ALEA, Lat. Am. J. Probab. Math. Stat. **21**, 245–265 (2024)
DOI: 10.30757/ALEA.v21-10

Quantitative bounds in the central limit theorem for m-dependent random variables

Svante Janson, Luca Pratelli and Pietro Rigo

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

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
E-mail address: `Luca_pratelli@marina.difesa.it`

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

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 (1948) and Diananda (1955) (for m-dependent sequences), and Orey (1958) (more generally, and also for triangular arrays). These results were then extended to the case of increasing $m = m_n$; see e.g. Bergström (1970), Berk (1973), Rio (1995), Romano and Wolf (2000), and Utev (1990a,b).

Received by the editors August 23rd, 2022; accepted December 27th, 2023.
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.
SJ supported by the Knut and Alice Wallenberg Foundation.

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 Bradley (2007), Dedecker et al. (2022), Peligrad (1996), Rio (1995), Utev (1990a,b) 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.

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

$$(X_{n,1}, \ldots, X_{n,N_n}) \text{ is } m_n\text{-dependent for every } n, \tag{1.1}$$

$$E(X_{ni}) = 0, \quad E(X_{ni}^2) < \infty, \quad \sigma_n^2 := E(S_n^2) > 0 \quad \text{for all } n \text{ and } i, \tag{1.2}$$

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 3.1) is that

$$d_W\Big(\frac{S_n}{\sigma_n}, \, Z\Big) \leq 30\,\big\{c^{1/3} + 12\, U_n(c/2)^{1/2}\big\} \qquad \text{for all } n \geq 1 \text{ and } c > 0, \tag{1.3}$$

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

Inequality (1.3) provides a quantitative estimate of $d_W\big(S_n/\sigma_n, \, Z\big)$. The connections between (1.3) and other analogous results are discussed in Remark 3.9 and Section 4. To our knowledge, however, no similar estimate of $d_W\big(S_n/\sigma_n, \, Z\big)$ is available under conditions (1.1)–(1.2) only. In addition, inequality (1.3) implies the following useful result:

Theorem 1.1 (Utev (1990a,b)). $S_n/\sigma_n \overset{dist}{\longrightarrow} Z$ provided conditions (1.1)–(1.2) hold and $U_n(c) \to 0$ for every $c > 0$.

Based on inequality (1.3), 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 2.1. 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 Pratelli and Rigo (2018) (see Theorem 2.2 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 (1.3). For instance, in addition to (1.1)–(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

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

The rest of this paper is organized as follows. Section 2 just recalls some definitions and known results, Section 3 is devoted to proving inequality (1.3), while Section 4 investigates $d_{TV}\big(S_n/\sigma_n,\, Z\big)$ and the convergence rate provided by (1.3). Section 5 contains some examples that illustrate the main results. Section 6 ends the paper with an extension that does not require (m_n)-dependence (but uses some other conditions).

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 (1.3) and similar powers in other results; we do not believe that these are optimal. This is discussed in Section 4. How far (1.3) 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 $n \ge 1$, the variables $(X_{ni} : 1 \le i \le N_n)$ are defined on the same probability space (which may depend on n).

Let $k \ge 0$ be an integer. A (finite or infinite) sequence (Y_i) of random variables is k-dependent if $(Y_i : i \le j)$ 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 $(Y_{ni} : n \ge 1,\, 1 \le i \le N_n)$ is said to be (k_n)-dependent if $(Y_{ni} : 1 \le i \le N_n)$ 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,

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

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,

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

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 2.1. Suppose $EX^2 \le 1$, $EY^2 \le 1$ and $EY = 0$. Then,

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

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. Chen et al. (2011, 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 2.2 (A version of Pratelli and Rigo (2018, 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 Pratelli and Rigo (2018, Theorem 1), with $\beta = 2$ and the constant k made explicit. Also, the assumption $d_W(X_n, VZ) \to 0$ in Pratelli and Rigo (2018, Theorem 1) is not needed; we use instead $d_W(X_n, VZ) \leq \sqrt{2}$ from Lemma 2.1. Using this and $EX_n^2 = 1$, the various constants appearing in the proof can be explicitly evaluated. In fact, improving the argument in Pratelli and Rigo (2018) 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 Pratelli and Rigo (2018) 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 3.1. Under conditions (1.1)–(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.

Before proceeding, we note a simple special case for bounded random variables.

Corollary 3.2. Suppose that conditions (1.1)–(1.2) hold and

$$\max_i |X_{n,i}| \le \sigma_n \gamma_n \qquad \text{a.s. for some constants } \gamma_n. \tag{3.1}$$

Then,

$$d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) \le 30 \cdot 2^{1/3}\,(m_n\gamma_n)^{1/3} \le 40\,(m_n\gamma_n)^{1/3},$$

where Z denotes a standard normal random variable.

Proof: Take $c = 2m_n\gamma_n$ in Theorem 3.1 and note that $U_n(c/2) = 0$. □

In turn, Theorem 3.1 follows from the following result, which is a sharper version of the special case $m_n = 1$.

Theorem 3.3. Let $X_1, \dots, X_N$ be real random variables and $S = \sum_{i=1}^N X_i$. Suppose $(X_1, \dots, X_N)$ 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 3.1 from Theorem 3.3, define $M_n = \lceil N_n/m_n \rceil$, $X_{n,i} = 0$ for $i > N_n$, and

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

Since $(Y_{n,1}, \dots, Y_{n,M_n})$ is 1-dependent and $\sum_i Y_{n,i} = \sum_i X_{n,i} = S_n$, Theorem 3.3 implies

$$d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) \le 30\,\big\{c^{1/3} + 6\,L_n(c)^{1/2}\big\} \tag{3.2}$$

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 3.1, it suffices to note the following inequality:

Lemma 3.4. 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 3.4 and Theorem 3.3. We also obtain a (very small) improvement of Utev's Theorem 1.1.

3.1. *Proof of Lemma 3.4 and Utev's theorem.*

Proof of Lemma 3.4: 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 (3.2), Theorem 3.3 implies:

Corollary 3.5. $S_n/\sigma_n \xrightarrow{dist} Z$ if conditions (1.1)–(1.2) hold and $L_n(c) \to 0$ for every $c > 0$.

Corollary 3.5 slightly improves Theorem 1.1. In fact, $U_n(c) \to 0$ for all $c > 0$ implies $L_n(c) \to 0$ for all $c > 0$, because of Lemma 3.4, but the converse is not true.

Example 3.6. ($L_n(c) \to 0$ 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 $T_n \xrightarrow{dist} N(0,3)$ 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,

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

for each n such that $c\,\sigma_n/m_n \geq 1$ and $m_n \leq t_n$. 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 3.3.* Our proof of Theorem 3.3 requires three lemmas. A result by Röllin (2018) plays a crucial role in one of them (Lemma 3.8).

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

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

provided $\max_i |X_i| \leq \sigma\,\gamma/3$ 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 $\max_i |Y_i| \leq \gamma\,\sigma$ a.s., one obtains

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

Therefore,

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

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 $(T_1,\ldots,T_N)$ is a martingale,

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

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

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

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

$$
d_W\Big(\frac{S}{\sigma}\,,\,Z\Big)\le 15\,\gamma^{1/3}.
$$

Proof: By Lemma 2.1, $d_W\big(S/\sigma,\,Z\big)\le\sqrt{2}$. Hence, it can be assumed that $\gamma\le 1$.

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 (2018, 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 $1 \le i \le N$,

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

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 3.7. It follows that

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

for each $a > 0$. Choosing $a = 3\gamma^{1/3}$, this yields

$$
d_W\Big(\sum_{i=1}^{N+1} J_i\,,\,Z\Big) \le \Big(6 + \frac{11}{3}\Big)\gamma^{1/3} \le 10\,\gamma^{1/3}.
$$

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,

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

where the last inequality is because of Lemma 3.7. 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 3.7 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 yields, using again $\gamma \le 1$,

$$\begin{aligned}
d_W\Big(\frac{S}{\sigma}, Z\Big) &\le d_W\Big(\frac{S}{\sigma}, \sum_{i=1}^N J_i\Big) + d_W\Big(\sum_{i=1}^N J_i, \sum_{i=1}^{N+1} J_i\Big) + d_W\Big(\sum_{i=1}^{N+1} J_i, Z\Big) \\
&\le \sqrt{6\,\gamma} + 2\sqrt{\gamma} + 10\,\gamma^{1/3} \le 15\,\gamma^{1/3}.
\end{aligned}$$

This concludes the proof. □

Remark 3.9. If we do not care about the value of the constant in the estimate, the proof of Lemma 3.8 could be shortened by exploiting a result by Fan and Ma (2020); this result, however, does not provide explicit values of the majorizing constants. We also note that, under the conditions of Lemma 3.8, Heyde–Brown's inequality (Heyde and Brown, 1970) 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 2.1 and 3.7, 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 3.8 works better than Heyde–Brown's inequality to estimate $d_W\big(S/\sigma, Z\big)$.

Recall $L(c)$ defined in Theorem 3.3.

Lemma 3.10. 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

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

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

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

Moreover,

$$\begin{aligned} \mathrm{Var}\,\Big(\sum_{i=1}^N T_i\Big) &= \sum_{i=1}^N \mathrm{Var}(T_i) + 2 \sum_{i=1}^{N-1} \mathrm{Cov}(T_i, T_{i+1}) \\ &\le \sum_{i=1}^N \mathrm{Var}(T_i) + \sum_{i=1}^{N-1} \Big(\mathrm{Var}(T_i) + \mathrm{Var}(T_{i+1})\Big) \le 3\, L(c). \end{aligned} \tag{3.3}$$

Similarly,

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

and

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

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, $\Big|\mathrm{Cov}\,(T_i, V_{i+1})\Big| \le 2\,\sigma^{-2} E\big(X_i^2\, 1_{A_i}\big)$. 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 $|\sigma_c - 1| \le |\sigma_c^2 - 1|$. $\square$

Having proved the previous lemmas, we are now ready to attack Theorem 3.3.

Proof of Theorem 3.3: Fix $c > 0$. We have to show that

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

Since $d_W\big(S/\sigma, Z\big) \leq \sqrt{2}$, this inequality is trivially true if $L(c) \geq 1/100$ or if $c \geq 1$. Hence, it can be assumed $L(c) < 1/100$ and $c < 1$. Then, Lemma 3.10 implies $\sigma_c > 0$.

Define T_i and V_i as in the proof of Lemma 3.10. Then $|V_i| \leq 2c$ for every i, and thus $(V_1, \ldots, V_N)$ satisfies the conditions of Lemma 3.8 with σ replaced by σ_c and $\gamma = 6\,c/\sigma_c$. Hence,

$$d_W\Big(\frac{\sum_{i=1}^N V_i}{\sigma_c}\,,\, Z\Big) \leq 15\,\big(6c/\sigma_c\big)^{1/3}.$$

Now, recall from (3.3) that $\mathrm{Var}\big(\sum_{i=1}^N T_i\big) \leq 3\,L(c)$. Hence, using Lemma 3.10 again, and the assumptions $L(c) < 1$ and $c < 1$,

$$\begin{aligned}
d_W\Big(\frac{S}{\sigma}\,,\, Z\Big) &\leq d_W\Big(\frac{S}{\sigma}\,,\, \sum_{i=1}^N V_i\Big) + d_W\big(\sum_{i=1}^N V_i,\, \sigma_c Z\big) + d_W(\sigma_c Z,\, Z)\\
&\leq E\Big|\frac{S}{\sigma} - \sum_{i=1}^N V_i\Big| + \sigma_c\, d_W\Big(\frac{\sum_{i=1}^N V_i}{\sigma_c}\,,\, Z\Big) + |\sigma_c - 1|\\
&\leq \sqrt{\mathrm{Var}\,\big(\sum_{i=1}^N T_i\big)} + 15\,\big(6\,c\,\sigma_c^2\big)^{1/3} + 13\,L(c)\\
&\leq \sqrt{3\,L(c)} + 15\,(6\,c)^{1/3}\,\big(1 + 13\,L(c)\big)^{2/3} + 13\,L(c)\\
&\leq \big(\sqrt{3} + 13\big)\,L(c)^{1/2} + 15\,(6\,c)^{1/3}\,\Big(1 + \big(13\,L(c)\big)^{2/3}\Big)\\
&\leq 15\,(6\,c)^{1/3} + \Big(\sqrt{3} + 13 + 15\cdot 6^{1/3}\cdot (13)^{2/3}\Big)\,L(c)^{1/2}\\
&\leq 30\,c^{1/3} + 170\,L(c)^{1/2}.
\end{aligned}$$

This concludes the proof of Theorem 3.3. □

4. Total variation distance and rate of convergence

Theorems 2.2 and 3.1 immediately imply the following result.

Theorem 4.1. 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 (1.1)–(1.2) hold, then

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

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

Proof: First apply Theorem 2.2, with $V = 1$ and $X_n = \frac{S_n}{\sigma_n}$, and then use Theorem 3.1. □

Obviously, Theorem 4.1 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 (1.1)–(1.2) hold together with (3.1) (as in Corollary 3.2). Then, by taking $c = 2m_n\gamma_n$ so that $U_n(c/2) = 0$, Theorem 4.1 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 5.1 below.

We next discuss the convergence rate provided by Theorem 3.1 and we compare it with some existing results.

A first remark is that Theorem 3.1 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 4.2. Suppose $X_{n1}, \ldots, X_{nN_n}$ are independent and conditions (1.2) and (3.1) hold. Define $m_n = 1$ for all n. Then, $U_n(\gamma_n) = 0$ and Theorem 3.1 (or Corollary 3.2) yields $d_W\big(S_n/\sigma_n,\, Z\big) \le 30\,(2\,\gamma_n)^{1/3}$. 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. DasGupta (2008, 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 Dedecker et al. (2022, Corollary 4.3). 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 (1.2) and (3.1), one obtains

$$d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) \le b\,\gamma_n\,\Big(1 + c_n\, \log\big(1 + c_n\, \sigma_n^2\big)\Big) \tag{4.1}$$

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

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

Inequality (4.1) is actually sharp. However, if compared with Theorem 3.1, it has three drawbacks. First, unlike Theorem 3.1, it requires condition (3.1). 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 (3.1) 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 3.1 works while inequality (4.1) does not.

Example 4.3. 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

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

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 (3.1) holds with $\gamma_n = 4\,n^{-\alpha}$. Consequently, Corollary 3.2 yields

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

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 Dedecker et al. (2022, 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 (2020), Röllin (2018) and Van Dung et al. (2014) (Röllin's result has been used for proving Lemma 3.8). There are also a number of estimates of $d_K\big(S_n/\sigma_n,\, Z\big)$ which, through Lemma 2.1, can be turned into upper bounds for $d_W\big(S_n/\sigma_n,\, Z\big)$; see Dedecker et al. (2022), Fan and Ma (2020) and references therein. However, to our knowledge, none of these estimates implies Theorem 3.1. Typically, they require further conditions (in addition to (1.1)–(1.2)) and/or they yield a worse convergence rate; see e.g. Remark 3.9 and Example 4.3. This is the current state of the art. Our conjecture is that, under conditions (1.1)–(1.2) and possibly (3.1), the rate of Theorem 3.1 can be improved. To this end, one possibility could be using an upper bound provided by Haeusler and Joos (1988) in the martingale CLT. Whether the rate of Theorem 3.1 can be improved, however, is currently an *open problem*.

5. Further examples and applications

To illustrate the results above, we give some applications of Theorems 3.1 and 4.1. As usual, Z denotes a standard normal random variable. We begin with a CLT in total variation distance.

Example 5.1. Let $(X_{n,i})$ and $(V_{n,i})$ be as in Example 4.3. 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 4.3), Theorem 4.1 yields (taking again $c = 2m_n\gamma_n = 8n^{-\alpha}$)

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

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

The next two examples are connected to the Breuer-Major theorem (henceforth, BMT); see Breuer and Major (1983). In both the examples, $g : \mathbb{R} \to \mathbb{R}$ is a Borel function with Hermite degree $d \geq 1$. This means that $E\big(g^2(Z)\big) < \infty$ with a series expansion of the type

$$g = \sum_{j=d}^{\infty} c_j\,H_j, \quad c_d \neq 0,$$

where H_j is the Hermite polynomial of degree j.

Example 5.2. There is recently a certain interest on the asymptotic behavior of

$$Q_n = \frac{\sum_{i=0}^{n-1} g(Y_i)}{\sqrt{\mathrm{Var}\big[\sum_{i=0}^{n-1} g(Y_i)\big]}},$$

where $(Y_n : n \geq 0)$ is a stationary Gaussian sequence of standard normal random variables; see e.g. Campese et al. (2020), Nourdin and Nualart (2020) and references therein. Because of BMT, $Q_n \xrightarrow{dist} Z$ provided $\sum_n |E(Y_nY_0)|^d < \infty$ (recall that $d \geq 1$ is the Hermite degree of g). To obtain a quantitative estimate of $d_W(Q_n, Z)$, some further conditions are needed. Essentially, g must belong to a suitable Sobolev space.

At the price of assuming (m_n)-dependence, Theorem 3.1 allows to improve BMT. Among other things, the stationarity assumption is dropped, sequences are replaced by arrays, and the conditions on g are much more general.

For each $n \geq 1$, suppose

$$(X_{n,1}, \ldots, X_{n,N_n}) \text{ is Gaussian}, \quad X_{n,i} \sim \mathcal{N}(0,1) \text{ for all } i,$$
$$\text{and} \quad E\big(X_{n,i}X_{n,j}\big) = 0 \text{ whenever } |i-j| > m_n.$$

Moreover, fix any Borel function $g_n : \mathbb{R} \to \mathbb{R}$ such that $E\big(g_n(Z)\big) = 0$ and $E\big(g_n^2(Z)\big) < \infty$ and suppose

$$\sigma_n^2 := \mathrm{Var}\left[\sum_{i=1}^{N_n} g_i(X_{n,i})\right] > 0.$$

Then, Theorem 3.1 yields

$$d_W(Q_n^*,\, Z) \leq 30\,c^{1/3} + \frac{360\,\sqrt{m_n}}{\sigma_n}\left(\sum_{i=1}^{N_n} E\Big[g_i^2(Z)\,1\big\{|g_i(Z)| > c\,\sigma_n/2\,m_n\big\}\Big]\right)^{1/2}$$

for all $n \geq 1$ and $c > 0$, where

$$Q_n^* = \frac{1}{\sigma_n}\sum_{i=1}^{N_n} g_i(X_{n,i}).$$

This upper bound is effective if the sequence $(g_n^2(Z) : n \geq 1)$ is uniformly integrable. Note also that, if $g_n(Z) \in L_\infty$ for all n, Corollary 3.2 yields

$$d_W(Q_n^*,\, Z) \leq 40\, \left(\frac{m_n}{\sigma_n}\, \max_{1\leq i\leq N_n} \|g_i(Z)\|_\infty\right)^{1/3}.$$

Example 5.3. Let $Y = (Y_t : t \geq 0)$ be a real cadlag process. To begin with, suppose Y is stationary, Gaussian, $Y_0 \sim \mathcal{N}(0,1)$, and define

$$Z_\epsilon(t) = \sqrt{\epsilon}\int_0^{t/\epsilon} g(Y_s)\, ds \qquad \text{for all } \epsilon > 0 \text{ and } t \geq 0.$$

If $\int |E(Y_t Y_0)|^d\, dt < \infty$ then, as $\epsilon \to 0$, the finite dimensional distributions of Z_ϵ converge weakly to those of $\sigma\, W$, where σ is an explicit constant and W a standard Brownian motion. This is BMT in continuous-time. By a result in Campese et al. (2020), if $E\big(|g(Z)|^p\big) < \infty$ for some $p > 2$, one also obtains $Z_\epsilon \xrightarrow{dist} \sigma\, W$ in the space $C([0,\infty),\, \mathbb{R})$ (equipped with the topology of uniform convergence on compacta).

Next, suppose Y is a Levy process. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function and $\lambda : (0,\infty) \to (0,\infty)$ a non-increasing function such that

$$a := \sup|f| < \infty \quad \text{and} \quad b := \sup \lambda < \infty.$$

Roughly speaking, λ should be regarded as a delay in observing Y. Given $\epsilon > 0$ and $s \geq \lambda(\epsilon)$, the actual value of Y at time $s - \lambda(\epsilon)$ is not $Y_{s-\lambda(\epsilon)}$ but Y_s. Hence, $Y_s - Y_{s-\lambda(\epsilon)}$ may be seen as an observation error. Let

$$Z_\epsilon^*(t) = \sqrt{\epsilon}\int_0^{t/\epsilon} f\Big(Y_s - Y_{(s-\lambda(\epsilon))^+}\Big)\, ds.$$

In order to apply Theorem 3.1 to Z_ϵ^*, fix $t > 0$ and define

$$n_\epsilon(t) = \left\lfloor \frac{t}{\epsilon\lambda(\epsilon)} \right\rfloor - 1 \quad \text{and} \quad I_t = \big\{\epsilon > 0 :\, n_\epsilon(t) \geq 1\big\}.$$

For $\epsilon \in I_t$ and $i \geq 1$, define also

$$X_{\epsilon,i} = \sqrt{\epsilon}\int_{i\lambda(\epsilon)}^{(i+1)\lambda(\epsilon)} f\Big(Y_s - Y_{s-\lambda(\epsilon)}\Big)\, ds, \quad V_\epsilon(t) = \sum_{i=1}^{n_\epsilon(t)} X_{\epsilon,i}, \quad \sigma_\epsilon^2(t) = E\big(V_\epsilon^2(t)\big).$$

Assume $E\big[f\big(Y_{\lambda(\epsilon)}\big)\big] = 0$ (for example, this holds if f is odd and $Y_{\lambda(\epsilon)}$ is symmetric), and also $\sigma_\epsilon^2(t) > 0$ for $\epsilon \in I_t$. Then, $E\big[f\big(Y_s - Y_{s-\lambda(\epsilon)}\big)\big] = E\big[f\big(Y_{\lambda(\epsilon)}\big)\big] = 0$ for $s \geq \lambda(\epsilon)$, so that $E\big(X_{\epsilon,i}\big) = 0$, and since the array

$$\big(X_{\epsilon,i} : \epsilon \in I_t,\; i = 1, \ldots, n_\epsilon(t)\big)$$

is 1-dependent, Theorem 3.1 yields, for any $c > 0$,

$$d_W\left(\frac{V_\epsilon(t)}{\sigma_\epsilon(t)},\, Z\right) \leq 30\, c^{1/3} \,+\, 360\,\sqrt{\frac{n_\epsilon(t)}{\sigma_\epsilon^2(t)}\;\, E\Big[X_{\epsilon,1}^2\, 1\big\{|X_{\epsilon,1}| > c\,\sigma_\epsilon(t)/2\big\}\Big]}.$$

Moreover, since $|X_{\epsilon,i}| \leq a\, b\, \sqrt{\epsilon}$, Corollary 3.2 yields

$$d_W\left(\frac{V_\epsilon(t)}{\sigma_\epsilon(t)},\, Z\right) \leq 40\, \left(\frac{a\, b\, \sqrt{\epsilon}}{\sigma_\epsilon(t)}\right)^{1/3}.$$

Since f is continuous and the Y-paths are cadlag, one also obtains

$$\lim_{\epsilon\to 0} \sigma_\epsilon^2(t) = \lim_{\epsilon\to 0}\Big\{n_\epsilon(t)\, E\big(X_{\epsilon,1}^2\big) + 2(n_\epsilon(t) - 1)\, E\big(X_{\epsilon,1} X_{\epsilon,2}\big)\Big\} = \frac{t\, r}{b}$$

where

$$r = E\Big[\Big(\int_b^{2b} f\Big(Y_s - Y_{s-b}\Big)\,ds\Big)^2\Big] + 2\,E\Big[\int_b^{2b} f\Big(Y_s - Y_{s-b}\Big)\,ds \int_{2b}^{3b} f\Big(Y_s - Y_{s-b}\Big)\,ds\Big].$$

Hence, if $r > 0$, then $\lim_{\epsilon\to 0} d_W\left(\frac{V_\epsilon(t)}{\sigma_\epsilon(t)}, Z\right) = 0$. Since

$$\Big|Z^*_\epsilon(t) - V_\epsilon(t)\Big| \le 2\,a\,b\,\sqrt{\epsilon}, \tag{5.1}$$

it follows that

$$Z^*_\epsilon(t) \overset{dist}{\longrightarrow} \sqrt{\frac{t\,r}{b}}\, Z \sim \sqrt{\frac{r}{b}}\, W_t, \qquad \text{as } \epsilon \to 0,$$

where W is a standard Brownian motion. Moreover, with exactly the same argument, one also obtains

$$\big(Z^*_\epsilon(t_1), \dots, Z^*_\epsilon(t_k)\big) \overset{dist}{\to} \sqrt{\frac{r}{b}}\, \big(W_{t_1}, \dots, W_{t_k}\big) \tag{5.2}$$

for all $k \ge 1$ and all $0 \le t_1 < t_2 < \dots < t_k$. Finally,

$$Z^*_\epsilon \overset{dist}{\to} \sqrt{\frac{r}{b}}\, W \qquad \text{in the space } C([0,\infty), \mathbb{R}). \tag{5.3}$$

We just give a sketch of the proof of (5.3). Let D be the space of real cadlag functions on $[0,\infty)$ endowed with the Skorohod topology. First, one proves that

$$E\big[\big(V_\epsilon(s) - V_\epsilon(t)\big)^4\big] \le \alpha\,\epsilon^2 \left(\lfloor\frac{t}{\epsilon\lambda(\epsilon)}\rfloor - \lfloor\frac{s}{\epsilon\lambda(\epsilon)}\rfloor\right)^2$$

for all $0 \le s < t$, all $\epsilon > 0$, and some constant α. Based on Nourdin and Nualart (2020, Lemma 3.1), this and the finite-dimensional convergence following from (5.1) and (5.2) imply $V_\epsilon \overset{dist}{\to} \sqrt{\frac{r}{b}}\, W$ in the space D. Because of (5.1), one also obtains $Z^*_\epsilon \overset{dist}{\to} \sqrt{\frac{r}{b}}\, W$ in the space D. Finally, (5.3) follows since Z^*_ϵ and $\sqrt{\frac{r}{b}}\, W$ have continuous paths.

Our last example may be useful as regards the CLT for high dimensional data.

Example 5.4. For $i = 1, \dots, N$, let

$$X_i = (X_{i,1}, \dots, X_{i,p})$$

be a p-dimensional random vector. Suppose:

(i) The vectors $X_1, \dots, X_N$ are m-dependent and $X_{i,j} \in L_\infty$ for all i, j;
(ii) $E(X_{i,j}) = 0$ and $E\big(X_{i,j}X_{h,k}\big) = 0$ for all i, j, h, k with $j \ne k$;
(iii) $\sigma_j^2 = E\big[\big(\sum_{i=1}^N X_{i,j}\big)^2\big] > 0$ for all $j = 1, \dots, p$.

Define

$$Y = \sum_{i=1}^N \sum_{j=1}^p \frac{a_j X_{i,j}}{\sigma_j},$$

where the a_j are constants satisfying $\sum_{j=1}^p a_j^2 = 1$, and note that $\mathrm{Var}(Y) = 1$. Upper bounds for $d_W(Y,Z)$ allow to estimate the goodness of the normal approximation for the distribution of Y. For instance, they are involved in the study of the dependence graph of high-dimensional time series; see Chang et al. (2024) and references therein. Under conditions (i)-(iii), Corollary 3.2 yields

$$d_W(Y,Z) \le 40\left(m\,\sqrt{p}\,\max_{i,j} \frac{\|X_{i,j}\|_\infty}{\sigma_j}\right)^{1/3}.$$

This can be compared to the related estimate in Chang et al. (2024, Corollary 1) (which is for the Kolmogorov distance, and among other differences includes a different power $m^{2/3}$).

6. Final comment: beyond (m_n)-dependence

We close with a result which enlarges the scope of Theorem 3.1. It is motivated by the following (natural) question. Let $(X_{n,i})$ be an arbitrary array of real random variables. Under what conditions $(X_{n,i})$ can be approximated by a (m_n)-dependent array? Sometimes, this approximation is possible. As suggested by an anonymous referee, for instance, it is actually possible if $(X_{n,i})$ satisfies a suitable mixing condition or some form of physical dependence. Generally, however, the approximation of $(X_{n,i})$ by a (m_n)-dependent array requires strong conditions. Therefore, we focus on a related problem, that is, we look for a version of Theorem 3.1 where $(X_{n,i})$ is not required to be (m_n)-dependent. To this end, we need some notation. Define

$$W_{n,i} = E\big(X_{n,i} + X_{n,i+1} \mid \mathcal{F}_{n,i}\big) - E\big(X_{n,i} + X_{n,i+1} \mid \mathcal{F}_{n,i-1}\big)$$

where $\mathcal{F}_{n,i} = \sigma(X_{n,1}, \ldots, X_{n,i})$ and $\mathcal{F}_{n,0}$ is the trivial σ-field. Define also

$$\gamma_n = \frac{1}{\sigma_n} \max_i \|X_{n,i}\|_\infty, \quad a_n^2 = E\Big[\Big(\sum_{i=2}^{N_n} E(X_{n,i} \mid \mathcal{F}_{n,i-2})\Big)^2\Big], \quad w_n^2 = \sum_{i=1}^{N_n-1} E\big(W_{n,i}^2\big).$$

Proposition 6.1. Suppose:

- $(X_{n,i})$ satisfies condition (1.2) and $X_{n,i} \in L_\infty$ for all n and i;
- There are constants α and β such that

$$\sigma_n \le \alpha\, w_n \quad \text{and} \quad \Big|\sum_{1 \le i < j < N_n} \operatorname{Cov}\big(W_{n,i}^2,\, W_{n,j}^2\big)\Big| \le \beta\, \gamma_n^2 \sigma_n^4$$

for all $n \ge 1$. Then, there is a constant q (independent of n) such that

$$d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) \le q\,\Big(\frac{a_n}{\sigma_n} + \gamma_n^{1/3}\Big) \qquad \text{for all } n \ge 1.$$

Proof: Letting

$$A_n = \sum_{i=2}^{N_n} E\big(X_{n,i} \mid \mathcal{F}_{n,i-2}\big), \quad W_n = \sum_{i=1}^{N_n-1} W_{n,i}, \quad L_n = X_{n,N_n} - E\big(X_{n,N_n} \mid \mathcal{F}_{n,N_n-1}\big),$$

one obtains

$$S_n = A_n + W_n + L_n.$$

Note also that $\big(W_{n,i} : 1 \le i < N_n\big)$ is a martingale difference sequence, and thus

$$E(W_n^2) = w_n^2.$$

Hence,

$$\begin{aligned} d_W\Big(\frac{S_n}{\sigma_n},\, Z\Big) &\le d_W\Big(\frac{S_n}{\sigma_n},\, \frac{A_n + W_n}{\sigma_n}\Big) + d_W\Big(\frac{A_n + W_n}{\sigma_n},\, \frac{W_n}{\sigma_n}\Big) + d_W\Big(\frac{W_n}{\sigma_n},\, \frac{W_n}{w_n}\Big) + d_W\Big(\frac{W_n}{w_n},\, Z\Big) \\ &\le 2\,\gamma_n + \frac{a_n}{\sigma_n} + \Big|1 - \frac{w_n}{\sigma_n}\Big| + d_W\Big(\frac{W_n}{w_n},\, Z\Big). \end{aligned}$$

Since $\big(W_{n,i} : 1 \le i < N_n\big)$ is a martingale difference sequence and

$$\max_i |W_{n,i}| \le 4 \max_i \|X_{n,i}\|_\infty \le 4\,\sigma_n\,\gamma_n \le 4\,\alpha\, w_n\, \gamma_n \qquad \text{a.s.},$$

the arguments of Lemmas 3.7 and 3.8 can be applied to $Y_i = W_{n,i}$ (with σ_n replaced by w_n). Therefore, $d_W\big(\frac{W_n}{w_n},\, Z\big) \le q^*\, \gamma_n^{1/3}$ for some constant q^* that depends on α and β (but nothing else).

In addition,

$$\begin{aligned}\Big|1-\frac{w_n^2}{\sigma_n^2}\Big| &= \frac{1}{\sigma_n^2}\Big|E\big[(A_n+W_n+L_n)^2\big]-E(W_n^2)\Big| \\ &= \frac{1}{\sigma_n^2}\Big|E\big[(A_n+L_n)^2\big]+2\,E\big[W_n(A_n+L_n)\big]\Big| \\ &\le \frac{2}{\sigma_n^2}\Big\{E(A_n^2)+E(L_n^2)+w_n\sqrt{E(A_n^2)}+w_n\sqrt{E(L_n^2)}\Big\} \\ &= \frac{2}{\sigma_n^2}\Big\{a_n^2+E(L_n^2)+w_n\left(a_n+\sqrt{E(L_n^2)}\right)\Big\},\end{aligned}$$

so that

$$\Big|1-\frac{w_n}{\sigma_n}\Big| = \frac{\Big|1-\frac{w_n^2}{\sigma_n^2}\Big|}{1+\frac{w_n}{\sigma_n}} \le 2\left(\frac{a_n^2}{\sigma_n^2}+\frac{a_n}{\sigma_n}+4\,\gamma_n^2+2\,\gamma_n\right).$$

Collecting all these facts together,

$$d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big) \le 3\,\frac{a_n}{\sigma_n}+2\,\frac{a_n^2}{\sigma_n^2}+8\,\gamma_n^2+6\,\gamma_n+q^*\,\gamma_n^{1/3}.$$

Hence, with $q=14+q^*$, if $\frac{a_n}{\sigma_n}\le 1$ and $\gamma_n\le 1$, one obtains

$$d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big) \le 5\,\frac{a_n}{\sigma_n}+(14+q^*)\,\gamma_n^{1/3} \le q\left(\frac{a_n}{\sigma_n}+\gamma_n^{1/3}\right), \tag{6.1}$$

and otherwise (6.1) is trivial since $d_W\Big(\frac{S_n}{\sigma_n},\,Z\Big)\le\sqrt{2}$ by Lemma 2.1. $\square$

It is worth noting that Proposition 6.1 deviates from some analogous results available in the literature (such as Cuny and Merlevède (2015) and Shao (1993)) for it does not require either stationarity-mixing assumptions or martingale assumptions. Furthermore, Proposition 6.1 provides a quantitative bound as well.

References

Bergström, H. A comparison method for distribution functions of sums of independent and dependent random variables. *Teor. Verojatnost. i Primenen.*, **15**, 442–468 (1970). In Russian; English translation in *Theor. Probab. Appl.* **15**, 430–457 (1970). MR283850; see also MR281245.

Berk, K. N. A central limit theorem for m-dependent random variables with unbounded m. *Ann. Probab.*, **1**, 352–354 (1973). MR350815.

Bradley, R. C. *Introduction to Strong Mixing Conditions. Vol. 1–3.* Kendrick Press, Heber City, UT (2007). ISBN 0-9740427-6-5; 0-9740427-7-3; 0-9740427-8-1. MR2325294–MR2325296.

Breuer, P. and Major, P. Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.*, **13** (3), 425–441 (1983). MR716933.

Campese, S., Nourdin, I., and Nualart, D. Continuous Breuer-Major theorem: tightness and nonstationarity. *Ann. Probab.*, **48** (1), 147–177 (2020). MR4079433.

Chang, J., Chen, X., and Wu, M. Central limit theorems for high dimensional dependent data. *Bernoulli*, **30** (1), 712–742 (2024). MR4665595.

Chen, L. H. Y., Goldstein, L., and Shao, Q.-M. *Normal approximation by Stein's method.* Probability and its Applications (New York). Springer, Heidelberg (2011). ISBN 978-3-642-15006-7. MR2732624.

Cuny, C. and Merlevède, F. Strong invariance principles with rate for "reverse" martingale differences and applications. *J. Theoret. Probab.*, **28** (1), 137–183 (2015). MR3320963.

DasGupta, A. *Asymptotic theory of statistics and probability.* Springer Texts in Statistics. Springer, New York (2008). ISBN 978-0-387-75970-8. MR2664452.

Dedecker, J., Merlevède, F., and Rio, E. Rates of convergence in the central limit theorem for martingales in the non stationary setting. *Ann. Inst. Henri Poincaré Probab. Stat.*, **58** (2), 945–966 (2022). MR4421614.

Dianandа, P. H. The central limit theorem for m-dependent variables. *Proc. Cambridge Philos. Soc.*, **51**, 92–95 (1955). MR67396.

Fan, X. and Ma, X. On the Wasserstein distance for a martingale central limit theorem. *Statist. Probab. Lett.*, **167**, 108892, 6 (2020). MR4138415.

Haeusler, E. and Joos, K. A nonuniform bound on the rate of convergence in the martingale central limit theorem. *Ann. Probab.*, **16** (4), 1699–1720 (1988). MR958211.

Heyde, C. C. and Brown, B. M. On the departure from normality of a certain class of martingales. *Ann. Math. Statist.*, **41**, 2161–2165 (1970). MR293702.

Hoeffding, W. and Robbins, H. The central limit theorem for dependent random variables. *Duke Math. J.*, **15**, 773–780 (1948). MR26771.

Nourdin, I. and Nualart, D. The functional Breuer-Major theorem. *Probab. Theory Related Fields*, **176** (1-2), 203–218 (2020). MR4055189.

Orey, S. A central limit theorem for m-dependent random variables. *Duke Math. J.*, **25**, 543–546 (1958). MR97841.

Peligrad, M. On the asymptotic normality of sequences of weak dependent random variables. *J. Theoret. Probab.*, **9** (3), 703–715 (1996). MR1400595.

Pratelli, L. and Rigo, P. Convergence in Total Variation to a Mixture of Gaussian Laws. *Mathematics*, **6** (2018). DOI: 10.3390/math6060099.

Rio, E. About the Lindeberg method for strongly mixing sequences. *ESAIM Probab. Statist.*, **1**, 35–61 (1995). MR1382517.

Röllin, A. On quantitative bounds in the mean martingale central limit theorem. *Statist. Probab. Lett.*, **138**, 171–176 (2018). MR3788734.

Romano, J. P. and Wolf, M. A more general central limit theorem for m-dependent random variables with unbounded m. *Statist. Probab. Lett.*, **47** (2), 115–124 (2000). MR1747098.

Shao, Q. M. Almost sure invariance principles for mixing sequences of random variables. *Stochastic Process. Appl.*, **48** (2), 319–334 (1993). MR1244549.

Utev, S. A. Central limit theorem for dependent random variables. In *Probability theory and mathematical statistics, Vol. II (Vilnius, 1989)*, pp. 519–528. Mokslas, Vilnius (1990a). ISBN 90-6764-129-4. MR1153906.

Utev, S. A. On the central limit theorem for φ-mixing arrays of random variables. *Theory Probab. Appl.*, **35** (1), 131–139 (1990b). MR1050059.

Van Dung, L., Son, T. C., and Tien, N. D. L_1 bounds for some martingale central limit theorems. *Lith. Math. J.*, **54** (1), 48–60 (2014). MR3189136.