Deviation inequalities for dependent sequences with applications to strong approximations

J Dedecker, F Merlevède, Emmanuel Rio

Deviation inequalities for dependent sequences with applications to strong approximations

J. Dedecker, F. Merlevède and E. Rio

July 5, 2023

Abstract

In this paper, we give precise rates of convergence in the strong invariance principle for stationary sequences of bounded real-valued random variables satisfying weak dependence conditions. One of the main ingredients is a new Fuk-Nagaev type inequality for a class of weakly dependent sequences. We describe also several classes of processes to which our results apply.

MSC2020 subject classifications: 60F17; 60E15; 37E05.
Keywords: Invariance principles, rates of convergence, deviation inequality, dependent sequences.
Running head: Deviation inequalities and strong approximations.

1 Introduction

Let $(X_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of real-valued random variables (r.v.) defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with mean zero and positive finite variance. Set $S_n = X_1 + X_2 + \cdots + X_n$. In this paper, we assume furthermore that the series $\sigma^2 = \sum_{k\in\mathbb{Z}} \mathrm{Cov}(X_0, X_k)$ is convergent (under this assumption $\lim_n n^{-1}\mathrm{Var}(S_n) = \sigma^2$). We are interested in obtaining sharp rates in the invariance principle (both in the almost sure sense and in the $\mathbb{L}^2$-sense). Recall that such invariance principles consist in constructing, on a possible larger probability space, a sequence $(Z_i)_{i\geq 1}$ of i.i.d. centered Gaussian variables with variance σ^2 in such a way that, setting $T_k = \sum_{i=1}^k Z_i$,

$$\sup_{1\leq k\leq n} \big|S_k - T_k\big| = O(a_n) \text{ a.s. or in } \mathbb{L}^2, \tag{1.1}$$

where $(a_n)_{n\geq 1}$ is a nondecreasing sequence of positive reals tending to infinity, satisfying $a_n = o(\sqrt{n})$. In the independent setting, the almost sure rates are $a_n = o(n^{1/p})$ when $\mathbb{E}(|X_0|^p) < \infty$ for $p > 2$ and are $a_n = O(\log n)$ when X_0 has a finite Laplace transform

in a neighborhood of 0 (see [18, 21]). Furthermore these rates are optimal according to Breiman [4] in the first case and to Bártfai [1] in the second one.

In the dependent setting, even when the random variables are bounded the almost sure rates can be arbitrarily large. More precisely, let us consider the class of irreducible aperiodic and positively recurrent Markov chains (ξ_n) with an atom denoted by A (see the definition page 286 in [3]). Let τ_A be the first return time in A, $\mathbb{P}_A$ be the probability of the chain starting from the atom and $\mathbb{E}_A$ be the expectation under $\mathbb{P}_A$. Let then π be the unique invariant distribution, (ξ_n) be the Markov chain starting from π, and (X_k) be the strictly stationary sequence defined by $X_k = f(\xi_k)$ with f a bounded function. Theorem 2.2 in [13] asserts that, for any $p > 2$ there exists an irreducible aperiodic and positively recurrent Markov chain $(\xi_k)_{k\geq 0}$ with uniform distribution over $[0, 1]$ satisfying $\mathbb{P}_A(\tau_A > x) = O(x^{-p})$ and such that for any absolutely continuous function f on $[0, 1]$ with $\pi(f) = 0$ and a strictly positive derivative,

$$\limsup_{n\to\infty} (n \log n)^{-1/p} \big| S_n - \sum_{k=1}^{n} g_k \big| > 0 \text{ a.s.}$$

for any stationary and Gaussian centered sequence $(g_k)_{k\in\mathbb{Z}}$ with convergent series of covariances. This shows that for this type of Markov chains, the rates in the almost sure invariance principle are linked to the moments of the return times in A.

For $p \in]2, 4]$, under the slightly stronger condition $\mathbb{E}_A(\tau_A^p) < \infty$, Csáki and Csörgő [5] proved the almost sure rates $a_n = O(n^{1/p} \log n)$ in the strong invariance principle (see their Theorem 2.1). Our main objective is to extend this result to the case of general stationary sequences of bounded random variables including the case of bounded variation functions of non irreducible Markov chains. We shall then consider the case of θ-dependent sequences whose coefficients are defined as follows:

Definition 1.1. Let $\Gamma_{p,q} = \{(a_i)_{1\leq i\leq p} \in \mathbb{N}^p : a_1 \geq 1 \text{ and } \sum_{i=1}^p a_i \leq q\}$, for p and q positive integers. Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary sequence of centered and bounded real-valued random variables and $\mathcal{F}_0 = \sigma(X_i, i \leq 0)$. For $k \geq 0$, set

$$\theta_{X,p,q}(k) = \sup_{\substack{k_p > k_{p-1} > \ldots > k_2 > k_1 \geq k \\ (a_1,\ldots,a_p)\in\Gamma_{p,q}}} \Big\| \mathbb{E}\Big(\prod_{i=1}^{p} X_{k_i}^{a_i} | \mathcal{F}_0 \Big) - \mathbb{E}\Big(\prod_{i=1}^{p} X_{k_i}^{a_i} \Big) \Big\|_1 .$$

These coefficients are suitable for non irreducible Markov chains (see the examples given in Section 4.2). In addition, in case of bounded additive functionals of irreducible aperiodic and positively recurrent Markov chains with an atom A, for any $p \geq 2$, the condition $\mathbb{E}_A(\tau_A^p) < \infty$ implies that $\sum_{k\geq 1} k^{p-2}\theta_{X,4,4}(k) < \infty$ (see Section 4.1 for more details). Our aim is then to show that for any stationary sequences of bounded random variables satisfying the later weak dependence condition for some p in $]2, 4]$, the rate in the almost sure invariance principle is $n^{1/p}$ up to some power of $\log n$.

To obtain such rates, a possible approach is to use a martingale approximation and the Skorokhod embedding theorem. Recall that with this method, the rate cannot be better than $n^{1/4}$ up to some power of $\log n$. The closest results in this direction are given in Wu [28] and Doukhan et al [9]. For instance, for p in $]2,4]$, Corollary 3.9 in [9] provides the rate $a_n = o(n^{1/p}(\log n)^{1/2+1/p+\varepsilon})$ for any $\varepsilon > 0$ in the almost sure invariance principle under the condition $\sum_{k>0} k^{p-1-2/p}\theta_{X,4,4}(k) < \infty$, which is suboptimal. Still by means of the Skorokhod embedding theorem, it is possible to get rates for the $\mathbb{L}^2$-norm of the error in the invariance principle with suboptimal conditions (see Liu and Wang [19]).

Let $\theta(k) = \theta_{X,4,4}(k)$, for any $k \geq 0$ (recall that the random variables are bounded in Definition 1.1). If p is in $]2,3]$ and $\sum_{k\geq 1} k^{p-2}\theta(k) < \infty$, one can obtain the rates $a_n = O(n^{1/p})$ (up to some power of $\log n$) both for the almost sure invariance principle and the $\mathbb{L}^2$-norm of the error, by adapting the proof of [23, Theorem 2.1]. The main ingredients used in [23] are a Fuk-Nagaev type inequality to control the fluctuations and an estimate of the quadratic cost in the conditional central limit theorem. However, compared to inequality (2.1) below, the second term in their Fuk-Nagaev type inequality cannot be better than Cn/x^3. In addition, their estimate of the quadratic cost in the conditional central limit theorem cannot be better than $n^{-1/2}$ (this follows from their inequality (A.6)). These upper bounds induce a limitation of the rates in the invariance principle at the level $n^{1/3}$ (up to some power of $\log n$). Recently, the authors proved that for p in $]3,4]$ and under the condition $\sum_{k\geq 1} k^{p-2}\theta(k) < \infty$, the estimate of the quadratic cost in the conditional central limit theorem is of order $n^{-(p-2)/2}$. This together with the Fuk-Nagaev type inequality stated in Section 2 are the main ingredients to go beyond the rate $n^{1/3}$ in the almost sure invariance principle and to get the rates $n^{1/p}$ (up to some power of $\log n$) for $p \in]3,4]$.

Our paper is organized as follows. In Section 2, we state a Fuk-Nagaev type inequality for partial sums associated with stationary sequences of bounded random variables satisfying the condition $\sum_{k\geq 1} k\theta(k) < \infty$. In this inequality, the second term is of order n/x^4 under the condition $\sum_{k\geq 1} k^2\theta(k) < \infty$. In Section 3, we give our main results concerning the rates in the invariance principles (in $\mathbb{L}^2$ and almost surely). In Section 4, we present several classes of examples to which our results apply, including the example of BV observables of the Liverani-Saussol-Vaienti map (see [20]). An application to rates of convergence in the functional central limit theorem for the quadratic cost associated with the uniform deviation between the Donsker line and the Brownian motion is provided in Section 4.4. The proof of the Fuk-Nagaev type inequality is given in Section 5. Section 6 is devoted to the proof of our results concerning the rates in the invariance principles in the non degenerate case, whereas the degenerate case is considered in Section 7.

In this paper, we shall use the following notations: $a \ll b$ means that there exists a

numerical positive constant C such that $a \leq Cb$, and $X^{(0)}$ means $X - \mathbb{E}(X)$.

2 Deviation inequalities

We set $S_0 = 0$ and $S_n = X_1 + X_2 + \cdots + X_n$ for any positive integer n. In all the paper, except in Section 7, we denote $\theta_{X,4,4}(k)$ by $\theta(k)$ for all k in $\mathbb{N}$ and we assume that $\|X_0\|_\infty = 1$. The general case follows by dividing the random variables by $\|X_0\|_\infty$.

Theorem 2.1. *Assume that $\sum_{k\geq 1} k\theta(k) < \infty$. Let $S_n^* = \max_{0\leq k\leq n} S_k$. Then the series $\mathbb{E}(X_0^2) + 2\sum_{k\geq 1}\mathbb{E}(X_0X_k)$ converges to some nonnegative real σ^2. Moreover, for any positive real x and any positive integer n,*

$$\mathbb{P}\big(S_n^* \geq x\big) \leq c_1 \mathbf{1}_{\{\sigma^2>0\}} \Big(\frac{n\sigma^2}{x^2}\Big)^4 \exp\Big(-\frac{x^2}{16n\sigma^2}\Big) + c_2 \frac{n}{x^4}\Big(\Theta_1\Theta_2 + \sum_{k\geq 1} k(k\wedge x)\theta(k)\Big)\,, \quad (2.1)$$

where $\Theta_1 = 1 + \sum_{k\geq 1}\theta(k)$, $\Theta_2 = 1 + \sum_{k\geq 1} k\theta(k)$ and c_1, c_2 are positive numerical constants.

Remark 2.1. Recall that $\lim_{n\to\infty} n^{-1}\mathbb{E}(S_n^2) = \sigma^2$ as soon as $\sum_{k\geq 1}\theta(k) < \infty$.

Let $p \in]3,4[$ and assume that $\sum_{k\geq 1} k^{p-2}\theta(k) < \infty$. An application of Theorem 2.1 gives that for any $\alpha \in]1/2, 1]$ and any $\varepsilon > 0$, $\sum_{n\geq 1} n^{\alpha p-2}\mathbb{P}\big(S_n^* \geq n^\alpha\varepsilon\big) < \infty$. Note that when $p \in]2,3[$ the same result holds by using Proposition A.2 in [23].

3 Application to strong approximations

Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary sequence of centered and bounded real-valued random variables such that $\sum_{k>0}\theta(k) < \infty$. In this situation, the series $\mathbb{E}(X_0^2) + 2\sum_{k\geq 1}\mathbb{E}(X_0X_k)$ is absolutely convergent. In this section, we are interested in strong approximations in the non degenerate case, meaning that the sum σ^2 of this series is positive. We will consider the case $\sigma^2 = 0$ in Section 7. In the sequel, we assume that the underlying probability space is rich enough to contain a random variable δ uniformly distributed over $[0,1]$, independent of the sequence $(X_k)_{k\in\mathbb{Z}}$. From δ, we construct a sequence $(\delta_i)_{i\in\mathbb{Z}}$ of i.i.d. random variables uniformly distributed over $[0,1]$ and independent of $(X_k)_{k\in\mathbb{Z}}$.

As explained in [23], having a suitable bound for the quadratic transportation cost in the conditional central limit theorem allows to derive strong approximation results. Indeed, let us recall the construction given in [23] which is inspired from Bártfai [1]. For $L \in \mathbb{N}$, let $m(L) \in \mathbb{N}$ be such that $m(L) \leq L$. Let

$$I_{k,L} =]2^L + (k-1)2^{m(L)}, 2^L + k2^{m(L)}] \cap \mathbb{N} \text{ and } U_{k,L} = \sum_{i\in I_{k,L}} X_i\,,\ k \in \{1,\cdots,2^{L-m(L)}\}\,.$$

For $k \in \{1, \cdots, 2^{L-m(L)}\}$, let $V_{k,L}$ be the $N(0, \sigma^2 2^{m(L)})$-distributed random variable defined from $U_{k,L}$ via the conditional quantile transformation, that is

$$V_{k,L} = \sigma 2^{m(L)/2} \Phi^{-1}(\widetilde{F}_{k,L}(U_{k,L} - 0) + \delta_{2^L + k2^{m(L)}}(\widetilde{F}_{k,L}(U_{k,L}) - \widetilde{F}_{k,L}(U_{k,L} - 0))) \,, \quad (3.1)$$

where $\widetilde{F}_{k,L} := F_{U_{k,L}|\mathcal{F}_{2^L+(k-1)2^{m(L)}}}$ is the d.f. of $P_{U_{k,L}|\mathcal{F}_{2^L+(k-1)2^{m(L)}}}$ (the conditional law of $U_{k,L}$ given $\mathcal{F}_{2^L+(k-1)2^{m(L)}}$) and Φ^{-1} is the inverse of the standard Gaussian distribution function Φ. Since $\delta_{2^L+k2^{m(L)}}$ is independent of $\mathcal{F}_{2^L+(k-1)2^{m(L)}}$, the random variable $V_{k,L}$ is independent of $\mathcal{F}_{2^L+(k-1)2^{m(L)}}$, and has the Gaussian distribution $N(0, \sigma^2 2^{m(L)})$ (see [26, Lemma F.1]. By induction on k, the random variables $(V_{k,L})_k$ are mutually independent and independent of $\mathcal{F}_{2^L}$. In addition

$$\begin{aligned} \mathbb{E}(U_{k,L} - V_{k,L})^2 &= \mathbb{E} \int_0^1 \left(F^{-1}_{U_{k,L}|\mathcal{F}_{2^L+(k-1)2^{m(L)}}}(u) - \sigma 2^{m(L)/2} \Phi^{-1}(u)\right)^2 du \\ &:= \mathbb{E}\big(W_2^2(P_{U_{k,L}|\mathcal{F}_{2^L+(k-1)2^{m(L)}}}, G_{\sigma^2 2^{m(L)}})\big) \,, \end{aligned} \quad (3.2)$$

where $G_{\sigma^2 2^{m(L)}}$ is the Gaussian distribution $N(0, \sigma^2 2^{m(L)})$.

This construction together with Theorem 2.1 and estimates of the quadratic conditional cost in the central limit theorem given in [14] are the main ingredients to get the next strong approximations results.

Theorem 3.1.

(i) Assume that $\theta(k) = O(k^{1-p})$ if $p \in]2,3[\cup]3,4[$ and $\sum_{k>0} k^{p-2}\theta(k) < \infty$ if $p \in \{3,4\}$. Then, one can construct a sequence of i.i.d. Gaussian random variables $(Z_i)_{i\geq 1}$ centered and with variance σ^2, such that, setting $T_k = \sum_{i=1}^k Z_i$, we have

$$\left\| \sup_{k\leq n} |S_k - T_k| \right\|_2 = O(n^{1/p}(\log n)^{(1/2-1/p)}) \,.$$

(ii) Let $p \in \{3,4\}$ and assume that $\theta(k) = O(k^{1-p})$. Then, one can construct a sequence of i.i.d. Gaussian random variables $(Z_i)_{i\geq 1}$ centered and with variance σ^2, such that, setting $T_k = \sum_{i=1}^k Z_i$,

$$\left\| \sup_{k\leq n} |S_k - T_k| \right\|_2 = O(n^{1/p}(\log n)^{1/2}) \,.$$

We now give almost sure rates for strong approximations.

Theorem 3.2.

(i) Let $p \in]2,3[\cup]3,4[$ and assume that $\sum_{k>0} k^{p-2}\theta(k) < \infty$. Then, one can construct a sequence of i.i.d. Gaussian random variables $(Z_i)_{i\geq 1}$ centered and with variance σ^2, such that, setting $T_k = \sum_{i=1}^k Z_i$,

$$\sup_{k\leq n} |S_k - T_k| = o(n^{1/p}(\log n)^{1/2-1/p}) \ a.s.$$

(ii) Assume that $\theta(k) = O(k^{1-p})$ *if* $p \in]2,3[\cup]3,4[$ *and* $\sum_{k>0} k^{p-2}\theta(k) < \infty$ *if* $p \in \{3,4\}$. *Then, for any* $\eta > 1/2$, *one can construct a sequence of i.i.d. Gaussian random variables* $(Z_i)_{i\geq 1}$ *centered and with variance* σ^2, *such that, setting* $T_k = \sum_{i=1}^k Z_i$,

$$\sup_{k\leq n} |S_k - T_k| = o(n^{1/p}(\log n)^\eta) \text{ a.s.}$$

(iii) Let $p \in \{3,4\}$ *and assume that* $\theta(k) = O(k^{1-p})$. *Then, for any* $\eta > 1/2 + 1/p$, *one can construct a sequence of i.i.d. Gaussian random variables* $(Z_i)_{i\geq 1}$ *centered and with variance* σ^2, *such that, setting* $T_k = \sum_{i=1}^k Z_i$,

$$\sup_{k\leq n} |S_k - T_k| = o(n^{1/p}(\log n)^\eta) \text{ a.s.}$$

Remark 3.1. Let $(\beta(k))_{k\geq 1}$ be the sequence of the usual β-mixing coefficients. According to the definition of $\beta(k)$ given page 147 in [22], $\theta(k) \leq 2\|X_0\|_\infty^4 \beta(k)$. Hence, by Item (b) of Theorem 2.2 in [13], for any $p > 2$ there exists a stationary Markov chain $(X_k)_{k\geq 0}$ with uniform distribution such that $\theta(k) \leq Ck^{1-p}$ and

$$\limsup_{n\to\infty} (n\log n)^{-1/p} \big| S_n - \sum_{k=1}^n g_k \big| > 0 \text{ a.s.}$$

for any stationary and Gaussian centered sequence $(g_k)_{k\in\mathbb{Z}}$ with convergent series of covariances. Consequently, the rates in Theorem 3.2 (and then also in Theorem 3.1) are optimal up to a power of $\log n$.

4 Examples and applications

4.1 α-mixing sequences

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathcal{U}$ and $\mathcal{V}$ be two σ-algebras of $\mathcal{A}$. The strong mixing coefficient $\alpha(\mathcal{U}, \mathcal{V})$ between these σ-algebras is defined as follows:

$$\alpha(\mathcal{U}, \mathcal{V}) = \sup\{|\ \mathbb{P}(U \cap V) - \mathbb{P}(U)\mathbb{P}(V)| : U \in \mathcal{U}, V \in \mathcal{V}\}\,.$$

Next, for a stationary sequence $(Y_i)_{i\in\mathbb{Z}}$ of random variables with values in a Polish space S, define its strong mixing (or α-mixing) coefficients of order 4 as follows: Let

$$\alpha_{\infty,4}(n) = \sup_{i_4>i_3>i_2>i_1\geq n} \alpha(\mathcal{F}_0, \sigma(Y_{i_1}, Y_{i_2}, Y_{i_3}, Y_{i_4}))\,,$$

where $\mathcal{F}_0 = \sigma(Y_i, i \leq 0)$. As page 146 in [22], these coefficients can be rewritten in the following form: Let B_1 be the class of measurable functions from S^4 to $\mathbb{R}$ and bounded by one. Then

$$\alpha_{\infty,4}(n) = \frac{1}{4} \sup_{f\in B_1} \sup_{i_4>i_3>i_2>i_1\geq n} \big\| \mathbb{E}(f(Y_{i_1}, Y_{i_2}, Y_{i_3}, Y_{i_4})|\mathcal{F}_0) - \mathbb{E}(f(Y_{i_1}, Y_{i_2}, Y_{i_3}, Y_{i_4})) \big\|_1\,.$$

Let now f be a bounded measurable numerical function and let $X_k = f(Y_k) - \mathbb{E}(f(Y_k))$. Then Theorems 3.1 and 3.2 apply to the partial sum $S_n = \sum_{k=1}^n X_k$, replacing the conditions on $\theta(k)$ by the same conditions on $\alpha_{\infty,4}(k)$.

Consequently, our results apply to irreducible stationary S-valued Markov chains $(Y_i)_{i\geq 0}$ with invariant probability π and transition kernel $P(\cdot,\cdot)$ satisfying the following minorization condition: there exists a positive integer m such that

$$P^m(x,A) \geq s(x)\nu(A) \quad \text{for } x \in S,\ A \in \mathcal{B}(S)\,,$$

where s is a measurable function with values in $[0,1]$ such that $\pi(s) > 0$ and ν is a probability measure absolutely continuous with respect to π (see [25, Chapter 2] and [16, Th. 9.2.15] for the fact that the invariant probability measure is a maximal irreducibility measure).

Indeed, let us explain how the strong mixing coefficients of the chain $(Y_i)_{i\geq 0}$ can be computed (we simply denote by $\alpha(k)$ these coefficients, because in this Markovian setting, $\alpha(\sigma(Y_0),\sigma(Y_k)) = \alpha(\sigma(Y_0),\mathcal{G}_k)$, where $\mathcal{G}_k = \sigma(Y_i, i \geq k)$). Let $Z_k = Y_{km}$. It follows that $(Z_k)_{k\geq 0}$ is an irreducible stationary Markov chain satisfying the minorization condition with $m = 1$ and then the conditions of [26, Proposition 9.7]. Let $\alpha_Z(k)$ denote the α-mixing coefficients of $(Z_i)_{i\geq 0}$. According to [26, Page 165] (see also [3]), if τ is one of the return times (i.e. the difference between two regeneration times) of the extended chain constructed from $(Z_i)_{i\geq 0}$, then for $q \geq 1$, $\sum_{k\geq 1} k^{q-1}\alpha_Z(k) < \infty$ iff $\mathbb{E}(\tau^{q+1}) < \infty$. Now, since $\alpha(k) \leq \alpha_Z([k/m])$, we infer that $\mathbb{E}(\tau^{q+1}) < \infty$ also implies that $\sum_{k\geq 1} k^{q-1}\alpha(k) < \infty$. In particular, if τ is such that $\mathbb{E}(\tau^p) < \infty$, then $\sum_{k\geq 1} k^{p-2}\alpha(k) < \infty$, and Theorems 3.1 and 3.2 apply to $S_n = \sum_{k=1}^n (f(Y_k) - \mathbb{E}(f(Y_k)))$.

In their paper, Merlevède and Rio [24] proved the almost sure invariance principle with rate $O(\log n)$ when the Markov chain is geometrically ergodic and the minorization condition holds with $m = 1$. This last condition allows the use of the regeneration technique. However in some situations one can only prove that the minorization condition holds for $m > 1$ (see for instance Exemple 2.3 (f) in Nummelin [25]), or even that it cannot hold with $m = 1$ (even with the additional assumption that there exists a joint density for (X_0, X_1), see Kendall and Montana [17]). Nevertheless, as explained before, for this class of Markov chain we do not need to assume $m = 1$ since our conditions are only expressed in terms of the α-mixing coefficients of the chain.

4.2 α-dependent sequences

We start by recalling the definition of the α-dependence coefficients as considered in [10].

Definition 4.1. For any random variable $Y = (Y_1, \cdots, Y_k)$ with values in $\mathbb{R}^k$ and any σ-algebra $\mathcal{F}$, let $\alpha(\mathcal{F}, Y) = \sup_{(x_1,\ldots,x_k)\in\mathbb{R}^k} \big\| \mathbb{E}\big(\prod_{j=1}^k (\mathbb{1}_{Y_j \leq x_j})^{(0)} \mid \mathcal{F}\big)^{(0)} \big\|_1$, where we recall

that $Z^{(0)}$ means $Z - \mathbb{E}(Z)$. For the sequence $\mathbf{Y} = (Y_i)_{i\in\mathbb{Z}}$, let $\mathcal{F}_0 = \sigma(Y_i, i \le 0)$,

$$\alpha_{\mathbf{Y},4}(0) = 1 \text{ and } \alpha_{\mathbf{Y},4}(n) = \max_{1\le l\le 4} \sup_{n\le i_1\le\cdots\le i_l} \alpha(\mathcal{F}_0, (Y_{i_1}, \dots, Y_{i_l})) \text{ for } n > 0\,. \tag{4.1}$$

Let BV_1 be the space of bounded variation functions f such that $\|df\| \le 1$, where $\|df\|$ is the variation norm on $\mathbb{R}$ of the measure df. As mentioned in [10], $\alpha(\mathcal{F}, Y)$ can also be defined by

$$\alpha(\mathcal{F}, Y) = \sup_{(f_1,\dots,f_k)\in \mathrm{BV}_1} \left\| \mathbb{E}\Big(\prod_{j=1}^{k} f_j(Y_j)^{(0)} \Big| \mathcal{F} \Big)^{(0)} \right\|_1 ,$$

It follows that, if f is a bounded variation function such that $\|df\| \le C$, and $X_k = f(Y_k) - \mathbb{E}(Y_k)$, then $\theta(k) \le C^4 \alpha_{\mathbf{Y},4}(k)$. Then Theorem 3.1 and 3.2 apply to the partial sum $S_n = \sum_{k=1}^n X_k$, replacing the condition on $\theta(k)$ by the same conditions on $\alpha_{\mathbf{Y},4}(k)$.

From this result and proceeding as in [23, Section 3], we can derive rates in the strong approximation results for the partial sums associated with BV observables of the LSV map. More precisely, for $\gamma \in]0,1[$, let T_γ defined from $[0,1]$ to $[0,1]$ by

$$T_\gamma(x) = \begin{cases} x(1 + 2^\gamma x^\gamma) & \text{if } x \in [0, 1/2[\\ 2x - 1 & \text{if } x \in [1/2, 1]\,. \end{cases}$$

This is the so-called LSV [20] map with parameter γ. Recall, that there exists a unique T_γ-invariant measure ν_γ on $[0,1]$, which is absolutely continuous with respect to the Lebesgue measure with positive density denoted by h_γ. From [10, Prop. 1.17], we know that the coefficients $\alpha_{\mathbf{Y},4}(k)$ of the Markov chain associated with T_γ are exactly of order $1/k^{(1-\gamma)/\gamma}$. Consequently, if f is a BV observable, we get that:

- For any $\gamma \in]0, 1/2[$ and any $\varepsilon > 0$, one can construct on the probability space $([0,1]\times[0,1], \nu_\gamma \otimes \lambda)$ a sequence of i.i.d. Gaussian random variables $(Z_i)_{i\ge1}$ centered and with variance σ^2, such that

$$\sup_{k\le n} \left| \sum_{i=1}^{k} (f \circ T_\gamma^i(x) - \nu_\gamma(f)) - \sum_{i=1}^{k} Z_i(x,y) \right| = o(n^{\max(\gamma,1/4)} (\log n)^{\eta+\varepsilon}) \ \nu_\gamma \otimes \lambda \text{ a.e.} \tag{4.2}$$

  where $\eta = 3/4$ for $\gamma \le 1/4$, $\eta = 5/6$ if $\gamma = 1/3$ and $\eta = 1/2$ otherwise.

- For any $\gamma \in]0, 1/2[$, one can construct on the probability space $([0,1]\times[0,1], \nu_\gamma \otimes \lambda)$ a sequence of i.i.d. Gaussian random variables $(Z_i)_{i\ge1}$ centered and with variance σ^2, such that

$$\left(\int \sup_{k\le n} \left| \sum_{i=1}^{k} (f \circ T_\gamma^i(x) - \nu_\gamma(f)) - \sum_{i=1}^{k} Z_i(x,y) \right|^2 \nu_\gamma(dx)dy \right)^{\frac{1}{2}} = o(n^{\max(\gamma,1/4)} (\log n)^{\eta}) \,, \tag{4.3}$$

  where $\eta = 1/4$ if $\gamma < 1/4$, $\eta = 1/2$ if $\gamma = 1/4$ or $\gamma = 1/3$ and $\eta = 1/2 - \gamma$ otherwise.

4.3 Maps modelled by Young towers

Let (M, d) be a complete bounded separable metric space with the Borel σ-algebra. Let $T : M \mapsto M$ be a map that can be modelled by a Young tower (see [29], or for instance [6]), and denote by ν the T-invariant probability measure on M induced by this Young tower. Let φ be an Hölder observable from M to $\mathbb{R}$. Then, as explained in [14], one can compute the coefficients $\theta(k)$ of the sequence $(\varphi(X_i))_{i\geq 0}$, where $(X_i)_{i\geq 0}$ is the stationary Markov chain associated with T, whose transition kernel is the Perron-Frobenius operator of the composition by T with respect to ν. We infer that Theorems 3.1 and 3.2 apply to $(\varphi(X_i) - \nu(\varphi))_{i\geq 0}$, and also (proceeding as in [23, Section 3]) to the sequence $(\varphi(T^i) - \nu(\varphi))_{i\geq 0}$ on the probability space (M, ν).

More precisely, the behaviour of the coefficients $\theta(k)$ of the sequence $(\varphi(X_i))_{i\geq 0}$ depends on the behaviour of the return time to the base of the tower. For instance, if the return time has a moment of order $p > 1$, then $\sum_{k>0} k^{p-2}\theta(k) < \infty$; if it has a weak moment of order $p > 1$ then $\theta(k) = O(k^{1-p})$; if it has an exponential moment, then $\theta(k) = O(a^k)$, for some $a \in]0, 1[$ (see Proposition 2.1 and its remark 2.1 in [7] in case of strong moments and Proposition 5.3 in [11] in case of weak moments). Our results apply in particular to Hölder observables of the LSV map, leading to the same upper bounds (4.2)-(4.3) as for BV observables.

Note that, for Hölder observables of maps that can be modelled by a Young tower, optimal almost sure rates can be obtained via another method described in the paper [6] and inspired by [2]. In particular, for Hölder observables of the LSV map of parameter $\gamma < 1/2$, the optimal almost sure rate $o(n^\gamma(\log n)^{\gamma+\varepsilon})$ is given in [6]. However, no results similar to the $\mathbb{L}^2$ control (4.3) are given in [6]. Moreover, with our approach, we can also obtain rates for a larger class of continuous observables (including Hölder observables of any exponent) when the return time to the base as an exponential moment (see [14, Corollary 3.4]).

4.4 Rates of convergence in the functional CLT

Let $(X_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of centered and bounded random variables, and define the Donsker line

$$B_n(t) = \frac{1}{\sqrt{n}}\Big(\sum_{k=1}^{[nt]} X_k + (nt - [nt])X_{[nt]}\Big).$$

It is well known that, if $\sum_{k>0}\theta(k) < \infty$, then B_n converges in distribution to σB on the space $C([0, 1])$ with the uniform distance, where B is a standard Brownian motion, and σ^2 is the covariance series defined in Theorem 2.1.

One can now ask for the convergence rate in this functional CLT, with respect to the Wasserstein distance of order $p \geq 1$, that is rates for the quantity $W_p(P_{B_n}, P_{\sigma B})$, where

P_{B_n} and $P_{\sigma B}$ are the distributions of B_n and σB, and the cost function is $|\cdot|_\infty^p$, $|\cdot|_\infty$ being the supremum norm on $C([0,1])$. Note that, by definition of W_p,

$$W_p(P_{B_n}, P_{\sigma B}) \le \Big\| \sup_{0\le t\le 1} |B_n(t) - \sigma B(t)| \Big\|_p ,$$

for any standard Brownian motion B. Consequently Theorem 3.1 applies when $p = 2$. For instance, if $\sum_{k>0} k^2\theta(k) < \infty$, Item (i) of Theorem 3.1 implies that

$$W_2(P_{B_n}, P_{\sigma B}) = O\big(n^{-1/4}(\log n)^{1/4}\big) . \tag{4.4}$$

Note that, according to Section 4.3, the upper bound (4.4) applies to sequences $(\varphi(T^i) - \nu(\varphi))_{i\ge 0}$, where φ is an Hölder observable and T can be modelled by a Young tower with return time to the base having a moment of order 4. Let us compare this result with some recent results obtained in this context by Liu and Wang [19].

Assume that the return time to the base has a finite moment of order $q \ge 4$. Using the Skorohod embedding theorem, Liu and Wang [19] (see their Theorem 3.4) proved that

$$W_{q/2}(P_{B_n}, P_{\sigma B}) = O(n^{-(q-2)/(4(q-1))}) . \tag{4.5}$$

Therefore, for $q = 4$, (4.5) gives the rate $O\big(n^{-1/6}\big)$, while the upper bound (4.4) gives the rate $O\big(n^{-1/4}(\log n)^{1/4}\big)$.

5 Proof of Theorem 2.1

Starting from inequality (A.42) in [23] (in the bounded case) together with the fact that $\|\mathbb{E}_{i-q}(X_i)\|_1 \le \theta(q)$, we infer that for any nondecreasing, non negative and convex function φ and any $x > 0$,

$$\mathbb{P}\big(S_n^* \ge 4x\big) \le \frac{\mathbb{E}(\varphi(S_n))}{\varphi(x)} + nx^{-1}\theta([x]) .$$

Now, since $x^3 \le 8\sum_{k=1}^{[x]} k^2$ when $x \ge 1$, $(\theta(k))_{k\ge 0}$ is non increasing and $\theta(0) \le 1$ (since $\|X_0\|_\infty = 1$),

$$x^3\theta([x]) \le 1 + 8\sum_{k=1}^{[x]} k^2\theta(k) \le 1 + 8\sum_{k\ge 1} k(k\wedge x)\theta(k) .$$

Hence

$$\mathbb{P}\big(S_n^* \ge 4x\big) \le \frac{\mathbb{E}(\varphi(S_n))}{\varphi(x)} + \frac{8n}{x^4}\Big(1 + \sum_{k\ge 1} k(k\wedge x)\theta(k)\Big) . \tag{5.1}$$

Next we handle the first term in the right-hand side of (5.1) with the following selection of φ: For any real t,

$$\varphi(t) = \begin{cases} 0 & \text{if } t \le x/2 \\ \frac{1}{24}(t - \frac{x}{2})^4 & \text{if } x/2 \le t \le x \\ \frac{x^4}{24\times 2^4} + \frac{x}{12}(t-x)^3 + \frac{x^2}{16}(t-x)^2 + \frac{x^3}{48}(t-x) & \text{if } t \ge x . \end{cases}$$

This is a nondecreasing and convex function such that $\|\varphi^{(3)}\|_\infty = x/2$ and $\|\varphi^{(4)}\|_\infty = 1$. Furthermore $384\,\varphi(x) = x^4$, whence $\mathbb{E}(\varphi(S_n))/\varphi(x) = 384\,x^{-4}\mathbb{E}(\varphi(S_n))$.

To bound up $\mathbb{E}(\varphi(S_n))$, we start by a symmetrization argument. Let $(X'_k)_{k\in\mathbb{Z}}$ be a stationary sequence independent of the sequence $(X_i)_{i\in\mathbb{Z}}$ and with the same joint law as $(X_i)_{i\in\mathbb{Z}}$. Set $S'_n = X'_1 + X'_2 + \cdots + X'_n$. Since S'_n is centered and independent of S_n, it follows from the conditional version of the Jensen inequality that

$$\mathbb{E}\big(\varphi(S_n)\big) \le \mathbb{E}\big(\varphi(S_n - S'_n)\big)\,. \tag{5.2}$$

Hence Inequality (2.1) will follow from (5.1) if we prove that

$$\mathbb{E}\big(\varphi(S_n - S'_n)\big) \ll \mathbf{1}_{\{\sigma^2>0\}}\Big(\frac{n\sigma^2}{x}\Big)^4 \exp\Big(-\frac{x^2}{16n\sigma^2}\Big) + n\Big(\Theta_1\Theta_2 + \sum_{k\ge1} k(k\wedge x)\theta(k)\Big)\,. \tag{5.3}$$

Define then the stationary sequence $(Z_i)_{i\in\mathbb{Z}}$ of centered and bounded random variables by $Z_i = X_i - X'_i$ for any integer i, and set

$$\tilde S_0 = 0 \ \text{ and } \tilde S_n = Z_1 + Z_2 + \cdots + Z_n = S_n - S'_n \text{ for any integer } n > 0. \tag{5.4}$$

From the definition of $(Z_i)_{i\in\mathbb{Z}}$,

$$\mathbb{E}(Z_0Z_k) = 2\mathbb{E}(X_0X_k) \text{ for any } k\in\mathbb{N}, \text{ whence } \mathbb{E}(Z_0^2) + 2\sum_{k\ge1}\mathbb{E}(Z_0Z_k) = 2\sigma^2. \tag{5.5}$$

To prove (5.3), we shall apply the Lindeberg method to $(Z_i)_{i\in\mathbb{Z}}$: we consider a sequence $(Y_k)_{k\ge1}$ of i.i.d. random variables with normal law $N(0, 2\sigma^2)$, independent of $(Z_i)_{i\in\mathbb{Z}}$. Set $T_0 = 0$ and $T_n = Y_1 + Y_2 + \cdots + Y_n$ for $n > 0$. Clearly

$$\mathbb{E}(\varphi(\tilde S_n)) = \mathbb{E}(\varphi(T_n)) + \big(\mathbb{E}(\varphi(\tilde S_n)) - \mathbb{E}(\varphi(T_n))\big). \tag{5.6}$$

We start by computing $\mathbb{E}(\varphi(T_n))$. If $\sigma^2 = 0$, then $T_n = 0$ and $\mathbb{E}(\varphi(T_n)) = 0$. If $\sigma^2 > 0$, then T_n has the normal law $N(0, 2n\sigma^2)$ and consequently

$$\mathbb{P}\big(T_n \ge t + x/2\big) \le \exp\big(-(t+x/2)^2/(4n\sigma^2)\big)\,,$$

Since $(t+x/2)^2 \ge tx + x^2/4$, we derive that

$$\int_0^{+\infty} t^3\mathbb{P}\big(T_n \ge t + x/2\big)dt \le e^{-x^2/(16n\sigma^2)}\int_0^{+\infty} t^3 e^{-tx/(4n\sigma^2)}dt \le 6\Big(\frac{4n\sigma^2}{x}\Big)^4 e^{-x^2/(16n\sigma^2)}\,.$$

So, overall,

$$\mathbb{E}(\varphi(T_n)) \le \mathbb{E}((T_n - x/2)_+^4) \le 4\int_0^{+\infty} t^3\mathbb{P}\big((T_n \ge t + x/2\big)dt \ll \Big(\frac{n\sigma^2}{x}\Big)^4 e^{-x^2/(16n\sigma^2)}\,.$$

According to (5.6) and the above inequality, to end the proof of Inequality (5.3), it remains to prove that

$$\mathbb{E}(\varphi(\tilde{S}_n)) - \mathbb{E}(\varphi(T_n)) \ll n\big(\Theta_1\Theta_2 + \sum_{k\geq 1} k(k\wedge x)\theta(k)\big)\,. \tag{5.7}$$

With this aim, let $\varphi_{n-k}(t) = \mathbb{E}(\varphi(t + T_n - T_k))$ and define, for $k \geq 1$,

$$\Delta_{n,k} = \varphi_{n-k}(\tilde{S}_{k-1} + Z_k) - \varphi_{n-k}(\tilde{S}_{k-1} + Y_k)\,.$$

The functions $\varphi_{n-k}(\cdot)$ are $\mathcal{C}^\infty$, $\|\varphi^{(3)}_{n-k}\|_\infty = b_3 \leq x/2$ and $\|\varphi^{(4)}_{n-k}\|_\infty = b_4 \leq 1$. Since the sequence $(Y_i)_{i\in\mathbb{N}}$ is independent of the sequence $(Z_i)_{i\in\mathbb{Z}}$,

$$\mathbb{E}(\varphi(\tilde{S}_n)) - \mathbb{E}(\varphi(T_n)) = \sum_{k=1}^{n} \mathbb{E}(\Delta_{n,k}). \tag{5.8}$$

Notation 5.1. Set $\Delta^{(1)}_{n,k} = \varphi'_{n-k}(\tilde{S}_{k-1})(Z_k - Y_k)$, $\Delta^{(2)}_{n,k} = \varphi''_{n-k}(\tilde{S}_{k-1})(Z_k^2 - Y_k^2)$ and $\Delta^{(3)}_{n,k} = \varphi^{(3)}_{n-k}(\tilde{S}_{k-1})(Z_k^3 - Y_k^3)$. Let $\Delta^*_{n,k} = \Delta^{(1)}_{n,k} + \frac{1}{2}\Delta^{(2)}_{n,k} + \frac{1}{6}\Delta^{(3)}_{n,k}$.

With the above notations, from the Taylor integral formula at order 4,

$$\Delta_{n,k} = \Delta^*_{n,k} + R_{n,k}\,, \tag{5.9}$$

with $R_{n,k} = \dfrac{1}{6}Z_k^4\displaystyle\int_0^1 (1-s)^3\varphi^{(4)}_{n-k}(\tilde{S}_{k-1} + sZ_k)ds + \frac{1}{6}Y_k^4\int_0^1 (1-s)^3\varphi^{(4)}_{n-k}(\tilde{S}_{k-1} + sY_k)ds\,.$

Since $0 \leq \varphi^{(4)}_{n-k} \leq 1$, $\|Z_k\|_\infty \leq 2$ and $\mathbb{E}(Y_k^4) = 12\sigma^4$, we derive that

$$0 \leq \sum_{k=1}^{n} \mathbb{E}(R_{n,k}) \leq n(1+\sigma^4) \leq 4n(1+\Theta_1\Theta_2)\,. \tag{5.10}$$

Indeed $\sigma^4 \leq 4\Theta_1^2 \leq 4\Theta_1\Theta_2$. Next

$$\mathbb{E}(\Delta^*_{n,k}) = \mathbb{E}(\Delta_{n,k} - R_{n,k}) = \mathbb{E}(\Delta^{(1)}_{n,k}) + \tfrac{1}{2}\mathbb{E}(\Delta^{(2)}_{n,k}) + \tfrac{1}{6}\mathbb{E}(\Delta^{(3)}_{n,k})\,. \tag{5.11}$$

From the fact that $(Y_k)_{k\geq 1}$ is independent of $(Z_i)_{i\in\mathbb{Z}}$, $\mathbb{E}(\Delta^{(1)}_{n,k}) = \mathbb{E}(\varphi'_{n-k}(\tilde{S}_{k-1})Z_k)$, $\mathbb{E}(\Delta^{(2)}_{n,k}) = \mathbb{E}(\varphi''_{n-k}(\tilde{S}_{k-1})(Z_k^2 - 2\sigma^2))$ and $\mathbb{E}(\Delta^{(3)}_{n,k}) = \mathbb{E}(\varphi^{(3)}_{n-k}(\tilde{S}_{k-1})Z_k^3)$. We now develop each term in the right-hand side of (5.11) with the help of the Lindeberg method.

Since $\mathbb{E}(\varphi'_{n-k}(0)Z_k) = \varphi'_{n-k}(0)\mathbb{E}(Z_k) = 0$, we have

$$\begin{aligned}\mathbb{E}(\Delta^{(1)}_{n,k}) &= \sum_{i=1}^{k-1} \mathbb{E}\big(\{\varphi'_{n-k}(\tilde{S}_{k-i}) - \varphi'_{n-k}(\tilde{S}_{k-i-1})\}Z_k\big) \\ &= \mathbb{E}(\Delta^{(1)}_{n,k,2}) + \tfrac{1}{2}\mathbb{E}(\Delta^{(1)}_{n,k,3}) + A^{(1)}_{n,k,2} + \tfrac{1}{2}A^{(1)}_{n,k,3} + B^{(1)}_{n,k}\,,\end{aligned} \tag{5.12}$$

where the following notations have been used: for $j = 2, 3$,

$$\Delta^{(1)}_{n,k,j} = \sum_{i=1}^{k-1} \varphi^{(j)}_{n-k}(\tilde{S}_{k-i-1})(Z^{j-1}_{k-i}Z_k)^{(0)},\ A^{(1)}_{n,k,j} = \sum_{i=1}^{k-1} \mathbb{E}\{\varphi^{(j)}_{n-k}(\tilde{S}_{k-i-1})\}\mathbb{E}(Z^{j-1}_{k-i}Z_k),$$

$$\text{and } B^{(1)}_{n,k} = \frac{1}{2}\sum_{i=1}^{k-1}\int_0^1 (1-s)^2\mathbb{E}\big(\varphi^{(4)}_{n-k}(\tilde{S}_{k-i-1} + sZ_{k-i})Z^3_{k-i}Z_k\big)ds\,.$$

In the decomposition (5.12), $A^{(1)}_{n,k,3} = 0$. This is due to the lemma below, whose proof uses the symmetry properties of $(Z_i)_{i\in\mathbb{Z}}$.

Lemma 5.1. *For any integers i, j and k, $\mathbb{E}(Z_iZ_jZ_k) = 0$.*

Proof of Lemma 5.1. From the definition of $(Z_i)_{i\in\mathbb{Z}}$, $(-Z_i, -Z_j, -Z_k)$ has the same joint law as (Z_i, Z_j, Z_k). Hence $\mathbb{E}(-Z_iZ_jZ_k) = \mathbb{E}(Z_iZ_jZ_k)$, which implies Lemma 5.1. □

Next, concerning the coefficients $(\theta_{Z,p,q}(k))_{k\geq 0}$ associated with the sequence $(Z_i)_{i\in\mathbb{Z}}$, they can be compared with the former coefficients $(\theta_{X,p,q}(k))_{k\geq 0}$ as follows.

Lemma 5.2. *For any integer $k \geq 0$ and any positive integers p and q,*

$$\theta_{Z,p,q}(k) \leq 2^{q+1}\theta_{X,p,q}(k)\,.$$

Proof of Lemma 5.2. Let $\mathcal{F}'_0 = \sigma(X'_i, i \leq 0)$ and $\tilde{\mathcal{F}}_0 = \mathcal{F}_0 \vee \mathcal{F}'_0$. By definition,

$$\theta_{Z,p,q}(k) = \sup_{\substack{k_p > k_{p-1} > \ldots > k_2 > k_1 \geq k \\ (a_1,\ldots,a_p)\in\Gamma_{p,q}}} \Big\|\mathbb{E}\Big(\prod_{i=1}^p Z^{a_i}_{k_i}|\tilde{\mathcal{F}}_0\Big) - \mathbb{E}\Big(\prod_{i=1}^p Z^{a_i}_{k_i}\Big)\Big\|_1\,.$$

Let $(a_i)_{1\leq i\leq p}$ in $\Gamma_{p,q}$ and $k_1, \cdots, k_p$ such that $k_p > k_{p-1} > \ldots > k_2 > k_1 \geq k$. Note that $Z^{a_i}_{k_i} = \sum_{\ell=0}^{a_i} C^\ell_{a_i} X^\ell_{k_i}(-X'_{k_i})^{a_i-\ell}$. Next, let $(\alpha_i)_{1\leq i\leq p}$, $(\beta_i)_{1\leq i\leq p}$ in $\Gamma_{p,q}$ and $k_1, \cdots, k_p$ such that $k_p > k_{p-1} > \ldots > k_2 > k_1 \geq k$. Setting $U = \prod_{i=1}^p X^{\alpha_i}_{k_i}$, $V = \prod_{i=1}^p (X'_{k_i})^{\beta_i}$ and $\sigma(X') = \sigma(X'_k, k \in \mathbb{Z})$, note that

$$\begin{aligned}\big\|\mathbb{E}\big(UV|\tilde{\mathcal{F}}_0\big) - \mathbb{E}\big(UV\big)\big\|_1 &= \big\|\mathbb{E}\big(\mathbb{E}\big(UV|\sigma(X')\vee\mathcal{F}_0\big)|\tilde{\mathcal{F}}_0\big) - \mathbb{E}\big(UV\big)\big\|_1 \\ &= \big\|\mathbb{E}\big(V\mathbb{E}\big(U|\sigma(X')\vee\mathcal{F}_0\big)|\tilde{\mathcal{F}}_0\big) - \mathbb{E}(V)\mathbb{E}(U)\big\|_1\,.\end{aligned}$$

Next we use the following well-known fact. Let Y be an integrable random variable, and $\mathcal{G}_1$ and $\mathcal{G}_2$ be two σ-algebras such that $\sigma(Y)\vee\mathcal{G}_1$ is independent of $\mathcal{G}_2$, then

$$\mathbb{E}\big(Y|\mathcal{G}_1\vee\mathcal{G}_2\big) = \mathbb{E}\big(Y|\mathcal{G}_1\big) \quad \text{a.s.}$$

From the above fact and since we assume that the X_i's are uniformly bounded by one,

$$\begin{aligned}\big\|\mathbb{E}\big(UV|\tilde{\mathcal{F}}_0\big) - \mathbb{E}\big(UV\big)\big\|_1 &= \big\|\mathbb{E}\big(V\big\{\mathbb{E}\big(U|\mathcal{F}_0\big) - \mathbb{E}(U)\big\}|\tilde{\mathcal{F}}_0\big) + \big\{\mathbb{E}\big(V|\tilde{\mathcal{F}}_0\big) - \mathbb{E}(V)\big\}\mathbb{E}(U)\big\|_1 \\ &\leq \big\|\mathbb{E}\big(U|\mathcal{F}_0\big) - \mathbb{E}(U)\big\|_1 + \big\|\mathbb{E}\big(V|\mathcal{F}'_0\big) - \mathbb{E}(V)\big\|_1 \leq 2\theta_{X,p,q}(k)\,.\end{aligned}$$

So, overall,

$$\theta_{Z,p,q}(k) \le 2 \sup_{(a_1,\dots,a_p)\in\Gamma_{p,q}} \prod_{i=1}^{p} 2^{a_i}\theta_{X,p,q}(k)\,,$$

proving the lemma. □

We now handle the rests $B_{n,k}^{(1)}$. Since $\|Z_{k-i}^3\varphi_{n-k}^{(4)}(\tilde{S}_{k-i-1}+sZ_{k-i})\|_\infty \le 2$,

$$\sum_{k=1}^{n} |B_{n,k}^{(1)}| \ll n\Theta_1\,. \tag{5.13}$$

To handle the second term in the right-hand side of (5.11), we introduce the following additional notations.

Notation 5.2. Let $\gamma_i = \mathbb{E}(Z_0Z_i)$. Define $\beta_{2,k} = 2\sum_{i=1}^{k-1}\gamma_i$ and $\beta_2^{(k)} = 2\sum_{i\ge k}\gamma_i$.

Since $\mathbb{E}(Z_{k-i}Z_k) = \gamma_i$, note that

$$\begin{aligned}\frac{1}{2}\mathbb{E}\{\varphi''_{n-k}(\tilde{S}_{k-1})\}\beta_{2,k} - A_{n,k,2}^{(1)} &= \sum_{i=1}^{k-1}\gamma_i\sum_{j=1}^{i}\mathbb{E}\{\varphi''_{n-k}(\tilde{S}_{k-j}) - \varphi''_{n-k}(\tilde{S}_{k-j-1})\}\\ &= \sum_{i=1}^{k-1}\gamma_i\sum_{j=1}^{i}\mathbb{E}\big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}\big\} + r_{n,k,2}^{(1)}\,,\end{aligned} \tag{5.14}$$

where

$$r_{n,k,2}^{(1)} := \int_0^1 (1-t)\sum_{i=1}^{k-1}\gamma_i\sum_{j=1}^{i}\mathbb{E}\big\{\varphi^{(4)}_{n-k}(\tilde{S}_{k-j-1}+tZ_{k-j})Z_{k-j}^2\big\}dt\,.$$

From the fact that $\|\varphi^{(4)}_{n-k}(\tilde{S}_{k-j-1}+tZ_{k-j})Z_{k-j}^2\|_1 \le b_4\mathbb{E}(Z_0^2) \le 4b_4$, it follows that

$$\sum_{k=1}^{n}|r_{n,k,2}^{(1)}| \ll n\sum_{i=1}^{n} i\gamma_i \ll n\Theta_2. \tag{5.15}$$

Starting from (5.11) and taking into account (5.12), (5.13), (5.15) and the definition of σ^2, we get

$$\begin{aligned}\mathbb{E}(\Delta_{n,k}^*) = \mathbb{E}(\Delta_{n,k,2}^{(1)}) &+ \frac{1}{2}\mathbb{E}\big(\varphi''_{n-k}(\tilde{S}_{k-1})(Z_k^2)^{(0)}\big) - \frac{1}{2}\mathbb{E}\big(\varphi''_{n-k}(\tilde{S}_{k-1})\big)\beta_2^{(k)} + \frac{1}{2}\mathbb{E}(\Delta_{n,k,3}^{(1)})\\ &- \sum_{i=1}^{k-1}\gamma_i\sum_{j=1}^{i}\mathbb{E}\big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}\big\} + \frac{1}{6}\mathbb{E}\big(\varphi^{(3)}_{n-k}(\tilde{S}_{k-1})Z_k^3\big) + \Gamma_{n,k}^{(1)}\,,\end{aligned} \tag{5.16}$$

where $\Gamma_{n,k}^{(1)}$ satisfies

$$\sum_{k=1}^{n}|\Gamma_{n,k}^{(1)}| \ll n\Theta_2\,. \tag{5.17}$$

Note that

$$\big|\mathbb{E}\big(\varphi''_{n-k}(\tilde{S}_{k-1})\big)\big| \le \mathbb{E}\big(\tilde{S}_{k-1} + T_n - T_k\big)^2 = 2\mathbb{E}(S_{k-1}^2) + 2(n-k)\sigma^2 \le 4n\Theta_1\,.$$

Hence

$$\sum_{k=1}^{n} |\mathbb{E}\big(\varphi''_{n-k}(\tilde{S}_{k-1})\big)\beta_2^{(k)}| \ll n\Theta_1 \sum_{k=1}^{n}\sum_{i\ge k}\theta(i) \ll n\Theta_1\Theta_2\,. \tag{5.18}$$

We now handle the quantity $\mathbb{E}\big(\varphi''_{n-k}(\tilde{S}_{k-i-1})(Z_{k-i}Z_k)^{(0)}\big)$ appearing in the first two terms of the right hand side of (5.16) for any integer i in $[0, k-1]$. With this aim, noticing that $\mathbb{E}\big(\varphi''_{n-k}(0)(Z_{k-i}Z_k)^{(0)}\big) = 0$,

$$\mathbb{E}\big\{\varphi''_{n-k}(\tilde{S}_{k-i-1})(Z_{k-i}Z_k)^{(0)}\big\} = \sum_{j=i+1}^{k-1} \mathbb{E}\big\{(\varphi''_{n-k}(\tilde{S}_{k-j}) - \varphi''_{n-k}(\tilde{S}_{k-j-1}))(Z_{k-i}Z_k)^{(0)}\big\},$$

whence

$$\begin{aligned}\mathbb{E}\big\{\varphi''_{n-k}(\tilde{S}_{k-i-1})(Z_{k-i}Z_k)^{(0)}\big\} &= \sum_{j=i+1}^{k-1} \mathbb{E}\big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\} \\ &\quad + \sum_{j=i+1}^{k-1}\int_0^1 (1-t)\mathbb{E}\big\{\varphi^{(4)}_{n-k}(\tilde{S}_{k-j-1} + tZ_{k-j})Z_{k-j}^2(Z_{k-i}Z_k)^{(0)}\big\}dt\,. \end{aligned}\tag{5.19}$$

Since $\|\varphi^{(4)}_{n-k}(\tilde{S}_{k-j-1} + tZ_{k-j})\|_\infty \le 1$, it follows that

$$\begin{aligned}\sum_{k=1}^{n}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1} &|\mathbb{E}\big\{\varphi^{(4)}_{n-k}(\tilde{S}_{k-j-1} + tZ_{k-j})Z_{k-j}^2(Z_{k-i}Z_k)^{(0)}\big\}| \\ &\ll \sum_{k=1}^{n}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}\theta(j-i)\wedge\theta(i) \ll n\Theta_2\,. \end{aligned}\tag{5.20}$$

Next, set $\Delta^{(1,3)}_{n,k,2}(i) = \sum_{j=i+1}^{k-1}\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}(Z_{k-i}Z_k)^{(0)}$. Starting from (5.16) and taking into account (5.17), (5.18), (5.19) and (5.20), we get

$$\begin{aligned}\mathbb{E}(\Delta^*_{n,k}) &= \frac{1}{2}\mathbb{E}(\Delta^{(1,3)}_{n,k,2}(0)) + \sum_{i=1}^{k-1}\mathbb{E}(\Delta^{(1,3)}_{n,k,2}(i)) + \frac{1}{2}\mathbb{E}(\Delta^{(1)}_{n,k,3}) \\ &\quad - \sum_{i=1}^{k-1}\gamma_i\sum_{j=1}^{i}\mathbb{E}\big(\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}\big) + \frac{1}{6}\mathbb{E}\big(\varphi^{(3)}_{n-k}(\tilde{S}_{k-1})Z_k^3\big) + \Gamma^{(2)}_{n,k}\,, \end{aligned}\tag{5.21}$$

where the rests $\Gamma^{(2)}_{n,k}$ satisfy

$$\sum_{k=1}^{n}|\Gamma^{(2)}_{n,k}| \ll n\Theta_2 + n\Theta_1\Theta_2\,. \tag{5.22}$$

Introduce now the following notations: for any integer i, let

$$\tilde{\Delta}^{(1,3)}_{n,k,2}(i) := \sum_{j=i+1}^{k-1} \big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})(Z_{k-j}(Z_{k-i}Z_k)^{(0)})^{(0)}\big\}\,.$$

Starting from (5.21), and using the fact that $\mathbb{E}\big\{Z_{k-j}Z_{k-i}Z_k\big\} = \mathbb{E}\big\{Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\} = 0$ thanks to Lemma 5.1, we get that

$$\begin{aligned}\mathbb{E}(\Delta^*_{n,k}) = \frac{1}{2}\mathbb{E}(\tilde{\Delta}^{(1,3)}_{n,k,2}(0)) + \sum_{i=1}^{k-1}\mathbb{E}(\tilde{\Delta}^{(1,3)}_{n,k,2}(i)) + \frac{1}{6}\mathbb{E}\big(\varphi^{(3)}_{n-k}(\tilde{S}_{k-1})Z_k^3\big)\big)\\ + \frac{1}{2}\mathbb{E}(\Delta^{(1)}_{n,k,3}) - \sum_{i=1}^{k-1}\gamma_i\sum_{j=1}^{i}\mathbb{E}\big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}\big\} + \Gamma^{(2)}_{n,k}. \quad (5.23)\end{aligned}$$

We handle now the first two terms in the right-hand side of (5.23). Since $\|\varphi^{(3)}_{n-k}\|_\infty \le x/2$,

$$\big|\mathbb{E}\big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\}\big| \ll x\{\theta(j-i)\wedge\theta(i)\}\,,$$

whence

$$|\mathbb{E}(\tilde{\Delta}^{(1,3)}_{n,k,2}(i))| \ll x\Big(i\theta(i) + \sum_{j=i}^{k-i}\theta(j)\Big)\,. \quad (5.24)$$

On another hand, since $\mathbb{E}\big\{\varphi^{(3)}_{n-k}(0)Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\} = 0$, we write

$$\begin{aligned}&\mathbb{E}\big\{\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-1})Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\}\\ &\quad= \sum_{u=1}^{k-j-1}\mathbb{E}\big\{\big(\varphi^{(3)}_{n-k}(\tilde{S}_{k-j-u}) - \varphi^{(3)}_{n-k}(\tilde{S}_{k-j-u-1})\big)Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\}\\ &\quad= \sum_{u=1}^{k-j-1}\int_0^1\mathbb{E}\big\{\varphi^{(4)}_{n-k}(\tilde{S}_{k-j-u-1}+tZ_{k-j-u})Z_{k-j-u}Z_{k-j}(Z_{k-i}Z_k)^{(0)}\big\}dt\,.\end{aligned}$$

Since $\|\varphi^{(4)}_{n-k}(S_{k-j-u-1}+tX_{k-j-u})\|_\infty \le 1$, it follows that

$$|\mathbb{E}(\tilde{\Delta}^{(1,3)}_{n,k,2}(i))| \ll \sum_{j=1}^{k}\sum_{u=1}^{k}\big(\theta(u)\wedge\theta(j)\wedge\theta(i)\big) \ll \sum_{j=1}^{k} j\big(\theta(j)\wedge\theta(i)\big)\,. \quad (5.25)$$

Therefore, using the upper bound (5.24) when $i \ge x$ and the upper bound (5.25) when $i < x$, we derive

$$\begin{aligned}\sum_{i=0}^{n}|\mathbb{E}(\tilde{\Delta}^{(1)}_{2,3,i})| &\ll x\sum_{i\ge[x]} i\theta(i) + \sum_{i=0}^{[x]}\sum_{j=1}^{[x]} j\big(\theta(j)\wedge\theta(i)\big) + \sum_{i=0}^{[x]}\sum_{j\ge[x]} j\big(\theta(j)\wedge\theta(i)\big)\\ &\ll x\sum_{i\ge[x]} i\theta(i) + \sum_{i=0}^{[x]}(i+1)^2\theta(i) + \sum_{j\ge1} j(j\wedge x)\theta(j)\,.\end{aligned}$$

Hence

$$\sum_{k=1}^{n}\sum_{i=0}^{k}|\mathbb{E}(\tilde{\Delta}_{2,3,i}^{(1)})| \ll n\Big(1+\sum_{k\geq 1}k(k\wedge x)\theta(k)\Big)\,. \tag{5.26}$$

With similar arguments, we infer that

$$\sum_{k=1}^{n}\big\{\big|\mathbb{E}\big(\varphi_{n-k}^{(3)}(\tilde{S}_{k-1})Z_k^3\big)\big|+\big|\mathbb{E}(\Delta_{n,k,3}^{(1)})\big|\big\} \ll n\Theta_2\,. \tag{5.27}$$

Starting from (5.23) and taking into account (5.22), (5.26) and (5.27), it follows that, if one can prove that

$$\sum_{k=1}^{n}\sum_{i=1}^{k}\sum_{j=1}^{i}\big|\gamma_i\mathbb{E}\big\{\varphi_{n-k}^{(3)}(\tilde{S}_{k-j-1})Z_{k-j}\big\}\big| \ll n\Theta_1\Theta_2, \tag{5.28}$$

then

$$\sum_{k=1}^{n}\mathbb{E}(\Delta_{n,k}^{*}) \ll n\big(\Theta_1\Theta_2+\sum_{k=1}^{n}k(k\wedge x)\theta(k)\big)\,.$$

This last upper bound together with (5.8), (5.9), (5.10) and (5.11) will end the proof of (5.7). The rest of the proof is devoted to the proof of (5.28). Since $\mathbb{E}\big\{\varphi_{n-k}^{(3)}(0)Z_{k-j}\big\}=0$, we have

$$\begin{aligned}\big|\mathbb{E}\big\{\varphi_{n-k}^{(3)}(\tilde{S}_{k-j-1})Z_{k-j}\big\}\big| &= \Big|\sum_{u=1}^{k-j-1}\mathbb{E}\big\{\big(\varphi_{n-k}^{(3)}(\tilde{S}_{k-j-u})-\varphi_{n-k}^{(3)}(\tilde{S}_{k-j-u-1})\big)Z_{k-j}\big\}\Big| \\ &= \Big|\sum_{u=1}^{k-j-1}\int_0^1\mathbb{E}\big\{\varphi_{n-k}^{(4)}(\tilde{S}_{k-j-u-1}+tZ_{k-j-u})Z_{k-j-u}Z_{k-j}\big\}dt\Big| \ll \sum_{u=1}^{k-j-1}\theta(u)\,.\end{aligned}$$

Therefore

$$\sum_{k=1}^{n}\sum_{i=1}^{k}\sum_{j=1}^{i}\big|\gamma_i\mathbb{E}\big\{\varphi_{n-k}^{(3)}(\tilde{S}_{k-j-1})Z_{k-j}\big\}\big| \ll \sum_{k=1}^{n}\sum_{i=1}^{k}\sum_{j=1}^{i}\sum_{u=1}^{k-j-1}\theta(u)\theta(i) \ \ll n\Theta_1\Theta_2\,,$$

which ends the proof of (5.28) and then of (5.7). This ends the proof of the theorem. □

6 Proof of Theorems 3.1 and 3.2 (case $\sigma^2>0$)

Starting from the construction of the $V_{k,L}$ given in (3.1), we now construct a suitable sequence $(Z_i)_{i\geq 1}$ of i.i.d. Gaussian random variables, centered and with variance σ^2. Let $Z_1=\sigma\Phi^{-1}(\delta_1)$. For any $L\in\mathbb{N}$ and any $k\in\{1,\cdots,2^{L-m(L)}\}$ the random variables $(Z_{2^L+(k-1)2^{m(L)}+1},\ldots,Z_{2^L+k2^{m(L)}})$ are defined in the following way. If $m(L)=0$, then $Z_{2^L+k2^{m(L)}}=V_{k,L}$. If $m(L)>0$, then by the Skorohod lemma [27], there exists a measurable function g from $\mathbb{R}\times[0,1]$ in $\mathbb{R}^{2^{m(L)}}$ such that, for any pair (V,δ) of independent

random variables with respective laws $N(0,\sigma^2 2^{m(L)})$ and the uniform distribution over $[0,1]$, $g(V,\delta)=(N_1,\dots,N_{2^{m(L)}})$ is a Gaussian random vector with i.i.d. components such that $V=N_1+\cdots+N_{2^{m(L)}}$. We then set

$$(Z_{2^L+(k-1)2^{m(L)}+1},\dots,Z_{2^L+k2^{m(L)}})=g(V_{k,L},\delta_{2^L+(k-1)2^{m(L)}+1})\,.$$

The so defined sequence (Z_i) has the prescribed distribution.

Set $S_j=\sum_{i=1}^j X_i$ and $T_j=\sum_{i=1}^j Z_i$. Let $D_L=\sup_{\ell\le 2^L}|\sum_{i=2^L+1}^{2^L+\ell}(X_i-Z_i)|$. Then, proceeding exactly as in [23], page 394, for any N in $\mathbb{N}^*$,

$$\sup_{1\le k\le 2^{N+1}}|S_k-B_k|\le|X_1-Z_1|+D_0+D_1+\cdots+D_N\,. \tag{6.1}$$

It remains to bound up the random variables D_L. We first notice that the following decomposition is valid:

$$D_L\le D_{L,1}+D_{L,2}\,, \tag{6.2}$$

where, recalling that $U_{k,L}=\sum_{i\in I_{k,L}}X_i$, we set

$$D_{L,1}:=\sup_{k\le 2^{L-m(L)}}\Big|\sum_{\ell=1}^k(U_{\ell,L}-V_{\ell,L})\Big| \text{ and } D_{L,2}:=\sup_{k\le 2^{L-m(L)}}\sup_{\ell\in I_{k,L}}\Big|\sum_{i=\inf I_{k,L}}^{\ell}(X_i-Z_i)\Big|\,.$$

In order to bound up $D_{L,1}$ and $D_{L,2}$ we shall use the two lemmas below.

Lemma 6.1. *Assume that $\sum_{k\ge1}k\theta(k)<\infty$. Then there exists a positive constant C such that, for any integer $m(L)$ in $[0,L]$,*

$$\|D_{L,1}\|_2^2\le C2^{L-m(L)}\big(1+\sum_{k\ge1}k(k\wedge 2^{m(L)/2})\theta(k)\big)\,. \tag{6.3}$$

Lemma 6.2. *Assume that $\sum_{k\ge1}k\theta(k)<\infty$. Then there exist positive constants c and C such that, for any positive λ and any integer $m(L)$ in $[0,L]$,*

$$\mathbb{P}(D_{L,2}\ge 2\lambda)\le C2^L\exp\big(-2^{-m(L)}\lambda^2/c\big)+C2^L\lambda^{-4}\big(1+\sum_{k\ge1}k(k\wedge\lambda)\theta(k)\big)\,. \tag{6.4}$$

Proof of Lemma 6.1. For any $\ell\in\{1,\cdots,2^{L-m(L)}\}$, let $\widetilde{U}_{\ell,L}=U_{\ell,L}-\mathbb{E}_{2^L+(\ell-1)2^{m(L)}}(U_{\ell,L})$. Then $(\widetilde{U}_{\ell,L})_{\ell\ge1}$ is a strictly stationary sequence of martingale differences adapted to the filtration $(\mathcal{F}_{2^L+\ell 2^{m(L)}})_{\ell\ge1}$. Notice first that

$$\|D_{L,1}\|_2\le\Big\|\sup_{k\le 2^{L-m(L)}}\Big|\sum_{\ell=1}^k(\widetilde{U}_{\ell,L}-V_{\ell,L})\Big|\Big\|_2+\Big\|\sup_{k\le 2^{L-m(L)}}\Big|\sum_{\ell=1}^k(\widetilde{U}_{\ell,L}-U_{\ell,L})\Big|\Big\|_2\,. \tag{6.5}$$

Let us deal with the first term on right hand. Proceeding as in the proof of Lemma 4.12 in [23], we have

$$\Big\|\sup_{k\le 2^{L-m(L)}}\Big|\sum_{\ell=1}^k(\widetilde{U}_{\ell,L}-V_{\ell,L})\Big|\Big\|_2^2\le 16\sum_{\ell=1}^{2^{L-m(L)}}\|U_{\ell,L}-V_{\ell,L}\|_2^2\,.$$

By using Theorem 2.1(b) in [14], it follows that

$$\Big\| \sup_{k\le 2^{L-m(L)}} \Big| \sum_{\ell=1}^{k} (\widetilde{U}_{\ell,L} - V_{\ell,L}) \Big| \Big\|_2^2 \ll 2^{L-m(L)} \Big(1 + \sum_{k\ge 1} k(k \wedge 2^{m(L)/2}) \theta(k) \Big) \,. \tag{6.6}$$

We deal now with the second term in the right hand side of (6.5). Using Proposition 1 in [15], we obtain that

$$\begin{aligned} \Big\| \sup_{k\le 2^{L-m(L)}} \Big| \sum_{\ell=1}^{k} (\widetilde{U}_{\ell,L} - U_{\ell,L}) \Big| \Big\|_2^2 &\le 4 \sum_{k=1}^{2^{L-m(L)}} \|\mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{k,L})\|_2^2 \\ &+ 8 \sum_{k=1}^{2^{L-m(L)}-1} \Big\| \mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{k,L}) \Big(\sum_{i=k+1}^{2^{L-m(L)}} \mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{i,L}) \Big) \Big\|_1 \,. \end{aligned} \tag{6.7}$$

Stationarity leads to

$$\|\mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{k,L})\|_2^2 = \|\mathbb{E}_0(S_{2^{m(L)}})\|_2^2 \le 2 \sum_{i=1}^{2^{m(L)}} \sum_{j=1}^{i} \mathbb{E}|X_j \mathbb{E}_0(X_i)| \le 2 \sum_{i=1}^{2^{m(L)}} i\theta(i) \ll 1 \,.$$

Consequently,

$$\sum_{k=1}^{2^{L-m(L)}} \|\mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{k,L})\|_2^2 \ll 2^{L-m(L)} \,. \tag{6.8}$$

We now bound up the second term in the right hand side of (6.7). Stationarity yields

$$\Big\| \mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{k,L}) \Big(\sum_{i=k+1}^{2^{L-m(L)}} \mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{i,L}) \Big) \Big\|_1 \le \sum_{j=1}^{2^{m(L)}} \sum_{i=2^{m(L)}+1}^{2^L-(k-1)2^{m(L)}} \mathbb{E}|X_j \mathbb{E}_0(X_i)| \,.$$

Therefore

$$\sum_{k=1}^{2^{L-m(L)}-1} \Big\| \mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{k,L}) \Big(\sum_{i=k+1}^{2^{L-m(L)}} \mathbb{E}_{2^L+(k-1)2^{m(L)}}(U_{i,L}) \Big) \Big\|_1 \ll 2^{L-m(L)} \sum_{i=2^{m(L)}+1}^{2^L} i\theta(i) \,. \tag{6.9}$$

Starting from (6.7) and considering the bounds (6.8) and (6.9), we get that

$$\Big\| \sup_{k\le 2^{L-m(L)}} \Big| \sum_{\ell=1}^{k} (\widetilde{U}_{\ell,L} - U_{\ell,L}) \Big| \Big\|_2^2 \ll 2^{L-m(L)} \,. \tag{6.10}$$

Starting from (6.5) and considering the bounds (6.6) and (6.10), we then get (6.3), which ends the proof of Lemma 6.1. □

Proof of Lemma 6.2. It follows the lines of the proof of [23, Lemma 4.1] with the difference that Theorem 2.1 is used instead of [23, Proposition A.2]. □

End of the proof of Theorem 3.1. Let us start by proving Item (i). In case $p \in]2,3[$ with the condition $\theta(k) = O(k^{1-p})$ and in case $p = 3$ with the condition $\sum_{k>0} k\theta(k) < \infty$

the result can be proved exactly as in [23] (indeed, in the bounded case, these authors could have used the coefficient $\theta(k)$ instead of their coefficient $\alpha_{2,\mathbf{X}}(k)$). Now, we turn to the case $p \in]3,4]$. Note that under the condition $\theta(k) = O(k^{1-p})$ if $p \in]3,4[$ and $\sum_{k>0} k^2\theta(k) < \infty$ if $p = 4$,

$$\sum_{k\geq 1} k(k \wedge \lambda)\theta(k) \leq C(1 + \lambda^{4-p})\,. \tag{6.11}$$

Therefore, by Lemma 6.2, simple computations lead to

$$\mathbb{E}(D_{L,2}^2) = 2\int_0^\infty x\mathbb{P}(D_{L,2} \geq x)dx \leq C\big(L2^{m(L)} + 2^{2L/p}\big)\,, \tag{6.12}$$

and, by Lemma 6.1,

$$\mathbb{E}(D_{L,1}^2) \leq C2^L 2^{(1-p/2)m(L)}\,. \tag{6.13}$$

Choosing

$$m(L) = [2(L - \log_2 L)/p] \text{ so that } 2^{-1+2L/p}L^{-2/p} \leq 2^{m(L)} \leq 2^{2L/p}L^{-2/p}\,, \tag{6.14}$$

for the construction of the Gaussian sequence, Item (i) follows (above square brackets designate as usual the integer part and $\log_2(x) = (\log x)/(\log 2)$).

We turn now to Item (ii). Let us complete the proof when $p = 4$, meaning that $\theta(k) = O(k^{-3})$. In this case,

$$\sum_{k\geq 1} k(k \wedge \lambda)\theta(k) \leq C(1 + \log \lambda)\,, \tag{6.15}$$

for any $\lambda \geq 1$, and then the term $2^{2L/p}$ appearing in (6.12) has to be replaced by $(1 + m(L))2^{m(L)}$ and the right-hand side of (6.13) will be $CL2^{L-m(L)}$. Choosing $m(L) = [L/2]$ completes the proof. Finally, the case $p = 3$ with the condition $\theta(k) = O(k^{-2})$ can be handled similarly by taking into account Lemmas 4.2 and 4.1 in [23] instead of our Lemmas 6.1 and 6.2.

End of the proof of Theorem 3.2. In case $p \in]2,3[$ and $\sum_{k>0} k^{p-2}\theta(k) < \infty$ the result can be proved exactly as in Theorem 2.1 Item 1a) in [23] (using the coefficient $\theta(k)$ instead of their coefficient $\alpha_{2,\mathbf{X}}(k)$). Similarly, in case $p \in]2,3[$ and $\theta(k) = O(k^{1-p})$ or $p = 3$ and $\sum_{k>0} k\theta(k) < \infty$, the result can be proved exactly as in Theorem 2.1 Item 1b) in [23] (with $\theta(k)$ instead of $\alpha_{2,\mathbf{X}}(k)$).

Let us now complete the proof of Item (i) when $p \in]3,4[$ and $\sum_{k>0} k^{p-2}\theta(k) < \infty$. In this case we select $m(L)$ as in (6.14) and set

$$\lambda_L = \kappa 2^{m(L)/2}\sqrt{L}\,, \tag{6.16}$$

with $\kappa = \sqrt{2c\log 2}$ where c is the positive constant of Lemma 6.2. For this choice,

$$\sum_{L>0} 2^L \exp\big(-\lambda_L^2 2^{-m(L)}/c\big) = \sum_{L\geq 0} 2^{L-2L} < \infty \text{ and } \sum_{L>0} 2^L\lambda_L^{-4} < \infty\,. \tag{6.17}$$

In addition, using the fact that $\sum_{k\geq 1} k(k\wedge(a\lambda))\theta(k) \leq a\sum_{k\geq 1} k(k\wedge\lambda)\theta(k)$, for any $a\geq 1$ and any positive λ, we get, since $\sum_{k>0} k^{p-2}\theta(k)<\infty$,

$$\sum_{L>0} 2^L\lambda_L^{-4}\sum_{k\geq 1} k(k\wedge\lambda_L)\theta(k) \leq \sum_{L>0} 2^L L^{1/2}\lambda_L^{-4}\sum_{k\geq 1} k(k\wedge 2^{m(L)/2})\theta(k)<\infty\,.$$

Therefore Lemma 6.2 entails that $\sum_{L>0}\mathbb{P}(D_{L,2}\geq 2\lambda_L)<\infty$ implying, via the Borel-Cantelli lemma, that

$$D_{L,2} = O(2^{L/p}L^{1/2-1/p}) \text{ a.s.} \tag{6.18}$$

On another hand, from (6.3) together with the Markov inequality,

$$\sum_{L>0}\mathbb{P}(D_{L,1}\geq\lambda_L) \leq C\sum_{L>0} 2^L L^{-1}2^{-2m(L)}\sum_{k\geq 1} k(k\wedge 2^{m(L)/2})\theta(k)<\infty\,,$$

since $\sum_{k>0} k^{p-2}\theta(k)<\infty$. Hence, by the Borel-Cantelli lemma,

$$D_{L,1} = O(2^{L/p}L^{1/2-1/p}) \text{ a.s.} \tag{6.19}$$

Finally Item (i) when $p\in]3,4[$ follows from (6.1), (6.2), (6.18) and (6.19).

We complete now the proof of Item (ii) when $p\in]3,4[$ and $\theta(k)=O(k^{1-p})$ or when $p=4$ and $\sum_{k>0} k^2\theta(k)<\infty$. In these cases we select $m(L)$ as follows: let $\varepsilon>0$ and set

$$m(L) = [2(L+\varepsilon\log_2 L)/p] \text{ so that } 2^{-1+2L/p}L^{2\varepsilon/p} \leq 2^{m(L)} \leq 2^{2L/p}L^{2\varepsilon/p}\,. \tag{6.20}$$

We still define λ_L by (6.16). For this choice of λ_L, the convergences in (6.17) still hold. In addition, taking into account (6.11), for any $p\in]3,4]$, under the conditions on $(\theta(k))_{k>0}$, we get

$$\sum_{L>0} 2^L\lambda_L^{-4}\sum_{k\geq 1} k(k\wedge\lambda_L)\theta(k) \leq C\sum_{L>0} 2^L\lambda_L^{-p}<\infty\,. \tag{6.21}$$

Therefore Lemma 6.2 entails that $\sum_{L>0}\mathbb{P}(D_{L,2}\geq 2\lambda_L)<\infty$ implying, via the Borel-Cantelli lemma, that

$$D_{L,2} = O(\lambda_L) = O(2^{L/p}L^{1/2+\varepsilon/p}) \text{ a.s.} \tag{6.22}$$

On another hand, from (6.3) and (6.11) together with the Markov inequality,

$$\sum_{L>0}\mathbb{P}(D_{L,1}\geq\lambda_L) \leq C\sum_{L>0}\big(2^L L^{-1}2^{-2m(L)}\big)2^{m(L)(2-p/2)}<\infty\,. \tag{6.23}$$

Hence, by the Borel-Cantelli lemma,

$$D_{L,1} = O(\lambda_L) = O(2^{L/p}L^{1/2+\varepsilon/p}) \text{ a.s.} \tag{6.24}$$

Finally Item (ii) when $p\in]3,4]$ follows from (6.1), (6.2), (6.22) and (6.24).

We turn now to the proof of Item (iii). When $p = 4$, meaning that $\theta(k) = O(k^{-3})$, instead of (6.11), we use (6.15). We still select λ_L by (6.16) but with the following choice of $m(L)$: for $\varepsilon > 0$

$$m(L) = [2(L + (1+\varepsilon)\log_2 L)/p] \text{ so that } 2^{-1+2L/p} L^{2(1+\varepsilon)/p} \le 2^{m(L)} \le 2^{2L/p} L^{2(1+\varepsilon)/p} . \tag{6.25}$$

The computations (6.21) and (6.23) are then replaced by the following ones:

$$\sum_{L>0} \frac{2^L}{\lambda_L^4} \sum_{k>0} (k^2 \wedge k\lambda_L)\theta(k) \le C \sum_{L>0} \frac{2^L}{\lambda_L^4} \log \lambda_L \,, \ \sum_{L>0} \mathbb{P}(D_{L,1} \ge \lambda_L) \le C \sum_{L>0} \frac{L 2^L}{2^{m(L)} \lambda_L^2} \,.$$

Note that the above upper bounds are finite for these selections of $m(L)$ and λ_L. The rest of the proof is unchanged compared to the previous cases.

It remains to prove Item (iii) when $p = 3$, meaning that $\theta(k) = O(k^{-2})$. The differences with the case $p = 4$ are that [23, Lemma 4.1] is used instead of our Lemma 6.1 and $\|D_{L,1}\|_2^2 \le C 2^{L-m(L)/2} m(L)$. This upper bound on $\|D_{L,1}\|_2^2$ comes from a slight modification of the proof of [23, Lemma 4.2], taking into account that $\sum_{k\ge1}(k \wedge \lambda)\theta(k) \le C(1 + \log \lambda)$ for any $\lambda \ge 1$. In addition, $m(L)$ is selected by (6.25) with $p = 3$. □

7 The degenerate case

In all this section, we shall denote $\theta_{X,1,1}(k)$ by $\theta(k)$ for all $k \in \mathbb{N}$.

Proposition 7.1. *Assume that $\|X_0\|_\infty \le M < \infty$ and $\sum_{k\ge0} \theta(k) < \infty$. Suppose in addition that $\sigma^2 = 0$. Then, for any $q \ge 1$,*

$$\mathbb{E}(|S_n|^q) \le q(2M)^q \sum_{k\ge0} (k+1)^{q-1} \theta(k) \,.$$

Theorem 7.1. *Let $p > 2$ and $S_n^* = \max_{1\le k\le n} |S_k|$. Assume that $\|X_0\|_\infty < \infty$ and $\theta(k) = O(k^{1-p})$. Suppose in addition that $\sigma^2 = 0$. Then, for any $r \in [1, p[$, $\|S_n^*\|_r = O(n^{1/p})$.*

Remark 7.1. Since for any increasing sequence $(b_n)_{n\ge1}$, $\sum_{n>0} n^{-1} \mathbb{P}(S_n^* > b_n) < \infty$ implies that $S_n = o(b_n)$ almost surely, it follows that, under the assumptions of the theorem, $S_n = o(n^{1/p}(\log n)^{\varepsilon+1/p})$ almost surely for any $\varepsilon > 0$.

Theorem 7.2. *Let $p \ge 2$ and $S_n^* = \max_{1\le k\le n} |S_k|$. Assume that $\|X_0\|_\infty < \infty$ and $\sum_{k\ge0}(k+1)^{p-2}\theta(k) < \infty$. Suppose in addition that $\sigma^2 = 0$. Then, for any $\alpha \in]0, 1[$ and any $\varepsilon > 0$, $\sum_{n>0} n^{\alpha p-2} \mathbb{P}(S_n^* > \varepsilon n^\alpha) < \infty$. Consequently $S_n = o(n^{1/p})$ almost surely.*

Proof of Proposition 7.1 . We start the proof with the following lemma.

Lemma 7.1. *Assume that* $X_0 \in \mathbb{L}^2$,

(a) $\mathbb{E}_0(S_n)$ *converges in* $\mathbb{L}^1$,

(b) $\lim_{n\to\infty} \mathbb{E}(X_0(X_0 + 2S_n)) = 0$,

(c) $\limsup_{N\to\infty} \limsup_{n\to\infty} \mathbb{E}(\mathbb{E}_{-N}(X_0)(X_0 + 2S_n)) \le 0$.

Then, for any integer i, $X_i = g_{i-1} - g_i$ *almost surely where* $g_i = \sum_{k\ge i+1} \mathbb{E}_i(X_k)$.

Proof of Lemma 7.1. Let N be a fixed positive integer. Set $d_{i,N} = \sum_{k=0}^{N-1} P_i(X_{k+i})$ where $P_i = \mathbb{E}_i - \mathbb{E}_{i-1}$. Define also $g_{i,N} = \sum_{k=1}^{N} \mathbb{E}_i(X_{k+i})$ and $Y_{i,N} = \mathbb{E}_i(X_{N+i})$. With these notations, the following decomposition is valid:

$$X_i = d_{i,N} + g_{i-1,N} - g_{i,N} + Y_{i,N}\,. \tag{7.1}$$

Since N is fixed, all the random variables in the above decomposition are in $\mathbb{L}^2$. Moreover, from the fact that $(d_{i,N})_{i\in\mathbb{Z}}$ is a stationary sequence of martingale differences,

$$\|d_{0,N}\|_1 \le \|d_{0,N}\|_2 \le \limsup_{n\to\infty} n^{-1/2} \Big\| \sum_{i=1}^{n} d_{i,N} \Big\|_2 . \tag{7.2}$$

Next, from (7.1), $\sum_{i=1}^{n} d_{i,N} = S_n + g_{n,N} - g_{0,N} - \sum_{i=1}^{n} Y_{i,N}$, which implies that

$$n^{-1/2} \Big\| \sum_{i=1}^{n} d_{i,N} \Big\|_2 \le n^{-1/2} \|S_n\|_2 + 2n^{-1/2} \|g_{0,N}\|_2 + n^{-1/2} \Big\| \sum_{i=1}^{n} Y_{i,N} \Big\|_2 .$$

Now, by item (b), $\lim_{n\to\infty} n^{-1/2} \|S_n\|_2 = 0$. Hence

$$\limsup_{n\to\infty} n^{-1/2} \Big\| \sum_{i=1}^{n} d_{i,N} \Big\|_2 \le \limsup_{n\to\infty} n^{-1/2} \Big\| \sum_{i=1}^{n} Y_{i,N} \Big\|_2 . \tag{7.3}$$

Next, by stationarity and the properties of the conditional expectation, we infer that

$$\Big\| \sum_{i=1}^{n} Y_{i,N} \Big\|_2^2 = \sum_{k=0}^{n-1} \mathbb{E}\big(\mathbb{E}_{-N}(X_0)(X_0 + 2S_k)\big)\,.$$

Now Item (c) combined with the Cesaro Lemma entails that

$$\lim_{N\to\infty} \limsup_{n\to\infty} n^{-1} \Big\| \sum_{i=1}^{n} Y_{i,N} \Big\|_2^2 = 0\,. \tag{7.4}$$

Taking into account (7.2)-(7.4), we derive that $d_{i,N}$ converges to 0 in $\mathbb{L}^1$ as N tends to ∞. Therefore the lemma follows by taking into account the decomposition (7.1) and noting that, by item (a), $g_{i,N}$ converges to g_i in $\mathbb{L}^1$ and $Y_{i,N}$ converges to 0 in $\mathbb{L}^1$. □

We now prove Proposition 7.1. We start by noticing that $\sum_{k\geq 1}\theta(k)<\infty$ and $\sigma^2=0$ imply items (a), (b) and (c) of Lemma 7.1. More precisely, to prove item (c), starting from Inequality (4.7) in [26], we obtain that $\big|\mathbb{E}(\mathbb{E}_{-N}(X_0)X_k)\big|\leq M\theta(k+N)$, implying that $\big|\mathbb{E}(\mathbb{E}_{-N}(X_0)(X_0+2S_n))\big|\leq 2M\sum_{\ell\geq N}\theta(\ell)$. So, overall, under the conditions of Theorem 7.1,

$$S_n=g_0-g_n \text{ in } \mathbb{L}^1 \text{ where } g_m=\sum_{k>0}\mathbb{E}(X_{k+m}|\mathcal{F}_m)\,. \tag{7.5}$$

Now, let $q\geq 1$ and r in $]1,+\infty]$ be the conjugate exponent, that is $1/q+1/r=1$. Note first that, by the Riesz-Fisher theorem, $\|g_0\|_q=\sup\big\{\mathbb{E}(g_0Z)\,:\,Z\in\mathcal{L}_r\big\}$, where $\mathcal{L}_r$ is the class of nonnegative random variables Z such that $\|Z\|_r=1$. But, by [8, Proposition 1],

$$\big|\mathbb{E}(g_0Z)\big|\leq\sum_{k>0}|\mathbb{E}\big(ZX_k\big)|\leq M\sum_{k>0}\int_0^{\theta(k)}Q_Z(u)du\leq M\int_0^1\theta^{-1}(u)Q_Z(u)du\,,$$

where Q_Z is the generalized inverse of $t\mapsto\mathbb{P}(Z>t)$ and $\theta^{-1}(u)=\sum_{i\geq 0}\mathbf{1}_{u<\theta(i)}$. Therefore, by Hölder's inequality,

$$\big|\mathbb{E}(g_0Z)\big|\leq M\Big(\int_0^1\big(\theta^{-1}(u)\big)^q du\Big)^{1/q}.$$

Hence, using the coboundary decomposition (7.5) and the stationarity of $(g_m)_{m\in\mathbb{Z}}$,

$$\|S_n\|_q\leq 2\|g_0\|_q\leq 2M\Big(\int_0^1\big(\theta^{-1}(u)\big)^q du\Big)^{1/q}.$$

Now, by inequality (C.5), page 184, in [26], $\int_0^1\big(\theta^{-1}(u)\big)^q du\leq q\sum_{k\geq 0}(k+1)^{q-1}\theta(k)$. □

Proof of Theorem 7.1. Let $q=\max(1,r(p-1)/p)$. We apply inequality (5.1) with $\varphi(x)=x^q$. Since $\sum_{k>0}k^{q-1}\theta(k)<\infty$, Proposition 7.1 ensures that $\|S_n\|_q\ll 1$. Therefore, for any $x>0$,

$$\mathbb{P}\big(S_n^*\geq x\big)\ll x^{-q}+\min(1,nx^{-p})\,. \tag{7.6}$$

Next, since $S_n^*\leq nM$, $\|S_n^*\|_r^r=r\int_0^{nM}x^{r-1}\mathbb{P}\big(S_n^*\geq x\big)dx$. Hence, applying (7.6) and taking into account the selection of q, we get

$$\|S_n^*\|_r^r\ll\int_0^{nM}x^{r-1-q}dx+\int_0^{n^{1/p}}x^{r-1}dx+n\int_{n^{1/p}}^{\infty}x^{r-1-p}dx\ll n^{r/p}\,.\quad\square$$

Proof of Theorem 7.2 . We start from inequality (5.1) applied with $\varphi(x)=x^{p-1}$. By Proposition 7.1, $\mathbb{E}(\varphi(S_n))\ll 1$. Therefore

$$\mathbb{P}\big(S_n^*\geq 4\varepsilon n^\alpha\big)\ll(\varepsilon n^\alpha)^{1-p}+n^{1-\alpha}\varepsilon^{-1}\theta([\varepsilon n^\alpha])\,.$$

Hence

$$\sum_{n>0}n^{\alpha p-2}\mathbb{P}\big(S_n^*\geq 4\varepsilon n^\alpha\big)\ll\varepsilon^{1-p}\sum_{n>0}n^{\alpha-2}+\varepsilon^{-1}\sum_{n\geq 1}n^{\alpha(p-1)-1}\theta([\varepsilon n^\alpha])\,.$$

The first series converges if $\alpha<1$ and the second one also converges as soon as $\sum_{k>0}k^{p-2}\theta(k)<\infty$. This ends the proof of the theorem. □

References

[1] Bártfai, P. Die Bestimmung der zu einem wiederkehrenden Prozess gehörenden Verteilungsfunktion aus den mit Fehlern behafteten Daten einer einzigen Realisation. *Studia Sci. Math. Hungar.* 1 (1966), 161–168.

[2] Berkes, I., Liu, W. and Wu, W.B. Komlós-Major-Tusnády approximation under dependence. *Ann. Probab.* **42** (2014), no. 2, 794–817.

[3] Bolthausen, E. The Berry-Esseen theorem for strongly mixing Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **60** (1982), 283–289.

[4] Breiman, L. On the tail behavior of sums of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** (1967), 20–25.

[5] Csáki, E. and Csörgő, M. On additive functionals of Markov chains. *J. Theoret. Probab.* **8** (1995), no. 4, 905–919.

[6] Cuny, C., Dedecker, J., Korepanov, A. and Merlevède, F. Rates in almost sure invariance principle for slowly mixing dynamical systems. *Ergodic Theory Dynam. Systems* **40** (2020), no. 9, 2317–2348.

[7] Cuny, C., Dedecker, J. and Merlevède, F. Deviation and concentration inequalities for dynamical systems with subexponential decay of correlations. *Stoch. Dyn.* **23** (2023), no. 3, Paper No. 2350025, 18 pp.

[8] Dedecker, J. and Doukhan, P. A new covariance inequality and applications. *Stochastic Process. Appl.* **1**06 (2003), no. 1, 63–80.

[9] Dedecker, J., Doukhan, P. and Merlevède, F. Rates of convergence in the strong invariance principle under projective criteria. *Electron. J. Probab.* **17** (2012), no. 16, 31 pp.

[10] Dedecker, J., Gouëzel, S. and Merlevède, F. Some almost sure results for unbounded functions of intermittent maps and their associated Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** (2010), no. 3, 796–821.

[11] Dedecker, J. and Merlevède, F. Moment bounds for dependent sequences in smooth Banach spaces. *Stochastic Process. Appl.* **1**25 (2015), no. 9, 3401-3429.

[12] Dedecker, J., Merlevède, F. and Rio, E. Rates of convergence for minimal distances in the central limit theorem under projective criteria. *Electron. J. Probab.* **14** (2009), no. 35, 978–1011.

[13] Dedecker, J., Merlevède, F. and Rio, E. Strong approximation of the empirical distribution function for absolutely regular sequences in $\mathbb{R}^d$. *Electron. J. Probab.* **19** (2014), no. 9, 56 pp.

[14] Dedecker, J., Merlevède, F. and Rio, E. Quadratic transportation cost in the conditional central limit theorem for dependent sequences. 2022. *hal-03890107*. To appear in *Annales Henri Lebesgue.*

[15] Dedecker, J. and Rio, E. On the functional central limit theorem for stationary processes. *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000), no. 1, 1–34.

[16] Douc, R., Moulines, E., Priouret, P. and Soulier, P. *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2018. xviii+757 pp.

[17] Kendall, W.S. and Montana, G. Small sets and Markov transition densities. *Stochastic Process. Appl.* 99 (2002), no. 2, 177–194.

[18] Komlós, J., Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent RV's, and the sample DF. II. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34** (1976), 33–58.

[19] Liu, Z. and Wang, Z. Wasserstein convergence rate in the invariance principle for deterministic dynamical systems. *arXiv:2204.00263v2.*

[20] Liverani, C., Saussol, B. and Vaienti, S. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** (1999), no. 3, 671–685.

[21] Major, P. The approximation of partial sums of independent RV's. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **35** (1976), 213–220.

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

[23] Merlevède, F. and Rio, E. Strong approximation of partial sums under dependence conditions with application to dynamical systems. *Stochastic Process. Appl.* **122** (2012), no. 1, 386–417.

[24] Merlevède, F. and Rio, E. Strong approximation for additive functionals of geometrically ergodic Markov chains. *Electron. J. Probab.* **20** (2015), no. 14, 27 pp.

[25] Nummelin, E. (1984). *General irreducible Markov chains and nonnegative operators*. Cambridge Tracts in Mathematics, **83**. Cambridge University Press, Cambridge.

[26] Rio, E. *Asymptotic theory of weakly dependent random processes*. Translated from the 2000 French edition. Probability Theory and Stochastic Modelling, 80. Springer, Berlin, 2017.

[27] Skorohod, A. V. On a representation of random variables. *Teor. Verojatnost. i Primenen.* **21** (1976), 645-648.

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

[29] Young, L.-S. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999), 153-188.