On the Wassertein distance for a martingale central limit theorem

Xiequan Fan*, Xiaohui Ma

Center for Applied Mathematics, Tianjin University, 300072 Tianjin, China

---

Abstract

We prove an upper bound on the Wassertein distance between normalized martingales and the standard normal random variable, which extends a result of Röllin [7]. The proof is based on a method of Bolthausen [1].

Keywords: Martingales; Central limit theorem; Wassertein metric

2000 MSC: 60G42; 60E15; 60F25

---

1. Introduction and main result

Let $n \geq 1$. Assume that $\mathbf{X} = (X_i)_{1 \leq i \leq n}$ is a sequence of martingale differences defined on probability space $(\Omega, \mathcal{F}, \mathbf{P})$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and that X_i is adapted to the σ-fields $\mathcal{F}_i, i = 1, 2, \ldots, n$, where $\mathcal{F}_i$ is the σ-algebra generated by $X_1, X_2, \ldots, X_i$. In other words, X_i satisfies $\mathbf{E}[X_i|\mathcal{F}_{i-1}] = 0$. Let M_n denote the class of all such sequences of length n. If $\mathbf{X} \in M_n$, write

$$S_n = X_1 + X_2 + \ldots + X_n,$$

$$\sigma_i^2 = \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right], i = 1, 2, \ldots, n,$$

$$s_n^2 = \sum_{i=1}^{n} \mathbf{E}X_i^2,$$

$$V_n^2 = \frac{\sum_{i=1}^{n} \sigma_i^2}{s_n^2},$$

$$\|\mathbf{X}\|_p = \max_{1 \leq i \leq n} \|X_i\|_p, \quad p \in [1, +\infty],$$

where $\|X_i\|_p = (\mathbf{E}|X_i|^p)^{1/p}$ for $1 \leq p < \infty$ and $\|X_i\|_\infty = \inf\{a : \mathbf{P}(|X_i| \leq a) = 1\}$.

According to the martingale central limit theory, it is well known that if, as $n \to \infty$,

$$V_n^2 \xrightarrow{\mathbf{P}} 1$$

---

*Corresponding author.
E-mail: fanxiequan@hotmail.com (X. Fan).

and the conditional Lindeberg condition

$$\sum_{i=1}^{n} \mathbf{E}\left[X_i^2 \mathbf{1}_{\{|X_i|\geq\varepsilon\}}|\mathcal{F}_{i-1}\right] \xrightarrow{\mathbf{P}} 0 \quad \text{for each} \ \ \varepsilon > 0$$

are satisfied, then S_n/s_n converges to the standard Gaussian random variable in distribution, that is,

$$\forall t \in \mathbf{R}, \quad \mathbf{P}\left(S_n/s_n \leq t\right) \longrightarrow \Phi(t), \quad \text{as} \ \ n \to \infty,$$

where $\Phi(t) = (2\pi)^{-1/2} \int_{-\infty}^{t} e^{-x^2/2} dx$. Define

$$D\left(S_n/s_n\right) = \sup_{t} |\mathbf{P}(S_n/s_n \leq t) - \Phi(t)|,$$

then

$$D\left(S_n/s_n\right) \longrightarrow 0, \quad \text{as} \ \ n \to \infty.$$

The Kolmogorov distance in central limit theorem for martingales has been intensely studied under various conditions. For instance, we recall the following result due to Heyde and Brown [12]: for a constant $p \in (1, 2]$, they proved that

$$D\left(S_n/s_n\right) \leq C_p \left(s_n^{-4p} \left\|V_n^2 - 1\right\|_p^p + s_n^{-2p} (\sum_{i=1}^{n} \mathbf{E}\left|X_i\right|^{2p}) \right)^{1/(2p+1)}, \tag{1}$$

where, here and after, C_p is a constant depending only on p. Bolthausen [1] proved that if $\|\mathbf{X}\|_\infty \leq \gamma$ a.s., then

$$D\left(S_n/s_n\right) \leq C_\gamma \left(\frac{n\log n}{s_n^3} + \min\left\{ \left\|V_n^2 - 1\right\|_\infty^{1/2}, \left\|V_n^2 - 1\right\|_1^{1/3} \right\} \right). \tag{2}$$

Using a modification of the method developed by Bolthausen, Haeusler [10] gave an extension of (1) to all $p > 1$; El Machkouri and Ouchti [8] replaced the term $n\log n/s_n^3$ of (2) by $\max_{1\leq i\leq n} \gamma_i \log n/s_n$ to a large class of martingale difference sequences satisfying

$$\mathbf{E}\left[|X_i|^3|\mathcal{F}_{i-1}\right] \leq \gamma_i \mathbf{E}\left[|X_i|^2|\mathcal{F}_{i-1}\right] a.s., \tag{3}$$

where γ_i, $i = 1, 2, \ldots, n$ are constants. Following Bolthausen again, Mourrat [2] has extended the term $\min\left\{\left\|V_n^2 - 1\right\|_\infty^{1/2}, \left\|V_n^2 - 1\right\|_1^{1/3}\right\}$ of (2) to the more general term $\left(\left\|V_n^2 - 1\right\|_p^p + s_n^{-2p}\right)^{1/(2p+1)}$, for $p \geq 1$. Recently, with the methods of Grama and Hausler [9] and Bolthausen [1] (see also Fan et al. [4, 5]), Fan [3] proved that if there exist two positive numbers ρ and γ, such that

$$\mathbf{E}\left[|X_i|^{2+\rho}|\mathcal{F}_{i-1}\right] \leq \gamma^\rho \mathbf{E}\left[|X_i|^2|\mathcal{F}_{i-1}\right] a.s. \tag{4}$$

for all $i = 1, 2, \ldots, n$, then for $p \geq 1$,

$$D\left(S_n/s_n\right) \leq C_{p,\gamma,\rho} \left(\alpha_n + \left(\left\|V_n^2 - 1\right\|_p^p + \frac{1}{s_n^{2p}} \mathbf{E} \max_{1\leq i\leq n} |X_i|^{2p} \right)^{1/(2p+1)} \right), \tag{5}$$

2

where

$$\alpha_n = \begin{cases} \frac{1}{s_n^\rho} & \text{if } \rho \in (0,1), \\[2mm] \frac{1}{s_n} \log s_n & \text{if } \rho \geq 1. \end{cases}$$

Despite the Kolmogorov distance in central limit theorem has been intensely studied, research of bounds with respect to the Wassertein distance are rare. To the best of our knowledge, we only aware the articles of Van Dung et al. [6] and Röllin [7]. Denote by $d_w(S_n/s_n) = \int_{-\infty}^{+\infty} |\mathbf{P}(S_n/s_n \leq x) - \Phi(x)| dx$ the Wassertein distance between the distributions of S_n/s_n and the standard normal random variable. Van Dung et al. [6] extended Mourrat's bounds to the L^1-bound in the mean central limit theorem, that is, if $\|\mathbf{X}\|_\infty \leq \gamma$ a.s., then for $p > 1/2$,

$$d_w(S_n/s_n) \leq C_{p,\gamma} \left(\frac{n \log n}{s_n^3} + \left(\|V_n^2 - 1\|_p^p + s_n^{-2p} \right)^{1/2p} \right). \tag{6}$$

Röllin [7] provided a new proof of already known result by using combination of both Lindeberg's and Stein's methods. His result states that if $V_n^2 = 1$ a.s., then for any $a \geq 0$,

$$d_w(S_n/s_n) \leq \frac{3}{s_n} \sum_{i=1}^n \mathbf{E} \frac{|X_i|^3}{\rho_i^2 + a^2} + \frac{2a}{s_n}, \tag{7}$$

where $\rho_i^2 = \sum_{k=i}^n \sigma_k^2$. The aim of this article is to extend (7) by relaxing the condition $V_n^2 = 1$ a.s. to $\|V_n^2 - 1\|_p \longrightarrow 0$ for any $p \geq 1$.

The following theorem is our main result.

Theorem 1.1. *For any $p \geq 1$, there exists a constant $C_p > 0$ such that*

$$d_w(S_n/s_n) \leq C_p \left(\frac{1}{s_n} \Big(\sum_{i=1}^n \mathbf{E} |X_i|^3 \Big)^{1/3} + \left(\|V_n^2 - 1\|_p^p + \frac{1}{s_n^{2p}} \mathbf{E} \max_{1 \leq i \leq n} |X_i|^{2p} \right)^{1/2p} \right). \tag{8}$$

Notice that if $\|\mathbf{X}\|_\infty \leq \gamma$ a.s., then the order of the term $\left(\|V_n^2 - 1\|_p^p + \frac{1}{s_n^{2p}} \mathbf{E} \max_{1 \leq i \leq n} |X_i|^{2p} \right)^{1/2p}$ is less than that of $\left(\|V_n^2 - 1\|_p^p + s_n^{-2p} \right)^{1/2p}$. Thus the second term of bound (8) implies the second term of bound (6).

Let $p = 3/2$. The following corollary is an immediate consequence of Theorem 1.1.

Corollary 1.1. *There exists a constant $C > 0$ such that*

$$d_w(S_n/s_n) \leq C \left(\frac{1}{s_n} \Big(\sum_{i=1}^n \mathbf{E} |X_i|^3 \Big)^{1/3} + \|V_n^2 - 1\|_{3/2}^{1/2} \right). \tag{9}$$

Though our bound may not be smaller than the bound of Röllin (cf. (7)), the advantage of (9) is that the term ρ_i^2 appearing in (7) does not appear any more. Moreover, inequality (9) gives a rate of convergence under weaker condition, that is the condition $V_n^2 = 1$ a.s. has been relaxed to $\|V_n^2 - 1\|_{3/2}^{1/2} \longrightarrow 0$.

2. Proof of Theorem 1.1

In the proof of the Theorem 1.1, we shall use the following lemma which is a consequence of Röllin's inequality (7) by dropping ρ_i^2.

Lemma 2.1. *Assume that $V_n^2 = 1$ a.s., there exists a constant $C > 0$ such that*

$$d_w(S_n/s_n) \le \frac{C}{s_n} \left(\sum_{i=1}^n \mathbf{E}\,|X_i|^3 \right)^{1/3}. \tag{10}$$

Proof of Lemma 2.1. From (7), by the fact $\rho_i^2 \ge 0$, it is easy to see that

$$d_w(S_n/s_n) \le \frac{3}{s_n} \sum_{i=1}^n \frac{\mathbf{E}\,|X_i|^3}{a^2} + \frac{2a}{s_n}. \tag{11}$$

Let $f(a) = \sum_{i=1}^n \frac{\mathbf{E}|X_i|^3}{a^2} + 2a$. Clearly if $a = \left(\sum_{i=1}^n \mathbf{E}\,|X_i|^3 \right)^{1/3}$, then $f(a)$ achieves the minimum, which gives the desired inequality (10). $\square$

Now, we are in position to prove Theorem 1.1. The idea is to construct a new sequence of martingale differences $(\hat{X}_i, \hat{\mathcal{F}}_i)_{1 \le i \le N}$ based on $\mathbf{X}$ such that $\sum_{i=1}^N \mathbf{E}[\hat{X}_i^2|\hat{\mathcal{F}}_{i-1}]/s_n^2 = 1$ a.s., where the exact value of N is given later, and then apply Lemma 2.1 to $\hat{\mathbf{X}}$. Consider the stopping time

$$\tau = \sup \left\{ 1 \le k \le n : \sum_{i=1}^k \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1} \right] \le s_n^2 \right\}.$$

Assume that $\varepsilon > 0$. Let $r = \left\lfloor \frac{s_n^2 - \sum_{i=1}^\tau \mathbf{E}[X_i^2|\mathcal{F}_{i-1}]}{\varepsilon^2} \right\rfloor$, where $\lfloor x \rfloor$ stands for the largest integer not exceeding x. Clearly $r \le \lfloor s_n^2/\varepsilon^2 \rfloor$. Let $N = n + \lfloor s_n^2/\varepsilon^2 \rfloor + 1$. Conditionally on $\mathcal{F}_\tau$, and for $\tau + 1 \le i \le \tau + r$, let Y_i be independent random variables such that $\mathbf{P}(Y_i = \pm\varepsilon) = 1/2$, with Y_i independent of $(X_i)_{i \ge \tau+1}$. When $i = \tau + r + 1$, let $Y_{\tau+r+1}$ be such that

$$\mathbf{P}\left(Y_{\tau+r+1} = \pm \left(s_n^2 - \sum_{i=1}^\tau \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1} \right] - r\varepsilon^2 \right)^{1/2} \right) = \frac{1}{2},$$

with the sign determined independent of everything else. Finally, if $\tau + r + 1 < i \le N$, then we let $Y_i = 0$. By the definition of $\hat{X}_i$, we have

$$\hat{X}_i = X_i \mathbf{1}_{\{i \le \tau\}} + Y_i \mathbf{1}_{\{\tau < i \le \tau+r\}} + Y_{\tau+r+1} \mathbf{1}_{\{i=\tau+r+1\}}, \quad 1 \le i \le N.$$

Define $\hat{\mathcal{F}}_i = \mathcal{F}_i$ for $i \le \tau$, $\hat{\mathcal{F}}_i = \sigma\{\mathcal{F}_i,\ Y_j, \tau + 1 \le j \le i\}$ for $\tau + 1 \le i \le N$. Notice that $\mathbf{1}_{\{\tau < i \le \tau+r\}}$ and $\mathbf{1}_{\{i=\tau+r+1\}}$ are $\hat{\mathcal{F}}_{i-1}$−measurable for all $\tau < i \le N$. Then we deduce that

$$\begin{aligned}
\mathbf{E}[\hat{X}_i|\hat{\mathcal{F}}_{i-1}] &= \mathbf{E}[X_i \mathbf{1}_{\{i \le \tau\}}|\hat{\mathcal{F}}_{i-1}] + \mathbf{E}[Y_i \mathbf{1}_{\{\tau < i \le \tau+r\}}|\hat{\mathcal{F}}_{i-1}] + \mathbf{E}[Y_{\tau+r+1}\mathbf{1}_{\{i=\tau+r+1\}}|\hat{\mathcal{F}}_{i-1}] \\
&= \mathbf{E}[X_i \mathbf{1}_{\{i \le \tau\}}|\hat{\mathcal{F}}_{i-1}] + \mathbf{1}_{\{\tau < i \le \tau+r\}}\mathbf{E}[Y_i|\hat{\mathcal{F}}_{i-1}] + \mathbf{1}_{\{i=\tau+r+1\}}\mathbf{E}[Y_{\tau+r+1}|\hat{\mathcal{F}}_{i-1}].
\end{aligned}$$

Since $\mathbf{1}_{\{i\le\tau\}} = 1 - \mathbf{1}_{\{\tau\le i-1\}}$ is $\hat{\mathcal{F}}_{i-1}$−measurable and $\hat{\mathcal{F}}_i = \mathcal{F}_i$ for $i \le \tau$, we have

$$\mathbf{E}[X_i\mathbf{1}_{\{i\le\tau\}}|\hat{\mathcal{F}}_{i-1}] = \mathbf{1}_{\{i\le\tau\}}\mathbf{E}[X_i|\hat{\mathcal{F}}_{i-1}] = \mathbf{1}_{\{i\le\tau\}}\mathbf{E}[X_i|\mathcal{F}_{i-1}] = 0.$$

For $\tau < i \le \tau + r$, by the definitions of Y_i and $\hat{\mathcal{F}}_{i-1}$, we get

$$\mathbf{E}[Y_i|\hat{\mathcal{F}}_{i-1}] = \mathbf{E}[Y_i|\mathcal{F}_\tau] = \varepsilon\mathbf{P}(Y_i = \varepsilon|\mathcal{F}_\tau) - \varepsilon\mathbf{P}(Y_i = -\varepsilon|\mathcal{F}_\tau) = 0.$$

Similarly, for $i = \tau + r + 1$, it holds

$$
\begin{aligned}
\mathbf{E}[Y_{\tau+r+1}|\hat{\mathcal{F}}_{i-1}] &= \mathbf{E}[Y_{\tau+r+1}|\mathcal{F}_\tau] \\
&= \frac{1}{2}\left(s_n^2 - \sum_{i=1}^\tau \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right] - r\varepsilon^2\right)^{1/2} - \frac{1}{2}\left(s_n^2 - \sum_{i=1}^\tau \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right] - r\varepsilon^2\right)^{1/2} \\
&= 0.
\end{aligned}
$$

Thus $\mathbf{E}[\hat{X}_i|\hat{\mathcal{F}}_{i-1}] = 0$, which implies that $\hat{\mathbf{X}} = (\hat{X}_i, \hat{\mathcal{F}}_i)_{1\le i\le N}$ is also a martingale difference sequence. Moreover, it holds

$$\sum_{i=\tau+1}^N \mathbf{E}[\hat{X}_i^2|\hat{\mathcal{F}}_{i-1}] = \sum_{i=\tau+1}^{\tau+r} \varepsilon^2 + s_n^2 - \sum_{i=1}^\tau \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right] - r\varepsilon^2 = s_n^2 - \sum_{i=1}^\tau \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right],$$

which implies that

$$\sum_{i=1}^N \mathbf{E}[\hat{X}_i^2|\hat{\mathcal{F}}_{i-1}] = s_n^2 \qquad a.s.$$

Consequently, $\sum_{i=1}^N \mathbf{E}[\hat{X}_i^2|\hat{\mathcal{F}}_{i-1}]/s_n^2 = 1$ a.s., by Lemma 2.1 it is easy to see that:

$$d_w(\hat{S}_N/s_n) \le \frac{C}{s_n}\left(\sum_{i=1}^N \mathbf{E}|\hat{X}_i|^3\right)^{1/3}. \tag{12}$$

According to the construction of $\hat{\mathbf{X}}$, we can easily give an estimation of $\sum_{i=1}^N \mathbf{E}|\hat{X}_i|^3$ as following:

$$\sum_{i=1}^N \mathbf{E}|\hat{X}_i|^3 = \sum_{i=1}^\tau \mathbf{E}\left|X_i\right|^3 + \sum_{i=\tau+1}^N \mathbf{E}|\hat{X}_i|^3 \le \sum_{i=1}^n \mathbf{E}\left|X_i\right|^3 + \left(1 + s_n^2/\varepsilon^2\right)\varepsilon^3.$$

Thus we get,

$$d_w(\hat{S}_N/s_n) \le \frac{C}{s_n}\left(\sum_{i=1}^n \mathbf{E}\left|X_i\right|^3 + \varepsilon^3 + s_n^2\varepsilon\right)^{1/3}. \tag{13}$$

For any $x > 0$, we have

$$
\begin{aligned}
d_w(S_n/s_n) &\le d_w(\hat{S}_N/s_n) + \int_{-\infty}^{+\infty}\left|\mathbf{P}(\hat{S}_N/s_n \le t+x) - \mathbf{P}(S_n/s_n \le t)\right|dt + \int_{-\infty}^{+\infty}\left|\Phi(t+x) - \Phi(t)\right|dt \\
&\le d_w(\hat{S}_N/s_n) + \frac{C}{x^{2p-1}s_n^{2p}}\mathbf{E}|\hat{S}_N - S_n|^{2p} + 2x, \tag{14}
\end{aligned}
$$

5

see (5.3) of Van Dung et al. [6]. We give an estimation for the second term in the right-hand side of (14). First we note that

$$S_n - \hat{S}_N = \sum_{i=\tau+1}^{N} (X_i - \hat{X}_i),$$

where we put $X_i = 0$ for $i > n$. As τ is a stopping time, conditionally on $\mathcal{F}_\tau$, the $(X_i - \hat{X}_i)_{\tau+1 \leq i \leq N}$ still forms a martingale difference sequence. Indeed, the following equalities hold

$$\mathbf{E}[(X_{\tau+1} - \hat{X}_{\tau+1})|\hat{\mathcal{F}}_\tau] = \mathbf{E}[(X_{\tau+1} - \hat{X}_{\tau+1})|\mathcal{F}_\tau] = \mathbf{E}[X_{\tau+1}|\mathcal{F}_\tau] - \mathbf{E}[Y_{\tau+1}|\mathcal{F}_\tau] = 0$$

and, for $2 \leq i \leq N - \tau$,

$$\mathbf{E}[(X_{\tau+i} - \hat{X}_{\tau+i})|\sigma\{\mathcal{F}_\tau, ..., \mathcal{F}_{\tau+i-1}, Y_{\tau+1}, ..., Y_{\tau+i-1}\}] = \mathbf{E}[X_{\tau+i}|\mathcal{F}_{\tau+i-1}] - \mathbf{E}[Y_{\tau+i}|\mathcal{F}_\tau] = 0,$$

where the last line follows by the fact that, conditionally on $\mathcal{F}_\tau$, $X_{\tau+i}$ are independent of $(Y_i)_{\tau+1 \leq i \leq N}$ and $(Y_i)_{\tau+1 \leq i \leq N}$ are independent random variables. Hence $(X_i - \hat{X}_i, \hat{\mathcal{F}}_i)_{1 \leq i \leq N}$ is also a martingale difference sequence. By Theorem 2.11 in Hall and Heyde [11], we deduce that

$$\frac{1}{C_p} \mathbf{E}\left[|\hat{S}_N - S_n|^{2p}\right] \leq \mathbf{E}\left[\left(\sum_{i=\tau+1}^{N} \mathbf{E}\left[(X_i - \hat{X}_i)^2 \Big| \hat{\mathcal{F}}_{i-1}\right]\right)^p\right] + \mathbf{E}\left[\max_{\tau+1 \leq i \leq N} \left|X_i - \hat{X}_i\right|^{2p}\right]$$

$$= \mathbf{E}\left[\left(\sum_{i=\tau+1}^{N} \mathbf{E}\left[(X_i - \hat{X}_i)^2 \Big| \mathcal{F}_{i-1}\right]\right)^p\right] + \mathbf{E}\left[\max_{\tau+1 \leq i \leq N} \left|X_i - \hat{X}_i\right|^{2p}\right], \qquad (15)$$

where the last line follows by the fact that when $\tau + 1 \leq i \leq N$, conditionally on $\mathcal{F}_{i-1}$, $X_i - \hat{X}_i$ is independent of $(Y_j)_{\tau+1 \leq j \leq i-1}$. As $\mathbf{E}[X_i \hat{X}_i|\mathcal{F}_{i-1}] = 0$ for $i \geq \tau + 1$, we have

$$\sum_{i=\tau+1}^{N} \mathbf{E}\left[(X_i - \hat{X}_i)^2 \Big| \mathcal{F}_{i-1}\right] = \sum_{i=\tau+1}^{N} \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right] + \sum_{i=\tau+1}^{N} \mathbf{E}\left[\hat{X}_i^2|\mathcal{F}_{i-1}\right]$$

$$= s_n^2 V_n^2 + s_n^2 - 2\sum_{i=1}^{\tau+1} \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right] + 2\mathbf{E}\left[X_{\tau+1}^2|\mathcal{F}_\tau\right]. \qquad (16)$$

Notice that $\sum_{i=1}^{\tau+1} \mathbf{E}\left[X_i^2|\mathcal{F}_{i-1}\right] > s_n^2$. Hence, we get

$$\sum_{i=\tau+1}^{N} \mathbf{E}\left[(X_i - \hat{X}_i)^2 \Big| \mathcal{F}_{i-1}\right] \leq s_n^2 V_n^2 - s_n^2 + 2\mathbf{E}\left[X_{\tau+1}^2|\mathcal{F}_\tau\right]. \qquad (17)$$

Using the inequality $|a + b|^k \leq 2^{k-1}\left(|a|^k + |b|^k\right), k \geq 1$, we have

$$\left(\sum_{i=\tau+1}^{N} \mathbf{E}\left[(X_i - \hat{X}_i)^2 \Big| \mathcal{F}_{i-1}\right]\right)^p \leq 2^{p-1}|s_n^2 V_n^2 - s_n^2|^p + 2^{2p-1}\left(\mathbf{E}\left[X_{\tau+1}^2|\mathcal{F}_\tau\right]\right)^p. \qquad (18)$$

For the second term in the right-hand side of (18), by Jensen's inequality, we obtain

$$\left(\mathbf{E}\left[X_{\tau+1}^2|\mathcal{F}_\tau\right]\right)^p \leq \mathbf{E}\left[|X_{\tau+1}|^{2p}|\mathcal{F}_\tau\right]. \qquad (19)$$

6

Taking expectations on both sides of (18), we deduce that

$$
\mathbf{E}\left[\left(\sum_{i=\tau+1}^{N} \mathbf{E}\left[(X_i - \hat{X}_i)^2 \Big| \mathcal{F}_{i-1}\right]\right)^p\right] \leq 2^{p-1} s_n^{2p} \left\|V_n^2 - 1\right\|_p^p + 2^{2p-1} \mathbf{E}\left[|X_{\tau+1}|^{2p}\right]
$$

$$
\leq 2^{p-1} s_n^{2p} \left\|V_n^2 - 1\right\|_p^p + 2^{2p-1} \mathbf{E} \max_{1 \leq i \leq n} |X_i|^{2p}. \qquad (20)
$$

By an argument similar to that of (18), we get

$$
\mathbf{E}\left[\max_{\tau+1 \leq i \leq N}\left|X_i - \hat{X}_i\right|^{2p}\right] \leq 2^{2p-1} \mathbf{E}\left[\max_{\tau+1 \leq i \leq n} |X_i|^{2p} + \varepsilon^{2p}\right]
$$

$$
\leq 2^{2p-1} \mathbf{E}\left[\max_{1 \leq i \leq n} |X_i|^{2p} + \varepsilon^{2p}\right]. \qquad (21)
$$

Combining the inequalities (15), (20) and (21) together, we have

$$
\mathbf{E}\left[|\hat{S}_N - S_n|^{2p}\right] \leq C_p \left(s_n^{2p} \left\|V_n^2 - 1\right\|_p^p + \mathbf{E} \max_{1 \leq i \leq n} |X_i|^{2p} + \varepsilon^{2p}\right). \qquad (22)
$$

Now we combine this with the inequalities (13) and (14) together, and then let $\varepsilon \to 0$, we get

$$
d_w(S_n/s_n) \leq C \frac{1}{s_n}\left(\sum_{i=1}^{n} \mathbf{E}\,|X_i|^3\right)^{1/3} + C x^{1-2p}\left(\left\|V_n^2 - 1\right\|_p^p + \frac{1}{s_n^{2p}}\mathbf{E} \max_{1 \leq i \leq n} |X_i|^{2p}\right) + 2x. \qquad (23)
$$

Putting $x = \left(\left\|V_n^2 - 1\right\|_p^p + \frac{1}{s_n^{2p}}\mathbf{E}\max_{1 \leq i \leq n} |X_i|^{2p}\right)^{1/2p}$, the final bound follows. $\qquad \square$

References

[1] E. Bolthausen, Exact convergence rates in some martingale central limit theorems, Ann. Probab. 10 (1982) 672-688.

[2] J.C. Mourrat, On the rate of convergence in the martingale central limit theorem, Bernoulli 19 (2) (2013) 633-645.

[3] X. Fan, Exact rates of convergence in some martingale central limit theorems, J. Math. Anal. Appl. 469 (2019) 1028-1044.

[4] X. Fan, I. Grama, Q. Liu, Non-uniform Berry-Esseen bounds for martingales with applications to statistical estimation, Statistics 51 (1) (2017) 105–122.

[5] X. Fan, I. Grama, Q. Liu, Q.M. Shao, Self-normalized Cramr type moderate deviations for martingales, Bernoulli 25 (4A) (2019) 2793-2823.

[6] L. Van Dung, T.C. Son, N.D. Tien, L_1-bounds for some martingale central limit theorems, Lith. Math. J. 54 (1) (2014) 48-60.

[7] A. Röllin, On quantitative bounds in the mean martingale central limit theorem, Statist. Probabil. Lett. 138 (2018) 171-176.

[8] M. El Machkouri, L. Ouchti, Exact convergence rates in the central limit theorem for a class of martingales, Bernoulli 13 (4) (2007) 981-999.

[9] I. Grama, E. Haeusler, Large deviations for martingales via Cramérs method, Stochastic Process. Appl. 85 (2000) 279-293.

[10] E. Haeusler, On the rate of convergence in the central limit theorem for martingales with discrete and continuous time, Ann. Probab. 16 (1) (1988) 275-299.

[11] P. Hall, C.C. Heyde, *Martingale Limit Theory and Its Application*. Academic Press, 1980.

[12] C.C. Heyde, B.M. Brown, On the departure from normality of a certain class of martingales, Ann. Math. Statist. 41 (1970) 2161-2165.