Invariance principles for adaptive self-normalized partial sums processes

Alfredas Račkauskas, Charles Suquet

Abstract

Let $\xi_n$ be the adaptive polygonal process of self-normalized partial sums $S_n = \sum_{1 \leq i \leq n} X_i$ of i.i.d. random variables defined by linear interpolation between the points $(\sqrt{\frac{\xi_n}{\xi_n^2}}, S_i/\xi_n)$, $k \leq n$, where $\xi_n^2 = \sum_{1 \leq i \leq n} X_i^2$. We investigate the weak Hölder convergence of $\xi_n^{\text{se}}$ to the Brownian motion $W$. We prove particularly that when $X_1$ is symmetric, $\xi_n^{\text{se}}$ 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.

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 $\xi_n^{\text{se}}$. 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, $\xi_n^{\text{se}}$ behaves better than the classical Donsker–Prohorov partial sum processes $\xi_n$. 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 $V_k^2/V_n^2$ ($0 \leq k \leq n$) instead of the deterministic $\sqrt{\frac{\xi_n}{\xi_n^2}}$.

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).
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 notations.

By \( \zeta_n \) (respectively \( \xi_n \)) we denote the random polygonal partial sums process defined on \([0,1]\) by linear interpolation between the vertices \((V^2_k/V^2_n, S_k), k=0,1,\ldots,n\) (respectively \((k/n,S_k), k=0,1,\ldots,n\)), where

\[
S_k = X_1 + \cdots + X_k, \quad V^2_k = X_1^2 + \cdots + X_k^2.
\]

For the special case \( k=0 \), we put \( S_0 = 0, V_0 = 0 \).

The upper scripts \( sr \) or \( se \) mean, respectively, normalization by square root of \( n \) or self-normalization. Hence,

\[
\zeta_n^{sr} = \frac{\zeta_n}{\sqrt{n}}, \quad \zeta_n^{se} = \frac{\zeta_n}{V_n}, \quad r_n^{sr} = \frac{r_n}{\sqrt{n}}, \quad r_n^{se} = \frac{r_n}{V_n}.
\]

By convention the random functions \( \zeta_n^{se} \) and \( \xi_n^{se} \) are defined to be the null function on the event \( \{V_n = 0\} \). Finally, the step partial sums processes \( \Xi_n, Z_n, \Xi_n^{se}, \) 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

\[
\zeta_n^{sr} \xrightarrow{D} W,
\]

in \( C[0,1] \), where \((W(t), t \in [0,1])\) is a standard Wiener process and \( \xrightarrow{D} \) 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), Rackauskas and Suquet (1999c).

Considering a symmetric random variable \( X_1 \) such that \( P(X_1 > u) = 1/(2u^\alpha) \), \( u \geq 1 \), Lamperti (1962) noticed that the corresponding sequence \( (\zeta_n^{sa}, r_n^{sa}) \) 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, Rackauskas and Suquet (1999c) proved more precisely that if \( (\zeta_n^{sa}, r_n^{sa}) \) satisfies the invariance principle in \( H_\alpha[0,1] \) for some \( 0 < \alpha < 1/2 \), then necessarily

\[
\sup_{t > 0} t^\alpha P(|X_1| > t) < \infty
\]

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”
(denoted by $X_1 \in \text{DAN}$) means that there exists a sequence $b_n \uparrow \infty$ such that
\[ b_n^{-1} S_n \xrightarrow{D} N(0,1). \]  
(3)

According to O’Brien’s (1980) result: $X_1 \in \text{DAN}$ if and only if
\[ V_n^{-1} \max_{1 \leq k \leq n} |X_k|^p \xrightarrow{P} 0, \]  
(4)

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

5 Theorem 1. The convergence
\[ \varepsilon_n \xrightarrow{D} W \]  
holds in the space $C[0,1]$ if and only if $X_1 \in \text{DAN}$.

7 Theorem 2. The convergence
\[ \varepsilon_n \xrightarrow{D} W \]  
holds in the space $C[0,1]$ if and only if $X_1 \in \text{DAN}$.

9 Let us remark that the necessity of $X_1 \in \text{DAN}$ 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 (Račkūkas and Suquet, 2000) for the step processes $\Xi_n$ and $\Xi_n^*$ within the Skorohod space $D(0,1)$.

For a modulus of continuity $\rho : [0,1] \to \mathbb{R}$, denote by $H_\rho[0,1]$ the set of continuous functions $x : [0,1] \to \mathbb{R}$ such that $\omega_\rho(x,1) < \infty$, where
\[ \omega_\rho(x,\delta) := \sup_{t,s \in [0,1]} \frac{|x(t) - x(s)|}{\rho|t-s|}. \]

The set $H_\rho[0,1]$ is a Banach space when endowed with the norm
\[ \|x\|_\rho := |x(0)| + \omega_\rho(x,1). \]

Define
\[ H_\rho^n[0,1] = \{ x \in H_\rho[0,1] : \lim_{\delta \to 0} \omega_\rho(x,\delta) = 0 \}. \]

Then $H_\rho^n[0,1]$ is a closed separable subspace of $H_\rho[0,1]$. In what follows we assume that the function $\rho$ satisfies technical conditions (12) to (16) (see Section 2). These assumptions are fulfilled particularly when $\rho = \rho_{x,\beta}$, $0 < x < 1$, $\beta \in \mathbb{R}$, defined by
\[ \rho_{x,\beta}(h) := h^\alpha \ln^{\beta}(c/h), \quad 0 < h \leqslant 1 \]
for a suitable constant $c$. We write $H_{x,\beta}$ and $H_{x,\beta}^n$ for $H_\rho[0,1]$ and $H_\rho^n[0,1]$, respectively, when $\rho = \rho_{x,\beta}$ and we abbreviate $H_{x,0}$ in $H_x$.

With respect to this Hölder scale $H_{x,\beta}$, we obtain an optimal result when $X_1$ is symmetric.
Theorem 3. Assume that \( \rho \) satisfies conditions (12)–(16) and
\[
\lim_{j \to \infty} \frac{2^j \rho^2(2^{-j})}{j} = \infty.
\] (7)
If \( X_1 \) is symmetric and \( X_1 \in \mathcal{DAN} \) then
\[
\mathcal{CL}_n \overset{D}{\to} W, \quad (8)
\]
in \( H_\rho^\beta[0,1] \).

Corollary 4. If \( X_1 \) is symmetric and \( X_1 \in \mathcal{DAN} \) then (8) holds in the space \( H_\rho^\beta \) for any \( \beta > 1/2 \).

It is well known that the Wiener process has a version in the space \( H_1^{1/2,1/2} \) but none in \( H_1^{1/2,1/2} \). Hence Corollary 4 gives the best result possible in the scale of the separable Hölder spaces \( H_\rho^\beta \). In Račkauskas and Suquet (1999c) it is proved that if the classical partial sums process \( \mathcal{CL}_n \) converges in \( H_1^{1/2,1/2} \) for some \( \beta > 1/2 \), then \( \|X_1\|_\phi < \infty \), where \( \|X_1\|_\phi \) is the Orlicz norm related to the Young function \( \phi(r) = \exp(r^\gamma) - 1 \) 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 \( \mathcal{CL}_n \overset{D}{\to} W \) in \( H_\rho^\beta \), then one can prove that \( \mathbf{E}X_1^2 < \infty \). Therefore \( \mathcal{CL}_n \overset{D}{\to} W \) in \( H_\rho^\beta \) 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
\[
P\left( \max_{1 \leq k \leq n} X_k^2 \leq \delta_n \right) \to 0 
\]
\[
P\left( \max_{1 \leq k \leq n} \frac{V_k^2}{V_n^2} \geq \delta_n \right) \to 0,
\]
with
\[
\delta_n = c \frac{2^{-(\log n)^\gamma}}{\log n} \quad \text{for some} \quad \frac{1}{2\beta} < \gamma < 1 \quad \text{and some} \quad c > 0.
\]
Then
\[
\mathcal{CL}_n \overset{D}{\to} W \quad \text{in} \quad H_\rho^\beta[0,1].
\]

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

Corollary 6. If for some \( \varepsilon > 0 \), \( \mathbf{E}|X_1|^{2+m\varepsilon} < \infty \), then for any \( \beta > 1/2 \), \( \mathcal{CL}_n \) converges weakly to \( W \) in the space \( H_\rho^\beta[0,1] \).

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 Götz (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 Christiaikov and Götz (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 $H^s[0, 1]$ 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 $\rho$ satisfies the following technical conditions where $c_1$, $c_2$ and $c_3$ are positive constants:

$$\rho(0) = 0, \quad \rho(\delta) > 0, \quad 0 < \delta \leq 1,$$

(12)

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

(13)

$$\rho(2\delta) \leq c_1 \rho(\delta), \quad 0 \leq \delta \leq 1/2,$$

(14)

$$\int_0^1 \rho\left(\frac{u}{u}\right) \frac{du}{u} \leq c_2 \rho(\delta), \quad 0 < \delta \leq 1,$$

(15)

$$\delta \int_0^{1/\delta} \rho\left(\frac{u}{u^2}\right) \frac{du}{u^2} \leq c_3 \rho(\delta), \quad 0 < \delta \leq 1.$$

(16)

For instance, elementary computations show that the functions

$$\rho(\delta):=\delta^z \ln^\beta \left(\frac{\epsilon}{\delta}\right), \quad 0 < z < 1, \quad \beta \in \mathbb{R},$$

(17)

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

2
Write $D_j$ for the set of dyadic numbers of level $j$ in $[0, 1]$, i.e. $D_0 = \{0, 1\}$ and for $j \geq 1$,
$$D_j = \{(2k + 1)2^{-j}; 0 \leq k < 2^{j-1}\}.$$

For any continuous function $x : [0, 1] \to \mathbb{R}$, define
$$\lambda_{0,j}(x) := x(t), \quad t \in D_0$$
and for $j \geq 1$,
$$\lambda_{j,i}(x) := x(t) - \frac{1}{4}(x(t + 2^{-j}) + x(t - 2^{-j})), \quad t \in D_j.$$

The $\lambda_{j,i}(x)$ are the coefficients of the expansion of $x$ in a series of triangular functions.

The $j$th partial sum $E_jx$ of this series is exactly the polygonal line interpolating $x$ between the dyadic points $k2^{-j}(0 \leq k \leq 2^j)$. Under (12)–(16), the norm $\|x\|_p$ is equivalent to the sequence norm
$$\|x\|_{p, \text{seq}} := \sup_{j \geq 0, t \in D_j} \frac{1}{\rho(2^{-j})} \max |\lambda_{j,i}(x)|.$$

In particular, both norms are finite if and only if $x$ belongs to $H^p$. It is easy to check that
$$\|x - E_jx\|_{p, \text{seq}} = \sup_{i \geq j} \frac{1}{\rho(2^{-i})} \max |\lambda_{i,i}(x)|.$$

**Proposition 7.** The sequence $(Y_n)$ of random elements in $H^p_0$ is tight if and only if the following two conditions are satisfied:
(i) For each $t \in [0, 1]$, the sequence $(Y_n(t))_{n \geq 1}$ is tight on $\mathbb{R}$.
(ii) For each $\varepsilon > 0$,
$$\lim_{j \to \infty} \sup_{n \geq 1} P(\|Y_n - E_jY_n\|_{p, \text{seq}} > \varepsilon) = 0.$$

**Remark 8.** Condition (ii) in Proposition 7 may be replaced by
$$\lim_{j \to \infty} \lim_{n \to \infty} P(\|Y_n - E_jY_n\|_{p, \text{seq}} > \varepsilon) = 0. \quad (17)$$

**2.2. Adaptive time and DAN**

We establish here the technical results on the adaptive time when $X_1 \in \text{DAN}$ 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 $X_1 \in \text{DAN}$, 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
$$b_n^{-2}V_n^2 \xrightarrow{p} 1. \quad (18)$$
We have moreover for each \( \tau > 0 \), putting \( b_n = n^{-1/2} \ell_n \),
\[
n\mathbb{P}(|X_1| > \tau \ell_n \sqrt{n}) \to 0,
\]
(19)
\[
n^{-1} \mathbb{E}(X_1^2; |X_1| \leq \tau \ell_n \sqrt{n}) \to 1,
\]
(20)
\[
n\mathbb{E}(X_1; |X_1| \leq \tau \ell_n \sqrt{n}) \to 0,
\]
(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 \( X_1 \in \text{DAN} \), then
\[
\sup_{0 \leq t \leq 1} \left| \frac{V_{[n]}^2 - t}{V_n^2} \right| \xrightarrow{p} 0.
\]
(22)

**Proof.** Consider the truncated random variables
\[
X_{n,i} := b_n^{-1}(X_i; X_i^2 \leq b_n^2), \quad i = 1, \ldots, n.
\]
Define \( V_{n,0} = 0 \) and \( V_{n,k} = X_{n,1}^2 + \cdots + X_{n,k}^2 \) for \( k = 1, \ldots, n \). Set
\[
\nu_n = \sup_{0 \leq t \leq 1} \left| \frac{V_{[n]}^2 - t}{V_n^2} \right| \quad \text{and} \quad \tilde{\nu}_n = \sup_{0 \leq t \leq 1} \left| \frac{V_{[n]}^2 - t}{V_n^2} \right|.
\]
Then we have for \( \lambda > 0 \),
\[
P(\nu_n > \lambda) \leq P(\tilde{\nu}_n > \lambda) + nP(X_1^2 > b_n^2).
\]
Due to (19) the proof of (22) reduces to the proof of
\[
\tilde{\nu}_n \xrightarrow{p} 0.
\]
(23)
Since \( V_{n,k}^2 \leq V_{n,n}^2 \) for \( k = 0, \ldots, n \), the elementary estimate
\[
\left| \frac{V_{n,k}^2 - k}{n} \right| \leq \frac{V_{n,k}^2}{V_{n,n}^2} |1 - V_{n,n}^2| + \frac{V_{n,k}^2 - k}{n}
\]
leads to
\[
\tilde{\nu}_n \leq \max_{0 \leq k \leq n} \left| \frac{V_{n,k}^2 - k}{n} \right| + |1 - V_{n,n}^2| + \frac{1}{n}.
\]
(24)
Noting that \( V_{n,n}^2 = b_n^{-2} V_n^2 R_n \) with
\[
R_n := \frac{1}{V_n^2} \sum_{i=1}^n (X_i^2; X_i^2 \leq b_n^2),
\]
we clearly have \( R_n \leq 1 \) a.s. and
\[
P(R_n < 1) = P \left( \max_{1 \leq i \leq n} |X_i| > b_n \right) \leq nP(|X_1| > b_n),
\]
which goes to zero by (19). This together with (18) gives
\[ \lim_{n \to \infty} \frac{V_{n,n}^2}{n} = 1. \]
Hence the proof of (23) reduces to
\[ \max_{0 \leq k \leq n} \left| \frac{V_{n,k}^2 - k/n}{n} \right| \overset{p}{\to} 0. \] (25)

For this convergence we have
\[ \max_{0 \leq k \leq n} |V_{n,k}^2 - k/n| \leq \max_{0 \leq k \leq n} |V_{n,k}^2 - \mathbb{E}V_{n,k}^2| + \max_{0 \leq k \leq n} |\mathbb{E}V_{n,k}^2 - k/n|. \]
Noting that
\[ \mathbb{E}V_{n,k}^2 - \frac{k}{n} = \frac{k}{n} (nb_n^{-2} \mathbb{E} (X_1^2; X_1^2 \leq b_n^2) - 1) \]
gives
\[ \max_{0 \leq k \leq n} \left| \mathbb{E}V_{n,k}^2 - \frac{k}{n} \right| \leq \left| nb_n^{-2} \mathbb{E} (X_1^2; X_1^2 \leq b_n^2) - 1 \right|, \]
which goes to zero by (20). Hence it remains to prove
\[ \max_{0 \leq k \leq n} |V_{n,k}^2 - \mathbb{E}V_{n,k}^2| \overset{p}{\to} 0. \] (26)

Putting \( T_{n,k} := V_{n,k}^2 - \mathbb{E}V_{n,k}^2 \), we have by Ottaviani inequality
\[ P \left( \max_{1 \leq k \leq n} |T_{n,k}| > 2\lambda \right) \leq 1 - \max_{1 \leq k \leq n} P(|T_{n,n} - T_{n,k}| > \lambda). \] (28)
Due to (25), we are left with the control of \( I := \max_{1 \leq k \leq n} P(|T_{n,k}| > \lambda) \). By Chebyshev’s inequality
\[ I \leq \lambda^{-2} \max_{1 \leq k \leq n} \mathbb{E}T_{n,k}^2 \leq \lambda^{-2} n \mathbb{E}X_{n,1}^4 \]
and we have to consider \( I_1 = n \mathbb{E}X_{n,1}^4 = nb_n^{-4} \mathbb{E}(X_1^4; |X_1| \leq b_n) \). For any \( 0 < \tau < 1 \),
\[ \mathbb{E}(X_1^4; |X_1| \leq b_n) \leq \mathbb{E}(X_1^4; |X_1| \leq \tau b_n) + \mathbb{E}(X_1^4; \tau b_n \leq |X_1| \leq b_n) \]
\[ \leq \tau^2 b_n^2 \mathbb{E}(X_1^2; |X_1| \leq \tau b_n) + b_n^4 P(|X_1| \geq \tau b_n). \]

So
\[ I_1 \leq \tau^2 nb_n^{-2} \mathbb{E}(X_1^2; |X_1| \leq \tau b_n) + n P(|X_1| \geq \tau b_n). \]
Choosing \( \tau = \sqrt{\lambda}/2 \) in (19) and (20), we can achieve \( I \leq 1/2 \) for \( n \) large enough and the proof is complete. \( \square \)

**Remark 10.** If \( X_1 \in \mathcal{DAN} \), we also have
\[ \sup_{0 \leq t \leq 1} \left| \frac{V_{[nt]+1}^2}{V_n^2} - t \right| \overset{p}{\to} 0. \] (29)
Indeed, recalling (4), it suffices to write
\[ V_{2n+1}^2 - V_n^2 \leq \left( \frac{1}{V_{n+1}^2} \max_{k \leq n+1} X_k^2 \right) \frac{V_{n+1}^2}{V_n^2}, \]
and observe that \( V_{n+1}^2/V_n^2 \) converges to 1 in probability since by Lemma 9,
\[ \sup_{0 \leq t \leq 1} \left| \frac{V_{2n+1}^2}{V_n^2} - t \right| \xrightarrow{p} 0. \]

3 Remark 11. For each \( t \in [0, 1] \),
\[ \frac{b_{2[nt]}}{b_n^2} \to t. \] (30)

This is a simple by-product of Lemma 9, writing
\[ \frac{b_{2[nt]}^2}{b_n^2} = \frac{V_{2n}^2}{b_n^2} \times \frac{b_{2[nt]}^2}{V_{2[nt]}^2} \times \frac{V_{2[nt]}^2}{V_n^2} \]
and noting that for fixed \( t > 0 \) and \( n \geq n_0 \) large enough \([nt] < [(n+1)t]\) so the sequence \((b_{2[nt]}^2/V_{2[nt]}^2)_{n \geq n_0}\) is a subsequence of \((b_n^2/V_n^2)_{n \geq n_0}\) which converges in probability to 1 by (18).

Define the random variables
\[ \tau_n(t) = \max\{k = 0, \ldots, n; V_k^2 \leq tV_n^2\}, \quad t \in [0, 1], \] (31)
so that we have \( \tau_n(1) = n \) and for \( 0 \leq t < 1 \),
\[ \frac{V_{\tau_n(t)}^2}{V_n^2} \leq t < \frac{V_{\tau_n(t)+1}^2}{V_n^2}. \] (32)

Lemma 12. If \( X_1 \in DAN \) then
\[ \sup_{t \in [0,1]} |n^{-1}\tau_n(t) - t| \xrightarrow{p} 0. \] (33)

11 Proof. The result will follow from Remark 10, if we check the inclusion of events
\[ \left\{ \sup_{t \in [0,1]} |n^{-1}\tau_n(t) - t| > \varepsilon \right\} \subset \left\{ \sup_{u \in [0,1]} \left| \frac{V_{[nu]+1}^2}{V_n^2} - u \right| \geq \varepsilon \right\}. \] (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
\[ \tau_n(s) > n(s + \varepsilon) \quad \text{or} \quad \tau_n(s) < n(s - \varepsilon) \] (35)
(36)

Observe that under (35), \( s + \varepsilon < 1 \), while under (36), \( s - \varepsilon > 0 \). From the definition of \( \tau_n \), (35) gives an integer \( k > n(s + \varepsilon) \) such that \( V_k^2/V_n^2 \leq s \), whence
\[ \frac{V_{[n(s+\varepsilon)]+1}^2}{V_n^2} \leq s. \] (37)
On the other hand, under (36), we have \( V_k^2/V_n^2 > s \) for every \( k \geq n(s-\varepsilon) \) and in particular

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

Recasting (37) and (38) under the form

\[
\frac{V_{n(s+\varepsilon)+1}^2}{V_n^2} - (s + \varepsilon) \leq -\varepsilon
\]

\[
\frac{V_{n(s-\varepsilon)+1}^2}{V_n^2} - (s - \varepsilon) > \varepsilon,
\]

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 \( \xi_n = (S_{[nt]}, t \in [0, 1]) \). By (4) applied to the obvious bound

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

the convergence of f.d.d. of \( \xi_n \) follows from those of the process \( \Xi_n \).

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

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

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 \( \Xi_n \). The convergence of finite dimensional distributions of the process \( \xi_n \) is thus established.

To prove the tightness we shall use Theorem 8.3 from Billingsley (1968). Since \( \xi_n(0) = 0 \), the proof reduces in showing that for all \( \varepsilon, \eta > 0 \) there exist \( n_0 \geq 1 \) and \( \delta, 0 < \delta < 1 \), such that

\[
\frac{1}{\delta} P \left\{ \sup_{1 \leq k \leq n} V_n^{-1} |S_{k+i} - S_k| \geq \varepsilon \right\} \leq \eta, \quad n \geq n_0
\]

for all \( 1 \leq k \leq n \).

Let us introduce the truncated variables

\[
Y_i := \ell_n^{-1}(X_i; X_i^2 \leq \tau^2 b_n^2), \quad i = 1, \ldots, n
\]
with $\ell_n = n^{-1/2}b_n$ as above and $\tau$ 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

$$P\left\{ \sup_{1 \leq i \leq n} |S_{k+i} - S_k| \geq \varepsilon \right\} \leq A + B + C,$$

where

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

$$B := P\{ V_n < \sqrt{n}/2 \},$$

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

Due to (21) we can choose $n_1$ such that $\sqrt{n}|EY_1| \leq 1/4$ for $n \geq n_1$. Then with $n \geq n_1$ and $\delta \leq \varepsilon$ we have

$$A \leq P\left\{ \max_{1 \leq i \leq n} \left| \sum_{j=k+1}^{k+i} (Y_j - EY_j) \right| + n\delta |EY_1| \geq \varepsilon \sqrt{n}/2 \right\} \leq P\left\{ \max_{1 \leq i \leq n} \left| \sum_{j=k+1}^{k+i} (Y_j - EY_j) \right| \geq \sqrt{n}/4 \right\}.$$ 

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

$$P\left\{ n^{-1/2} \left| \sum_{j=k+1}^{k+n} (Y_j - EY_j) \right| \geq \varepsilon \right\} \leq \frac{8^p}{\varepsilon^p n^{p/2}} E \left( \sum_{j=k+1}^{k+n} (Y_j - EY_j)^p \right)^{1/p} \leq \frac{8^p}{\varepsilon^p n^{p/2}} \left[ (n\delta)^{p/2} (EY_1^2)^{p/2} + n\delta E|Y_1|^p \right].$$

By Ottaviani inequality we find

$$A \leq \frac{\delta n}{3},$$

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

Now by Ottaviani inequality we find

$$A \leq \frac{\delta n}{3},$$

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


Next we consider $\mathcal{B}$. Since $n^{-1} E \hat{V}^2_n = E Y^2_1$ we have by (41) $n^{-1} E \hat{V}^2_n \geq 3/4$, for $n \geq n_2$. Furthermore,

\[ B \leq P \{ n^{-1} | \hat{V}^2_n - E \hat{V}^2_n | \geq 1/2 \} \leq 4n^{-1} E Y^4_1 \leq 4 \varepsilon^2 E Y^2_1 \leq \delta \eta / 3, \quad (43) \]

provided $n \geq n_2$ and $\tau^2 \leq \delta \eta / 18$.

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

**Proof of Theorem 2.** Due to Theorem 1, it suffices to check that $\| V_n^{-1}(z_n - \zeta_n) \|_\infty$ goes to zero in probability, where $\| f \|_\infty := \sup_{0 \leq t \leq 1} | f(t) |$. To this end let us introduce the random change of time $\theta_n$ defined as follows. When $V_n > 0$, $\theta_n$ is the map from $[0, 1]$ onto $[0, 1]$ which interpolates linearly between the points $(k/n, V^2_k / V^2_n)$, $k = 0, 1, \ldots, n$. 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

\[ \zeta^e_n(\theta_n(t)) = \zeta^e_n(t), \quad 0 \leq t \leq 1. \quad (44) \]

Clearly for each $t \in [0, 1], \quad \left| \frac{V^2_{[t]}}{V^2_n} - \theta_n(t) \right| \leq \max_{1 \leq k \leq n} \left| \frac{X^2_k}{V^2_n} \right. \]

It follows by (4) that

\[ \sup_{0 \leq t \leq 1} \left| \frac{V^2_{[t]}}{V^2_n} - \theta_n(t) \right|^p \to 0 \]

and this together with Lemma 9 gives

\[ \| \theta_n - I \|_\infty \to 0. \quad (45) \]

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

\[ \| \zeta^e_n - \zeta^e_n \|_\infty \leq \sup_{0 \leq t \leq 1} | \zeta^e_n(\theta_n(t)) - \zeta^e_n(\theta(t)) | \leq \omega(\zeta^e_n; \| \theta_n - I \|_\infty). \]

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

\[ P(\| \zeta^e_n - \zeta^e_n \|_\infty \geq \lambda) \leq P(\| \theta_n - I \|_\infty \geq \delta) + P(\omega(\zeta^e_n; \delta) \geq \lambda). \quad (46) \]

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

\[ \lim_{n \to \infty} \sup_{\lambda > 0} P(\omega(\zeta^e_n; \delta) \geq \lambda) \leq P(\omega(W; \delta) \geq \lambda). \]

Hence for $n \geq n_1$ we have $P(\omega(\zeta^e_n; \delta) \geq \lambda) < 2\varepsilon$. Having in mind (45) and (46) we see that the proof is complete. \( \square \)

**Proof of Theorem 3.** The convergence of finite dimensional distributions is already established in the proof of Theorem 2.
It remains to prove tightness of $\zeta_n^{se}$ in the space $H_p[0,1]$. To this aim, we have to check the second condition of Proposition 7 only.

Let $\xi_1, \ldots, \xi_n, \ldots$ be an independent Rademacher sequence which is independent on $(X_i)$. By symmetry of $X_1$, both sequences $(X_i)$ and $(\xi_i X_i)$ have the same distribution.

Noting also that $\xi_i^2 = 1$ a.s., we have that $\zeta_n^{se}$ has the same distribution as the random process $\tilde{\zeta}_n$ which is defined linearly between the points

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

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

$$\limsup_{j \to \infty} \sup_{n \geq 2} \max_{0 \leq k < 2^j} P(\{|\zeta_n^{se}(t + h) - \zeta_n^{se}(t)| > \epsilon p(2^{-j})\}) = 0. \quad (47)$$

To this aim we shall estimate

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

uniformly in $n$. First consider the case, where

$$0 \leq \frac{V^2_{k-1}}{V_n^2} \leq t < h \leq \frac{V^2_k}{V_n^2}.$$

so

$$0 \leq h \leq \frac{V^2_k}{V_n^2} - \frac{V^2_{k-1}}{V_n^2} = \frac{X^2_k}{V_n^2}.$$

We have then by linear interpolation

$$|\zeta_n^{se}(t + h) - \zeta_n^{se}(t)| = \frac{|\xi_k X_k|}{V_n} \frac{V^2_k}{X^2_k} h = \left(\frac{V_n}{|X_k|}\right)^{1/2} \leq \sqrt{h}. \quad (48)$$

Next consider the following configuration:

$$0 \leq \frac{V^2_{k-1}}{V_n^2} \leq t < \frac{V^2_k}{V_n^2} \leq \frac{V^2_{l+1}}{V_n^2} \leq t + h < \frac{V^2_{l+1}}{V_n^2}.$$

Then we have

$$|\zeta_n^{se}(t + h) - \zeta_n^{se}(t)| \leq \delta_1 + \delta_2 + \delta_3,$$

where

$$\delta_1 := |\zeta_n^{se}(t + h) - \zeta_n^{se}(V^2_k/V_n^2)| \leq \sqrt{t + h - V^2_l/V_n^2} \leq \sqrt{h},$$

$$\delta_2 := |\zeta_n^{se}(V^2_l/V_n^2) - \zeta_n^{se}(V^2_{l+1}/V_n^2)| = V^{-1}_n|U_l - U_k| \leq \frac{|U_l - U_k|}{\sqrt{V^2_l - V^2_{l+1}}} \sqrt{h},$$

$$\delta_3 := |\zeta_n^{se}(V^2_{l+1}/V_n^2) - \zeta_n^{se}(t)| \leq \sqrt{V^2_{l+1}/V_n^2 - t} \leq \sqrt{h}.$$

Hence, for any configuration we obtain

$$|\zeta_n^{se}(t + h) - \zeta_n^{se}(t)| \leq \frac{|U_l - U_k|}{\sqrt{V^2_l - V^2_{l+1}}} \sqrt{h} + 2\sqrt{h}, \quad (49)$$
if we agree that \(|U_l - U_k| (V_l^2 - V_k^2)^{-1/2} = 0\) when \(k = l\). Therefore,
\[
\delta(t, h, r) \leq P\left(\frac{|U_l - U_k|}{\sqrt{V_l^2 - V_k^2}} > \frac{r}{(2\sqrt{h})}\right),
\]
(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 \((\epsilon_i)\).

Thus conditioning on \(X_1, \ldots, X_n\) and applying the well known Hoeffding’s inequality we obtain
\[
\delta(t, h, r) \leq c \exp\left\{ -\frac{r^2}{(8h)} \right\}.
\]
(51)
Now (47) clearly follows if for every \(\epsilon > 0\),
\[
\sum_{j=1}^{\infty} 2^j \exp\{-\epsilon 2^j (2^{-j})\} < \infty,
\]
(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 \(\text{DAN}_n\), \(X_1\) 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
\[
\lim_{J \to \infty} \limsup_{n \to \infty} P\left(\left\|\xi_n - E_n \xi_n\right\|_{\rho p} > 4\epsilon\right) = 0.
\]
(53)
To this end, it suffices to prove that with some sequence \(J_n \uparrow \infty\) to be precised later,
\[
\limsup_{n \to \infty} \left( \sup_{J_n < J < \infty} \max_{0 < k < 2^J} \frac{1}{\rho(2^{-J})} |\lambda_{j,k}^{(\epsilon_n)}(\psi_n)| > \epsilon \right) = 0
\]
(54)
and
\[
\limsup_{n \to \infty} \left( \sup_{J_n < J < \infty} \max_{0 < k < 2^J} \frac{1}{\rho(2^{-J})} |\lambda_{j,k}^{(\epsilon_n)}(\psi_n)| > 3\epsilon \right) = 0,
\]
(55)
where
\[
\lambda_{j,k}^{(\epsilon_n)}(\psi_n) := \epsilon_n (k + 1)2^{-j} - \epsilon_n (k2^{-j}), \quad 0 \leq k < 2^J.
\]
To start with (54), following the same steps which led to (49) we obtain with \(k, l\) such that
\[
\frac{V_{k-1}^2}{V_n^2} < t \leq \frac{V_k^2}{V_n^2}, \quad \frac{V_{l-1}^2}{V_n^2} < t + h \leq \frac{V_l^2}{V_n^2},
\]
the upper bound
\[
|\psi_n(t + h) - \psi_n(t)| \leq \left( 2 + \frac{|S_{(i,j)}|}{V_{(i,j)}} \right) \sqrt{n},
\]
where we use the notations
\[
S_{(i,j)} := \sum_{i < k \leq j} X_k, \quad V_{(i,j)} := \left( \sum_{i < k \leq j} X_k^2 \right)^{1/2}
\]
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

$$
|\zeta_n(t + h) - \zeta_n(t)| \leq \sqrt{h} \max_{1 \leq k \leq l \leq n} T_{k,l}.
$$

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 \leq i \leq N) are subgaussian, then so is $\max_{1 \leq i \leq N} |Y_i|$, which more precisely satisfies

$$
\left\| \max_{1 \leq i \leq N} |Y_i| \right\|_{\phi_2} \leq a (\log N)^{1.2} \max_{1 \leq i \leq N} ||Y_i||_{\phi_2},
$$

where $a$ is an absolute constant and $\| \cdot \|_{\phi_2}$ 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 occurrence)

$$
\begin{align*}
\Pr \left( \sup_{j > J_n} \max_{0 \leq k < 2^j} {1 \over \rho(2^{-j})} |J_n\zeta_n^{(c)}(t)| > \epsilon \right) & \leq \sum_{j > J_n} \Pr \left( \max_{1 \leq k \leq l \leq n} T_{k,l} > c\epsilon 2^j \right) \\
& \leq \sum_{j > J_n} C \exp \left( -c \epsilon 2^j \log n \right),
\end{align*}
$$

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

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

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

To prove (55), we start with

$$
\begin{align*}
\Pr \left( \sup_{j \in [0,1]} {V_{2^n}^{(j)}} - {V_{2^n}^{[j]}} \leq \delta_n \right) & \leq P_1 + P_2 + P_3
\end{align*}
$$

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

$$
A_n = \left\{ \sup_{t \in [0,1]} \left| {V_{2^n}^{(j)}} - {V_{2^n}^{[j]}} \right| \leq \delta_n \right\} \cap \left\{ \sup_{t \in [0,1]} \left| {V_{2^n}^{[j]}} - t \right| \leq \delta_n \right\}.
$$

where $\delta_n$ is chosen as in (11), keeping the freedom of choice of the constant $c$. Now we define

$$
\begin{align*}
P_1 & := \Pr(A_n^c), \\
P_2 & := \Pr \left( A_n \cap \left\{ \max_{j \in J_n} \max_{0 \leq k < 2^j} {1 \over \rho(2^{-j})} \left| S_{(k+1)2^{-j} - n} - S_{(k-2^{-j} - n)} \right| > \epsilon \right\} \right), \\
P_3 & := \Pr \left( A_n \cap \left\{ \max_{j \in J_n} \max_{0 \leq k < 2^j} {1 \over \rho(2^{-j})} \left| \left| S_{j - |k2^{-j} - n|} \right| \leq n \delta_n \right| V_n \right| V_n \right) > 2\epsilon \right) \right\}.
\end{align*}
$$
The following easy estimates
\[
\sup_{t \in [0,1]} \left| \frac{V^2_{\lfloor nt \rfloor}}{V^2_n} - t \right| \leq \max_{1 \leq k \leq n} \left| \frac{V^2_k}{V^2_n} - \frac{k}{n} \right| + \frac{1}{n},
\]
\[
\sup_{t \in [0,1]} \left| \frac{V^2_{\lfloor nt \rfloor}}{V^2_n} - \frac{V^2_{\lfloor nt \rfloor}}{V^2_n} \right| \leq \max_{1 \leq k \leq n} \frac{X^2_k}{V^2_n} + \max_{1 \leq k \leq n} \left| \frac{V^2_k}{V^2_n} - \frac{k}{n} \right| + \frac{1}{n},
\]
lead by (9) and (10) to
\[
P(A_n^c) \to 0.
\] (60)

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

To control \( P_2 \), first write with self-explanatory notations
\[
\left| S_{[(k+1)2^{-j}-n]} - S_{[k2^{-j}-n]} \right| = \left| S_{[(k+1)2^{-j}-n]} - S_{[k2^{-j}-n]} \right| \times \frac{V_{[(k+1)2^{-j}-n],[k2^{-j}-n]]}}{V_n}.
\]

Observing that on the event \( A_n \), we have
\[
\frac{V_{[(k+1)2^{-j}-n],[k2^{-j}-n]]}}{V_n} \leq \sqrt{2^{-j} + \delta_n}
\]
and assuming that
\[
\delta_n \leq 2^{-J},
\] (61)
we get
\[
P_2 \leq \sum_{J \leq j < J_n} P \left( \max_{0 \leq k < 2^j} \frac{1}{\rho(2^{-j})} \left| \frac{S_{[(k+1)2^{-j}-n]} - S_{[k2^{-j}-n]} \right| V_{[(k+1)2^{-j}-n],[k2^{-j}-n]]} > \sqrt{2^{-j} + \delta_n} \right).
\]

Since we are dealing now with the maximum of \( 2^j \) uniformly subgaussian random variables (their \( \phi_2 \) norms are bounded by a constant which depends only on the distribution of \( X_1 \)), this leads to
\[
P_2 \leq \sum_{J \leq j < J_n} C \exp(-c2^{2j-1}) \leq \sum_{j=J}^{\infty} C \exp(-c2^{2j-1}).
\] (62)

To control \( P_3 \), we first get rid of the residual term by noting that
\[
\frac{2}{\rho(2^{-j})2^{j/2}} = \frac{c}{j^\beta} < \epsilon \quad \text{for} \quad J \geq J(\epsilon),
\]
uniformly in \( n \). So for \( J \geq J(\epsilon) \),
\[
P_3 \leq P \left( A_n \cap \left\{ \max_{J \leq j < J_n} \max_{0 \leq k < 2^j} \frac{1}{\rho(2^{-j}) \max \left| \frac{S_l - S_{[k2^{-j}-n]} \right| V_n} > \epsilon \right\} \right).
\]

On the event \( A_n \) we have for any \( l \) such that \( |l - [k2^{-j}-n]| \leq n\delta_n \),
\[
\frac{|V^2_{[k2^{-j}-n]} - V_{[l]}^2|}{V^2_n} \leq 2\delta_n.
\]

It follows that
\[
P_3 \leq P \left( \max_{J \leq j < J_n} \max_{0 \leq k < 2^j, |l - [k2^{-j}-n]| \leq n\delta_n} \left| \frac{S_l - S_{[k2^{-j}-n]} \right| V^2_{[k2^{-j}-n]} - V_{[l]}^2 > \epsilon \rho(2^{-j}) / \sqrt{2\delta_n} \right).
\]
Using the invariance of distributions under translations on $k_w$, we get
\[ P_3 \leq \sum_{J \leq j < J_0} 2^{j} P\left( \max_{0 < l < [2n \delta \alpha]} \frac{S_l}{V_l} > \frac{c \rho(2^{-j})}{\sqrt{2 \delta_n}} \right) \]
\[ \leq \sum_{J \leq j < J_0} 2^{j} C \exp\left( -\frac{c 2^{-j} \delta \beta}{\delta_n \log n} \right) \]
\[ \leq C \sum_{J \leq j < J_0} 2^{j} \exp\left( -\frac{c \delta \beta}{\delta_n \log n} \frac{2^{-j}}{J} \right) . \]

Now we see that the following convergence rate (stronger than (61))
\[ \delta_n = \frac{1}{2^\gamma \log n} = \frac{2^{-\gamma \log n}}{\log n}, \quad \text{with} \quad \frac{1}{2^\gamma} < \gamma < 1, \]
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 $\delta_n$, we have $n^{-c} = o(\delta_n)$ for any $\varepsilon > 0$. □

**Lemma 13.** If $E|X_1|^{2+\delta} < \infty$ for some $\delta > 0$, then almost surely
\[ n^{-c} \max_{1 \leq k \leq n} \left| \frac{V_k^2}{V_n^2} - \frac{k}{n} \right| \to 0, \quad (63) \]
where $c = \delta/(2 + 2\delta)$.

**Proof.** By Marcinkiewicz SLLN, if the i.i.d. sequence $(Y_k)$ satisfies $E|Y_1|^p < \infty$ for some $1 \leq p < 2$, then $n^{-1/p} \sum_{k=1}^n Y_k - nEY_1$ goes to 0 almost surely. Applying this to $Y_1 = X_1^2$ and $p = 1 + \delta/2$ gives
\[ \frac{V_k^2}{V_n^2} = 1 + n^{1/p-1} e_n, \quad n \geq 1, \]
where the random sequence $(e_n)$ goes to zero almost surely. Since we assume $P(X_1 = 0) < 1$, we have $P(\forall n \geq 1; V_n = 0) = 0$. On each event $\{V_n^2 > 0\}$, we may write with $a = 1 - 1/p$,
\[ \frac{V_k^2}{V_n^2} - \frac{k}{n} = \frac{k}{n} \left( \frac{V_k^2}{V_n^2} - 1 \right) = \frac{k}{n} \times \frac{k^{-a} e_k - n^{-a} e_n}{1 + n^{-a} e_n}. \]

For each $n \geq n_0 = n_0(\alpha)$ large enough, $n^{-a} e_n > -1/2$. Now for an exponent $0 < b < 1$ to be precised later, we have
\[ \left| \frac{V_k^2}{V_n^2} - \frac{k}{n} \right| \leq 4n^{b-1} \sup_{i \geq 1} |e_i| \quad \text{for} \quad n \geq n_0, 1 \leq k \leq n^b \]
and
\[ \left| \frac{V_k^2}{V_n^2} - \frac{k}{n} \right| \leq 4n^{-ab} \sup_{i \geq n^b} |e_i| \quad \text{for} \quad n \geq n_0, n^b < k \leq n. \]

The optimal choice of $b$ given by $1 - b = ab$ leads to the announced conclusion with
\[ c = a/(a + 1) = \delta/(2 + 2\delta). \]
Lemma 14. If $E|X_1|^{2+\delta} < \infty$ for some $\delta > 0$, then almost surely
\begin{equation}
\eta^d \max_{1 \leq k \leq n} \frac{X_k^2}{V_n^2} \rightarrow 0 \tag{64}
\end{equation}
for any $d < \delta/(2 + \delta)$.

Proof. We use the same trick as in O’Brien (1980, p. 542). For any positive $\epsilon$ we have (noting the key role of i.o. in the following inequalities)
\begin{align*}
P \left( \max_{1 \leq k \leq n} \frac{X_k^2}{V_n^2} > \epsilon n^{-d}, \text{ i.o.} \right) &\leq P \left( V_n^2 < \frac{n}{2}, \text{ i.o.} \right) + P \left( \max_{1 \leq k \leq n} \frac{X_k^2}{V_n^2} > \frac{n}{2} \epsilon n^{-d}, \text{ i.o.} \right) \\
&= 0 + P \left( X_n^2 > \frac{n}{2} \epsilon n^{-d}, \text{ i.o.} \right).
\end{align*}

Now observe that
\begin{equation*}
\sum_{n=1}^{\infty} P \left( X_n^2 > \frac{n}{2} \epsilon n^{-d} \right) \leq \left( \frac{2}{\epsilon} \right)^{1+d/2} E|X_1|^{2+\delta} \sum_{n=1}^{\infty} \frac{1}{n^{1-d(1+\delta/2)}}.
\end{equation*}

For any $d$ such that $(1 - d)(1 + \delta/2) > 1$, Borel–Cantelli’s Lemma leads to
\begin{equation*}
P \left( \max_{1 \leq k \leq n} \frac{X_k^2}{V_n^2} > \epsilon n^{-d}, \text{ i.o.} \right) = 0.
\end{equation*}

As $\epsilon$ is arbitrary, the result is proved. \qed

Uncited Reference

Račauskas 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


