QUANTITATIVE BOUNDS IN THE CENTRAL LIMIT THEOREM FOR m-DEPENDENT RANDOM VARIABLES

SVANTE JANSON, LUCA PRATELLI, AND PIETRO RIGO

ABSTRACT. For each $n \geq 1$, let $X_{n,1}, \ldots, X_{n,N_n}$ be real random variables and $S_n = \sum_{i=1}^{N_n} X_{n,i}$. Let $m_n \geq 1$ be an integer. Suppose $(X_{n,1}, \ldots, X_{n,N_n})$ is m_n-dependent, $E(X_{ni}) = 0$, $E(X_{ni}^2) < \infty$ and $\sigma_n^2 := E(S_n^2) > 0$ for all n and i. Then,

$$d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) \leq 30\,\big\{c^{1/3} + 12\,U_n(c/2)^{1/2}\big\} \qquad \text{for all } n \geq 1 \text{ and } c > 0,$$

where d_W is Wasserstein distance, Z a standard normal random variable and

$$U_n(c) = \frac{m_n}{\sigma_n^2} \sum_{i=1}^{N_n} E\Big[X_{n,i}^2\, 1\big\{|X_{n,i}| > c\,\sigma_n/m_n\big\}\Big].$$

Among other things, this estimate of $d_W\big(S_n/\sigma_n,\, Z\big)$ yields a similar estimate of $d_{TV}\big(S_n/\sigma_n,\, Z\big)$ where d_{TV} is total variation distance.

1. Introduction

Central limit theorems (CLTs) for m-dependent random variables have a long history. Pioneering results, for a fixed m, were given by Hoeffding and Robbins [11] and Diananda [7] (for m-dependent sequences), and Orey [12] (more generally, and also for triangular arrays). These results were then extended to the case of increasing $m = m_n$; see e.g. Bergström [1], Berk [2], Rio [15], Romano and Wolf [17], and Utev [18], [19].

Obviously, CLTs for m-dependent random variables are often corollaries of more general results obtained under mixing conditions. A number of CLTs under mixing conditions are actually available. Without any claim of being exhaustive, we mention [3], [6], [13], [15], [18], [19] and references therein. However, mixing conditions are not directly related to our purposes (as stated below) and they will not be discussed further.

This paper deals with an (m_n)-dependent array of random variables, where (m_n) is any sequence of integers, and provides an upper bound for the Wasserstein distance between the standard normal law and the distribution of the normalized partial sums. A related bound for the total variation distance is obtained as well. To be more precise, we need some notation.

Date: 11 August, 2022.

2020 *Mathematics Subject Classification.* 60F05.

Key words and phrases. Central limit theorem; Lindeberg condition; m-dependent random variables; Quantitative bound; Total variation distance; Wasserstein distance.

For each $n \geq 1$, let $1 \leq m_n \leq N_n$ be integers, $(X_{n,1}, \ldots, X_{n,N_n})$ a collection of real random variables, and

$$S_n = \sum_{i=1}^{N_n} X_{n,i}.$$

Suppose

$$(X_{n,1}, \ldots, X_{n,N_n}) \text{ is } m_n\text{-dependent for every } n, \tag{1}$$

$$E(X_{ni}) = 0, \quad E(X_{ni}^2) < \infty, \quad \sigma_n^2 := E(S_n^2) > 0 \quad \text{for all } n \text{ and } i, \tag{2}$$

and define

$$U_n(c) = \frac{m_n}{\sigma_n^2} \sum_{i=1}^{N_n} E\Big[X_{n,i}^2\, 1\big\{|X_{n,i}| > c\,\sigma_n/m_n\big\}\Big] \qquad \text{for all } c > 0.$$

Our main result (Theorem 4) is that

$$d_W\Big(\frac{S_n}{\sigma_n}, Z\Big) \leq 30\,\big\{c^{1/3} + 12\,U_n(c/2)^{1/2}\big\} \qquad \text{for all } n \geq 1 \text{ and } c > 0, \tag{3}$$

where d_W is Wasserstein distance and Z a standard normal random variable.

Inequality (3) provides a quantitative estimate of $d_W\big(S_n/\sigma_n, Z\big)$. The connections between (3) and other analogous results are discussed in Remark 11 and Section 4. To our knowledge, however, no similar estimate of $d_W\big(S_n/\sigma_n, Z\big)$ is available under conditions (1)–(2) only. In addition, inequality (3) implies the following useful result:

Theorem 1 (Utev [18, 19]). *$S_n/\sigma_n \overset{dist}{\longrightarrow} Z$ provided conditions (1)–(2) hold and $U_n(c) \to 0$ for every $c > 0$.*

Based on inequality (3), we also obtain quantitative bounds for $d_K\big(S_n/\sigma_n, Z\big)$ and $d_{TV}\big(S_n/\sigma_n, Z\big)$, where d_K and d_{TV} are Kolmogorov distance and total variation distance, respectively. As to d_K, it suffices to recall that

$$d_K\Big(\frac{S_n}{\sigma_n}, Z\Big) \leq 2\sqrt{d_W\Big(\frac{S_n}{\sigma_n}, Z\Big)};$$

see Lemma 2. To estimate d_{TV}, define

$$l_n = 2\int_0^\infty t\,|\phi_n(t)|\,dt$$

where ϕ_n is the characteristic function of S_n/σ_n. By a result in [14] (see Theorem 3 below),

$$d_{TV}\Big(\frac{S_n}{\sigma_n}, Z\Big) \leq 2\,d_W\Big(\frac{S_n}{\sigma_n}, Z\Big)^{1/2} + l_n^{2/3}\,d_W\Big(\frac{S_n}{\sigma_n}, Z\Big)^{1/3}.$$

Hence, $d_{TV}\big(S_n/\sigma_n, Z\big)$ can be upper bounded via inequality (3). For instance, in addition to (1)–(2), suppose $X_{ni} \in L_\infty$ for all n and i and define

$$c_n = \frac{2\,m_n}{\sigma_n}\,\max_i \|X_{ni}\|_\infty.$$

On noting that $U_n(c_n/2) = 0$, one obtains

$$d_{TV}\Big(\frac{S_n}{\sigma_n}, Z\Big) \leq \sqrt{120}\;c_n^{1/6} + 30^{1/3}\,l_n^{2/3}\,c_n^{1/9}.$$

The rest of this paper is organized into three sections. Section 2 just recalls some definitions and known results, Section 3 is devoted to proving inequality (3), while Section 4 investigates $d_{TV}\big(S_n/\sigma_n,\, Z\big)$ and the convergence rate provided by (3).

The numerical constants in our results are obviously not best possible; we have not tried to optimize them. More important are the powers, $c^{1/3}$ and $U_n(c/2)^{1/2}$ in (3) and similar powers in other results; we do not believe that these are optimal. This is discussed in Section 4. How far (3) can be improved, however, is essentially an open problem.

2. Preliminaries

The same notation as in Section 1 is adopted in the sequel. It is implicitly assumed that, for each $n \geq 1$, the variables $(X_{ni} : 1 \leq i \leq N_n)$ are defined on the same probability space (which may depend on n).

Let $k \geq 0$ be an integer. A (finite or infinite) sequence (Y_i) of random variables is k-dependent if $(Y_i : i \leq j)$ is independent of $(Y_i : i > j + k)$ for every j. In particular, 0-dependent is the same as independent. Given a sequence (k_n) of non-negative integers, an array $(Y_{ni} : n \geq 1,\ 1 \leq i \leq N_n)$ is said to be (k_n)-dependent if $(Y_{ni} : 1 \leq i \leq N_n)$ is k_n-dependent for every n.

Let X and Y be real random variables. Three well known distances between their probability distributions are Wasserstein's, Kologorov's and total variation. Kolmogorov distance and total variation distance are, respectively,

$$d_K(X,Y) = \sup_{t\in\mathbb{R}} |P(X \leq t) - P(Y \leq t)| \quad \text{and}$$
$$d_{TV}(X,Y) = \sup_{A\in\mathcal{B}(\mathbb{R})} |P(X \in A) - P(Y \in A)|.$$

Under the assumption $E|X| + E|Y| < \infty$, Wasserstein distance is

$$d_W(X,Y) = \inf_{U\sim X,V\sim Y} E|U - V|$$

where inf is over the real random variables U and V, defined on the same probability space, such that $U \sim X$ and $V \sim Y$. Equivalently,

$$d_W(X,Y) = \int_{-\infty}^{\infty} |P(X \leq t) - P(Y \leq t)|\, dt = \sup_f |Ef(X) - Ef(Y)|$$

where sup is over the 1-Lipschitz functions $f : \mathbb{R} \to \mathbb{R}$. The next lemma is certainly known, but we give a proof since we do not know of any reference for the first claims.

Lemma 2. *Suppose $EX^2 \leq 1$, $EY^2 \leq 1$ and $EY = 0$. Then,*

$$d_W(X,Y) \leq \sqrt{2},$$
$$d_W(X,Y) \leq 4\sqrt{d_K(X,Y)}.$$

If $Y \sim N(0,1)$, we also have

$$d_K(X,Y) \leq 2\sqrt{d_W(X,Y)}.$$

Proof. Take U independent of V with $U \sim X$ and $V \sim Y$. Then,

$$d_W(X,Y) \leq E|U-V| \leq \big\{E\big((U-V)^2\big)\big\}^{1/2} \leq \sqrt{2}.$$

Moreover, for each $c > 0$,

$$\begin{aligned} d_W(X,Y) &= \int_{-\infty}^{\infty} |P(X \leq t) - P(Y \leq t)|\, dt \\ &\leq 2\,c\,d_K(X,Y) + \int_c^{\infty} |P(X > t) - P(Y > t)|\, dt \\ &+ \int_c^{\infty} |P(-X > t) - P(-Y > t)|\, dt \\ &\leq 2\,c\,d_K(X,Y) + \int_c^{\infty} \big\{P(|X| > t) + P(|Y| > t)\big\}\, dt \\ &\leq 2\,c\,d_K(X,Y) + \int_c^{\infty} \frac{2}{t^2}\, dt = 2\,c\,d_K(X,Y) + \frac{2}{c}. \end{aligned}$$

Hence, letting $c = d_K(X,Y)^{-1/2}$, one obtains $d_W(X,Y) \leq 4\sqrt{d_K(X,Y)}$.

Finally, if $Y \sim N(0,1)$, it is well known that $d_K(X,Y) \leq 2\sqrt{d_W(X,Y)}$; see e.g. [4, Theorem 3.3]. □

Finally, under some conditions, d_{TV} can be estimated through d_W. We report a result which allows this; in our setting we simply take $V = 1$ below.

Theorem 3 (A version of [14, Theorem 1])**.** *Let X_n, V, Z be real random variables, and suppose that $Z \sim N(0,1)$, $V > 0$, $EV^2 = EX_n^2 = 1$ for all n, and V is independent of Z. Let ϕ_n be the characteristic function of X_n, and*

$$l_n = 2\int_0^{\infty} t\,|\phi_n(t)|\, dt.$$

Then,

$$d_{TV}(X_n,\, VZ) \leq \big\{1 + E(1/V)\big\}\, d_W(X_n,\, VZ)^{1/2} + l_n^{2/3}\, d_W(X_n,\, VZ)^{1/3}$$

for each n.

Proof. This is essentially a special case of [14, Theorem 1], with $\beta = 2$ and the constant k made explicit. Also, the assumption $d_W(X_n,\, VZ) \to 0$ in [14, Theorem 1] is not needed; we use instead $d_W(X_n,\, VZ) \leq \sqrt{2}$ from Lemma 2. Using this and $EX_n^2 = 1$, the various constants appearing in the proof can be explicitly evaluated. In fact, improving the argument in [14] slightly by using $P(|X_n| > t) \leq EX_n^2/t^2 = t^{-2}$, and as just said using $d_W(X_n,\, VZ) \leq \sqrt{2}$, we can take $k^* = 5 + 4\sqrt{2}$ in the proof. After simple calculations, this implies that the constant k in [14] can be taken as

$$k = \frac{1}{2} \cdot \frac{3}{2} \cdot 2^{1/3}(5 + 4\sqrt{2})^{1/3}\pi^{-2/3} < 1.$$

□

3. An upper bound for Wasserstein distance

As noted in Section 1, our main result is:

Theorem 4. *Under conditions* (1)–(2),

$$d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big)\le 30\,\big\{c^{1/3}+12\,U_n(c/2)^{1/2}\big\}$$

for all $n\ge 1$ and $c>0$, where Z denotes a standard normal random variable.

In turn, Theorem 4 follows from the following result, which is a sharper version of the special case $m_n=1$.

Theorem 5. *Let $X_1,\ldots,X_N$ be real random variables and $S=\sum_{i=1}^N X_i$. Suppose $(X_1,\ldots,X_N)$ is* 1*-dependent and*

$$E(X_i)=0,\;\;E(X_i^2)<\infty \text{ for all } i \text{ and } \sigma^2:=E(S^2)>0.$$

Then,

$$d_W\Big(\frac{S}{\sigma},\,Z\Big)\le 30\,\big\{c^{1/3}+6\,L(c)^{1/2}\big\}\qquad \text{for all } c>0,$$

where Z is a standard normal random variable and

$$L(c)=\frac{1}{\sigma^2}\sum_{i=1}^N E\Big[X_i^2\,1\big\{|X_i|>c\,\sigma\big\}\Big].$$

To deduce Theorem 4 from Theorem 5, define $M_n=\lceil N_n/m_n\rceil$, $X_{n,i}=0$ for $i>N_n$, and

$$Y_{n,i}=\sum_{j=(i-1)m_n+1}^{im_n}X_{n,j}\qquad \text{for } i=1,\ldots,M_n.$$

Since $(Y_{n,1},\ldots,Y_{n,M_n})$ is 1-dependent and $\sum_i Y_{n,i}=\sum_i X_{n,i}=S_n$, Theorem 5 implies

$$d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big)\le 30\,\big\{c^{1/3}+6\,L_n(c)^{1/2}\big\}\tag{4}$$

where

$$L_n(c)=\frac{1}{\sigma_n^2}\sum_{i=1}^{M_n}E\Big[Y_{n,i}^2\,1\big\{|Y_{n,i}|>c\,\sigma_n\big\}\Big].$$

Therefore, to obtain Theorem 4, it suffices to note the following inequality:

Lemma 6. *With notations as above, for every $c>0$,*

$$L_n(2c)\le 4\,U_n(c).$$

In the rest of this section, we prove Lemma 6 and Theorem 5. We also obtain a (very small) improvement of Utev's Theorem 1.

3.1. **Proof of Lemma 6 and Utev's theorem.**

Proof of Lemma 6. Fix $c > 0$ and define

$$V_{n,i} = \sum_{j=(i-1)m_n+1}^{im_n} X_{n,j}\, 1\{|X_{n,j}| > c\,\sigma_n/m_n\}.$$

Since $|Y_{n,i}| \le |V_{n,i}| + c\,\sigma_n$, one obtains

$$|Y_{n,i}|\, 1\{|Y_{n,i}| > 2\,c\,\sigma_n\} \le \big(|V_{n,i}| + c\,\sigma_n\big)\, 1\{|V_{n,i}| > c\,\sigma_n\} \le 2\,|V_{n,i}|.$$

Therefore,

$$\begin{aligned}
\sigma_n^2\, L_n(2c) &= \sum_{i=1}^{M_n} E\big[Y_{n,i}^2\, 1\{|Y_{n,i}| > 2\,c\,\sigma_n\}\big] \le 4 \sum_{i=1}^{M_n} E(V_{n,i}^2) \\
&\le 4\, m_n \sum_{i=1}^{M_n} \sum_{j=(i-1)m_n+1}^{im_n} E\big[X_{n,j}^2\, 1\{|X_{n,j}| > c\,\sigma_n/m_n\}\big] \\
&= 4\, m_n \sum_{i=1}^{N_n} E\big[X_{n,i}^2\, 1\{|X_{n,i}| > c\,\sigma_n/m_n\}\big] = 4\sigma_n^2\, U_n(c).
\end{aligned}$$

□

We also note that, because of (4), Theorem 5 implies:

Corollary 7. $S_n/\sigma_n \xrightarrow{dist} Z$ *if conditions* (1)–(2) *hold and* $L_n(c) \to 0$ *for every* $c > 0$.

Corollary 7 slightly improves Theorem 1. In fact, $U_n(c) \to 0$ for all $c > 0$ implies $L_n(c) \to 0$ for all $c > 0$, because of Lemma 6, but the converse is not true.

Example 8. ($L_n(c) \to 0$ does not imply $U_n(c) \to 0$). Let $(V_n : n \ge 1)$ be an i.i.d. sequence of real random variables such that V_1 is absolutely continuous with density $f(x) = (3/2)\, x^{-4}\, 1_{[1,\infty)}(|x|)$. Let m_n and t_n be positive integers such that $m_n \to \infty$. Define $N_n = m_n\,(t_n + 1)$ and

$$X_{n,i} = V_i \;\text{ if } 1 \le i \le m_n t_n \quad \text{and} \quad X_{n,i} = V_{m_n t_n + 1} \;\text{ if } m_n t_n < i \le m_n(t_n+1).$$

Define also

$$T_n = \frac{\sum_{j=1}^{m_n} V_j}{\sqrt{m_n}}.$$

Then, $EV_1^2 = 3$, $\sigma_n^2 = 3\,(m_n t_n + m_n^2)$ and

$$\begin{aligned}
L_n(c) &= \frac{1}{\sigma_n^2} \sum_{i=1}^{M_n} E\big[Y_{n,i}^2\, 1\{|Y_{n,i}| > c\,\sigma_n\}\big] \le \frac{1}{\sigma_n^2} \sum_{i=1}^{t_n} E\big[Y_{n,i}^2\, 1\{|Y_{n,i}| > c\,\sigma_n\}\big] + \frac{3\,m_n^2}{\sigma_n^2} \\
&= \frac{m_n t_n}{\sigma_n^2}\, E\big[T_n^2\, 1\{|T_n| > c\,\sigma_n/\sqrt{m_n}\}\big] + \frac{3\,m_n^2}{\sigma_n^2}.
\end{aligned}$$

If $m_n = \mathrm{o}(t_n)$, then $m_n^2/\sigma_n^2 \to 0$, $m_n t_n/\sigma_n^2 \to 1/3$ and $\sigma_n/\sqrt{m_n} \to \infty$. Moreover, the sequence (T_n^2) is uniformly integrable (since $T_n \overset{dist}{\longrightarrow} N(0,3)$ with (trivial) convergence of second moments). Hence, if $m_n = \mathrm{o}(t_n)$, one obtains, for every $c > 0$,

$$\limsup_n L_n(c) \le \frac{1}{3} \limsup_n E\big[T_n^2\, 1\{|T_n| > c\,\sigma_n/\sqrt{m_n}\}\big] = 0.$$

However,

$$\begin{aligned} U_n(c) &= \frac{m_n}{\sigma_n^2} \sum_{i=1}^{N_n} E\big[X_{n,i}^2\, 1\{|X_{n,i}| > c\,\sigma_n/m_n\}\big] = \frac{m_n N_n}{\sigma_n^2}\, E\big[V_1^2\, 1\{|V_1| > c\,\sigma_n/m_n\}\big] \\ &= \frac{3\, m_n N_n}{\sigma_n^2} \int_{c\,\sigma_n/m_n}^{\infty} x^{-2} dx = \frac{3 N_n}{c\,\sigma_n^2}\, \frac{m_n^2}{\sigma_n} \ge \frac{3 t_n m_n^3}{c(6 m_n t_n)^{3/2}} \end{aligned}$$

for each n such that $c\,\sigma_n/m_n \ge 1$ and $m_n \le t_n$. Therefore, $L_n(c) \to 0$ and $U_n(c) \to \infty$ for all $c > 0$ whenever $m_n = \mathrm{o}(t_n)$ and $t_n = \mathrm{o}(m_n^3)$. This happens, for instance, if $m_n \to \infty$ and $t_n = m_n^2$.

3.2. **Proof of Theorem 5.** Our proof of Theorem 5 requires three lemmas. A result by Röllin [16] plays a crucial role in one of them (Lemma 10).

In this subsection, $X_1, \ldots, X_N$ are real random variables and $S = \sum_{i=1}^N X_i$. We assume that $(X_1, \ldots, X_N)$ is 1-dependent and

$$E(X_i) = 0,\ E(X_i^2) < \infty \text{ for all } i \text{ and } \sigma^2 := E(S^2) > 0.$$

Moreover, Z is a standard normal random variable *independent of* $(X_1, \ldots, X_N)$.

For each $i = 1, \ldots, N$, define

$$Y_i = X_i - E(X_i \mid \mathcal{F}_{i-1}) + E(X_{i+1} \mid \mathcal{F}_i)$$

where $\mathcal{F}_0$ is the trivial σ-field, $\mathcal{F}_i = \sigma(X_1, \ldots, X_i)$ and $X_{N+1} = 0$. Then,

$$E(Y_i \mid \mathcal{F}_{i-1}) = 0 \text{ for all } i \text{ and } \sum_{i=1}^N Y_i = \sum_{i=1}^N X_i = S \text{ a.s.}$$

Lemma 9. *Let $\gamma > 0$ be a constant and $V^2 = \sum_{i=1}^N E(Y_i^2 \mid \mathcal{F}_{i-1})$. Then,*

$$E\Big\{\Big(\frac{V^2}{\sigma^2} - 1\Big)^2\Big\} \le 16\,\gamma^2$$

provided $\max_i |X_i| \le \sigma\,\gamma/3$ *a.s.*

Proof. First note that

$$\sigma^2 = E(S^2) = E\big\{\big(\sum_{i=1}^N Y_i\big)^2\big\} = \sum_{i=1}^N E(Y_i^2) = E\big(\sum_{i=1}^N Y_i^2\big).$$

Moreover, since $\max_i |Y_i| \le \gamma\,\sigma$ a.s., one obtains

$$\sum_{i=1}^N E(Y_i^4) \le \gamma^2 \sigma^2 \sum_{i=1}^N E(Y_i^2) = \gamma^2 \sigma^4.$$

Therefore,

$$\begin{aligned}
E\Big\{\Big(\frac{V^2}{\sigma^2}-1\Big)^2\Big\} &\le \frac{2}{\sigma^4}\Big\{E\Big[\Big(\sum_{i=1}^N (E(Y_i^2\mid\mathcal{F}_{i-1})-Y_i^2)\Big)^2\Big]+\operatorname{Var}\Big(\sum_{i=1}^N Y_i^2\Big)\Big\}\\
&=\frac{2}{\sigma^4}\Big\{\sum_{i=1}^N E\Big(Y_i^4-E(Y_i^2\mid\mathcal{F}_{i-1})^2\Big)+\sum_{i=1}^N \operatorname{Var}(Y_i^2)\\
&\qquad+2\sum_{1\le i<j\le N}\operatorname{Cov}(Y_i^2,\,Y_j^2)\Big\}\\
&\le \frac{4}{\sigma^4}\Big\{\sum_{i=1}^N E(Y_i^4)+\sum_{1\le i<j\le N}\operatorname{Cov}(Y_i^2,\,Y_j^2)\Big\}\\
&\le 4\,\gamma^2+\frac{4}{\sigma^4}\sum_{1\le i<j\le N}\operatorname{Cov}(Y_i^2,\,Y_j^2).
\end{aligned}$$

To estimate the covariance part, define

$$Q_i=Y_i^2-E(Y_i^2)\quad\text{and}\quad T_i=\sum_{k=1}^i Y_k=\sum_{k=1}^i X_k+E(X_{i+1}\mid\mathcal{F}_i).$$

For each fixed $1\le i<N$, since $(T_1,\dots,T_N)$ is a martingale,

$$\begin{aligned}
\sum_{j>i}\operatorname{Cov}(Y_i^2,\,Y_j^2) &= \sum_{j>i}E\big(Q_iY_j^2\big)=E\Big\{Q_i\sum_{j>i}Y_j^2\Big\}=E\Big\{Q_i\,(T_N-T_i)^2\Big\}\\
&=E\Big\{Q_i\,(T_N-T_{i+1})^2\Big\}+E\big(Q_iY_{i+1}^2\big)\\
&\le E\Big\{Q_i\,(T_N-T_{i+1})^2\Big\}+E(Y_i^4)+E(Y_{i+1}^4).
\end{aligned}$$

Finally, since $(X_1,\dots,X_N)$ is 1-dependent, $EQ_i=0$ and $EX_j=0$,

$$\begin{aligned}
E\Big\{Q_i\,(T_N-T_{i+1})^2\Big\} &= E\Big\{Q_i\,\Big(\sum_{k=i+2}^N X_k-E(X_{i+2}\mid\mathcal{F}_{i+1})\Big)^2\Big\}\\
&=E\Big\{Q_i\,\Big(E(X_{i+2}\mid\mathcal{F}_{i+1})^2-2\,X_{i+2}\,E(X_{i+2}\mid\mathcal{F}_{i+1})\Big)\Big\}\\
&=-E\Big\{Q_i\,E(X_{i+2}\mid\mathcal{F}_{i+1})^2\Big\}\\
&\le E(Y_i^2)\,E\big\{E(X_{i+2}\mid\mathcal{F}_{i+1})^2\big\}\le\gamma^2\sigma^2E(Y_i^2).
\end{aligned}$$

To sum up,

$$E\Big\{\Big(\frac{V^2}{\sigma^2}-1\Big)^2\Big\}\le 4\,\gamma^2+\frac{4}{\sigma^4}\sum_{i=1}^{N-1}\Big(E(Y_i^4)+E(Y_{i+1}^4)+\gamma^2\sigma^2E(Y_i^2)\Big)\le 16\,\gamma^2.$$

□

Lemma 10. *If* $\max_i|X_i|\le\sigma\,\gamma/3$ *a.s., then*

$$d_W\Big(\frac{S}{\sigma}\,,\,Z\Big)\le 16\,\gamma^{1/3}.$$

Proof. By Lemma 2, $d_W\big(S/\sigma,\,Z\big)\le\sqrt{2}$. Hence, it can be assumed that $\gamma\le1$.

Define

$$\tau = \max\Big\{m : 1 \le m \le N, \sum_{k=1}^{m} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1}) \le 1\Big\},$$

$$J_i = 1\{\tau \ge i\}\,\frac{Y_i}{\sigma} + 1\{\tau = i-1\}\,\Big(1 - \sum_{k=1}^{i-1} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^{1/2} Z \quad \text{for } i = 1, \dots, N,$$

$$J_{N+1} = 1\{\tau = N\}\,\Big(1 - \sum_{k=1}^{N} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^{1/2} Z.$$

Since τ is a stopping time, Z is independent of $(X_1, \dots, X_N)$, and $E(Y_i \mid \mathcal{F}_{i-1}) = 0$, one obtains

$$E(J_i \mid \mathcal{F}_{i-1}) = 0 \text{ for all } i \text{ and } \sum_{k=1}^{N+1} E(J_k^2 \mid \mathcal{F}_{k-1}) = 1 \text{ a.s.}$$

Therefore, for each $a > 0$, a result by Röllin [16, Theorem 2.1] implies

$$d_W\Big(\sum_{i=1}^{N+1} J_i\,,\, Z\Big) \le 2a + \frac{3}{a^2}\,\sum_{i=1}^{N+1} E|J_i|^3.$$

To estimate $E|J_i|^3$ for $i \le N$, note that $E|Z|^3 \le 2$ and $(1/\sigma)\,\max_i|Y_i| \le \gamma$ a.s. Therefore, for $1 \le i \le N$,

$$\begin{aligned}
E|J_i|^3 &= E\Big\{1\{\tau \ge i\}\,\frac{|Y_i|^3}{\sigma^3}\Big\} + E\Big\{1\{\tau = i-1\}\,\Big(1 - \sum_{k=1}^{i-1} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^{3/2} |Z|^3\Big\} \\
&\le \gamma\, E\Big\{1\{\tau \ge i\}\,\frac{Y_i^2}{\sigma^2}\Big\} \\
&\quad + E\Big\{1\{\tau = i-1\}\,\Big(1 - \sum_{k=1}^{i-1} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^{1/2}\Big\}\, E|Z|^3 \\
&\le \gamma\, E\Big\{1\{\tau \ge i\}\,\frac{Y_i^2}{\sigma^2}\Big\} \\
&\quad + 2\, E\Big\{1\{\tau = i-1\}\,\Big(\sum_{k=1}^{i} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1}) - \sum_{k=1}^{i-1} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^{1/2}\Big\} \\
&= \gamma\, E\Big\{1\{\tau \ge i\}\,\frac{Y_i^2}{\sigma^2}\Big\} + 2\, E\Big\{1\{\tau = i-1\}\, E(Y_i^2/\sigma^2 \mid \mathcal{F}_{i-1})^{1/2}\Big\} \\
&\le \gamma\, E\Big\{1\{\tau \ge i\}\,\frac{Y_i^2}{\sigma^2}\Big\} + 2\,\gamma\, P(\tau = i-1).
\end{aligned}$$

Hence,

$$\sum_{i=1}^{N} E|J_i|^3 \le \gamma\, E\Big[\sum_{i=1}^{N} \frac{Y_i^2}{\sigma^2}\Big] + 2\gamma = 3\gamma.$$

Similarly,

$$\begin{aligned}
E|J_{N+1}|^3 &= E\Big\{1\{\tau = N\}\,\Big(1-\sum_{k=1}^{N} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^{3/2}\Big\}\,E|Z|^3 \\
&\le 2\,E\Big\{1\{\tau = N\}\,\Big(1-\sum_{k=1}^{N} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)\Big\} \\
&\le 2\,E\Big\{\Big(1-\sum_{k=1}^{N} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})\Big)^2\Big\}^{1/2} \\
&= 2\,E\Big\{\Big(1-\frac{V^2}{\sigma^2}\Big)^2\Big\}^{1/2} \le 8\,\gamma
\end{aligned}$$

where the last inequality is due to Lemma 9. It follows that

$$d_W\Big(\sum_{i=1}^{N+1} J_i\,,\,Z\Big) \le 2a + \frac{3}{a^2}(3\gamma+8\gamma) = 2a + \frac{33\,\gamma}{a^2}.$$

Next, we estimate $d_W\big(S/\sigma,\,\sum_{i=1}^{N} J_i\big)$. To this end, we let

$$W_i = \sum_{k=1}^{i} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1})$$

and we note that

$$\begin{aligned}
\frac{S}{\sigma} - \sum_{i=1}^{N} J_i &= \sum_{i=1}^{N}\Big(\frac{Y_i}{\sigma} - J_i\Big) = \sum_{i=1}^{N} 1\{\tau < i\}\,\Big(\frac{Y_i}{\sigma} - J_i\Big) \\
&= \sum_{i=1}^{N-1} 1\{\tau = i\}\,\Big\{\sum_{k=i+1}^{N}\frac{Y_k}{\sigma} - \big(1-W_i\big)^{1/2} Z\Big\}.
\end{aligned}$$

Therefore, recalling the definition of τ,

$$\begin{aligned}
d_W\Big(\frac{S}{\sigma}\,,\,\sum_{i=1}^{N} J_i\Big)^2 &\le \Big(E\Big|\frac{S}{\sigma} - \sum_{i=1}^{N} J_i\Big|\Big)^2 \le E\Big\{\Big(\frac{S}{\sigma} - \sum_{i=1}^{N} J_i\Big)^2\Big\} \\
&= \sum_{i=1}^{N-1} E\Big\{1\{\tau = i\}\,\Big\{\sum_{k=i+1}^{N}\frac{Y_k}{\sigma} - \big(1-W_i\big)^{1/2} Z\Big\}^2\Big\} \\
&= \sum_{i=1}^{N-1} E\Big\{1\{\tau = i\}\,\Big\{\sum_{k=i+1}^{N} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1}) + 1 - W_i\Big\}\Big\} \\
&\le \sum_{i=1}^{N-1} E\Big\{1\{\tau = i\}\,\Big\{\sum_{k=i+2}^{N} E(Y_k^2/\sigma^2 \mid \mathcal{F}_{k-1}) + 2\,E(Y_{i+1}^2/\sigma^2 \mid \mathcal{F}_i)\Big\}\Big\} \\
&\le \sum_{i=1}^{N-1} E\Big\{1\{\tau = i\}\,\Big\{V^2/\sigma^2 - 1 + 2\,\gamma^2\Big\}\Big\} \le E|V^2/\sigma^2 - 1| + 2\,\gamma^2 \\
&\le 4\,\gamma + 2\,\gamma^2
\end{aligned}$$

where the last inequality is because of Lemma 9. Since we assumed $\gamma \le 1$, we obtain

$$d_W\Big(\frac{S}{\sigma}\,,\,\sum_{i=1}^N J_i\Big) \le \sqrt{6\,\gamma}.$$

Finally, using Lemma 9 again, one obtains

$$d_W\Big(\sum_{i=1}^N J_i\,,\,\sum_{i=1}^{N+1} J_i\Big) \le E|J_{N+1}| \le E\Big\{\Big|\frac{V^2}{\sigma^2}-1\Big|^{1/2}\Big\} \le E\Big\{\Big(\frac{V^2}{\sigma^2}-1\Big)^2\Big\}^{1/4} \le 2\sqrt{\gamma}.$$

Collecting all these facts together,

$$\begin{aligned} d_W\Big(\frac{S}{\sigma}\,,\,Z\Big) &\le d_W\Big(\frac{S}{\sigma}\,,\,\sum_{i=1}^N J_i\Big) + d_W\Big(\sum_{i=1}^N J_i\,,\,\sum_{i=1}^{N+1} J_i\Big) + d_W\Big(\sum_{i=1}^{N+1} J_i\,,\,Z\Big) \\ &\le \sqrt{6\,\gamma} + 2\sqrt{\gamma} + 2a + \frac{33\,\gamma}{a^2} \le 5\,\sqrt{\gamma} + 2a + \frac{33\,\gamma}{a^2}. \end{aligned}$$

For $a = 4\gamma^{1/3}$, the above inequality yields, using again $\gamma \le 1$,

$$d_W\Big(\frac{S}{\sigma}\,,\,Z\Big) \le 5\,\sqrt{\gamma} + \Big(8 + \frac{33}{16}\Big)\,\gamma^{1/3} \le 16\,\gamma^{1/3}.$$

This concludes the proof. □

Remark 11. If we do not care about the value of the constant in the estimate, the proof of Lemma 10 could be shortened by exploiting a result by Fan and Ma [8]; this result, however, does not provide explicit values of the majorizing constants. We also note that, under the conditions of Lemma 10, Heyde–Brown's inequality [10] yields

$$d_K\Big(\frac{S}{\sigma}\,,\,Z\Big) \le b\,\Big\{E\Big(\Big(\frac{V^2}{\sigma^2}-1\Big)^2\Big) + \frac{1}{\sigma^4}\sum_{i=1}^N EY_i^4\Big\}^{1/5}$$

for some constant b independent of N. By Lemmas 2 and 9, this implies

$$d_W\Big(\frac{S}{\sigma}\,,\,Z\Big) \le 4\sqrt{d_K\Big(\frac{S}{\sigma}\,,\,Z\Big)} \le 4\,\sqrt{b}\,\Big\{16\,\gamma^2 + \frac{\gamma^2}{\sigma^2}\sum_{i=1}^N EY_i^2\Big\}^{1/10} = 4\,\sqrt{b}\,17^{1/10}\,\gamma^{1/5}.$$

Hence, in this case, Lemma 10 works better than Heyde–Brown's inequality to estimate $d_W\big(S/\sigma\,,\,Z\big)$.

Recall $L(c)$ defined in Theorem 5.

Lemma 12. *Letting* $\sigma_c^2 = \mathrm{Var}\Big(\sum_{i=1}^N \frac{X_i}{\sigma}\,1\big\{|X_i| \le c\sigma\big\}\Big)$, *we have*

$$|\sigma_c - 1| \le |\sigma_c^2 - 1| \le 13\,L(c) \qquad \textit{for all } c > 0.$$

Proof. Fix $c > 0$ and define

$$A_i = \big\{|X_i| > c\sigma\big\}, \quad T_i = \frac{X_i}{\sigma}\,1_{A_i} - E\Big(\frac{X_i}{\sigma}\,1_{A_i}\Big), \quad V_i = \frac{X_i}{\sigma}\,1_{A_i^c} - E\Big(\frac{X_i}{\sigma}\,1_{A_i^c}\Big).$$

On noting that $\sigma_c^2 = \mathrm{Var}\,\Big(\sum_{i=1}^N V_i\Big)$, one obtains

$$1 = \mathrm{Var}\,\Big(\sum_{i=1}^N (T_i + V_i)\Big) = \mathrm{Var}\,\Big(\sum_{i=1}^N T_i\Big) + \sigma_c^2 + 2\,\mathrm{Cov}\,\Big(\sum_{i=1}^N T_i\,,\,\sum_{i=1}^N V_i\Big).$$

Since $(X_1,\ldots,X_N)$ is 1-dependent, it follows that

$$\begin{aligned}|\sigma_c^2-1| &\le \operatorname{Var}\Big(\sum_{i=1}^N T_i\Big)+2\,\Big|\operatorname{Cov}\Big(\sum_{i=1}^N T_i\,,\,\sum_{i=1}^N V_i\Big)\Big| \\ &= \operatorname{Var}\Big(\sum_{i=1}^N T_i\Big)+2\,\Big|\sum_{i=1}^N \operatorname{Cov}(T_i,V_i) \\ &\qquad + \sum_{i=1}^{N-1}\operatorname{Cov}(T_i,V_{i+1})+\sum_{i=2}^{N}\operatorname{Cov}(T_i,V_{i-1})\Big|.\end{aligned}$$

Moreover,

$$\begin{aligned}\operatorname{Var}\Big(\sum_{i=1}^N T_i\Big) &= \sum_{i=1}^N \operatorname{Var}(T_i)+2\sum_{i=1}^{N-1}\operatorname{Cov}(T_i,T_{i+1}) \\ &\le \sum_{i=1}^N \operatorname{Var}(T_i)+\sum_{i=1}^{N-1}\Big(\operatorname{Var}(T_i)+\operatorname{Var}(T_{i+1})\Big)\le 3\,L(c).\end{aligned} \tag{5}$$

Similarly,

$$\operatorname{Cov}(T_i,V_i) = -E\Big(\frac{X_i}{\sigma}\,1_{A_i}\Big)\,E\Big(\frac{X_i}{\sigma}\,1_{A_i^c}\Big) = E\Big(\frac{X_i}{\sigma}\,1_{A_i}\Big)^2 \le E\Big(\frac{X_i^2}{\sigma^2}\,1_{A_i}\Big)$$

and

$$\begin{aligned}\Big|\operatorname{Cov}(T_i,V_{i-1})\Big| &\le E\Big(\frac{|X_iX_{i-1}|}{\sigma^2}\,1_{A_i}\,1_{A_{i-1}^c}\Big)+E\Big(\frac{|X_i|}{\sigma}\,1_{A_i}\Big)\,E\Big(\frac{|X_{i-1}|}{\sigma}\,1_{A_{i-1}^c}\Big) \\ &\le 2\,c\,E\Big(\frac{|X_i|}{\sigma}\,1_{A_i}\Big)\le 2\,E\Big(\frac{X_i^2}{\sigma^2}\,1_{A_i}\Big)\end{aligned}$$

where the last inequality is because

$$\frac{c\,|X_i|}{\sigma}\,1_{A_i}\le\frac{|X_i^2|}{\sigma^2}\,1_{A_i}.$$

By the same argument, $\Big|\operatorname{Cov}(T_i,V_{i+1})\Big| \le 2\,\sigma^{-2}E\big(X_i^2\,1_{A_i}\big)$. Collecting all these facts together, one finally obtains

$$|\sigma_c^2-1|\le 3\,L(c)+10\sum_{i=1}^N E\Big(\frac{X_i^2}{\sigma^2}\,1_{A_i}\Big)=13\,L(c).$$

This completes the proof, since obviously $|\sigma_c-1|\le|\sigma_c^2-1|$. □

Having proved the previous lemmas, we are now ready to attack Theorem 5.

Proof of Theorem 5. Fix $c>0$. We have to show that

$$d_W\Big(\frac{S}{\sigma},\,Z\Big)\le 30\,\big\{c^{1/3}+6\,L(c)^{1/2}\big\}.$$

Since $d_W\big(S/\sigma,Z\big)\le\sqrt{2}$, this inequality is trivially true if $L(c)\ge 1/100$ or if $c\ge 1$. Hence, it can be assumed $L(c)<1/100$ and $c<1$. Then, Lemma 12 implies $\sigma_c>0$.

Define T_i and V_i as in the proof of Lemma 12. Then $|V_i| \le 2c$ for every i, and thus $(V_1, \dots, V_N)$ satisfies the conditions of Lemma 10 with σ replaced by σ_c and $\gamma = 6\,c/\sigma_c$. Hence,

$$d_W\Big(\frac{\sum_{i=1}^N V_i}{\sigma_c}\,,\,Z\Big) \le 16\,\big(6c/\sigma_c\big)^{1/3}.$$

Now, recall from (5) that $\mathrm{Var}\big(\sum_{i=1}^N T_i\big) \le 3\,L(c)$. Hence, using Lemma 12 again, and the assumptions $L(c) < 1$ and $c < 1$,

$$\begin{aligned}
d_W\Big(\frac{S}{\sigma}\,,\,Z\Big) &\le d_W\Big(\frac{S}{\sigma}\,,\,\sum_{i=1}^N V_i\Big) + d_W\big(\sum_{i=1}^N V_i,\,\sigma_c Z\big) + d_W(\sigma_c Z,\,Z)\\
&\le E\Big|\frac{S}{\sigma} - \sum_{i=1}^N V_i\Big| + \sigma_c\,d_W\Big(\frac{\sum_{i=1}^N V_i}{\sigma_c}\,,\,Z\Big) + |\sigma_c - 1|\\
&\le \sqrt{\mathrm{Var}\,\big(\sum_{i=1}^N T_i\big)} + 16\,\big(6\,c\,\sigma_c^2\big)^{1/3} + 13\,L(c)\\
&\le \sqrt{3\,L(c)} + 16\,(6\,c)^{1/3}\,\big(1 + 13\,L(c)\big)^{2/3} + 13\,L(c)\\
&\le \big(\sqrt{3} + 13\big)\,L(c)^{1/2} + 16\,(6\,c)^{1/3}\,\Big(1 + \big(13\,L(c)\big)^{2/3}\Big)\\
&\le 16\,(6\,c)^{1/3} + \Big(\sqrt{3} + 13 + 16\cdot 6^{1/3}\cdot(13)^{2/3}\Big)\,L(c)^{1/2}\\
&\le 30\,c^{1/3} + 180\,L(c)^{1/2}.
\end{aligned}$$

This concludes the proof of Theorem 5. □

4. Total variation distance and rate of convergence

Theorems 3 and 4 immediately imply the following result.

Theorem 13. *Let ϕ_n be the characteristic function of S_n/σ_n and*

$$l_n = 2\int_0^\infty t\,|\phi_n(t)|\,dt.$$

If conditions (1)–(2) *hold, then*

$$d_{TV}\Big(\frac{S_n}{\sigma_n},\,Z\Big) \le \sqrt{120}\,\Big\{c^{1/3}{+}12\,U_n(c/2)^{1/2}\Big\}^{1/2} + 30^{1/3}\,l_n^{2/3}\,\Big\{c^{1/3}{+}12\,U_n(c/2)^{1/2}\Big\}^{1/3}$$

for all $n \ge 1$ and $c > 0$, where Z is a standard normal random variable.

Proof. First apply Theorem 3, with $V = 1$ and $X_n = \frac{S_n}{\sigma_n}$, and then use Theorem 4. □

Obviously, Theorem 13 is non-trivial only if $l_n < \infty$. In this case, the probability distribution of S_n is absolutely continuous. An useful special case is when conditions (1)–(2) hold and

(6) $$\max_i |X_{n,i}| \le \sigma_n \gamma_n \qquad \text{a.s. for some constants } \gamma_n.$$

Under (6), since $U_n(m_n\,\gamma_n)=0$, Theorem 13 yields

$$d_{TV}\Big(\frac{S_n}{\sigma_n},\,Z\Big)\le\sqrt{120}\,\,(2\,m_n\,\gamma_n)^{1/6}+\,30^{1/3}\,l_n^{2/3}\,(2\,m_n\,\gamma_n)^{1/9}.$$

Sometimes, this inequality allows to obtain a CLT in total variation distance; see Example 16 below.

Finally, we discuss the convergence rate provided by Theorem 4 and we compare it with some existing results.

A first remark is that Theorem 4 is calibrated to the dependence case, and that it is not optimal in the independence case. To see this, it suffices to recall that we assume $m_n\ge 1$ for all n. If $X_{n1},\ldots,X_{nN_n}$ are independent, the best one can do is to let $m_n=1$, but this choice of m_n is not efficient as is shown by the following example.

Example 14. Suppose $X_{n1},\ldots,X_{nN_n}$ are independent and conditions (2) and (6) hold. Define $m_n=1$ for all n. Then, $U_n(\gamma_n)=0$ and Theorem 4 implies $d_W\big(S_n/\sigma_n,\,Z\big)\le 30\,(2\,\gamma_n)^{1/3}$. However, the Bikelis nonuniform inequality yields

$$\Big|P(S_n/\sigma_n\le t)-P(Z\le t)\Big|\le\frac{b}{(1+|t|)^3}\sum_{i=1}^{N_n}E\Big\{\frac{|X_{n,i}|^3}{\sigma_n^3}\Big\}\le\frac{b\,\gamma_n}{(1+|t|)^3}$$

for all $t\in\mathbb{R}$ and some universal constant b; see e.g. [5, p. 659]. Hence,

$$d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big)=\int_{-\infty}^{\infty}|P(S_n/\sigma_n\le t)-P(Z\le t)|\,dt\le\int_{-\infty}^{\infty}\frac{b\,\gamma_n}{(1+|t|)^3}\,dt=b\,\gamma_n.$$

Leaving independence aside, a recent result to be mentioned is [6, Corollary 4.3] by Dedecker, Merlevede and Rio. This result applies to sequences of random variables and requires a certain mixing condition (denoted by (H_1)) which is automatically true when $m_n=m$ for all n. In this case, under conditions (2) and (6), one obtains

$$d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big)\le b\,\gamma_n\,\Big(1+c_n\,\log\big(1+c_n\,\sigma_n^2\big)\Big)\tag{7}$$

where b and c_n are suitable constants with b independent of n. Among other conditions, the c_n must satisfy

$$c_n\,\sigma_n^2\ge\sum_{i=1}^{N_n}EX_{n,i}^2.$$

Inequality (7) is actually sharp. However, if compared with Theorem 4, it has three drawbacks. First, unlike Theorem 4, it requires condition (6). Secondly, the mixing condition (H_1) is not easily verified unless $m_n=m$ for all n. Thirdly, as seen in the next example, even if (6) holds and $m_n=m$ for all n, it may be that

$$\gamma_n\to 0\quad\text{but}\quad\gamma_n\,c_n\,\log\big(1+c_n\,\sigma_n^2\big)\to\infty\qquad\text{as } n\to\infty.$$

In such situations, Theorem 4 works while inequality (7) does not.

Example 15. Let (a_n) be a sequence of numbers in $(0,1)$ such that $\lim_n a_n=0$. Let $(T_i:i\ge 0)$ and $(V_{n,i}:n\ge 1,\,1\le i\le n)$ be two independent collections of real random variables. Suppose (T_i) is i.i.d. with $P(T_0=\pm 1)=1/2$ and $V_{n,1},\ldots,V_{n,n}$ are i.i.d. with $V_{n,1}$ uniformly distributed on the set $(-1,-1+a_n)\cup(1-a_n,1)$.

Fix a constant $\alpha \in (0,\, 1/3)$ and define $N_n = n$ and

$$X_{n,i} = n^{-1/2} V_{n,i} + n^{-\alpha}(T_i - T_{i-1})$$

for $i = 1, \ldots, n$. The array $(X_{n,i})$ is centered and 1-dependent (namely, $m_n = 1$ for all n). In addition, $S_n = n^{-1/2} \sum_{i=1}^n V_{n,i} + n^{-\alpha}(T_n - T_0)$ and

$$\sigma_n^2 = EV_{n,1}^2 + 2\, n^{-2\alpha}, \quad \sum_{i=1}^n EX_{n,i}^2 = EV_{n,1}^2 + 2\, n^{1-2\alpha}.$$

Since $\lim_n \sigma_n^2 = \lim_n EV_{n,1}^2 = 1$, one obtains

$$\max_i \frac{|X_{n,i}|}{\sigma_n} \le \frac{n^{-1/2} + 2\, n^{-\alpha}}{\sigma_n} \le \frac{3\, n^{-\alpha}}{\sigma_n} < 4\, n^{-\alpha} \quad \text{for large } n.$$

Hence, for large n, condition (6) holds with $\gamma_n = 4\, n^{-\alpha}$. Since $U_n(4\, n^{-\alpha}) = 0$, Theorem 4 implies (taking $c = 8n^{-\alpha}$)

$$d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) \le 60\, n^{-\alpha/3} \qquad \text{for large } n.$$

However,

$$\begin{aligned} 4\, n^{-\alpha}\, c_n\, \log\big(1 + c_n\, \sigma_n^2\big) &\ge 4\, n^{-\alpha}\, \frac{1}{\sigma_n^2} \sum_{i=1}^n EX_{n,i}^2\, \log\Big(1 + \sum_{i=1}^n EX_{n,i}^2\Big) \\ &\ge 4\,(1 - 2\alpha)\, \frac{n^{1-3\alpha}}{\sigma_n^2}\, \log n \longrightarrow \infty. \end{aligned}$$

In addition to [6, Corollary 4.3], there are some other estimates of $d_W\big(S_n/\sigma_n,\, Z\big)$. Without any claim of exhaustivity, we mention Fan and Ma [8], Röllin [16] and Van Dung, Son and Tien [20] (Röllin's result has been used for proving Lemma 10). There are also a number of estimates of $d_K\big(S_n/\sigma_n,\, Z\big)$ which, through Lemma 2, can be turned into upper bounds for $d_W\big(S_n/\sigma_n,\, Z\big)$; see [6], [8] and references therein. However, to our knowledge, none of these estimates implies Theorem 4. Typically, they require further conditions (in addition to (1)–(2)) and/or they yield a worse convergence rate; see e.g. Remark 11 and Example 15. This is the current state of the art. Our conjecture is that, under conditions (1)–(2) and possibly (6), the rate of Theorem 4 can be improved. To this end, one possibility could be using an upper bound provided by Haeusler and Joos [9] in the martingale CLT. Whether the rate of Theorem 4 can be improved, however, is currently an *open problem*.

We conclude the paper with a CLT in total variation distance obtained via Theorem 13.

Example 16. Let $(X_{n,i})$ and $(V_{n,i})$ be as in Example 15. Denote by ψ_n the characteristic function of $\sum_{i=1}^n V_{n,i}$. Then, for each $t \in \mathbb{R}$,

$$\psi_n(t) = \Big(\frac{1}{a_n} \int_{1-a_n}^1 \cos(t\, x)\, dx\Big)^n \qquad \text{and}$$

$$|\phi_n(t)| \le \Big|\psi_n\big[t\,(n\,\sigma_n^2)^{-1/2}\big]\Big| = \Big|\frac{1}{a_n} \int_{1-a_n}^1 \cos\big[t\,(n\,\sigma_n^2)^{-1/2}\, x\big]\, dx\Big|^n.$$

After some algebra (we omit the explicit calculations) it can be shown that

$$l_n = 2 \int_0^\infty t\, |\phi_n(t)|\, dt \le b\, a_n^{-2}$$

for some constant b independent of n. Recalling that $m_n = 1$ and $\gamma_n = 4\,n^{-\alpha}$ for large n (see Example 15), Theorem 13 yields (taking again $c = 2m_n\gamma_n = 8n^{-\alpha}$)

$$d_{TV}\Big(\frac{S_n}{\sigma_n},\, Z\Big) \le \sqrt{120}\;(2\,m_n\,\gamma_n)^{1/6} + 30^{1/3}\,l_n^{2/3}\,(2\,m_n\,\gamma_n)^{1/9}$$
$$\le \sqrt{120}\;8^{1/6}\,n^{-\alpha/6} \,+\, 30^{1/3}\,b^{2/3}\,8^{1/9}\,\Big(a_n^4\,n^{\alpha/3}\Big)^{-1/3}$$

for large n. Therefore, the probability distribution of S_n/σ_n converges to the standard normal law, in total variation distance, provided $a_n^4\,n^{\alpha/3} \to \infty$.

References

[1] Bergström H. (1970) A comparison method for distribution functions of sums of independent and dependent random variables, (Russian), *Teor. Verojatnost. i Primenen* 15, 442–468, 750; English transl. *Theor. Probab. Appl.* 15 (1970), 430–457, 727. MR 0283850, MR 0281245

[2] Berk K.N. (1973) A central limit theorem for m-dependent random variables with unbounded m, *Ann. Probab.* 1, 352–354. MR 0350815

[3] Bradley R.C. (2007) *Introduction to strong mixing conditions*, Vol. 1–3, Kendrick Press, Heber City, UT. MR 2325294–2325296

[4] Chen L.H.Y., Goldstein L., Shao Q.-M. (2011) *Normal approximation by Stein's method*, Springer-Verlag, Berlin. MR 2732624

[5] DasGupta A. (2008) *Asymptotic theory of statistics and probability*, Springer, New York.

[6] Dedecker J., Merlevede F., Rio E. (2022) Rates of convergence in the central limit theorem for martingales in the non stationary setting, *Ann. Inst. H. Poincare Probab. Statist.* 58, 945–966.

[7] Diananda P.H. (1955) The central limit theorem for m-dependent variables, *Proc. Cambridge Philos. Soc.* 51, 92–95. MR 0067396

[8] Fan X., Ma X. (2020) On the Wasserstein distance for a martingale central limit theorem, *Statist. Probab. Lett.* 167, 108892. MR 4138415

[9] Haeusler E., Joos K. (1988) A nonuniform bound on the rate of convergence in the martingale central limit theorem, *Ann. Probab.* 16, 1699–1720.

[10] Heyde C.C., Brown B.M. (1970) On the departure from normality of a certain class of martingales, *Ann. Math. Stat.* 41, 2161–2165. MR 0293702

[11] Hoeffding W., Robbins H. (1948) The central limit theorem for dependent random variables, *Duke Math. J.* 15, 773–780. MR 0026771

[12] Orey S. (1958) A central limit theorem for m-dependent random variables, *Duke Math. J.* 25, 543–546. MR 0097841

[13] Peligrad M. (1996) On the asymptotic normality of sequences of weak dependent random variables, *J. Theoret. Probab.* 9, 703–715. MR 1400595

[14] Pratelli L., Rigo P. (2018) Convergence in total variation to a mixture of Gaussian laws, *Mathematics* 6, 99.

[15] Rio E. (1995) About the Lindeberg method for strongly mixing sequences, *ESAIM: Probab. Statist.* 1, 35–61. MR 1382517

[16] Röllin A. (2018) On quantitative bounds in the mean martingale central limit theorem, *Statist. Probab. Lett.* 138, 171–176. MR 3788734

[17] Romano J.P., Wolf M. (2000) A more general central limit theorem for m-dependent random variables with unbounded m, *Statist. Probab. Lett.* 47, 115–124. MR 1747098

[18] Utev S.A. (1987) On the central limit theorem for φ-mixing arrays of random variables, *Theor. Probab. Appl.* 35, 131–139. MR 1050059

[19] Utev S.A. (1990) Central limit theorem for dependent random variables, *Probability theory and mathematical statistics, Vol. II* (Vilnius, 1989), 519–528, Mokslas, Vilnius. MR 1153906

[20] Van Dung L., Son T.C., Tien N.D. (2014) L_1-Bounds for some martingale central limit theorems, *Lith. Math. J.* 54, 48–60. MR 3189136

Svante Janson, Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden, URL: http://www.math.uu.se/svante-janson
Email address: svante.janson@math.uu.se

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
Email address: pratel@mail.dm.unipi.it

Pietro Rigo, Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, via delle Belle Arti 41, 40126 Bologna, Italy
Email address: pietro.rigo@unibo.it