Rates of convergence in the central limit theorem for martingales in the non stationary setting

Jérôme Dedecker*, Florence Merlevède † and Emmanuel Rio ‡

July 13, 2021

Abstract

In this paper, we give rates of convergence, for minimal distances and for the uniform distance, between the law of partial sums of martingale differences and the limiting Gaussian distribution. More precisely, denoting by P_X the law of a random variable X and by G_a the normal distribution $\mathcal{N}(0, a)$, we are interested by giving quantitative estimates for the convergence of $P_{S_n/\sqrt{V_n}}$ to G_1, where S_n is the partial sum associated with either martingale differences sequences or more general dependent sequences, and $V_n = \mathrm{Var}(S_n)$. Applications to linear statistics, non stationary ρ-mixing sequences and sequential dynamical systems are given.

Keywords. Minimal distances, ideal distances, Gaussian approximation, Berry-Esseen type inequalities, martingales, ρ-mixing sequences, sequential dynamical systems.

Mathematics Subject Classification (2020). 60F05, 60G42, 60G48

1 Introduction and Notations

Let $(\xi_i)_{i\in\mathbb{N}}$ denote a sequence of martingale differences in $\mathbb{L}^2$, with respect to the increasing filtration $(\mathcal{F}_i)_{i\in\mathbb{N}}$. Let $M_n = \sum_{k=1}^n \xi_k$ and $V_n = \sum_{k=1}^n \mathbb{E}(\xi_k^2)$. If

$$V_n^{-1/2}\mathbb{E}\left(\max_{1\le i\le n} |\xi_i|\right) \to 0 \ \text{ and } \ V_n^{-1}\sum_{k=1}^n \xi_k^2 \to^{\mathbb{P}} 1 \quad \text{as } n \to \infty, \tag{1.1}$$

*Jérôme Dedecker, Université de Paris, CNRS, MAP5, UMR 8145, 45 rue des Saints-Pères, F-75006 Paris, France.

†Florence Merlevède, LAMA, Univ Gustave Eiffel, Univ Paris Est Créteil, UMR 8050 CNRS, F-77454 Marne-La-Vallée, France.

‡Emmanuel Rio, Université de Versailles, LMV UMR 8100 CNRS, 45 avenue des Etats-Unis, F-78035 Versailles, France.

then $V_n^{-1/2}M_n$ converges in distribution to a standard normal variable (see [14]). Other sets of conditions implying the central limit theorem can be found in [12]. In particular, under the first part of condition (1.1), its second part is implied by

$$V_n^{-1}\langle M\rangle_n \to^{\mathbb{P}} 1 \;\text{ as } n\to\infty, \text{ where } \;\langle M\rangle_n := \sum_{k=1}^n \mathbb{E}(\xi_k^2|\mathcal{F}_{k-1})\,.$$

We are interested in bounds on the speed of convergence in this central limit theorem and in particular by giving upper bounds for the $\mathbb{L}_1$ and $\mathbb{L}_\infty$ distances defined respectively as

$$\Delta_{n,1} := \|F_n-\Phi\|_1 \;\text{ and }\; \Delta_{n,\infty} := \|F_n-\Phi\|_\infty\,, \tag{1.2}$$

where F_n is the cdf of $M_n/\sqrt{V_n}$ and Φ is the cdf of a standard normal variable. Both of these distances have their own interests. For instance, $\Delta_{n,\infty}$ provides useful estimates of the quantile $F_n^{-1}(u)$ of $M_n/\sqrt{V_n}$ when $\min(u,1-u)$ is large enough, whereas the $\mathbb{L}^1$-distance provides estimates of the super quantile (also called the conditional value at risk) as stated in [23, Theorem 2].

Concerning the $\mathbb{L}_\infty$-distance $\Delta_{n,\infty}$ for martingales, several results have been obtained under different kinds of assumptions.

One of the first results is due to Heyde and Brown [13] and can be stated as follows. For $p\in]2,4]$, there exists a positive constant C_p such that for any $n\geq 1$,

$$\Delta_{n,\infty} \leq C_p\Big(\|V_n^{-1}\langle M\rangle_n - 1\|_{p/2}^{p/2} + V_n^{-p/2}\sum_{k=1}^n \mathbb{E}(|\xi_k|^p)\Big)^{1/(p+1)}\,. \tag{1.3}$$

This result has been extended to any $p\in(2,\infty)$ by Haeusler [11]. See also Mourrat [19] for an improvement of (1.3) in the bounded case. If the conditional variances are constant meaning that $\mathbb{E}(\xi_k^2|\mathcal{F}_{k-1}) = \mathbb{E}(\xi_k^2)$ a.s. for any k, and if

$$\sup_{i\geq 1}\frac{\mathbb{E}(|\xi_i|^p)}{\mathbb{E}(|\xi_i|^2)} < \infty\,, \tag{1.4}$$

the rates in the central limit theorem in terms of the $\mathbb{L}_\infty$-distance are of order $V_n^{-(p-2)/(2p+2)}$. For $p=3$ this gives the rate $V_n^{-1/8}$. However in that case, under the additional assumption that there exist two positive constants α and β such that for any $i\geq 1$, $\alpha\leq\mathbb{E}(|\xi_i|^2)\leq\beta$, Grams [15] proved that the rate is of order $V_n^{-1/4}$ (see Theorem 1 in Bolthausen [2]). Even if this rate can appear to be poor compared with the iid case, it cannot be improved without additional assumptions as shown in [2, Section 6, Example 1]. More generally, when $p\in(2,3)$, under the same condition on the conditional variances and assuming (1.4), one can reach the rate $V_n^{-(p-2)/(2p-2)}$ (see our Corollary 3.1). Again this rate cannot be improved without additional assumptions as shown by our Proposition 3.1. The papers [8] and [9] are in this direction. For instance, still in the case where the conditional variances are constant, Theorem 2.1 in [9] states that $\Delta_{n,\infty} \le C V_n^{-1/2} \log n$ provided that there exists $\gamma > 0$ such that $\mathbb{E}(|\xi_k|^3 | \mathcal{F}_{k-1}) \le \gamma \mathbb{E}(\xi_k^2 | \mathcal{F}_{k-1})$ a.s. for any k (see [8] for related results).

Let us now comment on the quantity $\|V_n^{-1}\langle M\rangle_n - 1\|_{p/2}$ appearing in the right hand side of (1.3) when it is not equal to zero. For stationary sequences (except in some degenerate cases), $\|V_n^{-1}\langle M\rangle_n - 1\|_{p/2}$ is typically of order $V_n^{-1/2}$ which leads at best to the rate $V_n^{-p/(4p+4)}$. It is therefore clear that, in these non-degenerate situations, the rate $V_n^{-1/4}$ cannot be reached, whatever the value of p.

One of the goals of this paper is to give tractable conditions (not assuming that $\mathbb{E}(\xi_k^2 | \mathcal{F}_{k-1}) = \mathbb{E}(\xi_k^2)$ a.s. or $V_n^{-1}\langle M\rangle_n = 1$ a.s.) for $p \in (2, 3]$ under which the rate $V_n^{-(p-2)/(2p-2)}$ can be reached for $\Delta_{n,\infty}$ (up to a logarithmic term when $p = 3$). These conditions will be expressed with the help of quantities involving a sum of conditional expectations and allow to use martingale approximations techniques, as introduced by Gordin [10] (see also Volný [27]), to get rates when the sequence is not a martingale differences sequence. Applications via martingale approximations are provided in Section 4. The case of sequential dynamical systems as developed by Conze and Raugi [4] is considered in Subsection 4.3.

To derive the rates concerning $\Delta_{n,\infty}$, we shall rather work with minimal distances also called Wasserstein distances of order r (see Inequality (3.1) below for the connection between $\Delta_{n,\infty}$ and these distances). In particular, we shall also exhibit rates for the minimal distance $\Delta_{n,1}$ (see the equality (1.8) below).

Let us recall the definitions of these minimal distances. Let $\mathcal{L}(\mu, \nu)$ be the set of probability laws on $\mathbb{R}^2$ with marginals μ and ν. Let us consider the following minimal distances: for any $r > 0$,

$$W_r(\mu, \nu) = \inf\Big\{ \Big(\int |x - y|^r P(dx, dy) \Big)^{1/\max(1,r)} : P \in \mathcal{L}(\mu, \nu) \Big\}.$$

We consider also the following ideal distances of order r (Zolotarev distances of order r). For two probability measures μ and ν, and r a positive real, let

$$\zeta_r(\mu, \nu) = \sup\Big\{ \int f d\mu - \int f d\nu : f \in \Lambda_r \Big\},$$

where Λ_r is defined as follows: denoting by l the natural integer such that $l < r \le l + 1$, Λ_r is the class of real functions f which are l-times continuously differentiable and such that

$$|f^{(l)}(x) - f^{(l)}(y)| \le |x - y|^{r-l} \text{ for any } (x, y) \in \mathbb{R} \times \mathbb{R}. \tag{1.5}$$

For $r \in]0, 1]$, applying the Kantorovich-Rubinstein theorem (see for instance [7, Theorem 11.8.2]) to the metric $d(x, y) = |x - y|^r$, we infer that

$$W_r(\mu, \nu) = \zeta_r(\mu, \nu). \tag{1.6}$$

For $r > 1$ and for probability laws on the real line, the following inequality holds

$$W_r(\mu,\nu) \le c_r\big(\zeta_r(\mu,\nu)\big)^{1/r}, \tag{1.7}$$

where c_r is a constant depending only on r (see [22, Theorem 3.1]). Note that for $r = 1$, (1.6) ensures that

$$W_1(P_{M_n/\sqrt{V_n}}, G_1) = \zeta_1(P_{M_n/\sqrt{V_n}}, G_1) = \Delta_{n,1}, \tag{1.8}$$

where $P_{M_n/\sqrt{V_n}}$ is the law of $M_n/\sqrt{V_n}$ and G_1 the $\mathcal{N}(0,1)$ distribution.

The paper is organized as follows. In Section 2, we give rates in terms of Zolotarev and then in terms of Wasserstein distances between the law of the martingale having a moment of order $p \in (2,3]$ and the Gaussian distribution with the same variance. Upper and lower bounds for the uniform distance $\Delta_{n,\infty}$ are provided in Section 3. Applications to linear statistics associated with stationary sequences, ρ-mixing sequences in the sense of Kolmogorov and Rozanov [17] and sequential dynamical systems are presented in Section 4. All the proofs are postponed to Section 5.

In the rest of the paper, we shall use the following notations: we will denote by P_X the law of a r.v. X and by G_a the $\mathcal{N}(0,a)$ distribution, and for two sequences $(a_n)_{n\ge1}$ and $(b_n)_{n\ge1}$ of positive reals, $a_n \ll b_n$ means there exists a positive constant C not depending on n such that $a_n \le Cb_n$ for any $n \ge 1$. Moreover, given a filtration $\mathcal{F}_\ell$, we shall often use the notation $\mathbb{E}_\ell(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_\ell)$.

2 Rates for Zolotarev and Wasserstein distances

In this section $(\xi_i)_{i\in\mathbb{N}}$ will denote a sequence of martingale differences in $\mathbb{L}^2$, with respect to the increasing filtration $(\mathcal{F}_i)_{i\in\mathbb{N}}$ and with $\mathbb{E}(\xi_i^2) = \sigma_i^2$. We shall use the following notations:

$$M_n = \sum_{i=1}^n \xi_i\ ,\ V_n = \sum_{i=1}^n \sigma_i^2\ ,\ \delta_n = \max_{1\le i\le n}|\sigma_i|\ ,\ v_n(a) = a^2\delta_n^2 + \alpha V_n\,,$$

where a is a positive real and $\alpha = (1+a^2)/a^2$. Moreover, for $p \ge 2$ and $\ell \ge 2$, we denote by

$$U_{\ell,n}(p) = \Big\|(|\xi_{\ell-1}| \vee \sigma_{\ell-1})^{p-2}\Big|\sum_{k=\ell}^n(\mathbb{E}_{\ell-1}(\xi_k^2) - \sigma_k^2)\Big|\Big\|_1\,. \tag{2.1}$$

Theorem 2.1. *Let $p \in]2,3]$ and $r \in (0,p]$. There exists a positive constant $C_{r,p}$ depending on (r,p) such that for every positive integer n and any $a \ge 1$,*

$$\zeta_r(P_{M_n}, G_{V_n}) \le C_{r,p}\Big(\delta_n^r\int_a^{\sqrt{v_n(a)/\delta_n^2}}\frac{1}{x^{3-r}}dx + \delta_n^{r-1}\int_a^{\sqrt{v_n(a)/\delta_n^2}}\frac{\psi_n(x)}{x^{2-r}}dx + L_n(p,r,a\delta_n)\Big) + 4\sqrt{2}a^r\delta_n^r\,, \tag{2.2}$$

where

$$\psi_n(t) = \sup_{1\le k\le n} \frac{\mathbb{E}\inf(t\delta_n \xi_k^2, |\xi_k|^3)}{\sigma_k^2} \tag{2.3}$$

and

$$L_n(p,r,a\delta_n) = \sum_{\ell=2}^{n} \frac{U_{\ell,n}(p)}{(V_n - V_{\ell-1} + a^2\delta_n^2)^{(p-r)/2}}\,. \tag{2.4}$$

Remark 2.1. Let $p \in]2,3]$ and $r \in (0,p]$. Using (1.6) or (1.7), the fact that

$$\zeta_r(P_{M_n/\sqrt{V_n}}, G_1) = V_n^{-r/2}\zeta_r(P_{M_n}, G_{V_n})$$

and inequality (2.2), one can derive upper bounds for $W_r(P_{M_n/\sqrt{V_n}}, G_1)$ and then rates in the central limit theorem.

In case $r \in (0,1]$, starting from the dual equality (1.6), the following corollary holds.

Corollary 2.1. *Let $p \in]2,3]$ and $r \in (0,1]$. Under the assumptions and notations of Theorem 2.1, there exists a positive constant $C_{r,p}$ depending on (r,p) such that*

$$W_r(P_{M_n}, G_{V_n}) \le 4\sqrt{2}(a\delta_n)^r + C_{r,p}\left(\int_a^{\sqrt{v_n(a)/\delta_n^2}} \frac{\psi_n(x)}{x}dx + L_n(p,r,a\delta_n)\right).$$

In particular if the ξ_i's are in $\mathbb{L}^p$ with $p \in]2,3]$ and $(r,p) \ne (1,3)$,

$$W_r(P_{M_n}, G_{V_n}) \le 4\sqrt{2}(a\delta_n)^r + \tilde{C}_{r,p}\left(\sup_{1\le k\le n} \frac{\mathbb{E}(|\xi_k|^p)}{\sigma_k^2}(v_n(a))^{(2+r-p)/2} + L_n(p,r,a\delta_n)\right),$$

and if the ξ_i's are in $\mathbb{L}^3$,

$$W_1(P_{M_n}, G_{V_n}) \le 4\sqrt{2}a\delta_n + \tilde{C}_3\left(\sup_{1\le k\le n} \frac{\mathbb{E}(|\xi_k|^3)}{\sigma_k^2}\log(\sqrt{v_n(a)}/\delta_n) + L_n(3,1,a\delta_n)\right).$$

Remark 2.2. Note that if $(\xi_i)_{i\ge1}$ is a sequence of integer valued random variables then, whatever its dependence structure, setting $S_n = \sum_{k=1}^n \xi_i$ and proceeding as in the proof of [22, Theorem 5.1] we derive that for any $r > 0$,

$$\liminf_{n\to\infty}\Big(W_r(P_{S_n}, G_{\mathrm{Var}(S_n)})\Big)^{\max(1,r)} \ge 2^{-r}/(r+1)$$

provided $\mathrm{Var}(S_n) \to \infty$ as $n \to \infty$. Hence, in the case of martingale differences, if $p \in (2,3)$, $\sup_{1\le k\le n}\sigma_k^{-2}\mathbb{E}(|\xi_k|^p) \le C_1$ and $L_n(p,p-2,\delta_n) \le C_2$, we get

$$2^{-(p-2)}/(p-1) \le \liminf_{n\to\infty} W_{p-2}(P_{M_n}, G_{V_n}) \le \limsup_{n\to\infty} W_{p-2}(P_{M_n}, G_{V_n}) \le K$$

for some positive constant K. In addition, if $p=3$, $\sup_{1\le k\le n}\sigma_k^{-2}\mathbb{E}(|\xi_k|^3) \le C_1$ and $L_n(3,1,\delta_n) \le C_2$, we have

$$W_1(P_{M_n}, G_{V_n}) \ll \log(\sqrt{v_n(1)}/\delta_n)\,.$$

3 Berry-Esseen type results

Using [6, Remark 2.4] stating that, for any $p \in]2,3]$ and any integrable real-valued random variable Z,

$$\sup_{x\in\mathbb{R}} \big|\mathbb{P}(Z \le x) - \Phi(x)\big| \le (1 + (2\pi)^{-1/2})\big(W_{p-2}(P_Z, G_1)\big)^{1/(p-1)} , \tag{3.1}$$

Corollary 2.1 also leads to Berry-Esseen type upper bounds. More precisely, the following result holds

Corollary 3.1. *Assume that $(\xi_i)_{i\in\mathbb{Z}}$ is a sequence of martingale differences in $\mathbb{L}^p$ with $p \in]2,3]$. Let $\Delta_{n,\infty}$ be defined by (1.2). Then, with the notations of Section 2, one has*

$$\Delta_{n,\infty} \ll \begin{cases} V_n^{-\frac{(p-2)}{2(p-1)}} \left(\sup\limits_{1\le k\le n} \dfrac{\mathbb{E}(|\xi_k|^p)}{\sigma_k^2} + L_n(p, p-2, \delta_n) \right)^{1/(p-1)} & \text{if } p \in (2,3) \\ V_n^{-1/4} \left(\sup\limits_{1\le k\le n} \dfrac{\mathbb{E}(|\xi_k|^3)}{\sigma_k^2} \log\big(1 + 2V_n/\delta_n^2\big) + L_n(3, 1, \delta_n) \right)^{1/2} & \text{if } p = 3. \end{cases}$$

In particular if

$$\sup_{k\ge 1} \frac{\mathbb{E}(|\xi_k|^p)}{\sigma_k^2} \le C \quad \text{and for any } k \ge 1, \quad \mathbb{E}(\xi_k^2|\mathcal{F}_{k-1}) = \sigma_k^2 \text{ a.s.} \tag{3.2}$$

it follows that

$$\Delta_{n,\infty} \ll \begin{cases} V_n^{-\frac{(p-2)}{2(p-1)}} & \text{if } p \in (2,3) \\ V_n^{-1/4}\sqrt{\log(1 + 2V_n/\delta_n^2)} & \text{if } p = 3. \end{cases}$$

It turns out that one can construct a non stationary sequence of martingale differences satisfying (3.2) with $\sigma_k^2 = 1$ and such that there exists a positive constant c for which $\Delta_n \ge cn^{-\frac{(p-2)}{2(p-1)}}$ for any $p > 2$ and any $n \ge 20$. This shows that for $p \in (2,3)$ the rate given in Corollary 3.1 is optimal and quasi optimal (up to $\sqrt{\log n}$) in case $p = 3$.

Proposition 3.1. *Let $p > 2$ and $n \ge 20$. There exists $(X_1, \dots, X_n)$ such that*

1. $\mathbb{E}(X_k|\sigma(X_1, \dots, X_{k-1})) = 0$ *and* $\mathbb{E}(X_k^2|\sigma(X_1, \dots, X_{k-1})) = 1$ *a.s.,*
2. $\sup_{1\le k\le n} \mathbb{E}(|X_k|^p) \le \mathbb{E}(|Y|^p) + 5^{p-2}$ *where* $Y \sim \mathcal{N}(0, 1)$*,*
3. $\sup_{t\in\mathbb{R}} \big|\mathbb{P}(S_n \le t\sqrt{n}) - \Phi(t)\big| \ge 0.06\, n^{-(p-2)/(2p-2)}$*, where* $S_n = \sum_{k=1}^n X_k$*.*

Note that in case $p = 3$, Example 1 in [2] also shows that even for martingales with conditional variances equal to one and moments of order 3 uniformly bounded, the rate $n^{-1/4}$ cannot be improved in general.

Proof of Proposition 3.1. Let n be an integer satisfying $n \ge 20$. Let a be a real in $[1, \sqrt{n}/4[$, to be fixed later, and $k = \inf\{j \in \mathbb{N} : j \ge 4a^2\}$. Then $k < 1 + (n/4)$, which ensures that

$k < n$. Set $m = n - k$. We now define the sequence $(X_j)_{j \in [1,n]}$ of martingale differences as follows.

(i) The random variables $(X_j)_{j\in[1,m]}$ are independent and identically distributed with common law the standard normal law.

(ii) Let $U_{m+1}, \dots, U_n$ be a sequence of independent random variables with uniform distribution over $[0,1]$, independent of $(X_1, X_2, \dots, X_m)$. Let $S_m = X_1 + X_2 + \cdots + X_m$. If $|S_m| \notin [a, 2a]$, set $X_j = \Phi^{-1}(U_j)$ for any j in $[m+1, n]$. If $|S_m| \in [a, 2a]$, set

$$X_j = -(S_m/k)\mathbb{I}_{U_j \le k^2/(S_m^2+k^2)} + (k/S_m)\mathbb{I}_{U_j > k^2/(S_m^2+k^2)}. \tag{3.3}$$

From the definition of the random variables X_j, if $|S_m| \in [a, 2a]$ and $U_j \le k^2/(S_m^2 + k^2)$ for any j in $[m+1, n]$, then $S_n = 0$. It follows that

$$\mathbb{P}(S_n = 0) \ge \exp\bigl(-k \log(1 + 4a^2/k^2)\bigr) \frac{2}{\sqrt{2\pi m}} \int_a^{2a} \exp(-x^2/2m)dx. \tag{3.4}$$

We now estimate the conditional moments of the random variables X_j for $j > m$. From the definition of these random variables, for any measurable function f such that $f(X_j)$ is integrable

$$\mathbb{E}(f(X_j) \mid \mathcal{F}_{j-1}) = \mathbb{E}(f(X_j) \mid S_m). \tag{3.5}$$

Now, if $(S_m = x)$ for some x such that $|x| \notin [a, 2a]$, then $X_j = \Phi^{-1}(U_j)$ and consequently

$$\mathbb{E}(X_j \mid S_m = x) = 0 \ , \ \mathbb{E}(X_j^2 \mid S_m = x) = 1 \ \text{ and } \mathbb{E}(|X_j|^p \mid S_m = x) = \mathbb{E}(|Y|^p) \tag{3.6}$$

for any $p > 0$. Here Y is a random variable with law $\mathcal{N}(0, 1)$. Next, if $(S_m = x)$ for some x such that $|x| \in [a, 2a]$, then, according to (3.3),

$$\mathbb{E}(X_j \mid S_m = x) = 0 \ , \ \mathbb{E}(X_j^2 \mid S_m = x) = 1 \tag{3.7}$$

and, for any $p > 2$,

$$\mathbb{E}(|X_j|^p \mid S_m = x) = \frac{|x|^p k^{2-p} + k^p |x|^{2-p}}{x^2 + k^2}. \tag{3.8}$$

In that case, since $k \in [4a^2, 5a^2]$ and $|S_n| \in [a, 2a]$,

$$\mathbb{E}(|X_j|^p \mid S_m = x) \le |x|^p k^{-p} + k^{p-2}|x|^{2-p} \le 1 + (5a)^{p-2} \le 2\,(5a)^{p-2}. \tag{3.9}$$

From (3.6), the above upper bound and the fact that, since $n \ge 20$, $m \ge (3n/4) - 1 \ge (7n/10)$ and then

$$\mathbb{E}(|X_j|^p) \le \mathbb{E}(|Y|^p) + 2\,(5a)^{p-2}\mathbb{P}(|S_m| \in [a, 2a]) \le \mathbb{E}(|Y|^p) + 5^{p-2} 2a^{p-1} n^{-1/2}. \tag{3.10}$$

Now, for $p > 2$, choosing $a = (n/4)^{1/(2p-2)}$ in the above inequality, we get that

$$\mathbb{E}(|X_j|^p) \le \ \mathbb{E}(|Y|^p) + 5^{p-2}. \tag{3.11}$$

Consequently, for this choice of a, the absolute moments of order p of the random variables X_j are bounded by some positive constant depending only on p.

Now, using (3.4) we bound from below $\mathbb{P}(S_n = 0)$. First $4a^2 \leq k$, which ensures that $\exp\big(-k \log(1 + 4a^2/k^2)\big) \geq 1/e$, and second, for x in $[a, 2a]$,

$$\exp(-x^2/2m) \geq \exp(-2a^2/m) \geq \exp(-n/8m) \geq \exp(-10/56)$$

since $a^2 \leq n/16$ and $m \geq 7n/10$. Hence

$$\mathbb{P}(S_n = 0) \geq 0.24\ an^{-1/2} \geq 0.12\ n^{-(p-2)/(2p-2)} . \tag{3.12}$$

Therefrom, Item 3 of the proposition follows. □

4 Applications

Proposition 5.1 of Section 5 (which is the main ingredient for proving Theorem 2.1), combined with a suitable martingale approximation, can also be used to derive upper bounds for the Wasserstein distances between the law of partial sums of non necessarily stationary sequences and the corresponding limiting Gaussian distribution. This leads to new results for linear statistics, ρ-mixing sequences and sequential dynamical systems. Note that for these non stationary dynamical systems, a reversed martingale version of our Theorem 2.1 will be needed.

4.1 Linear statistics

Let $p \in]2, 3]$ and $(Y_i)_{i \in \mathbb{Z}}$ be a strictly stationary sequence of centered real-valued random variables in $\mathbb{L}^p$. Let $\mathcal{G}_k = \sigma(Y_i, i \leq k)$. Define $\gamma_k = \mathrm{Cov}(Y_0, Y_k)$ and

$$\lambda_k = \max\Big(\|Y_0 \mathbb{E}(Y_k|\mathcal{G}_0)\|_{p/2}, \sup_{j \geq i \geq k} \|\mathbb{E}(Y_i Y_j|\mathcal{G}_0)) - \mathbb{E}(Y_i Y_j)\|_{p/2}\Big) .$$

Let also

$$\Lambda_n = \sum_{i=1}^{n} i\lambda_i \quad \text{and} \quad \eta_n = \sum_{i=0}^{n} \|\mathbb{E}(Y_i|\mathcal{G}_0)\|_p . \tag{4.1}$$

Let $(\alpha_{i,n})_{i \geq 1}$ be a triangular array of real numbers and define

$$m_n = \max_{1 \leq \ell \leq n} |\alpha_{\ell,n}| \, , \, X_{i,n} = \alpha_{i,n} Y_i \, , \, S_n = \sum_{i=1}^{n} X_{i,n} \quad \text{and} \quad V_n = \mathrm{Var}(S_n) \, .$$

We refer to S_n as a “linear statistic” based on the stationary sequence $(Y_i)_{i \in \mathbb{Z}}$. Such linear statistics appear in many statistical contexts, for instance when considering least square estimators in a regression model with stationary errors (see for instance [5]).

In the two corollaries below we shall assume that $\sum_{k\geq 0}|\gamma_k| < \infty$ which implies in particular that $(Y_i)_{i\in\mathbb{Z}}$ has a bounded spectral density $f_Y(\theta) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\gamma_k \mathrm{e}^{\mathrm{i}k\theta}$ on $[-\pi,\pi]$. Moreover, in the first corollary, we assume in addition that the spectral density is bounded away from 0 (we refer to [3] for conditions ensuring such a fact). To state these corollaries, it is convenient to introduce the following quantity:

$$B(n,p) := \begin{cases} m_n^{p-2}\eta_n^{p-2}(\Lambda_n + \eta_n^2)\Big(\sum_{\ell=1}^n \alpha_{\ell,n}^2\Big)^{(3-p)/2} & \text{if } p \in (2,3) \\ m_n\eta_n(\Lambda_n+\eta_n^2)\log\Big(m_n^{-1}\sum_{\ell=1}^n \alpha_{\ell,n}^2\Big) & \text{if } p=3. \end{cases} \tag{4.2}$$

Corollary 4.1. *Let $p \in (2,3]$. Assume that $\sum_{k\geq 0}|\gamma_k| < \infty$ and that $\inf_{t\in[-\pi,\pi]}|f_Y(t)| = m > 0$. Then*

$$W_1(P_{S_n}, G_{V_n}) \ll m_n \sum_{k=0}^n \|\mathbb{E}(Y_k|\mathcal{G}_0)\|_2 + B(n,p)\,.$$

Note that if

$$\sum_{i\geq 1}\|\mathbb{E}(Y_i|\mathcal{G}_0)\|_2 < \infty\,, \tag{4.3}$$

then $\sum_{k\geq 0}|\gamma_k| < \infty$ (see for instance [18, p. 106]). If in addition to (4.3), we assume that $\sup_{n\geq 0}(\Lambda_n + \eta_n) < \infty$, then we get

$$W_1(P_{S_n}, G_{V_n}) \ll \begin{cases} m_n^{p-2}\Big(\sum_{\ell=1}^n \alpha_{\ell,n}^2\Big)^{(3-p)/2} & \text{if } p\in(2,3) \\ m_n\log\Big(m_n^{-1}\sum_{\ell=1}^n\alpha_{\ell,n}^2\Big) & \text{if } p=3. \end{cases} \tag{4.4}$$

For additional results in the special case where $(Y_i)_{i\in\mathbb{Z}}$ is a stationary sequence of martingale differences, we refer to [5].

Remark 4.1. If, for any positive k,

$$\lim_{n\to\infty}\frac{\sum_{\ell=1}^{n-k}\alpha_{\ell,n}\alpha_{\ell+k,n}}{\sum_{\ell=1}^n \alpha_{\ell,n}^2} = c_k\,,$$

and $\sum_{k\geq 0}|\gamma_k| < \infty$, then

$$\frac{V_n}{\sum_{\ell=1}^n\alpha_{\ell,n}^2} \to \sigma^2 = \gamma_0 + 2\sum_{k\geq 1}c_k\gamma_k\,, \text{ as } n\to\infty\,. \tag{4.5}$$

Moreover if $\inf_{t\in[-\pi,\pi]}|f_Y(t)| = m > 0$, then $\sigma^2 > 0$. Let $T_n = S_n/\sqrt{\sum_{\ell=1}^n\alpha_{\ell,n}^2}$. Under (4.3) and if f_Y is bounded away from zero, $\sup_{n\geq 0}(\Lambda_n+\eta_n) < \infty$ and (4.5) holds, it follows that

$$W_1(P_{T_n}, G_{\sigma^2}) \ll \left| \frac{V_n^{1/2}}{\sqrt{\sum_{\ell=1}^n \alpha_{\ell,n}^2}} - \sigma \right| + \begin{cases} \left(\frac{m_n}{\sqrt{\sum_{\ell=1}^n \alpha_{\ell,n}^2}} \right)^{p-2} & \text{if } p \in (2,3) \\ \frac{m_n}{\sqrt{\sum_{\ell=1}^n \alpha_{\ell,n}^2}} \log \Big(m_n^{-1} \sum_{\ell=1}^n \alpha_{\ell,n}^2 \Big) & \text{if } p = 3. \end{cases}$$

In case where $\alpha_{k,n} = \kappa k^\alpha$ with $\alpha > -1/2$, then $m_n (\sum_{\ell=1}^n \alpha_{\ell,n}^2)^{-1/2}$ is exactly of order $n^{-(\alpha+1/2)} \mathbf{1}_{-1/2<\alpha<0} + n^{-1/2} \mathbf{1}_{\alpha \geq 0}$ and we can show (since $\sum_{i\geq 1} i|\gamma_i| < \infty$ and $\sigma > 0$), that

$$\left| \frac{V_n^{1/2}}{\sqrt{\sum_{\ell=1}^n \alpha_{\ell,n}^2}} - \sigma \right| = O(1/n) \,.$$

Hence, for instance if $\alpha \geq 0$,

$$W_1(P_{T_n}, G_{\sigma^2}) \ll \begin{cases} n^{-(p-2)/2} & \text{if } p \in (2,3) \\ n^{-1/2} \log(n) & \text{if } p = 3. \end{cases}$$

Remark 4.2. Let $(\alpha_{\mathbf{Y}}(k))_{k>0}$ be the usual Rosenblatt strong mixing coefficients [25] of the sequence $(Y_i)_{i\in\mathbb{Z}}$. If we assume that

$$\mathbb{P}(|Y_0| \geq t) \leq Ct^{-s} \text{ for some } s > p \text{ and } \sum_{k\geq 1} k(\alpha_{\mathbf{Y}}(k))^{2/p-2/s} < \infty \,,$$

then condition (4.3) holds and $\sup_{n\geq 0}(\Lambda_n + \eta_n) < \infty$. Hence in this case (4.4) holds and Remark 4.1 applies.

If we do not require the spectral density bounded away from 0 but only that $f_Y(0) > 0$ then an additional term appears in the bound of the Wasserstein distance between P_{S_n} and G_{V_n}.

Corollary 4.2. *Let $p \in (2,3]$. Assume that $\sum_{k\geq 1} k^2 |\gamma_k| < \infty$ and $f_Y(0) > 0$. Then*

$$W_1(P_{S_n}, G_{V_n}) \ll m_n \sum_{k=0}^n \|\mathbb{E}(Y_k|\mathcal{G}_0)\|_2 + B(n,p) + \Big(\sum_{k=1}^{n+1} (\alpha_{k,n} - \alpha_{k-1,n})^2 \Big)^{1/2} ,$$

where $B(n,p)$ is defined in (4.2).

4.2 ρ-mixing sequences

In this section we consider a sequence $(X_i)_{i\geq 1}$ of centered ($\mathbb{E}(X_i) = 0$ for all i), real-valued bounded random variables, which are ρ-mixing in the sense that

$$\rho(k) = \sup_{j\geq 1} \sup_{v>u\geq j+k} \rho\big(\sigma(X_i, 1 \leq i \leq j), \sigma(X_u, X_v)\big) \to 0 \,, \text{ as } k \to \infty \,,$$

where $\sigma(X_t, t \in A)$ is the σ-field generated by the r.v.'s X_t with indices in A and we recall that the maximal correlation coefficient $\rho(\mathcal{U}, \mathcal{V})$ between two σ-algebras is defined by

$$\rho(\mathcal{U}, \mathcal{V}) = \sup\{|\mathrm{corr}(X, Y)| : X \in \mathbb{L}^2(\mathcal{U}), Y \in \mathbb{L}^2(\mathcal{V})\} .$$

In this section we shall also assume that the r.v.'s $(X_i)_{i\geq 1}$ satisfies the following set of assumptions

$$(H) := \begin{cases} 1)\ \Theta = \sum_{k\geq 1} k\rho(k) < \infty . \\ 2)\ \text{For any } n \geq 1,\ C_n := \max\limits_{1\leq \ell\leq n} \dfrac{\sum_{i=\ell}^n \mathbb{E}(X_i^2)}{\mathbb{E}(S_n - S_{\ell-1})^2} < \infty . \end{cases}$$

Remark 4.3. Note that in point 2) in (H) necessarily $C_n \geq 1$. In many cases of interest the sequence $(C_n)_n$ is bounded: for example, when $X_i = f_i(Y_i)$ where Y_i is a Markov chain satisfying $\rho_Y(1) < 1$, then according to [20, Proposition 13], $C_n \leq (1+\rho_Y(1))(1-\rho_Y(1))^{-1}$. Here $(\rho_Y(k))_{k\geq 0}$ is the sequence of ρ-mixing coefficients of the Markov chain $(Y_i)_i$.

Corollary 4.3. *Let $(X_i)_{i\geq 1}$ be a sequence of centered bounded real-valued random variables such that (H) is satisfied. Let $V_n = \mathrm{Var}(S_n)$ and $K_n = \max_{1\leq i\leq n} \|X_i\|_\infty$. Then for any positive integer n,*

$$W_1(P_{S_n}, G_{V_n}) \ll K_n(1 + C_n \log(1 + C_n V_n)) .$$

Remark 4.4. If the sequences $(C_n)_n$ and $(K_n)_n$ are bounded and $V_n \to \infty$, then Corollary 4.3 provides a rate in the central limit theorem for $S_n/\sqrt{V_n}$. More precisely,

$$W_1(P_{S_n/\sqrt{V_n}}, G_1) = O(V_n^{-1/2} \log(V_n)) \quad \text{and} \quad \|F_n - \Phi\|_\infty = O(V_n^{-1/4} \sqrt{\log(V_n)}) .$$

where F_n is the c.d.f. of $S_n/\sqrt{V}_n$ (the second inequality follows from (3.1)). Note that the above upper bounds hold even if we do not require a linear growth of the variance V_n as it is imposed for instance in [28, Theorem 3.1] and of course, in the stationary case, in [29, 21, 26].

4.3 Sequential dynamical systems

The term sequential dynamical system, introduced by Berend and Bergelson [1], refers to a non-stationary system defined by the composition of deterministic maps $T_k \circ T_{k-1} \circ \cdots \circ T_1$ acting on a space X.

More precisely, we consider here the setting described by Conze and Raugi [4] and Haydn et al. [16]. Let $(T_k)_{k\geq 1}$ be a sequence of maps from X to X, where X is either a compact subset of $\mathbb{R}^d$ or the d-dimensional torus $\mathbb{T}^d$. Let also m be the Lebesgue measure defined on the Borel σ-algebra $\mathcal{B}$ of X, normalized in such a way that $m(X) = 1$. We assume that each T_k is non singular with respect to m i.e. $m(A) > 0 \Longrightarrow m(T(A)) > 0$.

Let P_k be the Perron-Frobenius operator, that is the adjoint of the composition by T_k: for any $f \in \mathbb{L}_1(m), g \in \mathbb{L}_\infty(m)$,

$$\int_X f(x)\, g \circ T_k(x)\, m(dx) = \int_X (P_k f)(x)\, g(x)\, m(dx)\,.$$

Let also $\tau_k = T_k \circ T_{k-1} \circ \ldots \circ T_1$ and $\pi_k = P_k \circ P_{k-1} \circ \ldots \circ P_1$, and note that π_k is the Perron-Frobenius operator of τ_k.

Let $\mathcal{V} \subset \mathbb{L}_\infty(m)$, $(1 \in \mathcal{V})$, be a Banach space of functions from X to $\mathbb{R}$ with norm $\|\cdot\|_v$, such that $\|\phi\|_\infty \leq \kappa_1 \|\phi\|_v$ for some $\kappa_1 > 0$. We assume moreover that if ϕ_1, ϕ_2 are two functions in $\mathcal{V}$, then the usual product $\phi_1\phi_2$ belongs to $\mathcal{V}$ and satisfies $\|\phi_1\phi_2\|_v \leq \kappa_2 \|\phi_1\|_v \|\phi_2\|_v$ for some $\kappa_2 > 0$. In what follows, we set $\kappa = \max(\kappa_1, \kappa_2)$. Typical examples of Banach spaces $\mathcal{V}$ are the space BV of functions with bounded variation on a compact interval of $\mathbb{R}$, or the space $\mathcal{H}_\alpha$ of α-Hölder function on a compact set of $\mathbb{R}^d$, equipped with their usual norms.

We now recall the properties (DEC) and (MIN) introduced in [4] (we use the formulation of [16]):

Property (DEC): There exist two constants $C > 0$ and $\gamma \in (0,1)$ such that: for any positive integer n, any n-tuple $(j_1, \ldots, j_n)$ of positive integers, and any $f \in \mathcal{V}$,

$$\|P_{j_n} \circ \cdots \circ P_{j_1}(f - m(f))\|_v \leq C\gamma^n \|f - m(f)\|_v\,.$$

Property (MIN): There exist $\delta > 0$ and $\gamma \in (0,1)$ such that: for any positive integer n, and any n-tuple $(j_1, \ldots, j_n)$ of positive integers, we have the uniform lower bound

$$\inf_{x \in X} P_{j_n} \circ \cdots \circ P_{j_1} 1(x) \geq \delta\,.$$

The main result of this subsection is the following corollary.

Corollary 4.4. *Let $(\phi_n)_{n\geq 1}$ be a sequence of functions in $\mathcal{V}$ such that $\sup_{n\geq 1} \|\phi_n\|_v < \infty$. Let*

$$S_n = \sum_{k=1}^n (\phi_k(\tau_k) - m(\phi_k(\tau_k)))\,, \quad \text{and} \quad V_n = \int_X S_n^2(x)\, m(dx)\,.$$

Assume that the properties (DEC) and (MIN) are satisfied. Then, on the probability space $(X, \mathcal{B}, m)$,

$$W_1(P_{S_n}, G_{V_n}) \ll \log(n+1)\log(2+V_n)\,.$$

Remark 4.5. Under the assumptions of Corollary 4.4, we derive that

$$W_1(P_{S_n/\sqrt{V_n}}, G_1) \ll V_n^{-1/2}\log(n+1)\log(2+V_n)$$

and

$$\|F_n - \Phi\|_\infty \ll \Big(V_n^{-1/2}\log(n+1)\log(2+V_n)\Big)^{1/2}\,,$$

where F_n is the cdf of $S_n/\sqrt{V_n}$ (the second inequality follows from (3.1)). In particular, Corollary 4.4 provides a rate in the central limit theorem for $S_n/\sqrt{V_n}$ as soon as $(\log n \log\log n)/\sqrt{V_n} \to 0$ as $n \to \infty$.

5 Proofs

5.1 Proof of Theorem 2.1

The proof is based on the following proposition:

Proposition 5.1. *Let δ be a positive real and denote by $t_{\ell,n} = \big(V_n - V_\ell + \delta^2\big)^{1/2}$. Let $p \in]2,3]$ and $r \in (0,p]$. Then, there exist positive constants $\gamma_{r,p}$ depending on (r,p) and κ_r depending on r such that for every positive integer n,*

$$\zeta_r(P_{M_n}, G_{V_n}) \leq 4\sqrt{2}\delta^r + \gamma_{r,p}\Big\{\sum_{k=1}^n \Big(\frac{1}{t_{k,n}^{3-r}}\mathbb{E}\big(\xi_k^2 \min(\kappa_r t_{k,n}, |\xi_k|)\big) + \frac{\sigma_k^4}{t_{k,n}^{4-r}}\Big) + \sum_{\ell=2}^n \frac{U_{\ell,n}(p)}{(t_{\ell-1,n})^{p-r}}\Big\}, \tag{5.1}$$

where, for $\ell \geq 2$, $U_{\ell,n}(p)$ is defined in (2.1)*.*

Remark 5.1. When $r = 1$, $p = 3$ and $U_{\ell,n}(p) = 0$ for any ℓ, our bound is similar to the one stated in [24, Theorem 2.1]. However our quantity $\sum_{\ell=2}^n (t_{\ell-1,n})^{r-p} U_{\ell,n}(p)$ can be handled in many cases (see Section 4) while his condition $V_n^{-1}\langle M\rangle_n = 1$ a.s. is very restrictive.

We end the proof of the theorem with the help of this proposition taking $\delta = a\delta_n$. Hence we shall give an upper bound for

$$\sum_{k=1}^n \Big(\frac{1}{t_{k,n}^{3-r}}\mathbb{E}\big(\xi_k^2 \min(\kappa_r t_{k,n}, |\xi_k|)\big) + \frac{\sigma_k^4}{t_{k,n}^{4-r}}\Big),$$

where $t_{k,n} = (a^2\delta_n^2 + \sigma_{k+1}^2 + \cdots + \sigma_n^2)^{1/2}$. With this aim note first that

$$\frac{1}{t_{k,n}^{3-r}}\mathbb{E}\big(\xi_k^2 \min(\kappa_r t_{k,n}, |\xi_k|)\big) \leq \frac{\sigma_k^2}{t_{k,n}^{3-r}}\psi_n(\kappa_r \delta_n^{-1} t_{k,n}),$$

where $\psi_n(t)$ is defined in (2.3). Note that $t \mapsto \psi_n(t)$ is non decreasing and that for any $t \geq 0$ and any $\alpha \geq 1$, $\psi_n(\alpha t) \leq \alpha\psi_n(t)$. Hence,

$$\psi_n(\kappa_r t) \leq \psi_n(\max(1,\kappa_r)t) \leq \max(1,\kappa_r)\psi_n(t).$$

Next, let $\tilde\sigma_k = \sigma_k/\delta_n$. Note that since $\tilde\sigma_k \leq 1$,

$$\frac{\sigma_k^2}{t_{k,n}^2} = \frac{\tilde\sigma_k^2}{a^2 + \tilde\sigma_{k+1}^2 + \cdots + \tilde\sigma_n^2} \leq \frac{\alpha\tilde\sigma_k^2}{a^2 + \tilde\sigma_k^2 + \alpha\sum_{\ell=k+1}^n \tilde\sigma_\ell^2},$$

where $\alpha = (a^2+1)/a^2$. Let $u_k = a^2 + \alpha \sum_{\ell=k+1}^n \tilde\sigma_\ell^2$. It follows that

$$\frac{\sigma_k^2}{t_{k,n}^2} \le \frac{u_{k-1}-u_k}{(u_{k-1}-u_k)/\alpha + u_k} = \frac{\alpha(u_{k-1}-u_k)}{(u_{k-1}-u_k)+\alpha u_k} = \frac{\alpha a_k}{a_k+\alpha}$$

where

$$a_k = (u_{k-1}-u_k)/u_k\,.$$

But since $a^2 \ge 1$ we have $\alpha \le 2$. Hence, for any $x \ge 0$,

$$\frac{\alpha x}{x+\alpha} \le \log(1+x)\,,$$

implying that

$$\frac{\sigma_k^2}{t_{k,n}^2} \le \log(1+a_k) = \log(u_{k-1}/u_k)\,. \tag{5.2}$$

It follows that, if $r \ge 1$, since $t \mapsto \psi_n(t)$ is non decreasing and $t_{k,n}^2 \le \delta_n^2 u_k$ (since $\alpha \ge 1$),

$$\begin{aligned}\frac{\sigma_k^2}{t_{k,n}^{3-r}}\psi_n(\delta_n^{-1}t_{k,n}) = \frac{\sigma_k^2}{t_{k,n}^2}\psi_n(\delta_n^{-1}t_{k,n})t_{k,n}^{r-1} &\le 2\log(\sqrt{u_{k-1}}/\sqrt{u_k})\psi_n(\sqrt{u_k})\delta_n^{r-1}u_k^{(r-1)/2} \\ &\le 2\psi_n(\sqrt{u_k})\delta_n^{r-1}u_k^{(r-1)/2}\int_{\sqrt{u_k}}^{\sqrt{u_{k-1}}}\frac{1}{x}dx \le 2\delta_n^{r-1}\int_{\sqrt{u_k}}^{\sqrt{u_{k-1}}}\frac{\psi_n(x)}{x^{2-r}}dx\,.\end{aligned}$$

Hence, if $r \ge 1$,

$$\begin{aligned}\sum_{k=1}^n \frac{\sigma_k^2}{t_{k,n}^{3-r}}\psi_n(\delta_n^{-1}t_{k,n}) &\le 2\delta_n^{r-1}\int_a^{\sqrt{a^2+\alpha\sum_{\ell=1}^n\tilde\sigma_\ell^2}}\frac{\psi_n(x)}{x^{2-r}}dx \\ &\le 2\delta_n^{r-1}\int_a^{\sqrt{v_n(a)/\delta_n^2}}\frac{\psi_n(x)}{x^{2-r}}dx\,.\end{aligned} \tag{5.3}$$

We study now the case $r < 1$. With this aim, note first that taking into account that $\tilde\sigma_k^2 \le 1$, $\alpha \le 2$ and that $a \ge 1$, we have

$$t_{k,n}^2 = \delta_n^2\Big(a^2 + \sum_{\ell=k+1}^n \tilde\sigma_\ell^2\Big) \ge a^2(a^2+\alpha)^{-1}\delta_n^2 u_{k-1} \ge \delta_n^2 u_{k-1}/3\,, \tag{5.4}$$

(for the first inequality, use the fact that $a^2(a^2+\alpha)^{-1} \le \alpha^{-1}$). When $r < 1$, taking into account the upper bound (5.4), we then derive

$$\frac{\sigma_k^2}{t_{k,n}^{3-r}}\psi_n(\delta_n^{-1}t_{k,n}) \le 2\times 3^{(1-r)/2}\delta_n^{r-1}u_{k-1}^{(r-1)/2}\psi_n(\sqrt{u_k})\log(\sqrt{u_{k-1}}/\sqrt{u_k})\,.$$

Hence, when $r < 1$,

$$\sum_{k=1}^n \frac{\sigma_k^2}{t_{k,n}^{3-r}}\psi_n(\delta_n^{-1}t_{k,n}) \le 2\times 3^{(1-r)/2}\delta_n^{r-1}\int_a^{\sqrt{v_n(a)/\delta_n^2}}\frac{\psi_n(x)}{x^{2-r}}dx\,.$$

The bound (5.4) and (5.2) also implies that, for any $r \le 2$,

$$\sum_{k=1}^n \frac{\sigma_k^4}{t_{k,n}^{4-r}} \le \delta_n^2 \sum_{k=1}^n \frac{\sigma_k^2}{t_{k,n}^2} \times \frac{1}{t_{k,n}^{2-r}} \le 3^{(2-r)/2} \delta_n^r \sum_{k=1}^n \frac{\sigma_k^2}{t_{k,n}^2} \times \frac{1}{u_{k-1}^{(2-r)/2}}$$
$$\le 2 \times 3^{(2-r)/2} \delta_n^r \sum_{k=1}^n \log(\sqrt{u_{k-1}}/\sqrt{u_k}) \times \frac{1}{u_{k-1}^{(2-r)/2}} = 2 \times 3^{(2-r)/2} \delta_n^r \sum_{k=1}^n \frac{1}{u_{k-1}^{(2-r)/2}} \int_{\sqrt{u_k}}^{\sqrt{u_{k-1}}} \frac{1}{x} dx$$
$$\le 2 \times 3^{(2-r)/2} \delta_n^r \sum_{k=1}^n \int_{\sqrt{u_k}}^{\sqrt{u_{k-1}}} \frac{1}{x^{3-r}} dx \le 2 \times 3^{(2-r)/2} \delta_n^r \int_a^{\sqrt{u_0}} \frac{1}{x^{3-r}} dx \,.$$

When $r > 2$, we use the fact that $t_{k,n}^2 \le \delta_n^2 u_k$ to derive that

$$\sum_{k=1}^n \frac{\sigma_k^4}{t_{k,n}^{4-r}} \le 2\delta_n^r \int_a^{\sqrt{u_0}} \frac{1}{x^{3-r}} dx \,.$$

All these considerations end the proof of Theorem 2.1. It remains to prove Proposition 5.1.

Proof of Proposition 5.1. Let $(Y_i)_{i\in\mathbb{N}}$ be a sequence of $\mathcal{N}(0,\sigma_i^2)$-distributed independent random variables, independent of the sequence $(\xi_i)_{i\in\mathbb{N}}$. For $n > 0$, let $T_n = \sum_{j=1}^n Y_j$. Let also Z be a $\mathcal{N}(0,\delta^2)$-distributed random variable independent of $(\xi_i)_{i\in\mathbb{N}}$ and $(Y_i)_{i\in\mathbb{N}}$. Using Lemma 5.1 in [6] together with the fact that, for any real c, $\zeta_r(P_{cX}, P_{cY}) = |c|^r \zeta_r(P_X, P_Y)$, we derive that for any r in $]0,p]$,

$$\zeta_r(P_{M_n}, P_{T_n}) \le 2\zeta_r(P_{M_n} * P_Z, P_{T_n} * P_Z) + 4\sqrt{2}\delta^r \,. \tag{5.5}$$

Consequently it remains to bound up

$$\zeta_r(P_{M_n} * P_Z, P_{T_n} * P_Z) = \sup_{f\in\Lambda_r} \mathbb{E}(f(M_n+Z) - f(T_n+Z)) \,.$$

Recall that $V_n = \sum_{i=1}^n \sigma_i^2$ and, for any $k \le n$, set

$$f_{V_n-V_k}(x) = \mathbb{E}(f(x + T_n - T_k + Z)).$$

Then, from the independence of the above sequences,

$$\mathbb{E}(f(M_n+Z) - f(T_n+Z)) = \sum_{k=1}^n D_k \,,$$

where

$$D_k = \mathbb{E}\big(f_{V_n-V_k}(M_{k-1}+\xi_k) - f_{V_n-V_k}(M_{k-1}+Y_k)\big) \,.$$

By the Taylor formula, we get

$$f_{V_n-V_k}(M_{k-1}+\xi_k) - f_{V_n-V_k}(M_{k-1}+Y_k)$$
$$= f'_{V_n-V_k}(M_{k-1})(\xi_k - Y_k) + \frac{1}{2} f''_{V_n-V_k}(M_{k-1})(\xi_k^2 - Y_k^2) - \frac{1}{6} f^{(3)}_{V_n-V_k}(M_{k-1})(Y_k^3) + R_k \,,$$

where

$$R_k \le \xi_k^2 \Big(\|f''_{V_n - V_k}\|_\infty \wedge \frac{1}{6} \|f^{(3)}_{V_n - V_k}\|_\infty |\xi_k| \Big) + \frac{1}{24} \|f^{(4)}_{V_n - V_k}\|_\infty Y_k^4 \,.$$

Using the fact that $(\xi_k)_{k\in\mathbb{N}}$ is a sequence of martingale differences independent of the sequence of iid Gaussian random variables $(Y_k)_{k\in\mathbb{N}}$, we then get

$$\mathbb{E}(f(M_n+Z)-f(T_n+Z)) = \frac{1}{2}\sum_{k=1}^n \mathbb{E}\big(f''_{V_n-V_k}(M_{k-1})(\xi_k^2-Y_k^2)\big) + \sum_{k=1}^n \mathbb{E}(R_k) \,. \qquad (5.6)$$

Note first that

$$\mathbb{E}(R_k) \le \mathbb{E}\Big(\xi_k^2\Big(\|f''_{V_n-V_k}\|_\infty \wedge \frac{1}{6}\|f^{(3)}_{V_n-V_k}\|_\infty|\xi_k|\Big)\Big) + \frac{\sigma_k^4}{8}\|f^{(4)}_{V_n-V_k}\|_\infty \,.$$

Recall the notation $t_{k,n} = (\delta^2+\sigma_{k+1}^2+\cdots+\sigma_n^2)^{1/2}$. By Lemma 6.1 in [6], we have that for any integer $i\ge 1$,

$$\|f^{(i)}_{V_n-V_k}\|_\infty \le c_{r,i}t_{k,n}^{r-i} \,. \qquad (5.7)$$

Hence, setting $\kappa_r = 6c_{r,2}/c_{r,3}$, we get

$$\mathbb{E}(R_k) \le \frac{c_{r,3}}{6}\times\frac{1}{t_{k,n}^{3-r}}\mathbb{E}\Big(\xi_k^2\min(\kappa_r t_{k,n},|\xi_k|)\Big) + \frac{c_{r,4}}{8}\frac{\sigma_k^4}{t_{k,n}^{4-r}} \,. \qquad (5.8)$$

For $r=1$, we can take $\kappa_r=6$, $c_{r,3}=1$ and $c_{r,4}=8/5$.

We study now the quantity $\sum_{k=1}^n \mathbb{E}\big(f''_{V_n-V_k}(M_{k-1})(\xi_k^2-Y_k^2)\big)$. With this aim let us consider a sequence (Y'_k) of real-valued random variables independent of (Y_k) and (ξ_k) and such that $\mathcal{L}(Y'_k)=\mathcal{L}(Y_k)$. Note first that

$$\mathbb{E}\big((f''_{V_n-V_k}(M_{k-1}+Y'_k)-f''_{V_n-V_k}(M_{k-1}))(\xi_k^2-Y_k^2)\big) = \mathbb{E}\big(f^{(3)}_{V_n-V_k}(M_{k-1})Y'_k(\xi_k^2-Y_k^2)\big)+\mathbb{E}(R'_k) \,,$$

where, by taking into account (5.7) and the independence between $(Y'_k)_k$ and $(\xi_k,Y_k)_k$,

$$\mathbb{E}(|R'_k|)) \le \|f^{(4)}_{V_n-V_k}\|_\infty \mathbb{E}\big|(Y'_k)^2(\xi_k^2-Y_k^2)\big| \le 2c_{r,4}\frac{\sigma_k^4}{t_{k,n}^{4-r}} \,.$$

Since $\mathbb{E}(Y'_k)=0$ and $(Y'_k)_k$ is independent of $(\xi_k,Y_k)_k$, we get

$$\sum_{k=1}^n \Big|\mathbb{E}\big((f''_{V_n-V_k}(M_{k-1}+Y'_k)-f''_{V_n-V_k}(M_{k-1}))(\xi_k^2-Y_k^2)\big)\Big| \le 2c_{r,4}\sum_{k=1}^n\frac{\sigma_k^4}{t_{k,n}^{4-r}} \,. \qquad (5.9)$$

Now

$$\mathbb{E}\big(f''_{V_n-V_k}(M_{k-1}+Y'_k)(\xi_k^2-Y_k^2)\big) = \mathbb{E}\big(f''_{V_n-V_{k-1}}(M_{k-1})(\xi_k^2-Y_k^2)\big)$$
$$= \sum_{\ell=2}^k \mathbb{E}\Big(\big(f''_{V_n-V_{k-1}}(M_{\ell-1}+T_{k-1}-T_{\ell-1}) - f''_{V_n-V_{k-1}}(M_{\ell-2}+T_{k-1}-T_{\ell-2})\big)(\xi_k^2-Y_k^2)\Big)$$
$$= \sum_{\ell=2}^k \mathbb{E}\Big(\big(f''_{V_n-V_{\ell-1}}(M_{\ell-1}) - f''_{V_n-V_{\ell-1}}(M_{\ell-2}+T_{\ell-1}-T_{\ell-2})\big)(\xi_k^2-Y_k^2)\Big) \,.$$

Hence, by using Lemma 6.1 in [6], there exists a positive constant $c_{r,p}$ depending on (r,p) such that for any $n \geq 1$,

$$\sum_{k=1}^{n} \mathbb{E}\big(f''_{V_n-V_k}(M_{k-1}+Y'_k)(\xi_k^2-Y_k^2)\big)$$
$$= \sum_{\ell=2}^{n} \mathbb{E}\Big(\big(f''_{V_n-V_{\ell-1}}(M_{\ell-1}) - f''_{V_n-V_{\ell-1}}(M_{\ell-2}+T_{\ell-1}-T_{\ell-2})\big)\sum_{k=\ell}^{n}(\mathbb{E}_{\ell-1}(\xi_k^2)-\sigma_k^2)\Big)$$
$$\leq c_{r,p}\sum_{\ell=2}^{n}\frac{1}{(V_n-V_{\ell-1}+\delta^2)^{(p-r)/2}}\Big\||\xi_{\ell-1}-Y_{\ell-1}|^{p-2}\Big|\sum_{k=\ell}^{n}(\mathbb{E}_{\ell-1}(\xi_k^2)-\sigma_k^2)\Big|\Big\|_1 . \quad (5.10)$$

Starting from (5.6) and taking into account the upper bounds (5.8), (5.9) and (5.10), the desired inequality follows since for any integer $\ell \in [2,n]$ and any $p \in [2,3]$, we have $\mathbb{E}(|Y_{\ell-1}|^{p-2}) \leq (\mathbb{E}|Y_{\ell-1}|)^{p-2} \leq \sigma_{\ell-1}^{p-2}$. □

5.2 Proof of Corollary 4.1

For any $k \geq 1$, let $\mathcal{F}_k = \sigma(X_1,\dots,X_k)$ and $\mathcal{F}_0 = \{\emptyset,\Omega\}$. Write first

$$S_n = \sum_{k=1}^{n}(\mathbb{E}_k(S_n)-\mathbb{E}_{k-1}(S_n)) =: \sum_{k=1}^{n} d_{k,n} .$$

Note that $(d_{k,n})_{1\leq k\leq n}$ is a triangular array of martingale differences with respect to $(\mathcal{F}_k)_{k\geq 1}$ and that $V_n = \sum_{k=1}^{n}\mathbb{E}(d_{k,n}^2) = \mathbb{E}(S_n^2)$. Hence, setting $\delta_n = \max_{1\leq k\leq n}\|d_{k,n}\|_2$ and applying Proposition 5.1 we get that, for any $a \geq 1$,

$$W_1(P_{S_n},G_{V_n}) \ll a\delta_n + \sum_{k=1}^{n}\Big(\frac{\mathbb{E}(|d_{k,n}|^p)}{B_{k+1,n}^{(p-1)/2}(a)} + \frac{\sigma_{k,n}^4}{B_{k+1,n}^{3/2}(a)}\Big) + \sum_{\ell=2}^{n}\frac{1}{B_{\ell,n}^{(p-1)/2}(a)}U_{\ell,n}(p) , \quad (5.11)$$

where $\sigma_{k,n} = \|d_{k,n}\|_2$ and

$$B_{\ell,n}(a) = \sum_{k=\ell}^{n}\mathbb{E}(d_{k,n}^2)+a^2\delta_n^2 \ \text{ and } \ U_{\ell,n}(p) = \Big\|\big(|d_{\ell-1,n}|\vee\sigma_{\ell-1,n}\big)\Big|^{p-2}\sum_{k=\ell}^{n}(\mathbb{E}_{\ell-1}(d_{k,n}^2)-\sigma_{k,n}^2)\Big|\Big\|_1 .$$

Proceeding as in the proof of Theorem 2.1, we get that

$$\sum_{k=1}^{n}\frac{\sigma_{k,n}^4}{B_{k+1,n}^{3/2}(a)} \ll \delta_n . \quad (5.12)$$

Next, setting $\alpha = (a^2+1)/a^2$, note that

$$B_{k+1,n}(a) \geq \alpha^{-1}\Big(a^2\delta_n^2 + \sigma_{k,n}^2 + \alpha\sum_{\ell=k+1}^{n}\sigma_{\ell,n}^2\Big) \geq 2^{-1}B_{k,n}(a) .$$

Note also that

$$U_{\ell,n}(p) \le 2\|d_{\ell-1,n}\|_p^{p-2}\Big\|\sum_{k=\ell}^{n}(\mathbb{E}_{\ell-1}(d_{k,n}^2)-\sigma_{k,n}^2)\Big\|_{p/2}\,.$$

But, setting $A_{k,n}=\mathbb{E}_k(S_n-S_k)$, note that the following decomposition is valid:

$$d_{k,n}=X_{k,n}+A_{k,n}-A_{k-1,n}\,. \tag{5.13}$$

Hence

$$3^{1-p}\|d_{\ell,n}\|_p^p \le \|X_{\ell,n}\|_p^p+\|A_{\ell,n}\|_p^p+\|A_{\ell-1,n}\|_p^p \ \le |\alpha_{n,\ell}|^p\|Y_0\|_p^p+2\Big(\sum_{i=\ell}^{n}|\alpha_{i,n}|\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\Big)^p\,.$$

But, by convexity, setting $\beta_i=\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\big(\sum_{u=\ell}^{n}\|\mathbb{E}(Y_u|\mathcal{G}_{\ell-1})\|_p\big)^{-1}$, we get

$$\begin{aligned}\Big(\sum_{i=\ell}^{n}|\alpha_{i,n}|\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\Big)^p &\le \sum_{i=\ell}^{n}|\alpha_{i,n}|^p\beta_i^{1-p}\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p^p\\ &\le \Big(\sum_{u=1}^{n}\|\mathbb{E}(Y_u|\mathcal{G}_0)\|_p\Big)^{p-1}\sum_{i=\ell}^{n}|\alpha_{i,n}|^p\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\,,\end{aligned}$$

implying that

$$\|d_{\ell,n}\|_p^p \ll \Big(\sum_{u=0}^{n}\|\mathbb{E}(Y_u|\mathcal{G}_0)\|_p\Big)^{p-1}\sum_{i=\ell}^{n}|\alpha_{i,n}|^p\|\mathbb{E}(Y_i|\mathcal{G}_\ell)\|_p\,. \tag{5.14}$$

It follows that

$$\max_{1\le\ell\le n}\|d_{\ell,n}\|_p^{p-2} \ll \max_{1\le i\le n}|\alpha_{i,n}|^{p-2}\Big(\sum_{u=0}^{n}\|\mathbb{E}(Y_u|\mathcal{G}_0)\|_p\Big)^{p-2} := \max_{1\le i\le n}|\alpha_{i,n}|^{p-2}\eta_n^{p-2}\,. \tag{5.15}$$

On another hand

$$\begin{aligned}\Big\|\sum_{k=\ell}^{n}(\mathbb{E}_{\ell-1}(d_{k,n}^2)-\sigma_{k,n}^2)\Big\|_{p/2} &= \Big\|\mathbb{E}_{\ell-1}\Big(\sum_{k=\ell}^{n}d_{k,n}\Big)^2-\mathbb{E}\Big(\sum_{k=\ell}^{n}d_{k,n}\Big)^2\Big\|_{p/2}\\ &= \|\mathbb{E}_{\ell-1}(S_n-\mathbb{E}_{\ell-1}(S_n))^2-\mathbb{E}(S_n-\mathbb{E}_{\ell-1}(S_n))^2\|_{p/2}\\ &\le \|\mathbb{E}_{\ell-1}(S_n-S_{\ell-1})^2-\mathbb{E}(S_n-S_{\ell-1})^2\|_{p/2}+2\|\mathbb{E}_{\ell-1}(S_n-S_{\ell-1})\|_p^2\,.\end{aligned} \tag{5.16}$$

Note that

$$\begin{aligned}\|\mathbb{E}_{\ell-1}(S_n-S_{\ell-1})^2-\mathbb{E}(S_n-S_{\ell-1})^2\|_{p/2} &\le 2\sum_{i=\ell}^{n}\sum_{j=i}^{n}\|\mathbb{E}_{\ell-1}(X_{i,n}X_{j,n})-\mathbb{E}(X_{i,n}X_{j,n})\|_{p/2}\\ &\le 2\sum_{i=\ell}^{n}\sum_{j=i}^{n}|\alpha_{i,n}\alpha_{j,n}|\|\mathbb{E}(Y_iY_j|\mathcal{G}_{\ell-1})-\mathbb{E}(Y_iY_j)\|_{p/2}\,.\end{aligned}$$

But, note that, for any $j \geq i \geq \ell$, $\|\mathbb{E}(Y_iY_j|\mathcal{G}_{\ell-1}) - \mathbb{E}(Y_iY_j)\|_{p/2} \leq 2\|\mathbb{E}(Y_i\mathbb{E}(Y_j|\mathcal{G}_i))\|_{p/2}$. It follows that

$$\|\mathbb{E}_{\ell-1}(S_n - S_{\ell-1})^2 - \mathbb{E}(S_n - S_{\ell-1})^2\|_{p/2} \leq 2\sum_{i=\ell}^{n}\sum_{j=i}^{2i-\ell}|\alpha_{i,n}\alpha_{j,n}|\|\mathbb{E}(Y_iY_j|\mathcal{G}_{\ell-1}) - \mathbb{E}(Y_iY_j)\|_{p/2} + 4\sum_{i=\ell}^{n}\sum_{j=2i-\ell+1}^{n}|\alpha_{i,n}\alpha_{j,n}|\|Y_i\mathbb{E}(Y_j|\mathcal{G}_i)\|_{p/2}\,. \quad (5.17)$$

Hence by stationarity,

$$\|\mathbb{E}_{\ell-1}(S_n - S_{\ell-1})^2 - \mathbb{E}(S_n - S_{\ell-1})^2\|_{p/2} \leq 4\Big(\sum_{i=\ell}^{n}\sum_{j=i}^{n\wedge(2i-\ell)}|\alpha_{i,n}\alpha_{j,n}|\lambda_{i-\ell+1} + \sum_{i=\ell}^{n}\sum_{j=2i-\ell+1}^{n}|\alpha_{i,n}\alpha_{j,n}|\lambda_{j-i}\Big)\,.$$

It follows that

$$\|\mathbb{E}_{\ell-1}(S_n - S_{\ell-1})^2 - \mathbb{E}(S_n - S_{\ell-1})^2\|_{p/2} \leq 2\Big(\sum_{i=\ell}^{n}\alpha_{i,n}^2(i-\ell+1)\lambda_{i-\ell+1} + 2\sum_{j=\ell}^{n}\alpha_{n,j}^2\sum_{u=[(j-\ell)/2]}^{j-\ell}\lambda_u + \sum_{i=\ell}^{n}\alpha_{i,n}^2\sum_{u=i-\ell+1}^{n-i}\lambda_u\Big)\,.$$

In addition, recalling that $\beta_i = \|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\big(\sum_{u=\ell}^{n}\|\mathbb{E}(Y_u|\mathcal{G}_{\ell-1})\|_p\big)^{-1}$, we get by convexity,

$$\|\mathbb{E}_{\ell-1}(S_n - S_{\ell-1})\|_p^2 = \Big(\sum_{i=\ell}^{n}|\alpha_{i,n}|\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\Big)^2 \leq \sum_{i=\ell}^{n}\alpha_{i,n}^2\beta_i^{-1}\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p^2 \leq \sum_{u=1}^{n}\|\mathbb{E}(Y_u|\mathcal{G}_0)\|_p\sum_{i=\ell}^{n}\alpha_{i,n}^2\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\,. \quad (5.18)$$

So, overall, recalling that $\eta_n = \sum_{i=0}^{n}\|\mathbb{E}(Y_i|\mathcal{G}_0)\|_p$, we get

$$U_{\ell,n}(p) \ll \max_{1\leq\ell\leq n}|\alpha_{\ell,n}|^{p-2}\eta_n^{p-2}\Big(\sum_{i=\ell}^{n}\alpha_{i,n}^2(i-\ell+1)\lambda_{i-\ell+1} + \sum_{j=\ell}^{n}\alpha_{n,j}^2\sum_{u=[(j-\ell)/2]}^{n-j}\lambda_u + \eta_n\sum_{i=\ell}^{n}\alpha_{i,n}^2\|\mathbb{E}(Y_i|\mathcal{G}_{\ell-1})\|_p\Big)\,.$$

Hence, setting

$$\Lambda_{i,\ell} = (i-\ell+1)\lambda_{i-\ell+1} + \sum_{u=[(i-\ell)/2]}^{n-i}\lambda_u + \eta_n\|\mathbb{E}(Y_{i-\ell+1}|\mathcal{G}_0)\|_p\,,$$

we get

$$\sum_{\ell=2}^{n} \frac{1}{B_{\ell,n}^{(p-1)/2}(a)} U_{\ell,n}(p) \ll \max_{1\le \ell\le n} |\alpha_{\ell,n}|^{p-2}\eta_n^{p-2} \sum_{i=1}^{n} \frac{\alpha_{i,n}^2}{B_{i,n}^{(p-1)/2}(a)} \sum_{\ell=1}^{i} \Lambda_{i,\ell}\,.$$

Since, for any $i \le n$,

$$\sum_{\ell=1}^{i} \Lambda_{i,\ell} \ll \sum_{u=0}^{n}((u+1)\lambda_u + \eta_n \|\mathbb{E}(Y_u|\mathcal{G}_0)\|_p) \le \Lambda_n + \eta_n^2\,,$$

it follows that

$$\sum_{\ell=2}^{n} \frac{1}{B_{\ell,n}^{(p-1)/2}(a)} U_{\ell,n}(p) \ll \max_{1\le \ell\le n} |\alpha_{\ell,n}|^{p-2}\eta_n^{p-2}(\Lambda_n + \eta_n^2) \sum_{i=1}^{n} \frac{\alpha_{i,n}^2}{B_{i,n}^{(p-1)/2}(a)}\,. \tag{5.19}$$

Let

$$a = \frac{\max_{1\le k\le n} |\alpha_{k,n}| \max(\|Y_0\|_2, \sqrt{2\pi m}) + 2\max_{1\le k\le n-1} \|A_{k,n}\|_2}{\max_{1\le k\le n} \|d_{k,n}\|_2}\,,$$

where $m = \inf_{t\in[-\pi,\pi]} f_Y(t)$. The decomposition (5.13) entails that $a \ge 1$. On another hand, for any integer ℓ in $[1, n]$,

$$\begin{aligned} B_{\ell,n}(a) &= \mathbb{E}(S_n - S_{\ell-1} - A_{\ell-1})^2 + a^2\delta_n^2 = \mathbb{E}(S_n - S_{\ell-1})^2 - \mathbb{E}(A_{\ell-1})^2 + a^2\delta_n^2 \\ &\ge \|S_n - S_{\ell-1}\|_2^2 + \max_{1\le k\le n} |\alpha_{k,n}|^2 \max(\|Y_0\|_2^2, 2\pi m)\,. \end{aligned}$$

But

$$\mathrm{Var}(S_n - S_{\ell-1}) = \int_{-\pi}^{\pi} \Big|\sum_{k=\ell}^{n} \alpha_{k,n} \mathrm{e}^{\mathrm{i}tk}\Big|^2 f_Y(t)dt \ge m \int_{-\pi}^{\pi} \Big|\sum_{k=\ell}^{n} \alpha_{k,n} \mathrm{e}^{\mathrm{i}tk}\Big|^2 dt = 2\pi m \sum_{k=\ell}^{n} \alpha_{k,n}^2\,.$$

It follows that, for any integer ℓ in $[1, n]$,

$$B_{\ell,n}(a) \ge 2\pi m\Big(\sum_{i=\ell}^{n} \alpha_{i,n}^2 + \max_{1\le k\le n} \alpha_{k,n}^2\Big)\,. \tag{5.20}$$

Starting from (5.19) and taking into account (5.20) and the fact that $m > 0$, it follows that

$$\sum_{\ell=2}^{n} \frac{1}{B_{\ell,n}^{(p-1)/2}(a)} U_{\ell,n}(p) \ll \max_{1\le \ell\le n} |\alpha_{\ell,n}|^{p-2}\eta_n^{p-2}(\Lambda_n + \eta_n^2) \sum_{i=1}^{n} \frac{\alpha_{i,n}^2}{\Big(\sum_{j=i}^{n} \alpha_{j,n}^2 + \max_{1\le k\le n} \alpha_{k,n}^2\Big)^{(p-1)/2}}\,.$$

Hence proceeding as in the proof of Theorem 2.1, we get

$$\sum_{\ell=2}^{n} \frac{1}{B_{\ell,n}^{(p-1)/2}(a)} U_{\ell,n}(p) \ll \begin{cases} \max_{1\le \ell\le n} |\alpha_{\ell,n}|^{p-2}\eta_n^{p-2}(\Lambda_n + \eta_n^2)\Big(\sum_{\ell=1}^{n} \alpha_{\ell,n}^2\Big)^{(3-p)/2} & \text{if } p \in (2,3) \\ \max_{1\le \ell\le n} |\alpha_{\ell,n}|\eta_n(\Lambda_n + \eta_n^2) \log\Big(m_n^{-1} \sum_{\ell=1}^{n} \alpha_{\ell,n}^2\Big) & \text{if } p = 3. \end{cases} \tag{5.21}$$

On another hand, taking into account (5.14) and proceeding as before we get

$$\sum_{\ell=2}^{n} \frac{1}{B_{\ell,n}^{(p-1)/2}(a)} \|d_{\ell,n}\|_p^p \ll \begin{cases} \max\limits_{1\le \ell\le n} |\alpha_{\ell,n}|^{p-2}\eta_n^p \Big(\sum\limits_{\ell=1}^{n} \alpha_{\ell,n}^2 \Big)^{(3-p)/2} & \text{if } p\in(2,3) \\ \max\limits_{1\le \ell\le n} |\alpha_{\ell,n}|\eta_n^3 \log \Big(m_n^{-1} \sum\limits_{\ell=1}^{n} \alpha_{\ell,n}^2 \Big) & \text{if } p=3. \end{cases} \tag{5.22}$$

Starting from (5.11) and taking into account (5.12), (5.21) and (5.22) together with the fact that

$$a\delta_n \ll \max_{1\le \ell\le n} |\alpha_{\ell,n}| \Big(\sqrt{m} + \sum_{k=0}^{n} \|\mathbb{E}(Y_k|\mathcal{G}_0)\|_2 \Big) ,$$

the corollary follows.

5.3 Proof of Corollary 4.2

The proof follows the lines of the proof of Corollary 4.1. The only difference is in the choice of a. We take here

$$a = \frac{\max_{1\le k\le n} |\alpha_{k,n}| \big(\max(\|Y_0\|_2, \sqrt{2\pi f_Y(0)}\big) + 2\max_{1\le k\le n-1} \|A_{k,n}\|_2 + \sqrt{K(n)}}{\max_{1\le k\le n} \|d_{k,n}\|_2} ,$$

where $K(n) = \big(\sum_{k\ge1} k^2|\gamma_k|\big) \sum_{i=1}^{n+1} |\alpha_{i,n} - \alpha_{n,i-1}|^2$. Once again, the decomposition (5.13) entails that $a \ge 1$. On another hand,

$$\begin{aligned} B_{\ell,n}(a) &= \mathbb{E}(S_n - S_{\ell-1} - A_{\ell-1})^2 + a^2\delta_n^2 = \mathbb{E}(S_n - S_{\ell-1})^2 - \mathbb{E}(A_{\ell-1})^2 + a^2\delta_n^2 \\ &\ge \|S_n - S_{\ell-1}\|_2^2 + \max_{1\le k\le n} |\alpha_{k,n}|^2 \big(\max(\|Y_0\|_2^2, 2\pi f_Y(0)\big) + K(n) . \end{aligned}$$

But, setting $\tilde\alpha_u = \alpha_{u,n}$ if $u \in [\ell, n]$ and 0 otherwise, we get

$$\mathrm{Var}(S_n - S_{\ell-1}) = \sum_{k\in\mathbb{Z}} \gamma_k \sum_{i\in\mathbb{Z}} \tilde\alpha_i \tilde\alpha_{i+k} = 2\pi f_Y(0) \sum_{i=\ell}^{n} \alpha_{i,n}^2 - 2^{-1} \sum_{k\in\mathbb{Z}} \gamma_k \sum_{i\in\mathbb{Z}} (\tilde\alpha_i - \tilde\alpha_{i+k})^2 .$$

Setting $K = \sum_{k\ge1} k^2|\gamma_k|$, it follows that

$$\|S_n - S_{\ell-1}\|_2^2 + K \sum_{i=1}^{n+1} |\alpha_{i,n} - \alpha_{i-1,n}|^2 \ge 2\pi f_Y(0) \sum_{i=\ell}^{n} \alpha_{i,n}^2 ,$$

implying that

$$B_{\ell,n}(a) \ge 2\pi f_Y(0) \Big(\sum_{i=\ell}^{n} \alpha_{i,n}^2 + \max_{1\le k\le n} \alpha_{k,n}^2 \Big) .$$

Using the fact that $f_Y(0) > 0$, the rest of the proof is the same as that of Corollary 4.1.

5.4 Proof of Corollary 4.3

We start as in the proof of Corollary 4.1 and use the notation introduced there. So we have the upper bound (5.11) with $p = 3$. Recalling the notation $A_{k,n} = \mathbb{E}(S_n - S_k|\mathcal{F}_k)$, we select

$$a = \frac{\max_{1\le k\le n} \|X_k\|_2 + 2\max_{1\le k\le n-1} \|A_{k,n}\|_2}{\max_{1\le k\le n} \|d_{k,n}\|_2} .$$

The decomposition (5.13) entails that $a \ge 1$ and also that

$$\sum_{i=k}^{n} \|d_{i,n}\|_2^2 = \Big\| \sum_{i=k}^{n} d_{i,n} \Big\|_2^2 = \|S_n - S_{k-1} - A_{k-1,n}\|_2^2 .$$

It follows that

$$\begin{aligned} B_{k,n}(a) &= \sum_{i=k}^{n} \|d_{i,n}\|_2^2 + a^2\delta_n^2 \\ &= \|S_n - S_{k-1}\|_2^2 - \|A_{k-1,n}\|_2^2 + \big(\max_{1\le k\le n} \|X_k\|_2 + 2 \max_{1\le k\le n-1} \|A_{k,n}\|_2 \big)^2 \\ &\ge \|S_n - S_{k-1}\|_2^2 + \max_{1\le k\le n} \|X_k\|_2^2 . \end{aligned}$$

Using point 2) in (H) and the fact that $C_n \ge 1$, we derive

$$\frac{1}{B_{k,n}(a)} \le \frac{C_n}{\sum_{\ell=k}^{n} \|X_\ell\|_2^2 + \max_{1\le k\le n} \|X_k\|_2^2} := \frac{C_n}{\widetilde{V}_{k,n}} . \tag{5.23}$$

On another hand, for any $1 \le k \le n-1$ and any $\eta > 1/2$, by the definition of the ρ-mixing coefficients,

$$\begin{aligned} \|A_{k,n}\|_2^2 \le \Big(\sum_{\ell=k+1}^{n} \|\mathbb{E}(X_\ell|\mathcal{F}_k)\|_2 \Big)^2 &\ll \sum_{\ell=k+1}^{n} (\ell-k)^{2\eta} \|\mathbb{E}(X_\ell|\mathcal{F}_k)\|_2^2 \\ &\ll \sum_{\ell=k+1}^{n} (\ell-k)^{2\eta} \|X_\ell\|_2^2 \rho^2(\ell-k) . \end{aligned}$$

According to point 1) in (H) we can take $\eta > 1/2$ such that $\sum_{\ell\ge 1} \ell^{2\eta}\rho^2(\ell) < \infty$. Hence

$$\|A_{k,n}\|_2 \ll \max_{1\le k\le n} \|X_k\|_2 ,$$

implying that

$$a\delta_n \ll \max_{1\le k\le n} \|X_k\|_2 .$$

On another hand, from decomposition (5.13),

$$\|d_{k,n}\|_3^3 \le 9\Big(K_n \|X_k\|_2^2 + \|A_{k,n}\|_3^3 + \|A_{k-1,n}\|_3^3 \Big) .$$

But, for any $1 \le k \le n-1$ and any $\eta > 2/3$, by the definition of the ρ-mixing coefficients,

$$\|A_{k,n}\|_3^3 \le \Big(\sum_{\ell=k+1}^{n} \|\mathbb{E}(X_\ell|\mathcal{F}_k)\|_3 \Big)^3 \ll \sum_{\ell=k+1}^{n} (\ell-k)^{3\eta} \|\mathbb{E}(X_\ell|\mathcal{F}_k)\|_3^3$$
$$\ll K_n \sum_{\ell=k+1}^{n} (\ell-k)^{3\eta} \|\mathbb{E}(X_\ell|\mathcal{F}_k)\|_2^2 \ll K_n \sum_{\ell=k+1}^{n} (\ell-k)^{3\eta} \|X_\ell\|_2^2 \rho^2(\ell-k)\,.$$

So, overall,

$$a\delta_n + \sum_{k=1}^{n} \frac{\mathbb{E}(|d_{k,n}|^3)}{B_{k+1,n}(a)} \ll \max_{1\le k\le n} \|X_k\|_2 + K_n C_n \sum_{k=1}^{n} \frac{\|X_k\|_2^2}{\widetilde{V}_{k,n}}$$
$$+ K_n C_n \sum_{k=1}^{n} \frac{\sum_{\ell=k}^{n} (\ell-k+1)^{3\eta} \|X_\ell\|_2^2 \rho^2(\ell-k+1)}{\widetilde{V}_{k,n}}$$
$$\ll K_n + K_n C_n \sum_{k=1}^{n} \frac{\|X_k\|_2^2}{\widetilde{V}_{k,n}}$$
$$+ K_n C_n \sum_{\ell=1}^{n} \frac{\|X_\ell\|_2^2}{\widetilde{V}_{\ell,n}} \sum_{k=1}^{\ell} (\ell-k+1)^{3\eta} \rho^2(\ell-k+1)$$

According to point 1) in (H) we can take $\eta > 2/3$ such that $\sum_{\ell\ge 1} \ell^{3\eta}\rho^2(\ell) < \infty$. Hence, it follows that

$$a\delta_n + \sum_{k=1}^{n} \frac{\mathbb{E}(|d_{k,n}|^3)}{B_{k+1,n}(a)} \ll K_n + K_n C_n \sum_{\ell=1}^{n} \frac{\|X_\ell\|_2^2}{\widetilde{V}_{\ell,n}}\,.$$

With similar arguments as those leading to (5.3), we get

$$a\delta_n + \sum_{k=1}^{n} \frac{\mathbb{E}(|d_{k,n}|^3)}{B_{k+1,n}(a)} \ll K_n + K_n C_n \log\Big(1 + \sum_{k=1}^{n} \|X_k\|_2^2\Big)\,.$$

On another hand, we have

$$U_{\ell,n}(3) \le 2\|d_{\ell-1,n}\|_2 \Big\| \sum_{k=\ell}^{n} (\mathbb{E}_{\ell-1}(d_{k,n}^2) - \sigma_{k,n}^2) \Big\|_2\,.$$

To give an upper bound of this quantity we start from (5.16) with $p = 4$. Note first that

$$\|\mathbb{E}_{\ell-1}^2(S_n - S_{\ell-1})\|_2 \le 2 \sum_{i=\ell}^{n} \sum_{j=i}^{n} \|\mathbb{E}_{\ell-1}(X_i)\mathbb{E}_{\ell-1}(X_j)\|_2$$
$$\le 2 \sum_{i=\ell}^{n} \sum_{j=i}^{2i-\ell} \|\mathbb{E}_{\ell-1}(X_i) X_j\|_2 + 2 \sum_{i=\ell}^{n} \sum_{j=2i-\ell+1}^{n} \|X_i \mathbb{E}_{\ell-1}(X_j)\|_2\,.$$

Hence, by the definition of the ρ-mixing coefficients, we get

$$\|\mathbb{E}_{\ell-1}^2(S_n - S_{\ell-1})\|_2 \le 4K_n \sum_{i=\ell}^{n} (i-\ell)\|X_i\|_2 \rho(i-\ell)\,. \tag{5.24}$$

On another hand, by the definition of the ρ-mixing coefficients, we have: for $j \geq i \geq \ell$,

$$\|\mathbb{E}(X_iX_j|\mathcal{F}_{\ell-1}) - \mathbb{E}(X_iX_j)\|_2 \leq \|X_iX_j\|_2\rho(i-\ell) \leq K_n\|X_i\|_2\rho(i-\ell)\,, \tag{5.25}$$

and

$$\|X_i\mathbb{E}(X_j|\mathcal{F}_i)\|_2^2 = \mathbb{E}(\mathbb{E}(X_i^2X_j|\mathcal{F}_i)X_j) \leq K_n\|X_i\mathbb{E}(X_j|\mathcal{F}_i)\|_2\|X_j\|_2\rho(j-i)\,. \tag{5.26}$$

Hence starting from (5.16) with $p=4$ and taking into account (5.24) and the upper bounds (5.17) and (5.18) together with (5.25) and (5.26), we derive

$$U_{\ell,n}(3) \ll K_n\|d_{\ell-1,n}\|_2\sum_{i=\ell}^n \|X_i\|_2(i-\ell+1)\rho([i-\ell]/2)\,.$$

Hence, taking into account point 1) in (H) and (5.23),

$$\begin{aligned}\sum_{\ell=2}^n \frac{1}{B_{\ell,n}(a)}U_{\ell,n}(3) &\ll K_n\sum_{2\leq \ell\leq i\leq n}\frac{\|d_{\ell-1,n}\|_2\|X_i\|_2}{B_{\ell,n}(a)}(i-\ell+1)\rho([i-\ell]/2)\\ &\ll K_n\Big(\sum_{\ell=2}^n\frac{\|d_{\ell-1,n}\|_2^2}{B_{\ell,n}(a)}+\sum_{i=2}^n\frac{\|X_i\|_2^2}{B_{i,n}(a)}\Big)\sum_{k=0}^n(k+1)\rho(k/2)\\ &\ll \Theta K_n\Big(\sum_{\ell=2}^n\frac{\|d_{\ell-1,n}\|_2^2}{B_{\ell-1,n}(a)}+C_n\sum_{i=2}^n\frac{\|X_i\|_2^2}{\widetilde{V}_{i,n}(a)}\Big)\,,\end{aligned}$$

since $B_{\ell-1,n}(a) \leq 2B_{\ell,n}(a)$. With similar arguments as those leading to (5.3), we get

$$\sum_{\ell=2}^n \frac{1}{B_{\ell,n}(a)}U_{\ell,n}(3) \ll \Theta K_nC_n\log\Big(1+\sum_{k=1}^n\|X_k\|_2^2\Big)\,.$$

This ends the proof of the corollary since $\sum_{k=1}^n \|X_k\|_2^2 \leq C_nV_n$.

5.5 Proof of Corollary 4.4

As we shall see the result will use an approximation by a "reversed" martingale differences sequence. Hence, as a preliminary, we first state the following fact:

Fact 5.1. *[Reversed martingale differences sequences] Let $p \in (2,3]$. Assume that $(d_n)_{n\in\mathbb{N}}$ is a real-valued sequence of reversed martingale differences in $\mathbb{L}^p$ with respect to a non-increasing sequence $(\mathcal{G}_n)_{n\in\mathbb{N}}$ of σ-algebras. This means that for any integer n, d_n is $\mathcal{G}_n$-adapted and $\mathbb{E}(d_n|\mathcal{G}_{n+1})=0$ a.s. Let $M_n=\sum_{k=1}^n d_k$. Note that $M_n=\sum_{k=1}^n\xi_{n,k}$ with $\xi_{n,k}=d_{n-k+1}$. Clearly $(\xi_{n,k})_{1\leq k\leq n}$ is a sequence of martingale differences with respect to the increasing sequence $(\mathcal{F}_{k,n})_k$ of σ-algebras with $\mathcal{F}_{k,n}=\mathcal{G}_{n-k+1}$. Hence, applying*

Proposition 5.1, it follows that (5.1) holds with $\widetilde{t}_{k,n} = \big(\sum_{i=1}^{k-1}\mathbb{E}(d_i^2)+\delta^2\big)^{1/2}$ *replacing* $t_{k,n}$*,* d_k *in place of* ξ_k*,* $\widetilde{t}_{\ell+1,n}$ *in place of* $t_{\ell-1,n}$ *and*

$$\tilde{U}_{\ell,n}(p) = \Big\|(|d_{\ell+1}|\vee\sigma_{\ell+1})^{p-2}\Big|\sum_{i=1}^{\ell}(\mathbb{E}(d_i^2|\mathcal{G}_{\ell+1})-\sigma_i^2)\Big|\Big\|_1 \tag{5.27}$$

in place of $U_{\ell,n}(p)$*. In particular, the following "reversed" version of Theorem 2.1 holds: setting* $\mathbb{E}(d_i^2)=\sigma_i^2$ *and* $\psi_n(t)=\sup_{1\le k\le n}\sigma_k^{-2}\mathbb{E}\inf(t\delta_n d_k^2,|d_k|^3)$*, there exist positive constants* $C_{r,p}$ *depending on* (r,p) *and* κ_r *depending on* r *such that for every positive integer* n *and any real* $a\ge 1$*,*

$$\zeta_r(P_{M_n},G_{V_n}) \le C_{r,p}\Big(\delta_n^r\int_a^{\sqrt{v_n(a)/\delta_n^2}}\frac{1}{x^{3-r}}dx+\delta_n^{r-1}\int_a^{\sqrt{v_n(a)/\delta_n^2}}\frac{\psi_n(\kappa_r x)}{x^{2-r}}dx \\ +\sum_{k=1}^{n-1}\frac{\tilde{U}_{k,n}(p)}{(a^2\delta_n^2+\sum_{i=1}^{k}\sigma_i^2)^{(p-r)/2}}\Big)+4\sqrt{2}a^r\delta_n^r\,. \tag{5.28}$$

We go back to the proof of Corollary 4.4. Let $\mathcal{B}_n=\tau_n^{-1}\mathcal{B}$ and $\tilde{\phi}_k=\phi_k-m(\phi_k(\tau_k))$. As quoted by Conze and Raugi [4], the following martingale-coboundary decomposition is valid: for any $n\in\mathbb{N}$,

$$\tilde{\phi}_n=\psi_n-h_n+h_{n+1}\circ T_{n+1}\,, \tag{5.29}$$

where $(d_n)_{n\ge0}$ defined by $d_n=\psi_n\circ\tau_n$ is a sequence of reversed martingale differences with respect to the filtration $(\mathcal{B}_n)_{n\ge0}$ and $(h_n)_{n\ge0}$ is such that $m(h_n(\tau_n))=0$, and there exists a positive constant K such that $\sup_{n\ge0}\|h_n\|_\infty\le K$.

Set $M_n=\sum_{k=1}^n d_n$ and $V(M_n)=\int_X M_n^2(x)\,m(dx)=\sum_{k=1}^n\int_X d_n^2(x)\,m(dx)$. We have

$$W_1(P_{S_n},G_{V_n})\le W_1(P_{S_n},P_{M_n})+W_1(P_{M_n},G_{V(M_n)})+W_1(G_{V(M_n)},G_{V_n})\,.$$

Using that $W_1(G_{V(M_n)},G_{V_n}))\le\big|\sqrt{V(M_n)}-\sqrt{V_n}\big|\le\|S_n-M_n\|_2$ and the martingale-coboundary decomposition (5.29), it follows that

$$W_1(P_{S_n},G_{V_n})\le W_1(P_{M_n},G_{V(M_n)})+4\sup_{n\ge0}\|h_n\|_\infty\le W_1(P_{M_n},G_{V(M_n)})+4K\,. \tag{5.30}$$

Since $\sup_{n\ge0}\|d_n\|_\infty<\infty$, Corollary 4.4 will follow from Fact 5.1 provided we can suitably handle the quantities $\Big\|\sum_{i=1}^{\ell}(\mathbb{E}(d_i^2|\mathcal{B}_{\ell+1})-\mathbb{E}(d_i^2))\Big\|_1$. With this aim, note that by (5.29), we have

$$d_i^2=\tilde{\phi}_i^2(\tau_i)+2\tilde{\phi}_i(\tau_i)(h_i(\tau_i)-h_{i+1}(\tau_{i+1}))+\big(h_i(\tau_i)-h_{i+1}(\tau_{i+1})\big)^2\,,$$

implying that

$$\|\mathbb{E}(d_i^2|\mathcal{B}_{\ell+1})-\mathbb{E}(d_i^2)\|_\infty\le\|\mathbb{E}(\tilde{\phi}_i^2(\tau_i)-m(\tilde{\phi}_i^2(\tau_i))|\mathcal{B}_{\ell+1})\|_\infty+\|\mathbb{E}(h_i^2(\tau_i)-m(h_i^2(\tau_i))|\mathcal{B}_{\ell+1})\|_\infty \\ +\|\mathbb{E}(h_{i+1}^2(\tau_{i+1})-m(h_{i+1}^2(\tau_{i+1}))|\mathcal{B}_{\ell+1})\|_\infty+2\|\mathbb{E}(\tilde{\phi}_i(\tau_i)h_i(\tau_i)-m(\tilde{\phi}_i(\tau_i)h_i(\tau_i))|\mathcal{B}_{\ell+1})\|_\infty \\ +2\|\mathbb{E}(h_i(\tau_i)h_{i+1}(\tau_{i+1})-m(h_i(\tau_i)h_{i+1}(\tau_{i+1}))|\mathcal{B}_{\ell+1})\|_\infty \\ +2\|\mathbb{E}(\tilde{\phi}_i(\tau_i)h_{i+1}(\tau_{i+1})-m(\tilde{\phi}_i(\tau_i)h_{i+1}(\tau_{i+1}))|\mathcal{B}_{\ell+1})\|_\infty\,. \tag{5.31}$$

From Relations (1.8) and (1.10) in [4], we get that for any function f in $\mathcal{V}$ and any $i \le \ell$,

$$\mathbb{E}(f(\tau_i) - m(f(\tau_i))|\mathcal{B}_{\ell+1}) = \Big(\frac{P_{\ell+1} \circ \cdots \circ P_{i+1}(\tilde{f}_i \pi_i 1)}{\pi_{\ell+1} 1}\Big) \circ \tau_{\ell+1} \,, \tag{5.32}$$

where $\tilde{f}_i = f - m(f\pi_i 1)$. Hence taking into account the properties (DEC) and (MIN), we get that

$$\begin{aligned} \|\mathbb{E}(f(\tau_i) - m(f(\tau_i))|\mathcal{B}_{\ell+1})\|_\infty &\le \kappa\delta^{-1} \|P_{\ell+1} \circ \cdots \circ P_{i+1}(\tilde{f}_i \pi_i 1)\|_v \\ &\le \kappa\delta^{-1} C \gamma^{\ell+1-i} \|\tilde{f}_i \pi_i 1\|_v \,. \end{aligned}$$

Hence, using Relation (3.10) in [4], we get overall that there exists a positive constant M such that, for any function f in $\mathcal{V}$ and any $i \le \ell$,

$$\|\mathbb{E}(f(\tau_i) - m(f(\tau_i))|\mathcal{B}_{\ell+1})\|_\infty \le M \gamma^{\ell+1-i} \|f\|_v \,. \tag{5.33}$$

Taking into account (5.33), it follows that the sum of the four first terms in the right-hand side of (5.31) can be bounded by a positive constant times

$$\gamma^{\ell-i}\Big(\sup_{n\ge 0} \|h_n\|_v^2 + \sup_{n \ge 0} \|\phi_n\|_v^2\Big) \,. \tag{5.34}$$

To take care of the two last terms in (5.31), we shall use the following fact: for any functions f and g in $\mathcal{V}$, by using twice (5.32) and setting

$$Q_{i+1} f = \frac{P_{i+1}(f\pi_i 1)}{\pi_{i+1} 1} \,,$$

the following relation holds: for any $i \le \ell$,

$$\begin{aligned} &\mathbb{E}(f(\tau_i) g(\tau_{i+1})|\mathcal{B}_{\ell+1}) = \mathbb{E}(g(\tau_{i+1}) \mathbb{E}(f(\tau_i)|\mathcal{B}_{i+1})|\mathcal{B}_{\ell+1}) \\ &\quad = \mathbb{E}\Big(g \circ \tau_{i+1}\Big(\frac{P_{i+1}(f\pi_i 1)}{\pi_{i+1} 1}\Big) \circ \tau_{i+1} \Big| \mathcal{B}_{\ell+1}\Big) = \Big(\frac{P_{\ell+1} \circ \cdots \circ P_{i+2}(g Q_{i+1} f \pi_{i+1} 1)}{\pi_{\ell+1} 1}\Big) \circ \tau_{\ell+1} \,. \end{aligned}$$

Therefore, for any functions f and g in $\mathcal{V}$ and any $i \le \ell$,

$$\begin{aligned} &\mathbb{E}(f(\tau_i) g(\tau_{i+1}) - m(f(\tau_i) g(\tau_{i+1}))|\mathcal{B}_{\ell+1}) \\ &\qquad\qquad = \Big(\frac{P_{\ell+1} \circ \cdots \circ P_{i+2}((g Q_{i+1} f - m(g Q_{i+1} f))\pi_{i+1} 1)}{\pi_{\ell+1} 1}\Big) \circ \tau_{\ell+1} \,. \end{aligned}$$

Hence, taking into account the properties (DEC) and (MIN), we get that for any $i \le \ell$,

$$\begin{aligned} &\|\mathbb{E}(f(\tau_i) g(\tau_{i+1}) - m(f(\tau_i) g(\tau_{i+1}))|\mathcal{B}_{\ell+1})\|_\infty \\ &\qquad \le \kappa\delta^{-1} \|P_{\ell+1} \circ \cdots \circ P_{i+2}((g Q_{i+1} f - m(g Q_{i+1} f))\pi_{i+1} 1)\|_v \\ &\qquad \le \kappa\delta^{-1} C \gamma^{\ell-i} \|(g Q_{i+1} f - m(g Q_{i+1} f))\pi_{i+1} 1\|_v \,. \end{aligned}$$

But

$$\|(gQ_{i+1}f - m(gQ_{i+1}f))\pi_{i+1}1\|_v \le \|(gQ_{i+1}f)\pi_{i+1}1\|_v + \|m(gQ_{i+1}f)\pi_{i+1}1\|_v$$
$$\le \|gP_{i+1}(f\pi_i 1)\|_v + \|gQ_{i+1}f\|_\infty \|\pi_{i+1}1\|_v \le \kappa\|g\|_v\|P_{i+1}(f\pi_i 1)\|_v + \|gQ_{i+1}f\|_\infty\|\pi_{i+1}1\|_v\,.$$

By the property (DEC) we have $\|P_{i+1}(f\pi_i 1)\|_v \le \kappa_3\|f\|_v$ where κ_3 is a positive constant not depending on i and on f. On another hand, by the properties (DEC) and (MIN), we have

$$\|gQ_{i+1}f\|_\infty \le \kappa\delta^{-1}\|g\|_\infty\|P_{i+1}(f\pi_i 1)\|_v \le \kappa_4\|f\|_v\|g\|_v\,,$$

where κ_4 is a positive constant not depending on (i, f, g). So overall, there exists a positive constant M such that, for any functions f and g in $\mathcal{V}$ and any $i \le \ell$,

$$\|\mathbb{E}(f(\tau_i)g(\tau_{i+1}) - m(f(\tau_i)g(\tau_{i+1}))|\mathcal{B}_{\ell+1})\|_\infty \le M\gamma^{\ell-i}\|f\|_v\|g\|_v\,. \tag{5.35}$$

Taking into account (5.35), it follows that the sum of the two last terms in the right-hand side of (5.31) can be bounded by a positive constant times the quantity (5.34). So, overall, for any $i \le \ell$,

$$\Big\| |d_{\ell+1}| \big|\mathbb{E}(d_i^2|\mathcal{B}_{\ell+1}) - \mathbb{E}(d_i^2)\big| \Big\|_1 \ll \sup_{n\ge 0}\|d_n\|_\infty \min(\mathbb{E}(d_i^2), \gamma^{\ell-i})\,. \tag{5.36}$$

Therefore, recalling the notation (5.27) and setting $\delta_n^2 = \max_{1\le i\le n}\mathbb{E}(d_i^2)$ and $a^2 = 1+\delta_n^{-2}$, we get

$$\sum_{\ell=1}^{n-1}\frac{\tilde U_{\ell,n}(3)}{a^2\delta_n^2 + \sum_{k=1}^{\ell}\mathbb{E}(d_k^2)} \ll \sum_{\ell=1}^{n-1}\sum_{i=1}^{\ell}\frac{\min(\mathbb{E}(d_i^2),\gamma^{\ell-i})}{1+\delta_n^2+\sum_{k=1}^{i}\mathbb{E}(d_k^2)}\,.$$

Let α be a positive real and $\varphi_\alpha(\ell) = [\alpha\log(\ell)]$. Let $\ell_0 = \inf\{\ell \ge 1 \,:\, \ell - \varphi_\alpha(\ell) \ge 1\}$. We then have

$$\sum_{\ell=1}^{n-1}\frac{\tilde U_{\ell,n}(3)}{a^2\delta_n^2 + \sum_{k=1}^{\ell}\mathbb{E}(d_k^2)} \ll \sum_{\ell=1}^{n-1}\sum_{i=1}^{\ell-\varphi_\alpha(\ell)}\gamma^{\ell-i} + \sum_{\ell=1}^{n-1}\sum_{i=\ell-\varphi_\alpha(\ell)+1}^{\ell}\frac{\mathbb{E}(d_i^2)}{1+\delta_n^2+\sum_{k=1}^{i}\mathbb{E}(d_k^2)}$$
$$\ll \sum_{\ell=1}^{n-1}(1-\gamma)^{-1}\gamma^{\varphi_\alpha(\ell)} + (\log n)\sum_{i=1}^{n-1}\frac{\mathbb{E}(d_i^2)}{1+\delta_n^2+\sum_{k=1}^{i}\mathbb{E}(d_k^2)}\,.$$

Selecting α such that $\alpha\log(1/\gamma) > 1$ and using similar arguments as those developed in Theorem 2.1, it follows that

$$\sum_{\ell=1}^{n-1}\frac{\tilde U_{\ell,n}(3)}{a^2\delta_n^2 + \sum_{k=1}^{\ell}\mathbb{E}(d_k^2)} \ll 1 + (\log n)\log(1+V(M_n))\,.$$

Hence by taking into account this upper bound in (5.28) (with $r = 1$ and $p = 3$), we derive that

$$W_1(P_{M_n}, G_{V(M_n)}) + 4\sup_{n\ge 0}\|h_n\|_\infty$$
$$\ll 1 + \sqrt{\max_{1\le i\le n}\mathbb{E}(d_i^2)} + \Big(\max_{1\le i\le n}\|d_i\|_\infty + \log n\Big)\log(1+V(M_n))\,. \tag{5.37}$$

Starting from (5.30) and considering (5.37) together with the fact that $\sup_{i\geq 1} \|d_i\|_\infty < \infty$ and that there exists a positive constant B such that $V(M_n) \leq 2V_n + B$, the result follows.

Acknowledgement. The authors are indebted to the referee for carefully reading the manuscript and for helpful comments.

References

[1] Berend, D., and Bergelson, V. (1984). Ergodic and Mixing Sequences of Transformations, *Ergod. Th. and Dynam. Syst.* **4** 353-366.

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

[3] Bradley, R. (2002). On positive spectral density. *Bernoulli* **8** 175-193.

[4] Conze, J.-P., and Raugi, A. (2007). Limit theorems for sequential expanding dynamical systems on $[0, 1]$, *Ergodic theory and related fields* 89-121, Contemp. Math., **430** Amer. Math. Soc., Providence, RI.

[5] Dedecker, J., and Merlevède, F. (2011). Rates of convergence in the central limit theorem for linear statistics of martingale differences. *Stoch. Process. Appl.* **121** 1013-1043.

[6] Dedecker, J., Merlevède, F. and Rio, E. (2009). Rates of convergence for minimal distances in the central limit theorem under projective criteria. *Electron. J. Probab.* **14** 978-1011.

[7] Dudley, R. M. (1989). *Real analysis and probability.* Chapman & Hall Mathematics Series. Wadsworth, Inc. Chapman & Hall, New York, London.

[8] El Machkouri, M. and Ouchti, L. (2007). Exact convergence rates in the central limit theorem for a class of martingales. *Bernoulli* **13** no. 4, 981–999.

[9] Fan, X. (2019). Exact rates of convergence in some martingale central limit theorems. *J. Math. Anal. Appl.* **469** no. 2, 1028–1044.

[10] Gordin, M. I. (1969). The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188** 739–741.

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

[12] Helland, I.S. (1982). Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.* **9** no. 2, 79–94.

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

[14] Gänssler, P. and Häusler, E. (1979). Remarks on the functional central limit theorem for martingales. *Z. Wahrsch. verw. Gebiete* **50** no. 3, 237–243.

[15] Grams, W.F. (1972). Rates of convergence in the central limit theorem for dependent variables. *Thesis Florida state University.*

[16] Haydn, N., Nicol, M., Török, A., and Vaienti, S. (2017). Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc.* **369** 5293-5316.

[17] Kolmogorov, A. N.; Rozanov, Ju. A. (1960). On a strong mixing condition for stationary Gaussian processes. *Teor. Verojatnost. i Primenen.* **5** 222–227.

[18] Merlevède, F., Peligrad, M. and Utev, S. *Functional Gaussian Approximation for Dependent Structures.* (2019). Oxford Studies in Probability **6** Oxford University Press.

[19] Mourrat, J.C. (2013). On the rate of convergence in the martingale central limit theorem. *Bernoulli* **19** no. 2, 633–645.

[20] Peligrad, M. (2012). Central limit theorem for triangular arrays of non-homogeneous Markov chains. *Probab. Theory Related Fields* **154** no. 3-4, 409–428

[21] Rio, E. (1995). About the Lindeberg method for strongly mixing sequences. *ESAIM Probab. Statist.* **1** 3561.

[22] Rio, E. (2009). Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** no. 3, 802–817.

[23] Rio, E. (2017). About the conditional value at risk of partial sums. *C. R. Math. Acad. Sci. Paris* **355** no. 11, 1190–1195.

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

[25] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci.U. S. A.* **42** 43–47.

[26] Tikhomirov, A. N. (1980). Convergence rate in the central limit theorem for weakly dependent random variables. *Teor. Veroyatnost. i Primenen.* **25** no. 4, 800–818.

[27] Volný, D. (1993). Approximating martingales and the central limit theorem for strictly stationary processes. *Stoch. Process. Appl.* **44** 41-74.

[28] Wang, X. J. and Hu, S. H. (2018). The Berry-Esseen bound for ρ-mixing random variables and its applications in nonparametric regression model. *Teor. Veroyatnost. i Primenen.* **63** no. 3, 584–608.

[29] Zuparov, T. M. (1991). The rate of convergence in the central limit theorem for weakly dependent variables. *Teor. Veroyatnost. i Primenen.* **36** no. 4, 635–644.