Accepted Manuscript

On the Komlós, Major and Tusnády strong approximation for some classes of random iterates

Christophe Cuny, Jérôme Dedecker, Florence Merlevède

Please cite this article as: C. Cuny, J. Dedecker, F. Merlevède, On the Komlós, Major and Tusnády strong approximation for some classes of random iterates, *Stochastic Processes and their Applications* (2017), http://dx.doi.org/10.1016/j.spa.2017.07.011

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

On the Komlós, Major and Tusnády strong approximation for some classes of random iterates

Christophe Cuny,[*] Jérôme Dedecker[†] and Florence Merlevède[‡]

March 16, 2017

Abstract

The famous results of Komlós, Major and Tusnády (see [15] and [17]) state that it is possible to approximate almost surely the partial sums of size n of i.i.d. centered random variables in $\mathbb{L}^p$ ($p > 2$) by a Wiener process with an error term of order $o(n^{1/p})$. Very recently, Berkes, Liu and Wu [3] extended this famous result to partial sums associated with functions of an i.i.d. sequence, provided a condition on a functional dependence measure in $\mathbb{L}_p$ is satisfied. In this paper, we adapt the method of Berkes, Liu and Wu to partial sums of functions of random iterates. Taking advantage of the Markovian setting, we shall give new dependent conditions, expressed in terms of a natural coupling (in $\mathbb{L}^\infty$ or in $\mathbb{L}^1$), under which the strong approximation result holds with rate $o(n^{1/p})$. As we shall see our conditions are well adapted to a large variety of models, including left random walks on $GL_d(\mathbb{R})$, contracting iterated random functions, autoregressive Lipschitz processes, and some ergodic Markov chains. We also provide some examples showing that our $\mathbb{L}^1$-coupling condition is in some sense optimal.

1 Introduction

In this paper we shall adapt the approach of Berkes-Liu-Wu [3] to certain classes of Markov chains. To motivate this work, let us describe in detail the example of the left random walk on $GL_d(\mathbb{R})$, $d \geq 2$ (the group of invertible d-dimensional real matrices).

Let $(\varepsilon_n)_{n \geq 1}$ be independent random matrices taking values in $G = GL_d(\mathbb{R})$, with common distribution μ. Let $\| \cdot \|$ be the euclidean norm on $\mathbb{R}^d$. We shall say that μ has a moment of order $p \geq 1$ if

$$\int_G (\log N(g))^p \mu(dg) < \infty, \tag{1}$$

where $N(g) := \max(\|g\|, \|g^{-1}\|)$.

Let $A_0 = \mathrm{Id}$ and for every $n \geq 1$, $A_n = \varepsilon_n \cdots \varepsilon_1$. Recall that if μ admits a moment of order 1 then

$$\lim_{n \to \infty} \frac{1}{n} \log \|A_n\| = \lambda_\mu \quad \mathbb{P}\text{-a.s.,} \tag{2}$$

---

[*]Université de la Nouvelle-Calédonie, Equipe ERIM. Email: christophe.cuny@univ-nc.nc

[†]Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5 (UMR 8145). Email: jerome.dedecker@parisdescartes.fr

[‡]Université Paris-Est, LAMA (UMR 8050), UPEM, CNRS, UPEC. Email: florence.merlevede@u-pem.fr

1

where $\lambda_\mu := \lim_{n \to +\infty} n^{-1} \mathbb{E}(\log \|Y_n \cdots Y_1\|)$ is the so-called first Lyapunov exponent (see for instance [14]). For any $x \in S^{d-1}$, we want to describe as precisely as possible the asymptotic behavior of the quantity

$$\log \|A_n x\| \,. \tag{3}$$

The left random walk of law μ started at $x \in S^{d-1}$ is the Markov chain defined by $W_{0,x} := x$ and $W_{n,x} = \varepsilon_n W_{n-1,x}$ for $n \geq 1$. As usual, to handle the quantity (3), we consider the partial sums associated with the random variables $(X_{n,x})_{n \geq 1}$ given by

$$X_{n,x} := h(\varepsilon_n, W_{n-1,x}) \,, \ n \geq 1 \,,$$

where for every $g \in G$ and every $y \in \mathbb{R}^d - \{0\}$,

$$h(g,y) = \log \left(\frac{\|g \cdot y\|}{\|y\|} \right) \,.$$

By definition of h, and since $X_n = h(\varepsilon_n, A_{n-1}x)$, we easily see that, for any $x \in S^{d-1}$,

$$S_{n,x} = \sum_{k=1}^n X_{k,x} = \log \|A_n x\| \,. \tag{4}$$

Hence, the asymptotic behavior of (3) can be deduced from the asymptotic behavior of partial sums of functions of the Markov chain $W_{n,x}$.

This problem can be tackled under some assumptions on μ (strong irreducibility and proximality, see subsection 3.1 for more details) which implies that the chain $(W_n)_{n \geq 0}$ admits an unique invariant measure ν defined on the projective space $X := P_{d-1}(\mathbb{R}^d)$ of $\mathbb{R}^d - \{0\}$. Under these assumptions on μ, and assuming moreover that μ has a moment of order $p \in (2,4)$, Cuny-Dedecker-Jan [7] proved the following strong approximation result: there exists $\sigma^2 \geq 0$ such that, for every (fixed) $x \in S^{d-1}$, one can redefine $(\log \|A_n x\|)_{n \geq 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i \geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,

$$\log \|A_n x\| - n\lambda_\mu - \sum_{i=1}^n N_i = o(n^{1/p}\sqrt{\log n}) \ \text{a.s.} \tag{5}$$

If μ has a moment of order $p = 4$, the same authors showed that this strong approximation holds with a rate of order $O(n^{1/4}\sqrt{\log(n)}(\log \log n)^{1/4})$.

To prove (5), Cuny-Dedecker-Jan used a martingale approximation (as described for instance in Cuny-Merlevède [9]), together with some appropriate upper bounds on the quantities

$$\sup_{\|x\|=1, \|y\|=1} \mathbb{E}\left(|X_{k,x} - X_{k,y}|\right) \,. \tag{6}$$

The main drawback of this approach is that it cannot give a better rate than $n^{1/4}$, because it is based on the Skorokhod representation theorem for martingales.

On another hand, since the stationary Markov chain W_n is a function of the starting point W_0 and of the "innovations" $\varepsilon_1, \cdots, \varepsilon_n$, one can also apply the approximation results by Berkes-Liu-Wu (in fact, this is not completely immediate because it does not fit exactly into the framework described by these authors, and some extra work is required there). Doing so, one can reach a

2

rate of order $n^{1/p}$ for any $p > 2$, but only by assuming that μ has a moment of order $q(p) > p$. More precisely, their functional measure of dependence in $\mathbb{L}_p$, say $\delta_{k,p}$, can be bounded by $\sup_{\|x\|=1,\|y\|=1} \|X_{k,x} - X_{k,y}\|_p$. Hence, applying Proposition 3 in [7], one can see that condition (2.3) in [3] is satisfied provided μ has at least a moment of order $(5p/2) - 1$. This is somewhat surprising: on the one hand, one can go beyond the rate of order $n^{1/4}$, and on the other hand we need stronger assumptions than in Cuny-Dedecker-Jan [7] to get the rate $n^{1/p}$ when $p \in (2, 4)$.

This gave us a strong motivation to understand completely the proof by Berkes-Liu-Wu [3], and to see whether it is possible to take advantage of the Markovian setting to get the rate $n^{1/p}$ in (5) under a moment of order p, for any $p > 2$. As we shall see in this paper, the answer is positive.

As already mentioned, in the case of the left random walk on $GL_d(\mathbb{R})$, one can get a control on the quantities defined in (6). However, in many other cases of random iterates, such a control is not possible, while one can get some upper bounds on

$$\iint \mathbb{E}\left(|X_{k,x} - X_{k,y}|\right) \nu(dx)\nu(dy), \tag{7}$$

where ν is the invariant distribution of the chain $(W_n)_{n \geq 1}$.

Consequently, we shall establish two distinct results, with different range of applicability. In Theorem 1, we give a strong approximation result under conditions involving some quantities similar to (6). In Theorem 2 the conditions are expressed in terms of the quantities (7). The second Theorem applies to a large variety of examples, including some well known examples of irreducible and aperiodic Markov Chains with countable or continuous state space. These examples of ergodic Markov chains will allow us to prove that the conditions given in Theorem 2 are in some sense optimal.

In all the paper, we shall use the notation $a_n \ll b_n$, which means that there exists a positive constant C not depending on n such that $a_n \leq Cb_n$, for all positive integers n.

2 Main results

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and let $(\varepsilon_i)_{i \geq 1}$ be iid random variables defined on Ω, with values in a measurable space G and with common distribution μ. Let W_0 be a random variable defined on Ω with values in a measurable space X, independent of $(\varepsilon_i)_{i \geq 1}$, and let F be a measurable function from $G \times X$ to X. For any $n \geq 1$, define

$$W_n = F(\varepsilon_n, W_{n-1}),$$

and assume that $(W_n, n \geq 1)$ has a stationary distribution ν. Let now h be a measurable function from $G \times X$ to $\mathbb{R}$ and define, for any $n \geq 1$,

$$X_n = h(\varepsilon_n, W_{n-1}). \tag{8}$$

Then $(X_n)_{n \geq 1}$ forms a stationary sequence with stationary distribution, say π. Let $(\mathcal{G}_i)_{i \in \mathbb{Z}}$ be the non-decreasing filtration defined as follows: for any $i < 0$, $\mathcal{G}_i = \{\emptyset, \Omega\}$, $\mathcal{G}_0 = \sigma(W_0)$ and for any $i \geq 1$, $\mathcal{G}_i = \sigma(\varepsilon_i, \ldots, \varepsilon_1, W_0)$. It follows that for any $n \geq 1$, X_n is $\mathcal{G}_n$-measurable.

Our first result proves that the strong approximation result holds with rate $n^{1/p}$ when the stationary distribution π has a moment of order $p > 2$ and we impose that the sequence of coupling coefficients $(\delta_\infty(n))_{n \geq 1}$ defined in (10) decreases arithmetically to zero plus the condition

3

(12). As we shall see in Section 3, these conditions are satisfied for instance for the left random walk on $GL_d(\mathbb{R})$.

Let W_0 and W_0^* be random variables with law ν, and such that W_0^* is independent of $(W_0, (\varepsilon_i)_{i \geq 1})$. For any $n \geq 1$, let

$$X_n^* = h(\varepsilon_n, W_{n-1}^*) \text{ with } W_n^* = F(\varepsilon_n, W_{n-1}^*). \tag{9}$$

Define then

$$\delta_\infty(n) = \|\mathbb{E}(|X_n - X_n^*| \,|(W_0, W_0^*))\|_\infty \,, n \geq 1 \,, \tag{10}$$

where, above and in all the rest of the paper, the infinite norm is the usual essential supremum norm.

Theorem 1 *Let $(X_n, n \geq 1)$ be the stationary sequence defined by (8) and assume that its stationary distribution π has moment of order $p > 2$. Assume in addition that there exists a positive constant c such that for any $n \geq 1$,*

$$\delta_\infty(n) \leq cn^{-q} \text{ with } q > (p-1)/2 \,, \tag{11}$$

where $(\delta_\infty(n))_{n \geq 1}$ is defined in (10), and that

$$\sup_{n \geq 1} \|\mathbb{E}(X_n^2 | \mathcal{G}_{n-1})\|_\infty \leq c \,. \tag{12}$$

Let $S_n = \sum_{k=1}^n X_k$. Then $n^{-1}\mathbb{E}\big((S_n - n\mathbb{E}(X_1))^2\big) \to \sigma^2$ as $n \to \infty$ and one can redefine $(X_n)_{n \geq 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i \geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,

$$S_n - n\mathbb{E}(X_1) - \sum_{i=1}^n N_i = o(n^{1/p}) \ \mathbb{P}\text{-a.s.} \tag{13}$$

In the rest of this section, we shall give conditions expressed in terms of the quantities $\|X_n - X_n^*\|_1$ for the strong approximation (13) to hold. Before stating the result, we need to introduce some notations:

For any $n \geq 0$, let us define the sequence $(\delta(n))_{n \geq 0}$ as follows

$$\delta(0) = \delta(1) = \mathbb{E}(|X_1|) \text{ and } \delta(n) = 2^{-1} \sup_{k \geq n-1} \|X_k - X_k^*\|_1 \,, n \geq 2 \,.$$

These quantities are finite if π has a moment of order 1.

For any $x \geq 0$, denote by

$$\delta(x) = \delta([x])$$

and, for any $u \in [0, \mathbb{E}(|X_1|)]$, let

$$\delta^{-1}(u) = \inf\{q \in \mathbb{N} \,:\, \delta(q) \leq u\} = \sum_{n \geq 0} \mathbf{1}_{u < \delta(n)} \,.$$

Denote also by Q the quantile function associated with $|X|$ where X is a random variable with law π: it is then the generalized inverse of the tail function $t \mapsto \mathbb{P}(|X| > t) = \pi((-\infty, -t]) + \pi([t, \infty))$. Let H be the function from $[0,1]$ to $\mathbb{R}^+$ defined by $H(x) = \int_0^x Q(u)du$. We shall assume the following condition

$$\sum_{n \geq 1} n^{p-2} \int_0^{\delta(n)} Q^{p-1} \circ H^{-1}(u)du < \infty \,. \tag{14}$$

4

Theorem 2 *Let $(X_n, n \geq 1)$ be a stationary sequence defined by (8) and assume that its stationary distribution π has a moment of order $p > 2$. Assume in addition that condition (14) holds. Let $S_n = \sum_{k=1}^{n} X_k$. Then $n^{-1}\mathbb{E}\big((S_n - n\mathbb{E}(X_1))^2\big) \to \sigma^2$ as $n \to \infty$ and one can redefine $(X_n)_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,*

$$S_n - n\mathbb{E}(X_1) - \sum_{i=1}^{n} N_i = o(n^{1/p}) \ \ \mathbb{P}\text{-}a.s.$$

Remark 3 *If we define*

$$\gamma(x) = H^{-1}(\delta([x])) \ \text{for any } x \geq 0 \ \text{and} \ \gamma^{-1}(u) = \delta^{-1} \circ H(u) \ \text{for any } u \in [0, 1] \,, \qquad (15)$$

then condition (14) can be rewritten as

$$\sum_{n\geq 1} n^{p-2} \int_0^{\gamma(k)} Q^p(u) du < \infty \,, \qquad (16)$$

which also reads as

$$\int_0^1 R^{p-1}(u) Q(u) du < \infty \ \text{ where } R(u) = \gamma^{-1}(u) Q(u) \,, \qquad (17)$$

Remark 4 *Sufficient conditions for (14) to hold in terms of moments (or weak moments) of π can be given by using Lemma 2 in Dedecker and Doukhan [10]. For instance, if*

$$\|X_1\|_r \ \text{for some } r > p, \ \text{and} \ \sum_{n\geq 1} n^{(pr-2r+1)/(r-p)} \delta(n) < \infty \,, \qquad (18)$$

then condition (14) is satisfied. Note that in the case where $\|X_1\|_\infty < \infty$, condition (14) is equivalent to $\sum_{n\geq 1} n^{p-2}\delta(n) < \infty$.

If we define the following meeting time

$$T^* = \inf\{k \in \mathbb{N} \,:\, W_k = W_k^*\} \,, \qquad (19)$$

it follows that, for any $n \geq 2$,

$$\delta(n) \leq \int_0^{\mathbb{P}_{\nu\otimes\nu}(T^* \geq n)} Q(u) du \,.$$

Therefore the following corollary holds.

Corollary 5 *Let $(X_n, n \geq 1)$ be the stationary sequence defined by (8) and assume that its stationary distribution π has a moment of order $p > 2$. Assume in addition that*

$$\sum_{n\geq 0} (n+1)^{p-2} \int_0^{\mathbb{P}_{\nu\otimes\nu}(T^* \geq n)} Q^p(u) du < \infty \,. \qquad (20)$$

Then the conclusions of Theorem 2 hold.

5

According to the computations given in Annex C of Rio [23], if

$$\|X_1\|_r \text{ for some } r > p, \text{ and } \sum_{n \geq 1} n^{(pr-2r+p)/(r-p)} \mathbb{P}_{\nu \otimes \nu}(T^* \geq n) < \infty, \tag{21}$$

then condition (20) is satisfied. In the case where $\|X_1\|_\infty < \infty$, condition (20) is equivalent to

$$\sum_{n \geq 1} n^{p-2} \mathbb{P}_{\nu \otimes \nu}(T^* \geq n) < \infty. \tag{22}$$

Propositions 15 and 18 in Section 3.3 will show that condition (22) is optimal in some sense.

3 Applications

3.1 Left random walk on $GL_d(\mathbb{R})$

As in the introduction, let $(\varepsilon_n)_{n \geq 1}$ be independent random matrices taking values in $G = GL_d(\mathbb{R})$, $d \geq 2$, with common distribution μ. let $A_0 = \text{Id}$ and for every $n \geq 1$, $A_n = \varepsilon_n \cdots \varepsilon_1$.

Let $\| \cdot \|$ be the euclidean norm on $\mathbb{R}^d$. Recall that μ has a moment of order $p \geq 1$ if (1) holds. Recall also that if μ admits a moment of order 1 then (2) holds, and the quantity λ_μ is well defined.

Let $X := P_{d-1}(\mathbb{R}^d)$ be the projective space of $\mathbb{R}^d - \{0\}$ and write $\bar{x}$ as the projection of $x \in \mathbb{R}^d - \{0\}$ to X. We assume that μ is strongly irreducible (i.e. that no proper finite union of subspaces of $\mathbb{R}^d$ are invariant by Γ_μ, the closed semi-group generated by the support of μ) and proximal (i.e. that there exists a matrix in Γ_μ admitting a unique (with multiplicity one) eigenvalue with maximum modulus). Under those assumptions (see e.g. Bougerol-Lacroix [4] or Benoist-Quint [2]) it is well-known that there exists a unique invariant measure ν on $\mathcal{B}(X)$, meaning that for any continuous and bounded function f from X to $\mathbb{R}$,

$$\int_X f(x)\nu(dx) = \int_G \int_X f(g \cdot x)\mu(dg)\nu(dx).$$

The left random walk of law μ is the process defined by $W_0 := \varepsilon_0$ and $W_n = \varepsilon_n W_{n-1}$ for $n \geq 1$ where we assume that ε_0 is independent of $(\varepsilon_n)_{n \geq 1}$. As explained in the introduction, our aim is to study the partial sums associated with the random sequence $(X_n)_{n \geq 1}$ given by

$$X_n := h(\varepsilon_n, W_{n-1}), \, n \geq 1,$$

where for every $g \in G$ and every $\bar{x} \in X$,

$$h(g, \bar{x}) = \log\left(\frac{\|g \cdot x\|}{\|x\|}\right).$$

As usual, we shall denote by $X_{n,\bar{x}}$ the random variable for which $W_0 = \bar{x}$. We then define $S_{n,\bar{x}} = \sum_{k=1}^n X_{n,\bar{x}}$ and recall that the identity (4) holds: for any $x \in S^{d-1}$,

$$S_{n,\bar{x}} = \sum_{k=1}^n X_{k,\bar{x}} = \log\|A_n x\|.$$

Applying Theorem 1, the following strong approximation with rate holds.

6

Corollary 6 *Let μ be a proximal and strongly irreducible probability measure on $\mathcal{B}(G)$. Assume that μ has a moment of order $p > 2$. Then $n^{-1}\mathbb{E}_\nu\big((S_n - n\lambda_\mu)^2\big) \to \sigma^2$ as $n \to \infty$ and for every (fixed) $\bar{x} \in X$, one can redefine $(S_{n,\bar{x}})_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,*

$$S_{n,\bar{x}} - n\lambda_\mu - \sum_{i=1}^{n} N_i = o(n^{1/p}) \, a.s.$$

Remark 7 *It follows from item c) of Theorem 4.11 of Benoist-Quint [2] that $\sigma > 0$ if the image of Γ_μ in $PGL_d(\mathbb{R})$ is unbounded.*

Proof of Corollary 6. Using the same arguments as in Cuny-Dedecker-Jan [7] (see the proof of their Theorem 1), we infer that it suffices to prove the result on stationary regime. More precisely, it suffices to prove that one can redefine $(S_n)_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,

$$S_n - n\lambda_\mu - \sum_{i=1}^{n} N_i = o(n^{1/p}) \ \mathbb{P}_\nu\text{-a.s.} \tag{23}$$

Note also that the fact that $n^{-1}\mathbb{E}_\nu\big((S_n - n\lambda_\mu)^2\big) \to \sigma^2$ as $n \to \infty$ comes from Theorem 2 (ii) in [7]. Now the strong invariance principle (23) is a direct application of Theorem 1. To see this, note first that the following estimate is valid (see Proposition 3 in [7]):

$$\sum_{k\geq 1} k^{p-2} \sup_{\bar{x},\bar{y}\in X} \mathbb{E}\big(\big|X_{k,\bar{x}} - X_{k,\bar{y}}\big|\big) < \infty \,.$$

Since $\big(\sup_{\bar{x},\bar{y}\in X} \mathbb{E}\big(\big|X_{k,\bar{x}} - X_{k,\bar{y}}\big|\big)\big)_{k\geq 1}$ is non increasing, $\sup_{\bar{x},\bar{y}\in X} \mathbb{E}\big(\big|X_{k,\bar{x}} - X_{k,\bar{y}}\big| \ll k^{-(p-1)}$. Hence condition (11) holds with $q = p - 1$. To end the proof it suffices to notice that condition (12) also holds since, for any $k \geq 1$,

$$\big\|\mathbb{E}(X_k^2|\mathcal{G}_{k-1})\big\|_\infty \leq \int_G (\log N(g))^2 \mu(dg) < \infty \,.$$

$\square$

3.2 Contracting iterated random functions

3.2.1 Uniform contraction

Assume that there is a distance d on X, and that there exist $\kappa > 0$ and $\rho \in (0, 1)$ such that, for any $n \geq 1$,

$$\|\mathbb{E}(d(W_n, W_n^*) | (W_0, W_0^*))\|_\infty \leq \kappa \rho^n \,, \tag{24}$$

where W_n^* is defined in (9). Note that condition (24) holds if the chain is "one step contracting" in the following sense

$$\|d(W_0, W_0^*)\|_\infty < \infty \text{ and } \mathbb{E}(d(W_{1,x}, W_{1,y})) \leq \rho d(x, y) \text{ for any } (x, y) \in X \times X \,.$$

7

Let us now define a class of observables from $G \times X$ to $\mathbb{R}$ for which one can easily compute the coefficient $\delta_\infty(n)$. Let η be a measurable function from G to $\mathbb{R}^+$ such that $\mathbb{E}(\eta(\varepsilon_0)) < \infty$, and let c be a concave non-decreasing function from $\mathbb{R}^+$ to $\mathbb{R}^+$ such that $c(0) = 0$.

One says that $h : G \times X \to \mathbb{R}$ belongs to the class $\mathcal{L}(\eta, c)$ if,

$$|h(z, x) - h(z, y)| \leq \eta(z)c(d(x, y)) \text{ for any } (x, y, z) \in X \times X \times G.$$

Lemma 8 *Assume that the stationary Markov chain $(W_n)_{n \geq 0}$ satisfies the contraction condition (24), and let $(X_n)_{n \geq 1}$ be defined by (8) for some $h \in \mathcal{L}(\eta, c)$. Then, there exists a constant $A > 0$ such that, for any $n \geq 1$,*

$$\delta_\infty(n) \leq Ac(\kappa\rho^{n-1}).$$

Proof. Let $A = \mathbb{E}(\eta(\varepsilon_0))$. Since h belongs to $\mathcal{L}(\eta, c)$, and since c is concave,

$$\mathbb{E}(|X_{n+1,x} - X_{n+1,y}|) \leq A\mathbb{E}(c(d(W_{n,x}, W_{n,y}))) \leq Ac(\mathbb{E}(d(W_{n,x}, W_{n,y}))).$$

Hence, since c is non-decreasing and $(W_n)_{n \geq 0}$ satisfies (24),

$$\|\mathbb{E}(|X_{n+1} - X_{n+1}^*| \,|(W_0, W_0^*))\|_\infty \leq Ac(\|\mathbb{E}(d(W_n, W_n^*) \,|(W_0, W_0^*))\|_\infty) \leq Ac(\kappa\rho^n).$$

$\square$

Applying Theorem 1, the following result holds:

Corollary 9 *Assume that the stationary Markov chain $(W_n)_{n \geq 0}$ satisfies the contraction condition (24), and let $(X_n)_{n \geq 1}$ be defined by (8) for some $h \in \mathcal{L}(\eta, c)$. Assume moreover that $\mathbb{E}(\eta(\varepsilon_1)^p) < \infty$ for some $p > 2$, and that there exists $x_0 \in X$ such that $\|c(d(W_0, x_0))\|_\infty < \infty$ and $\mathbb{E}(|h(\varepsilon_1, x_0)|^p) < \infty$. If $c(\kappa\rho^n) = O(n^{-q})$ for some $q > (p-1)/2$, then the conclusion of Theorem 1 holds.*

Remark 10 *Note that Corollary 9 applies to a large class of continuous observales (as functions of x), including all Hölder observables (case where $c(x) = x^\alpha$ for some $\alpha \in (0, 1)$). More precisely it applies to any concave non-decreasing function c such that $c(x) \leq C|\ln(x)|^{-\gamma}$ in a neighborhood of 0, for some $\gamma > (p-1)/2$.*

Proof of Corollary 9. Applying Lemma 8, we infer that $\delta_\infty(n) = O(n^{-q})$ for some $q > (p-1)/2$. Hence, if one can prove that

$$\sup_{n \geq 1} \|\mathbb{E}(|X_n|^p|\mathcal{G}_{n-1})\|_\infty \leq M, \tag{25}$$

for some finite constant M, the result will follow directly from Theorem 1. To prove (25), we note that

$$\mathbb{E}(|h(\varepsilon_n, W_{n-1})|^p|\mathcal{G}_{n-1}) \leq 2^{p-1}\mathbb{E}(|h(\varepsilon_n, W_{n-1}) - h(\varepsilon_n, x_0)|^p|\mathcal{G}_{n-1}) + 2^{p-1}\mathbb{E}(|h(\varepsilon_1, x_0)|^p). \tag{26}$$

For the first term on the right-hand side of (26), we use the fact that $h \in \mathcal{L}(\eta, c)$, which gives

$$\mathbb{E}(|h(\varepsilon_n, W_{n-1}) - h(\varepsilon_n, x_0)|^p|\mathcal{G}_{n-1}) \leq \mathbb{E}(\eta(\varepsilon_1)^p)\|c(d(W_0, x_0))\|_\infty^p. \tag{27}$$

Under the assumptions of Corollary 9, it follows from (26) and (27) that the upper bound (25) holds. $\square$

8

3.2.2 $\mathbb{L}^1$-contraction

Assume that there is a distance d on X, and that there exist $\kappa > 0$ and $\rho \in (0,1)$ such that, for any $n \geq 1$,

$$\|d(W_n, W_n^*)\|_1 \leq \kappa \rho^n,\qquad (28)$$

where W_n^* is defined in (9). Note that condition (28) holds if the chain is "one step contracting" in the following sense:

$$\mathbb{E}(d(x_0, F(\varepsilon_1, x_0))) < \infty \text{ for some } x_0 \in X$$

and

$$\mathbb{E}(d(W_{1,x}, W_{1,y})) \leq \rho d(x,y) \text{ for any } (x,y) \in X \times X.$$

Note also that, under the two conditions above, there exists an unique stationary distribution ν (see Theorem 2 of [25]).

Let us now define a class of observables from $G \times X$ to $\mathbb{R}$ for which one can easily compute the coefficients $\delta(n)$. Let c be a concave non-decreasing function from $\mathbb{R}^+$ to $\mathbb{R}^+$ such that $c(0) = 0$.

One says that $h : G \times X \to \mathbb{R}$ belongs to the class $\mathcal{L}(c)$ if,

$$\mathbb{E}(|h(\varepsilon_1, x) - h(\varepsilon_1, y)|) \leq c(d(x,y)) \text{ for any } (x,y) \in X \times X.$$

Lemma 11 *Assume that the stationary Markov chain $(W_n)_{n\geq 0}$ satisfies the contraction condition* (28), *and let $(X_n)_{n\geq 1}$ be defined by* (8) *for some $h \in \mathcal{L}(c)$. Then, for $n \geq 2$,*

$$\delta(n) \leq 2^{-1} c(\kappa \rho^{n-2}).$$

Proof. Let $k \geq n \geq 2$. Since h belongs to $\mathcal{L}(c)$, and since c is concave,

$$\|X_k - X_k^*\|_1 \leq \iint \|c(d(W_{k-1}, W_{k-1}^*))\|_1 \leq c\left(\|d(W_{k-1}, W_{k-1}^*)\|_1\right).$$

Hence, since c is non-decreasing and $(W_n)_{n\geq 0}$ satisfies (28),

$$\|X_k - X_k^*\|_1 \leq c\left(\kappa \rho^{k-1}\right).$$

The result follows from the definition of $\delta(n)$ and the fact that c is non-decreasing. $\square$

Recall that the function Q and H related to the tail function $t \mapsto \mathbb{P}(|X_1| > t)$ have been defined in Section 2. Combining Theorem 2 and Lemma 11, the following result holds:

Corollary 12 *Assume that the stationary Markov chain $(W_n)_{n\geq 0}$ satisfies the contraction condition* (28), *and let $(X_n)_{n\geq 1}$ be defined by* (8) *for some $h \in \mathcal{L}(c)$. Assume moreover that*

$$\sum_{n\geq 1} n^{p-2} \int_0^{c(\kappa \rho^n)} Q^{p-1} \circ H^{-1}(u) du < \infty.\qquad (29)$$

Then the conclusion of Theorem 2 holds.

Remark 13 *From Remark 4, it follows that* (29) *holds as soon as*

$$\|X_1\|_r \text{ for some } r > p, \text{ and } \sum_{n\geq 1} n^{(pr-2r+1)/(r-p)} c\left(\kappa \rho^n\right) < \infty.\qquad (30)$$

The condition (30) *is equivalent to the following integral condition on the function c*

$$\int_0^{1/2} \frac{1}{t} c(t) |\ln(t)|^{(pr-2r+1)/(r-p)} dt < \infty.$$

9

3.3 Ergodic Markov chains

3.3.1 A discrete ergodic Markov chain example

Let $(\varepsilon_i)_{i \in \mathbb{Z}}$ be a sequence of iid real-valued random variables distributed as ε with

$$\mathbb{P}(\varepsilon = k) = p_k \,, \ k \in \mathbb{N}^* \,,$$

Let W_0 be a random variable with values in $\mathbb{N}$ independent of $(\varepsilon_i)_{i \in \mathbb{Z}}$, and define for any $k \geq 1$,

$$W_k = (W_{k-1} - 1)\mathbf{1}_{W_{k-1} \neq 0} + (\varepsilon_k - 1)\mathbf{1}_{W_{k-1} = 0} \,. \tag{31}$$

Hence $(W_k, k \in \mathbb{N})$ is a Markov chain with state space $\mathbb{N}$, initial distribution $\mathcal{L}(W_0)$ and transition probabilities satisfying

$$P_{i,i-1} = 1 \text{ and } P_{0,i-1} = p_i \,, \ i \geq 1 \,.$$

Assume that $p_1 > 0$ and $p_{n_j} > 0$ along $n_j \to \infty$. Then the chain $\{W_k; k \geq 0\}$ is irreducible and aperiodic. Moreover, the stationary distribution exists if and only if $\mathbb{E}(\varepsilon) < \infty$ and is given by

$$\nu_0 = 1/\mathbb{E}(\varepsilon) \text{ and } \nu_j = \nu_0 \sum_{i=j+1}^{\infty} p_i \,, \ j \in \mathbb{N}^* \,.$$

Corollary 14 *Let $p > 2$ and f be a function from $\mathbb{N}$ to $\mathbb{R}$ such that $\nu(|f|^r) < \infty$ with $r > p$. Assume that*

$$\sum_{n \geq 1} n^{\frac{p(r-1)}{r-p}} p_n < \infty \,. \tag{32}$$

Then condition (21) is satisfied and the conclusions of Theorem 2 hold for $X_n = f(W_n)$ where $(W_n)_{n \geq 0}$ is the Markov chain defined by (31) with $\mathcal{L}(W_0) = \nu$.

For bounded observables (case $r = \infty$), condition (32) reads as $\sum_{n \geq 1} n^p p_n < \infty$. As we shall see in the proof of the next proposition (see (39)), $\sum_{n \geq 1} n^p p_n < \infty$ is equivalent to $\sum_{n \geq 1} n^{p-2}\mathbb{P}_{\nu \otimes \nu}(T^* \geq n) < \infty$, where T^* is the meeting time defined in (19). The next proposition shows that this latter condition is in some sense optimal.

Proposition 15 *Let $p > 2$ and $(W_k)_{k \geq 0}$ be the Markov chain described above with $p_k := 1/(\zeta(p+1)k^{p+1})$, $k \in \mathbb{N}^*$, where $\zeta(p+1) = \sum_{k \geq 1} k^{-(p+1)}$. Then*

$$\sum_{n \geq 1} n^{p-2}\mathbb{P}_{\nu \otimes \nu}(T^* \geq n) = \infty, \text{ and for any } \varepsilon > 0, \ \sum_{n \geq 2} n^{p-2}(\log n)^{-(1+\varepsilon)}\mathbb{P}_{\nu \otimes \nu}(T^* \geq n) < \infty \,. \tag{33}$$

Moreover, for any stationary and Gaussian centered sequence $(g_k)_{k \in \mathbb{Z}}$ with convergent series of covariances,

$$\limsup_{n \to \infty} (n \log n)^{-1/p}\Big| \sum_{k=1}^{n} (\mathbf{1}_{\{W_k = 0\}} - \nu_0) - \sum_{k=1}^{n} g_k \Big| > 0 \text{ almost surely.} \tag{34}$$

Proof of Corollary 14. Define

$$T_0^* = \inf\{k \in \mathbb{N} \,:\, W_k = W_k^* = 0\} \,.$$

10

By definition, $T^* \leq T_0^*$. Hence for any $n \in \mathbb{N}$,

$$\mathbb{P}_{\nu \otimes \nu}(T^* \geq n) \leq \mathbb{P}_{\nu \otimes \nu}(T_0^* \geq n). \tag{35}$$

Next, it is esay to see that for any $n \in \mathbb{N}$,

$$\mathbb{P}_{\nu \otimes \nu}(T_0^* = n) = \mathbb{P}_{\nu \otimes \nu}(T_0' = n) \tag{36}$$

with

$$T_0' = \inf\{k \geq 0 \,:\, W_k = W_k' = 0\},$$

where $(W_k', k \in \mathbb{N})$ is the Markov chain defined as follows: Let $(\varepsilon_k')_{k \in \mathbb{Z}}$ be an independent copy of $(\varepsilon_k)_{k \in \mathbb{Z}}$ and independent of W_0. Let W_0' be independent of $(W_0, (\varepsilon_k)_{k \in \mathbb{Z}}, (\varepsilon_k')_{k \in \mathbb{Z}})$ and, for any $k \geq 1$, set

$$W_k' = (W_{k-1}' - 1)\mathbf{1}_{W_{k-1}' \neq 0} + \varepsilon_k' \mathbf{1}_{W_{k-1}' = 0}.$$

According to Lindvall [16], if $\mathbb{E}_\nu(\psi(\tau)) < \infty$ where $\tau = \inf\{k \geq 1 \,:\, W_k = 0\}$ and ψ is a non-decreasing function from $\mathbb{N}$ to $[2, \infty[$ such that $((\log(\psi(n))/n)_n$ is non-increasing and converges to 0, then $\mathbb{E}_{\nu \otimes \nu}(\psi(T_0')) < \infty$. Note now that

$$\mathbb{P}_\nu(\tau \geq n) = \sum_{\ell=0}^{n-1} \nu_\ell \sum_{j \geq n-\ell} p_j + \sum_{\ell \geq n} \nu_\ell \leq 2 \sum_{\ell \geq [n/2]} \nu_\ell + \sum_{j \geq [n/2]+1} p_j. \tag{37}$$

Hence under (32), $\mathbb{E}_\nu(\psi_{r,p}(\tau)) < \infty$ with $\psi_{r,p}(x) = x^{r(p-1)/(r-p)}$. It follows that $\mathbb{E}_\nu(\psi_{r,p}(T_0')) < \infty$ which in turn implies that $\mathbb{E}_\nu(\psi_{r,p}(T^*)) < \infty$ by taking into account (35) and (36). Therefore condition (21) is satisfied and Corollary 5 applies. $\qquad \square$

Proof of Proposition 15. Note first that the following coupling inequality holds: for any $n \geq 1$,

$$\beta(n) := \frac{1}{2} \int \|\delta_x P^n - \nu\|_v \nu(dx) \leq \mathbb{P}_{\nu \otimes \nu}(T^* \geq n), \tag{38}$$

where $\|\mu\|_v$ denotes the total variation norm of a signed measure μ and P is the transition function of the Markov chain $(W_k)_{k \in \mathbb{N}}$. But for any $n \geq 1$, $\beta(n) \geq 2\alpha(n)$ where $(\alpha(n))_{n \geq 1}$ is the sequence of strong mixing coefficients of the chain which starts from the stationary distribution. As quoted in Chapter 30 of Bradley [6],

$$\alpha(n) \geq \left| \mathbb{P}_\nu(W_0 \geq n+1, W_n = 0) - \mathbb{P}_\nu(W_0 \geq n+1)\mathbb{P}_\nu(W_n = 0) \right| = \nu_0 \sum_{k \geq n+1} \nu_k.$$

It follows that for any $s \geq 0$,

$$\sum_{n \geq 1} n^s \sum_{k \geq n+1} \nu_k = \infty \Rightarrow \sum_{n \geq 1} n^s \mathbb{P}_{\nu \otimes \nu}(T^* \geq n) = \infty,$$

which together with the arguments developed in the proof of Corollary 14 show that

$$\sum_{n \geq 1} n^{s+2} p_n < \infty \iff \sum_{n \geq 1} n^s \mathbb{P}_{\nu \otimes \nu}(T^* \geq n) < \infty. \tag{39}$$

This proves the first part of (33). To prove its second part, it suffices to use again the arguments developed in the proof of Corollary 14 and to notice that, for $p_k := 1/(\zeta(p+1)k^{p+1})$, $k \in \mathbb{N}^*$,

11

the upper bound (37) entails that $\mathbb{E}_\nu(\psi_p(\tau)) < \infty$ with $\psi_p(x) = \frac{x^{p-1}}{(\log(1+x))^{1+\varepsilon}}$ where $\varepsilon > 0$. This ends the proof of (33).

To prove the second part of the proposition, we shall use similar arguments as those developed in the proof of Theorem 2.2 in Dedecker-Merlevède-Rio [11] and adopt the following notations: the regeneration times $(T_k)_{k\geq 0}$ of the Markov chain $(W_k)_{k\geq 0}$ are defined by induction as follows: $R_0 = \inf\{n > 0 : W_n = 0\}$ and $R_k = \inf\{n > R_{k-1} : W_n = 0\}$. Let $\tau_k = R_{k+1} - R_k$ for $k \geq 0$. Note that $(\tau_k)_{k\geq 0}$ are iid and that their common law is the law of R_0 when the chain starts at zero. Note that

$$\mathbb{P}(\tau_k > (k \ln k)^{1/p}) \sim c_p/(k \log k).$$

Since the regeneration times τ_k are independent, by the converse Borel-Cantelli lemma, it follows that

$$\mathbb{P}(\tau_k > (k \log k)^{1/p} \text{ infinitely often }) = 1.$$

Now we take

$$f(x) := \mathbf{1}_{x=0} \text{ and } g(x) := f(x) - \nu_0.$$

f is obviously a bounded function and $\nu(g) = 0$. Note that, for any $\ell \geq 0$,

$$\sum_{k=R_\ell+1}^{R_\ell+m} (\nu_0/2 - f(W_k)) = m\nu_0/2 \text{ for any } 1 \leq m < \tau_\ell.$$

Since R_n/n converges to $\mathbb{E}(\tau_0)$ almost surely, it follows that, for some positive constant c depending on $\mathbb{E}(\tau_0)$,

$$\limsup_n \sum_{i=n+1}^{n+[c(n\log n)^{1/p}]} (\nu_0/2 - f(W_i)) \geq 0 \text{ almost surely.} \tag{40}$$

Consider now a stationary and Gaussian centered sequence $(g_k)_{k\in\mathbb{Z}}$ with convergent series of covariances. If follows from both the Borel-Cantelli lemma and the usual tail inequality for Gaussian random variables that, for any positive θ,

$$\liminf_n \sum_{i=n+1}^{n+[c(n\log n)^{1/p}]} (g_i + \theta) \geq 0 \text{ almost surely.}$$

Taking $\theta = \nu_0/4$ in the above inequality and using (40), we then infer that

$$\limsup_{n\to\infty} \frac{1}{[c(n\log n)^{1/p}]} \sum_{i=n+1}^{n+[c(n\log n)^{1/p}]} \left(g_i + \nu_0 - f(W_i)\right) \geq \nu_0/4 \text{ almost surely,}$$

which implies (34). $\square$

12

3.3.2 An example of ergodic Markov chain with continuous state space

In this section, we consider an homogenous Markov chain with state space $[0,1]$ and transition probability kernel $P(x,\cdot)$ given by

$$P(x,A) = (1-x)\delta_x(A) + x\pi(A), \tag{41}$$

where δ_x denotes the Dirac measure at point x and

$$\pi(dx) = (a+1)x^a dx \text{ with } a > 1.$$

Note that the chain is irreducible and aperiodic and admits a unique invariant probability measure ν given by

$$\nu(dx) = ax^{a-1}dx.$$

As in Section 9.3 in Rio [23], we now construct a stationary Markov chain $(W_n)_{n\in\mathbb{N}}$ with initial law ν and transition probability measure $P(x,\cdot)$. Let ξ_0 be a random variable with law ν. We assume that the underlying probability space is rich enough to contain a sequence $(\varepsilon_i)_{i\in\mathbb{Z}} := (U_i,V_i)_{i\in\mathbb{Z}}$ of independent random variables with uniform law over $[0,1]^2$, and that this random sequence is independent of ξ_0. The stationary Markov chain $(W_n)_{n\in\mathbb{N}}$ is then constructed via the following recursive equation: $W_0 = \xi_0$ and, for any $k \geq 1$,

$$W_k = W_{k-1}\mathbf{1}_{U_k \geq W_{k-1}} + F_\pi^{-1}(V_k)\mathbf{1}_{U_k < W_{k-1}}, \tag{42}$$

where F_π^{-1} is the inverse of the cumulative function of π. It is easy to see that $(W_n)_{n\in\mathbb{N}}$ is a Markov chain with initial distribution ν and transition probability kernel given by (41).

Corollary 16 *Let $p > 2$ and $(W_k)_{k\in\mathbb{N}}$ be the stationary Markov chain defined by (42) with $a > p-1$. Then condition (22) is satisfied and the conclusions of Theorem 2 hold for $X_n = f(W_n)$, for any bounded function f defined on $[0,1]$.*

The proof of this corollary is a direct application of Corollary 5 by taking into account the following lemma whose proof is postponed to the Appendix (see Section 5.2).

Lemma 17 *For any $a > 1$ there exist positive constants $c(a)$ and $C(a)$ depending only on a such that for any $n \geq 1$,*

$$\frac{c(a)}{n^a} \leq \mathbb{P}_{\nu\otimes\nu}(T^* > n) \leq \frac{C(a)}{n^a}, \tag{43}$$

where T^* is the meeting time defined in (19).

In addition, this lemma together with Theorem 2.2 in Dedecker-Merlevède-Rio [11] proves the sharpness of condition (22) also in case of Markov chains with continuous state space. This is summarized in the next proposition.

Proposition 18 *Let $p > 2$ and $(W_k)_{k\in\mathbb{N}}$ be the stationary Markov chain defined by (42) with $a = p-1$. Then condition (22) fails. In addition, for any map f from $[0,1]$ to $\mathbb{R}$ with continuous and strictly positive derivative f' on $[0,1]$, and any stationary and Gaussian centered sequence $(g_k)_{k\in\mathbb{Z}}$ with convergent series of covariances,*

$$\limsup_{n\to\infty}(n\log n)^{-1/p}\Big|\sum_{k=1}^n f(W_k^a) - n\int_0^1 f(t)dt - \sum_{k=1}^n g_k\Big| > 0 \text{ almost surely.} \tag{44}$$

3.4 Lipschitz autoregressive models

We consider the autoregressive Lipschitz model as in Dedecker-Rio [13]. Let $\tau \in [0,1)$, $C \in (0,1]$ and $f : \mathbb{R} \to \mathbb{R}$ a 1-Lipschitz function such that

$$f(0) = 0 \quad \text{and} \quad |f'(t)| \leq 1 - \frac{C}{(1 + |t|)^\tau} \quad \text{for almost every } t \,.$$

Let $(\varepsilon_i)_{i \geq 1}$ be iid real-valued random valued with common law μ and define for any $n \geq 1$

$$W_n = f(W_{n-1}) + \varepsilon_n \,, \quad \text{with } W_0 \text{ independent of } (\varepsilon_i)_{i \geq 1} \,. \tag{45}$$

Let $S_n(h) = \sum_{k=1}^n h(W_i)$ for any measurable function h.

The model above corresponds to the previously considered situation with $G = \mathbb{R}$ and $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $F(x,y) = x + f(y)$, for every $x, y \in \mathbb{R}$.

Let $S \geq 1$ and assume that μ admits a moment of order S. It follows from Dedecker-Rio [13] that there exists a unique invariant probability ν on $\mathbb{R}$, such that

$$\int_{\mathbb{R}} |x|^{S-\tau} \nu(dx) < \infty \,. \tag{46}$$

The following strong approximation with rates holds.

Corollary 19 *Let $\tau \in (0,1)$ and assume that μ admits a moment of order $S = p + \tau p$ for some $p > 2$. Let $(W_n)_{n \geq 0}$ be defined by (45) with $\mathcal{L}(W_0) = \nu$. Then, for any Lipschitz function h such that $\nu(h) = 0$, $n^{-1}\mathrm{Var}(S_n(h)) \to \sigma^2(h)$ as $n \to \infty$ and one can redefine $(W_n)_{n \geq 0}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i \geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2(h))$, such that,*

$$S_n(h) - \sum_{i=1}^n N_i = o(n^{1/p}) \;\; \mathbb{P}\text{-a.s.}$$

Proof. The result comes from an application of Theorem 2 by taking into account Remark 4. As already mentionned, ν admits a moment of order $S - \tau = p + (p-1)\tau$. Hence, one can prove that condition (18) holds with $r = p + \tau(p-1)$, by using the following lemma (taking $t = 1$ and $\gamma = (pr - 2r + 1)/(r - p) = -2 + (S-1)/\tau$).

Lemma 20 *Let $\gamma > -1$ and $t > 0$. Assume that $S \geq t + (\gamma + 2)\tau$. Then $\sum_{n \geq 1} n^\gamma \delta(n) < \infty$.*

The proof of the lemma above is postponed to the Appendix (see Section 5.3).

4 Proofs of Theorems 1 and 2

The proofs of Theorems 1 and 2 follow the scheme of proof of Theorem 2.1 in Berkes-Liu-Wu [3] by applying the following general Proposition 21, which comes from a careful analysis of the proof of their strong approximation result. To state this general proposition several preliminary notations are needed.

14

A Preliminary result. For Proposition 21 below, we consider $(X_k)_{k\geq 1}$ a strictly stationary sequence of real-valued random variables in $\mathbb{L}^p$ $(p>2)$ and $(\varepsilon_i)_{i\geq 0}$ a sequence of iid random variables. Let $(M_k)_{k\geq 1}$ be a sequence of positive real numbers and define

$$\varphi_k(x) = (x \wedge M_k) \vee (-M_k) \text{ and } g_k(x) = x - \varphi_k(x). \tag{47}$$

Then, define

$$X_{k,j} = \varphi_k(X_j) - \mathbb{E}\varphi_k(X_j) \text{ and } W_{k,\ell} = \sum_{i=1+3^{k-1}}^{\ell+3^{k-1}} X_{k,i}. \tag{48}$$

Let now $(m_k)_{k\geq 1}$ be a non-decreasing sequence of positive integers such that $m_k = o(3^k)$, as $k \to \infty$, and define

$$\tilde{X}_{k,j} = \mathbb{E}\big(\varphi_k(X_j)|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}\big) - \mathbb{E}\varphi_k(X_j) \text{ for any } j \geq m_k \text{ and } \widetilde{W}_{k,\ell} = \sum_{i=1+3^{k-1}}^{\ell+3^{k-1}} \tilde{X}_{k,i}. \tag{49}$$

Finally set $k_0 := \inf\{k \geq 1 : m_k \leq 2^{-1}3^{k-2}\}$ and define

$$\nu_k = m_k^{-1}\big\{\mathbb{E}(\widetilde{W}_{k,m_k}^2) + 2\mathbb{E}(\widetilde{W}_{k,m_k}(\widetilde{W}_{k,2m_k} - \widetilde{W}_{k,m_k}))\big\}. \tag{50}$$

The general proposition coming from a careful analysis of the proof of Theorem 2.1 in Berkes-Liu-Wu [3] reads as follows

Proposition 21 (Berkes-Liu-Wu [3]) *Let $p > 2$. Assume that we can find a sequence of positive real numbers $(M_k)_{k\geq 1}$ a non-decreasing sequence of positive integers $(m_k)_{k\geq 1}$ such that $m_k = o(3^{2k/p}k^{-1})$, as $k \to \infty$, in such a way that the following conditions are satisfied:*

$$\sum_{k\geq 1} 3^{k(p-1)/p}\mathbb{E}(|g_k(X_1)|) < \infty, \tag{51}$$

there exists $\alpha \geq 1$ such that

$$\sum_{k\geq k_0} 3^{-\alpha k/p}\Big\| \max_{1\leq \ell \leq 3^k - 3^{k-1}} |W_{k,\ell} - \widetilde{W}_{k,\ell}|\Big\|_\alpha^\alpha < \infty, \tag{52}$$

and there exists $r \in]2, \infty[$ such that

$$\sum_{k\geq k_0} \frac{3^k}{3^{kr/p}m_k}\mathbb{E}\Big(\max_{1\leq \ell \leq 3m_k} |\widetilde{W}_{k,\ell}|^r\Big) < \infty. \tag{53}$$

Assume in addition that

$$\text{the series } \sigma^2 = \text{Var}(X_1) + 2\sum_{i\geq 1} \text{Cov}(X_1, X_{i+1}) \text{ converge}, \tag{54}$$

and

$$3^k(\nu_k^{1/2} - \sigma)^2 = o(3^{2k/p}(\log k)^{-1}), \text{ as } k \to \infty. \tag{55}$$

Then, one can redefine $(X_n)_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,

$$S_n - n\mathbb{E}(X_1) - \sum_{i=1}^n N_i = o(n^{1/p}) \;\; \mathbb{P}\text{-a.s.} \tag{56}$$

15

Note that (54) implies that $n^{-1}\mathrm{Var}(S_n)$ converges to σ^2 (which is therefore non-negative). Let us now briefly explain how the proposition follows from the work of Berkes-Liu-Wu [3].

Condition (51) together with condition (52) prove that it is enough to show (56) with

$$\tilde{S}_n = \sum_{k=1}^{h_n-1} \widetilde{W}_{k,3^k-3^{k-1}} + \widetilde{W}_{h_n,n-3^{h_n-1}}$$

instead of $S_n - n\mathbb{E}(X_1)$, where, for $n \geq 2$, $h_n := \lceil (\log n)/(\log 3) \rceil$ (so that h_n is the unique integer such that $3^{h_n-1} < n \leq 3^{h_n}$). Next, condition (53) allows first to show that the proof of the proposition is reduced to prove (56) with $S_n^\diamond$ replacing $S_n - n\mathbb{E}(X_1)$ where

$$S_n^\diamond = \sum_{k=K_0}^{h_n-1} \sum_{j=1}^{q_k} B_{k,j} + \sum_{j=1}^{\tau_n} B_{h_n,j}, \ \text{ where } \tau_n = \Big[\frac{n-3^{h_n-1}}{3m_{h_n}}\Big] - 2\,,$$

with $B_{k,j} = 0$ if $k < k_0$ and for $k \geq k_0$,

$$B_{k,j} = \sum_{i=1+3jm_k+3^{k-1}}^{3(j+1)m_k+3^{k-1}} \tilde{X}_{k,i}\,, \ j = 1, 2, \ldots, q_k = [2 \times 3^{k-2}/m_k] - 2\,.$$

A careful analysis of the steps 3.2 and 3.3 of the proof of Theorem 2.1 in Berkes-Liu-Wu [3] reveals that condition (53) is also sufficient to apply Theorem 1 in Sakhanenko [24] (at different steps of their proof) and this leads to the following strong approximation result: one can redefine $(X_n)_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exists a standard Brownian motion $B = \{B(t), t \in \mathbb{R}^+\}$ such that,

$$\max_{i \leq n} \big|S_i^\diamond - B(\sigma_i^2)\big| = o(n^{1/p}) \ \mathbb{P}\text{-a.s.}$$

where

$$\sigma_n^2 = \sum_{k=k_0}^{h_n-1} 3m_k q_k \nu_k + 3m_{\tau_n}\tau_n\nu_{h_n}\,.$$

The last step 3.4 of their proof then consists in showing that one can construct another standard Brownian motion $W = \{W(t), t \in \mathbb{R}^+\}$ (depending on B) such that

$$B(\sigma_n^2) - \sigma W(n) = o(n^{1/p}) \ \mathbb{P}\text{-a.s.}$$

This step is achieved provided that we can prove that $\nu_k \to \sigma^2$, $m_k = o(3^{2k/p}k^{-1})$, as $k \to \infty$, and condition (55) holds.

Some preliminary considerations. The following considerations allowing to extend the stationary sequence $(X_n)_{n\geq 1}$ defined by (8) to a stationary sequence on $\mathbb{Z}$ will be useful.

For any $n \geq 1$, let $V_n = (\varepsilon_n, W_{n-1})$. Hence $(X_n)_{n\geq 1}$ is a functional of the Markov chain $(V_n)_{n\geq 1}$ with state space $G \times X$ and stationary distribution $\mu \otimes \nu$. The Markov chain $(V_n)_{n\geq 1}$ being stationary, by Kolmogorov's theorem, there exists a probability $\hat{\mathbb{P}}$ on the measurable space $(\hat{\Omega}, \hat{\mathcal{F}}) = ((G \times X)^{\mathbb{Z}}, (\mathcal{B}(G) \times \mathcal{B}(X))^{\mathbb{Z}})$ invariant by the shift $\hat{\eta}$ on $\hat{\Omega}$ and such that the law of the coordinate process $(\hat{V}_n = (\hat{\varepsilon}_n, \hat{W}_{n-1}))_{n\in\mathbb{Z}}$ (with values in $G \times X$) under $\hat{\mathbb{P}}$ is the same as

16

the one of $(V_n)_{n\geq 1}$ under $\mathbb{P}_\nu$. Hence, if we define for any integer n, $\hat{X}_n := h(\hat{V}_0) \circ \hat{\eta}^n$, it follows that $(\hat{X}_n)_{n\in\mathbb{Z}}$ forms a stationary sequence with stationary distribution π, whose law under $\hat{\mathbb{P}}$ is the same as the one of $(X_n)_{n\geq 1}$ under $\mathbb{P}_\nu$. To prove the theorem, it suffices then to prove that it holds for the extended sequence $(\hat{X}_n)_{n\in\mathbb{Z}}$ which is a stationary sequence adapted to the stationary filtration $(\widehat{\mathcal{F}}_n)_{n\in\mathbb{Z}}$ where $\widehat{\mathcal{F}}_n = \sigma(\hat{V}_k, k \leq n)$ To avoid additional notations, in the rest of the proof we write $(X_n)_{n\in\mathbb{Z}}$ for $(\hat{X}_n)_{n\in\mathbb{Z}}$, $(V_n)_{n\in\mathbb{Z}}$ for $(\hat{V}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ for $(\widehat{\mathcal{F}}_n)_{n\in\mathbb{Z}}$.

4.1 Proof of Theorem 1

By the reverse martingale convergence theorem and stationarity, $\|\mathbb{E}(X_n|\mathcal{F}_0) - \mathbb{E}(X_n)\|_2$ is decreasing to $\|\mathbb{E}(X_0|\mathcal{F}_{-\infty}) - \mathbb{E}(X_0)\|_2$, as $n \to \infty$. Hence, by condition (11), $\mathbb{E}(X_0|\mathcal{F}_{-\infty}) = \mathbb{E}(X_0)$ a.s. Applying Lemma 22 of the Appendix and taking into account condition (11), we get (since $q > 1/2$),

$$\sum_{k\geq 1} |\text{Cov}(X_1, X_{k+1})| \ll \Big(\sum_{k\geq 0}(k+1)^{-1/2}\|\mathbb{E}(X_k|V_0) - \mathbb{E}(X_k)\|_2\Big)^2 < \infty \,.$$

This proves that the series $\sigma^2 = \text{Var}(X_1) + 2\sum_{i\geq 1}\text{Cov}(X_1, X_{i+1})$ converge absolutely and condition (54) of Proposition 21 holds.

Assume first that $\sigma^2 = 0$. To prove that $S_n - n\mathbb{E}(X_1) = o(n^{1/p})$ a.s., we shall use Theorem 4.7 in Cuny-Merlevède [8]. Hence, it suffices to prove that

$$\sum_{n\geq 1} \frac{\|S_n - n\mathbb{E}(X_1)\|_p}{n^{1+1/p}} < \infty \,. \tag{57}$$

With this aim, we start by noticing that by condition (11),

$$\sum_{k\geq 1} k^{-1/p}\|\mathbb{E}(X_k|V_0) - \mathbb{E}(X_k)\|_p < \infty \text{ and therefore } \sum_{k\geq 1} \frac{\|\mathbb{E}(S_k|V_0) - k\mathbb{E}(X_1)\|_p}{k^{1+1/p}} < \infty \,.$$

Theorem 2.3 in [8] then asserts that there exists a stationary sequence $(D_k)_{k\in\mathbb{Z}}$ of martingale differences in $\mathbb{L}^p$, adapted to $(\mathcal{F}_k)_{k\in\mathbb{Z}}$ and such that $n^{-1/2}\|S_n - n\mathbb{E}(X_1) - \sum_{k=1}^n D_k\|_p \to 0$, as $n \to \infty$. Together with the fact that $\lim_{n\to\infty} n^{-1}\text{Var}(S_n) = \sigma^2 = 0$, it follows that $D_k = 0$ a.s, for any k. Therefore, the upper bound (4) in [8] and condition (11) entail that

$$\sum_{n\geq 1} \frac{\|S_n - n\mathbb{E}(X_1)\|_p}{n^{1+1/p}} \ll \sum_{n\geq 1} \frac{n^{1/2}}{n^{1+1/p}} \sum_{k\geq[n^{p/2}]} \frac{\|\mathbb{E}(S_k|V_0) - k\mathbb{E}(X_1)\|_p}{k^{1+1/p}}$$

$$\ll \sum_{k\geq 1} \frac{1}{k^{1+2/p^2}}\|\mathbb{E}(S_k|V_0) - k\mathbb{E}(X_1)\|_p \ll \sum_{k\geq 1} \frac{1}{k^{2/p^2}}\|\mathbb{E}(X_k|V_0) - \mathbb{E}(X_1)\|_p \ll \sum_{k\geq 1} \frac{1}{k^{q+2/p^2}} \,,$$

which proves (57) since $q + 2/p^2 - 1 > (2p^2)^{-1}(p^3 - 3p^2 + 4) = (2p^2)^{-1}(p-2)^2(p+1) > 0$. The theorem is then proved in the case where $\sigma^2 = 0$.

Assume from now that $\sigma^2 > 0$. We choose

$$M_k = 3^{k/p} \text{ and } m_k = [3^{2(1-\varepsilon)k/p}] \text{ with } 0 < \varepsilon < \min\Big(1 - \frac{p-1}{2q}, \frac{1}{2}\Big),$$

17

Note that the sequence $(m_k)_{k\geq 0}$ satisfies $m_k = o(3^{2k/p}k^{-1})$, as $k \to \infty$. We prove below that conditions (51), (52), (53) and (55) of Proposition 21 are satisfied with the above choices of $(M_k)_{k\geq 0}$ and $(m_k)_{k\geq 0}$.

Since the X_i's are in $\mathbb{L}^p$, it is easy to see that with the choice of M_k, condition (51) is satisfied (it suffices to write that $\mathbb{E}(|g_k(X_1)|) \leq \mathbb{E}(|X_1|\mathbf{1}_{|X_1|>M_k})$ and to use Fubini's Theorem). Next, for $k \geq k_0$, Lemma 24 of the Appendix combined with condition (11) implies that

$$\Big\| \max_{1\leq \ell \leq 3^k-3^{k-1}} \big|W_{k,\ell} - \widetilde{W}_{k,\ell}\big| \Big\|_1 \leq \sum_{i=1+3^{k-1}}^{3^k} \|X_{k,i} - \tilde{X}_{k,i}\|_1 \leq \frac{C3^k}{m_k^q}.$$

Therefore,

$$\sum_{k\geq k_0} 3^{-k/p}\Big\| \max_{1\leq \ell \leq 3^k-3^{k-1}} \big|W_{k,\ell} - \widetilde{W}_{k,\ell}\big| \Big\|_1 \ll \sum_{k\geq 1} \frac{3^{k(p-1)/p}}{3^{2q(1-\varepsilon)k/p}} < \infty \,,$$

since $2q(1-\varepsilon) > p-1$. Condition (52) is then satisfied with $\alpha = 1$. We prove now that we can find a real number $r \in]2,\infty[$ such that (53) holds. Let $r \geq 2$,

$$Y_{k,i} = \tilde{X}_{k,i+3^{k-1}} \,, \ \mathcal{G}_{k,i} := \mathcal{G}_{i+3^{k-1}} = \sigma(\varepsilon_{i+3^{k-1}}, \ldots, \varepsilon_1, W_0) \,,$$

and

$$d_{k,i} = Y_{k,i} - \mathbb{E}(Y_{k,i}|\mathcal{G}_{k,i-1}) \,.$$

With these notations, we have

$$\Big\| \max_{1\leq \ell \leq 3m_k} \big|\widetilde{W}_{k,\ell}\big| \Big\|_r \leq \Big\| \max_{1\leq \ell \leq 3m_k} \Big|\sum_{i=1}^{\ell} d_{k,i}\Big| \Big\|_r + \Big\| \max_{1\leq \ell \leq 3m_k} \Big|\sum_{i=1}^{\ell} \mathbb{E}(Y_{k,i}|\mathcal{G}_{k,i-1})\Big| \Big\|_r \,. \tag{58}$$

By Rosenthal's inequality for martingales,

$$\Big\| \max_{1\leq \ell \leq 3m_k} \Big|\sum_{i=1}^{\ell} d_{k,i}\Big| \Big\|_r^r \ll \Big(\sum_{i=1}^{3m_k} \|\mathbb{E}(d_{k,i}^2|\mathcal{G}_{k,i-1})\|_{r/2}\Big)^{r/2} + \sum_{i=1}^{3m_k} \|d_{k,i}\|_r^r \,.$$

Note that

$$\mathbb{E}(d_{k,i}^2|\mathcal{G}_{k,i-1}) \leq \mathbb{E}(\mathbb{E}^2(\varphi_k(X_{i+3^{k-1}})|\mathcal{H}_{k,i})|\mathcal{G}_{k,i-1}) \leq \mathbb{E}(\mathbb{E}(X_{i+3^{k-1}}^2|\mathcal{H}_{k,i})|\mathcal{G}_{k,i-1}) \,,$$

where $\mathcal{H}_{k,i} = \sigma(\varepsilon_{i+3^{k-1}}, \ldots, \varepsilon_{i+3^{k-1}-m_k})$. Here, recall the following well known fact: if Y is an integrable random variable, and $\mathcal{G}_1$ and $\mathcal{G}_2$ are two σ-algebras such that $\sigma(Y) \vee \mathcal{G}_1$ is independent of $\mathcal{G}_2$, then

$$\mathbb{E}(Y|\mathcal{G}_1 \vee \mathcal{G}_2) = \mathbb{E}(Y|\mathcal{G}_1) \quad \text{a.s.} \tag{59}$$

Applying (59) with $\mathcal{G}_1 = \sigma(\varepsilon_{i+3^{k-1}-1}, \ldots, \varepsilon_{i+3^{k-1}-m_k})$, $\mathcal{G}_2 = \mathcal{G}_{k,i-m_k-1}$ and $Y = \mathbb{E}(X_{i+3^{k-1}}^2|\mathcal{H}_{k,i})$, we get

$$\mathbb{E}(d_{k,i}^2|\mathcal{G}_{k,i-1}) \leq \mathbb{E}(X_{i+3^{k-1}}^2|\sigma(\varepsilon_{i+3^{k-1}-1}, \ldots, \varepsilon_{i+3^{k-1}-m_k})) \,.$$

Hence, by assumption (12),

$$\|\mathbb{E}(d_{k,i}^2|\mathcal{G}_{k,i-1})\|_{r/2} \leq \|\mathbb{E}(X_{i+3^{k-1}}^2|\mathcal{G}_{k,i-1})\|_{r/2} \leq \|\mathbb{E}(X_{i+3^{k-1}}^2|\mathcal{G}_{k,i-1})\|_\infty \leq c \,.$$

18

On another hand, by stationarity,

$$\|d_{k,i}\|_r^r \ll \|\varphi_k(X_0)\|_r^r \ll \|X_0 \mathbf{1}_{|X_0| \le M_k}\|_r^r + M_k^r \mathbb{P}(|X_0| > M_k) \,.$$

So, overall,

$$\Big\| \max_{1 \le \ell \le 3m_k} \Big| \sum_{i=1}^{\ell} d_{k,i} \Big| \Big\|_r^r \ll m_k^{r/2} + m_k \big(\|X_0 \mathbf{1}_{|X_0| \le M_k}\|_r^r + M_k^r \mathbb{P}(|X_0| > M_k) \big) \,. \tag{60}$$

We handle now the second term in the right-hand side of (58). We apply Proposition 23 of the Appendix with $\alpha = r$, $r = r_k$ where r_k is the unique positive integer such that $2^{r_k - 1} \le 3m_k < 2^{r_k}$,

$$Z_i = 0 \text{ for } i \le 0 \,, \ Z_i := Z_{k,i} = \mathbb{E}(Y_{k,i} | \mathcal{G}_{k,i-1}) \text{ for } i \ge 1$$

and

$$\mathcal{F}_i = \{\emptyset, \Omega\} \text{ for } i \le 0 \,, \ \mathcal{F}_i = \sigma(\varepsilon_{i-1+3^{k-1}}, \dots, \varepsilon_1, W_0) = \mathcal{G}_{k,i-1} \text{ for } i \ge 1 \,.$$

We then get

$$\Big\| \max_{1 \le \ell \le 3m_k} \Big| \sum_{i=1}^{\ell} \mathbb{E}(Y_{k,i} | \mathcal{G}_{k,i-1}) \Big| \Big\|_r$$

$$\ll \Big(\sum_{j=1}^{3m_k} \|\mathbb{E}(Y_{k,j} | \mathcal{G}_{k,j-1})\|_r^2 \Big)^{1/2} + \sum_{\ell=0}^{r_k-1} \Big(\sum_{m=1}^{2^{r_k - \ell}} \|\mathbb{E}(T_{m2^\ell} - T_{(m-1)2^\ell} | \mathcal{F}_{(m-2)2^\ell + 1})\|_r^2 \Big)^{1/2} \,, \tag{61}$$

where $T_\ell = \sum_{i=1}^{\ell} \mathbb{E}(Y_{k,i} | \mathcal{G}_{k,i-1})$. By fact (59), we note that, for any $i \ge 1$,

$$\|\mathbb{E}(Y_{k,i} | \mathcal{G}_{k,i-1})\|_\infty = \|\mathbb{E}(X_{k,i+3^{k-1}} | \sigma(\varepsilon_{i+3^{k-1}-1}, \dots, \varepsilon_{i+3^{k-1}-m_k}))\|_\infty$$

$$\le 2\|\mathbb{E}(\varphi_k(X_{i+3^{k-1}}) | \mathcal{G}_{k,i-1})\|_\infty \le 2\|\mathbb{E}(\varphi_k^2(X_{i+3^{k-1}}) | \mathcal{G}_{k,i-1})\|_\infty^{1/2} \,.$$

Therefore, by condition (12),

$$\|\mathbb{E}(Y_{k,i} | \mathcal{G}_{k,i-1})\|_\infty \le 2\sqrt{c} \,. \tag{62}$$

Next, since $\mathcal{F}_i = \{\emptyset, \Omega\}$ for $i \le 0$ and the Z_i's are centered, for any $\ell \ge 0$,

$$\|\mathbb{E}(T_{2^\ell} | \mathcal{F}_{-2^\ell + 1})\|_r = 0 \,.$$

Moreover, for any $m \ge 2$ and any $\ell \ge 0$,

$$\|\mathbb{E}(T_{m2^\ell} - T_{(m-1)2^\ell} | \mathcal{F}_{(m-2)2^\ell + 1})\|_r \le \sum_{i=(m-1)2^\ell + 1}^{m2^\ell} \|\mathbb{E}(\tilde{X}_{k,i+3^{k-1}} | \mathcal{G}_{k,(m-2)2^\ell})\|_r \,.$$

But, for any $m \ge 2$, any $\ell \ge 0$ and any $i \ge (m-1)2^\ell + 1$,

$$\|\mathbb{E}(\tilde{X}_{k,i+3^{k-1}} | \mathcal{G}_{k,(m-2)2^\ell})\|_r$$

$$= \|\mathbb{E}(\mathbb{E}(\varphi_k(X_{i+3^{k-1}}) | \varepsilon_{i+3^{k-1}}, \dots, \varepsilon_{i+3^{k-1}-m_k}) | \mathcal{G}_{k,(m-2)2^\ell}) - \mathbb{E}(\varphi_k(X_{i+3^{k-1}}))\|_r \,.$$

19

Hence, if $2^\ell \geq m_k$,

$$\|\mathbb{E}(\tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell})\|_r = 0\,,$$

and if $2^\ell \leq m_k - 1$, by using (59),

$$\|\mathbb{E}(\tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell})\|_r \leq \|\mathbb{E}(X_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell})\|_\infty\,.$$

But, by using stationarity, the Markov property and the fact that φ_k is 1-Lipschitz,

$$\|\mathbb{E}(X_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell})\|_\infty \leq \sup_{x \in X}\left|\mathbb{E}(\varphi_k(X_{i-((m-2)2^\ell,x)})) - \int \mathbb{E}(\varphi_k(X_{i-((m-2)2^\ell,y)}))d\nu(y)\right|$$

$$\leq \sup_{x \in X}\int\left|\mathbb{E}(\varphi_k(X_{i-((m-2)2^\ell,x)})) - \mathbb{E}(\varphi_k(X_{i-((m-2)2^\ell,y)}))\right|d\nu(y)\,.$$

Hence, for any $m \geq 2$, any $\ell \geq 0$ and any $i \geq (m-1)2^\ell + 1$,

$$\|\mathbb{E}(\tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell})\|_r \leq \|\mathbb{E}(X_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell})\|_\infty \ll \frac{1}{(i-(m-2)2^\ell)^q}\,. \tag{63}$$

Since $q > 1/2$, the above considerations imply that

$$\sum_{\ell=0}^{r_k-1}\Big(\sum_{m=1}^{2^{r_k-\ell}}\|\mathbb{E}(T_{m2^\ell} - T_{(m-1)2^\ell}|\mathcal{F}_{(m-2)2^\ell+1})\|_r^2\Big)^{1/2} \ll 2^{r_k/2}\sum_{\ell=0}^{r_k-1}2^{\ell/2}2^{-\ell q} \ll m_k^{1/2}\,.$$

Combined with (61) and (62), the upper bound above implies that

$$\Big\|\max_{1\leq\ell\leq 3m_k}\Big|\sum_{i=1}^\ell\mathbb{E}(Y_{k,i}|\mathcal{G}_{k,i-1})\Big|\Big\|_r^r \ll m_k^{r/2}\,. \tag{64}$$

Hence, starting from (58) and taking into account (60) and (64), we get that for any $r \geq 2$,

$$\Big\|\max_{1\leq\ell\leq 3m_k}\big|\widetilde{W}_{k,\ell}\big|\Big\|_r^r \ll m_k^{r/2} + m_k\big(\|X_0\mathbf{1}_{|X_0|\leq M_k}\|_r^r + M_k^r\mathbb{P}(|X_0| > M_k)\big)\,.$$

This implies that (53) holds with $r > \max\big\{2, \varepsilon^{-1}\big(p - 2(1-\varepsilon)\big)\big\}$.

To end the proof it remains to prove condition (55). Note first that since σ^2 is assumed to be strictly positive, we have

$$|\nu_k^{1/2} - \sigma| = \sigma^{-1}\sigma|\nu_k^{1/2} - \sigma| \leq \sigma^{-1}(\nu_k^{1/2} + \sigma)|\nu_k^{1/2} - \sigma| = \sigma^{-1}|\nu_k - \sigma^2|\,,$$

and therefore condition (55) reads as

$$3^k(\nu_k - \sigma^2)^2 = o(3^{2k/p}(\log k)^{-1})\,, \quad \text{as } k \to \infty\,. \tag{65}$$

To verify condition (65), let us define, for $i \geq 0$,

$$\tilde{c}_{k,i} = \text{cov}(\tilde{X}_{k,m_k+1}, \tilde{X}_{k,i+m_k+1}) \quad \text{and} \quad \hat{c}_{k,i} = \text{cov}(X_{k,0}, X_{k,i})\,.$$

20

Using stationarity, we have

$$\mathbb{E}(\widetilde{W}_{k,m_k}^2) = m_k \tilde{c}_{k,0} + 2 \sum_{\ell=1}^{m_k-1} (m_k - \ell)\tilde{c}_{k,\ell} \ \text{and} \ \mathbb{E}(\widetilde{W}_{k,m_k}(\widetilde{W}_{k,2m_k} - \widetilde{W}_{k,m_k})) = \sum_{\ell=1}^{m_k} \ell\tilde{c}_{k,\ell}.$$

Therefore

$$\nu_k = \tilde{c}_{k,0} + 2 \sum_{\ell=1}^{m_k} \tilde{c}_{k,\ell}. \tag{66}$$

We first prove that

$$\left| \nu_k - \left(\hat{c}_{k,0} + 2 \sum_{\ell \geq 1} \hat{c}_{k,\ell}\right) \right| \ll 3^{k/(2p)} m_k^{-q/2} + \left(m_k^{-1/2}(\log m_k)\mathbf{1}_{q=1} + m_k^{-q+1/2}\mathbf{1}_{q<1}\right). \tag{67}$$

With this aim we use the arguments developed in [3] to get their inequality (3.56). Hence, we start by noting that since φ_k is 1-Lipschitz, $\left(\|\mathbb{E}(\varphi_k(X_n)|\mathcal{F}_0 - \mathbb{E}(\varphi_k(X_n)\|_2)\right)_{n\geq 0}$ is a decreasing sequence such that $\|\mathbb{E}(\varphi_k(X_n)|\mathcal{F}_0 - \mathbb{E}(\varphi_k(X_n)\|_2 \leq \delta_\infty(n)$. Hence, by the same arguments as those developed in the first lines of the proof of Theorem 1, we infer that, under condition (11), there exists a constant C not depending on k such that $\sum_{\ell \geq 1} |\hat{c}_{k,\ell}| \leq C$. Therefore, $\lim_{j\to\infty} j^{-1}\mathbb{E}(W_{k,j}^2) = \hat{c}_{k,0} + 2\sum_{\ell \geq 1} \hat{c}_{k,\ell}$. On another hand, the following convergence clearly holds: $\lim_{j\to\infty} j^{-1}\mathbb{E}(\widetilde{W}_{k,j}^2) = \nu_k$. In addition, for all $j \geq 1$,

$$|\mathbb{E}(\widetilde{W}_{k,j}^2) - \mathbb{E}(W_{k,j}^2)| \leq \|\widetilde{W}_{k,j} - W_{k,j}\|_2 \|\widetilde{W}_{k,j} + W_{k,j}\|_2.$$

The above considerations imply

$$\left| \nu_k - \left(\hat{c}_{k,0} + 2 \sum_{\ell \geq 1} \hat{c}_{k,\ell}\right) \right| \leq \limsup_{j\to\infty} j^{-1}\|\widetilde{W}_{k,j} - W_{k,j}\|_2^2 + C \limsup_{j\to\infty} j^{-1/2}\|\widetilde{W}_{k,j} - W_{k,j}\|_2. \tag{68}$$

To take care of $\|\widetilde{W}_{k,j} - W_{k,j}\|_2$, we apply Proposition 23 of the Appendix with, this time, $\alpha = 2$, $r = r_j$ where r_j is the unique positive integer such that $2^{r_j-1} \leq j < 2^{r_j}$,

$$Z_i = 0 \text{ for } i \leq 0, \ Z_i := Z_{i,k} = X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}} \text{ for } i \geq 1,$$

and

$$\mathcal{F}_i = \{\emptyset, \Omega\} \text{ for } i \leq 0, \ \mathcal{F}_i = \sigma(\varepsilon_{i+3^{k-1}}, \ldots, \varepsilon_1, W_0) = \mathcal{G}_{k,i} \text{ for } i \geq 1.$$

Hence

$$\|\widetilde{W}_{k,j} - W_{k,j}\|_2 \ll \left(\sum_{i=1}^{j} \|X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}\|_2^2\right)^{1/2}$$
$$+ \sum_{\ell=0}^{r_k-1} \left(\sum_{m=1}^{2^{r_k-\ell}} \|\mathbb{E}(T_{m2^\ell} - T_{(m-1)2^\ell}|\mathcal{F}_{(m-2)2^\ell+1})\|_2^2\right)^{1/2}, \tag{69}$$

where $T_\ell = \sum_{i=1}^{\ell}(X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}})$. Lemma 24 of the Appendix combined with condition (11) implies that

$$\sum_{i=1}^{j} \|X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}\|_2^2 \leq 2 \times 3^{k/p} \sum_{i=1+3^{k-1}}^{j+3^{k-1}} \|X_{k,i} - \tilde{X}_{k,i}\|_1 \leq \frac{Cj3^{k/p}}{m_k^q}. \tag{70}$$

21

Next, since $\mathcal{F}_i = \{\emptyset, \Omega\}$ for $i \leq 0$, for any $\ell \geq 0$,

$$\|\mathbb{E}(T_{2^\ell}|\mathcal{F}_{-2^\ell+1})\|_2 = |\mathbb{E}(T_{2^\ell})| = 0 \,.$$

Moreover, by (63), we infer that for any $m \geq 2$, any $\ell \geq 0$ and any $i \geq (m-1)2^\ell + 1$,

$$\|\mathbb{E}(X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell+1})\|_2$$
$$\leq \|\mathbb{E}(X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell+1})\|_\infty \ll \frac{1}{(i-(m-2)2^\ell)^q} \ll 2^{-\ell q} \,.$$

On another hand, Lemma 24 of the Appendix combined with condition (11) implies that

$$\|\mathbb{E}(X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell+1})\|_1 \leq \|X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}\|_1 \ll m_k^{-q} \,.$$

Hence, for any $m \geq 2$, any $\ell \geq 0$ and any $i \geq (m-1)2^\ell + 1$, we also have

$$\|\mathbb{E}(X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell+1})\|_2^2 \leq 2^{-\ell q} m_k^{-q} \,.$$

The considerations above imply that, for any $m \geq 2$, any $\ell \geq 0$ and any $i \geq (m-1)2^\ell + 1$,

$$\|\mathbb{E}(X_{k,i+3^{k-1}} - \tilde{X}_{k,i+3^{k-1}}|\mathcal{G}_{k,(m-2)2^\ell+1})\|_2 \leq 2^{-\ell q/2} \min(2^{-\ell q/2}, m_k^{-q/2}) \,.$$

Hence, since $q > 1/2$,

$$\sum_{\ell=0}^{r_j-1} \Big(\sum_{m=1}^{2^{r_j-\ell}} \|\mathbb{E}(T_{m2^\ell} - T_{(m-1)2^\ell}|\mathcal{F}_{(m-2)2^\ell+1})\|_2^2 \Big)^{1/2} \ll 2^{r_j/2} \sum_{\ell=0}^{r_j-1} 2^{\ell/2} 2^{-\ell q/2} \min(2^{-\ell q/2}, m_k^{-q/2})$$
$$\ll j^{1/2} \big(m_k^{-q/2} \mathbf{1}_{q>1} + m_k^{-1/2}(\log m_k)\mathbf{1}_{q=1} + m_k^{-q+1/2}\mathbf{1}_{q<1} \big) \,. \quad (71)$$

Starting from (69) and considering the upper bounds (70) and (71), we get

$$j^{-1/2}\|\widetilde{W}_{k,j} - W_{k,j}\|_2 \ll 3^{k/(2p)} m_k^{-q/2} + \big(m_k^{-1/2}(\log m_k)\mathbf{1}_{q=1} + m_k^{-q+1/2}\mathbf{1}_{q<1} \big) \,. \quad (72)$$

Hence starting from (68) and taking into account (72) together with the fact that $3^{k/(2p)} m_k^{-q/2} \leq 2^{q/2}$, the upper bound (67) follows.

Let now $c_i = \text{Cov}(X_0, X_i)$ and note that (see Relation (3.54) in [3], where the same truncation level is used)

$$\sup_{i \geq 0} |\hat{c}_{k,i} - c_i| = o(3^{-k(p-2)/p}) \,.$$

Let

$$\ell_k = [3^{k(p-2)/(2p)}(\log k)^{-1/2}] \,.$$

Since $\sigma^2 = c_0 + 2\sum_{i \geq 1} c_i$, it follows that

$$\Big| \sigma^2 - \big(\hat{c}_{k,0} + 2\sum_{i \geq 1} \hat{c}_{k,i} \big) \Big| \leq o(\ell_k 3^{-k(p-2)/p}) + 2\sum_{i > \ell_k} |c_i - \hat{c}_{k,i}| \,. \quad (73)$$

But

$$|c_i - \hat{c}_{k,i}| = |\text{Cov}(X_0 - \varphi_k(X_0), X_i) + \text{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))| \,.$$

22

Set $g_k(x) = x - \varphi_k(x)$ and note that, by the reverse martingale convergence theorem and condition (11), $\mathbb{E}(g_k(X_0)|\mathcal{F}_{-\infty}) = \mathbb{E}(g_k(X_0))$ a.s. and $\mathbb{E}(X_0|\mathcal{F}_{-\infty}) = \mathbb{E}(X_0)$ a.s. Hence, applying Lemma 22 of the Appendix and taking into account condition (11), we get

$$\sum_{i > \ell_k} |\operatorname{Cov}(X_0 - \varphi_k(X_0), X_i)| \ll \sum_{j=0}^{\infty} \|P_0(g_k(X_j))\|_2 \sum_{i \geq [2^{-1}(\ell_k+j)]+1} i^{-1/2} \|\mathbb{E}(X_i|V_0)\|_2$$
$$\ll \sum_{j=0}^{\infty} \|P_0(g_k(X_j))\|_2 (\ell_k + j)^{1/2-q} ,$$

where $P_0(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_0) - \mathbb{E}(\cdot|\mathcal{F}_{-1})$. But, by Lemma 22 of the Appendix,

$$\sum_{j=0}^{\infty} \|P_0(g_k(X_j))\|_2 (\ell_k + j)^{1/2-q} \ll \ell_k^{1/2-q} \sum_{j=0}^{\ell_k} \|P_0(g_k(X_j))\|_2 + \sum_{j \geq \ell_k} j^{1/2-q} \|P_0(g_k(X_j))\|_2$$

$$\ll \ell_k^{1/2-q} \sum_{j=0}^{\ell_k} (j+1)^{-1/2} \|\mathbb{E}(g_k(X_j)|V_0) - \mathbb{E}(g_k(X_j))\|_2 + \sum_{j \geq \ell_k} j^{1/2-q} \|P_0(g_k(X_j))\|_2 .$$

Note now that, since $q > 1/2$,

$$\sum_{j \geq \ell_k} j^{1/2-q} \|P_0(g_k(X_j))\|_2 = \sum_{j \geq \ell_k} j^{1/2-q} \|P_0(g_k(X_j))\|_2 \Big(j^{-1} \sum_{m=1}^{j} 1\Big)$$

$$= \sum_{m \geq 1} \sum_{j \geq \max(m, \ell_k)} j^{-(1/2+q)} \|P_0(g_k(X_j))\|_2$$

$$\ll \ell_k^{1-q} \Big(\sum_{j \geq \ell_k} \|P_0(g_k(X_j))\|_2^2\Big)^{1/2} + \sum_{m > \ell_k} m^{-q} \Big(\sum_{j \geq m} \|P_0(g_k(X_j))\|_2^2\Big)^{1/2}$$

$$\ll \ell_k^{1-q} \|\mathbb{E}(g_k(X_{\ell_k})|V_0) - \mathbb{E}(g_k(X_{\ell_k}))\|_2 + \sum_{m > \ell_k} m^{-q} \|\mathbb{E}(g_k(X_m)|V_0) - \mathbb{E}(g_k(X_m))\|_2 . \quad (74)$$

So, overall,

$$\sum_{i > \ell_k} |\operatorname{Cov}(X_0 - \varphi_k(X_0), X_i)| \ll \ell_k^{1/2-q} \sum_{j=0}^{\ell_k} (j+1)^{-1/2} \|\mathbb{E}(g_k(X_j)|V_0) - \mathbb{E}(g_k(X_j))\|_2$$
$$+ \ell_k^{1-q} \|\mathbb{E}(g_k(X_{\ell_k})|V_0) - \mathbb{E}(g_k(X_{\ell_k}))\|_2 + \sum_{m > \ell_k} m^{-q} \|\mathbb{E}(g_k(X_m)|V_0) - \mathbb{E}(g_k(X_m))\|_2 .$$

Next, we note that

$$\|\mathbb{E}(g_k(X_0)|V_0) - \mathbb{E}(g_k(X_0))\|_2 \ll o(3^{-k(p-2)/(2p)}) ,$$

and that, for $j \geq 1$, by condition (11),

$$\|\mathbb{E}(g_k(X_j)|V_0) - \mathbb{E}(g_k(X_j))\|_2 \ll \min\left(j^{-q}, j^{-q/2} 3^{-k(p-1)/(2p)}\right) . \quad (75)$$

23

Hence, since $q > 1/2$, we infer that

$$\sum_{i>\ell_k} |\operatorname{Cov}(X_0 - \varphi_k(X_0), X_i)| \ll \ell_k^{1/2-q} 3^{-k(p-2)/(2p)} + \ell_k^{1-2q} \mathbf{1}_{q>1} + k\, \ell_k^{-1/2} 3^{-k(p-1)/(2p)} \mathbf{1}_{q=1}$$

$$+ \left(k\, 3^{-k(p-1)/(2p)} + \ell_k^{1-3q/2} 3^{-k(p-1)/(2p)} + 3^{k(1-2q)(p-1)/(pq)} \right) \mathbf{1}_{q<1}. \quad (76)$$

We handle now the series

$$\sum_{i>\ell_k} |\operatorname{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))|.$$

Applying again Lemma 22 of the Appendix, we first write that

$$\sum_{i>\ell_k} |\operatorname{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))|$$

$$\ll \sum_{\ell=0}^{\infty} \|P_0(\varphi_k(X_\ell))\|_2 \sum_{i \geq [2^{-1}\ell_k]+1} i^{-1/2} \|\mathbb{E}(g_k(X_i)|\mathcal{F}_0) - \mathbb{E}(g_k(X_i))\|_2.$$

By condition (11) and since $q > 1/2$,

$$\sum_{\ell \geq 0} \|P_0(\varphi_k(X_\ell))\|_2 \ll \sum_{\ell \geq 0} (\ell+1)^{-1/2} \|\mathbb{E}(\varphi_k(X_j)|V_0) - \mathbb{E}(\varphi_k(X_j))\|_2 \leq C.$$

So, taking into account (75) and the fact that $q > 1/2$,

$$\sum_{i>\ell_k} |\operatorname{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))|$$

$$\ll 3^{-k(p-1)/(2p)} \ell_k^{(1-q)/2} \mathbf{1}_{q>1} + 3^{-k(p-1)(2q-1)/(2pq)} \mathbf{1}_{q<1} + k\, 3^{-k(p-1)/(2p)} \mathbf{1}_{q=1}. \quad (77)$$

Considering the upper bounds (73), (76) and (77), we then derive

$$\left| \sigma^2 - (\hat{c}_{k,0} + 2\sum_{i \geq 1} \hat{c}_{k,i}) \right| \leq o(\ell_k 3^{-k(p-2)/p}) + \ell_k^{1/2-q} 3^{-k(p-2)/(2p)} + \left(\ell_k^{1-2q} + 3^{-k(p-1)/(2p)} \ell_k^{(1-q)/2} \right) \mathbf{1}_{q>1}$$

$$+ k 3^{-k(p-1)/(2p)} \mathbf{1}_{q=1} + \left(k\, 3^{-k(p-1)/(2p)} + \ell_k^{1-3q/2} 3^{-k(p-1)/(2p)} + 3^{-k(p-1)(2q-1)/(2pq)} \right) \mathbf{1}_{q<1},$$

which combined with (67) gives

$$\left| \nu_k - \sigma^2 \right| \ll o(\ell_k 3^{-k(p-2)/p}) + \ell_k^{1/2-q} 3^{-k(p-2)/(2p)} + \left(\ell_k^{1-2q} + 3^{-k(p-1)/(2p)} \ell_k^{(1-q)/2} \right) \mathbf{1}_{q>1}$$

$$+ k 3^{-k(p-1)/(2p)} \mathbf{1}_{q=1} + \left(k\, 3^{-k(p-1)/(2p)} + \ell_k^{1-3q/2} 3^{-k(p-1)/(2p)} + 3^{-k(p-1)(2q-1)/(2pq)} \right) \mathbf{1}_{q<1}$$

$$+ 3^{k/(2p)} m_k^{-q/2} + \left(m_k^{-1/2} (\log m_k) \mathbf{1}_{q=1} + m_k^{-q+1/2} \mathbf{1}_{q<1} \right). \quad (78)$$

Let us verify that (65) holds, namely:

$$3^{k(p-2)/(2p)} (\nu_k - \sigma^2) = o((\log k)^{-1/2}).$$

The choice of ℓ_k implies that $\ell_k 3^{-k(p-2)/p} = 3^{-k(p-2)/(2p)} (\log k)^{-1/2}$ and $\ell_k^{1/2-q} = o((\log k)^{-1/2})$ (since $q > 1/2$). Moreover, when $q > 1$, we clearly have $3^{k(p-2)/(2p)} \ell_k^{1-2q} = o((\log k)^{-1/2})$ and

24

$3^{k(p-2)/(2p)}3^{-k(p-1)/(2p)}\ell_k^{(1-q)/2} = o((\log k)^{-1/2})$. It is also clear that $3^{k(p-2)/(2p)}k3^{-k(p-1)/(2p)} = o((\log k)^{-1/2})$. Next, since $q > (p-1)/2$,

$$3^{k(p-2)/(2p)}3^{-k(p-1)(2q-1)/(2pq)}\mathbf{1}_{q<1} \le 3^{k(p-2)/(2p)}3^{-k(p-1)(p-2)/(2p)},$$

proving (since $p > 2$) that $3^{k(p-2)/(2p)}3^{-k(p-1)(2q-1)/(2pq)}\mathbf{1}_{q<1} = o((\log k)^{-1/2})$. Also, since $p > 2$,

$$3^{k(p-2)/(2p)}\ell_k^{1-3q/2}3^{-k(p-1)/(2p)}\mathbf{1}_{(p-1)/2<q<1} \ll 3^{k(p-2)/(2p)}\ell_k^{1/4}3^{-k(p-1)/(2p)}\mathbf{1}_{2<p<3},$$

which proves that $3^{k(p-2)/(2p)}\ell_k^{1-3q/2}3^{-k(p-1)/(2p)}\mathbf{1}_{(p-1)/2<q<1} = o((\log k)^{-1/2})$. Next, we note that

$$3^{k(p-2)/(2p)}\frac{3^{k/(2p)}}{m_k^{q/2}} = 3^{k(p-1)/(2p)}3^{-qk(1-\varepsilon)/p} = o((\log k)^{-1/2}),$$

since $\varepsilon < 1 - \frac{p-1}{2q}$.

Now, if $q = 1$ then $p < 3$ (since $q > (p-1)/2$). Hence since $\varepsilon < 1/2$, we get that $3^{k(p-2)/(2p)}m_k^{-1/2}(\log m_k)\mathbf{1}_{q=1} = o((\log k)^{-1/2})$. Finally, using again that $q > (p-1)/2$ and that $\varepsilon < 1/2$, we derive that $3^{k(p-2)/(2p)}m_k^{-q+1/2}\mathbf{1}_{q<1} = o((\log k)^{-1/2})$. This ends the proof of (65) and then of the theorem. $\square$

4.2 Proof of Theorem 2

By Remark 3, we know that condition (14) is equivalent to (17), namely:

$$\int_0^1 R^{p-1}(u)Q(u)du < \infty,$$

where, for any $u \in [0,1]$, $\gamma^{-1}(u) = \delta^{-1} \circ H(u)$ and $R(u) = \gamma^{-1}(u)Q(u)$.

Notice first that, by Proposition 1 in Dedecker-Doukhan [10],

$$\sum_{i\ge 1}|\text{Cov}(X_0,X_i)| \le 2\sum_{i\ge 1}\int_0^{2^{-1}\|\mathbb{E}(X_i|V_0)-\mathbb{E}(X_i)\|_1} Q \circ H^{-1}(u)du \le \int_0^1 R(u)Q(u)du < \infty,$$

by condition (17). Hence the series $\sigma^2 = \text{Var}(X_1) + 2\sum_{i\ge 1}\text{Cov}(X_1,X_{i+1})$ converge absolutely and condition (54) of Proposition 21 holds.

Assume first that $\sigma^2 > 0$. To prove the theorem, we shall verify that the other conditions of Proposition 21 are satisfied and with this aim we need to define suitable sequences (m_k) and (M_k). Since we have $\text{Var}(S_n)/n \to \sigma^2 > 0$, it follows that $\text{Var}(S_n) \to \infty$. Hence $\mathbb{P}(|X_1| > 0) > 0$ since otherwise we would have $X_1 = 0$ a.s. and then $S_n = 0$ a.s. for all $n \ge 1$, contradicting the fact that $\text{Var}(S_n) \to \infty$. Let $u_1 = (1/2)\mathbb{P}(|X_1| > 0)$ (hence $u_1 > 0$) and define

$$K_0 = \inf\{k \in \mathbb{N} : R(u_1) \le 3^{k/p}\}.$$

Obviously $K_0 < \infty$ since $u_1 > 0$ which implies that $Q(u_1) < \infty$ and $\gamma^{-1}(u_1) < \infty$. Next, for any $k \ge K_0$, let

$$v_k = \inf\{u \in [0,u_1] : R(u) \le 3^{k/p}\} := R^{-1}(3^{k/p}), \quad M_k = Q(v_k),$$

25

and $M_k = 1$ for $0 \le k < K_0$. Since $u_1 < \mathbb{P}(|X_1| > 0)$, it follows that $Q(u_1) > 0$ and therefore since Q is non-increasing and $v_k \le u_1$, $M_k \ge Q(u_1) > 0$, for $k \ge K_0$. Let now, for any $k \ge K_0$,

$$m_k = \inf\{n \ge 0 : \gamma(n) \le v_k\},$$

and $m_k = 1$ for any $1 \le k < K_0$. Since v_k is assumed to be strictly less than 1 (since $v_k \le u_1 \le 1/2$), $m_k \ge 1$ (indeed $\gamma(0) = H^{-1}(\mathbb{E}(|X_1|)) = 1$). In addition, since R is right continuous and non-increasing, $u < R^{-1}(x) \iff R(u) > x$. Hence, $R(R^{-1}(u)) \le u$ for all $u \in [0, 1]$, implying that

$$m_k M_k \le R(v_k) \le 3^{k/p}.$$

Therefore, for any $k \ge K_0$, since $M_k \ge Q(u_1) > 0$,

$$m_k \le (Q(u_1))^{-1} m_k M_k \le 3^{k/p},$$

which proves that $m_k = o(3^{2k/p} k^{-1})$, as $k \to \infty$.

To prove now that the conditions (51), (52), (53) and (55) of Proposition 21 are satisfied, we first notice the following useful facts:

$$Q_{|\varphi_k(X_j)|}(x) := \bar{Q}_k(u) = Q(x \vee v_k) \quad \text{and} \quad Q_{|g_k(X_j)|}(u) := \tilde{Q}_k(u) = Q(x) \mathbf{1}_{x \le v_k}. \tag{79}$$

Let us start by proving that condition (51) holds. By using (79), we get

$$\sum_{k \ge 1} 3^{k(p-1)/p} \mathbb{E}(|g_k(X_1)|) = \sum_{k \ge 1} 3^{k(p-1)/p} \int_0^1 Q_{|g_k(X_1)|}(u) du = \sum_{k \ge 1} 3^{k(p-1)/p} \int_0^1 Q(u) \mathbf{1}_{u < v_k} du.$$

But

$$\sum_{k \ge 1} 3^{k(p-1)/p} \int_0^1 Q(u) \mathbf{1}_{u < v_k} du = \int_0^1 Q(u) \sum_{k \ge 1} 3^{k(p-1)/p} \mathbf{1}_{R(u) > 3^{k/p}} du \ll \int_0^1 R^{p-1}(u) Q(u) du < \infty, \tag{80}$$

by condition (17) (which is equivalent to condition (14)). Hence condition (51) is satisfied. Next we note that by Lemma 24 of the Appendix,

$$\Big\| \max_{1 \le \ell \le 3^k - 3^{k-1}} |W_{k,\ell} - \widetilde{W}_{k,\ell}| \Big\|_1 \le \sum_{i=1+3^{k-1}}^{3^k} \|X_{k,i} - \tilde{X}_{k,i}\|_1 \le 2 \times 3^k \delta(m_k).$$

Therefore, by using (80),

$$\sum_{k \ge 1} 3^{-k/p} \Big\| \max_{1 \le \ell \le 3^k - 3^{k-1}} |W_{k,\ell} - \widetilde{W}_{k,\ell}| \Big\|_1 \le 2 \sum_{k \ge 1} 3^{k(p-1)/p} \delta(m_k)$$

$$\le 2 \sum_{k \ge 1} 3^{k(p-1)/p} H(\gamma(m_k)) \le 2 \sum_{k \ge 1} 3^{k(p-1)/p} \int_0^{v_k} Q(u) du < \infty. \tag{81}$$

Hence, condition (52) is satisfied with $\alpha = 1$. We prove now that we can find a real number $r \in]2, \infty[$ such that (53) holds. With this aim we start by noticing that, for any $r \ge 1$, by

26

Lemma 24 of the Appendix,

$$\Big\| \max_{1 \leq \ell \leq 3m_k} |\widetilde{W}_{k,\ell} - W_{k,\ell}| \Big\|_r \leq \sum_{i=3^{k-1}+1}^{3m_k+3^{k-1}} \|X_{k,i} - \tilde{X}_{k,i}\|_r$$

$$\leq (2M_k)^{(r-1)/r} \sum_{i=3^{k-1}+1}^{3m_k+3^{k-1}} \|X_{k,i} - \tilde{X}_{k,i}\|_1^{1/r} \ll m_k M_k^{(r-1)/r} (\delta(m_k))^{1/r}.$$

Hence, since $m_k M_k \leq 3^{k/p}$, for any $r \geq 1$,

$$\sum_{k \geq \ell_0} \frac{3^k}{3^{kr/p} m_k} \mathbb{E}\Big(\max_{1 \leq \ell \leq 3m_k} |\widetilde{W}_{k,\ell} - W_{k,\ell}|^r \Big) \ll \sum_{k \geq \ell_0} \frac{(m_k M_k)^{r-1}}{3^{k(r-p)/p}} \delta(m_k) \ll \sum_{k \geq \ell_0} 3^{k(p-1)/p} \delta(m_k),$$

which is finite by taking into account (81). Hence to prove that condition (53) holds, it suffices to prove that we can find a real number $r \in]2, \infty[$ such that

$$\sum_{k \geq \ell_0} \frac{3^k}{3^{kr/p} m_k} \mathbb{E}\Big(\max_{1 \leq \ell \leq 3m_k} |W_{k,\ell}|^r \Big) < \infty. \tag{82}$$

To prove (82), we apply the Rosenthal inequality for τ-dependent sequences as given in Corollary 1 in Dedecker-Prieur [12]. Let us first recall the definition of the τ-dependence coefficients: for any random variable Y with values in $\mathbb{R}^\ell$ and any σ-algebra $\mathcal{F}$,

$$\tau(\mathcal{F}, Y) = \sup_{f \in \Lambda_1(\mathbb{R}^\ell)} \Big\| \sup_{h \in \Lambda_1(\mathbb{R})} \Big| \int h(x) P_{f(Y)|\mathcal{F}}(dx) - \int h(x) P_{f(Y)}(dx) \Big| \Big\|_1,$$

where, for any integer $\ell \geq 1$, $\Lambda_1(\mathbb{R}^\ell)$ is the set of 1-Lipschitz function from $\mathbb{R}^\ell$ to $\mathbb{R}$ with respect to the norm $|x - y|_1 \leq \sum_{k=1}^\ell |x_i - y_i|$. Taking $\mathcal{F}_p = \sigma(X_i, i \leq p)$, the coefficients $\tau(i)$ of the sequence $(\varphi_k(X_i))_{i \in \mathbb{Z}}$ are then defined by: for any $i \geq 0$,

$$\tau(i) = \sup_{u > 0} \max_{1 \leq \ell \leq u} \frac{1}{\ell} \sup \big\{ \tau(\mathcal{F}_p, (\varphi_k(X_{j_1}), \dots, \varphi_k(X_{j_\ell}))), \, p + i \leq j_1 < \dots < j_\ell \big\}.$$

In the stationary case, Corollary 1 in Dedecker-Prieur [12] implies that, for any $r > 2$,

$$\mathbb{E}\Big(\max_{1 \leq \ell \leq 3m_k} |W_{k,\ell}|^r \Big) \ll \Big(m_k \int_0^{\|\varphi_k(X_1)\|_1} ((\tau/2)^{-1}(u) \wedge m_k) Q_{|\varphi_k(X_1)|} \circ H_{|\varphi_k(X_1)|}^{-1}(u) du \Big)^{r/2}$$

$$+ m_k \int_0^{\|\varphi_k(X_1)\|_1} ((\tau/2)^{-1}(u) \wedge m_k)^{r-1} Q_{|\varphi_k(X_1)|}^{r-1} \circ H_{|\varphi_k(X_1)|}^{-1}(u) du,$$

where τ^{-1} is the generalized inverse of the function τ defined by $\tau(x) = \tau([x])$.

To compare the coefficients $\tau(i)$ with the coefficients $\delta(i)$, we consider $(W_0', (\varepsilon_j')_{j \geq 1})$ an independent copy of $(W_0, (\varepsilon_j)_{j \geq 1})$ and define $W_1' = F(\varepsilon_1', W_0')$ and $W_m' = F(\varepsilon_m, W_{m-1}')$ for any $m \geq 2$. Note that for any $j \geq 2$, by using the relation (97) of the Appendix, we have

$$X_j = h(\varepsilon_j, W_{j-1}) := h(\varepsilon_j, F_{j-2}(\varepsilon_{j-1}, \dots, \varepsilon_2, W_1)) := g_{j-2}(\varepsilon_j, \varepsilon_{j-1}, \dots, \varepsilon_2, W_1). \tag{83}$$

27

Define now, for any $j \geq 2$,

$$X'_j = h(\varepsilon_j, W'_{j-1}) = h(\varepsilon_j, F_{j-2}(\varepsilon_{j-1}, \ldots, \varepsilon_2, W'_1)) := g_{j-2}(\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_2, W'_1). \tag{84}$$

Clearly for any $2 \leq j_1 < \ldots < j_\ell$, $(\varphi_k(X'_{j_1}), \ldots, \varphi_k(X'_{j_\ell}))$ is distributed as $(\varphi_k(X_{j_1}), \ldots, \varphi_k(X_{j_\ell}))$ and is independent of (ε_0, W_{-1}). Hence, by stationarity and Lemma 3 in Dedecker-Prieur [12],

$$\tau\big(\mathcal{F}_0, (\varphi_k(X_{j_1}), \ldots, \varphi_k(X_{j_\ell}))\big)$$
$$\leq \sup_{f \in \Lambda_1(\mathbb{R}^\ell)} \|f(\varphi_k(X_{j_1}), \ldots, \varphi_k(X_{j_\ell})) - f(\varphi_k(X'_{j_1}), \ldots, \varphi_k(X'_{j_\ell}))\|_1$$
$$\leq \sum_{v=1}^{\ell} \|X_{j_v} - X'_{j_v}\|_1 = \sum_{v=1}^{\ell} \iint \mathbb{E}|X_{j_v-1,x} - X_{j_v-1,y}| d\nu(x) d\nu(y) \leq 2 \sum_{v=1}^{\ell} \delta(j_v),$$

where the second inequality comes from the fact that $f \in \Lambda_1(\mathbb{R}^\ell)$ and φ_k is 1-Lipschitz. Therefore, since δ is non-increasing, for any $i \geq 2$,

$$\tau(i) \leq 2\delta(i).$$

Moreover, for any $i \in \{0, 1\}$, we obviously get that $\tau(i) \leq 2\mathbb{E}(|X_1|) = 2\delta(0)$. It follows that for any $x \geq 0$,

$$\tau(x) := \tau([x]) \leq 2\delta(x).$$

Therefore, since both τ and δ are non-increasing,

$$(\tau/2)^{-1}(u) \leq \delta^{-1}(u).$$

In addition, since φ_k is 1-Lipschitz and such that $\varphi_k(0) = 0$,

$$H_{|\varphi_k(X_1)|}(x) \leq H(x) \text{ and then } H^{-1}(u) \leq H_{|\varphi_k(X)|}^{-1}(u), \tag{85}$$

since H is non-decreasing. Therefore, using additionally the fact that $u < v \iff Q_{|\varphi_k(X_1)|}(v) < Q_{|\varphi_k(X_1)|}(u)$, we get

$$\mathbb{E}\Big(\max_{1 \leq \ell \leq 3m_k} |W_{k,\ell}|^r\Big) \ll \Big(m_k \int_0^{\|X_1\|_1} (\delta^{-1}(u) \wedge m_k) Q_{|\varphi_k(X_1)|} \circ H^{-1}(u) du\Big)^{r/2}$$
$$+ m_k \int_0^{\|X_1\|_1} (\delta^{-1}(u) \wedge m_k)^{r-1} Q_{|\varphi_k(X_1)|}^{r-1} \circ H^{-1}(u) du,$$

and then, since $\gamma^{-1}(u) = \delta^{-1} \circ H(u)$,

$$\mathbb{E}\Big(\max_{1 \leq \ell \leq 3m_k} |W_{k,\ell}|^r\Big) \ll \Big(m_k \int_0^1 (\gamma^{-1}(u) \wedge m_k) Q_{|\varphi_k(X_1)|}(u) Q(u) du\Big)^{r/2}$$
$$+ m_k \int_0^1 (\gamma^{-1}(u) \wedge m_k)^{r-1} Q_{|\varphi_k(X_1)|}^{r-1}(u) Q(u) du.$$

Recall now that $m_k = \gamma^{-1}(v_k)$, therefore since γ^{-1} is non-increasing,

$$\gamma^{-1}(u) \wedge m_k = m_k \mathbf{1}_{u < v_k} + \gamma^{-1}(u) \mathbf{1}_{u \geq v_k}.$$

28

Using also the fact that $Q_{|\varphi_k(X_1)|}(x) = Q(x \vee v_k)$, we get

$$\mathbb{E}\Big(\max_{1 \leq \ell \leq 3m_k} |W_{k,\ell}|^r \Big) \ll \Big(m_k Q^{1-2/r}(v_k) \int_0^1 \gamma^{-1}(u) Q^{1+2/r}(u) du \Big)^{r/2}$$

$$+ m_k (m_k Q(v_k))^{r-1} \int_0^1 Q(u) \mathbf{1}_{u<v_k} du + m_k \int_0^1 R^{r-1}(u) Q(u) \mathbf{1}_{u \geq v_k} du$$

$$:= I_k^{(1)} + I_k^{(2)} + I_k^{(3)} .$$

Using the fact that $m_k Q(v_k) \leq 3^{k/p}$ and (80), we get that, for any $r > 2$,

$$\sum_{k \geq \ell_0} 3^{k(p-r)/p} m_k^{-1} I_k^{(2)} \leq \sum_{k \geq \ell_0} 3^{k(p-1)/p} \int_0^1 Q(u) \mathbf{1}_{u<v_k} du < \infty .$$

On another hand, for any $r > p$,

$$\sum_{k \geq \ell_0} 3^{k(p-r)/p} m_k^{-1} I_k^{(3)} \leq \sum_{k \geq \ell_0} 3^{k(p-r)/p} \int_0^1 R^{r-1}(u) Q(u) \mathbf{1}_{u \geq v_k} du$$

$$= \int_0^1 \sum_{k \geq \ell_0} 3^{k(p-r)/p} R^{r-1}(u) \mathbf{1}_{3^{k/p} \geq R(u)} Q(u) du \ll \int_0^1 R^{p-1} Q(u) du < \infty ,$$

by condition (17) (which is equivalent to condition (14)). Finally using again that $m_k Q(v_k) \leq 3^{k/p}$, we derive that, for any $r > 2(p-1)$,

$$\sum_{k \geq \ell_0} 3^{k(p-r)/p} m_k^{-1} I_k^{(1)} \leq \sum_{k \geq \ell_0} 3^{k(p-r)/p} \Big((m_k Q(v_k))^{1-2/r} \int_0^1 \gamma^{-1}(u) Q^{1+2/r}(u) du \Big)^{r/2}$$

$$\leq \sum_{k \geq \ell_0} 3^{k(p-r)/p} 3^{k(r-2)/(2p)} \Big(\int_0^1 \gamma^{-1}(u) Q^{1+2/r}(u) du \Big)^{r/2} ,$$

since condition (17) obviously implies that $\int_0^1 \gamma^{-1}(u) Q^{1+2/r}(u) du < \infty$. So, overall, (82) holds provided we select $r > 2(p-1)$.

To end the proof it remains to show that condition (55) holds. With this aim, we start by recalling the equation (66), namely:

$$\nu_k = \tilde{c}_{k,0} + 2 \sum_{\ell=1}^{m_k} \tilde{c}_{k,\ell} ,$$

where, for $i \geq 0$,

$$\tilde{c}_{k,i} = \operatorname{cov}(\tilde{X}_{k,m_k+1}, \tilde{X}_{k,i+m_k+1}) \text{ and } \hat{c}_{k,i} = \operatorname{cov}(X_{k,0}, X_{k,i}) .$$

But, by using Lemma 24 of the Appendix, we have, for any $i \geq 0$,

$$|\tilde{c}_{k,i} - \hat{c}_{k,i}| = |\operatorname{cov}(\tilde{X}_{k,m_k+1} - X_{k,m_k+1}, \tilde{X}_{k,i+m_k+1}) + \operatorname{cov}(X_{k,m_k+1}, \tilde{X}_{k,i+m_k+1} - X_{k,i+m_k+1})|$$

$$\leq Q(v_k) \|\tilde{X}_{k,m_k+1} - X_{k,m_k+1}\|_1 + Q(v_k) \|\tilde{X}_{k,i+m_k+1} - X_{k,i+m_k+1}\|_1$$

$$\leq 4 \times Q(v_k) \delta(m_k) \leq 4 \times Q(v_k) \int_0^{\gamma(m_k)} Q(u) du \leq 4 \times Q(v_k) \int_0^{v_k} Q(u) du .$$

29

Hence, since $m_k Q(v_k) \leq 3^{k/p}$,

$$\left| \nu_k - \hat{c}_{k,0} - 2\sum_{\ell=1}^{m_k} \hat{c}_{k,\ell} \right| \leq 8 \times m_k Q(v_k) \int_0^1 Q(u)\mathbf{1}_{u<v_k} du \leq 8 \times 3^{k/p} \int_0^1 Q(u)\mathbf{1}_{R(u)>3^{k/p}} du$$

$$\leq 8 \times 3^{k/p} 3^{-k(p-1)/p} \int_0^1 R^{p-1} Q(u) du \ll 3^{-k(p-2)/p}, \quad (86)$$

by condition (17) (which is equivalent to condition (14)). Taking into account (86) together with the fact that $\sigma^2 = \sum_{k\in\mathbb{Z}} \operatorname{cov}(X_0, X_k)$, we get

$$|\nu_k - \sigma^2| \ll 2\sum_{i=0}^{m_k} |\hat{c}_{k,i} - \operatorname{cov}(X_0, X_i)| + 2\sum_{i \geq m_k+1} |\operatorname{cov}(X_0, X_i)| + 3^{-k(p-2)/p}. \quad (87)$$

Next, by using Proposition 1 in Dedecker-Doukhan [10], we derive

$$\sum_{i \geq m_k} |\operatorname{cov}(X_0, X_i)| \leq 2\sum_{i \geq m_k} \int_0^{\delta(i)} Q \circ H^{-1}(u) du = 2\sum_{i \geq m_k} \int_0^{\gamma(i)} Q^2(u) du.$$

But, since $m_k = \gamma^{-1}(v_k)$, note that

$$\int_0^1 R(u)Q(u)1_{u<v_k} du = \int_0^1 \gamma^{-1}(u)Q^2(u)\mathbf{1}_{u<v_k} du = \sum_{i \geq 0} \int_0^1 Q^2(u)\mathbf{1}_{u<v_k}\mathbf{1}_{u<\gamma(i)} du$$

$$= m_k \int_0^1 Q^2(u)1_{u<v_k} du + \sum_{i \geq m_k} \int_0^1 Q^2(u)\mathbf{1}_{u<\gamma(i)} du. \quad (88)$$

Hence

$$\sum_{i \geq m_k} |\operatorname{cov}(X_0, X_i)| \leq 2\int_0^1 R(u)Q(u)1_{u<v_k} du = 2\int_0^1 R(u)Q(u)1_{R(u)>3^{k/p}} du$$

$$\leq 2 \times 3^{-k(p-2)/p} \int_0^1 R^{p-1}(u)Q(u)1_{R(u)>3^{k/p}} du \ll 3^{-k(p-2)/p}, \quad (89)$$

by condition (17). On another hand, by using inequality (1.11a) in [23] and (79), we derive that, for any $i \geq 0$,

$$|\hat{c}_{k,i} - \operatorname{cov}(X_0, X_i)| = \left| \operatorname{cov}(X_{k,0} - X_0, X_{k,i}) + \operatorname{cov}(X_0, X_{k,i} - X_i) \right|$$

$$\leq 2\int_0^{1/2} Q_{|\varphi_k(X)|}(u)Q_{|g_k(X)|}(u)du + 2\int_0^{1/2} Q_{|X_1|}(u)Q_{|g_k(X)|}(u)du \leq 4\int_0^1 Q^2(u)\mathbf{1}_{u<v_k} du.$$

Hence, by taking into account (88),

$$\sum_{i=0}^{m_k} |\hat{c}_{k,i} - \operatorname{cov}(X_0, X_i)| \leq 4(m_k+1)\int_0^1 Q^2(u)\mathbf{1}_{u<v_k} du \leq 8\int_0^1 R(u)Q(u)1_{u<v_k} du.$$

30

So, by the computations in (89),

$$\sum_{i=0}^{m_k} |\hat{c}_{k,i} - \mathrm{cov}(X_0, X_i)| \ll 3^{-k(p-2)/p}. \tag{90}$$

Hence, starting from (87) and taking into account (89) and (90), it follows that

$$|\nu_k - \sigma^2| \ll 3^{-k(p-2)/p},$$

implying, since $p > 2$, that

$$3^k(\nu_k - \sigma^2)^2 \ll 3^{k(4-p)/p} = o(3^{2k/p}(\log k)^{-1}), \text{ as } k \to \infty.$$

This proves that (65) holds and then that (55) is satisfied since $\sigma^2 > 0$. The proof is complete for the case $\sigma^2 > 0$.

Assume now that $\sigma^2 = 0$. Let M be a positive real number. According to inequality (5.42) in Merlevède-Rio [21], for any positive integer n, any real number λ, and any positive integer $q \le n$ and such that $qM \le \lambda$, we have

$$\mathbb{P}\Big(\max_{1 \le k \le n} |S_k - k\mathbb{E}(X_1)| \ge 5\lambda \Big) \le \frac{\mathrm{Var}(S_n)}{\lambda^2} + 2\lambda^{-1} \sum_{i=1}^{n} \mathbb{E}|g_M(X_i)|$$
$$+ \lambda^{-1} \sum_{i=1}^{n} \|\mathbb{E}(\varphi_M(X_i)|V_{i-q}) - \mathbb{E}(\varphi_M(X_i))\|_1. \tag{91}$$

Choose now $u = R^{-1}(\lambda)$, $q = \gamma^{-1}(u) \wedge n$ and $M = Q(u)$. Since R is right continuous, we have $R(u) \le \lambda$, hence $qM \le R(u) \le \lambda$. Note also that

$$\sum_{k=1}^{n} \mathbb{E}|g_M(X_k)| \le n \int_0^u Q(x)dx \le n \int_0^1 Q(x)\mathbf{1}_{R(x)>\lambda}dx. \tag{92}$$

In addition,

$$\|\mathbb{E}(\varphi_M(X_i)|V_{i-q}) - \mathbb{E}(\varphi_M(X_i))\|_1 \le 2\delta(q) = 2\int_0^{\gamma(q)} Q(x)dx.$$

Since $\gamma(q) \le u$, it follows that

$$\sum_{i=1}^{n} \|\mathbb{E}(\varphi_M(X_i)|V_{i-q}) - \mathbb{E}(\varphi_M(X_i))\|_1 \le 2n \int_0^1 Q(x)\mathbf{1}_{R(x)>\lambda}dx.$$

Starting from (91) and taking into account the considerations above, we get that, for any $\lambda > 0$,

$$\mathbb{P}\Big(\max_{1 \le k \le n} |S_k - k\mathbb{E}(X_1)| \ge 5\lambda \Big) \le \frac{\mathrm{Var}(S_n)}{\lambda^2} + \frac{4n}{\lambda} \int_0^1 Q(x)\mathbf{1}_{R(x)>\lambda}dx.$$

Hence, for any $\varepsilon > 0$, selecting $\lambda = \varepsilon n^{1/p}$, we derive

$$\sum_{n \ge 1} n^{-1}\mathbb{P}\Big(\max_{1 \le k \le n} |S_k - k\mathbb{E}(X_1)| \ge 5\varepsilon n^{1/p} \Big)$$
$$\le \varepsilon^{-2} \sum_{n \ge 1} \frac{\mathrm{Var}(S_n)}{n^{1+2/p}} + 4\varepsilon^{-1} \sum_{n \ge 1} n^{-1/p} \int_0^1 Q(x)\mathbf{1}_{R(x)>\varepsilon n^{1/p}}dx.$$

31

The second series in the right-hand side is finite under condition condition (17) (which is equivalent to condition (14)). Hence, if we can prove that

$$\sum_{n\geq 1}\frac{\text{Var}(S_n)}{n^{1+2/p}}<\infty\,,\tag{93}$$

then we will get that, for any $\varepsilon>0$,

$$\sum_{n\geq 1}n^{-1}\mathbb{P}\Big(\max_{1\leq k\leq n}|S_k-k\mathbb{E}(X_1)|\geq \varepsilon n^{1/p}\Big)<\infty$$

which will imply $S_n-n\mathbb{E}(X_1)=o(n^{1/p})$ a.s. and therefore the proof of the theorem will be complete. In the case where $p\geq 3$, (93) is almost immediate. To see this, we first note that condition (17) implies $\sum_{i\geq 1}i|\text{Cov}(X_0,X_i)|<\infty$. Indeed, by Proposition 1 in Dedecker-Doukhan [10],

$$\sum_{i\geq 1}i|\text{Cov}(X_0,X_i)|\leq 2\sum_{i\geq 1}i\int_0^{2^{-1}\|\mathbb{E}(X_i|V_0)-\mathbb{E}(X_i)\|_1}Q\circ H^{-1}(u)du\leq \int_0^1 R^2(u)Q(u)du\,,$$

which is finite under condition (17). Therefore, by Lemma 1 in Bradley [5], $\text{Var}(S_n)$ is bounded which obviously entails (93). To handle the case where $p\in]2,3[$, we first note that, by inequality (4.84) in [19],

$$\|\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k)\|_2^2\leq \mathbb{E}\big|X_k(\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k))\big|\leq 3\int_0^{\|\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k)\|_1}Q\circ H^{-1}(u)du\,.$$

But, $\|\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k)\|_1\leq 2\delta(k)$. Hence

$$\|\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k)\|_2^2\leq 6\int_0^{\delta(k)}Q\circ H^{-1}(u)du\,.$$

Hence condition (14) entails

$$\|\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k)\|_2\ll k^{-(p-1)/2}\,,$$

which implies (since $p>2$) that

$$\sum_{k\geq 1}k^{-1/2}\|\mathbb{E}(X_k|V_0)-\mathbb{E}(X_k)\|_2<\infty\ \text{ and therefore }\ \sum_{k\geq 1}\frac{\|\mathbb{E}(S_k|V_0)-k\mathbb{E}(X_1)\|_2}{k^{3/2}}<\infty\,.\tag{94}$$

We use now the same arguments as developed at the beginning of the proof of Theorem 1. The fact that the series in (94) converge implies that there exists a stationary sequence $(D_k)_{k\in\mathbb{Z}}$ of martingale differences in $\mathbb{L}^2$, adapted to $(\mathcal{F}_k)_{k\in\mathbb{Z}}$ and such that

$$\lim_{n\to\infty}n^{-1/2}\Big\|S_n-n\mathbb{E}(X_1)-\sum_{k=1}^n D_k\Big\|_2=0\,.$$

32

Together with the fact that $\lim_{n\to\infty} n^{-1}\mathrm{Var}(S_n) = \sigma^2 = 0$, it follows that $D_k = 0$ a.s, for any k. Hence, using the upper bound (4) in Cuny-Merlevède [8] (see also Proposition 1 in [18]), it follows that, for any $p \in]2,3[$,

$$\mathrm{Var}(S_n) \ll n\Big(\sum_{k\geq n} \frac{\|\mathbb{E}(S_k|V_0) - k\mathbb{E}(X_1)\|_2}{k^{3/2}}\Big)^2$$

$$\ll n\Big(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} \|\mathbb{E}(X_k|V_0) - \mathbb{E}(X_k)\|_2 + \sum_{k\geq n} \frac{\|\mathbb{E}(X_k|V_0) - \mathbb{E}(X_1)\|_2}{k^{1/2}}\Big)^2 \ll n^{3-p}\,.$$

Therefore, for any $p \in]2,3[$,

$$\sum_{n\geq 1} \frac{\mathrm{Var}(S_n)}{n^{1+2/p}} \ll \sum_{n\geq 1} \frac{1}{n^{p+2/p-2}}\,,$$

which is finite since $p + 2/p - 3 = p^{-1}(p-1)(p-2) > 0$. This ends the proof of the theorem. $\square$

5 Appendix

5.1 Some technical results

In this section, we collect some technical results that are useful for the proofs of Theorems 1 and 2.

Lemma 22 *Let $(Y_k)_{k\in\mathbb{Z}}$ be a stationary sequence of real-valued random variables adapted to an increasing and stationary filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}}$. Let f and g be two functions in $\mathbb{L}^2(\mathbb{R}, P_{Y_0})$ such that $\mathbb{E}(f(Y_0)|\mathcal{F}_{-\infty}) = \mathbb{E}(f(Y_0))$ a.s. and $\mathbb{E}(g(Y_0)|\mathcal{F}_{-\infty}) = \mathbb{E}(g(Y_0))$ a.s. Then, for any positive integer L,*

$$\sum_{i\geq L} |\mathrm{cov}(f(Y_0), g(Y_i))| \leq 3\sqrt{2} \sum_{\ell=0}^{\infty} \|P_0(f(Y_\ell))\|_2 \Big(\sum_{k\geq [(L+\ell)/2]+1} k^{-1/2} \|\mathbb{E}(g(Y_k)|\mathcal{F}_0) - \mathbb{E}(g(Y_k))\|_2\Big)$$

and

$$\sum_{i=0}^{L} \|P_0(g(Y_i))\|_2 \leq \sqrt{2} \sum_{k=0}^{L} (k+1)^{-1/2} \|\mathbb{E}(g(Y_k)|\mathcal{F}_0) - \mathbb{E}(g(Y_k))\|_2\,,$$

where $P_j(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_j) - \mathbb{E}(\cdot|\mathcal{F}_{j-1})$.

Proof. Since $\mathbb{E}(f(Y_0)|\mathcal{F}_{-\infty}) = \mathbb{E}(f(Y_0))$ a.s. and $\mathbb{E}(g(Y_0)|\mathcal{F}_{-\infty}) = \mathbb{E}(g(Y_0))$ a.s., we first write

$$f(Y_0) - \mathbb{E}(f(Y_0)) = \sum_{\ell=0}^{\infty} P_{-\ell}(f(Y_0)) \text{ and } g(Y_i) - \mathbb{E}(g(Y_i)) = \sum_{\ell=-i}^{\infty} P_{-\ell}(g(Y_i)) \text{ a.s.}$$

Hence, by orthogonality, for any $i \geq 0$,

$$|\mathrm{cov}(f(Y_0), g(Y_i))| \leq \sum_{\ell=0}^{\infty} |\mathbb{E}(P_{-\ell}(f(Y_0))P_{-\ell}(g(Y_i)))|\,,$$

33

and then, by Cauchy-Schwarz's inequality and stationarity,

$$\sum_{i\geq L}|\operatorname{cov}(f(Y_0),g(Y_i))|\leq\sum_{\ell=0}^{\infty}\|P_0(f(Y_\ell))\|_2\sum_{i\geq L+\ell}\|P_0(g(Y_i))\|_2. \tag{95}$$

But, for any $m\geq 1$, by Cauchy-Schwarz's inequality,

$$\sum_{i\geq m}\|P_0(g(Y_i))\|_2=\sum_{i\geq m}i^{-1}\|P_0(g(Y_i))\|_2\sum_{k=1}^{i}=\sum_{k\geq 1}\sum_{i\geq\max(m,k)}i^{-1}\|P_0(g(Y_i))\|_2$$

$$=\sum_{k=1}^{m}\sum_{i\geq m}i^{-1}\|P_0(g(Y_i))\|_2+\sum_{k>m}\sum_{i\geq k}i^{-1}\|P_0(g(Y_i))\|_2$$

$$\leq\sqrt{2m}\Big(\sum_{i\geq m}\|P_0(g(Y_i))\|_2^2\Big)^{1/2}+\sqrt{2}\sum_{k>m}k^{-1/2}\Big(\sum_{i\geq k}\|P_0(g(Y_i))\|_2^2\Big)^{1/2}, \tag{96}$$

giving

$$\sum_{i\geq m}\|P_0(g(Y_i))\|_2\leq\sqrt{2m}\|\mathbb{E}(g(Y_m)|\mathcal{F}_0)-\mathbb{E}(g(Y_m))\|_2+\sqrt{2}\sum_{k>m}k^{-1/2}\|\mathbb{E}(g(Y_k)|\mathcal{F}_0)-\mathbb{E}(g(Y_k))\|_2.$$

Since $(\|\mathbb{E}(g(Y_k)|\mathcal{F}_0)-\mathbb{E}(g(Y_k))\|_2)_{k\geq 0}$ is non-increasing, we get that for any $m\geq 1$,

$$\sum_{i\geq m}\|P_0(g(Y_i))\|_2\leq 3\sqrt{2}\sum_{k\geq [m/2]+1}k^{-1/2}\|\mathbb{E}(g(Y_k)|\mathcal{F}_0)-\mathbb{E}(g(Y_k))\|_2,$$

which combined with (95) gives the first inequality of the lemma. To prove the second one, it suffices to write that $\sum_{i=0}^{L}\|P_0(g(Y_i))\|_2=\sum_{i=0}^{L}(i+1)^{-1}\|P_0(g(Y_i))\|_2\big(\sum_{k=1}^{i+1}1\big)$ and to use Cauchy-Schwarz's inequality as in (96). □

The following proposition is a non stationary version of the Peligrad-Utev-Wu [22] inequality. As in [22], the proof can be done by induction (a complete proof appears in Section 3.2.1 of [20]).

Proposition 23 *Let $\alpha\geq 2$ and $(Z_k)_{k\in\mathbb{Z}}$ be a sequence of real-valued random variables in $\mathbb{L}^\alpha$ and adapted to a non-decreasing filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}}$. Then, for any $n\geq 1$,*

$$\Big\|\max_{1\leq k\leq n}\Big|\sum_{i=1}^{k}Z_i\Big|\Big\|_\alpha\leq(2c_\alpha+1)\Big(\sum_{j=1}^{n}\|Z_j\|_\alpha^2\Big)^{1/2}$$

$$+\sqrt{2}(2c_\alpha+1)\sum_{\ell=0}^{r-1}\Big(\sum_{m=1}^{2^{r-\ell}}\|\mathbb{E}(S_{m2^\ell}-S_{(m-1)2^\ell}|\mathcal{F}_{(m-2)2^\ell+1})\|_\alpha^2\Big)^{1/2},$$

where $S_k=\sum_{i=1}^{k}Z_i$, $c_\alpha=\frac{\alpha}{(\alpha-1)^{1/2}}$ if $\alpha>2$, $c_2=1$ and r is the unique positive integer such that $2^{r-1}\leq n<2^r$.

Lemma 24 *For any $q\in[1,p)$, for any $k\geq 1$ and any $j\geq m_k+1$,*

$$\|X_{k,j}-\tilde{X}_{k,j}\|_q^q\leq\iint\mathbb{E}(|X_{m_k+1,x}-X_{m_k+1,y}|^q)\nu(dx)\nu(dy),$$

where $X_{k,j}$ and $\tilde{X}_{k,j}$ are defined in (48) and (49) respectively.

34

Proof. Let $(W_0', (\varepsilon_j')_{j\geq 1})$ be an independent copy of $(W_0, (\varepsilon_j)_{j\geq 1})$ and define $W_j' = F(\varepsilon_j', W_{j-1}')$, $j \geq 1$. For $\ell \geq 1$, let F_ℓ be the function from $G^\ell \times X$ to X defined in an iterative way as follows

$$F_1 = F \text{ and } F_\ell(x_1, x_2, \ldots, x_{\ell+1}) = F_{\ell-1}(x_1, x_2, \ldots, x_{\ell-1}, F(x_\ell, x_{\ell+1})), \ell \geq 2\,.$$

Note that for any integer ℓ such that $1 \leq \ell \leq j-1$,

$$W_{j-1} = F_\ell(\varepsilon_{j-1}, \varepsilon_{j-2}, \ldots, \varepsilon_{j-\ell}, W_{j-\ell-1})\,. \tag{97}$$

Hence, for any $j \geq m_k + 1$,

$$\begin{aligned}
\mathbb{E}\big(&\varphi_k(X_j)|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}\big) \\
&= \mathbb{E}\big(\varphi_k(h(\varepsilon_j, F_{m_k}(\varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1})))|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}\big) \\
&= \mathbb{E}\big(\varphi_k(h(\varepsilon_j, F_{m_k}(\varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1}')))|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_1, W_0\big)\,.
\end{aligned}$$

On another hand, for any $j \geq 1$,

$$\begin{aligned}
\varphi_k(X_j) &= \mathbb{E}\big(\varphi_k(X_j)|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_1, W_0\big) \\
&= \mathbb{E}\big(\varphi_k(h(\varepsilon_j, F_{m_k}(\varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1})))|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_1, W_0\big)\,.
\end{aligned}$$

Hence, for any $j \geq m_k + 1$,

$$\begin{aligned}
\|X_{k,j} &- \tilde{X}_{k,j}\|_q \\
&\leq \|\varphi_k(h(\varepsilon_j, F_{m_k}(\varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1}))) - \varphi_k(h(\varepsilon_j, F_{m_k}(\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1}')))\|_q \\
&\leq \|h(\varepsilon_j, F_{m_k}(\varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1})) - h(\varepsilon_j, F_{m_k}(\varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}, W_{j-m_k-1}'))\|_q\,,
\end{aligned}$$

where the second inequality comes from the fact that φ_k is 1-Lipschitz. By stationarity, it follows that

$$\|X_{k,j} - \tilde{X}_{k,j}\|_q \leq \|h(\varepsilon_{m_k+1}, F_{m_k}(\varepsilon_{m_k}, \ldots, \varepsilon_1, W_0)) - h(\varepsilon_{m_k+1}, F_{m_k}(\varepsilon_{m_k}, \ldots, \varepsilon_1, W_0'))\|_q\,.$$

Hence, if we define $(X_n^*)_{n\geq 1}$ by

$$X_n^* = h(\varepsilon_n, W_{n-1}^*) \text{ where } W_n^* = F(\varepsilon_n, W_{n-1}^*)\,,$$

with W_0^* independent of $(W_0, (\varepsilon_k)_{k\geq 1})$ and such that $W_0^* =^{\mathcal{L}} W_0$, we get that for any $j \geq m_k + 1$,

$$\|X_{k,j} - \tilde{X}_{k,j}\|_q \leq \|X_{m_k+1} - X_{m_k+1}^*\|_q\,. \tag{98}$$

But,

$$\begin{aligned}
\|X_{m_k+1} - X_{m_k+1}^*\|_q^q &= \iint \mathbb{E}(|X_{m_k+1} - X_{m_k+1}^*|^q|W_0 = x, W_0^* = y)\nu(dx)\nu(dy) \\
&= \iint \mathbb{E}(|X_{m_k+1,x} - X_{m_k+1,y}|^q)\nu(dx)\nu(dy)\,,
\end{aligned}$$

which combined with (98) gives the lemma. $\square$

35

5.2 Proof of Lemma 17

The first inequality in (43) comes from the coupling inequality (38) and the fact that $\liminf_{n\to\infty} n^a\beta(n) > 0$ (see Theorem 9.4 in Rio [23]). We prove now the second inequality in (43).

Let $W_{n,x}$ be the chain starting at x. Note first that for any any $x, y \in [0, 1]$,

$$
\mathbb{P}_{x,y}(T^* > n) = \mathbb{P}_{x,y}\big(T^* > n, \{W_{n,x} = x \cup W_{n,y} = y\}\big) + \mathbb{P}_{x,y}\big(T^* > n, \{W_{n,x} \neq x, W_{n,y} \neq y\}\big)
$$
$$
\leq (1-x)^n + (1-y)^n + \mathbb{P}_{x,y}\big(T^* > n, \{W_{n,x} \neq x, W_{n,y} \neq y\}\big).
$$

But

$$
\mathbb{P}_{x,y}\big(T^* > n, \{W_{n,x} \neq x, W_{n,y} \neq y\}\big)
$$
$$
= \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \mathbb{P}_{x,y}\big(T^* > n, W_{n,x} = F_\pi^{-1}(V_i), W_{n,y} = F_\pi^{-1}(V_j)\big).
$$

For $j > i$, define $\mathcal{W}_{i,j} := \bigcap_{k=i}^{j}\{W_{k,x} \neq W_{k,y}\}$, $\mathcal{E}_{i,j}(x) := \bigcap_{k=i}^{j}\{W_{k,x} = F_\pi^{-1}(V_i)\}$, and note that

$$
\mathbb{P}_{x,y}\big(T^* > n, W_{n,x} = F_\pi^{-1}(V_i), W_{n,y} = F_\pi^{-1}(V_j)\big)
$$
$$
= (a+1)^2 \iint_{[0,1]^2} \mathbb{P}_{x,y}\Big(\mathcal{W}_{1,i-1}, \mathcal{E}_{i,n}(x), \mathcal{W}_{i,j-1}, \mathcal{E}_{j,n}(y)\big| F_\pi^{-1}(V_i) = u, F_\pi^{-1}(V_j) = v\Big) u^a v^a du dv
$$
$$
\leq (a+1)^2 \iint_{[0,1]^2} \mathbb{P}_{x,y}\Big(\mathcal{W}_{1,i}, \{U_i < W_{i-1,x}\}, \bigcap_{k=i+1}^{j} \{U_k \geq u\}, \bigcap_{k=j+1}^{n} \{U_k \geq u \vee v\}\Big) u^a v^a du dv
$$
$$
\leq (a+1)^2 \iint_{[0,1]^2} \mathbb{P}_{x,y}(T^* > i, W_{i,x} = F_\pi^{-1}(V_i)) \mathbb{P}\Big(\bigcap_{k=i+1}^{j} \{U_k \geq u\}, \bigcap_{k=j+1}^{n} \{U_k \geq u \vee v\}\Big) u^a v^a du dv.
$$

So, overall, setting $w_i(x,y) := \mathbb{P}_{x,y}(T^* > i, W_{i,x} = F_\pi^{-1}(V_i))$,

$$
\mathbb{P}_{\nu\otimes\nu}(T^* > n) \leq 2a \int_0^1 (1-x)^n x^{a-1} dx
$$
$$
+ 2(a+1)^2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \nu \otimes \nu(w_i) \iint_{[0,1]^2} \mathbb{P}\Big(\bigcap_{k=i+1}^{j} \{U_k \geq u\}, \bigcap_{k=j+1}^{n} \{U_k \geq u \vee v\}\Big) u^a v^a du dv.
$$

Using the fact that for any $b > -1$,

$$
\int_0^1 (1-x)^k x^b dx \leq k^{-(b+1)} \int_0^k e^{-x} x^b dx \leq k^{-(b+1)} \Gamma(b+1), \tag{99}
$$

we get that

$$
\mathbb{P}_{\nu\otimes\nu}(T^* > n) \leq 2a\Gamma(a) n^{-a} + 2(a+1)^2 \sum_{i=1}^{n-1} \nu \otimes \nu(w_i) \sum_{j=i+1}^{n} \iint_{[0,1]^2} (1-u)^{j-i} (1-u\vee v)^{n-j} u^a v^a du dv. \tag{100}
$$

By easy computations (that are left to the reader), we infer that Lemma 17 will hold provided one can prove that:

36

Lemma 25 *For any $a > 1$, there exists a positive constant $\kappa(a)$ depending only on a such that for any $n \geq 1$,*

$$\nu \otimes \nu(w_n) \leq \frac{\kappa(a)}{n^a} \,. \tag{101}$$

Obviously, inequality (101) holds for any positive integer $n \leq \kappa(a)$. It is then enough to prove it for $n > \kappa(a)$. Let us do it by recurrence. Hence we assume that for any $k \leq n-1$, $\nu \otimes \nu(w_k) \leq \kappa(a)k^{-a}$ and we want to prove it at step n. With this aim, we argue as above and infer that

$$w_n(x,y) \leq (1-y)^n + (a+1) \sum_{i=1}^{n-1} w_i(y,x) \int_{[0,1]} (1-u)^{n-i} u^a du \,.$$

Hence,

$$\nu \otimes \nu(w_n) \leq a\Gamma(a)n^{-a} + (a+1) \sum_{i=1}^{n-1} \nu \otimes \nu(w_i) \int_{[0,1]} (1-u)^{n-i} u^a du \,.$$

Using the recurrence assumption, it follows that

$$\sum_{i=1}^{n-1} \nu \otimes \nu(w_i) \int_{[0,1]} (1-u)^{n-i} u^a du \leq \sum_{i=1}^{[n/2]} \int_{[0,1]} (1-u)^{n-i} u^a du$$

$$+ \frac{\kappa(a)}{([n/2]+1)^a} \sum_{i=[n/2]+1}^{n-[\log n]} \int_{[0,1]} (1-u)^{n-i} u^a du$$

$$+ \frac{\kappa(a)}{(n-[\log n]+1)^a} \sum_{i=n-[\log n]+1}^{n-1} \int_{[0,1]} (1-u)^{n-i} u^a du \,.$$

Then, taking into account (99), we infer that

$$(a+1) \sum_{i=1}^{n-1} \nu \otimes \nu(w_i) \int_{[0,1]} (1-u)^{n-i} u^a du$$

$$\leq 2^a(a+1)\frac{\Gamma(a)}{n^a} + +2^a(a+1)\frac{\kappa(a)}{n^a[\log n]^a} + a^{-1}\frac{\kappa(a)}{(n-[\log n]+1)^a} \,.$$

So, overall, since $n \geq \kappa(a)$, we get

$$\nu \otimes \nu(w_n) \leq \kappa(a)\rho(a)n^{-a} \,,$$

where

$$\rho(a) := \big(a + 2^a(a+1)\big)\frac{a\Gamma(a)}{\kappa(a)} + \frac{2^a(a+1)}{[\log \kappa(a)]^a} + a^{-1}\Big(1 - \frac{\log \kappa(a)}{\kappa(a)}\Big)^{-a} \,.$$

So choosing $\kappa(a)$ large enough so that $\rho(a) \leq 1$ (which is always possible since $a^{-1} < 1$), inequality (101) is proved at step n which ends the recurrence. $\square$

37

5.3 Proof of Lemma 20

We do the proof in the case where h is 1-Lipschitz. We start by noting that, for any $n \geq 2$,

$$\delta(n) \leq \iint \mathbb{E}|W_{n-1,x} - W_{n-1,y}|\nu(dx)\nu(dy).$$

Recall now the inequality line 5 page 27 of Dedecker-Rio [13], which holds for every $x, y \in \mathbb{R}$, every $n \geq 1$ and any $t > 0$,

$$|W_{n,x} - W_{n,y}|^t \leq \alpha^{nt}(|x| + |y| + \Sigma_{n-1})|x - y|^t, \qquad (102)$$

where $\alpha(t) = 1 - \frac{C}{(1+t)^\tau}$, for every $t \geq 0$, $\Sigma_0 = 0$ and $\Sigma_n = |\varepsilon_1| + \cdots |\varepsilon_n|$, for every $n \geq 1$.

Denote $\upsilon := \mathbb{E}(|\varepsilon_1|)$ and let $0 < \eta \leq 1/\tau - 1$. Notice that α is non-decreasing and bounded by 1. Hence, for any $n \geq 1$, using that $n \leq n^{1/\tau - \eta}$, we get

$$\alpha^{nt}(|x| + |y| + \Sigma_{n-1}) \leq \mathbf{1}_{\{\Sigma_{n-1} > n\upsilon + n^{1/\tau - \eta}\}} + \alpha^{nt}(2(1 + \upsilon)n^{1/\tau - \eta})\mathbf{1}_{\{(1+\upsilon)n^{1/\tau - \eta} \geq |x| + |y|\}}$$
$$+ \alpha^{nt}(2(|x| + |y|))\mathbf{1}_{\{(1+\upsilon)n^{1/\tau - \eta} < |x| + |y|\}}.$$

By Theorems 3 and 4 in Baum and Katz [1], since μ has a moment of order S,

$$\sum_{n \geq 1} n^\gamma \mathbb{P}(\Sigma_{n-1} > n\upsilon + n^{1/\tau - \eta}) < \infty, \qquad (103)$$

provided that $\gamma \leq S(1/\tau - \eta) - 2$. Since $S/\tau - 2 \geq t/\tau + \gamma$, the latter holds as soon as $\eta \leq t/(S\tau)$. Hence, we choose $\eta = \min(t/(S\tau), 1/\tau - 1)$. On another hand,

$$\sum_{n \geq 1} n^\gamma \alpha^{nt}(2(1 + \upsilon)n^{1/\tau - \eta}) = \sum_{n \geq 1} n^\gamma \Big(1 - \frac{C}{(1 + 2(1 + \upsilon)n^{1/\tau - \eta})^\tau}\Big)^{nt} < \infty. \qquad (104)$$

Finally,

$$\sum_{n \geq 1} n^\gamma \alpha^{nt}(2(|x| + |y|))\mathbf{1}_{\{|x| + |y| > 1\}} = \sum_{n \geq 1} n^\gamma \Big(1 - \frac{C}{(1 + 2(|x| + |y|))^\tau}\Big)^{nt}\mathbf{1}_{\{|x| + |y| > 1\}}$$
$$\leq \sum_{n \geq 1} n^\gamma e^{-Cnt/(3(|x| + |y|))^\tau} \leq D(|x|^{\tau(\gamma+1)} + |y|^{\tau(\gamma+1)}), \qquad (105)$$

where D is a constant depending on γ, t and C. Starting from (102) and taking into account (103), (104) and (105) together with the fact that, by (46), ν has a moment of order $S - \tau$ and that $S - \tau \geq \tau(\gamma + 1) + t$, we get the desired result. $\qquad \square$

References

[1] Baum, L. and Katz, M. Convergence rates in the law of large numbers. *Bull. Amer. Math. Soc.* **69** (1963), 771-772.

[2] Y. Benoist and J.-F. Quint, *Central limit theorem for linear groups*, Ann. Probab. 44 (2016), no. 2, 1308-1340.

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

[4] P. Bougerol and J. Lacroix, Products of random matrices with applications to Schrödinger operators. Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, 1985

[5] Bradley, R. C. On quantiles and the central limit question for strongly mixing sequences. *J. Theor. Probab.* **10** (1997), 507-555.

[6] Bradley, R. C. *Introduction to strong mixing conditions*. Vol. 1,2,3. Kendrick Press, Heber City, UT, 2007.

[7] Cuny, C., Dedecker, J. and Jan, C. Limit theorems for the left random walk on $GL_d(\mathbb{R})$. (2017). *hal-01283929*. To appear in *Ann. Inst. H. Poincar Probab. Statist.*

[8] Cuny, C. and Merlevède, F. On martingale approximations and the quenched weak invariance principle. *Ann. Probab.* **42** (2014), no. 2, 760-793.

[9] Cuny, C. and Merlevède, F. Strong invariance principles with rate for "reverse" martingale differences and applications. *J. Theoret. Probab.* **28** (2015), no. 1, 137-183.

[10] Dedecker, J. and Doukhan, P. A new covariance inequality and applications. *Stochastic Process. Appl.* **106** (2003), no. 1, 63-80.

[11] 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.

[12] Dedecker, J. and Prieur, C. Coupling for τ-dependent sequences and applications. *J. Theoret. Probab.* **17** (2004), no. 4, 861-885.

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

[14] Furstenberg, H. and Kesten, H. Products of Random Matrices. *Ann. Math. Statist.* **31** (1960), no. 2, 457-469.

[15] Komlós, J., Major, P. and Tusnády, G. An approximation of partial sums of independent RVs, and the sample DF. II. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **34** (1976), 33-58.

[16] Lindvall, T. On Coupling of Discrete Renewal Processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **48** (1979), 57-70.

[17] Major, P. The approximation of partial sums of independent RVs. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **35** (1976), 213-220.

[18] Merlevède, F., Peligrad, C. and Peligrad, M. Almost sure invariance principles via martingale approximation. *Stochastic Process. Appl.* **122** (2012), no. 1, 170-190.

[19] Merlevède, F. and Peligrad, M. On the weak invariance principle for stationary sequences under projective criteria. *J. Theoret. Probab.* **19** (2006), no. 3, 647-689.

[20] Merlevède, F., Peligrad, M. and Utev, S. *Functional Gaussian Approximation for Dependent structures*, 2017, submitted.

[21] 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.

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

[23] Rio, E. *Théorie asymptotique des processus aléatoires faiblement dpendants*, Mathématiques & Applications (Berlin), vol. 31, Springer-Verlag, Berlin, 2000.

[24] Sakhanenko, A. I. Estimates in the invariance principle in terms of truncated power moments. *Sibirsk. Mat. Zh.* **47** (2006), 1355-1371.

[25] Shao, X. and Wu, W. B. Limit theorems for iterated random functions. *J. Appl. Prob.* **41** (2004), 425-436.