proceedings of the Fourth Hungarian Colloquium on Limit Theorems of Probability and Statistics, in preparation.

[20] Račkauskas, A., Suquet, Ch., 2000. Convergence of self-normalized partial sums processes in $C[0,1]$ and $D[0,1]$, Pub. IRMA Lille 53-VI, preprint.

[21] Shao, Q.-M., 1997. Self-normalized large deviations. Ann. Probab. 25, 285–328.

[22] Shao, Q.-M., 1999. A Cramér type large deviation result for Student's t-statistic. J. Theoret. Probab. 12, 387–398.

UNCORRECTED PROOF