An alternative to the coupling of Berkes-Liu-Wu for strong approximations

Christophe Cuny, Jérôme Dedecker, Florence Merlevède

An alternative to the coupling of Berkes-Liu-Wu for strong approximations

Christophe Cuny[a], Jérôme Dedecker[b] and Florence Merlevède[c]

July 20, 2017

Abstract

In this paper we propose an alternative to the coupling of Berkes, Liu and Wu [1] to obtain strong approximations for partial sums of dependent sequences. The main tool is a new Rosenthal type inequality expressed in terms of the coupling coefficients. These coefficients are well suited to some classes of Markov chains or dynamical systems, but they also give new results for smooth functions of linear processes.

[a] Université de la Nouvelle-Calédonie, Institut de Sciences Exactes et Appliquées.
Email: christophe.cuny@univ-nc.nc

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

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

1 Introduction

Let $(\varepsilon_i)_{i\in\mathbb{Z}}$ be a sequence of independent and identically distributed (iid) random variables, and let $(X_n)_{n\in\mathbb{Z}}$ be a strictly stationary sequence such that

$$X_n = f(\dots,\varepsilon_0,\dots,\varepsilon_{n-1},\varepsilon_n)\,, \tag{1}$$

for some real-valued measurable function f.

In 2005, Wu [9] introduced the so-called *physical dependence measure* defined in terms of the following coupling: let ε_0' be distributed as ε_0 and independent of $(\varepsilon_i)_{i\in\mathbb{Z}}$, and let

$$\tilde{X}_n = f(\dots,\varepsilon_{-1},\varepsilon_0',\varepsilon_1,\dots,\varepsilon_{n-1},\varepsilon_n)\,. \tag{2}$$

The physical dependence coefficents in $\mathbb{L}^p$ (assuming that $\|X_0\|_p^p = \mathbb{E}(|X_0|^p) < \infty$) are then given by

$$\delta_p(n) = \|X_n - \tilde{X}_n\|_p\,.$$

As pointed out by Wu, the coefficient $\delta_p(n)$ can be computed for a large variety of examples, including iterated random functions and functions of linear processes. As we shall see, it is particularly easy to compute when X_n is a smooth function of a linear process.

Let $S_n = \sum_{k=1}^n X_k$. In a recent paper, Berkes, Liu and Wu [1] use the coupling defined above to prove the following strong approximation result: under an appropriate polynomial decay of the coefficients $\delta_p(n)$, the sequence $n^{-1}\mathbb{E}\big((S_n - n\mathbb{E}(X_1))^2\big)$ converges to σ^2 as $n\to\infty$ and, if

$\sigma^2 > 0$, 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,

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

The proof (from which we extract Proposition 13, Section 4) is based on a approximation by m-dependent sequences combined with an application of a deep result by Sakhanenko [6].

This is a very important result, because it gives a full extension of the Komlos, Major and Tusnady (KMT) strong approximation [5] for partial sums of iid random variables in $\mathbb{L}^p$ (for which $\delta_p(n) = 0$ if $n \geq 1$). Most of the previous results in the dependent context were limitated to the rate $n^{1/4}$, because they were based on the Skorokhod representation theorem for martingales. As an exception, let us mention the paper [7], where the rate $O(\log n)$ is reached for bounded observables of geometrically ergodic Markov chains.

In this paper, we follow the main steps of the proof of Berkes, Liu and Wu [1], but we use a different coupling. Let $(\varepsilon_i')_{i\in\mathbb{Z}}$ be an independent copy of $(\varepsilon_i)_{i\in\mathbb{Z}}$ and let

$$X_n^* = f(\ldots, \varepsilon_{-2}', \varepsilon_{-1}', \varepsilon_0', \varepsilon_1, \ldots, \varepsilon_{n-1}, \varepsilon_n)\,.$$

Our coupling coefficient in $\mathbb{L}^p$ is then defined as

$$\tilde{\delta}_p(n) = \|X_n - X_n^*\|_p$$

As δ_p, this coefficient can be computed for a large class of examples (see Section 3).

Notice that $\mathbb{E}(X_n|\varepsilon_1,\ldots,\varepsilon_n) = \mathbb{E}(\tilde{X}_n|\varepsilon_1,\ldots,\varepsilon_n)$ a.s., and that $\|X_n - \mathbb{E}(X_n|\varepsilon_1,\ldots,\varepsilon_n)\|_p \leq \tilde{\delta}_p(n)$, from which we easily deduce that $\delta_p(n) \leq 2\tilde{\delta}_p(n)$. Moreover, it seems natural to think that, in many situations, the coefficient $\delta_p(n)$ should be much smaller than $\delta_p'(n)$, because X_n differs from $\tilde{X}_n$ by changing only the coordinate at point 0, while all the coordinates before time 0 have been changed in X_n^*. For Markov chains, however, the two coefficients should be of the same order (we shall give in Section 2 an alternative definition of $\tilde{\delta}_p(n)$, which is more adapted to the Markovian setting).

Since $\delta_p(n) \leq 2\tilde{\delta}_p(n)$, a reasonable question is then: what could be the interest to deal with $\tilde{\delta}_p(n)$? The answer is simple : the coupling X_n^* is often easier to handle than X_n' (because X_n^* is by definition independent of the past σ-algebra $\mathcal{F}_0 = \sigma(\varepsilon_i, i \leq 0)$) and we can develop specific tools involving the coefficent $\tilde{\delta}_p(n)$. In this paper, we shall prove and use a new Rosenthal-type inequality (see Section 5.1) expressed in terms of the coefficients $\tilde{\delta}_2$ and $\tilde{\delta}_p$. As a consequence, the conditions that we impose on $\tilde{\delta}_p(n)$ are weaker than the corresponding conditions on $\delta_p(n)$ in the paper by Berkes, Liu and Wu. The two results are not comparable, but we shall obtain better conditions in all the cases where $\delta_p(n)$ and $\tilde{\delta}_p(n)$ are exactly of the same order (for instance in the case of Markov chains).

Let us present a simple example where our conditions are less restrictive than those of Berkes, Liu and Wu. Assume that

$$X_n = g\left(\sum_{i=0}^{\infty} a_i \varepsilon_{n-i}\right)$$

where $(a_i)_{i\geq 0} \in \ell_1$, and $(\varepsilon_i)_{i\in\mathbb{Z}}$ is a sequence of iid random variables in $\mathbb{L}^p$. Here g is a continuous function such that

$$|g(x) - g(y)| \leq c(|x-y|)$$

where c is a non-decreasing concave function and $c(0) = 0$ (c is then a concave majorant of the modulus of continuity of g). In that case, using Lemma 5.1 in [4], it is easy to see that

$$\delta_p(n) \leq \|c(|a_n(\varepsilon_0 - \varepsilon_0')|)\|_p \leq c(2\|\varepsilon_0\|_p|a_n|)\,,$$

and

$$\tilde{\delta}_p(n) \le \left\| c\left(\left|\sum_{i=n}^{\infty} a_i(\varepsilon_{n-i}-\varepsilon'_{n-i})\right|\right)\right\|_p \le c\left(\left\|\sum_{i=n}^{\infty} a_i(\varepsilon_{n-i}-\varepsilon'_{n-i})\right\|_p\right) \le c\left(C_p\|\varepsilon_0\|_p\sqrt{\sum_{i=n}^{\infty} a_i^2}\right) ,$$

where we have used Burkholder's inequality for the last upper bound (the positive constant C_p depends only on p). As expected, we see that the upper bound for $\delta_p(n)$ is smaller than the upper bound for $\delta'_p(n)$.

Let us consider now the case where $|a_i| = O(i^{-a})$ for some $a > 1$, and $c(x) \le C|x|^\beta$ in a neighborhood of 0 for some $\beta \in (0, 1]$. In that case, the conditions of Berkes, Liu and Wu on δ_p hold provided

$$a > \frac{2}{\beta} \quad \text{for } p \in (2, 4], \text{ and } \quad a > \frac{\tau(p)+1}{\beta} \quad \text{for } p > 4,$$

where

$$\tau(p) := \frac{(p-2)\sqrt{p^2+20p+4}+p^2-4}{8p} ,$$

while our condition on $\tilde{\delta}_p$ are satisfied as soon as

$$a > \frac{\kappa(p)}{\beta} + \frac{1}{2} ,$$

where, for $p > 2$,

$$\kappa(p) := \frac{(p-2)\sqrt{p^2+12p+4}+p^2+4p-4}{8p} . \tag{3}$$

As one can see, our condition on a is always less restrictive: for $p \in (2, 4]$ it suffices to notice that κ is increasing and $\kappa(4) < 1.4$. For $p > 4$, it suffice to notice that $\kappa(p) < \tau(p) + 0.5$. Note that, since $\kappa(p) \to 1/2$ as $p \to 2$, in the case where $\beta = 1$ (Lipschitz observables), we only need $a > 1$ and a moment of order $b > 2$ for ε_0 to get a strong approximation of order $o(n^{1/(2+\epsilon)})$ for some $\epsilon > 0$.

In addition to this example of functions of linear processes, we shall apply our main results to some classes of Markov chains or dynamical systems. The Markov chains we shall consider are not (or have no reasons to be) irreducible, and some kind of regularity on the observables is required (as in the previous example). We shall express these regularity conditions in terms of the modulus of continuity (or $\mathbb{L}^p$-modulus of continuity) of the observables. These examples of Markov chains are different from the examples we considered in the previous paper [3], where we used the coefficient $\tilde{\delta}_1$. On the one hand, the coupling coefficient $\tilde{\delta}_1(n)$ can be computed for a larger class of examples, but on the other hand we need to impose a moment condition related to p and to the decay rate of $\tilde{\delta}_1(n)$ to get the strong approximation with rate $o(n^{1/p})$.

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 $a_n \le Cb_n$, for all positive integers n.

2 Main results

Before giving our first main result, let us give the appropriate definition of the coefficient $\tilde{\delta}_p$ when $(X_n)_{n\in\mathbb{Z}}$ is a stationary sequence such that

$$X_n := f(\varepsilon_n, \varepsilon_{n+1}, \ldots) \tag{4}$$

for some measurable real-valued function f. This representation will play an important role in the application to certain non-invertible dynamical systems.

Recall that $(\varepsilon_i)_{i\geq 0}$ is a sequence of iid random variables, and that $(\varepsilon'_i)_{i\geq 0}$ is an independent copy of $(\varepsilon_i)_{i\geq 0}$. Define then $\tilde{X}_{1,n} := f(\varepsilon_1, \ldots, \varepsilon_n, \varepsilon'_{n+1}, \varepsilon'_{n+2}, \ldots)$. Then, for every $p \geq 1$, the coefficient $\tilde{\delta}_p(n)$ is defined by:

$$\tilde{\delta}_p(n) := \|X_1 - \tilde{X}_{1,n}\|_p \,. \tag{5}$$

Recall also that, for any $p > 2$, the fonction $\kappa(p)$ has been defined in (3).

Theorem 1 *Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary sequence defined by either* (1) *or* (4)*, and assume that X_0 has a moment of order $p > 2$. Assume in addition that there exists a positive constant c such that for any $n \geq 1$,*

$$\tilde{\delta}_p(n) \leq cn^{-\gamma} \,, \tag{6}$$

for some $\gamma > \kappa(p)$. Let $S_n = \sum_{k=1}^n X_k$. Then $n^{-1}\mathbb{E}\big((S_n - n\mathbb{E}(Y_1))^2\big) \to \sigma^2$ as $n \to \infty$ and one can redefine $(Y_n)_{n\geq 0}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\geq 1}$ with common distribution $\mathcal{N}(0, \sigma^2)$, such that,

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

Remark 2 *Concerning the function κ, note that $\kappa(p) < (p+4)/4$ and that the function $p \to (p+4)/4$ is an asymptot of κ as $p \to \infty$.*

As quoted in the introduction, we shall now consider the case where the variables X_n are functions of random iterates. In that case the representation (1) is not necessarily appropriate (nor even easy to establish), and we need to define an appropriate coefficient δ'_p similar to $\tilde{\delta}_p$.

Let $(\varepsilon_i)_{i\geq 1}$ be iid random variables with values in a measurable space G and common distribution μ. Let W_0 be a random variable 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

$$W_n = F(\varepsilon_n, W_{n-1}) \,, \tag{7}$$

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 $(X_n)_{n\geq 1}$ is 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.

Let W_0 and W_0^* be two 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

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

We then define the coefficients $(\delta'_p(n))_{n\geq 0}$ as follows

$$\delta'_p(0) := \|X_1\|_p \text{ and } \delta'_p(n) := \sup_{k\geq n} \|X_k - X_k^*\|_p, \; n \geq 1 \,. \tag{9}$$

It is not difficult to see that, for any positive integer n,

$$\delta'_p(n) \leq 2\|X_n - X_n^*\|_p \,. \tag{10}$$

For such functions of random iterates, the following counterpart of Theorem 1 holds:

Theorem 3 *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 there exists a positive constant c such that for any $n \geq 1$,*

$$\delta'_p(n) \leq cn^{-\gamma}\,, \tag{11}$$

for some $\gamma > \kappa(p)$. 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,

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

3 Applications

3.1 Applications to contracting iterated random functions.

We use the notations from the second part of Section 2, with a Markov chain $(W_n)_{n\geq 1}$ defined by the recursive equation (7) and a sequence $(X_n)_{n\geq 1}$ defined by (8). Assume that X is equipped with a metric d and that it is endowed with the corresponding Borel σ-algebra. Let us fix a "base point" $x_0 \in X$. For every $x \in X$, write $\chi(x) := 1 + d(x_0, x)$.

Let us assume that there exists $C > 0$, $\rho \in (0, 1)$ and $\alpha \geq 1$, such that

$$\int_G \big(\chi(F(g, x_0))\big)^{\alpha} \mu(dg) < \infty\,, \tag{12}$$

and

$$\mathbb{E}\big(d(W_{n,x}, W_{n,y})^{\alpha}\big) \leq C\rho^n (d(x, y))^{\alpha}\,, \tag{13}$$

where $W_{n,x}$ is the chain defined by (7) starting from $W_0 = x$.

The next lemma is a combination of Theorem 2 and Lemma 1 of Shao and Wu [8].

Lemma 4 *Assume that* (12) *and* (13) *hold for some $\alpha \geq 1$, $C > 0$ and $\rho \in (0, 1)$. Then the Markov chain $(W_n)_{n\in\mathbb{N}}$ admits a stationary distribution ν such that $\int_X (\chi(x))^{\alpha}\nu(dx) < \infty$. Moreover, there exist $C(\alpha) > 0$ and $\rho(\alpha) \in (0, 1)$, such that for every $n \geq 1$,*

$$\iint \mathbb{E}\big((d(W_{n,x}, W_{n,y}))^{\alpha}\big)\nu(dx)\nu(dy) \leq C(\alpha)(\rho(\alpha))^n\,.$$

From now, the sequence $(X_n)_{n\geq 1}$ is defined by (8), where the chain $(W_n)_{n\geq 0}$ is strictly stationary, with stationary distribution ν.

We shall say that a function $h : G \times X \to \mathbb{R}$ satisfies the assumption $H_{s,t}$, for some $s, t \geq 0$ if there exist non-negative functions η and $\tilde{\eta}$ and a non-decreasing function $\beta : [0, 1] \to [0, +\infty)$ such that, for every $(g, x) \in G \times X$,

$$\big|h(g, x)\big| \leq \eta(g)(\chi(x))^s\,, \tag{14}$$

and for every $u \in (0, 1]$,

$$\sup_{y\in X\,:\,d(x,y)\leq u} \big|h(g, x) - h(g, y)\big| \leq \beta(u)\tilde{\eta}(g)(\chi(x))^t\,. \tag{15}$$

Lemma 5 *Let $p \geq 1$. Assume that* (12) *and* (13) *hold for some $\alpha \geq 1$. Assume that* (14) *and* (15) *hold for some $0 \leq s < \alpha/p$, $0 \leq t \leq \alpha/p$ and some $\eta, \tilde{\eta}$ such that $\int_G (\eta(g))^p \mu(dg) < \infty$ and $\int_G (\tilde{\eta}(g))^p \mu(dg) < \infty$. Then, there exist $0 < \omega_1, \omega_2 < 1$ and $C > 0$, such that*

$$\delta'_p(n) \leq C(\beta(\omega_1^n) + \omega_2^n) \qquad \forall n \geq 1\,. \tag{16}$$

Proof. Let $\varepsilon > 0$, and let $u_n(x,y) := \left|h(\varepsilon_{n+1}, W_{n,x}) - h(\varepsilon_{n+1}, W_{n,y})\right|$. Writing,

$$u_n^p(x,y) = u_n^p(x,y)\mathbf{1}_{\{d(W_{n,x},W_{n,y})\le\rho^{\varepsilon n}\}} + u_n^p(x,y)\mathbf{1}_{\{d(W_{n,x},W_{n,y})>\rho^{\varepsilon n}\}}\,, \tag{17}$$

we obtain the upper bound

$$\iint \mathbb{E}(u_n^p(x,y))\,\nu(dx)\nu(dy) \le I_n + II_n\,.$$

Clearly,

$$I_n \le (\beta(\rho^{\varepsilon n}))^p \left(\int_G (\tilde{\eta}(g))^p \mu(dg)\right)\left(\int_X (\chi(x))^{pt}\nu(dx)\right)\,.$$

Moreover, using Hölder's inequality and Lemma 4, we have

$$\begin{aligned} II_n &\\ \le \frac{2^{p-1}}{\rho^{\varepsilon(\alpha-sp)n}} &\left(\int_G \eta^p(g)\mu(dg)\right)\iint \mathbb{E}\left[(\chi^{sp}(W_{n,x}) + \chi^{sp}(W_{n,y}))\,(d(W_{n,x},W_{n,y}))^{\alpha-sp}\right]\,\nu(dx)\nu(dy)\\ &\le \frac{C}{\rho^{\varepsilon(\alpha-sp)n}}\left(\iint \mathbb{E}(d(W_{n,x},W_{n,y}))^{\alpha}\,\nu(dx)\nu(dy)\right)^{1-sp/\alpha} \le \tilde{C}\left(\frac{(\rho(\alpha))^{1/\alpha}}{\rho^{\varepsilon}}\right)^{(\alpha-sp)n}\,. \end{aligned}$$

The desired bound follows by taking ε small enough. □

Proposition 6 *Let $1 < p < \alpha$. Assume that there exist $C > 0$ and $\rho \in (0,1)$ such that (12) and (13) hold. Let $0 \le s < \alpha/p$ and $0 \le t \le p/\alpha$. Let h satisty $H_{s,t}$. Assume that $\int_G (\tilde{\eta}(g))^p \mu(dg) < \infty$, $\int_G (\eta(g))^p \mu(dg) < \infty$ and that $\beta(2^{-n}) = O(n^{-\gamma})$, with $\gamma > \kappa(p)$. Then, the conclusion of Theorem 3 holds.*

Proof. Starting from Theorem 3 and Lemma 5, it is enough to prove that our assumption implies that, for any $a \in (0,1)$, $\beta(a^n) = O(n^{-\gamma})$. Too see this, we note that there exists an integer $\ell \ge 1$ such that $a^\ell \le 1/2$. Let $n \ge 2\ell$. Notice that $n/(2\ell) \le [n/\ell] \le n/\ell$. Since β is non-decreasing, we have

$$\beta(a^n) \le \beta(2^{-[n/\ell]}) \le C([n/\ell])^{-\gamma} \le C(2\ell/n)^{\gamma}\,,$$

and the result follows. □

3.2 Applications to dilating endomorphisms of the torus

Let A be an $m \times m$ matrix with integral entries. Then, A induces a transformation θ_A of the m-dimensional torus $\mathbb{T}_m := \mathbb{R}^m/\mathbb{Z}^m$ preserving the Haar measure λ.

Assume that A is dilating, i.e. that all its eigenvalues have modulus strictly greater than one. Let Γ be a system of representative of $\mathbb{Z}^m/A\mathbb{Z}^m$. Then, θ_A admits a Perron-Frobenius operator P_A given by

$$P_A f(x) = \frac{1}{N}\sum_{\gamma\in\Gamma} f(A^{-1}x + A^{-1}\gamma)\,, \tag{18}$$

for every continuous function f on $\mathbb{T}_m$, where $N = |\det A| = \#\Gamma$.

Since P_A is markovian, there exists a Markov chain with state space $\mathbb{T}_m$ admitting λ as stationary distribution. This Markov chain may be realized as follows: let W_0 be a random variable taking values in $\mathbb{T}_m$ and $(\varepsilon_i)_{i\ge1}$ be iid variables uniformly distributed on Γ and independent of

W_0. For every $n \geq 1$, define $W_n := A^{-1}W_{n-1} + A^{-1}\varepsilon_n$. Denote by $(W_{n,x})_{n\geq 0}$ the Markov chain starting at $x \in \mathbb{T}_m$.

Let h be some measurable function from $\mathbb{T}_m$ to $\mathbb{R}$, and let $X_n = h(W_n)$ where W_0 has distribution λ. Let also $X_{n,x} = h(W_{n,x})$.

For every $p \geq 1$ and every $f \in \mathbb{L}^p(\lambda)$ the $\mathbb{L}^p$-modulus of continuity of f is given by

$$\omega_{p,f}(\delta) := \sup_{|x|\leq\delta} \|f(\cdot + x) - f\|_p \quad \forall\, 0 \leq \delta \leq 1\,,$$

where $|\cdot|$ stands for the euclidean norm.

Lemma 7 *Let $p \geq 1$ and $h \in \mathbb{L}^p(\lambda)$. The following upper bound holds:*

$$\left(\iint \mathbb{E}(|X_{n,x} - X_{n,y}|^p)\, \lambda(dx)\lambda(y)\right)^{1/p} \leq 2^{m/p}\omega_{p,h}\left(\Delta\left(A^{-n}([0,1]^m)\right)\right)\,,$$

where $\Delta\left(A^{-n}([0,1]^m)\right)$ stands for the diameter of $A^{-n}([0,1]^m)$. Consequently (using (10)*), the coefficients $\delta'_p(n)$ of the stationary sequence $(X_n)_{n\in\mathbb{Z}}$ satisfy*

$$\delta'_p(n) \leq 2^{m/p+1}\omega_{p,h}\left(\Delta\left(A^{-n}([0,1]^m)\right)\right)\,.$$

Proof. We start by some preliminary considerations. Iterating the recursive equation $W_{n,x} = A^{-1}W_{n-1,x} + A^{-1}\varepsilon_n$, we get that

$$W_{n,x} = A^{-n}x + \sum_{i=1}^{n} A^{-i}\varepsilon_{n-i+1}\,.$$

Note that the random variable $W_{n,x}$ has the same distribution as $Y_{1,x}$, where $Y_{1,x}$ is the first iteration of the Markov chain starting at x with transition $P_A^n = P_{A^n}$. As explained at the beginning of this section, this may be realized as

$$Y_{1,x} = A^{-n}x + A^{-n}\xi_1\,,$$

where ξ_1 is uniformly distributed over Γ_n (a system of representative of $\mathbb{Z}^m/A^n\mathbb{Z}^m$). Let $Z_{1,x} = h(Y_{1,x})$ It follows that

$$\iint \mathbb{E}(|X_{n,x} - X_{n,y}|^p)\, \lambda(dx)\lambda(y) = \iint \mathbb{E}(|Z_{1,x} - Z_{1,y}|^p)\, \lambda(dx)\lambda(y)\,.$$

From this last equality, we see that it suffices to prove Lemma 7 for $n = 1$, the general case then follows by considering A^n rather than A.

We refer to [2] for the results that we need about tiling. There exists a unique compact set $K \subset \mathbb{R}^m$, such that

$$K = \cup_{\gamma\in\Gamma}(A^{-1}K + A^{-1}\gamma) \tag{19}$$

and an integer $q \geq 1$ such that

$$\sum_{\underline{n}\in\mathbb{Z}^m} \mathbf{1}_{K+\underline{n}} = q \qquad \lambda\text{-almost everywhere.}$$

Moreover, for every $\gamma, \gamma' \in \Gamma$ with $\gamma \neq \gamma'$, $\lambda\big((A^{-1}K + A^{-1}\gamma) \cap (A^{-1}K + A^{-1}\gamma')\big) = 0$. Using that

$$\mathbf{1}_K = \sum_{\underline{n}\in\mathbb{Z}^m} \mathbf{1}_{(K\cap([0,1]^m-\underline{n})} = \sum_{\underline{n}\in\mathbb{Z}^m} \mathbf{1}_{((K+\underline{n})\cap[0,1]^m)-\underline{n}} \qquad \lambda\text{-almost everywhere}\,,$$

we then infer that for every $\mathbb{Z}^m$-periodic locally integrable function g on $\mathbb{R}^m$,

$$\int_K g\,d\lambda = \int_{\mathbb{R}^m} \Big(\sum_{\underline{n}\in\mathbb{Z}^m} \mathbf{1}_{((K+\underline{n})\cap[0,1]^m)-\underline{n}} \Big) g\,d\lambda = \sum_{\underline{n}\in\mathbb{Z}^m} \int_{(K+\underline{n})\cap[0,1]^m} g\,d\lambda = q \int_{\mathbb{T}_m} g\,d\lambda\,. \qquad (20)$$

Let $h \in \mathbb{L}^p(\lambda)$ (we identify h with a $\mathbb{Z}^m$-periodic function on $\mathbb{R}^m$). We have

$$\begin{aligned}
\iint \mathbb{E}(|X_{1,x} - X_{1,y}|^p)\,\lambda(dx)\lambda(y) &= \iint \mathbb{E}(|h(A^{-1}x + A^{-1}\varepsilon_1) - h(A^{-1}y + A^{-1}\varepsilon_1)|^p)\,\lambda(dx)\lambda(dy) \\
&= \int_{\mathbb{T}_m} \left(\int_{\mathbb{T}_m - x} \mathbb{E}(|h(A^{-1}x + A^{-1}\varepsilon_1) - h(A^{-1}(x+y) + A^{-1}\varepsilon_1)|^p)\,\lambda(dy) \right) \lambda(dx) \\
&= \int_{[-1,1]^m} \left(\int_{(\mathbb{T}_m - y)\cap\mathbb{T}_m} \mathbb{E}(|h(A^{-1}x + A^{-1}\varepsilon_1) - h(A^{-1}(x+y) + A^{-1}\varepsilon_1)|^p)\,\lambda(dx) \right) \lambda(dy)\,.
\end{aligned}$$

Set

$$\begin{aligned}
\psi_y(x) &:= \mathbb{E}(|h(A^{-1}x + A^{-1}\varepsilon_1) - h(A^{-1}(x+y) + A^{-1}\varepsilon_1)|^p) \\
&= \frac{1}{N} \sum_{\gamma\in\Gamma} |h(A^{-1}x + A^{-1}\gamma) - h(A^{-1}(x+y) + A^{-1}\gamma)|^p\,.
\end{aligned}$$

Notice that ψ_y is $\mathbb{Z}^m$-periodic. Hence, using (20) and (19), we have

$$\begin{aligned}
\int_{(\mathbb{T}_m - y)\cap\mathbb{T}_m} \psi_y(x)\lambda(dx) &\le \int_{\mathbb{T}_m} \psi_y(x)\lambda(dx) = \frac{1}{q} \int_K \psi_y(x)\lambda(dx) \\
&= \frac{1}{q} \int_{\cup_{\gamma\in\Gamma}(A^{-1}K + A^{-1}\gamma)} |h(x) - h(x + A^{-1}y)|^p \lambda(dx) = \frac{1}{q} \int_K |h(x) - h(x + A^{-1}y)|^p \lambda(dx) \\
&= \int_{\mathbb{T}_m} |h(x) - h(x + A^{-1}y)|^p \lambda(dx)\,,
\end{aligned}$$

and the result follows. □

We shall now explain how to obtain the strong approximation result with rate $o(n^{1/p})$ for the partial sums of the process $(h \circ \theta_A^n)_{n\in\mathbb{N}}$ for $h \in \mathbb{L}^p(\lambda)$. Let $(\varepsilon_n)_{n\ge0}$ be a sequence of iid variables uniformly distributed on Γ. We define a probability ν on $\mathbb{T}_m$ by setting, for every $f \in C([0,1])$,

$$\int_{\mathbb{T}_m} f\,d\nu := \mathbb{E}\left(f\left(\sum_{k\ge0} A^{-k-1}\varepsilon_k \right) \right)\,.$$

By construction, ν is P_A-invariant. Since A is dilating, the only P_A-invariant probability on $\mathbb{T}_m$ is λ.

Define $Z_0 := \sum_{k\ge0} A^{-k-1}\varepsilon_k$ and for every $n \ge 1$, (with equality in $\mathbb{T}_m$)

$$Z_n := A^n Z_0 = \sum_{k\ge0} A^{n-k-1}\varepsilon_k = \sum_{k\ge n} A^{n-k-1}\varepsilon_k = \sum_{k\ge0} A^{-k-1}\varepsilon_{k+n}\,.$$

Notice that for any $h \in \mathbb{L}^p(\lambda)$ the processes $(h \circ \theta_A^n)_{n\ge0}$ (under λ) and $(Y_n)_{n\in\mathbb{N}} := (h(Z_n))_{n\ge0}$ (under $\mathbb{P}$) have the same distribution.

Let $\tilde\delta_p(n)$ be the coefficients associated with $(Y_n)_{n\ge0}$ as in (5). The computations done in the proof of Lemma 7 yield to the following bound

$$\tilde{\delta}_p(n) \le 2^{m/p+1}\omega_{p,h}\left(\Delta\left(A^{-n}([0,1]^m)\right)\right) . \tag{21}$$

As a consequence of Lemma 7 and of (21), Theorem 1 (applied to $(h(Z_n))_{n\ge 0}$) or Theorem 3 (applied to $(h(W_n))_{n\ge 0}$), lead to the following proposition:

Proposition 8 *Let $p > 2$ and let $\kappa(p)$ be defined in (3). Let $h \in \mathbb{L}^p(\lambda)$ be such that $\omega_{p,h}(2^{-n}) \le O(n^{-\gamma})$ for some $\gamma > \kappa(p)$. Assume that, with the above notations, $S_n = X_1 + \cdots + X_n = h(W_1) + \cdots + h(W_n)$, or $S_n = h(Z_1) + \cdots + h(Z_n)$. Then $n^{-1}\mathbb{E}\big((S_n - n\lambda(h))^2\big) \to \sigma^2$ as $n \to \infty$ and for every (fixed) $x \in [0,1]$, one can redefine $(S_n)_{n\ge 1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\ge 1}$ with common distribution $\mathcal{N}(0,\sigma^2)$, such that,*

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

Remark 9 *Alternatively, one can also apply Theorem 2 in [3], by using the upper bound on $\delta_1'(n)$ given in Lemma 7. For instance, if h is bounded and such that $\sum_{n>0} n^{p-2}\omega_{1,h}(2^{-n}) < \infty$, then the conclusion of Proposition 8 holds. If $m = 1$ (for instance for the transformation $\theta(x) = 2x - [2x]$), this implies that, for BV-observables, the strong approximation holds with the rate $o(n^{1/p})$ for any $p > 2$.*

3.3 Applications to dilating piecewise affine maps

Let K be a countable set, with $|K| \ge 2$, and $(I_k)_{k\in K}$ be a collection of disjoint open subintervals of $[0,1]$ such that $\bigcup_{k\in K} \overline{I}_k = [0,1]$ (it is possible to have several accumulation points). Notice that $\sum_{k\in K} \lambda(I_k) = 1$, where λ stands for the Lebesgue measure on $[0,1]$.

Let $T : [0,1] \to [0,1]$ be a map such that $T_{|I_k}$ is affine and onto $(0,1)$, so that $T_{|I_k}$ extends in a trivial way to an affine map from $\overline{I}_k$ onto $[0,1]$, that we still denote by $T_{|I_k}$. The values of T on $[0,1]\setminus\bigcup_{k\in K} I_k$ will be irrelevant in the sequel. For every $k \in K$, denote by s_k the inverse of $T_{|I_k}$ from $(0,1)$ onto I_k. There exist reals α_k and β_k, such that for every $x \in (0,1)$, $s_k(x) = \alpha_k x + \beta_k$ (hence $T_{|I_k}(u) = (u - \beta_k)/\alpha_k$). Then, $|\alpha_k| = \lambda(I_k)$.

Such a map T admits a Perron-Frobenius operator P defined by

$$Pf(x) := \sum_{k\in K} |\alpha_k| f(\alpha_k x + \beta_k) ,$$

for every continuous function f on $[0,1]$.

Since P is Markovian and leaves λ invariant, there exists a Markov chain with state space $[0,1]$ admitting λ as stationary distribution. Since $|K| \ge 2$ then $0 < |\alpha_k| < 1$ for every $k \in K$ and one may easily prove that λ is the only P-invariant measure on $[0,1]$.

The above Markov chain may be realized as follows. Let W_0 be a random variable taking values in $[0,1]$. Let $(\varepsilon_i)_{i\ge 1}$ be iid random variables independent of W_0, taking values in K, such that $\mathbb{P}(\varepsilon_1 = k) = |\alpha_k|$ for every $k \in K$. For every $n \ge 1$, set $W_n := s_{\varepsilon_n}(W_{n-1})$ and denote by $(W_{n,x})_{n\ge 0}$ the Markov chain starting from $x \in [0,1]$. Notice that for every $n \ge 1$ and every $x \in [0,1]$,

$$W_{n,x} = s_{\varepsilon_n} \circ \cdots \circ s_{\varepsilon_1}(0) + \alpha_{\varepsilon_n} \ldots \alpha_{\varepsilon_1} x := B_n + A_n x .$$

Let h be some measurable function from $[0,1]$ to $\mathbb{R}$, and let $X_n = h(W_n)$ where W_0 has distribution λ. Let also $X_{n,x} = h(W_{n,x})$.

For every $f \in C([0,1])$ define

$$\omega_{\infty,f}(\delta) = \sup_{x,y\in[0,1],\,|x-y|\le\delta} |f(x) - f(y)| , \quad \forall \delta \in [0,1] .$$

Define also

$$\delta_\infty(n) := \sup_{x,y\in[0,1]} \mathbb{E}|X_{n,x} - X_{n,y}| .$$

Lemma 10 *Let $h \in C([0,1])$, and let $\bar\alpha := \max_{k\in K} |\alpha_k|$. For every integer $n \geq 1$, we have*

$$\sup_{x,y\in[0,1]} |X_{n,x} - X_{n,y}| \leq 2\omega_{\infty,h}(\bar\alpha^n) .$$

In particular for every $n \geq 1$, $\delta'_p(n) \leq 2\omega_{\infty,h}(\bar\alpha^n)$ for any $p \geq 1$, and $\delta_\infty(n) \leq 2\omega_{\infty,h}(\bar\alpha^n)$.

Proof. For every $x \in [0,1]$, we have

$$\begin{aligned}|X_{n,x} - X_{n,y}| &= |h(B_n + A_n x) - h(Z_n + A_n y)| \\ &\leq |h(B_n + A_n x) - h(B_n)| + |h(B_n + A_n y) - h(B_n)| ,\end{aligned}$$

and the result follows. □

Remark 11 *When $K = \{1, \ldots, r\}$ and $|\alpha_1| = \cdots = |\alpha_r|$, similar computations as those done in Section 3.2 allow to control $(\delta'_p(n))_{n\in\mathbb{N}}$ thanks to the $\mathbb{L}^p$-modulus of continuity. Notice that, actually, the case where $\alpha_1 = \cdots = \alpha_r$ is included in section 3.2 (taking $m = 1$).*

As in the previous subsection, let us also consider the process $(h \circ T^n)_{n\in\mathbb{N}}$. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be iid random variables taking values in K such that $\mathbb{P}(\varepsilon_1 = k) = |\alpha_k|$ for every $k \in K$. For every $n \in \mathbb{N}$, set $Z_n := \sum_{\ell\in\mathbb{N}} s_{\varepsilon_n} \circ \cdots \circ s_{\varepsilon_{\ell+n}}(0) = \sum_{\ell\in\mathbb{N}} \left(\prod_{j=0}^{\ell-1} \alpha_{\varepsilon_{j+n}}\right) \beta_{\varepsilon_{\ell+n}}$. Then, $(Z_n)_{n\in\mathbb{N}}$ is identically distributed and the common law is invariant by P, so it is the Lebesgue measure on $[0,1]$. Moreover, one can see that $Z_n = T^n Z_0$ for every $n \in \mathbb{N}$. Hence, for every $h \in C([0,1])$, the processes $(h \circ T^n)_{n\in\mathbb{N}}$ (under λ) and $(Y_n)_{n\in\mathbb{N}} := (h(Z_n))_{n\in\mathbb{N}}$ (under $\mathbb{P}$) have the same distribution. As above the following upper bound clearly holds

$$\tilde\delta_\infty(n) \leq 2\omega_{\infty,h}(\bar\alpha^n) . \tag{22}$$

Proposition 12 *Let $p > 2$, $\kappa(p)$ be defined by (3), and $h \in C([0,1])$. Let $S_n = X_1 + \cdots + X_n = h(W_1) + \cdots + h(W_n)$, or $S_n = h(Z_1) + \cdots + h(Z_n)$. Assume that $\omega_{\infty,h}(2^{-n}) \leq O(n^{-\gamma})$ for some γ such that*

$$\gamma > \frac{p-1}{2} \text{ if } p \in (2,3], \text{ and } \gamma > \kappa(p) \text{ if } p > 3.$$

Then $n^{-1}\mathbb{E}\big((S_n - n\lambda(h))^2\big) \to \sigma^2$ as $n \to \infty$ and for every (fixed) $x \in [0,1]$, one can redefine $(S_n)_{n\geq1}$ without changing its distribution on a (richer) probability space on which there exist iid random variables $(N_i)_{i\geq1}$ with common distribution $\mathcal{N}(0,\sigma^2)$, such that,

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

Proof. It suffices to combine Lemma 10 (or (22)) with either our Theorem 3 (or our Theorem 1) for $p > 3$ or Theorem 1 of [3] for $p \in (2,3]$ (which applies also to processes as defined by (4)). □

4 Proof of the results

As in [3], the proof is based on a general proposition that can be established by combining the arguments given in the paper by Berkes, Liu and Wu [1]. Let us now recall this proposition: it applies to a strictly stationary sequence $(X_k)_{k\geq1}$ of real-valued random variables in $\mathbb{L}^p$ $(p > 2)$

that can be well approximated by a sequence of m-dependent random random variables, with the help of an auxiliary sequence of iid random variables $(\varepsilon_i)_{i\geq 0}$. Let $(M_k)_{k\geq 1}$ be a sequence of positive real numbers and define

$$\varphi_k(x) = (x \wedge M_k) \vee (-M_k) \text{ and } g_k(x) = x - \varphi_k(x)\,. \tag{23}$$

Then, define

$$X_{k,j} = \varphi_k(X_j) - \mathbb{E}\varphi_k(X_j) \text{ and } W_{k,\ell} = \sum_{i=1+3^{k-1}}^{\ell+3^{k-1}} X_{k,i}\,. \tag{24}$$

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

$$\tilde{X}_{k,j} = \mathbb{E}\big(\varphi_k(X_j)|\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_{j-m_k}\big) - \mathbb{E}\varphi_k(X_j) \text{ for any } j \geq m_k+1 \text{ and } \widetilde{W}_{k,\ell} = \sum_{i=1+3^{k-1}}^{\ell+3^{k-1}} \tilde{X}_{k,i}\,. \tag{25}$$

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

$$\nu_k = m_k^{-1}\big\{\mathbb{E}(\widetilde{W}^2_{k,m_k}) + 2\mathbb{E}(\widetilde{W}_{k,m_k}(\widetilde{W}_{k,2m_k} - \widetilde{W}_{k,m_k}))\big\}\,. \tag{26}$$

Proposition 13 (Berkes, Liu and Wu [1]) *Let $p > 2$. Assume that we can find a sequence of positive reals $(M_k)_{k\geq 1}$, 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:*

$$\sum_{k\geq 1} 3^{k(p-1)/p}\mathbb{E}(|g_k(X_1)|) < \infty\,, \tag{27}$$

there exists $\alpha \geq 1$ such that

$$\sum_{k\geq k_0} 3^{-\alpha k/p} \Big\| \max_{1\leq \ell \leq 3^k - 3^{k-1}} \big| W_{k,\ell} - \widetilde{W}_{k,\ell} \big| \Big\|_\alpha^\alpha < \infty\,, \tag{28}$$

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

$$\sum_{k\geq k_0} \frac{3^k}{3^{kr/p}m_k} \mathbb{E}\left(\max_{1\leq \ell\leq 3m_k} \big|\widetilde{W}_{k,\ell}\big|^r\right) < \infty\,. \tag{29}$$

Assume in addition that

$$\text{the series} \quad \sigma^2 = \mathrm{Var}(X_1^2) + 2\sum_{i\geq 1}\mathrm{Cov}(X_1, X_{i+1}) \quad \text{converge,} \tag{30}$$

and

$$3^k(\nu_k^{1/2} - \sigma)^2 = o(3^{2k/p}(\log k)^{-1})\,, \quad \text{as } k \to \infty\,. \tag{31}$$

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,

$$\left| S_n - n\mathbb{E}(X_1) - \sum_{i=1}^n N_i \right| = o\left(n^{1/p}\right) \quad \mathbb{P}\text{-a.s.} \tag{32}$$

Theorem 3 is a consequence of the next proposition, whose proof follows from Proposition 13. This proposition applies to the stationary sequence $(X_n)_{n\geq 1}$ defined by (8) and the conditions are expressed in terms of the coefficient δ'_p defined in the second part of Section 2. In what follows, all the proofs will be written with the help of that coefficient, but the arguments are exactly the same for the sequence defined by (1) or (4) and the coefficients $\tilde{\delta}_p$.

Proposition 14 *Let $p > 2$. Assume that we can find 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:*

$$\sum_{k\geq k_0} 3^{k(p-2)/2}\left(\sum_{\ell\geq k}\delta_2'(m_\ell)m_{\ell+1}^{1/2}\right)^p < \infty\,, \qquad \sum_{k\geq k_0}\left(\sum_{\ell\geq k}\delta_p'(m_\ell)m_{\ell+1}^{1-1/p}\right)^p < \infty\,, \tag{33}$$

$$\sum_{\ell\geq k}\delta_2'(m_\ell)m_{\ell+1}^{1/2} = o(3^{k(2-p)/2p}/\sqrt{\log k})\,, \tag{34}$$

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

$$\sum_{k\geq 0} 3^{k(p-r)/p} m_k^{(r-2)/2} < \infty\,, \tag{35}$$

and

$$\sum_{j\geq 1}\frac{(\delta_p'(j))^{p/r}}{j^{1/r}} < \infty\,. \tag{36}$$

Then, (30) holds. Moreover, if $\sigma > 0$, 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,

$$\left|S_n - n\mathbb{E}(X_1) - \sum_{i=1}^n N_i\right| = o\left(n^{1/p}\right) \quad \mathbb{P}\text{-}a.s. \tag{37}$$

4.1 Proof of Proposition 14

We consider a process $(X_n)_{n\geq 1}$ satisfying (8), with stationary distribution π. We shall check that the assumptions of Proposition 13 are satisfied.

Set $V_0 := (\varepsilon_0, W_0)$. It is not difficult to see that $\|\mathbb{E}((X_k - \mathbb{E}(X_1))|V_0)\|_2 \leq \delta_2'(k) \leq \delta_p'(k)$. Now, since $r \geq p \geq 2$, it follows from (36) that

$$\sum_{n\geq 1}\frac{\delta_2'(n)}{\sqrt{n}} < \infty\,. \tag{38}$$

Then, the fact that (30) holds follows from the fact that (see e.g. Lemma 22 of [3])

$$\sum_{k\geq 1}|\mathrm{Cov}(X_1, X_{k+1})| \ll \left(\sum_{k\geq 0}(k+1)^{-1/2}\|\mathbb{E}((X_k - \mathbb{E}(X_1))|V_0)\|_2\right)^2 < \infty\,. \tag{39}$$

We choose $M_k = 3^{k/p}$. Since the X_i's are in $\mathbb{L}^p$, it is easy to see that with this choice of M_k, condition (27) 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).

We shall check the condition (28) with $\alpha = p$. To do so, we apply the Rosenthal-type inequality given in Proposition 15 of the appendix, to the process $(\tilde{X}_{k,\ell+3^{k-1}} - X_{k,\ell+3^{k-1}})_{\ell\geq 1}$, with the choice $\eta_0 = (\varepsilon_{3^{k-1}}, \ldots, \varepsilon_{1+3^{k-1}-m_k}, W_{3^{k-1}-m_k})$ and for every $\ell \geq 1$, $\eta_\ell = \varepsilon_{\ell+3^{k-1}}$. We have to bound, for $q \geq 1$, the coefficients $\delta_{q,k}^*(n)$ used in Proposition 15. When $1 \leq n \leq m_k$, we use the bound $\delta_{q,k}^*(n) \leq 2\|\tilde{X}_{k,3^{k-1}} - X_{k,3^{k-1}}\|_q \leq 2\delta_q'(m_k)$. When $n > m_k$ we notice that the contribution of $(\tilde{X}_{k,\ell+3^{k-1}})_{\ell\geq 1}$ to $\delta_{q,k}^*(n)$ is null, so that $\delta_{q,k}^*(n) \leq \delta_q'(n)$.

In particular we infer that, for $k \geq k_0$,

$$\Big\| \max_{1\le \ell\le 3^k-3^{k-1}} |\widetilde{W}_{k,\ell}-W_{k,\ell}| \Big\|_p$$

$$\ll \big(3^{k/2}\delta_2'(m_k)m_k^{1/2}+3^{k/p}m_k^{1-1/p}\delta_p'(m_k)\big)+\left(3^{k/2}\sum_{j\ge m_k}\frac{\delta_2'(j)}{\sqrt{j}}+3^{k/p}\sum_{j\ge m_k}\frac{\delta_p'(j)}{j^{1/p}}\right)$$

$$\ll 3^{k/2}\sum_{\ell\ge k}\delta_2'(m_\ell)m_{\ell+1}^{1/2}+3^{k/p}\sum_{\ell\ge k}\delta_p'(m_\ell)m_{\ell+1}^{1-1/p}\,.$$

Hence, (28) holds with $\alpha=p$, since (33) is satisfied.

We prove now that (29) holds for some $r>2$. We apply again Proposition 15, but now to the process $(\tilde X_{k,\ell+3^{k-1}})_{1\le \ell\le 3m_k}$ and with the choice $\eta_0=(\varepsilon_{3^{k-1}},\dots,\varepsilon_{1+3^{k-1}-m_k})$ and for every $\ell\ge 1$, $\eta_\ell=\varepsilon_{\ell+3^{k-1}}$.

For every $q\ge 1$, denote by $\delta^*_{q,k}(n)$ the n^{th} coefficient δ^* associated with the above choice, and notice that $\delta^*_{q,k}(n)=0$ as soon as $n>m_k$. For every $q\ge 1$, denote by $\delta'_{q,k}(n)$ the n^{th} coefficient δ' associated with the process $(X_{k,\ell+3^{k-1}})_{\ell\ge 1}$. One can see that for every $n\ge 0$, $\delta^*_{q,k}(n)\le \delta'_{q,k}(n)$.

For every $r\ge 2$, every $k\ge 1$, with d_k the unique integer such that $2^{d_k-1}<3m_k\le 2^{d_k}$, Proposition 15 gives

$$\Big\| \max_{1\le \ell\le 3m_k} \big|\widetilde{W}_{k,\ell}\big| \Big\|_r \ll 2^{d_k/2}\sum_{j=0}^{m_k}\delta'_{2,k}(j)/(j+1)^{1/2}+2^{d_k/r}\sum_{j=0}^{m_k}\delta'_{r,k}(j)/(j+1)^{1/r}\,. \tag{40}$$

Hence, (29) holds for some $r>2$, if

$$\sum_{k\ge 0}3^{k(p-r)/p}m_k^{(r-2)/2}<\infty \qquad\text{and}\qquad \sup_{k\ge 0}\sum_{j\ge 0}\delta'_{2,k}(j)/(j+1)^{1/2}<\infty\,, \tag{41}$$

and

$$\sum_{k\ge 0}3^{k(p-r)/p}\Big(\sum_{j\ge 0}\delta'_{r,k}(j)/(j+1)^{1/r}\Big)^r<\infty\,. \tag{42}$$

The first part of (41) is exactly (35). Moreover since φ_k is 1-Lipschitz, we have $\delta'_{2,k}(n)\le \delta_2'(n)\le \delta_p'(n)$. Hence the second part of (41) holds for some $r>2$ as soon as (36) does.

It remains to prove (42). By Hölder's inequality,

$$\sum_{j\ge 0}\frac{\delta'_{r,k}(j)}{(j+1)^{1/r}}\le \left(\sum_{j\ge 0}\frac{(\delta_p'(j))^{p/r}}{(j+1)^{1/r}}\right)^{(r-1)/r}\left(\sum_{j\ge 0}\frac{(\delta'_{r,k}(j))^r}{(j+1)}\frac{(j+1)^{(r-1)/r}}{(\delta_p'(j))^{p(r-1)/r)}}\right)^{1/r}.$$

Taking into account (36), we see that (42) holds as soon as

$$\sum_{j\ge 0}\frac{(\delta_p'(j))^{p(1-r)/r}}{(j+1)^{1/r}}\sum_{k\ge 0}3^{k(p-r)/p}(\delta'_{r,k}(j))^r<\infty\,.$$

Using the fact that $\delta'_{r,k}(j)\le 2\|\varphi_k(X_0)\|_r$, that for every non negative random variable Z,

$$\sum_{k\ge 0}3^{k(p-r)/p}\mathbb{E}((\varphi_k(Z))^r)\le C_{r,p}\mathbb{E}(Z^p)\,,$$

and (36) again, we see that (42) holds.

To end the proof, it remains to prove that (31) holds. Since $\sigma > 0$, it follows from equation (65) of [3] that (31) is satisfied as soon as

$$3^k(\nu_k - \sigma^2)^2 = o(3^{2k/p}(\log k)^{-1}), \text{ as } k \to \infty. \tag{43}$$

To prove (43), let us define, for $i \geq 0$,

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

We have

$$|\nu_k - \sigma^2| \leq \left| \sum_{i=-m_k}^{m_k} \tilde{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} \right| + \left| \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} c_{|i|} \right|. \tag{44}$$

Arguing as in [3] to obtain their equation (68), and making use of (38), we see that

$$\left| \sum_{i=-m_k}^{m_k} \tilde{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} \right| \leq C\Big(\limsup_{j\to\infty} j^{-1/2} \left\| \widetilde{W}_{k,j} - W_{k,j} \right\|_2 + \limsup_{j\to\infty} j^{-1} \left\| \widetilde{W}_{k,j} - W_{k,j} \right\|_2^2 \Big),$$

for some $C > 0$, independent of $k > 0$. Estimating the right-hand side thanks to Proposition 15 with $p = 2$, we infer that

$$\begin{aligned} \left| \sum_{i=-m_k}^{m_k} \tilde{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} \right| &\ll \left(1 + \sum_{\ell\geq 0} \frac{\delta'_2(\ell)}{\sqrt{\ell+1}} \right) \left(\sqrt{m_k}\delta'_2(m_k) + \sum_{\ell\geq m_k} \frac{\delta'_2(\ell)}{\sqrt{\ell+1}} \right) \\ &\ll \left(\sum_{\ell\geq 0} \frac{\delta'_2(\ell)}{\sqrt{\ell+1}} \right) \sum_{\ell\geq k} \delta'_2(m_\ell) m_{\ell+1}^{1/2}. \end{aligned}$$

Hence by (34) and (36),

$$3^k \left| \sum_{i=-m_k}^{m_k} \tilde{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} \right|^2 = o(3^{2k/p}(\log k)^{-1}). \tag{45}$$

Let now $c_i = \mathrm{Cov}(X_0, X_i)$ and note that (see Relation (3.54) in [1])

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

Let

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

It follows that

$$\left| \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} c_{|i|} \right| \leq o(\ell_k 3^{-k(p-2)/p}) + 2\sum_{i>\ell_k} |c_i - \hat{c}_{k,i}|.$$

Now

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

Therefore

$$\begin{aligned} \left| \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} c_{|i|} \right| \leq o(\ell_k 3^{-k(p-2)/p}) &+ 2\sum_{i>\ell_k} |\,\mathrm{Cov}(X_0 - \varphi_k(X_0), X_i)| \\ &+ 2\sum_{i>\ell_k} |\,\mathrm{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))|. \end{aligned} \tag{46}$$

Let us first handle the series

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

Set $g_k(x) = x - \varphi_k(x)$. Applying Lemma 22 of [3] and using the fact that $(W_k)_{k\geq 0}$ is a Markov chain, we infer that

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

We shall now use the following estimate, to be proved at the end of this subsection: let $\varphi_M(x) = (x \wedge M) \vee (-M)$ and $g_M = x - \varphi_M(x)$, then

$$\|\mathbb{E}(g_M(X_n)|V_0) - \mathbb{E}(g_M(X_n))\|_2 \ll \frac{1}{M^{(p-2)/2}} \big(\delta'_p(n)\big)^{p/(2(p-1))} . \qquad (47)$$

Taking into account (47) and the fact that $\|\mathbb{E}(X_i|V_0) - \mathbb{E}(X_i)\|_2 \leq \delta'_2(i) \leq \delta'_p(i)$, it follows that

$$\sum_{i>\ell_k} |\operatorname{Cov}(X_0 - \varphi_k(X_0), X_i)| \ll \frac{1}{M_k^{(p-2)/2}} \Big(\sum_{i\geq [\ell_k/2]} i^{-1/2} \delta'_p(i) \Big) \sum_{j=0}^{\infty} (j+1)^{-1/2} (\delta'_p(j))^{p/(2(p-1))} .$$

Now, using (36) and the fact that $(\delta'_p(j))_{j\geq 0}$ is non increasing, we see that $\delta'_p(j) = o(j^{(1-r)/p})$. In particular, $\sum_{j=0}^{\infty} (j+1)^{-1/2} (\delta'_p(j))^{p/(2(p-1))} < \infty$ and, since $r \geq p$,

$$\sum_{i\geq [\ell_k/2]} i^{-1/2} \delta'_p(i) = O(\ell_k^{1/p-1/2}) .$$

Hence

$$\sum_{i>\ell_k} |\operatorname{Cov}(X_0 - \varphi_k(X_0), X_i)| \ll \frac{\ell_k^{1/p-1/2}}{M_k^{(p-2)/2}} . \qquad (48)$$

Let us now handle the series

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

Applying again Lemma 22 of [3] and taking into account the fact that $(W_k)_{k\geq 0}$ is a Markov chain, we first infer that

$$\sum_{i>\ell_k} |\operatorname{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))|$$
$$\ll \sum_{\ell=0}^{\infty} (\ell+1)^{-1/2} \|\mathbb{E}(\varphi_k(X_\ell)|V_0) - \mathbb{E}(\varphi_k(X_\ell))\|_2 \sum_{i\geq [2^{-1}(\ell_k+\ell)]+1} i^{-1/2} \|\mathbb{E}(g_k(X_i)|\mathcal{F}_0) - \mathbb{E}(g_k(X_i))\|_2 .$$

Since φ_k is 1-Lipschitz, we have $\|\mathbb{E}(\varphi_k(X_\ell)|V_0) - \mathbb{E}(\varphi_k(X_\ell))\|_2 \leq \delta'_2(\ell) \leq \delta'_p(\ell)$. Therefore, since by assumption, $\sum_{\ell=0}^{\infty} (\ell+1)^{-1/2} \delta'_p(\ell) < \infty$,

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

Using (47), the fact that $\delta'_p(i) = o(i^{(1-r)/p})$ and that $r > p$, it follows

$$\sum_{i>\ell_k} |\operatorname{Cov}(\varphi_k(X_0), X_i - \varphi_k(X_i))| \ll \frac{1}{M_k^{(p-2)/2}} \sum_{i \geq [2^{-1}\ell_k]+1} i^{-1/2} (\delta'_p(i))^{p/(2(p-1))} \ll \frac{\ell_k^{(p-r)/(2(p-1))}}{M_k^{(p-2)/2}} . \quad (49)$$

Starting from (46) and using (48) and (49), we get

$$\Big| \sum_{i\in\mathbb{Z}} \hat{c}_{k,|i|} - \sum_{i\in\mathbb{Z}} c_{|i|} \Big| \ll o(\ell_k 3^{-k(p-2)/p}) + \frac{\ell_k^{1/p-1/2}}{M_k^{(p-2)/2}} + \frac{\ell_k^{(p-r)/(2(p-1))}}{M_k^{(p-2)/2}} . \quad (50)$$

Starting from (44) and taking into account (45) and (50), the condition (43) is satisfied (since $M_k = 3^{k/p}$, $\ell_k = 3^{k(p-2)/(2p)}(\log k)^{-1/2}$ and $r > p > 2$). □

Proof of (47). We start by noticing that, for any $\varepsilon > 0$,

$$\begin{aligned} \|\mathbb{E}(g_M(X_n)|V_0) - \mathbb{E}(g_M(X_n)\|_2^2 &= \int \big|\mathbb{E}(g_M(X_{n,x})) - \int \mathbb{E}(g_M(X_{n,y}))\nu(dy)\big|^2 \nu(dx) \\ &\leq \iint \big|\mathbb{E}(g_M(X_{n,x}) - g_M(X_{n,y}))\big|^2 \nu(dx)\nu(dy) \\ &\leq 2 \iint \big|\mathbb{E}\big((g_M(X_{n,x}) - g_M(X_{n,y}))\mathbf{1}_{\{|g_M(X_{n,x}) - g_M(X_{n,y})| \leq \varepsilon\}}\big)\big|^2 \nu(dx)\nu(dy) \\ &\quad + 2 \iint \big|\mathbb{E}\big((g_M(X_{n,x}) - g_M(X_{n,y}))\mathbf{1}_{\{|g_M(X_{n,x}) - g_M(X_{n,y})| > \varepsilon\}}\big)\big|^2 \nu(dx)\nu(dy) . \quad (51) \end{aligned}$$

Now,

$$\begin{aligned} &\iint \big|\mathbb{E}\big((g_M(X_{n,x}) - g_M(X_{n,y}))\mathbf{1}_{\{|g_M(X_{n,x}) - g_M(X_{n,y})| \leq \varepsilon\}}\big)\big|^2 \nu(dx)\nu(dy) \\ &\leq \varepsilon \iint \mathbb{E}\big|g_M(X_{n,x}) - g_M(X_{n,y})\big| \nu(dx) d\nu(dy) \\ &\leq 2\varepsilon \mathbb{E}(|g_M(X_n)|) \leq 2\varepsilon \mathbb{E}(|X_n| \mathbf{1}_{\{|X_n| > M\}}) \leq 2\varepsilon M^{1-p} \mathbb{E}(|X_1|^p) . \quad (52) \end{aligned}$$

On the other hand,

$$\begin{aligned} &\iint \big|\mathbb{E}\big((g_M(X_{n,x}) - g_M(X_{n,y}))\mathbf{1}_{\{|g_M(X_{n,x}) - g_M(X_{n,y})| > \varepsilon\}}\big)\big|^2 \nu(dx)\nu(dy) \\ &\leq \iint \mathbb{E}\big((g_M(X_{n,x}) - g_M(X_{n,y}))^2 \mathbf{1}_{\{|g_M(X_{n,x}) - g_M(X_{n,y})| > \varepsilon\}}\big) \nu(dx)\nu(dy) \\ &\leq \varepsilon^{2-p} \iint \mathbb{E}\big|g_M(X_{n,x}) - g_M(X_{n,y})\big|^p \nu(dx)\nu(dy) \\ &\leq 2^{p-1}\varepsilon^{2-p} \iint \mathbb{E}\big|X_{n,x} - X_{n,y}\big|^p \nu(dx)\nu(dy) + 2^{p-1}\varepsilon^{2-p} \iint \mathbb{E}\big|\varphi_M(X_{n,x}) - \varphi_M(X_{n,y})\big|^p \nu(dx)\nu(dy) \\ &\leq 2^p \varepsilon^{2-p} (\delta'_p(n))^p . \quad (53) \end{aligned}$$

Starting from (51), taking into account (52) and (53), and selecting $\varepsilon = M\big(\delta'_p(n)\big)^{p/(p-1)}$, the upper bound (47) follows.

4.2 Proof of Theorem 3

Assume that $\delta'_p(n) = O(n^{-\gamma})$, for some $\gamma > 0$.

We shall first assume that $\sigma > 0$ and apply Proposition 14. We shall take $m_k = O(3^{k\beta/p})$ for some $2 \geq \beta > 0$. Hence, we have to find $r > p$, $\gamma > 0$ and $2 \geq \beta > 0$ such that (33), (34), (35) and (36) hold. One easily sees that the condition (33) holds provided that

$$p - 2 < \beta(2\gamma - 1) \qquad \text{and} \qquad \gamma > 1 - 1/p\,;$$

the condition (34) holds provided that

$$p - 2 < \beta(2\gamma - 1)\,; \tag{54}$$

the condition (35) holds provided that

$$r < 1 + \gamma p\,; \tag{55}$$

and the condition (36) holds provided that

$$\beta(r - 2) < 2(r - p)\,. \tag{56}$$

Notice that the condition $p - 2 < \beta(2\gamma - 1)$ appears twice, that (56) implies that $\beta \leq 2$, and that the condition $\gamma > 1 - 1/p$ is realized as soon as $r < 1 + \gamma p$.

Hence, we have to find $\gamma, \beta > 0$ and $r > p$ such that (54), (55) and (56) hold. In particular, one has to find $\beta > 0$ such that $(p-2)/(2\gamma - 1) < \beta < 2(r-2)/(r-p)$, which is possible as soon as $(p-2)/(2\gamma-1) < 2(r-2)/(r-p)$. The latter condition is equivalent to (provided that $4\gamma > p$, a condition to be checked at the end):

$$r - 2 > \frac{(2p-4)(2\gamma-1)}{4\gamma - p}\,.$$

Now, one can find $r > p$ satisfying the latter condition and (55) as soon as

$$2 + \frac{(2p-4)(2\gamma-1)}{4\gamma - p} < 1 + \gamma p\,,$$

which is equivalent to

$$4p\gamma^2 - (p^2 + 4p - 4)\gamma + 3p - 4 > 0\,.$$

Then, one finds that

$$\gamma > \frac{(p-2)\sqrt{p^2 + 12p + 4} + p^2 + 4p - 4}{8p}$$

solves the problem.

Hence, Proposition 14 applies and the Theorem is proved in the case $\sigma > 0$, provided that our condition on γ imples that $4\gamma > p$, but this may be easily checked.

Assume now on that $\sigma = 0$. Proceeding as in the proof of Theorem 1 of [3] (see page 17), we see that it suffices to prove that

$$\sum_{n \geq 1} \frac{\delta'_p(n)}{n^{2/p^2}} < \infty\,.$$

Hence, it is enough to prove that

$$\frac{2}{p^2} + \frac{(p-2)\sqrt{p^2 + 12p + 4} + p^2 + 4p - 4}{8p} > 1\,,$$

which in turn is equivalent to (recall that $p > 2$)

$$p\sqrt{p^2 + 12p + 4} > -(p^2 - 2p - 8)\,. \tag{57}$$

The right-hand side of (57) is non-positive for $p \geq 1 + 2\sqrt{2}$. Taking the squares of (57), one can see that (57) holds for $p \in]2, 1 + 2\sqrt{2}]$, which ends the proof of the theorem. □

5 Appendix

5.1 A Rosenthal-type inequality under dependence

We shall state and prove our inequality in a more general framework than needed. It is not difficult to prove that the coefficients $(\delta_p^*(n))_{n\geq 0}$ defined by (58) below are precisely the ones introduced in (9), taking $\eta_0 = W_0$ and for every $k \geq 1$, $\eta_k = \varepsilon_k$.

Let $(\eta_k)_{k\geq 0}$ be independent random variables (not necessarily identically distributed) and define for every $k \geq 1$, $\mathcal{G}_{k,-1/2} := \{\emptyset, \Omega\}$ and $\mathcal{G}_{k,0} = \sigma\{\eta_k\}$ and for every $k \geq 2$ and every $0 \leq \ell \leq k-1$, $\mathcal{G}_{k,\ell} := \sigma\{\eta_k, \eta_{k-1}, \ldots, \eta_{k-\ell}\}$.

Let $(X_n)_{n\geq 1}$ be a process given by $X_n := f_n(\eta_n, \eta_{n-1}, \ldots, \eta_0)$, for $n \geq 1$, where f_n is a real-valued measurable function. Assume that for every $n \geq 1$, $\mathbb{E}(|X_n|) < \infty$ and that $\mathbb{E}(X_n) = 0$. We want to prove a Rosenthal-type inequality for $S_n := X_1 + \cdots + X_n$, $n \geq 1$.

We shall need the following measure of dependence. Let $(\eta'_k)_{k\geq 0}$ be an independent copy of $(\eta_k)_{k\geq 0}$. For every $k \geq m+1$ and every $m \geq 0$, set $X'_{k,m} = f_k(\eta_k, \ldots, \eta_{k-m}, \eta'_{k-m-1}, \ldots, \eta'_0)$ and, then, for every $n \geq 1$,

$$\delta_p^*(n) := \sup_{m\geq n-1} \sup_{k\geq m} \|X_k - X'_{k,m}\|_p \,. \tag{58}$$

Define also $\delta_p^*(0) = \sup_{k\geq 0} \|X_k\|_p$.

For every $\ell \geq 1$, set $T_{0,\ell} := \mathbb{E}(X_\ell|\mathcal{G}_{\ell,0})$. For every $d \geq 0$, every $0 \leq k \leq d$ and every $1 \leq \ell \leq 2^{d-k}$, set

$$U_{k,\ell} := \sum_{j=(\ell-1)2^k+1}^{\ell 2^k} \big(X_j - \mathbb{E}(X_j|\mathcal{G}_{\ell 2^k, 2^k-1})\big) = \sum_{j=(\ell-1)2^k+1}^{\ell 2^k} \big(X_j - \mathbb{E}(X_j|\mathcal{G}_{j, j-(\ell-1)2^k-1})\big)\,. \tag{59}$$

and

$$T_{k+1,\ell} = \mathbb{E}\left(\big(U_{k,2\ell-1} + U_{k,2\ell}\big)|\mathcal{G}_{\ell 2^{k+1}, 2^{k+1}-1}\right)\,.$$

Proposition 15 *For every $d \geq 0$, we have*

$$\max_{1\leq n\leq 2^d} |S_n| \leq \sum_{k=0}^d \max_{1\leq \ell\leq 2^{d-k}} |U_{k,\ell}| + \sum_{k=0}^d \max_{1\leq m\leq 2^{d-k}} \left|\sum_{\ell=1}^m T_{k,\ell}\right|\,. \tag{60}$$

In particular, if $X_n \in \mathbb{L}^p$, for every $n \geq 1$ and some $p \geq 2$, we have

$$\Big\|\max_{1\leq n\leq 2^d} |S_n|\Big\|_p \leq \sum_{k=0}^d \left(\sum_{\ell=1}^{2^{d-k}} \|U_{k,\ell}\|_p^p\right)^{1/p} + C_p \sum_{k=0}^d \left(\left(\sum_{\ell=1}^{2^{d-k}} \|T_{k,\ell}\|_2^2\right)^{1/2} + \left(\sum_{\ell=1}^{2^{d-k}} \|\, T_{k,\ell}\|_p^p\right)^{1/p}\right) \tag{61}$$

$$\leq C'_p 2^{d/2} \sum_{j=0}^{2^d} \frac{\delta_2^*(j)}{(j+1)^{1/2}} + C''_p 2^{d/p} \sum_{j=0}^{2^d} \frac{\delta_p^*(j)}{(j+1)^{1/p}}\,, \tag{62}$$

where C_p is the best constant in the Rosenthal inequality for independent random variables, $C'_p = \frac{C_p 2^{3/2}}{\sqrt{2}-1}$ and $C''_p = \frac{2^{1+1/p}(C_p+1)}{2^{1/p}-1}$.

Proof. The proof is done by induction on $d \geq 0$. The case where $d = 0$ follows from the decomposition

$$X_n = \big(X_n - \mathbb{E}(X_n|\sigma\{\varepsilon_n\})\big) + \mathbb{E}(X_n|\sigma\{\varepsilon_n\})\,.$$

Assume now that (60) holds for some $d \geq 0$. Let us prove that it holds for $d+1$.

For every $n \geq 1$, we have $S_n = \sum_{k=1}^n \big(X_k - \mathbb{E}(X_k|\mathcal{G}_{k,0})\big) + \sum_{k=1}^n \mathbb{E}(X_k|\mathcal{G}_{k,0}) := R_n + \sum_{\ell=1}^n T_{0,\ell}$. Hence

$$\max_{1\leq n\leq 2^{d+1}} |S_n| \leq \max_{1\leq n\leq 2^{d+1}} |R_n| + \max_{1\leq n\leq 2^{d+1}} \Big|\sum_{\ell=1}^n T_{0,\ell}\Big| \,.$$

Using that for every $m \geq 1$, we have $|R_{2m+1}| \leq |R_{2m}| + |X_{2m+1} - \mathbb{E}(X_{2m+1}|\mathcal{G}_{2m+1,0})|$, we infer that

$$\max_{1\leq n\leq 2^{d+1}} |S_n| \leq \max_{1\leq n\leq 2^{d}} |R_{2n}| + \max_{1\leq k\leq 2^{d+1}} |U_{0,k}| + \max_{1\leq n\leq 2^{d+1}} \Big|\sum_{\ell=1}^n T_{0,\ell}\Big| \,. \qquad (63)$$

We shall use our induction hypothesis to handle the first term in the right-hand side of (63).

For every $m \geq 1$, let

$$\tilde X_m := X_{2m-1} - \mathbb{E}(X_{2m-1}|\mathcal{G}_{2m-1,0}) + X_{2m} - \mathbb{E}(X_{2m}|\mathcal{G}_{2m,0}) \ \text{ and } \ \tilde S_m := R_{2m}\,.$$

Set $\tilde\eta_m := (\eta_{2m}, \eta_{2m-1})$ for $m \geq 1$ and $\tilde\eta_0 := \eta_0$. Let also $\tilde{\mathcal{G}}_{k,\ell} := \sigma\{\tilde\eta_k, \dots \tilde\eta_{k-\ell}\} = \mathcal{G}_{2k,2\ell+1}$. Then, for every $(\ell-1)2^k + 1 \leq j \leq \ell 2^k$, using that $\mathcal{G}_{2j-1,0} \subset \mathcal{G}_{\ell 2^{k+1},2^{k+1}-1} = \tilde{\mathcal{G}}_{\ell 2^k,2^k-1}$ and that $\mathcal{G}_{2j,0} \subset \mathcal{G}_{\ell 2^{k+1},2^{k+1}-1}$, we have

$$\tilde X_j - \mathbb{E}(\tilde X_j|\tilde{\mathcal{G}}_{\ell 2^k,2^k-1}) = (X_{2j-1} + X_{2j}) - \mathbb{E}(X_{2j-1} + X_{2j}|\mathcal{G}_{\ell 2^{k+1},2^{k+1}-1})\,.$$

Hence, for every $k \geq 0$,

$$\tilde U_{k,\ell} := \sum_{j=(\ell-1)2^k+1}^{\ell 2^k} \big(\tilde X_j - \mathbb{E}(\tilde X_j|\tilde{\mathcal{G}}_{\ell 2^k,2^k-1})\big) = \sum_{j=(\ell-1)2^{k+1}+1}^{\ell 2^{k+1}} \big(X_j - \mathbb{E}(X_j|\mathcal{G}_{\ell 2^{k+1},2^{k+1}-1})\big) = U_{k+1,\ell}\,,$$

and

$$\tilde T_{k+1,\ell} := \mathbb{E}\Big(\big(\tilde U_{k,2\ell-1} + \tilde U_{k,2\ell}\big)|\tilde{\mathcal{G}}_{\ell 2^{k+1},2^{k+1}-1}\Big) = \mathbb{E}\Big(\big(U_{k+1,2\ell-1} + U_{k+1,2\ell}\big)|\mathcal{G}_{\ell 2^{k+2},2^{k+2}-1}\Big) = T_{k+2,\ell}\,.$$

Notice that we also have

$$\tilde T_{0,\ell} = T_{1,\ell}.$$

Applying the induction hypothesis, we infer that

$$\begin{aligned}
\max_{1\leq n\leq 2^d} |\tilde S_n| &\leq \sum_{k=0}^d \max_{1\leq \ell\leq 2^{d-k}} \big|\tilde U_{k,\ell}\big| + \sum_{k=0}^d \max_{1\leq m\leq 2^{d-k}} \Big|\sum_{\ell=1}^m \tilde T_{k,\ell}\Big| \\
&\leq \sum_{k=0}^d \max_{1\leq \ell\leq 2^{d-k}} \big|U_{k+1,\ell}\big| + \sum_{k=0}^d \max_{1\leq m\leq 2^{d-k}} \Big|\sum_{\ell=1}^m T_{k+1,\ell}\Big| \\
&= \sum_{k=1}^{d+1} \max_{1\leq \ell\leq 2^{d+1-k}} \big|U_{k,\ell}\big| + \sum_{k=1}^{d+1} \max_{1\leq m\leq 2^{d+1-k}} \Big|\sum_{\ell=1}^m T_{k,\ell}\Big| \,,
\end{aligned}$$

which, combined with (63) yields (60) with $d+1$ in place of d.

To prove (61), we notice that on the one hand, $(\max_{1\leq \ell\leq 2^{d-k}} \|U_{k,\ell}\|)^p \leq \sum_{\ell=1}^{2^{d-k}} \|U_{k,\ell}\|^p$ and on the other hand, for every $0 \leq k \leq d$, the variables $(T_{k,\ell})_{1\leq \ell\leq 2^{d-k}}$ are independent. Then, it is a direct consequence from (60) and the Rosenthal inequality for independent variables.

Since $\|T_{0,\ell}\|_q \le \delta_q^*(0)$ and $\|T_{k+1,\ell}\|_q \le 2\sum_{j=1}^{2^k} \delta_q^*(j)$, we infer from (61) that

$$\Big\| \max_{1\le n\le 2^d} |S_n| \Big\|_p \le 2C_p \sum_{k=0}^{d} 2^{(d-k)/2} \sum_{j=0}^{2^{k-1}} \delta_2^*(j) + 2^{d/p}\delta_p^*(0) + (2C_p+1) \sum_{k=0}^{d} 2^{(d-k)/p} \sum_{j=1}^{2^k} \delta_p^*(j)\,, \quad (64)$$

and (62) easily follows.

□

Acknowledgement. The first author is very thankful to the laboratories MAP5 and LAMA for their invitations, that made possible the present collaboration.

References

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

[2] J.-P. Conze, L. Hervé, L. and A. Raugi, *Pavages auto-affines, opérateurs de transfert et critères de réseau dans* $\mathbb{R}^d$, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 1, 1-42.

[3] Cuny, C., Dedecker, J. and Merlevède, F., On the Komlós, Major and Tusnády strong approximation for some classes of random iterates, arXiv:1706.08282v1

[4] Dedecker, J., Inégalités de Hoeffding et théorème limite central pour des fonctions peu régulières de chaînes de Markov non irréductibles. Ann. I.S.U.P. 52 (2008), no. 1-2, 39-46.

[5] 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.

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

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

[8] Shao, X. and Wu, W.-B., Limit theorems for iterated random functions. *Journal of Applied Probability* **41** (2004), no. 2, 425–436.

[9] Wu, W. B. Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** (2005), no 40, 14150-14154.