On martingale approximations and the quenched weak invariance principle

Christophe Cuny[a] and Florence Merlevède[b]

[a] Laboratoire MAS, Ecole Centrale de Paris, Grande Voie des Vignes, 92295 Chatenay-Malabry cedex, FRANCE. E-mail: christophe.cuny@ecp.fr

[b] Université Paris Est, LAMA (UMR 8050), UPEMLV, CNRS, UPEC, Bâtiment Copernic, 5 Boulevard Descartes, 77435 Champs-Sur-Marne, FRANCE. E-mail: florence.merlevede@univ-mlv.fr

Abstract

In this paper, we obtain sufficient conditions in terms of projective criteria under which the partial sums of a stationary process with values in $\mathcal{H}$ (a real and separable Hilbert space) admits an approximation, in $\mathbb{L}^p(\mathcal{H})$, $p > 1$, by a martingale with stationary differences and we then estimate the error of approximation in $\mathbb{L}^p(\mathcal{H})$. The results are exploited to further investigate the behavior of the partial sums. In particular we obtain new projective conditions concerning the Marcinkiewicz-Zygmund theorem, the moderate deviations principle and the rates in the central limit theorem in terms of Wasserstein distances. The conditions are well suited for a large variety of examples including linear processes or various kinds of weak dependent or mixing processes. In addition, our approach suits well to investigate the quenched central limit theorem and its invariance principle via martingale approximation, and allows us to show that they hold under the so-called Maxwell-Woodroofe condition that is known to be optimal.

Running head: Martingale approximations in $\mathbb{L}^p$

Key words: martingale approximation, stationary process, quenched invariance principle, moderate deviations, Wasserstein distances, ergodic theorems.

AMS 2000 Subject Classification: Primary: 60F15 ; Secondary: 60F05

1 Introduction

Since the seminal paper of Gordin [15] in 1969, approximation via a martingale is known to be a nice method to derive limit theorems for stochastic processes. For instance, the martingale method has been used successfully by Heyde [20] and Gordin and Lifsic [16] to derive central limit theorems for the partial sums of a stationary sequence, and it has undergone substantial improvements. For recent contributions where the central limit theory and weak convergence problems are handled with the help of martingale approximations, let us mention the recent papers by Maxwell and Woodroofe [23], Wu and Woodroofe [35], Peligrad and Utev [28], Merlevède and Peligrad [24], Zhao and Woodroofe [38] and Gordin and Peligrad [17]. In all these papers, conditions are then imposed to be able to implement the martingale method; namely, to approximate in a suitable way the partial sums of a stationary process by a martingale. However to derive many other kinds of limit theorems from the martingale method, more precise estimates of the approximation error of partial sums by a martingale may be useful. We refer to the recent papers by Wu [34], Zhao and Woodroofe [37], Cuny [4], Dedecker, Doukhan and Merlevède [8] and Merlevède, Peligrad and Peligrad [26] where almost sure behaviors of the partial sums process have been addressed with the help of estimates of this approximation error.

In order to say more about these papers and to present our results, let us first introduce the following notation giving a way to define stationary processes.

Notation 1.1 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\theta : \Omega \mapsto \Omega$ be a bijective bi-measurable transformation preserving the probability $\mathbb{P}$. Let $\mathcal{F}_0$ be a σ-algebra of $\mathcal{A}$ satisfying $\mathcal{F}_0 \subseteq \theta^{-1}(\mathcal{F}_0)$. We then define a non-decreasing filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ by $\mathcal{F}_i = \theta^{-i}(\mathcal{F}_0)$, and a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ \theta^i$ where X_0 is a real-valued centered random variable (or possibly taking values in some real and separable Hilbert space). The sequence will be called adapted to the filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ if X_0 is $\mathcal{F}_0$-measurable. Define then the partial sum by $S_n = X_1 + X_2 + \cdots + X_n$. The following notations will also be used: $\mathcal{F}_{-\infty} = \bigcap_{i\in\mathbb{Z}} \mathcal{F}_i$, $\mathcal{F}_\infty = \bigvee_{i\in\mathbb{Z}} \mathcal{F}_i$, $\mathbb{E}_k(X) = \mathbb{E}(X|\mathcal{F}_k)$, $\mathcal{P}_k(X) = \mathbb{E}_k(X) - \mathbb{E}_{k-1}(X)$, and when X is real-valued, its $\mathbb{L}^p$ norm is denoted by $\|X\|_p = \big(\mathbb{E}(|X|^p)\big)^{1/p}$. We shall also use the notation $a_n \ll b_n$ to mean that there exists a numerical constant C not depending on n such that $a_n \le C b_n$, for all positive integers n.*

In all what follows the sequence $(X_i)_{i\in\mathbb{Z}}$ is assumed to be stationary and adapted to $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ and the variables are in $\mathbb{L}^p$, for some $p > 1$.

In [34] and [8], it is assumed that $D = \sum_{i\ge 0} \mathcal{P}_0(X_i)$ converges in $\mathbb{L}^p$, $p > 1$, and estimates of $\|S_n - M_n\|_p$ where $M_n = \sum_{i=1}^n D \circ \theta^i$ are provided involving either the terms $\sum_{k\ge n} \|\mathcal{P}_0(X_k)\|_p$ (see [34]) or the terms $\|\mathbb{E}_0(S_n)\|_p$ and $\|\sum_{k\ge n} \mathcal{P}_0(X_k)\|_p$ (see [8]). Those estimates are then exploited to derive explicit rates in the almost sure invariance principle under projective conditions that are well adapted to a large variety of examples. The paper by Merlevède *et al* [26] addresses different questions about the almost sure behavior of S_n such as quenched invariance principles or almost sure central limit theorems. Their proof is based under a precise estimate of the $\mathbb{L}^2$ approximation error between the partial sums process and their constructed approximating stationary martingale, provided that the Maxwell-Woodroofe condition (1) holds. More precisely, in the case where $p = 2$, they proved that if

$$\sum_{k=1}^{\infty} \frac{\|\mathbb{E}_0(S_k)\|_2}{k^{3/2}} < \infty\,, \tag{1}$$

then there is a martingale M_n with stationary and square integrable differences such that

$$\|S_n - M_n\|_2 \ll n^{1/2} \sum_{k\ge n} \frac{\|\mathbb{E}_0(S_k)\|_2}{k^{3/2}}\,. \tag{2}$$

To implement a martingale method for other questions related to the behavior of the partial sums, as for instance rates in the strong laws of large numbers or in the central limit theorem in terms of Wasserstein distances, or also moderate deviations principles, the first question that our paper addresses is the construction of a stationary martingale M_n in $\mathbb{L}^p$ ($p > 1$) in such a way that an estimate of $\|S_n - M_n\|_p$ can be given in the spirit of (2). Our Theorem 2.3 is in this direction. When $p \ge 2$, it states in particular that if

$$\sum_{k=1}^{\infty} \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}} < \infty\,, \tag{3}$$

then we can construct a stationary sequence $(D_k = D \circ \theta^k)_{k\in\mathbb{Z}}$ of martingale differences in $\mathbb{L}^p$ adapted to $(\mathcal{F}_k)_{k\in\mathbb{Z}}$ such that setting $M_n = \sum_{k=1}^n D \circ \theta^k$,

$$\|S_n - M_n\|_p \ll n^{1/2} \sum_{k\ge [n^{p/2}]} \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\,. \tag{4}$$

While (4) and (2) coïncide when $p = 2$, our method of proof is different from the one used in [26]. In Theorem 2.3, we shall consider also the case when $p \in]1, 2[$. The main tools to prove the martingale approximation with the bound (4) being algebraic computations and Burkholder's inequality, the

estimate also holds for variables taking values in a separable real Hilbert space. Hence Theorem 2.3 is stated in this setting. As we shall see, this martingale approximation result leads to new projective conditions allowing results concerning the moderate deviations principle or also estimates of Wasserstein distances in the CLT (see Sections 3.2 and 3.3). Notice that the projective conditions assumed all along the paper are general enough to contain a wide class of dependent sequences.

Another interesting point of our approach and of the approximating martingale we consider here, is that they lead not only to a useful estimate of $\|S_n - M_n\|_p$, but, together with a new ergodic theorem with rate (see Theorem 4.7), they allow also to show that, under the Maxwell-Woodroofe condition (1), $\mathbb{E}_0[(S_n - M_n)^2] = o(n)$ $\mathbb{P}$-a.s. (see our Proposition 4.9). This allows to give a definitive positive answer to the question whether the quenched central limit theorem for $n^{-1/2}S_n$ holds true under (1). As we shall see, we can even say more since, using a maximal inequality from Merlevède and Peligrad [25], we establish in Theorem 2.7 that the functional form of the quenched central limit theorem also holds under the Maxwell-Woodroofe condition.

Our paper is structured as follows. Section 2 contains our main results. More precisely, in Section 2.1 we construct an approximating martingale with stationary differences in $\mathbb{L}^p$ that leads to estimates of the $\mathbb{L}^p$ approximating error between the partial sums and the constructed martingale (see Theorem 2.3). In Section 2.2, we address the question of the quenched weak invariance principle under the Maxwell-Woodroofe's condition (1). Section 3 is devoted to some applications of the estimates given in Theorem 2.3 to various kind of limit behavior of the partial sums. In Section 4, we prove the results stated in Sections 2.1 and 2.2 and state a new ergodic theorem with rate (see Theorem 4.7) whose proof is postponed in Section A. Some technical results are given and proven in Section B.

2 Main results

In complement to Notation 1.1, we introduce additional notations used all along the paper.

Notation 2.1 *Let $\mathcal{H}$ be a real and separable Hilbert space equipped with the norm $|\cdot|_{\mathcal{H}}$. For a random variable X with values in $\mathcal{H}$, we denote its norm in $\mathbb{L}^p(\mathcal{H})$ by $\|X\|_{p,\mathcal{H}} = \big(\mathbb{E}(|X|_{\mathcal{H}})^p\big)^{1/p}$, and we simply denote $\mathbb{L}^p = \mathbb{L}^p(\mathbb{R})$.*

Notation 2.2 *Let $p' = \min(2,p)$, $p'' = \max(2,p)$ and $q = p''/p'$.*

2.1 Martingale approximation in $\mathbb{L}^p(\mathcal{H})$

Let $p > 1$. In this section, we shall establish conditions in order for S_n to be approximated by a martingale M_n with stationary differences in $\mathbb{L}^p(\mathcal{H})$ in such a way that the approximation error $\|S_n - M_n\|_{p,\mathcal{H}}$ is explicitly controlled.

Let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p(\mathcal{H})$ in the sense of Notation 1.1. When

$$D = \sum_{n\geq 0}\sum_{k\geq n}\frac{\mathcal{P}_0(X_k)}{k+1}\,, \tag{5}$$

converges in $\mathbb{L}^p(\mathcal{H})$, then $(D_k = D\circ\theta^k)_{k\in\mathbb{Z}}$ forms a stationary sequence of martingale differences in $\mathbb{L}^p(\mathcal{H})$ adapted to $(\mathcal{F}_k)_{k\in\mathbb{Z}}$. Notice that, by Lemma 4.1, the series $\sum_{k\geq 0}\frac{\mathcal{P}_0(X_k)}{k+1}$ converges in $\mathbb{L}^p(\mathcal{H})$ as soon as $X_0 \in \mathbb{L}^p(\mathcal{H})$. In addition, note that the series in (5) converges in $\mathbb{L}^p(\mathcal{H})$ as soon as the series $\sum_{k\geq 0}\mathcal{P}_0(X_k)$ does (see Lemma B.1).

Theorem 2.3 *Let $p > 1$ and let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p(\mathcal{H})$ in the sense of Notation 1.1. Assume that*

$$\sum_{n\geq 1}\frac{\|\mathbb{E}_0(S_n)\|_{p,\mathcal{H}}}{n^{1+1/p''}} < \infty\,. \tag{6}$$

Then $\sum_{n\geq 1}|\sum_{k\geq n}k^{-1}\mathcal{P}_0(X_{k-1})|_{\mathcal{H}}$ *converges in* $\mathbb{L}^p$ *and setting* $M_n=\sum_{k=1}^n D\circ\theta^k$ *where* D *is defined by* (5)*, the following inequality holds:*

$$\|S_n-M_n\|_{p,\mathcal{H}} \ll n^{1/p'}\sum_{k\geq[n^q]}\frac{\|\mathbb{E}_0(S_k)\|_{p,\mathcal{H}}}{k^{1+1/p''}}\,. \tag{7}$$

Remark 2.4 *Let* $p>1$ *and* $\alpha\in]0,1/p'']$*. Let us introduce the following assumption:*

$$\sum_{n\geq 1}\frac{\|\mathbb{E}_0(S_n)\|_{p,\mathcal{H}}}{n^{1+\alpha}}<\infty\,. \tag{8}$$

Assume that (8) *holds with* $\alpha=\min(1/2,2/p^2)$*. By combining (7) with Corollary 22 of [25] (with the norm* $|\cdot|_{\mathcal{H}}$ *replacing the absolute values) we have*

$$\|\max_{1\leq k\leq n}|S_k-M_k|_{\mathcal{H}}\|_p=o(n^{1/p})\,. \tag{9}$$

Notice also that if $p>2$ *and* (8) *holds with* $\alpha\in]2/p^2,1/p]$*, then (7) combined with the maximal inequality (7) of [25] (with the norm* $|\cdot|_{\mathcal{H}}$ *replacing the absolute values) implies that*

$$\|\max_{1\leq k\leq n}|S_k-M_k|_{\mathcal{H}}\|_p=o(n^{\alpha p/2})\,.$$

The fact that the maximal inequality (7) of [25] is still valid when the variables take values in a Hilbert space comes from the fact that its proof is only based on chaining arguments (still valid in functional spaces by replacing the absolute values by the corresponding norms) and on Doob's maximal inequality that also holds in Hilbert spaces. Since Corollary 22 of [25] is proved via their maximal inequality (7), it is still valid in the Hilbert space setting.

Comment 2.5 *Theorem 1 in [34] (still valid in the Hilbert space context) states the following martingale approximation: Let* $p>1$ *and assume that*

$$\mathbb{E}_{-\infty}(X_0)=0\ \mathbb{P}\text{-a.s. and }\sum_{k\geq 0}\|\mathcal{P}_0(X_k)\|_{p,\mathcal{H}}<\infty\,. \tag{10}$$

Then setting $D=\sum_{k\geq 0}\mathcal{P}_0(X_k)$ *and* $M_n=\sum_{i=1}^n D\circ\theta^i$*,*

$$\|S_n-M_n\|_{p,\mathcal{H}}^{p'}\ll\sum_{k=1}^n\Big(\sum_{i\geq k}\|\mathcal{P}_0(X_j)\|_{p,\mathcal{H}}\Big)^{p'}\,. \tag{11}$$

The approximations (7) *and* (11) *cannot be compared and cover distinct classes of dependent sequences. Indeed, there exist examples of processes in* $\mathbb{L}^2$ *satisfying one of the conditions* (1) *or* (10) *but not the other one, see e.g. [14].*

Comment 2.6 *Notice that the quantity* $\|\mathbb{E}_0(S_k)\|_{p,\mathcal{H}}$ *can be estimated in a large variety of examples such as linear processes or mixing sequences. To give an example, let us consider* $p\geq 2$ *and the so-called stationary* ρ*-mixing real sequences defined by the coefficient*

$$\rho(n)=\rho(\mathcal{F}_{-\infty}^0,\mathcal{F}_n^\infty)\ \text{where}\ \mathcal{F}_i^j=\sigma(X_i,\dots,X_j) \tag{12}$$

and

$$\rho(\mathcal{B},\mathcal{C})=\sup\Big\{\frac{\mathrm{Cov}(X,Y)}{\|X\|_2\|Y\|_2}\ :\ X\in L^2(\mathcal{B}),Y\in L^2(\mathcal{C})\Big\}\,.$$

Here $L^2(\mathcal{B})$ denotes the space of real-valued random variables in $\mathbb{L}^2$ that are $\mathcal{B}$-measurable. In the proof of Lemma 1 in [29], it has been proven that for any $p \geq 2$ and any $k \geq 0$,

$$\|\mathbb{E}_0(S_{2^{k+1}})\|_p \ll \sum_{i=0}^{k} 2^{i/2} \rho^{2/p}(2^i)\,, \tag{13}$$

provided that $\sum_{k\geq 0} \rho^{2/p}(2^k) < \infty$. On an other hand, since $\left(\|\mathbb{E}_0(S_n)\|_p\right)_{n\geq 1}$ is a subadditive sequence, it follows from Lemma 2.7 in [28] that, for any $\alpha > 0$, (8) is equivalent to $\sum_{k\geq 0} 2^{-\alpha k}\|\mathbb{E}_0(S_{2^k})\|_p < \infty$. By using (13), one can see that the latter convergence holds provided that, for $\alpha \in]0, 1/2]$, $\sum_{i\geq 0} 2^{i(1/2-\alpha)}\rho^{2/p}(2^i) < \infty$.

2.2 Martingale approximation under $\mathbb{P}_0$ and the quenched (weak) invariance principle

Limit theorems for stochastic processes that do not start from equilibrium are timely and motivated by evolutions in quenched random environment. Recent discoveries by Volný and Woodroofe [32] show that many of the central limit theorems satisfied by classes of stochastic processes in equilibrium, fail to hold when the processes are started from a point. In this section, we address the question whether the Maxwell-Woodroofe condition (1) is sufficient for the validity of the quenched central limit theorem since this condition is known to be optimal (see e.g. [28] or [31] where the optimality of this condition is discussed). This question starts with a result in Borodin and Ibragimov ([1], Ch 4) stating that if $\|\mathbb{E}_0(S_n)\|_2$ is bounded, then one has the CLT starting at a point in its functional form. Later, works by Derriennic and Lin (see [11], [12], [13]), Zhao and Woodroofe [37], Cuny and Lin [5], Cuny [4], Merlevède, Peligrad and Peligrad [26] improved on this result by imposing weaker and weaker conditions on $\|\mathbb{E}_0(S_n)\|_2$, but always stronger than (1). Let us mention that a result in Cuny and Peligrad [6] shows that the condition $\sum_{k=1}^{\infty} \|\mathbb{E}_0(X_k)\|_2/k^{1/2} < \infty$, is sufficient for the quenched CLT. It is also sufficient for the quenched weak invariance principle by a recent result of Cuny and Volny [7].

As we shall see in the proof of Theorem 2.7 below, the approximating martingale that we defined in Section 2.1 also allows to show that, under (1), $\lim_{n\to\infty} n^{-1}\mathbb{E}_0(|S_n - \mathbb{E}_0(S_n) - M_n|^2) = 0$ $\mathbb{P}$-a.s. Combined with a new ergodic theorem with rate (see our Theorem 4.7) and a maximal inequality from Merlevède and Peligrad [25], this implies that the quenched CLT in its functional form holds under the Maxwell-Woodroofe condition (1).

To state that result we need some further notations. Let us first assume the existence of a regular version of the conditional probability on $\mathcal{A}$ given $\mathcal{F}_0$, that is, we assume the existence of a transition probability $K(\cdot,\cdot)$ on $(\Omega, \mathcal{A})$, such that for every $A \in \mathcal{A}$, $K(\cdot, A)$ is a version of $\mathbb{E}(\mathbf{1}_A|\mathcal{F}_0)$. Then, we denote by $\mathbb{E}_\omega$ the expectation with respect to $K(\omega,\cdot)$. We also define the Donsker process W_n by $W_n(t) = n^{-1/2}(S_{[nt]} + (nt - [nt])X_{[nt]+1})$.

Theorem 2.7 *let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^2$ in the sense of Notation 1.1. Assume that (1) holds. Then $\sum_{n\geq 1} |\sum_{k\geq n} k^{-1}\mathcal{P}_0(X_{k-1})|$ converges in $\mathbb{L}^2$ and setting $M_n = \sum_{k=1}^{n} D\circ\theta^k$ where D is defined by (5), the following holds:*

$$\frac{\mathbb{E}_0(\max_{1\leq k\leq n} |S_k - M_k|^2)}{n} \underset{n\to+\infty}{\longrightarrow} 0 \qquad \mathbb{P}\text{-}a.s. \tag{14}$$

In particular, (S_n) satisfies the following quenched weak invariance principle: there exists $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for every $\omega \in \Omega_0$, for any continuous and bounded function f from $(C([0,1]), \|.\|_\infty)$

to $\mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}_\omega(f(W_n)) = \int f(z\sqrt{\eta(\omega)})W(dz)\,, \tag{15}$$

where $\eta = \lim_{n\to\infty} n^{-1}\mathbb{E}(S_n^2|\mathcal{I}) = \lim_{n\to\infty} n^{-1}\mathbb{E}_0(S_n^2)$ *in* $\mathbb{L}^1$*, and* W *is the distribution of a standard Wiener process. Here* $\mathcal{I}$ *is the invariant sigma field, i.e.* $\mathcal{I} = \{A \in \mathcal{A} \ : \ \theta^{-1}(A) = A\}$.

It follows from Comment 2.6 that if the ρ-mixing coefficients of $(X_n)_{n\in\mathbb{Z}}$ satisfy $\sum_{k\geq 0}\rho(2^k) < \infty$, then the quenched invariance principle holds. Hence the CLT from Ibragimov [21] for ρ-mixing sequences that is known to be essentially optimal, is also quenched.
A careful analysis of the proof of Theorem 2.7, shows that if the random variables are assumed to be in $\mathbb{L}^2(\mathcal{H})$, then under (6) with $p = 2$, the almost sure convergence (14) still holds with the norm $|\cdot|_{\mathcal{H}}$ replacing the absolute values.

Theorem 2.7 has an interesting interpretation in the terminology of additive functionals of Markov chains. Let $(\xi_n)_{n\geq 0}$ be a Markov chain with values in a Polish space S, so that there exists a regular transition probability $P_{\xi_1|\xi_0=x}$. Let P be the transition kernel defined by $P(g)(x) = P_{\xi_1|\xi_0=x}(g)$ for any bounded measurable function g from S to $\mathbb{R}$, and assume that there exists an invariant probability π for this transition kernel, that is a probability measure on S such that $\pi(g) = \pi(P(g))$ for any bounded measurable function g from S to $\mathbb{R}$. Let then $\mathbb{L}^2(\pi)$ be the set of functions from S to $\mathbb{R}$ such that $\pi(g^2) < \infty$. For $g \in \mathbb{L}^2(\pi)$ such that $\pi(g) = 0$, define $X_i = g(\xi_i)$. In this setting the condition (1) is $\sum_{n\geq 1} n^{-3/2}\|\sum_{k=1}^n P^k(g)\|_{\mathbb{L}^2(\pi)} < \infty$. In the context of Markov chain the conclusion of Theorem 2.7 is also known under the terminology of functional CLT started at a point. To rephrase it, let $\mathbb{P}^x$ be the probability associated to the Markov chain started from x and let $\mathbb{E}^x$ be the corresponding expectation. Then, for π-almost every $x \in S$, for any continuous and bounded function f from $(C([0,1]), \|.\|_\infty)$ to $\mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}^x(f(W_n)) = \int f(z\sqrt{\eta_x})W(dz)\,, \tag{16}$$

where $\eta_x := \lim_n \mathbb{E}^x(S_n^2)/n$. Note that Theorem 2.7 improves Corollary 5.10 of [4] stated for Markov chains with normal Markov operator. Let us mention that the convergence (16) has also been obtained recently in Dedecker, Merlevède and Peligrad [9] under the condition: $\sum_{k\geq 0}\pi(|gP^k(g)|) < \infty$. The latter condition and (1) are of independent interests (see Section 5.2 of [9]).

3 Applications

As we mentioned in the introduction, having estimates of the approximation error of partial sums by a martingale can be useful to derive different kinds of limit theorems for the partial sums associated with a stationary process. For instance, starting from (2), Merlevède *et al* [26] have obtained sufficient projective conditions in order for the partial sums to satisfy either the law of the iterated logarithm or the almost sure central limit theorem. In this section, we shall use our estimate (7), either to give new projective conditions under which the partial sums associated with a stationary process satisfy a moderate deviations type results, or to analyze the rates of convergence in the CLT in terms of Wasserstein distances. Before stating those results we provide a simple and direct application of our results, leading to new projective criteria to obtain rates in the SLLN.

3.1 Strong law of large numbers with rate

Our martingale approximation in $\mathbb{L}^p$ for $1 < p < 2$ combined with our new ergodic theorem with rate (see Theorem 4.7) allows us to derive very directly a projective condition for the Marcinkiewicz-Zygmund strong law of large numbers.

Theorem 3.1 *Let $1 < p < 2$ and let $(X_n)_{n\in\mathbf{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p(\mathcal{H})$ in the sense of Notation 1.1. Assume that*

$$\sum_{n\geq 2} \log n \frac{\|\mathbb{E}_0(S_n)\|_{p,\mathcal{H}}}{n^{3/2}} < \infty\,.$$

Then, there exists a stationary martingale $(M_n)_{n\geq 1}$ in $\mathbb{L}^p(\mathcal{H})$, such that $|S_n - M_n|_{\mathcal{H}} = o(n^{1/p})$ $\mathbb{P}$-a.s. In particular, we have $|S_n|_{\mathcal{H}} = o(n^{1/p})$ $\mathbb{P}$-a.s.

Proof of Theorem 3.1. Using Theorem 4.7, the first part of the result will follow if we can prove that $\sum_{n\geq 1} n^{-1-1/p}\|S_n - M_n\|_{p,\mathcal{H}} < \infty$. This convergence follows by using Theorem 2.3 to control $\|S_n - M_n\|_{p,\mathcal{H}}$. For the last part of the theorem, it suffices to notice that by the Marcinkiewicz-Zygmund strong law of large numbers for martingales $|M_n|_{\mathcal{H}} = o(n^{1/p})$ $\mathbb{P}$-a.s. for any $p \in]1,2[$ as soon as the martingales are in $\mathbb{L}^p(\mathcal{H})$ (see Woyczyński [33]). □

3.2 Moderate deviations

The aim of this section is to obtain asymptotic expansions for probabilities of moderate deviation for stationary adapted real-valued processes under projective criteria; more precisely we want to study the asymptotic behavior of $\mathbb{P}(S_n \geq \sigma\sqrt{n} r_n)$ where (r_n) is a sequence of positive numbers that diverges to infinity at an appropriate rate and $\sigma = \lim_{n\to\infty} \|S_n\|_2/\sqrt{n}$. Specifically, we aim to find the zone for x of the following moderate deviations principle:

$$\frac{\mathbb{P}(S_n \geq x\sigma\sqrt{n} r_n)}{1 - \Phi(x\, r_n)} = 1 + o(1)\,, \tag{17}$$

where $\Phi(x)$ is the standard normal distribution function. If $r_n = r > 0$ is fixed, then (17) is essentially the well-known central limit theorem. However, for the case when $r = r_n$ is allowed to tend to infinity, the problem of *moderate deviation* probabilities is to find all the possible speed of convergence of $r_n \to \infty$ such that (17) holds. It is a challenging problem to establish moderate deviations principle (MDP) for dependent variables. However starting from the deep results of Grama [18] and of Grama and Haeusler [19] for martingales, Wu and Zhao [36] showed that it is possible to obtain MDP results for a certain class of stationary processes such as functions of an iid sequence as soon as the partial sum process can be well approximated by a martingale. Using our Theorem 2.3, we shall give sufficient conditions for the MDP to hold that are different to the ones obtained by Wu and Zhao [36].

Let us first start with some notations and definitions.

Let $p \in (2,4]$. For $x > 1$, let $r_x > 0$ be the solution to the equation

$$x = (1 + r_x)^{\nu(p)} \exp(r_x^2/2) \text{ where } \nu(p) = \begin{cases} p+1 & \text{if } 2 < p \leq 3 \\ 3p-3 & \text{if } 3 < p \leq 4\,. \end{cases}$$

The function $\nu(p)$ results from the martingale MDP as obtained in [18] and in [19] (see also Theorem 2 and Remark 5 in [36]). In addition, by Remark 1 in [19], as $x \to \infty$, r_x has the asymptotic expansion $r_x^2 = 2\log x - 2[\nu(p) + o(1)]\log(1 + \sqrt{2\log x})$.

Let $\tau_n \to \infty$ be a positive sequence of numbers and (U_n) a sequence of real valued random variables such that $U_n \to^{\mathcal{D}} \mathcal{N}(0,1)$. We shall say that (U_n) satisfies the moderate deviation principle (MDP) with rate τ_n and exponent $p > 0$ if for every $a > 0$ there exists a positive constant $C = C_{a,p}$ depending neither on x nor on n such that

$$\max\left\{ \Big|\frac{\mathbb{P}(U_n \geq r_x)}{1 - \Phi(r_x)} - 1\Big|, \Big|\frac{\mathbb{P}(U_n \leq -r_x)}{1 - \Phi(-r_x)} - 1\Big| \right\} \leq C\Big(\frac{x}{\tau_n}\Big)^{1/(1+p)}\,,$$

holds uniformly in $x \in [1, a\tau_n]$. Therefore τ_n gives a range for which the MDP holds.

Theorem 3.2 *Let $2 < p \leq 4$ and let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p$ in the sense of Notation 1.1. Assume that*

$$\sum_{n\geq 1} \frac{\|\mathbb{E}_0(S_n)\|_p}{n^{1+2/p^2}} < \infty \text{ and } \sum_{n\geq 1} \frac{1}{n^{2/p}} \sum_{k\geq n} \frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^{3/2}} < \infty\,. \tag{18}$$

Assume in addition that

$$\sum_{n\geq 1} \frac{1}{n^{1+2/p}} \|\mathbb{E}_{-n}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} < \infty\,. \tag{19}$$

Then $n^{-1}\mathbb{E}(S_n^2)$ converges to some non-negative number σ^2 and if $\sigma > 0$, $\left(\frac{S_n}{\sigma\sqrt{n}}\right)_{n\geq 1}$ satisfies the MDP with rate $\tau_n = n^{p/2-1}$ and exponent p.

Proof. Analyzing the proof of Theorem 1 in [36], we infer that the theorem will be proven if we can show that there exists a $\mathbb{L}^p$ stationary sequence $(D_i)_{i\in\mathbb{Z}}$ of martingale differences with respect to $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ such that setting $M_n = \sum_{i=1}^n D_i$,

$$\|S_n - M_n\|_p = o(n^{1/p}) \tag{20}$$

and

$$\Big\| \sum_{i=1}^n \mathbb{E}_{i-1}(D_i^2) - \mathbb{E}(D_i^2) \Big\|_{p/2} = O(n^{2/p})\,. \tag{21}$$

According to Theorem 2.3 combined with Remark 2.4, the first part of condition (18) implies (20). On the other hand, since $1 < p/2 \leq 2$, according to Theorem 3 in [36] applied to the stationary sequence $(\mathbb{E}_{i-1}(D_i^2) - \mathbb{E}(D_i^2))_{i\geq 1}$ and using the fact that M_n is a martingale, (21) holds if

$$\sum_{k\geq 0} \frac{1}{2^{2k/p}} \|\mathbb{E}_0(M_{2^k}^2) - \mathbb{E}(M_{2^k}^2)\|_{p/2} < \infty\,. \tag{22}$$

We notice now that since M_n is a stationary martingale, for any $r \geq 1$,

$$\begin{aligned}
\|\mathbb{E}_0(M_{2^k}^2) - \mathbb{E}(M_{2^k}^2)\|_r &= \Big\| \sum_{i=0}^{k-1} \big(\mathbb{E}_{-2^i}(M_{2^i}^2) - \mathbb{E}(M_{2^i}^2)\big) \circ \theta^{2^i} + \big(\mathbb{E}_0(D_1^2) - \mathbb{E}(D_1^2)\big) \Big\|_r \\
&\leq \sum_{i=0}^{k-1} \|\mathbb{E}_{-2^i}(M_{2^i}^2) - \mathbb{E}(M_{2^i}^2)\|_r + \|\mathbb{E}_0(D_1^2) - \mathbb{E}(D_1^2)\|_r \\
&\leq 2\sum_{i=1}^{k-1} \|\mathbb{E}_{-2^i}(M_{2^{i-1}}^2) - \mathbb{E}(M_{2^{i-1}}^2)\|_r + 2\|\mathbb{E}_0(D_1^2) - \mathbb{E}(D_1^2)\|_r\,.
\end{aligned} \tag{23}$$

It follows that (22) is equivalent to: $\sum_{k\geq 0} 2^{-2k/p} \|\mathbb{E}_{-2^{k+1}}(M_{2^k}^2) - \mathbb{E}(M_{2^k}^2)\|_{p/2} < \infty$. Due to the subadditivity of the sequence $(\|\mathbb{E}_{-2n}(M_n^2) - \mathbb{E}(M_n^2)\|_{p/2})_{n\geq 1}$, the latter condition is equivalent to

$$\sum_{n\geq 1} \frac{1}{n^{1+2/p}} \|\mathbb{E}_{-2n}(M_n^2) - \mathbb{E}(M_n^2)\|_{p/2} < \infty\,, \tag{24}$$

(see Lemma 2.7 in [28]). Using now Proposition B.3, we infer that (24) holds if (19) and the second part of (18) do and if: $\sum_{n\geq 1} n^{-(1+4/p^2)} \|\mathbb{E}_0(S_n)\|_p^2 < \infty$. To end the proof, it suffices to notice that since $(\|\mathbb{E}_0(S_n)\|_p)_{n\geq 1}$ is a subadditive sequence, the latter condition is satisfied provided the first part of (18) is (see item 3 of Lemma 37 in [25]). □

The quantities involved in conditions (18) and (19) can be handled by controlling norms of individual summands which involve terms such as $\mathbb{E}_0(X_iX_j)$ and $\mathbb{E}_0(X_i)$. The latter quantities can be then in turn controlled by using various mixing or dependence coefficients (see e.g. [8]). For instance, as a corollary of Theorem 3.2, the following result holds (its proof is omitted since it follows the lines of the proof of Corollary 2.1 in [8]).

Corollary 3.3 *Let $2 < p \leq 4$ and let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p$ in the sense of Notation 1.1. Assume that there exists $\gamma \in]0,1]$ such that*

$$\sum_{n>0} \frac{n^{(p-2)/(\gamma p)}}{n^{1/p}} \|\mathbb{E}_0(X_n)\|_p < \infty \ \text{ and } \ \sum_{n>0} \frac{n^{\gamma}}{n^{2/p}} \sup_{i\geq j\geq n} \|\mathbb{E}_0(X_iX_j) - \mathbb{E}(X_iX_j)\|_{p/2} < \infty\,.$$

Then the conclusion of Theorem 3.2 holds with $\sigma^2 = \sum_{k\in\mathbb{Z}} \mathrm{Cov}(X_0, X_k)$.

As in [8], this result may be used, for instance, to derive under which conditions the partial sum of a function f of the stationary Markov chain $(\xi_k)_{k\in\mathbb{Z}}$ with transition $Kf(x) = \frac{1}{2}(f(x+a) + f(x-a))$, when a is irrational in $[0,1]$ and badly approximable by rationals, satisfy the conclusion of Theorem 3.2. For instance, one can prove that if f is three times differentiable, $\left(\frac{S_n(f)}{\sigma(f)\sqrt{n}}\right)_{n\geq 1}$ satisfies the MDP with rate $\tau_n = n$ and exponent 4 provided that $\sigma(f) > 0$. Here $S_n(f) = \sum_{k=1}^n \big(f(\xi_k) - m(f)\big)$ where m is the Lebesgue-Haar measure and $\sigma^2(f) = m((f - m(f))^2) + 2\sum_{n>0} m(fK^n(f - m(f)))$.

Since in Theorem 3.2 the conditions are expressed in terms of the conditional expectation of the partial sum or of its square, it is also possible to obtain applications for mixing sequences. As an example, the following corollary gives conditions in terms of ρ-mixing coefficients as defined in Comment 2.6.

Corollary 3.4 *Let $2 < p \leq 4$ and let $p \leq \alpha \leq 4$. Let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^\alpha$ in the sense of Notation 1.1. Let $(\rho(n))_{n\geq 1}$ be its associated rho-mixing coefficients as defined in* (12)*. Assume that*

$$\sum_{n\geq 1} \frac{\rho^{2/p}(n)}{n^{1/2+2/p^2}} < \infty \ \text{ and } \ \sum_{n\geq 1} \frac{\rho^{s}(n)}{n^{2/p}} < \infty \ \text{ where } s = 2(\alpha-2)/\alpha\,. \tag{25}$$

Then the conclusion of Theorem 3.2 holds with rate $\tau_n = n^{p/2-1}$ and exponent p.

Notice that if $\alpha = 4$, condition (25) reduces to its first part.

Proof. Let us prove that the first part of (18) holds. With this aim, we first notice that, due to the subadditivity of the sequence $(\|\mathbb{E}_0(S_n)\|_p)_{n\geq 1}$, this condition is equivalent to (see Lemma 2.7 in [28])

$$\sum_{k\geq 0} \frac{\|\mathbb{E}_0(S_{2^k})\|_p}{2^{2k/p^2}} < \infty\,. \tag{26}$$

Since $p > 2$, (25) implies that $\sum_{k\geq 0} \rho^{2/p}(2^k) < \infty$. Therefore, by using (13), it follows that (26) is satisfied as soon as $\sum_{k\geq 0} 2^{-2k/p^2} \sum_{i=0}^k 2^{i/2} \rho^{2/p}(2^i) < \infty$, which is equivalent to the first part of condition (25).

We prove now that the second part of (18) holds. Due to the monotonicity of the sequence $\big(\sum_{\ell\geq n} \ell^{-3/2} \|\mathbb{E}_{-n}(S_\ell)\|_2\big)_{n\geq 1}$, the second part of (18) is equivalent to:

$$\sum_{k\geq 0} \frac{2^k}{2^{2k/p}} \sum_{j\geq k} 2^{-3j/2} \sum_{\ell=2^j}^{2^{j+1}-1} \|\mathbb{E}_{-2^k}(S_\ell)\|_2 < \infty\,. \tag{27}$$

To prove the above condition, we first notice that by stationarity, for any $\ell \in \{2^j, \dots, 2^{j+1}-1\}$,

$$\begin{aligned}\|\mathbb{E}_{-2^k}(S_\ell)\|_2 &\le \|\mathbb{E}_{-2^k}(S_\ell - S_{2^j})\|_2 + \|\mathbb{E}_{-2^k}(S_{2^j})\|_2 \\ &\le \|\mathbb{E}_{-2^k-2^j}(S_{\ell-2^j})\|_2 + \sum_{s=0}^{j-1} \|\mathbb{E}_{-2^k-2^s}(S_{2^s})\|_2 + \|\mathbb{E}_{-2^k}(X_1)\|_2 \,.\end{aligned}$$

Since, for any positive integers r and t, $\|\mathbb{E}_{-r}(S_t)\|_2 \ll \rho(r)\sqrt{t}$, it follows that

$$\sum_{\ell=2^j}^{2^{j+1}-1} \|\mathbb{E}_{-2^k}(S_\ell)\|_2 \ll 2^{3j/2}\rho(2^j) + 2^j\rho(2^k) + 2^j \sum_{s=0}^{j-1} 2^{s/2}\rho(2^k+2^s)\,.$$

So overall, since $p > 2$, we infer that

$$\sum_{k\ge 0} \frac{2^k}{2^{2k/p}} \sum_{j\ge k} 2^{-3j/2} \sum_{\ell=2^j}^{2^{j+1}-1} \|\mathbb{E}_{-2^k}(S_\ell)\|_2 \ll \sum_{k\ge 0} 2^{k(1-2/p)}\rho(2^k)\,. \tag{28}$$

Noticing that (25) implies in particular that

$$\rho(2^k) = o(2^{-k(p^2-4)/(4p)}) \text{ as } k \to \infty\,, \tag{29}$$

and taking into account that $p > 2$, we then infer that the sums in the right-hand side of (28) are finite under (25). This ends the proof of (27), hence the second part of (18) holds.

It remains to show that (19) is satisfied. Note first that since $p \in]2,4]$ and $\alpha \ge p$,

$$\|\mathbb{E}_{-n}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} \le \|\mathbb{E}_{-n}(S_n^2) - \mathbb{E}(S_n^2)\|_{\alpha/2} \le \sup_{Z\in B^{\alpha/(\alpha-2)}(\mathcal{F}_{-n})} \mathrm{Cov}(Z, S_n^2)\,,$$

where $B^r(\mathcal{F}_{-n})$ stands for the set of $\mathcal{F}_{-n}$-measurable random variables such that $\|Z\|_r \le 1$. Using then Theorem 4.12 in [2], we get that

$$\|\mathbb{E}_{-n}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} \le 2^{1-s}\rho^s(n)\|S_n^2\|_{\alpha/2} = 2^{1-s}\rho^s(n)\|S_n\|_\alpha^2\,,$$

where $s = 2(\alpha-2)/\alpha$. Now the first part of (25) implies $\sum_{k>0}\rho^{1/2}(2^k) < \infty$ (see also (29)), therefore $\|S_n\|_\alpha \ll n^{1/2}$ (see [27] or [30]). Hence,

$$\|\mathbb{E}_{-n}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} \ll n\rho^s(n)\,, \tag{30}$$

which proves that (19) holds as soon as the second part of (25) does. This ends the proof of the corollary. □

3.3 Rates of convergence for Wasserstein distances in the CLT

Let $\mathcal{L}(\mu,\nu)$ be the set of probability laws on $\mathbb{R}^2$ with marginals μ and ν. Let us consider the Wasserstein distances of order $r \ge 1$ defined by

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

Let $p \in]2,3[$ and let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p$ in the sense of Notation 1.1. Denote by $P_{S_n/n^{1/2}}$ the law of $S_n/n^{1/2}$ and by G_{σ^2} the normal distribution $\mathcal{N}(0,\sigma^2)$ where $\sigma^2 = \lim_{n\to\infty} n^{-1}\mathbb{E}(S_n^2)$ provided the limit exists. Starting from Theorem 2.1 in [10] and using our Theorem 2.3, we get the following result concerning the order of $W_r^r(P_{n^{-1/2}S_n}, G_{\sigma^2})$ where $r \in [1,p]$.

Theorem 3.5 *Let $2 < p \leq 3$ and let $1 \leq r \leq p$. Let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p$ in the sense of Notation 1.1. Assume that (19) holds and that*

$$\sum_{n\geq 1} \frac{1}{n^{3-p/2}} \|\mathbb{E}_{-n}(S_n^2) - \mathbb{E}(S_n^2)\|_{1+\gamma} < \infty \text{ for some } \gamma > 0\,. \tag{31}$$

Assume in addition that

$$\sum_{n\geq 1} \frac{\|\mathbb{E}_0(S_n)\|_p^2}{n^{1+4/p^2}} < \infty\,, \tag{32}$$

and that

$$\sum_{n\geq 1} \frac{\|\mathbb{E}_0(S_n)\|_2}{n^{(5-p)/2}} < \infty \quad \text{if } r \in [1,2] \text{ and } \|\mathbb{E}_0(S_n)\|_r = O(n^{(3-p)/r}) \quad \text{if } r \in]2,p]\,. \tag{33}$$

Then $n^{-1}\mathbb{E}(S_n^2)$ converges to some non-negative number σ^2, and $W_r^r(P_{n^{-1/2}S_n}, G_{\sigma^2}) = O(n^{1-p/2})$.

The above result improves Theorem 3.1 in Dedecker, Merlevède and Rio [10] that imposes the series $\sum_{n>0} \mathbb{E}(X_n|\mathcal{F}_0)$ to converge in $\mathbb{L}^p$ instead of the weaker conditions (32) and (33).

When ρ-mixing sequences are considered, applying Theorem 3.5 we derive the following corollary (its proof is omitted since it uses similar bounds as those obtained in the proof of Corollary 3.4).

Corollary 3.6 *Let $2 < p \leq 3$ and let $p \leq \alpha \leq 4$. Let $(X_n)_{n\in\mathbb{Z}}$ be a adapted stationary sequence in $\mathbb{L}^\alpha$ in the sense of Notation 1.1. Let $(\rho(n))_{n\geq 1}$ be its associated rho-mixing coefficients as defined in* (12)*. Assume that*

$$\sum_{n\geq 1} \frac{\rho^s(n)}{n^{2-p/2}} < \infty \text{ where } s = 2(\alpha-2)/\alpha\,.$$

Then the conclusion of Theorem 3.5 holds for any $1 \leq r \leq 2$.

Proof of Theorem 3.5. Notice first that (32) implies in particular that $\|\mathbb{E}_0(S_n)\|_p = o(n^{2/p^2})$ (apply for instance Item 2 of Lemma 37 in [25] to the sequence $(\|\mathbb{E}_0(S_n)\|_p^2)_{n\geq 0}$). Now, since $p > 2$, (32) then entails that (6) holds true. Therefore, by Theorem 2.3, D defined by (5) is in $\mathbb{L}^p$. In addition, since $p > 2$, (6) implies that $\sum_{n>0} n^{-3/2}\|\mathbb{E}_0(S_n)\|_2 < \infty$ which is a sufficient condition for $n^{-1}\mathbb{E}(S_n^2)$ to converge (see Theorem 1 in [28]).

Let now $M_n = \sum_{k=1}^n D \circ \theta^k$ and $R_n = S_n - M_n$. According to the proof of Theorem 3.1 in [10] and to their remark 2.1, the theorem will follow if we can prove that

$$\|R_n\|_r = O(n^{(3-p)/2})\,. \tag{34}$$

and also that

$$\sum_{k\geq 0} \frac{\|\mathbb{E}_0(M_{2^k}^2) - \mathbb{E}(M_{2^k}^2)\|_{1+\gamma}}{2^{k(2-p/2)}} < \infty \text{ for a } \gamma > 0 \text{ and } \sum_{k\geq 0} \frac{\|\mathbb{E}_0(M_{2^k}^2) - \mathbb{E}(M_{2^k}^2)\|_{p/2}}{2^{2k/p}} < \infty\,.$$

Using (23) and the subadditivity of the sequence $(\|\mathbb{E}_{-2n}(M_n^2) - \mathbb{E}(M_n^2)\|_q)_{n\geq 1}$, for any $q \geq 1$, we infer that the latter conditions are equivalent to

$$\sum_{n\geq 1} \frac{\|\mathbb{E}_{-2n}(M_n^2) - \mathbb{E}(M_n^2)\|_{1+\gamma}}{n^{3-p/2}} < \infty \text{ for a } \gamma > 0 \text{ and } \sum_{n\geq 1} \frac{\|\mathbb{E}_{-2n}(M_n^2) - \mathbb{E}(M_n^2)\|_{p/2}}{n^{1+2/p}} < \infty\,. \tag{35}$$

Using Proposition B.3 we infer that (35) holds provided that (19) and (31) do, and that

$$\sum_{n\geq 1}\frac{\|\mathbb{E}_0(S_n)\|_p^2}{n^{1+4/p^2}}<\infty\,,\ \sum_{n\geq 1}\frac{\|\mathbb{E}_0(S_n)\|_{2(1+\gamma)}^2}{n^{1+(4-p)/(2+2\gamma)}}<\infty \text{ and } \sum_{n\geq 1}\frac{\|\mathbb{E}_0(S_n)\|_2}{n^{(5-p)/2}}<\infty\,. \tag{36}$$

Notice first that the third part of (36) holds provided that (33) does (notice that the second part of (33), for $r>2$ implies the first part of (33)), whereas the first part of (36) is exactly condition (32). Notice now that for any $p\in]2,3[$ and γ small enough, $(4-p)/(2+2\gamma)\geq 4/p^2$ and $p\geq 2+2\gamma$. Therefore the second part of (36) is implied by condition (32).

It remains to prove (34). By Lemma 2.7 of [28], the first part of (33) implies that $\|\mathbb{E}_0(S_n)\|_2=o(n^{(3-p)/2})$. Therefore by using Theorem 2.3, we infer that, since $p>2$, for any r in $[1,2]$, $\|R_n\|_r\leq\|R_n\|_2=o(n^{(3-p)/2})$ under the first part of (33). Now, since $p>2$, for any r in $]2,p]$, the second part of (33) implies that $\|R_n\|_r=O(n^{(3-p)/2})$ by Theorem 2.3. □

4 Proof of the martingale approximation results

In all the following lemmas, $p>1$ and $(X_n)_{n\in\mathbb{Z}}$ is an adapted stationary sequence in $\mathbb{L}^p(\mathcal{H})$ in the sense of Notation 1.1.

Lemma 4.1 *We have* $\sum_{k\geq 0}(k+1)^{-1}\|\mathcal{P}_0(X_k)\|_{p,\mathcal{H}}<\infty$.

Proof. We first prove the case $p\geq 2$. By Hölder's inequality, we have

$$\begin{aligned}\Big(\sum_{k\geq 0}\frac{\|\mathcal{P}_0(X_k)\|_{p,\mathcal{H}}}{k+1}\Big)^p &\ll \sum_{k\geq 0}\|\mathcal{P}_{-k}(X_0)\|_{p,\mathcal{H}}^p \ll \Big\|\Big(\sum_{k\geq 0}|\mathcal{P}_{-k}(X_0)|_{\mathcal{H}}^p\Big)^{1/p}\Big\|_p^p\\ &\ll \Big\|\Big(\sum_{k\geq 0}|\mathcal{P}_{-k}(X_0)|_{\mathcal{H}}^2\Big)^{1/2}\Big\|_p^p \ll \|X_0\|_{p,\mathcal{H}}^p\,,\end{aligned}$$

where we used $\|\cdot\|_{\ell^p}\leq\|\cdot\|_{\ell^2}$ and Burkholder's inequality for $\mathcal{H}$-valued martingales (see [3]).

Let prove the case $1<p<2$. By Hölder inequality

$$\begin{aligned}\Big(\sum_{k\geq 0}\frac{\|\mathcal{P}_0(X_k)\|_{p,\mathcal{H}}}{k+1}\Big)^p &\ll \sum_{k\geq 0}\frac{\|\mathcal{P}_{-k}(X_0)\|_{p,\mathcal{H}}^p}{(k+1)^{p/2}} = \mathbb{E}\Big(\sum_{k\geq 0}\frac{|\mathcal{P}_{-k}(X_0)|_{\mathcal{H}}^p}{(k+1)^{p/2}}\Big)\\ &\ll \Big\|\Big(\sum_{k\geq 0}|\mathcal{P}_{-k}(X_0)|_{\mathcal{H}}^2\Big)^{1/2}\Big\|_p^p \ll \|X_0\|_{p,\mathcal{H}}^p\,,\end{aligned}$$

where we used again Hölder's inequality and Burkholder's inequality for $\mathcal{H}$-valued martingales. □

Lemma 4.2 *Assume that*

$$\sum_{n\geq 1}\sum_{k\geq 0}\frac{\|\mathcal{P}_0(S_n\circ\theta^{k-1})\|_{p,\mathcal{H}}}{(n+k)^2}<\infty\,. \tag{37}$$

Then $\sum_{n\geq 0}|\sum_{k\geq n}\frac{\mathcal{P}_0(X_k)}{k+1}|_{\mathcal{H}}$ *converges in* $\mathbb{L}^p$ *and a.s. Moreover for any integer* $m\geq 0$,

$$\Big\|\sum_{n\geq m}\sum_{k\geq n}\frac{\mathcal{P}_0(X_k)}{k+1}\Big\|_{p,\mathcal{H}}\leq\sum_{k\geq m}\sum_{n\geq 1}\frac{\|\mathcal{P}_0(S_n\circ\theta^{k-1})\|_{p,\mathcal{H}}}{(n+k)^2}\,. \tag{38}$$

Proof. By assumption, the series

$$\sum_{k\geq 0}\Big|\sum_{n\geq 1}\frac{\sum_{l=0}^{n-1}\mathcal{P}_0(X_{l+k})}{(n+k)(n+k+1)}\Big|_{\mathcal{H}}$$

converges a.s. and in $\mathbb{L}^p$. On the other hand, using Lemma 4.1 to invert the order of summation, we have

$$\sum_{l\geq k}\frac{\mathcal{P}_0(X_l)}{l+1}=\sum_{l\geq 0}\frac{\mathcal{P}_0(X_{k+l})}{k+l+1}=\sum_{l\geq 0}\sum_{n\geq l+1}\frac{\mathcal{P}_0(X_{k+l})}{(k+n)(k+n+1)}=\sum_{n\geq 1}\frac{\sum_{l=0}^{n-1}\mathcal{P}_0(X_{l+k})}{(n+k)(n+k+1)}\,,$$

which gives the desired convergence. □

Lemma 4.3 *For every integer $r\geq 0$,*

$$\sum_{k\geq r}\sum_{m\geq 1}\frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k)^2}\ll\sum_{k\geq r+1}\frac{\|\mathbb{E}_{-r}(S_k)\|_{p,\mathcal{H}}}{k^{1+1/p''}}\,. \tag{39}$$

Proof. Let m be a positive integer. Assume first that $p\geq 2$. By Hölder's inequality and using that $\|.\|_{\ell^p}\leq\|.\|_{\ell^2}$, we have

$$\begin{aligned}\sum_{k\geq r}\frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k)^2}&\ll(m+r)^{-(1+1/p)}\big(\sum_{k\geq r}\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}^p\big)^{1/p}\\&\ll(m+r)^{-(1+1/p)}\Big(\mathbb{E}\Big(\sum_{k\geq r}|\mathcal{P}_{-k}(S_m)|_{\mathcal{H}}^2\Big)^{p/2}\Big)^{1/p}\ll\frac{\|\mathbb{E}_{-r}(S_m)\|_{p,\mathcal{H}}}{(m+r)^{1+1/p}}\,,\end{aligned}$$

where we used Burkholder's inequality for $\mathcal{H}$-valued martingales (see [3]), in the last step.

Assume now that $1<p<2$. We use Hölder's inequality twice and once again Burkholder's inequality for $\mathcal{H}$-valued martingales in the last step, to obtain

$$\begin{aligned}\sum_{k\geq r}\frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k)^2}&\ll\frac{1}{(m+r)^{1/p}}\Big(\sum_{k\geq r}\frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}^p}{(m+k)^p}\Big)^{1/p}\\&\ll\frac{1}{(m+r)^{1/p}}\left(\frac{1}{(m+r)^{3p/2-1}}\mathbb{E}\Big(\sum_{k\geq r}|\mathcal{P}_{-k}(S_m)|_{\mathcal{H}}^2\Big)^{p/2}\right)^{1/p}\ll\frac{\|\mathbb{E}_{-r}(S_m)\|_{p,\mathcal{H}}}{(m+r)^{3/2}}\,.\end{aligned}$$

From the above computations, we then derive that

$$\begin{aligned}\sum_{k\geq r}\sum_{m\geq 1}\frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k)^2}&\ll\sum_{m\geq 1}\frac{\|\mathbb{E}_{-r}(S_m)\|_{p,\mathcal{H}}}{(m+r)^{1+1/p''}}\\&\ll\frac{1}{(r+1)^{1/p''}}\max_{1\leq m\leq r}\|\mathbb{E}_{-r}(S_m)\|_{p,\mathcal{H}}+\sum_{m\geq r+1}\frac{\|\mathbb{E}_{-r}(S_m)\|_{p,\mathcal{H}}}{m^{1+1/p''}}\,.\end{aligned}$$

The lemma then follows by using Lemma B.2 with $\gamma=1/p''$ and $\ell=r$. □

Lemma 4.4 *For every $r\geq 0$,*

$$X_0=\sum_{k=0}^{r}\sum_{l\geq k}\frac{\mathcal{P}_0(X_l)}{l+1}-\sum_{k=0}^{r}\sum_{l\geq k}\frac{\mathbb{E}_0(X_{l+1})-\mathbb{E}_{-1}(X_l)}{l+1}+(r+1)\sum_{l\geq r}\frac{\mathbb{E}_0(X_{l+1})}{(l+1)(l+2)}\,. \tag{40}$$

In particular, if we assume (6)*, letting* $r \to \infty$*, we have*

$$X_0 = \sum_{k\geq 0}\sum_{l\geq k} \frac{\mathcal{P}_0(X_l)}{l+1} - \sum_{k\geq 0}\sum_{l\geq k} \frac{\mathbb{E}_0(X_{l+1}) - \mathbb{E}_{-1}(X_l)}{l+1}\,.$$

Proof. Let $m \geq k \geq 0$. We have

$$\sum_{l=k}^{m} \frac{\mathcal{P}_0(X_l)}{l+1} = \frac{\mathbb{E}_0(X_k)}{k+1} - \frac{\mathbb{E}_0(X_{m+1})}{m+2} + \sum_{l=k}^{m} \frac{\mathbb{E}_0(X_{l+1})}{l+2} - \sum_{l=k}^{m} \frac{\mathbb{E}_{-1}(X_l)}{l+1}\,.$$

Hence

$$\sum_{l=k}^{m} \frac{\mathcal{P}_0(X_l)}{l+1} = \frac{\mathbb{E}_0(X_k)}{k+1} - \frac{\mathbb{E}_0(X_{m+1})}{m+2} + \sum_{l=k}^{m} \frac{\mathbb{E}_0(X_{l+1}) - \mathbb{E}_{-1}(X_l)}{l+1} - \sum_{l=k}^{m} \frac{\mathbb{E}_0(X_{l+1})}{(l+1)(l+2)}\,.$$

Notice that $m^{-1}\|\mathbb{E}_0(X_m)\|_{p,\mathcal{H}} \to 0$ and that $\sum_{l\geq 0} \frac{\|\mathbb{E}_0(X_{l+1})\|_{p,\mathcal{H}}}{(l+1)(l+2)} < \infty$. Hence, using Lemma 4.1, we may and do let $m \to \infty$, to obtain

$$\sum_{l\geq k} \frac{\mathcal{P}_0(X_l)}{l+1} = \frac{\mathbb{E}_0(X_k)}{k+1} + \sum_{l\geq k} \frac{\mathbb{E}_0(X_{l+1}) - \mathbb{E}_{-1}(X_l)}{l+1} - \sum_{l\geq k} \frac{\mathbb{E}_0(X_{l+1})}{(l+1)(l+2)}\,.$$

Let $r \geq 0$. We then deduce that

$$\sum_{k=0}^{r}\sum_{l\geq k} \frac{\mathcal{P}_0(X_l)}{l+1} = \sum_{k=0}^{r} \frac{\mathbb{E}_0(X_k)}{k+1} + \sum_{k=0}^{r}\sum_{l\geq k} \frac{\mathbb{E}_0(X_{l+1}) - \mathbb{E}_{-1}(X_l)}{l+1} - \sum_{k=0}^{r}\sum_{l\geq k} \frac{\mathbb{E}_0(X_{l+1})}{(l+1)(l+2)}\,.$$

Hence, interverting the order of summation in the last term,

$$\sum_{k=0}^{r}\sum_{l\geq k} \frac{\mathcal{P}_0(X_l)}{l+1} = X_0 + \sum_{k=0}^{r}\sum_{l\geq k} \frac{\mathbb{E}_0(X_{l+1}) - \mathbb{E}_{-1}(X_l)}{l+1} - (r+1)\sum_{l\geq r} \frac{\mathbb{E}_0(X_{l+1})}{(l+1)(l+2)}\,.$$

Assume (6). In view of Lemmas 4.2 and 4.3, we see that the series on the left converges in $\mathbb{L}^p(\mathcal{H})$. On an other hand, Lemma B.2 (with $\gamma = 1$) implies that $n^{-1}\|\mathbb{E}_0(S_n)\|_{p,\mathcal{H}} \to 0$. Therefore by Abel summation, $\|(r+1)\sum_{l\geq r} \frac{\mathbb{E}_0(X_{l+1})}{(l+1)(l+2)}\|_{p,\mathcal{H}} \to 0$, when $r \to \infty$. □

4.1 Proof of Theorem 2.3

The first assertion comes from Lemma 4.2 combined with Lemma 4.3. Now, by Lemma 4.4, we have

$$X_1 = D \circ \theta - \sum_{k\geq 0}\sum_{l\geq k+1} \frac{\mathbb{E}_1(X_{l+1}) - \mathbb{E}_0(X_l)}{l}.$$

Hence, using that $\mathbb{E}_1(X_{l+1}) = \mathbb{E}_0(X_l) \circ \theta$, we obtain that for any positive integer n,

$$S_n - M_n = -\sum_{k\geq 0}\sum_{l\geq k+1} \frac{\mathbb{E}_0(X_l) \circ \theta^n - \mathbb{E}_0(X_l)}{l} = -\sum_{k\geq 0}\sum_{l\geq k+1} \frac{\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_l)}{l}\,.$$

Let N be a positive integer, fixed for the moment. Then writing

$$V_{n,N} = \sum_{k=0}^{N-1}\sum_{l\geq k+1} \frac{\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_{l+n})}{l}\,, \tag{41}$$

and

$$W_{n,N} = \sum_{k\geq N} \sum_{l\geq k+1} \frac{\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_{l+n})}{l} \,. \tag{42}$$

we obtain

$$S_n - M_n - \mathbb{E}_0(S_n) = -\sum_{k\geq 0} \sum_{l\geq k+1} \frac{\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_{l+n})}{l} = -(V_{n,N} + W_{n,N}) \,. \tag{43}$$

We first deal with $V_{n,N}$. We have

$$\begin{aligned} V_{n,N} &= \sum_{l=1}^{N} \big(\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_{l+n})\big) + N \sum_{l\geq N+1} \frac{\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_{l+n})}{l} \\ &= \mathbb{E}_0(S_N)\circ\theta^n - \mathbb{E}_0(S_N\circ\theta^n) + N \sum_{l\geq N+1} \frac{\mathbb{E}_n(X_l\circ\theta^n) - \mathbb{E}_0(X_l\circ\theta^n)}{l} \,. \end{aligned} \tag{44}$$

Let $j \in \{0, n\}$. By (6) and Lemma B.2 with $\gamma = 1$,

$$\frac{\|\mathbb{E}_0(S_N)\|_{p,\mathcal{H}}}{N} \ll \sum_{l\geq N} \frac{\|\mathbb{E}_0(S_l)\|_{p,\mathcal{H}}}{l^2} = o(1). \tag{45}$$

Using Abel summation we have, for every $s \geq N+1$,

$$\begin{aligned} \sum_{l=N+1}^{s} \frac{\mathbb{E}_j(X_l\circ\theta^n)}{l} &= \sum_{l=N+1}^{s} \frac{\mathbb{E}_j(S_l\circ\theta^n - S_{l-1}\circ\theta^n)}{l} \\ &= -\frac{\mathbb{E}_j(S_N\circ\theta^n)}{N+1} + \frac{\mathbb{E}_j(S_s\circ\theta^n)}{s+1} + \sum_{l=N+1}^{s} \frac{\mathbb{E}_j(S_l\circ\theta^n)}{l(l+1)} \,. \end{aligned}$$

Letting $s \to \infty$, it follows from (45) that

$$\sum_{l\geq N+1} \frac{\mathbb{E}_j(X_l\circ\theta^n)}{l} = -\frac{\mathbb{E}_j(S_N\circ\theta^n)}{N+1} + \sum_{l\geq N+1} \frac{\mathbb{E}_j(S_l\circ\theta^n)}{l(l+1)} \,. \tag{46}$$

Hence, starting from (44) and considering (46) and (45), we derive that

$$\begin{aligned} \|V_{n,N}\|_{p,\mathcal{H}} &\leq 2\frac{\|\mathbb{E}_0(S_N)\|_{p,\mathcal{H}}}{N} + N \sum_{l\geq N+1} \frac{\|\mathbb{E}_n(S_l\circ\theta^n) - \mathbb{E}_0(S_l\circ\theta^n)\|_{p,\mathcal{H}}}{l(l+1)} \\ &\ll N \sum_{l\geq N} \frac{\|\mathbb{E}_0(S_l)\|_{p,\mathcal{H}}}{l^2} \,. \end{aligned} \tag{47}$$

It remains to deal with $W_{n,N}$. Since $\mathbb{E}_0(W_{n,N}) = 0$, we have $W_{n,N} = \sum_{r=1}^{n} \mathcal{P}_r(W_{n,N})$. Using that $\mathcal{P}_r$ defines a continuous operator on $\mathbb{L}^p(\mathcal{H})$ and that the series in (42) converges in $\mathbb{L}^p(\mathcal{H})$, we infer that

$$W_{n,N} = \sum_{r=1}^{n} \sum_{k\geq N} \sum_{l\geq k+1} \frac{\mathbb{E}_r(X_{l+n}) - \mathbb{E}_{r-1}(X_{l+n})}{l} \,. \tag{48}$$

But, by Burkholder's inequality for $\mathcal{H}$-valued martingales (see [3]),

$$\|W_{n,N}\|_{p,\mathcal{H}}^{p'} \ll \sum_{r=1}^{n} \|\mathcal{P}_r(W_{n,N})\|_{p,\mathcal{H}}^{p'} \,. \tag{49}$$

Notice that for any $r \in \{1, \dots, n\}$,

$$\mathcal{P}_r(W_{n,N}) = \Big(\sum_{k \geq N} \sum_{l \geq 1} \frac{\mathcal{P}_0(X_{l+k+n-r})}{l+k} \Big) \circ \theta^r \,.$$

Now, using Lemma 4.1,

$$\sum_{l \geq 1} \frac{\mathcal{P}_0(X_{l+k+n-r})}{l+k} = \sum_{l \geq 1} \mathcal{P}_0(X_{l+k+n-r}) \sum_{m \geq l} \frac{1}{(m+k)(m+k+1)} = \sum_{m \geq 1} \frac{\mathcal{P}_0(S_m \circ \theta^{k+n-r})}{(m+k)(m+k+1)} \,.$$

Therefore,

$$\Big| \sum_{k \geq N} \sum_{l \geq 1} \frac{\mathcal{P}_0(X_{l+k+n-r})}{l+k} \Big| \leq \sum_{m \geq 1} \sum_{k \geq N} \frac{|\mathcal{P}_0(S_m \circ \theta^{k+n-r})|}{(m+k)^2} \,. \tag{50}$$

Hence, with $s = n - r$,

$$\|W_{n,N}\|_{p,\mathcal{H}} \ll n^{1/p'} \max_{0 \leq s \leq n-1} \sum_{k \geq N+s} \sum_{m \geq 1} \frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k-s)^2} \,.$$

Now we take $N = u_n \geq n$. We then infer that

$$\|W_{n,u_n}\|_{p,\mathcal{H}} \ll n^{1/p'} \sum_{k \geq u_n} \sum_{m \geq 1} \frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k)^2} \,. \tag{51}$$

Hence using (43), (47) with $N = u_n$, and (51), we get that

$$\|S_n - M_n\|_{p,\mathcal{H}} \ll \|\mathbb{E}_0(S_n)\|_{p,\mathcal{H}} + u_n \sum_{m \geq u_n} \frac{\|\mathbb{E}_0(S_m)\|_{p,\mathcal{H}}}{m^2} + n^{1/p'} \sum_{k \geq u_n} \sum_{m \geq 1} \frac{\|\mathcal{P}_{-k}(S_m)\|_{p,\mathcal{H}}}{(m+k)^2} \,. \tag{52}$$

Next using Lemma B.2 with $\gamma = 1$, we derive that

$$\|\mathbb{E}_0(S_n)\|_{p,\mathcal{H}} \leq \max_{1 \leq k \leq u_n} \|\mathbb{E}_0(S_k)\|_{p,\mathcal{H}} \ll u_n \sum_{m \geq u_n} \frac{\|\mathbb{E}_0(S_m)\|_{p,\mathcal{H}}}{m^2} \,. \tag{53}$$

Starting from (52) with $u_n = [n^q]$ and taking into account (53) and Lemma 4.3, Theorem 2.3 follows. □

4.2 Proof of Theorem 2.7

Part of the proof relies on a new ergodic theorem with rate. Hence we first recall some facts from ergodic theory and state our ergodic theorem, while we give its proof in Section A.

Let T be a Dunford-Schwartz operator on Ω, i.e. T is a contraction of $\mathbb{L}^1$ and $\mathbb{L}^\infty$. Let $\mathbf{T}$ be the linear modulus of T (see e.g. Theorem 1.1, chapter 4 of [22]). Recall that $\mathbf{T}$ is a positive Dunford-Schwartz operator such that $|Tf| \leq \mathbf{T}|f|$, for every $f \in \mathbb{L}^1$ and $|Tf|^p \leq \mathbf{T}(|f|^p)$, for every $f \in \mathbb{L}^p$.

We will make use, for $p \geq 1$, of the weak $\mathbb{L}^p$-spaces

$$\mathbb{L}^{p,w} := \{ f \in \mathbb{L}^0 \ : \ \sup_{\lambda > 0} \lambda^p \mathbb{P}\{|f| \geq \lambda\} < \infty \} \,,$$

where $\mathbb{L}^0$ is the space of all $\mathcal{A} - \mathcal{B}(\mathbb{R})$ measurable functions.

Recall that, when $p > 1$, there exists a norm $\|\cdot\|_{p,w}$ on $\mathbb{L}^{p,w}$ that makes $\mathbb{L}^{p,w}$ a Banach space and which is equivalent to the "pseudo"-norm $(\sup_{\lambda>0} \lambda^p \mathbb{P}\{|f| \geq \lambda\})^{1/p}$.

We define, for every $l \geq 0$, a maximal operator as follows. For any non-negative function $h \in \mathbb{L}^1$, let

$$\mathcal{M}_l(h) = \sup_{n\geq 1} \frac{h + \mathbf{T}^{2^l} h + \ldots + (\mathbf{T}^{2^l})^{n-1} h}{n}.$$

By the Dunford-Schwartz (or Hopf) ergodic theorem (see e.g. Krengel [22], Lemma 6.1 page 51 and Corollary 3.8 p. 131),

$$\sup_{\lambda>0} \lambda \mathbb{P}\{\mathcal{M}_l(h) \geq \lambda\} \leq \|h\|_1 .$$

In particular, for every $p > 1$, there exists $C_p > 0$ such that, for every $f \in \mathbb{L}^p$,

$$\|(\mathcal{M}_l(|f|^p))^{1/p}\|_{p,w} \leq C_p \|f\|_p . \tag{54}$$

Let $\mathcal{B}$ be a Banach space with norm $|\cdot|_{\mathcal{B}}$. For every $p \geq 1$, we denote by $\mathbb{L}^p(\mathcal{B})$ the Bochner space $\{f \ : \ \Omega \to \mathcal{B} \ , |f|_{\mathcal{B}} \in \mathbb{L}^p\}$. When T is induced by a measurable transformation θ preserving $\mathbb{P}$, $\mathcal{M}_l(|f|_{\mathcal{B}})$ is well-defined for every $f \in \mathbb{L}^1(\mathcal{B})$. We prove the following, where $U_n(f) = f + \ldots + T^{n-1} f$.

Proposition 4.5 *Let T be a Dunford-Schwartz operator on $(\Omega, \mathcal{A}, \mathbb{P})$ and $f \in \mathbb{L}^1$. We have*

$$\max_{1\leq n\leq 2^r} |U_n(f)| \leq 2^{r/p} \sum_{k=0}^{r} \frac{\left[\mathcal{M}_k(|U_{2^k}(f)|^p)\right]^{1/p}}{2^{k/p}} .$$

When T is induced by a measure preserving transformation θ and $\mathcal{B}$ is a Banach space, the result holds also for $f \in \mathbb{L}^1(\mathcal{B})$, replacing $|\cdot|$ with $|\cdot|_{\mathcal{B}}$.

Proof. The proof follows from the following lemma, using that $U_{2^k m}(f) - U_{2^k(m-1)}(f) = T^{2^k(m-1)} f + \ldots + T^{2^k m - 1} f = (T^{2^k})^{(m-1)} U_{2^k}(f)$. □

Lemma 4.6 *Let (a_n) be a sequence in a Banach space $\mathcal{B}$ with norm $|\cdot|_{\mathcal{B}}$. Write $s_n = a_1 + \ldots + a_n$ and $s_0 = 0$. Let $p \geq 1$. For every $r \geq 0$, we have*

$$\max_{1\leq n\leq 2^r} |s_n|_{\mathcal{B}} \leq \sum_{k=0}^{r} \Big(\sum_{m=1}^{2^{r-k}} |s_{2^k m} - s_{2^k(m-1)}|_{\mathcal{B}}^p \Big)^{1/p} . \tag{55}$$

Proof. We make the proof by induction on $r \geq 0$. The result is obvious for $r = 0$. Let $1 \leq n \leq 2^r$. We have $|s_{2n-1}|_{\mathcal{B}} \leq |s_{2n-2}|_{\mathcal{B}} + |a_{2n-1}|_{\mathcal{B}}$. Hence, writing $\tilde{a}_n = a_{2n-1} + a_{2n}$ and $\tilde{s}_n = \sum_{k=1}^{n} \tilde{a}_k = s_{2n}$, we get that

$$\max_{1\leq l\leq 2^{r+1}} |s_l|_{\mathcal{B}} \leq \max_{1\leq n\leq 2^r} |\tilde{s}_n|_{\mathcal{B}} + \Big(\sum_{n=1}^{2^r} |a_{2n-1}|_{\mathcal{B}}^p \Big)^{1/p} ,$$

and the result follows. □

Theorem 4.7 *Let T be a Dunford-Schwartz operator on $(\Omega, \mathcal{A}, \mathbb{P})$. Let $f \in \mathbb{L}^p$, $p > 1$. Let ψ be a positive non-decreasing function, such that there exists $C > 1$ such that $\psi(2x) \leq C\psi(x)$, for every $x \geq 1$. Assume that*

$$\sum_n \frac{\|f + \ldots + T^{n-1} f\|_p}{\psi(n) n^{1+1/p}} < \infty . \tag{56}$$

Then $\sup_{n\geq 1} \frac{|f+\ldots+T^{n-1}f|}{\psi(n)n^{1/p}} \in \mathbb{L}^{p,w}$ and $\frac{|f+\ldots T^{n-1}f|}{\psi(n)n^{1/p}} \to 0$ $\mathbb{P}$-a.s.
If T is induced by a measure-preserving transformation and $(\mathcal{B}, |\cdot|_{\mathcal{B}})$ is a Banach space, the result holds with $|\cdot|_{\mathcal{B}}$ instead of $|\cdot|$ for every $f \in \mathbb{L}^p(\mathcal{B})$ such that $\sum_n \frac{\| |f+\ldots+T^{n-1}f|_{\mathcal{B}} \|_p}{\psi(n)n^{1+1/p}} < \infty$.

Comment 4.8 *Take $\psi \equiv 1$, which is the relevant case in our applications. Then, condition* (56) *is weaker than condition (8) in [34] and also (slightly) improves condition (10) of [4] (obtained for $p = 2$). In [34] and [4], only the case where T is induced by a transformation is considered.*

We turn now to the proof of Theorem 2.7. It will follow from the next two propositions. Notice that the second one is a version of Corollary 22 of Merlevède-Peligrad [25] under $\mathbb{E}_0$.

Proposition 4.9 *Assume* (1). *Then $\mathbb{E}_0[(S_n - M_n - \mathbb{E}_0(S_n))^2] = o(n)$ $\mathbb{P}$-a.s. and $\mathbb{E}_0(S_n) = o(\sqrt{n})$ $\mathbb{P}$-a.s. In particular*

$$\mathbb{E}_0[(S_n - M_n)^2] = o(n) \qquad \mathbb{P}\text{-a.s.}$$

Proposition 4.10 *Assume* (1) *and that $\mathbb{E}_0(S_n^2) = o(n)$ $\mathbb{P}$-a.s. Then*

$$\mathbb{E}_0(\max_{1\le k\le n} S_k^2) = o(n) \qquad \mathbb{P}\text{-a.s.} \tag{57}$$

Before proving the above propositions, we indicate how they lead to Theorem 2.7. Using Proposition 4.9, we apply Proposition 4.10 with $S_n - M_n$ in place of S_n. This proves (14). Now the convergence (15) follows from (14) together with the quenched weak invariance principle for martingales (see for instance Derriennic and Lin [11] for the ergodic case). To be more precise, if we define $D_k = D \circ \theta^k$ and $\widetilde{W}_n$ by $\widetilde{W}_n(t) = n^{-1/2}(M_{[nt]} + (nt - [nt])D_{[nt]+1})$, then (15) holds with $\widetilde{W}_n$ in place of W_n, and $\eta = \mathbb{E}(D^2|\mathcal{I})$. To end the proof, we first notice that by Theorem 1 of Peligrad and Utev (2005), $\mathbb{E}(D^2|\mathcal{I}) = \lim_{n\to\infty} n^{-1}\mathbb{E}(S_n^2|\mathcal{I})$ in $\mathbb{L}^1$. It remains to prove that $\mathbb{E}(D^2|\mathcal{I}) = \lim_{n\to\infty} n^{-1}\mathbb{E}_0(S_n^2)$ in $\mathbb{L}^1$. But, by (1) and (7), $\|S_n^2 - M_n^2\|_1 = o(n)$. Hence it suffices to prove that $\mathbb{E}(D^2|\mathcal{I}) = \lim_{n\to\infty} n^{-1}\mathbb{E}_0(M_n^2)$ in $\mathbb{L}^1$.

With this aim, we will make use of the operator Q defined by

$$QZ = \mathbb{E}_0(Z \circ \theta) \qquad \forall Z \in \mathbb{L}^1 .$$

The operator Q is markovian, hence it is a Dunford-Schwartz operator. Notice that $Q^n Z = \mathbb{E}_0(Z \circ \theta^n)$. Moreover, by Lemma 7.1 in [9], if Z is additionnally assumed to be in $\mathcal{F}_\infty$,

$$(QZ + \ldots + Q^n Z)/n \text{ converges } \mathbb{P}\text{-a.s. and in } \mathbb{L}^1 \text{ to } \mathbb{E}(Z|\mathcal{I}). \tag{58}$$

To conclude we take $Z = D^2$ and we notice that, by orthogonality, $\mathbb{E}_0(M_n^2) = Q(D^2) + \cdots + Q^n(D^2)$.

It remains to prove Propositions 4.9 and 4.10.

Proof of Proposition 4.9. The fact that $\mathbb{E}_0(S_n) = o(\sqrt{n})$ $\mathbb{P}$-a.s. under (1) comes directly from an application of Theorem 4.7 with $T = Q$. We prove now that under (1), the following convergence holds: $\mathbb{E}_0[(S_n - M_n - \mathbb{E}_0(S_n))^2] = o(n)$ $\mathbb{P}$-a.s.

Let N be a positive integer fixed for the moment. By (43), we have

$$S_n - M_n - \mathbb{E}_0(S_n) = -(V_{n,N} + W_{n,N}) , \tag{59}$$

where $V_{n,N}$ and $W_{n,N}$ are given respectively by (41) and (42).

Let $\varphi_N := \mathbb{E}_0(S_N)$ and $\psi_N = \sum_{l\ge N+1} \frac{\varphi_l}{l(l+1)}$, where ψ_N is well-defined in $\mathbb{L}^2$, by (1).

Then, by (44) and (46),

$$|V_{n,N}| \ll |\varphi_N \circ \theta^n| + |Q^n \varphi_N| + |\psi_N \circ \theta^n| + |Q^n \psi_N| .$$

Hence, by using (58),

$$\mathbb{E}_0(V_{n,N}^2) \ll Q^n(\varphi_N^2) + Q^n(\psi_N^2) = o(n) \qquad \mathbb{P}\text{-a.s.}$$

Then, using that $\mathbb{E}_0(S_n) = o(\sqrt{n})$ $\mathbb{P}$-a.s. and (59), we obtain

$$\limsup_n \frac{\mathbb{E}_0((S_n - M_n)^2)}{n} \leq \limsup_n \frac{\mathbb{E}_0(W_{n,N}^2)}{n}\,.$$

It remains to deal with $W_{n,N}$. Recall that by (48),

$$W_{n,N} = \sum_{r=1}^n \mathcal{P}_r(W_{n,N}) = \sum_{r=1}^n \Big(\sum_{k\geq N} \sum_{l\geq 1} \frac{\mathcal{P}_0(X_{l+k+n-r})}{l+k} \Big) \circ \theta^r\,.$$

Hence, by orthogonality,

$$\mathbb{E}_0(W_{n,N}^2) = \sum_{r=1}^n \mathbb{E}_0(\mathcal{P}_r(W_{n,N})^2) = \sum_{r=1}^n Q^r \Big(\sum_{k\geq N} \sum_{l\geq 1} \frac{\mathcal{P}_0(X_{l+k+n-r})}{l+k} \Big)^2\,.$$

But, using (50) and Cauchy-Schwarz's inequality, we have

$$\Big| \sum_{k\geq N} \sum_{l\geq 1} \frac{\mathcal{P}_0(X_{l+k+n-r})}{l+k} \Big| \ll \sum_{m\geq 1} \frac{1}{(m+N)^{3/2}} \Big(\sum_{k\geq 0} |\mathcal{P}_{-k}(S_m)|^2 \circ \theta^k \Big)^{1/2}\,.$$

Let now $g_N := \sum_{m\geq 1} \frac{1}{(m+N)^{3/2}} (\sum_{k\geq 0} |\mathcal{P}_{-k}(S_m)|^2 \circ \theta^k)^{1/2}$. Then g_N is in $\mathbb{L}^2$ and

$$\|g_N\|_2 \leq \sum_{m\geq 1} \frac{\|\mathbb{E}_0(S_m)\|_2}{(m+N)^{3/2}} < \infty\,.$$

In particular, $\|g_N\|_2 \to 0$, as $N \to \infty$. So, finally, by using (58), we get that

$$\frac{\mathbb{E}_0(W_{n,N}^2)}{n} \ll \frac{\sum_{r=1}^n Q^r(g_N^2)}{n} \underset{n\to+\infty}{\longrightarrow} \mathbb{E}(g_N^2|\mathcal{I}) \qquad \mathbb{P}\text{-a.s.}$$

Since $\|\mathbb{E}(g_N^2|\mathcal{I})\|_1 \leq \|g_N^2\|_1 \to 0$, there exists a sub-sequence (N_j) such that $\mathbb{E}(g_{N_j}^2|\mathcal{I}) \to 0$ $\mathbb{P}$-a.s. as $j \to \infty$ and the result follows. □

To prove Proposition 4.10, we will make use of the following maximal inequality from Merlevède and Peligrad (2012). They did not state the result exactly in that context but it may be proved exactly the same way, applying Doob's maximal inequality conditionally, so the proof is omitted.

Proposition 4.11 *Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary sequence in $\mathbb{L}^2$ in the sense of Notation 1.1 and adapted to the filtration $(\mathcal{F}_n)$. We have,*

$$\begin{aligned}\Big(\mathbb{E}_0\big(\max_{1\leq i\leq 2^r} |S_i|^2\big)\Big)^{1/2} &\leq 2\big(\mathbb{E}_0(S_{2^r}^2)\big)^{1/2} + 2\sum_{l=0}^{r-1} \Bigg(\sum_{k=1}^{2^{r-l}-1} \mathbb{E}_0\big(\big(\mathbb{E}_{k2^l}(S_{(k+1)2^l}) - S_{k2^l}\big)^2\big) \Bigg)^{1/2} \\ &= 2\big(\mathbb{E}_0(S_{2^r}^2)\big)^{1/2} + 2\sum_{l=0}^{r-1} \Bigg(\sum_{k=1}^{2^{r-l}-1} Q^{k2^l}\big((\mathbb{E}_0(S_{2^l}))^2\big) \Bigg)^{1/2} \qquad \mathbb{P}\text{-}a.s. \qquad (60)\end{aligned}$$

Proof of Proposition 4.10.

Let $v \geq 0$ be an integer, fixed for the moment. Let $r > v$. Then we have

$$\max_{1\leq k\leq 2^r} |S_k| \leq \max_{1\leq s\leq 2^{r-v}} |S_{s2^v}| + 2^v \max_{1\leq j\leq 2^r} |X_j|.$$

Let $K \geq 1$, be fixed for the moment. We have $\max_{1\leq j\leq 2^r} |X_j|^2 \leq K^2 + \sum_{j=1}^{2^r} |X_j|^2 \mathbf{1}_{\{|X_j|\geq K\}}$.

Hence, applying Proposition 4.11 to the stationary sequence $(S_{(k+1)2^v} - S_{k2^v})_{k\geq 0}$ adapted to the filtration $(\mathcal{F}_{k2^v})_{k\geq 0}$, we obtain (with the convention that $S_0 = 0$)

$$\begin{aligned}\mathbb{E}_0\big(\max_{1\leq i\leq 2^r} |S_i|^2\big) &\ll 4^v K^2 + 4^v \sum_{j=1}^{2^r} Q^j(|X_0|^2 \mathbf{1}_{\{|X_0|\geq K\}}) + \mathbb{E}_0(S_{2^r}^2) \\ &\quad + \Bigg(\sum_{l=0}^{r-v-1} \Bigg(\sum_{k=1}^{2^{r-v-l}-1} Q^{k2^{l+v}}\big((\mathbb{E}_0(S_{2^{l+v}}))^2\big)\Bigg)^{1/2}\Bigg)^2 \\ &\ll 4^v K^2 + 4^v \sum_{j=1}^{2^r} Q^j(|X_0|^2 \mathbf{1}_{\{|X_0|\geq K\}}) + \mathbb{E}_0(S_{2^r}^2) + 2^r \Bigg(\sum_{l=0}^{r-v-1} \frac{\big(\mathcal{M}_{l+v}\big((\mathbb{E}_0(S_{2^{l+v}}))^2\big)\big)^{1/2}}{2^{(l+v)/2}}\Bigg)^2 .\end{aligned}$$

By assumption $\mathbb{E}_0(S_{2^r}^2) = o(2^r)$ $\mathbb{P}$-a.s. By (58), $(\sum_{j=1}^{2^r} Q^j(|X_0|^2 \mathbf{1}_{\{|X_0|\geq K\}}))/2^r \to \mathbb{E}(|X_0|^2 \mathbf{1}_{\{|X_0|\geq K\}}|\mathcal{I})$ $\mathbb{P}$-a.s. Since $\|\mathbb{E}(|X_0|^2 \mathbf{1}_{\{|X_0|\geq K\}}|\mathcal{I})\|_1 \leq \|X_0^2 \mathbf{1}_{\{|X_0|\geq K\}}\|_1 \to \infty$, there exists a subsequence (K_j) such that $\mathbb{E}(|X_0|^2 \mathbf{1}_{\{|X_0|\geq K_j\}}|\mathcal{I}) \to 0$ $\mathbb{P}$-a.s. as $j \to \infty$. Hence taking the $\limsup_r$ and letting $j \to \infty$, we obtain

$$\limsup_r \frac{\mathbb{E}_0(\max_{1\leq i\leq 2^r} |S_i|^2)}{2^r} \ll \Bigg(\sum_{l\geq v} \frac{\big(\mathcal{M}_l\big((\mathbb{E}_0(S_{2^l}))^2\big)\big)^{1/2}}{2^{l/2}}\Bigg)^2 \qquad \mathbb{P}\text{-a.s.}$$

To finish the proof, it suffices to prove that the random variable defined by the series in the right-hand side is $\mathbb{P}$-a.s. finite. But it is in $\mathbb{L}^{2,w}$ since, by (1),

$$\Big\| \sum_{l\geq 0} \frac{\big(\mathcal{M}_l\big((\mathbb{E}_0(S_{2^l}))^2\big)\big)^{1/2}}{2^{l/2}} \Big\|_{2,w} \leq \sum_{l\geq 0} \frac{\|\big(\mathcal{M}_l\big((\mathbb{E}_0(S_{2^l}))^2\big)\big)^{1/2}\|_{2,w}}{2^{l/2}} \ll \sum_{l\geq 0} \frac{\|\mathbb{E}_0(S_{2^l})\|_2}{2^{l/2}} < \infty .$$

□

A Proof of Theorem 4.7

We make the proof for T Dunford-Schwartz and f real-valued since the proof in the case where f is $\mathcal{B}$-valued is identical, replacing $|\cdot|$ with $|\cdot|_{\mathcal{B}}$ when necessary.

Write $U_n(f) = f + \ldots + T^{n-1} f$. Since ψ is monotonic, it follows from the subadditivity of $(\|U_n(f)\|_p)$ (see for instance [28] Lemma 2.7 and [25] equation (92)) that (56) is equivalent to

$$\sum_n \frac{\|f + \ldots + T^{2^n-1} f\|_p}{\psi(2^n) 2^{n/p}} = \sum_n \frac{\|U_{2^n}(f)\|_p}{\psi(2^n) 2^{n/p}} < \infty .$$

We proceed now as in the proof of Proposition 4.10, namely: we consider dyadic blocs. Let us give the hints. Let $v \geq 0$ be an integer. For $r > v$, write that

$$\max_{1\leq k\leq 2^r} |U_k(f)| \leq \max_{1\leq s\leq 2^{r-v}} |U_{s2^v}(f)| + 2^v \max_{1\leq j\leq 2^r} |T^j f| .$$

Using Proposition 4.5 to take care of the first term in the right hand side, it follows that

$$\max_{1\leq k\leq 2^r} |U_k(f)| \leq 2^v \max_{1\leq j\leq 2^r} |T^j f| + 2^{r/p} \sum_{k\geq 0} \frac{\big[\mathcal{M}_{k+v}(|U_{2^{k+v}}(f)|^p)\big]^{1/p}}{2^{(k+v)/p}} .$$

We finish the proof by using arguments developped in the proof of Proposition 4.10. □

B Auxiliary results

Lemma B.1 *Let $\mathcal{B}$ be a Banach space and $(a_n)_{n\geq 1}$ a $\mathcal{B}$-valued sequence. The following are equivalent:*

(i) the series $\sum_{n\geq 1} a_n$ converges,

(ii) $\lim_{n\to\infty} n \sum_{k\geq n}(k+1)^{-1}a_k = 0$ and the series $\sum_{n\geq 1}\sum_{k\geq n}(k+1)^{-1}a_k$ converges.

The proof is omitted since it follows from standard arguments based on Abel summation by part.

The next lemma is Lemma 19 in Merlevède, Peligrad and Peligrad [26]. In their paper, the lemma is stated with $\ell = 0$ and with $\mathcal{H} = \mathbb{R}$ but with similar arguments as done in their proof, it works for any non-negative integer ℓ and for adapted stationary sequences with values in a normed space by replacing the absolute values by the corresponding norms.

Lemma B.2 *Let $p \geq 1$ and let $(X_n)_{n\in\mathbb{Z}}$ be an adapted stationary sequence in $\mathbb{L}^p(\mathcal{H})$ in the sense of Notation 1.1. For every $\gamma > 0$, $n \geq 1$ and any integer $\ell \geq 0$,*

$$\frac{1}{n^\gamma}\max_{1\leq k\leq n}\|\mathbb{E}_{-\ell}(S_k)\|_{p,\mathcal{H}} \leq 2^{3\gamma+3}\sum_{k=n+1}^{6n}\frac{1}{k^{\gamma+1}}\|\mathbb{E}_{-\ell}(S_k)\|_{p,\mathcal{H}}\,.$$

Proposition B.3 *Let $p \in [2,4]$ and let $(X_n)_{n\in\mathbb{Z}}$ be an adapted and stationary sequence in $\mathbb{L}^p$ in the sense of Notation 1.1. Assume that (6) holds. Then setting $M_n = \sum_{k=1}^n D\circ\theta^k$ where D is defined by (5), the following inequality holds: for any non-negative integers r and n,*

$$\begin{aligned}\|\mathbb{E}_{-r}(M_n^2) - \mathbb{E}(M_n^2)\|_{p/2} \ll{}& \|\mathbb{E}_{-r}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} + \|\mathbb{E}_{-r}(S_{2n}^2) - \mathbb{E}(S_{2n}^2)\|_{p/2}\\ &+ n\left(\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\right)^2 + n\sum_{k\geq n}\frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^{3/2}}\,.\end{aligned}$$

In the statement of the proposition as well as in its proof, the constants arising from the symbol $\ll$ are independent from n and r.

Proof. Setting $R_n = S_n - M_n$, we start with the following inequality:

$$\|\mathbb{E}_{-r}(M_n^2) - \mathbb{E}(M_n^2)\|_{p/2} \ll \|\mathbb{E}_{-r}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} + 2\|R_n\|_p^2 + 2\|\mathbb{E}_{-r}(S_nR_n) - \mathbb{E}(S_nR_n)\|_{p/2}\,. \quad (61)$$

Using Theorem 2.3 with $p \geq 2$, we first get that

$$\|R_n\|_p^2 \ll n\left(\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\right)^2. \quad (62)$$

Now, starting from (43) and using the decompositions (41), (42), (44) and (46) with $N = 2n$, we write that

$$R_n = \mathbb{E}_0(S_n) + \frac{\mathbb{E}_0(S_{2n}\circ\theta^n)}{2n+1} - \frac{\mathbb{E}_n(S_{2n}\circ\theta^n)}{2n+1} - A_n - B_n\,, \quad (63)$$

where

$$A_n = 2n\sum_{l\geq 2n+1}\frac{\mathbb{E}_n(S_l\circ\theta^n) - \mathbb{E}_0(S_l\circ\theta^n)}{l(l+1)}\,, \quad (64)$$

and

$$B_n = \sum_{k\geq 2n}\sum_{l\geq k+1}\frac{\mathbb{E}_n(X_{l+n}) - \mathbb{E}_0(X_{l+n})}{l}\,. \quad (65)$$

Notice first that

$$\Big\|\mathbb{E}_{-r}\Big(S_n\Big(\mathbb{E}_0(S_n)+\frac{\mathbb{E}_0(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)-\mathbb{E}\Big(S_n\Big(\mathbb{E}_0(S_n)+\frac{\mathbb{E}_0(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)\Big\|_{p/2}$$
$$\leq 2\Big\|\mathbb{E}_0\Big(S_n\Big(\mathbb{E}_0(S_n)+\frac{\mathbb{E}_0(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)\Big\|_{p/2}\leq 2\|\mathbb{E}_0(S_n)\|_p^2+2(2n+1)^{-1}\|\mathbb{E}_0(S_n)\|_p\|\mathbb{E}_0(S_{2n})\|_p\,,$$

which combined with (53) with $u_n=[n^{p/2}]$ implies that

$$\Big\|\mathbb{E}_{-r}\Big(S_n\Big(\mathbb{E}_0(S_n)+\frac{\mathbb{E}_0(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)-\mathbb{E}\Big(S_n\Big(\mathbb{E}_0(S_n)+\frac{\mathbb{E}_0(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)\Big\|_{p/2}$$
$$\ll n\bigg(\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\bigg)^2. \tag{66}$$

Now writing that $S_{2n}\circ\theta^n=S_{2n}\circ\theta^n-S_n\circ\theta^n+S_n\circ\theta^n$ and using the fact that S_n is $\mathcal{F}_n$-measurable, we get

$$\Big\|\mathbb{E}_{-r}\Big(S_n\Big(\frac{\mathbb{E}_n(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)-\mathbb{E}\Big(S_n\Big(\frac{\mathbb{E}_n(S_{2n}\circ\theta^n)}{2n+1}\Big)\Big)\Big\|_{p/2}$$
$$\leq n^{-1}\|\mathbb{E}_{-r}(S_n(S_{2n}-S_n))-\mathbb{E}(S_n(S_{2n}-S_n))\|_{p/2}+n^{-1}\|\mathbb{E}_{-r}(S_n\mathbb{E}_n(S_{3n}-S_{2n}))\|_{p/2}\,. \tag{67}$$

Using the identity $2ab=(a+b)^2-a^2-b^2$ and the stationarity, we first obtain that

$$2\|\mathbb{E}_{-r}(S_n(S_{2n}-S_n))-\mathbb{E}(S_n(S_{2n}-S_n))\|_{p/2}\leq 2\|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}$$
$$+\|\mathbb{E}_{-r}(S_{2n}^2)-\mathbb{E}(S_{2n}^2)\|_{p/2}\,. \tag{68}$$

To bound up the second term in (67), we write $C_n:=n^{-1}\mathbb{E}_n(S_{3n}-S_{2n})$ and we follow the lines of the proof of Theorem 2.3 in [8] (see the displaylines between their equations (4.13) and (4.16)). Hence we first write that

$$\begin{aligned}\|\mathbb{E}_{-r}(S_nC_n)\|_{p/2}&\leq\|\mathbb{E}_{-r}^{1/2}(S_n^2)\mathbb{E}_{-r}^{1/2}(C_n^2)\|_{p/2}\\&\leq\|(\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2))^{1/2}\mathbb{E}_{-r}^{1/2}(C_n^2)\|_{p/2}+(\mathbb{E}(S_n^2))^{1/2}\|\mathbb{E}_{-r}^{1/2}(C_n^2)\|_{p/2}\\&\leq\|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}+\|C_n\|_p^2+(\mathbb{E}(S_n^2))^{1/2}\|\mathbb{E}_{-r}^{1/2}(C_n^2)\|_{p/2}\,.\end{aligned}$$

Notice that since (6) holds, by Theorem 2.3, we have in particular that $\|S_n\|_2=o(n^{1/2})+\|M_n\|_2$, implying that

$$\|S_n\|_2\ll n^{1/2}\,. \tag{69}$$

Using (69) and the fact that the function $x\mapsto|x|^{p/4}$ is concave, it follows that

$$\|\mathbb{E}_{-r}(S_nC_n)\|_{p/2}\ll\|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}+\|C_n\|_p^2+n^{1/2}\|C_n\|_2\,. \tag{70}$$

By stationarity and using (53) with $u_n=[n^{p/2}]$, we get that

$$\|C_n\|_p\ll n^{-1}\|\mathbb{E}_{-n}(S_n)\|_{p,\mathcal{H}}\ll n^{-1/2}\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\,. \tag{71}$$

On another hand, by using once again stationarity and Lemma B.2,

$$\|C_n\|_2\ll n^{-1}\|\mathbb{E}_{-n}(S_n)\|_2\ll\sum_{k\geq n}\frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^2}\,. \tag{72}$$

Therefore starting from (67) and using (68), (70), (71) and (72), we infer that

$$\Big\|\mathbb{E}_{-r}\Big(S_n\Big(\frac{\mathbb{E}_n(S_{2n}\circ\theta_n)}{2n+1}\Big)\Big)-\mathbb{E}\Big(S_n\Big(\frac{\mathbb{E}_n(S_{2n}\circ\theta_n)}{2n+1}\Big)\Big)\Big\|_{p/2}\ll\|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}$$
$$+n^{-1}\|\mathbb{E}_{-r}(S_{2n}^2)-\mathbb{E}(S_{2n}^2)\|_{p/2}+n^{-1}\bigg(\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\bigg)^2+n^{1/2}\sum_{k\geq n}\frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^2}\,. \tag{73}$$

We consider now the term $\|\mathbb{E}_{-r}(S_nA_n)-\mathbb{E}(S_nA_n)\|_{p/2}$. With this aim, we first define

$$\widetilde{A}_n=2n\mathbb{E}_n(S_n\circ\theta^n)\sum_{l\geq 2n+1}\frac{1}{l(l+1)}\,.$$

Since S_n is $\mathcal{F}_n$-measurable,

$$\|\mathbb{E}_{-r}(S_n\widetilde{A}_n)-\mathbb{E}(S_n\widetilde{A}_n)\|_{p/2}\leq\|\mathbb{E}_{-r}(S_n(S_{2n}-S_n))-\mathbb{E}(S_n(S_{2n}-S_n))\|_{p/2}\,.$$

Using then the identity $2ab=(a+b)^2-a^2-b^2$ and stationarity, it follows that

$$2\|\mathbb{E}_{-r}(S_n\widetilde{A}_n)-\mathbb{E}(S_n\widetilde{A}_n)\|_{p/2}\leq 2\|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}+\|\mathbb{E}_{-r}(S_{2n}^2)-\mathbb{E}(S_{2n}^2)\|_{p/2}\,. \tag{74}$$

Let now

$$D_n:=n\sum_{k\geq 2n+1}\frac{\mathbb{E}_n(S_k\circ\theta^n)-\mathbb{E}_n(S_n\circ\theta^n)}{k(k+1)}$$

and notice that, by stationarity,

$$\|\mathbb{E}_{-r}(S_n(A_n-\widetilde{A}_n))-\mathbb{E}(S_n(A_n-\widetilde{A}_n))\|_{p/2}$$
$$\ll n\|\mathbb{E}_0(S_n)\|_p\sum_{k\geq n+1}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^2}+\|\mathbb{E}_{-r}(S_nD_n)\|_{p/2}\,. \tag{75}$$

Using (53) with $u_n=n$, we first get that

$$n\|\mathbb{E}_0(S_n)\|_p\sum_{\ell\geq n+1}\frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^2}\ll n^2\bigg(\sum_{k\geq n}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^2}\bigg)^2.$$

But, by using Lemma B.2 and the fact that $p\geq 2$,

$$n\sum_{k\geq n}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^2}\leq\max_{1\leq k\leq[n^{p/2}]}\|\mathbb{E}_0(S_k)\|_p+n\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^2}$$
$$\ll n^{p/2}\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^2}\ll n^{1/2}\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\,. \tag{76}$$

Therefore

$$n\|\mathbb{E}_0(S_n)\|_p\sum_{\ell\geq n+1}\frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^2}\ll n\bigg(\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\bigg)^2. \tag{77}$$

We bound now the second term in the right-hand side of (75). Proceeding as to get (70), we infer that

$$\|\mathbb{E}_{-r}(S_nD_n)\|_{p/2}\ll\|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}+\|D_n\|_p^2+n^{1/2}\|D_n\|_2\,. \tag{78}$$

Stationarity and inequality (76) imply that

$$\|D_n\|_p \ll n \sum_{k\geq n} \frac{\|\mathbb{E}_{-n}(S_k)\|_p}{k^2} \ll n^{1/2} \sum_{k\geq [n^{p/2}]} \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}} \,. \tag{79}$$

On another hand, using once again stationarity,

$$\|D_n\|_2 \ll n \sum_{k\geq n} \frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^2} \,. \tag{80}$$

Overall, starting from (75) and considering the bounds (77), (78), (79) and (80), it follows that

$$\begin{aligned}\|\mathbb{E}_{-r}(S_n(A_n-\widetilde{A}_n)) - \mathbb{E}(S_n(A_n-\widetilde{A}_n))\|_{p/2} &\ll \|\mathbb{E}_{-r}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} \\ &+ n\left(\sum_{k\geq [n^{p/2}]} \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\right)^2 + n^{3/2}\sum_{k\geq n} \frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^2} \,. \end{aligned} \tag{81}$$

We consider now the term $\|\mathbb{E}_{-r}(S_nB_n) - \mathbb{E}(S_nB_n)\|_{p/2}$. Proceeding as to get (70), we infer that

$$\|\mathbb{E}_{-r}(S_nB_n) - \mathbb{E}(S_nB_n)\|_{p/2} \ll \|\mathbb{E}_{-r}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} + \|B_n\|_p^2 + n^{1/2}\|B_n\|_2 \,. \tag{82}$$

According to the bound (51) with $u_n = 2n$, followed by an application of Lemma 4.3,

$$\|B_n\|_2 \ll n^{1/2} \sum_{k\geq n}\sum_{m\geq 1} \frac{\|\mathcal{P}_{-k}(S_m)\|_2}{(m+k)^2} \ll n^{1/2} \sum_{k\geq n} \frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^{3/2}} \,. \tag{83}$$

To bound $\|B_n\|_p$, we use (63). By stationarity, we then infer that

$$\|B_n\|_p \leq \|R_n\|_p + 3\|\mathbb{E}_0(S_n)\|_p + 2n \sum_{\ell\geq n+1} \frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^2} \,.$$

Hence using Theorem 2.3 and inequality (53) with $u_n = n$, we get that

$$\|B_n\|_p \ll n^{1/2} \sum_{\ell\geq [n^{p/2}]} \frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^{1+1/p}} + n\sum_{\ell\geq n} \frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^2} \,,$$

which together with (76) implies that

$$\|B_n\|_p \ll n^{1/2} \sum_{\ell\geq [n^{p/2}]} \frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^{1+1/p}} \,. \tag{84}$$

Starting from (82) and using (83) and (84), we then obtain that

$$\begin{aligned}\|\mathbb{E}_{-r}(S_nB_n) - \mathbb{E}(S_nB_n)\|_{p/2} \ll &\|\mathbb{E}_{-r}(S_n^2) - \mathbb{E}(S_n^2)\|_{p/2} + n\left(\sum_{k\geq [n^{p/2}]} \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\right)^2 \\ &+ n\sum_{k\geq n} \frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^{3/2}} \,. \end{aligned} \tag{85}$$

Taking into account the decomposition (63) together with the bounds (66), (73), (74), (81) and (85), we then derive that

$$\|\mathbb{E}_{-r}(S_nR_n)-\mathbb{E}(S_nR_n)\|_{p/2} \ll \|\mathbb{E}_{-r}(S_n^2)-\mathbb{E}(S_n^2)\|_{p/2}+\|\mathbb{E}_{-r}(S_{2n}^2)-\mathbb{E}(S_{2n}^2)\|_{p/2}$$
$$+\,n\left(\sum_{k\geq[n^{p/2}]}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\right)^2+n\sum_{k\geq n}\frac{\|\mathbb{E}_{-n}(S_k)\|_2}{k^{3/2}}\,. \tag{86}$$

Starting from (61) and considering the inequalities (62) and (86), the proposition follows. □

Acknowledgments. C. Cuny would like to thank Dalibor Volný for helpful discussions while he was visiting the university of Rouen. The authors are also indebted to the referee for carefully reading the manuscript.

References

[1] A.N. Borodin and I. A. Ibragimov, *Limit theorems forfunctionals of random walks.* Trudy Mat. Inst. Steklov. 195 (1994). Transl. into English:Proc. Steklov Inst. Math. (1995).

[2] R. C. Bradley, Introduction to strong mixing conditions. *Vol. 1. Kendrick Press, Heber City, UT* (2007).

[3] D. L. Burkholder, *Sharp inequalities for martingales and stochastic integrals. Colloque Paul Lvy sur les Processus Stochastiques (Palaiseau, 1987).* Astérisque No. 157-158 (1988), 75-94.

[4] C. Cuny, *Pointwise ergodic theorems with rate and application to limit theorems for Markov chains*, Stoch. Dyn. 11 (2011), 135-155.

[5] C. Cuny and M. Lin, *Pointwise ergodic theorems with rate and application to the CLT for Markov chains*, Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009), no. 3, 710-733.

[6] C. Cuny and M. Peligrad, *Central limit theorem started at a point for stationary processes and additive functional of reversible Markov chains*, J. Theoret. Probab. 25 (2012), 171-188.

[7] C. Cuny and D. Volný, *A quenched invariance principle for stationary processes*, arXiv:1202.4875, accepted for publication in ALEA Lat. Am. J. Probab. Math. Stat.

[8] J. Dedecker, P. Doukhan and F. Merlevède, *Rates of convergence in the strong invariance principle under projective criteria*, Electron. J. Probab. 17 (2012), 1-31

[9] J. Dedecker, F. Merlevède and M. Peligrad, *A quenched weak invariance principle*, arXiv:1204.4554, accepted for publication in Ann. Inst. Henri Poincaré Probab. Stat.

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

[11] Y. Derriennic and M. Lin, *The central limit theorem for Markov chains with normal transition operators, started at a point*, Probab. Theory Relat. Fields. 119, (2001) 508-528.

[12] Y. Derriennic and M. Lin, *Fractional Poisson equations and ergodic theorems for fractional coboundaries*, Israel J. Math. 123 (2001), 93-130.

[13] Y. Derriennic and M. Lin, *The central limit theorem for Markov chains started at a point*, Probab. Theory Related Fields. 125 (2003), 73-76.

[14] O. Durieu, *Independence of four projective criteria for the weak invariance principle*, ALEA Lat. Am. J. Probab. Math. Stat. 5 (2009), 21-26.

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

[16] M. I. Gordin and B. A. Lifsic, *Central limit theorem for stationary Markov processes*, Dokl. Akad. Nauk SSSR 239 (1978), 766-767.

[17] M. Gordin and M. Peligrad, *On the functional CLT via martingale approximation*, Bernoulli 17 (2011), 424-440.

[18] I. G. Grama, *On moderate deviations for martingales*, Ann. Probab. 25 (1997), 152-183.

[19] I. G. Grama and E. Haeusler, *An asymptotic expansion for probabilities of moderate deviations for multivariate martingales*, J. Theoret. Probab. 19 (2006), 1-44.

[20] C. C. Heyde, *On the central limit theorem and iterated logarithm law for stationary processes*, Bull. Austral. Math. Soc. 12 (1975), 1-8.

[21] I. A. Ibragimov, *A remark on the central limit theorem for dependent random variables*, Teor. Verojatnost. i Primenen. 20 (1975), 134-140.

[22] U. Krengel, Ergodic theorems, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, (1985).

[23] M. Maxwell and M. Woodroofe. *C*entral limit theorems for additive functionals of Markov chains, Ann. Probab. 28 (2000), 713–724.

[24] F. Merlevède and M. Peligrad, *On the weak invariance principle for stationary sequences under projective criteria.* J. Theoret. Probab. 19 (2006), 647-689.

[25] F. Merlevède and M. Peligrad, *Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples*, Ann. Probab 41 (2013), 914-960.

[26] F. Merlevède, C. Peligrad and M. Peligrad, *Almost Sure Invariance Principles via Martingale Approximation*, Stochastic Process. Appl. 122 (2012), 170-190.

[27] M. Peligrad, *Convergence rates of the strong law for stationary mixing sequences*, Z. Wahrsch. Verw. Gebiete 70 (1985), 307-314.

[28] M. Peligrad and S. Utev, *A new maximal inequality and invariance principle for stationary sequences*, Ann. Probab. 33 (2005), no. 2, 798-815.

[29] M. Peligrad, S. Utev and W. B. Wu, *A maximal Lp-inequality for stationary sequences and its applications*, Proc. Amer. Math. Soc. 135 (2007), 541-550.

[30] Q. M. Shao, *Maximal inequalities for partial sums of -mixing sequences*, Ann. Probab. 23 (1995) 948-965.

[31] D. Volný, *Martingale approximation and optimality of some conditions for the central limit theorem*, J. Theo. Prob. 23 (2010) 888-903.

[32] D. Volný and M. Woodroofe, *An example of non-quenched convergence in the conditional central limit theorem for partial sums of a linear process*, Dependence in analysis, probability and number theory (The Philipp memorial volume), Kendrick Press. (2010) 317-323.

[33] W.A. Woyczyński, Asymptotic behavior of martingales in Banach spaces. II. *Martingale theory in harmonic analysis and Banach spaces (Cleveland, Ohio, 1981)*, pp. 216-225, Lecture Notes in Math., 939, Springer, Berlin-New York, (1982).

[34] W. B. Wu, *Strong invariance principles for dependent random variables*, Ann. Probab. 35 (2007), 2294-2320.

[35] W. B. Wu and M. Woodroofe, *Martingale approximations for sums of stationary processes*, Ann. Probab. 32 (2004), 1674-1690.

[36] W. B. Wu and Z. Zhao, *Moderate deviations for stationary processes*, Statis. Sinica 18 (2008), 1831-1847.

[37] O. Zhao and M. Woodroofe, *Law of the iterated logarithm for stationary processes*, Ann. Probab. 36 (2008), 127-142.

[38] O. Zhao and M. Woodroofe, *On martingale approximations*, Ann. Appl. Prob. 18, (2008), 1831-1847.