Strong invariance principles with rate for "reverse" martingale differences and applications.

Christophe Cuny[a] and Florence Merlevède[b]

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

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

Abstract

In this paper, we obtain almost sure invariance principles with rate of order $n^{1/p}\log^\beta n$, $2 < p \le 4$, for sums associated with a sequence of reverse martingale differences. Then, we apply those results to obtain similar conclusions in the context of some non-invertible dynamical systems. For instance we treat several classes of maps of the interval (for possibly unbounded observables) or smooth dynamical systems under a very weak regularity condition on the observables.

Mathematics Subject Classification (2010): 37E05, 37C30, 60F15.
Key words: Expanding maps, smooth dynamical systems, strong invariance principle, reverse martingale.

1 Introduction

The almost sure invariance principle (ASIP) is a powerful tool in both probability and statistics. It says that the partial sums of random variables can be approximated by those of independent Gaussian random variables, and that the approximation error between the trajectories of the two processes is negligible in a certain sense. The result can be more or less precise, depending on the specific error term one can obtain. For real-valued independent and identically distributed (iid) random variables with finite second moment, using the Skorohod embedding theorem, Strassen [36] proved that the approximation holds with an error of order $o((n\log\log n)^{1/2})$ in the almost sure sense. Still in the iid setting, the so-called Komlós, Major and Tusnády result states that this error of approximation is of order $o(n^{1/p})$ in the almost sure sense, provided that the variables are in $\mathbb{L}^p$ for a $p > 2$ (see [19] and [23]). Their method of proof makes use of an explicit construction of the approximating Brownian motion, based on quantile transformation.

Since the papers by Strassen and by Komlós, Major and Tusnády, there has been a great amount of works to extend this type of almost sure approximation results to dependent sequences under various conditions. In this paper, we are particularly interested by such extensions in the context of real-valued observables of the iterates of dynamical systems. In this context, the question of possible extensions of the Strassen or Komlós, Major and Tusnády strong approximation results can be formulated, in a very general way, as follows. Given (X, Σ, ν) a probability space, T a Σ-measurable and ergodic transformation on X, preserving the probability ν, and f a measurable function from X to $\mathbb{R}$ such that $\nu(|f|^p) < \infty$ for some $p \ge 2$, for what classes of dynamical systems (T, ν) and what classes of observables f, does there exist a sequence of independent identically distributed (iid) Gaussian random variables $(Z_i)_{i\ge1}$ such that

$$\sup_{1\le k\le n}\Big|\sum_{i=0}^{k-1}(f\circ T^i - \nu(f) - Z_i)\Big| = o(b_n) \quad \text{almost surely,} \tag{1.1}$$

where $(b_n)_{n\ge1}$ is a non-decreasing sequence of positive reals tending to infinity slower than $(n\log\log n)^{1/2}$? Obviously, the smaller the sequence (b_n) is, the better the approximation result is. Notice that when $b_n = (n\log\log n)^{1/2}$, (1.1) entails the functional law of the iterated logarithm and when $b_n = n^{1/2-\varepsilon}$ it leads to the weak invariance principle. Moreover, results of type (1.1) allow to control, in the almost sure sense, the increments of the partial sums process from the ones of the Brownian motion, which are well understood (see for instance Lemma 4.1).

To prove the central limit theorem (or its functional form) for dynamical systems, one popular method is the use of martingale approximation à la Gordin [13]. It should be emphasised that this martingale approximation method leads in many situations only to a reverse martingale increment sequence. This is obviously not a problem when one has in mind to prove results in distribution such as the central limit theorem. For almost sure results, the situation is different. However, as quoted in the introduction of Melbourne and Nicol [24], this approximation by a partial sum associated with a reverse martingale increment sequence is not an issue if the class of dynamical systems under consideration is closed under time reversal (see [2], [11] and [26]). But, for classes of dynamical systems that are intrinsically time-orientated such as those considered in Section 3.2, the situation can be more delicate. Up to our knowledge, the pioneering work concerning almost sure approximation results for such classes of dynamical systems is due to Hofbauer and Keller [17]. For a wide class of uniformly expanding maps of the interval, and for Hölder observables f, by exploiting a result of Philipp and Stout [32, Theorem 7.1], they obtained (1.1) with $b_n = n^{1/2-\varepsilon}$ for some $\varepsilon > 0$. For real-valued Hölder observables, a similar approach has also been successfully used by Melbourne and Nicol ([24],[25]) leading to some improvements in the rate obtained in [17]. Let us mention that other approaches have been recently exploited. For instance, the paper by Merlevède and Rio [29] is based on an explicit construction of the approximating Brownian motion with the help of the conditional quantile transform. For the class of expanding maps considered in [17], their result leads to the rate $b_n = n^{1/3}(\log n)^{1/2}(\log\log n)^{(1+\varepsilon)/3}$, for any $\varepsilon > 0$, in (1.1) when f is piecewise monotonic and in $\mathbb{L}^r(\nu)$ where $r > 3$. On another hand, Gouëzel [12] obtained a very general result concerning the rates in the ASIP for *vector-valued* observables of the iterates of dynamical systems by mean of spectral methods. In the particular case of expanding maps as those considered in [17], his result gives the rate $o(n^{1/4+\varepsilon})$ for any $\varepsilon > 0$ for some bounded vector-valued observables.

The starting point of this paper was to exploit directly the popular martingale approximation à la Gordin in the context of non-invertible maps to get almost sure approximation results as in (1.1). For invertible maps, it should be mentioned that such a martingale approximation method has been used in previous works (see [32], [35], [38], [4]). Now, since for non-invertible maps, the approximation leads to a reverse martingale increment sequence, we first prove almost sure invariance principles with rate for the partial sums associated with a reverse martingale increment sequence (see our Theorem 2.3). This result combined with an explicit control of the error in the martingale approximation (see our Proposition 5.1) allows to exhibit very general conditions on the dynamical system and on the observables under which a strong approximation principle of type (1.1) is valid with $b_n = n^{1/p}L(n)$ where $p \in]2, 4]$ and $L(n)$ is a slowly varying function. When applied for instance to the class of expanding maps considered in [17], our results lead to the rate $O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})$ in (1.1), when f is piecewise monotonic and in $\mathbb{L}^4(\nu)$. When f is piecewise monotonic and in $\mathbb{L}^p(\nu)$ for some $p \in]2, 4[$ the rate becomes $o(n^{1/p}L(n))$. Such results, that do not require the boundedness of f, are not very common in the context of dynamical systems. In case of a smooth transformation on a compact Riemann manifold, we obtain similar rates under a very weak regularity condition on the observables.

Our paper is organized as follows. In Section 2, we state some results concerning the almost sure invariance principle with rates for sums associated with a sequence of reverse martingale differences. The proofs of these results are postponed to Section 4. Section 3 is devoted to applications to dynamical systems. More precisely, in Section 3.1, we give general results concerning the rates of convergence in the ASIP for the partial sums associated with the observables of the iterates of non-invertible maps. Applications to different classes of dynamical systems are given in Section 3.2. Section 5 (respectively Section 6) is devoted to the proofs of the results stated in Section 3.1 (respectively in Section 3.2).

In this paper, we shall use sometimes the notation $a_n \ll b_n$ to mean that there exists a numerical constant C not depending on n such that $a_n \leq C b_n$, for all positive integers n.

2 ASIP with rates for sums of differences of reverse martingales

The next two results can be viewed as reverse martingales analogues to a result of Shao [35, Theorem 2.1]. The proof of Proposition 2.1 below follows from the Skorohod embedding of reverse martingales in Brownian motion as obtained in Scott and Huggins [34] together with an estimate concerning the increments of Brownian motions given in Hanson and Russo [14, Theorem 3.2A]. The proofs of all the results presented in this section are postponed to Section 4.

In this section we consider real-valued random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $(\mathcal{G}_n)_{n\in\mathbb{N}}$ be a non-increasing sequence of σ-algebras. We shall say that $(\xi_n)_{n\in\mathbb{N}}$ is a sequence of differences of reverse martingales (or a reverse martingale differences sequence) with respect to $(\mathcal{G}_n)_{n\in\mathbb{N}}$, if, for each n, ξ_n is integrable, $\mathcal{G}_n$-measurable and such that $\mathbb{E}(\xi_n|\mathcal{G}_{n+1}) = 0$ $\mathbb{P}$-a.s.

Proposition 2.1 *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of square integrable reverse martingale differences with respect to a non-increasing filtration $(\mathcal{G}_n)_{n\in\mathbb{N}}$. Assume that $\delta_n^2 := \sum_{k\geq n}\mathbb{E}(\xi_k^2) < \infty$. Then $V_n^2 := \sum_{k\geq n}\mathbb{E}(\xi_k^2|\mathcal{G}_{k+1})$ is well defined $\mathbb{P}$-a.s. and in $\mathbb{L}^2$ as well as $R_n = \sum_{k\geq n}\xi_k$. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of non-increasing positive numbers with $\alpha_n = O(\delta_n^2)$ and $\alpha_n/\delta_n^4 \to \infty$. Assume that*

$$V_n^2 - \delta_n^2 = o(\alpha_n) \qquad \mathbb{P}\text{-a.s.} \tag{2.1}$$

$$\sum_{n\geq 1}\alpha_n^{-\nu}\mathbb{E}(|\xi_n|^{2\nu}) \quad \text{for some } 1\leq\nu\leq 2\,. \tag{2.2}$$

Then, enlarging our probability space if necessary, it is possible to find a standard Brownian motion $(B_t)_{t\geq 0}$, such that

$$R_n - B_{\delta_n^2} = o\Big(\big(\alpha_n\big(|\log(\delta_n^2/\alpha_n)| + \log\log(\alpha_n^{-1})\big)\big)^{1/2}\Big) \qquad \mathbb{P}\text{-a.s.} \tag{2.3}$$

Remark 2.2 *It follows from the proof that if* (2.1) *holds with "big O" instead of "little o" then* (2.3) *holds with the same change.*

Now, we give a result for partial sums associated with a sequence of reverse martingale differences rather than for tail series. It may be viewed as the analogue of Theorem 2.1 in Shao [35] but in the context of reverse martingale differences.

Theorem 2.3 *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of square integrable reverse martingale differences with respect to a non-increasing filtration $(\mathcal{G}_n)_{n\in\mathbb{N}}$. Assume that $\sigma_n^2 := \sum_{k=1}^n\mathbb{E}(X_k^2) \to \infty$ and that $\sup_n\mathbb{E}(X_n^2) < \infty$. Let $(a_n)_{n\in\mathbb{N}}$ be a non-decreasing sequence of positive numbers such that $(a_n/\sigma_n^2)_{n\in\mathbb{N}}$ is non-increasing and $(a_n/\sigma_n)_{n\in\mathbb{N}}$ is non-decreasing. Assume that*

$$\sum_{k=1}^n(\mathbb{E}(X_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2)) = o(a_n) \qquad \mathbb{P}\text{-a.s.} \tag{2.4}$$

$$\sum_{n\geq 1}a_n^{-\nu}\mathbb{E}(|X_n|^{2\nu}) < \infty \quad \text{for some } 1\leq\nu\leq 2\,. \tag{2.5}$$

Then, enlarging our probability space if necessary, it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables with $\mathbb{E}(Z_k^2) = \mathbb{E}(X_k^2)$ such that

$$\sup_{1\leq k\leq n}\Big|\sum_{i=1}^k X_i - \sum_{i=1}^k Z_i\Big| = o\big(\big(a_n(|\log(\sigma_n^2/a_n)| + \log\log a_n)\big)^{1/2}\big) \qquad \mathbb{P}\text{-a.s.} \tag{2.6}$$

Remark 2.4 *An inspection of the proof allows to weaken slightly some of the assumptions as follows. Assume that $\mathbb{E}(X_n^2) = O(\sigma_n^{2s})$ for some $0\leq s<1$ instead of boundedness and assume that there exists $C>1$, such that for every $n\geq 1$, $\sup_{k\geq n}(a_k/\sigma_k^2) \leq Ca_n/\sigma_n^2$, and $\inf_{k\geq n}(a_k/\sigma_k) \geq a_n/(C\sigma_n)$, instead of the corresponding monotonicity assumptions.*

We derive now the functional LIL for the partial sums associated with a stationary and ergodic sequence of reverse martingale differences. In the next corollaries, we make use of $\theta : \Omega \mapsto \Omega$ a measurable transformation preserving the probability $\mathbb{P}$.

Corollary 2.5 *Let X_0 in $\mathbb{L}^2$ and for $n\geq 1$, $X_n = X_0\circ\theta^n$. Assume that θ is ergodic. Assume that $(X_n)_{n\geq 0}$ is a sequence of reverse martingale differences with respect to a non-increasing filtration $(\mathcal{G}_n)_{n\in\mathbb{N}}$. Enlarging our probability space if necessary, it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables with $\mathbb{E}(Z_k^2) = \mathbb{E}(X_1^2)$ such that*

$$\sup_{1\leq k\leq n}\Big|\sum_{i=1}^k X_i - \sum_{i=1}^k Z_i\Big| = o\big(\sqrt{n\log\log n}\big) \qquad \mathbb{P}\text{-a.s.} \tag{2.7}$$

Remark 2.6 *The Strassen functional law of the iterated logarithm (FLIL) follows from the corollary. Notice that a semi-FLIL has been proved by Wu [37].*

We now give rate results in the strong invariance principle for the partial sums associated with a stationary sequence of reverse martingale differences.

Corollary 2.7 *Let $2 < p < 4$. Let X_0 be in $\mathbb{L}^p$ and for $n \geq 1$, $X_n = X_0 \circ \theta^n$. Assume that $(X_n)_{n\geq 0}$ is a sequence of reverse martingale differences with respect to a non-increasing filtration $(\mathcal{G}_n)_{n\in\mathbb{N}}$. Let $b(\cdot)$ be a positive non-decreasing slowly varying function, such that $x \mapsto x^{2/p-1}b(x)$ is non-increasing. Assume that*

$$\sum_{k=1}^{n}(\mathbb{E}(X_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2)) = o(n^{2/p}b(n)) \qquad \mathbb{P}\text{-a.s.} \tag{2.8}$$

Enlarging our probability space if necessary, it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables with $\mathbb{E}(Z_k^2) = \mathbb{E}(X_1^2)$ such that

$$\sup_{1\leq k\leq n}\Big|\sum_{i=1}^{k} X_i - \sum_{i=1}^{k} Z_i\Big| = o\big(n^{1/p}\sqrt{b(n)\log n}\big) \qquad \mathbb{P}\text{-a.s.} \tag{2.9}$$

Corollary 2.8 *Let $\mathcal{G}_0$ be a sub-σ-algebra of $\mathcal{A}$ satisfying $\theta^{-1}(\mathcal{G}_0) \subseteq \mathcal{G}_0$. Let X_0 be a $\mathcal{G}_0$-measurable random variable in $\mathbb{L}^4$. For $n \geq 1$, let $X_n = X_0 \circ \theta^n$ and $\mathcal{G}_n = \theta^{-n}(\mathcal{G}_0)$. Assume that θ is ergodic and that $\mathbb{E}(X_n|\mathcal{G}_{n+1}) = 0$ $\mathbb{P}$-a.s. Assume that*

$$\sum_{k=1}^{n}(\mathbb{E}(X_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2)) = O((n\log\log n)^{1/2}) \qquad \mathbb{P}\text{-a.s.}$$

Enlarging our probability space if necessary, it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables with $\mathbb{E}(Z_k^2) = \mathbb{E}(X_1^2)$ such that

$$\sup_{1\leq k\leq n}\Big|\sum_{i=1}^{k} X_i - \sum_{i=1}^{k} Z_i\Big| = O\big(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}\big) \qquad \mathbb{P}\text{-a.s.} \tag{2.10}$$

3 ASIP with rates for non-invertible maps

All along the section, we consider (X, Σ, ν) a probability space and T a *non-invertible* Σ-measurable transformation on X, preserving the probability ν. We assume T to be ergodic, that is for every $A \in \Sigma$ such that $T^{-1}(A) = A$, $\nu(A) \in \{0, 1\}$.

We introduce now the Perron-Frobenius operator (or transfert operator) K associated with T, that is

$$\int_X f\, g \circ T\, d\nu = \int_X Kf\, g\, d\nu\,, \tag{3.1}$$

for every positive measurable functions f, g on X.

We assume that K is induced by a transition probability that we still denote by K. In particular, for every positive measurable f on X, $Kf = \int_X f(y)K(\cdot, dy)$. Recall that this assumption is always satified when X is a Polish space and Σ is the σ-algebra of its Borel sets (see Neveu [30, Proposition V.4.3]).

All along the paper, we shall denote $\|\cdot\|_{p,\nu}$ the $\mathbb{L}^p$-norm on (X, Σ, ν).

3.1 General results

The aim of this section is to give sufficient general conditions on K and f to obtain explicit rates in the strong invariance principle for the partial sums $\sum_{i=0}^{n-1}(f \circ T^i - \nu(f))$. The strategy to obtain such results is to approximate (with a suitable rate) the partial sum associated with the iterates of T by a sum of reverse martingale differences (see our propositions 5.1 and 5.2) and then to apply our results of Section 2.

Our first result is the following:

Theorem 3.1 *Let (X, Σ, ν), T and K be as above. Let $2 < p \le 4$ and $t > 2/p$. Let $f : X \to \mathbb{R}$ be a measurable function such that $\nu(|f|^p) < \infty$. Let $g = f - \nu(f)$. Assume that*

$$\sum_{n\ge 2} \frac{n^{p-1}}{n^{2/p}(\log n)^{\frac{(t-1)p}{2}}} \|K^n g\|_{p,\nu}^p < \infty \ \text{ and } \sum_{n\ge 2} \frac{n^{3p/4}}{n^2(\log n)^{\frac{(t-1)p}{2}}} \|K^n g\|_{2,\nu}^{p/2} < \infty\,, \tag{3.2}$$

and

$$\sum_{n\ge 2} \frac{1}{n^2(\log n)^{(t-1)p/2}} \Big(\sum_{i=1}^{n} \sum_{j=0}^{n-i} \|K^i(gK^j(g)) - \nu(gK^j(g))\|_{\frac{p}{2},\nu} \Big)^{p/2} < \infty\,. \tag{3.3}$$

Then, the series

$$\sigma^2 = \sigma^2(f) = \nu((f-\nu(f))^2) + 2\sum_{k>0} \nu((f-\nu(f)) f\circ T^k) \tag{3.4}$$

converges absolutely to some non negative number, and, enlarging the probability space (X, Σ, ν) if necessary, there exists a sequence $(Z_i)_{i\ge 0}$ of iid Gaussian random variables with mean zero and variance σ^2 such that

$$\sup_{1\le k\le n} \Big| \sum_{i=0}^{k-1} (f\circ T^i - \nu(f)) - Z_i) \Big| = o(n^{1/p}(\log n)^{(1+t)/2}) \ \ \nu\text{-a.s.} \tag{3.5}$$

Theorem 3.1 is the analogue of Theorem 2.3 in [4] in case of non-invertible maps.

When $p = 4$ and if we reinforce the conditions, we can obtain a better rate in the strong invariance principle (3.5). More precisely, the following result is valid.

Theorem 3.2 *Let (X, Σ, ν), T and K be as above. Let $f : X \to \mathbb{R}$ be a measurable function such that $\nu(|f|^4) < \infty$. Let $g = f - \nu(f)$. Assume that*

$$\sum_{n\ge 2} n^{5/2}(\log n)^3 \|K^n g\|_{4,\nu}^4 < \infty \ \text{ and } \sum_{n\ge 2} n(\log n)^3 \|K^n g\|_{2,\nu}^2 < \infty\,, \tag{3.6}$$

and

$$\sum_{n\ge 2} \frac{(\log n)^3}{n^2} \Big(\sum_{i=1}^{n} \sum_{j=0}^{n-i} \|K^i(gK^j(g)) - \nu(gK^j(g))\|_{2,\nu} \Big)^2 < \infty\,. \tag{3.7}$$

Then, the series σ^2 defined in (3.4) converges absolutely, and, enlarging the probability space (X, Σ, ν) if necessary, there exists a sequence $(Z_i)_{i\ge 0}$ of iid Gaussian random variables with mean zero and variance σ^2 such that

$$\sup_{1\le k\le n} \Big| \sum_{i=0}^{k-1} (f\circ T^i - \nu(f)) - Z_i) \Big| = O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}) \ \ \nu\text{-a.s.} \tag{3.8}$$

3.2 Applications

3.2.1 Intermittent maps of the interval

For ease of exposition, we consider the family of maps $T : X \to X$, where X is the interval $[0, 1]$, studied in [21]. For $\gamma \in (0, 1)$, let

$$T(x) = \begin{cases} x(1 + 2^\gamma x^\gamma) & \text{if } x \in [0, 1/2] \\ 2x - 1 & \text{if } x \in (1/2, 1]\ . \end{cases} \tag{3.9}$$

Recall that T admits a unique absolutely continuous invariant probability measure ν, with density h_ν. Moreover, it is ergodic, has full support, and $h_\nu(x)/x^{-\gamma}$ is bounded from above and below.

If f is Hölder continuous, the behavior of $\sum_{i=0}^{n-1}(f \circ T^i - \nu(f))$ is very well understood, thanks to Young [40] and Melbourne-Nicol [24]: these sums satisfy the almost sure invariance principle with rate $(n \log\log n)^{1/2}$ for $\gamma < 1/2$. In addition, still in case of Hölder continuous functions f, Melbourne and Nicol [25] obtained the following explicit error term in the almost sure invariance principle (see their Theorem 1.6 and their Remark 1.7): Let $p > 2$ and $0 < \gamma < 1/p$, then the error term in the almost sure invariance principle is $O(n^{\beta+\varepsilon})$ where $\varepsilon > 0$ is arbitrarily small and $\beta = \frac{\gamma}{2} + \frac{1}{4}$ if γ belongs to $]1/4, 1/2[$

and $\beta = \frac{3}{8}$ if $\gamma \leq 1/4$. When f is less regular the situation may be more delicate. Indeed, functions with countably many discontinuities are not easily amenable to the Young's tower method. Recently, Merlevède and Rio [29] have obtained explicit rates in the almost sure invariance principle when f belongs to weak $\mathbb{L}^p$-like spaces ($p > 2$) as defined in Definition 3.3 above (note that these spaces contain functions with countably many discontinuities). We refer also to the recent paper [6] for the boundary case $p = 2$.

Definition 3.3 *A function H from $\mathbb{R}_+$ to $[0, 1]$ is a tail function if it is non-increasing, right continuous, converges to zero at infinity, and $x \mapsto xH(x)$ is integrable. If ν is a probability measure on $\mathbb{R}$ and H is a tail function, let $\mathrm{Mon}(H, \nu)$ denote the set of functions $f : \mathbb{R} \to \mathbb{R}$ which are monotonic on some interval and null elsewhere and such that $\mu(|f| > t) \leq H(t)$. Let $\mathrm{Mon}^c(H, \nu)$ be the closure in $\mathbb{L}^1(\nu)$ of the set of functions which can be written as $\sum_{\ell=1}^{L} a_\ell f_\ell$, where $\sum_{\ell=1}^{L} |a_\ell| \leq 1$ and $f_\ell \in \mathrm{Mon}(H, \nu)$.*

In particular, for $\gamma < 1/3$ and f of bounded variation, Corollary 3.1 in [29] shows that $\sum_{i=0}^{n-1}(f \circ T^i - \nu(f))$ satisfies the almost sure principle with rate $O(n^{1/3}(\log n)^{1/2}(\log\log n)^{(1+\varepsilon)/3})$. Applying Theorem 3.1, we get the following result:

Corollary 3.4 *Let T be a map as defined in (3.9) with parameter $\gamma \in (0, 1/2)$. Let $p \in]2, 4]$ and $\delta = p + 1 - 2/p$. Assume that $\delta \leq 1/\gamma$ and let H be a tail function with*

$$\int_0^\infty x^{p-1}(H(x))^{\frac{1-\gamma\delta}{1-\gamma}} dx < \infty \,. \tag{3.10}$$

Then, for any $f \in \mathrm{Mon}^c(H, \nu)$, the conclusion of Theorem 3.1 holds with $t = 1$.

Note that Corollary 3.4 can be extended to generalized Pomeau-Manneville maps (or GPM maps) of parameter $\gamma \in (0, 1)$ as defined in Dedecker, Gouëzel and Merlevède [5]. Notice also that when f has bounded variation, (3.10) is trivially satisfied, hence, Corollary 3.4 applies as soon as $\gamma \leq \delta^{-1}$. Therefore when $\gamma < 3/10$, we obtain better rates than the one obtained in [29, Corollary 3.1]. In particular, if $\gamma \leq 2/9$, we obtain the rate $o(n^{1/4} \log n)$ in the almost sure invariance principle.

3.2.2 Expanding maps of the interval

Several classes of uniformly expanding maps of the interval are considered in the literature. We shall consider here the very general definition of [33] to allow infinitely many branches.

Definition 3.5 *A map $T : [0, 1] \to [0, 1]$ is uniformly expanding, mixing and with density bounded from below if it satisfies the following properties:*

1. *There is a (finite or countable) partition of T into subintervals I_n on which T is strictly monotonic, with a C^2 extension to its closure $\overline{I_n}$, satisfying Adler's condition $|T''|/|T'|^2 \leq C$, and with $|T'| \geq \lambda$ (where $C > 0$ and $\lambda > 1$ do not depend on I_n).*
2. *The length of $T(I_n)$ is bounded from below.*
3. *In this case, T has finitely many absolutely continuous invariant measures, and each of them is mixing up to a finite cycle. We assume that T has a single absolutely continuous invariant probability measure ν, and that it is mixing.*
4. *Finally, we require that the density h of ν is bounded from below on its support.*

We refer to [6] for some comments on this definition (see the few lines before their Definition 1.1).

As in case of intermittent maps, the almost sure behavior of $\sum_{i=0}^{n-1}(f \circ T^i - \nu(f))$ is well understood when f is Hölder continuous (see [17]), but strong almost sure approximation results can also be obtained when the observables belong to $\mathbb{L}^p$-like spaces as defined below.

Definition 3.6 *If μ is a probability measure on $\mathbb{R}$ and $p \in [2, \infty)$, $M \in (0, \infty)$, let $\mathrm{Mon}_p(M, \mu)$ denote the set of functions $f : \mathbb{R} \to \mathbb{R}$ which are monotonic on some interval and null elsewhere and such that $\mu(|f|^p) \leq M^p$. Let $\mathrm{Mon}_p^c(M, \mu)$ be the closure in $\mathbb{L}^p(\mu)$ of the set of functions which can be written as $\sum_{\ell=1}^{L} a_\ell f_\ell$, where $\sum_{\ell=1}^{L} |a_\ell| \leq 1$ and $f_\ell \in \mathrm{Mon}_p(M, \mu)$.*

Definition 3.6 deals with $\mathbb{L}^p$-like spaces, with an additional monotonicity condition. Note that, in previous papers (see for instance [6]), the closure in $\mathbb{L}^1(\mu)$ was used in Definition 3.6. It turns out that both definitions coincide. Indeed, a sequence bounded in $\mathbb{L}^p(\mu)$ and converging in $\mathbb{L}^1(\mu)$, converges for the weak topology in $\mathbb{L}^p(\mu)$. To conclude, recall that, by the Hahn-Banach theorem, in any Banach space, the weak closure of a convex set is equal to its strong closure.

Applying Theorem 3.2, we obtain the following result:

Corollary 3.7 *Let T be a uniformly expanding map as defined in Definition 3.5, with absolutely continuous invariant measure ν. Then, for any $f \in \mathrm{Mon}_4^c(M,\nu)$ for some $M > 0$, the conclusion of Theorem 3.2 holds.*

When $p \in]2,4[$ and $f \in \mathrm{Mon}_4^c(M,\nu)$ for some $M > 0$, Theorem 3.1 can be applied which would give rates of order $o(n^{1/p}(\log n)^{\eta})$, with $\eta > 1/2$, in the almost sure invariance principle. However, in this situation the power of the logarithm term can be weakened. More precisely, we can obtain the following result:

Theorem 3.8 *Let T be a uniformly expanding map as defined in Definition 3.5, with absolutely continuous invariant measure ν. Let $p \in]2,4[$ and $f \in \mathrm{Mon}_p^c(M,\nu)$ for some $M > 0$. Then, the series σ^2 defined in* (3.4) *converges absolutely and, enlarging the probability space $([0,1],\nu)$ if necessary, there exists a sequence $(Z_i)_{i\geq 0}$ of iid Gaussian random variables with mean zero and variance σ^2 such that*

$$\sup_{1\leq k\leq n}\Big|\sum_{i=0}^{k-1}(f\circ T^i - \nu(f)) - Z_i)\Big| = o(n^{1/p}(\log n)^{1-2/p})\ \nu\text{-a.s.}$$

Let us mention that the boundary case $p = 2$ has been handled in [6].

3.2.3 Smooth dynamic

We consider here smooth dynamical systems as defined by Baladi [1].

Let X be a compact connected Riemann manifold. Denote by d the distance induced by the Riemann metric and by m the Lebesgue measure on X.

Let $T : X \to X$ be a $C^{1,\alpha}$ map, i.e., T is differentiable, with α-Hölder (with respect to d) tangent map.

Assume moreover that T is uniformly expanding, i.e. that there exists $\gamma > 1$ such that for every $x \in X$ and every $v \in T_xX$,

$$\|(DT)_x v\| \geq \gamma\|v\|\,,$$

where $\|\cdot\|$ denotes the norm induced by the Riemann metric.

In all that subsection, we shall assume the above conditions. It follows then from Theorem 2.1 of [1] that there exists a unique absolutely continuous (with respect to m) and T-invariant Borel measure, say $\nu = \varphi m$. Moreover, φ is α-Hölder and $\inf_X \varphi > 0$.

We shall consider functions with a specific modulus of continuity.

Let $c : [0,+\infty[\to [0,+\infty[$ be a non-decreasing concave function with $c(0) = 0$.

Denote by Γ_c the space of real-valued (continuous) functions f on X, such that

$$|f(x) - f(y)| \leq c(d(x,y)) \qquad \forall x,y \in X\,.$$

Corollary 3.9 *Let $p \in]2,4]$. Let c be a non-decreasing and non-negative concave function such that*

$$\int_0^1 (\log(1/u))^{\delta}\frac{c^{p/2}(u)}{u}\,du < \infty \quad \text{where } \delta = \frac{2p+\sqrt{1+4p(p-2)}}{4} - \frac{7}{4}\,. \tag{3.11}$$

Then, for every $f \in \Gamma_c$, the conclusion of Theorem 3.1 holds with $t = 1$.

4 Proofs of the reverse martingale's results

We start by recalling the following estimate of Hanson and Russo [14, Theorem 3.2A]

Lemma 4.1 *Let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Then*

$$\lim_{a\to\infty}\sup_{t\geq 0}\sup_{0\leq s\leq a}\frac{|B_{t+s}-B_t|}{(2a[\log((t+a)/a))+\log\log a])^{1/2}}=1 \qquad \mathbb{P}\text{-}a.s. \tag{4.1}$$

We also recall the following convergence result for reverse martingales. The proof comes from an application of Burkholder's inequality in this context.

Lemma 4.2 *Let $(\xi_n)_{n\geq 1}$ be a sequence of reverse martingale differences in $\mathbb{L}^p$, $1\leq p\leq 2$, with respect to a non-increasing filtration $(\mathcal{G}_n)_{n\geq 1}$. Assume that $\sum_{n\geq 1}\mathbb{E}(|\xi_n|^p)<\infty$. Then $\sum_{n\geq 1}\xi_n$ converges $\mathbb{P}$-a.s. and in $\mathbb{L}^p$.*

4.1 Proof of Proposition 2.1.

The $\mathbb{L}^2$ and a.s. convergence of $\sum_{k\geq 1}\xi_k$ follows from Lemma 4.2. By Theorem 2 of Scott and Huggins [34], enlarging our probability space if necessary, there exists a Brownian motion $(B_t)_{t\geq 0}$, a non-increasing filtration $(\mathcal{H}_n)_{n\in\mathbb{N}}$ and a non-increasing process $(\tau_n)_{n\in\mathbb{N}}$ adapted to $(\mathcal{H}_n)_{n\in\mathbb{N}}$, such that

$$R_n = B_{\tau_n} \qquad \mathbb{P}\text{-a.s.}$$

Moreover, writing $t_n := \tau_n - \tau_{n+1}\geq 0$ $\mathbb{P}$-a.s., we have

$$\mathbb{E}(t_n|\mathcal{H}_{n+1}) = \mathbb{E}(\xi_n^2|\mathcal{G}_{n+1}) \qquad \mathbb{P}\text{-a.s.}\,, \tag{4.2}$$

$$\mathbb{E}(t_n^{p/2}|\mathcal{H}_{n+1}) \leq C_p\mathbb{E}(|\xi_n|^p|\mathcal{G}_{n+1}) \qquad \mathbb{P}\text{-a.s.} \quad \text{for every } p>1 \quad . \tag{4.3}$$

Hence, using (4.2) twice,

$$\tau_n - \mathbb{E}(\tau_n) = \tau_n - \delta_n^2 = \sum_{k\geq n}\big(t_k - \mathbb{E}(t_k|\mathcal{H}_{k+1})\big) + V_n^2 - \delta_n^2 \qquad \mathbb{P}\text{-a.s.}$$

But it follows from (2.2) and (4.3) that $\sum_{n\geq 1}\alpha_n^{-\nu}\mathbb{E}(t_n^\nu)<\infty$ which implies, by Lemma 4.2, that $\sum_{k\geq 1}\alpha_k^{-1}(t_k-\mathbb{E}(t_k|\mathcal{H}_{k+1}))$ converges $\mathbb{P}$-a.s. Then, by an analogue to the Kronecker lemma (see e.g. Heyde [16, Lemma 1]), $\sum_{k\geq n}(t_k-\mathbb{E}(t_k|\mathcal{H}_{k+1}))=o(\alpha_n)$ $\mathbb{P}$-a.s. Together with (2.1), this implies in particular that $\tau_n-\delta_n^2=o(\alpha_n)$ $\mathbb{P}$-a.s.

For every $t>0$ define $\widetilde{B}_t = tB_{1/t}$, and $\widetilde{B}_0=0$. It is well-known that $(\widetilde{B}_t)_{t\geq 0}$ is a standard Brownian motion. We have $B_{\tau_n}-B_{\delta_n^2}=\tau_n(\widetilde{B}_{1/\tau_n}-\widetilde{B}_{1/\delta_n^2})+(\tau_n-\delta_n^2)\widetilde{B}_{1/\delta_n^2}$. By the law of the iterated logarithm for $(\widetilde{B}_t)_{t\geq 0}$ (or using that the supremum in (4.1) is greater than what we have for $t=0$ and $s=1/\delta_n^2$), we see that $\widetilde{B}_{1/\delta_n^2}=O(\delta_n^{-1}(\log\log(1/\delta_n))^{1/2})$ $\mathbb{P}$-a.s. Hence, since $\alpha_n=O(\delta_n^2)$, $(\tau_n-\delta_n^2)\widetilde{B}_{1/\delta_n^2}=o((\alpha_n\log\log(1/\alpha_n))^{1/2})$ $\mathbb{P}$-a.s.
Let us deal now with $\tau_n(\widetilde{B}_{1/\tau_n}-\widetilde{B}_{1/\delta_n^2})$. With this aim, we shall use (4.1). Since $\alpha_n=O(\delta_n^2)$ and $\tau_n-\delta_n^2=o(\alpha_n)$ $\mathbb{P}$-a.s., we have $|1/\tau_n-1/\delta_n^2|=o(1/\alpha_n)$ $\mathbb{P}$-a.s. Define $u_n:=\alpha_n/(\tau_n\delta_n^2)$, $\varepsilon_n:=\max(|\delta_n^2-\tau_n|/\alpha_n, u_n^{-1/2})$, $s_n:=|1/\tau_n-1/\delta_n^2|$, $a_n:=\varepsilon_n u_n$ and $v_n:=\min(1/\delta_n^2,1/\tau_n)$. Notice that $a_n\to\infty$, $\varepsilon_n\to 0$, $v_n+s_n=\max(1/\delta_n^2,1/\tau_n)$ and $|\widetilde{B}_{1/\tau_n}-\widetilde{B}_{1/\delta_n^2}|=|\widetilde{B}_{v_n+s_n}-\widetilde{B}_{v_n}|$. By (4.1), we have

$$\begin{aligned}&\frac{|\widetilde{B}_{v_n+s_n}-\widetilde{B}_{v_n}|}{(2a_n[\log((v_n+a_n)/a_n)+\log\log a_n])^{1/2}}\\ &\leq \sup_{t\geq 0}\sup_{0\leq s\leq a_n}\frac{|\widetilde{B}_{t+s}-\widetilde{B}_t|}{(2a_n[\log((t+a_n)/a_n))+\log\log a_n])^{1/2}}\to 1 \qquad \mathbb{P}\text{-a.s.}\end{aligned}$$

In particular, we have $|\widetilde{B}_{1/\tau_n}-\widetilde{B}_{1/\delta_n^2}|=O\big([\varepsilon_n u_n(|\log(\delta_n^2/(\alpha_n\varepsilon_n))|+\log\log(u_n\varepsilon_n))]^{1/2}\big)$ $\mathbb{P}$-a.s. Then, using that $|\log(\delta_n^2/(\alpha_n\varepsilon_n))|\leq|\log(\delta_n^2/\alpha_n)|+|\log\varepsilon_n|$ and that $\varepsilon_n u_n\log\log(\varepsilon_n u_n)=o(u_n\log\log u_n)$, we obtain

$$\tau_n(\widetilde{B}_{1/\tau_n}-\widetilde{B}_{1/\delta_n^2})=o\big([\alpha_n(|\log(\delta_n^2/\alpha_n)|+\log\log(\alpha_n/\delta_n^4)]^{1/2}\big) \qquad \mathbb{P}\text{-a.s.}\,, \tag{4.4}$$

which proves the result, since $1/\delta_n^4=O(1/\alpha_n^2)$. □

Remark 4.3 *It follows from the proof that the assumption* (2.2) *may be replaced by* $\sum_{k\geq n}(t_k-\mathbb{E}(t_k|\mathcal{H}_{k+1}))=o(\alpha_n)$ *$\mathbb{P}$-a.s.*

4.2 Proof of Theorem 2.3.

Define $\xi_n := X_n/\sigma_n^2$. Then, since $\mathbb{E}(\xi_k^2) = (\sigma_k^2 - \sigma_{k-1}^2)\sigma_k^{-4}$, by comparing sums and integrals, it follows that $\sum_{k\geq 1}\mathbb{E}(\xi_k^2) < \infty$. Using the notations $V_n^2 = \sum_{k\geq n}(\mathbb{E}(\xi_k^2|\mathcal{G}_{k+1})$ and $\delta_n^2 = \sum_{k\geq n}\mathbb{E}(\xi_k^2)$, and writing $T_n := \sum_{k=1}^n (\mathbb{E}(X_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2))$, we have

$$V_n^2 - \delta_n^2 = \sum_{k\geq n}\frac{T_k - T_{k-1}}{\sigma_k^4} = \sum_{k\geq n} T_k\Big(\frac{1}{\sigma_k^4} - \frac{1}{\sigma_{k+1}^4}\Big) - \frac{T_{n-1}}{\sigma_n^4}\,.$$

Using (2.4) and that $(\sigma_n)_{n\in\mathbb{N}}$ and $(a_n/\sigma_n^2)_{n\in\mathbb{N}}$ are respectively non-decreasing and non-increasing, we obtain

$$|V_n^2 - \delta_n^2| = o\Big(\frac{a_n}{\sigma_n^2}\Big)\sum_{k\geq n}\Big(\frac{1}{\sigma_k^2} - \frac{1}{\sigma_{k+1}^2}\Big) + o\Big(\frac{a_n}{\sigma_n^4}\Big) = o\Big(\frac{a_n}{\sigma_n^4}\Big)\,.$$

We want to apply Proposition 2.1 to (ξ_n) with $\alpha_n := a_n/\sigma_n^4$. Using (2.5), we have

$$\sum_{i\geq 1}\alpha_i^{-\nu}\mathbb{E}(|\xi_i|^{2\nu}) = \sum_{n\geq 1} a_i^{-\nu}\mathbb{E}(|X_i|^{2\nu}) < \infty\,,$$

hence condition (2.2) holds. It remains to prove that $\alpha_n = O(\delta_n^2)$ and that $\alpha_n/\delta_n^4 \to \infty$. With this aim, we first notice that

$$\frac{1}{\sigma_{n-1}^2} - \delta_n^2 = \sum_{k\geq n}\int_{\sigma_{k-1}^2}^{\sigma_k^2}\Big(\frac{1}{x^2} - \frac{1}{\sigma_k^4}\Big)dx\,.$$

Hence, using that $\sup_n \mathbb{E}(X_n^2) < \infty$, it follows that $\sigma_n = O(\sigma_{n-1})$ and

$$0 \leq \frac{1}{\sigma_{n-1}^2} - \delta_n^2 \leq \sum_{k\geq n}\mathbb{E}(X_k^2)\Big(\frac{1}{\sigma_{k-1}^4} - \frac{1}{\sigma_k^4}\Big) = O\Big(\sum_{k\geq n}\frac{\sigma_k^2 - \sigma_{k-1}^2}{\sigma_k^6}\Big) = O\Big(\frac{\delta_n^2}{\sigma_n^2}\Big)\,.$$

In particular, since $\delta_n^2 = O(\sigma_{n-1}^{-2}) = O(\sigma_n^{-2})$ and $|\sigma_n^{-2} - \sigma_{n-1}^{-2}| = O(\sigma_n^{-4})$, we have

$$\Big|\frac{1}{\sigma_n^2} - \delta_n^2\Big| = O\Big(\frac{1}{\sigma_n^4}\Big)\,. \tag{4.5}$$

Since $a_n\sigma_n^{-2}$ is non-increasing, (4.5) implies that $\alpha_n = O(\delta_n^2)$. In addition since a_n is tending to infinity, (4.5) entails also that $\alpha_n/\delta_n^4 \to \infty$.

By Proposition 2.1, enlarging our probability space if necessary, there exists a standard Brownian motion $(B_t)_{t\geq 0}$, such that (2.3) holds with $\delta_n^2 = \sum_{k\geq n}(\sigma_k^2 - \sigma_{k-1}^2)\sigma_k^{-4}$. Now, for every $t > 0$ define $\widetilde{B}_t = tB_{1/t}$, and $\widetilde{B}_0 = 0$ (recall that $(\widetilde{B}_t)_{t\geq 0}$ is a standard Brownian motion). Notice that

$$B_{1/\sigma_n^2} - B_{\delta_n^2} = \sigma_n^{-2}(\widetilde{B}_{\sigma_n^2} - \widetilde{B}_{1/\delta_n^2}) + (\sigma_n^{-2} - \delta_n^2)\widetilde{B}_{1/\delta_n^2}\,.$$

By (4.5) and the law of the iterated logarithm for $(\widetilde{B}_t)_{t\geq 0}$, we derive that

$$(\sigma_n^{-2} - \delta_n^2)\widetilde{B}_{1/\delta_n^2} = O(\sigma_n^{-2}(\log\log(\sigma_n))^{1/2}) = o\bigg(\frac{\big(a_n(\log\log a_n)\big)^{1/2}}{\sigma_n^2}\bigg) \quad \mathbb{P}\text{-a.s.}$$

To deal now with $\sigma_n^{-2}(\widetilde{B}_{\sigma_n^2} - \widetilde{B}_{1/\delta_n^2})$, we use the same arguments as the ones used to derive (4.4) (with σ_n^{-2} replacing τ_n). Hence we infer that

$$\sigma_n^{-2}(\widetilde{B}_{\sigma_n^2} - \widetilde{B}_{1/\delta_n^2}) = o\bigg(\frac{\big(a_n(|\log(\sigma_n^2/a_n)| + \log\log a_n)\big)^{1/2}}{\sigma_n^2}\bigg) \qquad \mathbb{P}\text{-a.s.}$$

So, overall, it follows that

$$|R_n - B_{1/\sigma_n^2}| = o\bigg(\frac{\big(a_n(|\log(\sigma_n^2/a_n)| + \log\log a_n)\big)^{1/2}}{\sigma_n^2}\bigg) \qquad \mathbb{P}\text{-a.s.}\,, \tag{4.6}$$

where $R_n = \sum_{k \geq n} X_n/\sigma_n^2$.

Write $\widetilde{Z}_n := \sigma_n^2(B_{1/\sigma_n^2} - B_{1/\sigma_{n+1}^2})$. By independence of the increments, $(\widetilde{Z}_n)$ is a sequence of independent centered Gaussian variables. Notice that, by stationarity of the increments $\mathbb{E}(\widetilde{Z}_n^2) = \sigma_n^2 \mathbb{E}(X_n^2)/\sigma_{n+1}^2$.

We have

$$\begin{aligned}\sum_{k=1}^n X_k - \sum_{k=1}^n \widetilde{Z}_k &= \sum_{k=1}^n \sigma_k^2((R_k - B_{1/\sigma_k^2}) - (R_{k+1} - B_{1/\sigma_{k+1}^2})) \\ &= \sum_{k=2}^n (R_k - B_{1/\sigma_k^2})(\sigma_k^2 - \sigma_{k-1}^2) + \sigma_1^2(R_1 - B_{1/\sigma_1^2}) - \sigma_n^2(R_{n+1} - B_{1/\sigma_{n+1}^2}) \,.\end{aligned}$$

Using that (σ_n), (σ_n^2/a_n), (a_n/σ_n) and (a_n) are non-decreasing, and taking into account (4.6), we deduce that

$$\begin{aligned}\sum_{k=1}^n X_k - \sum_{k=1}^n \widetilde{Z}_k &= o\Big(\Big(\frac{a_n(|\log(\sigma_n^2/a_n)| + \log\log a_n)}{\sigma_n}\Big)^{1/2}\Big) \sum_{k=2}^n \frac{\sigma_k^2 - \sigma_{k-1}^2}{(\sigma_k^2)^{3/4}} \\ &= o\Big(\big(a_n(|\log(\sigma_n^2/a_n)| + \log\log a_n\big)^{1/2}\Big) \,,\end{aligned}$$

where we used that $\sum_{k=2}^n (\sigma_k^2 - \sigma_{k-1}^2)(\sigma_k^2)^{-3/4} = O(\int_0^{\sigma_n^2} dx/x^{3/4})$.

Finally, define $Z_n := \widetilde{Z}_n \sigma_{n+1}/\sigma_n$. Notice that $|Z_n - \widetilde{Z}_n| \leq C|\widetilde{Z}_n|/\sigma_n$ for some $C > 0$. Hence $\sum_{n \geq 1} \mathbb{E}((Z_n - \widetilde{Z}_n)^2)/a_n \leq \sum_{n \geq 1} (\sigma_n^2 - \sigma_{n-1}^2)/\sigma_n^3 < \infty$. So, by the Kolmogorov theorem (see also Lemma 4.2), $\sum_{n \geq 1} (Z_n - \widetilde{Z}_n)/\sqrt{a_n}$ converges $\mathbb{P}$-a.s., and (2.6) follows from the Kronecker lemma. □

4.3 Proof of Corollary 2.5.

Assume that $\mathbb{E}(X_1^2) \neq 0$, otherwise there is nothing to prove. Notice first that by stationarity and Fubini theorem,

$$\sum_{n \geq 1} \frac{\mathbb{E}(|X_n| \mathbf{1}_{\{|X_n| > \sqrt{n}\}})}{\sqrt{n}} = \mathbb{E}\Big(|X_1| \sum_{1 \leq n < X_1^2} \frac{1}{\sqrt{n}}\Big) \leq C\mathbb{E}(X_1^2) < \infty \,.$$

Hence,

$$\sum_{n \geq 1} n^{-1/2} |X_n| \mathbf{1}_{\{|X_n| > \sqrt{n}\}} < \infty \quad \mathbb{P}\text{-a.s. and} \quad \sum_{n \geq 1} n^{-1/2} \mathbb{E}(|X_n| \mathbf{1}_{\{|X_n| > \sqrt{n}\}} | \mathcal{G}_{n+1}) < \infty \quad \mathbb{P}\text{-a.s.} \tag{4.7}$$

and by the Kronecker lemma,

$$\sum_{k=1}^n |X_k| \mathbf{1}_{\{|X_k| > \sqrt{k}\}} = o(\sqrt{n}) \quad \mathbb{P}\text{-a.s. and} \quad \sum_{k=1}^n \mathbb{E}(|X_k| \mathbf{1}_{\{|X_k| > \sqrt{k}\}} | \mathcal{G}_{k+1}) = o(\sqrt{n}) \quad \mathbb{P}\text{-a.s.} \tag{4.8}$$

Define $Y_n := X_n \mathbf{1}_{\{|X_n| \leq \sqrt{n}\}} - \mathbb{E}(X_n \mathbf{1}_{\{|X_n| \leq \sqrt{n}\}} | \mathcal{G}_{n+1})$. Then, by the above, using that $\mathbb{E}(X_n | \mathcal{G}_{n+1}) = 0$ $\mathbb{P}$-a.s., we see that it suffices to prove (2.7) with (X_n) replaced with (Y_n).

We want to apply Theorem 2.3 to (Y_n) with $a_n = \sigma_n^2 = n$. We have to prove conditions (2.4) and (2.5). Let us prove (2.4). Clearly, $(\mathbb{E}(Y_1^2) + \cdots + \mathbb{E}(Y_n^2))/n \to \mathbb{E}(X_1^2)$. Hence, we only need to prove that

$$(\mathbb{E}(Y_1^2 | \mathcal{G}_2) + \cdots + \mathbb{E}(Y_n^2 | \mathcal{G}_{n+1}))/n \to \mathbb{E}(X_1^2) \qquad \mathbb{P}\text{-a.s.} \tag{4.9}$$

We first prove that

$$(\mathbb{E}(Y_1^2 | \mathcal{G}_2) + \cdots + \mathbb{E}(Y_n^2 | \mathcal{G}_{n+1})) - (Y_1^2 + \cdots + Y_n^2) = o(n) \qquad \mathbb{P}\text{-a.s.} \tag{4.10}$$

By Kronecker lemma, this will follow from the convergence of the series $\sum_n (\mathbb{E}(Y_n^2 | \mathcal{G}_{n+1}) - Y_n^2)/n$. By Lemma 4.2, this last convergence will hold true provided that $\sum_n \mathbb{E}(Y_n^4)/n^2 < \infty$. But, by stationarity and Fubini theorem, we have

$$\sum_{n \geq 1} \frac{\mathbb{E}(Y_n^4)}{n^2} \leq 16\, \mathbb{E}\Big(X_1^4 \sum_{n \geq X_1^2} \frac{1}{n^2}\Big) \leq C\mathbb{E}(X_1^2) < \infty \,. \tag{4.11}$$

Therefore (4.10) is proved. Now, by the ergodic theorem we have

$$\limsup_n \frac{\sum_{k=1}^n X_k^2 \mathbf{1}_{\{|X_k|\le\sqrt{k}\}}}{n} \le \lim_n \frac{\sum_{k=1}^n X_k^2}{n} = \mathbb{E}(X_1^2) \qquad \mathbb{P}\text{-a.s.}\,,$$

and for any A fixed,

$$\liminf_n \frac{\sum_{k=1}^n X_k^2 \mathbf{1}_{\{|X_k|\le\sqrt{k}\}}}{n} \ge \lim_n \frac{\sum_{k=1}^n X_k^2 \mathbf{1}_{\{|X_k|\le A\}}}{n} = \mathbb{E}(X_1^2 \mathbf{1}_{\{|X_1|\le A\}}) \qquad \mathbb{P}\text{-a.s.}$$

Letting $A \to \infty$, we see that the $\liminf$ and the $\limsup$ above are equal to $\mathbb{E}(X_1^2)$. Hence

$$\frac{\sum_{k=1}^n X_k^2 \mathbf{1}_{\{|X_k|\le\sqrt{k}\}}}{n} \to \mathbb{E}(X_1^2) \qquad \mathbb{P}\text{-a.s.} \tag{4.12}$$

On another hand using the fact that $\mathbb{E}(X_k|\mathcal{G}_{k+1}) = 0$ a.s. together with (4.8), we get that

$$n^{-1}\sum_{k=1}^n (\mathbb{E}(X_k \mathbf{1}_{\{|X_k|\le\sqrt{k}\}}|\mathcal{G}_{k+1}))^2 \le n^{-1/2}\sum_{k=1}^n \mathbb{E}(|X_k|\mathbf{1}_{\{|X_k|>\sqrt{k}\}}|\mathcal{G}_{k+1}) = o(1) \qquad \mathbb{P}\text{-a.s.} \tag{4.13}$$

Combining (4.12), (4.13) and (4.10), we see that (4.9) holds, which proves (2.4).

The fact that (2.5) holds with $\nu = 2$ follows from (4.11). By Theorem 2.3, there exists a sequence of independent centered Gaussian variables $(\widetilde{Z}_n)_{n\ge1}$ such that $\mathbb{E}(\widetilde{Z}_n^2) = \mathbb{E}(Y_n^2) = \mathbb{E}(X_1^2) + o(1)$ and $Y_1 + \cdots Y_n - (\widetilde{Z}_1 + \cdots + \widetilde{Z}_n) = o(\sqrt{n \log\log n})$ $\mathbb{P}$-a.s. Let $(\delta_k)_{k\ge1}$ be a sequence of iid Gaussian random variables with mean zero and variance $\mathbb{E}(X_1^2)$, independent of the sequence $(\widetilde{Z}_n)_{n\ge1}$. We now construct a sequence $(Z_n)_{n\ge1}$ as follows. If $\mathbb{E}(\widetilde{Z}_n^2) = 0$, then $Z_n = \delta_n$, else $Z_n = c_n\widetilde{Z}_n$ where $c_n = \sqrt{\frac{\mathbb{E}(X_1^2)}{\mathbb{E}(\widetilde{Z}_n^2)}}$. By construction, the Z_n's are iid Gaussian random variables with mean zero and variance $\mathbb{E}(X_1^2)$. Write $G_n := Z_n - \widetilde{Z}_n$ and $v_n^2 := \sum_{k=1}^n \mathbb{E}(G_k^2)$. By Lévy's inequality (see for instance Proposition 2.3 in [20]),

$$\mathbb{P}\Big(\max_{1\le k\le 2^r}\Big|\sum_{i=1}^k G_i\Big| > x\Big) \le 2\exp\Big(-\frac{x^2}{2v_{2^r}^2}\Big)\,. \tag{4.14}$$

Hence taking $x = 2v_{2^r}(\log\log 2^r)^{1/2}$, we get that

$$\sum_{r\ge0}\mathbb{P}\Big(\max_{1\le k\le 2^r}\Big|\sum_{i=1}^k G_i\Big| > 2v_{2^r}(\log\log 2^r)^{1/2}\Big) < \infty\,.$$

Therefore $\sup_{1\le k\le 2^r}\big|\sum_{i=1}^k G_i\big| = O\big(v_{2^r}(\log\log 2^r)^{1/2}\big)$ almost surely. $\mathbb{P}$-a.s. This ends the proof of Corollary 2.5 since $v_n^2 = o(n)$. □

4.4 Proof of Corollary 2.7.

Define $Y_n := X_n \mathbf{1}_{\{|X_n|\le n^{1/p}\}} - \mathbb{E}(X_n \mathbf{1}_{\{|X_n|\le n^{1/p}\}}|\mathcal{G}_{n+1})$. Since $\mathbb{E}(X_n|\mathcal{G}_{n+1}) = 0$ a.s.,

$$\sum_{k\ge1}\frac{\mathbb{E}|X_k - Y_k|}{k^{1/p}} \le 2\sum_{k\ge1}\frac{\mathbb{E}(|X_k|\mathbf{1}_{\{|X_n|\le k^{1/p}\}})}{k^{1/p}}\,.$$

Hence by stationary and Fubini theorem, $\sum_{k\ge1} k^{-1/p}\mathbb{E}|X_k - Y_k| < \infty$, implying via the Kronecker lemma that

$$\sum_{k=1}^n |X_k - Y_k| = o(n^{1/p}) \qquad \mathbb{P}\text{-a.s.}$$

Let us prove now that $(Y_n)_{n\ge1}$ satisfies the conditions of Theorem 2.3 with $a_n = n^{2/p}b(n)$ and $\sigma_n^2 = n$. With this aim, we first notice that since $\mathbb{E}(X_n|\mathcal{G}_{n+1}) = 0$ a.s.,

$$\mathbb{E}(X_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2) - \mathbb{E}(Y_k^2|\mathcal{G}_{k+1}) + \mathbb{E}(Y_k^2) = \mathbb{E}(X_k^2\mathbf{1}_{\{|X_k|>k^{1/p}\}}|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2\mathbf{1}_{\{|X_k|>k^{1/p}\}})\,.$$

Since by stationarity and Fubini theorem, $\sum_{k\geq 1} k^{-2/p}\mathbb{E}(X_k^2 \mathbf{1}_{\{|X_k|>k^{1/p}\}}) < \infty$, we conclude via the Kronecker lemma that

$$\sum_{k=1}^{n} |\mathbb{E}(X_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(X_k^2) - \mathbb{E}(Y_k^2|\mathcal{G}_{k+1}) + \mathbb{E}(Y_k^2)| = o(n^{2/p}) \qquad \mathbb{P}\text{-a.s.}$$

Together with condition (2.8), this implies that

$$\sum_{k=1}^{n} (\mathbb{E}(Y_k^2|\mathcal{G}_{k+1}) - \mathbb{E}(Y_k^2)) = o(n^{2/p} b(n)) \qquad \mathbb{P}\text{-a.s.}$$

Notice now that by stationarity and Fubini theorem,

$$\sum_{n\geq 1} \frac{\mathbb{E}(Y_n^4)}{n^{4/p}} \leq 16\, \mathbb{E}\Big(X_1^4 \sum_{n\geq |X_1|^p} \frac{1}{n^{4/p}} \Big) \leq C_p \mathbb{E}(|X_1|^p) < \infty,$$

Therefore $(Y_n)_{n\geq 1}$ satisfies (2.5) with $\nu = 2$. Applying Theorem 2.3, we conclude that enlarging our probability space if necessary, there exists a sequence a sequence of independent centered Gaussian variables $(\widetilde{Z}_n)_{n\geq 1}$ such that $\mathbb{E}(\widetilde{Z}_n^2) = \mathbb{E}(Y_n^2)$ and $Y_1 + \cdots Y_n - (\widetilde{Z}_1 + \cdots + \widetilde{Z}_n) = o(n^{1/p}\sqrt{b(n)\log n})$ $\mathbb{P}$-a.s. We consider now the sequence of iid centered Gaussian variables $(Z_n)_{n\geq 1}$ with variance $\mathbb{E}(X_1^2)$ as defined in the proof of Corollary 2.5. Notice then that

$$\mathbb{E}(Z_k - \tilde{Z}_k)^2 = \big((\mathbb{E}(X_k^2))^{1/2} - (\mathbb{E}(Y_k^2))^{1/2} \big)^2 \leq \mathbb{E}((X_k - Y_k)^2) \leq \mathbb{E}(X_k^2 \mathbf{1}_{\{|X_k|>k^{1/p}\}}) ,$$

where for the last inequality, we have used the fact that $\mathbb{E}(X_n|\mathcal{G}_{n+1}) = 0$ a.s. Hence by stationarity

$$\sum_{n\geq 1} \mathbb{E}(Z_n - \tilde{Z}_n)^2/n^{2/p} \leq \mathbb{E}(X_1^2 \sum_{1\leq n\leq |X_1|^p} 1/n^{2/p}) \leq C\mathbb{E}(|X_1|^p) < \infty .$$

Therefore by the Kolmogorov theorem (or Lemma 4.2), $\sum_{n\geq 1}(Z_n - \tilde{Z}_n)/n^{2/p}$ converges $\mathbb{P}$-a.s.and by the Kronecker lemma $Z_1 + \ldots + Z_n - (\widetilde{Z}_1 + \ldots + \widetilde{Z}_n) = o(n^{1/p}\sqrt{\log\log n})$ $\mathbb{P}$-a.s. This achieves the proof of Corollary 2.7. ☐

4.5 Proof of Corollary 2.8.

The proof relies more deeply on the construction of Scott and Huggins [34]. We want to use Theorem 2.3 without condition (2.5). Now the proof of Theorem 2.3 relies on Proposition 2.1 and (2.5) is used to ensure that condition (2.2) holds for an auxiliary process. Instead of (2.5) we will make use of Remark 4.3. We define a reverse martingale $(R_n)_{n\geq 1}$, by $R_n = \sum_{k\geq n} X_k/k$. Notice that R_n is well defined in $\mathbb{L}^2$ by Lemma 4.2.

For every $n \leq -1$ define $\widetilde{R}_n := R_{-n}$, $\widetilde{X}_n := X_{-n}$ and $\widetilde{\mathcal{G}}_n := \mathcal{G}_{-n}$. Then $(\widetilde{R}_n, \widetilde{\mathcal{G}}_n)_{n\leq -1}$ is a martingale.

Enlarging our probability space if necessary, we may consider a countable set of standard Brownian motions $(B_t^{(n)})_{t\geq 0}$, $n \leq -1$ that are independent of each others and of $(\widetilde{X}_n)_{n\leq -1}$. Notice that the process $(\widetilde{X}_n, (B_t^{(n)})_{t\geq 0})_{n\leq -1}$ with values in $\mathbb{R}\times\mathbb{R}^{\mathbb{R}^+}$, is stationary.

We now define a filtration $(\widetilde{\mathcal{H}}_t)_{t\leq -1}$ as follows. For $n \leq -1$ an integer, write $\widetilde{\mathcal{H}}_n = \widetilde{\mathcal{G}}_n \vee \sigma\{B_t^{(j)}, 0 \leq t < \infty, -\infty < j \leq n\}$. For every $t \leq -1$, not an integer, write $\widetilde{\mathcal{H}}_t = \widetilde{\mathcal{H}}_{[t]} \vee \{R_{[t]+1} + B^{([t]+1)}_{\phi(s)}, 0 < s \leq t - [t]\}$, where $[t]$ stands for the largest negative integer, not exceeding t, and ϕ is defined on $]0, 1]$ by $\phi(s) := 1/s - 1$. Then, we define a continuous martingale with respect to $(\widetilde{\mathcal{H}}_t)_{t\leq -1}$ interpolating $(\widetilde{R}_n)$, by $\widetilde{R}_t = \mathbb{E}(\widetilde{R}_{[t]+1}|\widetilde{\mathcal{H}}_t)$, for every $t \leq -1$. Notice that

$$\widetilde{R}_t = \widetilde{R}_{[t]} + \frac{\mathbb{E}(X_{-[t]-1}|\widetilde{\mathcal{H}}_t)}{-[t]-1} . \tag{4.15}$$

Using Theorem A of [34] as done page 451 of [34], there exists a continuous non-decreasing process $(\widetilde{\tau}_t)_{t\leq -1}$ and a Brownian motion $(B_t^*)_{t\geq 0}$ such that $\widetilde{R}_t = B^*_{\widetilde{\tau}_t}$ a.s. and $(\widetilde{R}_t^2 - \widetilde{\tau}_t)_{t\leq -1}$ is a martingale with respect to $(\widetilde{\mathcal{H}}_t)_{t\leq -1}$. In particular, $(\widetilde{\tau}_t)_{t\leq -1}$ must be the quadratic variation of $(\widetilde{R}_t)$ on $]-\infty, t]$.

For every $n \geq 1$, define $\tau_n := \widetilde{\tau}_{-n}$, $\mathcal{H}_n := \widetilde{\mathcal{H}}_{-n}$. These are exactly the quantities involved in the proof of Proposition 2.1. Then $t_n = \tau_n - \tau_{n+1}$ is nothing else but the quadratic variation of $(\widetilde{R}_t)$ on $[-n-1, -n]$. But it follows from (4.15) that $(n^2 t_n)_{n\geq 1}$ is a stationary and ergodic process.

By Remark 4.3 we need to prove

$$\sum_{k\geq n} (t_k - \mathbb{E}(t_k|\mathcal{H}_{k+1})) = O\big(n^{-3/2}\sqrt{\log\log n}\big) \quad \mathbb{P}\text{-a.s.} \tag{4.16}$$

Since $(n^2 t_n)_{n\geq 1}$ is a stationary and ergodic sequence in $\mathbb{L}^2$, it follows from Corollary 2.5 that $\sum_{k=1}^n k^2(t_k - \mathbb{E}(t_k|\mathcal{H}_{k+1})) = O(\sqrt{n\log\log n})$ $\mathbb{P}$-a.s., which proves (4.16) by an Abel summation.

□

5 Proof of the results of Section 3.1

All along this section, we shall make use of the following definitions and preliminary considerations.

For every $n \geq 0$, define $\mathcal{G}_n := T^{-n}(\Sigma)$. Then $(\mathcal{G}_n)_{n\geq 0}$ is a non-increasing filtration. Denote also $\mathcal{G}_\infty := \cap_{n\geq 0}\mathcal{G}_n$.

Recall that the Perron-Frobenius operator K associated with T is defined in (3.1). Denote by U the operator of composition with T and $\mathbb{E}_\nu$ the expectation on X with respect to ν. It is not hard to see that $KU = Id$ and that for every $n \geq 0$ and every positive measurable f,

$$\mathbb{E}_\nu(f|\mathcal{G}_n) = (K^n f)\circ T^n = U^n K^n f\,.$$

Let $\mathbf{Y} = (Y_n)_{n\in\mathbb{Z}}$ be a stationary Markov chain defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with transition probability K and invariant distribution ν. Let $\mathcal{F}_k = \sigma(Y_i, i \leq k)$ for any $k \in \mathbb{Z}$.

Recall then (see for instance Hennion-Hervé [15, Lemma XI.3]) that for every $n \geq 1$, we have the following equalities in law (where in the left-hand side the law is meant under ν and in the right-hand side the law is meant under $\mathbb{P}$)

$$(T^n, \dots, T) \stackrel{d}{=} (Y_1, \dots, Y_n) \tag{5.1}$$

$$\max_{1\leq k\leq n} |\sum_{\ell=k}^{n} f\circ T^\ell| \stackrel{d}{=} \max_{1\leq k\leq n} |\sum_{\ell=1}^{k} f(Y_\ell)|\,, \tag{5.2}$$

$$\mathbb{E}_\nu(f|\mathcal{G}_n) = U^n K^n f \stackrel{d}{=} K^n f(Y_0) = \mathbb{E}(f(Y_n)|\mathcal{F}_0)\,. \tag{5.3}$$

Finally, for any positive integer n, let $S_n(f) = \sum_{\ell=0}^{n-1}(f\circ T^\ell - \nu(f))$.

We shall denote $\|\cdot\|_p$ the $\mathbb{L}^p$-norm on $(\Omega, \mathcal{A}, \mathbb{P})$.

5.1 Proof of Theorem 3.1

Assume that we have proved that the series σ^2 defined in (3.4) converges absolutely. To prove Theorem 3.1, we shall then prove that under the conditions (3.2) and (3.3), there exists a stationary sequence $(D_\ell)_{\ell\geq 0}$ in $\mathbb{L}^p(\nu)$, of reverse martingale differences with respect to the non-increasing filtration $(\mathcal{G}_\ell)_{\ell\geq 0}$, such that setting $M_n(f) = \sum_{\ell=0}^{n-1} D_\ell$, the following approximations hold:

$$\sup_{1\leq k\leq n} \big|S_k(f) - M_k(f)\big| = o(n^{1/p}(\log n)^{(1+t)/2}) \;\; \nu\text{-a.s.}, \tag{5.4}$$

and, enlarging the probability space (X, Σ, ν) if necessary, there exists a sequence $(Z_i)_{i\geq 0}$ of iid Gaussian random variables with mean zero and variance σ^2 such that

$$\sup_{1\leq k\leq n} \Big|\sum_{i=1}^{k}(D_i - Z_i)\Big| = o(n^{1/p}(\log n)^{(1+t)/2}) \;\; \nu\text{-a.s.} \tag{5.5}$$

The proof will be divided in several steps.

5.1.1 Step 1. Proof of the absolute convergence of the series σ^2

Using (5.1), we have that for any $k \geq 0$, $\nu((f-\nu(f))f \circ T^k) = \mathbb{E}(g(Y_0)g(Y_k))$. Therefore to prove that the series σ^2 defined in (3.4) converges absolutely, it suffices to prove that

$$\sum_{k\geq 0} |\mathbb{E}(g(Y_0)g(Y_k))| < \infty\,. \tag{5.6}$$

This latter convergence holds provided that $\sum_{n\geq 1} n^{-1/2}\|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_2 < \infty$ (see for instance Section 1 in [8]). Since $(Y_k)_{k\in\mathbb{Z}}$ is a Markov chain with transition kernel K, (5.6) then follows provided that

$$\sum_{n\geq 1} \frac{\|K^n g\|_{2,\nu}}{n^{1/2}} < \infty\,. \tag{5.7}$$

Noticing that $(\|K^n g\|_{2,\nu})_{n\geq 0}$ is a non-increasing sequence, the second part of condition (3.2) implies that $\|K^n g\|_{2,\nu} = o(n^{2/p-3/2}(\log n)^{t-1})$. Since $p>2$, this implies (5.7) and then (5.6).

5.1.2 Step 2. Construction of the approximating sequence of reverse martingale differences

The aim of this step is to prove that under condition (5.8) below, there exists a stationary sequence $(D_\ell)_{\ell\geq 0}$ of reverse martingale differences in $\mathbb{L}^p(\nu)$, with respect to the non-increasing filtration $(\mathcal{G}_\ell)_{\ell\geq 0}$, and such that we can explicitely control the error in $\mathbb{L}^p(\nu)$ between $S_n(f)$ and $M_n(f) = \sum_{\ell=0}^{n-1} D_\ell$. The estimate (5.9) can be viewed as a non-invertible version of Proposition 2.1 in [4].

Proposition 5.1 *Let $p \geq 2$. Let $f \in \mathbb{L}^p(X,\Sigma,\nu)$ and $g = f-\nu(f)$. Assume that*

$$\sum_{n\geq 1} \frac{\|K^n g\|_{p,\nu}}{n^{1/p}} < \infty\,. \tag{5.8}$$

Then, the series $D := \sum_{n\geq 0}(K^n g - UK^{n+1}g)$ converges in $\mathbb{L}^p(X,\mathcal{G}_0,\nu)$, and $\mathbb{E}_\nu(D|\mathcal{G}_1) = 0$ ν-a.s. Moreover, $\nu(D^2)=\sigma^2$ where σ^2 is defined by (3.4) and, setting

$$M_n(f) = \sum_{\ell=0}^{n-1} D_\ell \quad \textit{where} \quad D_\ell = D\circ T^\ell \quad \textit{for } \ell \geq 0,$$

we have for any positive integer n,

$$\|S_n(f) - M_n(f)\|_{p,\nu} \ll \sum_{1\leq k\leq n^{p/2}} \|K^k g\|_{p,\nu} + n^{1/2} \sum_{k\geq n^{p/2}} \frac{\|K^k g\|_{p,\nu}}{k^{1/p}}\,, \tag{5.9}$$

and

$$\|\max_{1\leq k\leq 2^n} |S_k(f) - M_k(f)|\|_{p,\nu} \ll \|S_{2^n}(f) - M_{2^n}(f)\|_{p,\nu} + 2^{n/p}\sum_{\ell=1}^{2^n} \frac{\|K^\ell g\|_{p,\nu}}{\ell^{1/p}}\,. \tag{5.10}$$

Notice that since $(\|K^n g\|_{p,\nu})_{n\geq 0}$ is a non-increasing sequence, the first part of condition (3.2) implies that $\|K^n g\|_{p,\nu} = o(n^{2/p^2-1}(\log n)^{(t-1)/2})$. Therefore as soon as $p>2$, the first part of condition (3.2) entails that (5.8) is satisfied.

Proof of Proposition 5.1. By (5.3), condition (5.8) can be rewritten as $\sum_{n\geq 1} n^{-1/p}\|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_p < \infty$. Then it follows from Lemma 5.1 of [4] that $\sum_{n\geq 0}(K^n g(Y_1) - K^{n+1}g(Y_0))$ converges in $\mathbb{L}^p(\Omega,\mathcal{A},\mathbb{P})$. Hence the function

$$\varphi(x,y) := \sum_{n\geq 0}(K^n g(x) - K^{n+1}g(y)) \tag{5.11}$$

is well defined in $\mathbb{L}^p(X\times X, \Sigma\otimes\Sigma, \mathbb{P}\circ(Y_1,Y_0)^{-1})$. Whence, by (5.1), $\sum_{n\geq 0}(K^n g - UK^{n+1}g)$ converges in $\mathbb{L}^p(\nu)$ to $\varphi(Id,T)$. Now, by (5.2),

$$\|S_n(f) - M_n(f)\|_{p,\nu} = \Big\| \sum_{k=1}^{n}(g(Y_k) - \varphi(Y_k,Y_{k-1}))\Big\|_p\,.$$

But, by Proposition 2.1 of [4] (with $N = n^{p/2}$) together with their Lemma 5.1,

$$\Big\| \sum_{k=1}^{n} (g(Y_k) - \varphi(Y_k, Y_{k-1})) \Big\|_p \ll \sum_{1 \le k \le n^{p/2}} \|\mathbb{E}(g(Y_k)|\mathcal{F}_0)\|_p + n^{1/2} \sum_{k \ge n^{p/2}} \frac{\|\mathbb{E}(g(Y_k)|\mathcal{F}_0)\|_p}{k^{1/p}} ,$$

and the estimate (5.9) follows from (5.3). Using now (5.9) with $p = 2$, it follows that $\nu(D^2) = \lim_{n\to\infty} \mathbb{E}(S_n^2(f))/n$. Since, the series σ^2 defined in (3.4) is absolutely convergent, the latter limit is σ^2. This shows that $\nu(D^2) = \sigma^2$. To end the proof of the proposition it remains to prove (5.10). Using (5.2), we have that

$$\| \max_{1 \le k \le 2^n} |S_k(f) - M_k(f)| \|_{p,\nu} \le 2 \Big\| \max_{1 \le k \le 2^n} \Big| \sum_{\ell=1}^{k} (g(Y_\ell) - \varphi(Y_\ell, Y_{\ell-1})) \Big| \Big\|_p .$$

The estimate (5.10) then follows by an application of Corollary 3 in [28] by taking into account (5.3). □

5.1.3 Step 3. Proof of the almost convergence (5.4)

We prove here that for the sequence of reverse martingale differences defined in Proposition 5.1, the convergence (5.4) is satisfied. Keeping in mind that the first part of condition (3.2) implies (5.8), the convergence (5.4) follows from the above proposition applied with $\psi(n) = (\log n)^{(t+1)/2}$. Note that the forthcoming proposition can be viewed as a non-invertible version of Proposition 4.1 in [4].

Proposition 5.2 *Let $p \ge 2$ and $f \in \mathbb{L}^p(X, \Sigma, \nu)$ satisfying* (5.8)*. Then, for every positive and non-decreasing function ψ such that $\psi(2n) \le C\psi(n)$ and $\sum_{n\ge1} 1/\psi(2^n)^p < \infty$,*

$$S_n(f) - M_n(f) = o(n^{1/p}\psi(n)) \qquad \nu\text{-a.s.}$$

Proof. We make use of the function φ defined in (5.11). Let us apply inequality (8) of Proposition 5 of [28] with $x = \lambda 2^{r/p}\psi(2^r)$, where $\lambda > 0$. Using stationarity and that $\mathbb{E}(\varphi(Y_k, Y_{k-1})|\mathcal{F}_0) = 0$ for any $k \ge 1$, we obtain

$$\mathbb{P}\Big(\max_{1 \le i \le 2^n} \Big| \sum_{k=1}^{i} (g(Y_k) - \varphi(Y_k, Y_{k-1})) \Big| > \lambda 2^{n/p}\psi(2^n) \Big)$$
$$\ll \frac{\| \sum_{k=1}^{2^n} (g(Y_k) - \varphi(Y_k, Y_{k-1})) \|_1}{\lambda 2^{n/p}\psi(2^n)} + \frac{1}{\lambda^p \psi(2^n)^p} \Big(\sum_{\ell \ge 0} 2^{-\ell/p} \sum_{m=0}^{2^\ell} \|\mathbb{E}(g(Y_m)|\mathcal{F}_0)\|_p \Big)^p .$$

Using (5.2), we derive that

$$\sum_{n\ge0} \nu\big(\max_{1 \le i \le 2^n} |S_i(f) - M_i(f)| > \lambda 2^{n/p}\psi(2^n) \big)$$
$$\ll \sum_{n\ge0} \frac{\|S_{2^n}(f) - M_{2^n}(f)\|_{1,\nu}}{\lambda 2^{n/p}\psi(2^n)} + \Big(\sum_{n\ge0} \frac{1}{\lambda^p \psi(2^n)^p} \Big) \Big(\sum_{k\ge1} \frac{\|K^k g\|_{p,\nu}}{k^{1/p}} \Big)^p .$$

Since λ is arbitrary, by our assumptions on ψ, the corollary will follow if we can prove that

$$\sum_{n\ge0} \frac{\|S_{2^n}(g) - M_{2^n}(g)\|_{1,\nu}}{2^{n/p}} < \infty .$$

The convergence of that series follows from Proposition 5.1, using (5.9) with $p = 2$ and the fact that $\|S_{2^n}(f) - M_{2^n}(f)\|_{1,\nu} \le \|S_{2^n}(f) - M_{2^n}(f)\|_{2,\nu}$. □

5.1.4 Step 4. Proof of the almost convergence (5.5)

We prove here that the sequence of reverse martingale differences defined in Proposition 5.1 satisfies the almost sure invariance principle (5.5). Recall that the first part of condition (3.2) implying (5.8), the sequence of reverse martingale differences defined in Proposition 5.1 is well defined in $\mathbb{L}^p(X,\mathcal{G}_0,\nu)$ and satisfies $\nu(D^2)=\sigma^2$. Therefore, in view of Corollary 2.7, it suffices to prove that

$$\sum_{k=1}^n (\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})-\nu(D^2)) = o(n^{2/p}(\log n)^t) \qquad \nu\text{-a.s.} \tag{5.12}$$

But (5.12) will hold if

$$\sum_{n\geq 1}\frac{\|\max_{1\leq \ell\leq 2^n}|\sum_{k=1}^{\ell}(\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})-\nu(D^2))|\|_{p/2,\nu}^{p/2}}{2^n n^{tp/2}}<\infty\,. \tag{5.13}$$

To prove the latter convergence, we first write that

$$\Big\|\max_{1\leq \ell\leq 2^n}\Big|\sum_{k=1}^{\ell}(\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})-\nu(D^2))\Big|\Big\|_{p/2,\nu}\leq 2\Big\|\max_{1\leq \ell\leq 2^n}\Big|\sum_{k=\ell}^{2^n}(\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})-\nu(D^2))\Big|\Big\|_{p/2,\nu}\,.$$

By standard arguments, we have, with $h:=\int_X\varphi^2(y,\cdot)K(\cdot,dy)$,

$$\begin{aligned}&\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})=\mathbb{E}_\nu(D^2\circ T^k|T^{k+1})=\mathbb{E}_\nu(D^2|T)\circ T^k=h(T^k) \qquad \nu\text{-a.s.}\\ \text{and}\qquad &\mathbb{E}(\varphi^2(Y_k,Y_{k-1})|\mathcal{F}_{k-1})=\mathbb{E}(h(Y_{k-1}|\mathcal{F}_{k-1})=h(Y_{k-1}) \qquad \mathbb{P}\text{-a.s.}\end{aligned}$$

Therefore, by taking into account (5.1), it follows that

$$\begin{aligned}&\Big\|\max_{1\leq \ell\leq 2^n}\Big|\sum_{k=1}^{\ell}(\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})-\nu(D^2))\Big|\Big\|_{p/2,\nu}\\ &\leq 2\Big\|\max_{1\leq \ell\leq 2^n}\Big|\sum_{k=1}^{\ell}(\mathbb{E}(\varphi^2(Y_k,Y_{k-1})|\mathcal{F}_{k-1})-\mathbb{E}(\varphi^2(Y_k,Y_{k-1})))\Big|\Big\|_{p/2}\,.\end{aligned}$$

Hence, by using Theorem 3 of [39] and using the fact that $\mathbb{E}(\varphi(Y_k,Y_{k-1})|\mathcal{F}_{k-1})=0$ $\mathbb{P}$-a.s., it follows that there exists a constant C_p depending only on p such that

$$\begin{aligned}&\Big\|\max_{1\leq \ell\leq 2^n}\Big|\sum_{k=1}^{\ell}(\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1})-\nu(D^2))\Big|\Big\|_{p/2,\nu}\\ &\qquad\leq C_p\Big(2^{2n/p}\|\varphi(Y_1,Y_0)\|_p^2+2^{2n/p}\sum_{r=0}^{n}2^{-2r/p}\|\mathbb{E}(V_{2^r}^2|\mathcal{F}_0)-\mathbb{E}(V_{2^r}^2)\|_{p/2}\Big)\,,\end{aligned}$$

where we used the notation

$$V_n:=\sum_{k=1}^n\varphi(Y_k,Y_{k-1})\,. \tag{5.14}$$

So, overall, since $t>2/p$, (5.13) (and then (5.12)) will hold if we can prove that

$$\sum_{n\geq 1}n^{-tp/2}\Big(\sum_{r=0}^{n}2^{-2r/p}\|\mathbb{E}(V_{2^r}^2|\mathcal{F}_0)-\mathbb{E}(V_{2^r}^2)\|_{p/2}\Big)^{p/2}<\infty\,.$$

Using Holder's inequality, we infer that this last series is convergent as soon as

$$\sum_{r\geq 1}\frac{2^{-r}}{r^{(t-1)p/2}}\|\mathbb{E}(V_{2^r}^2|\mathcal{F}_0)-\mathbb{E}(V_{2^r}^2)\|_{p/2}^{p/2}<\infty\,. \tag{5.15}$$

Next, starting from (5.15) and using the subadditivity of the sequence $\big(\|\mathbb{E}(V_n^2|\mathcal{F}_0)-\mathbb{E}(V_n^2)\|_{p/2}\big)_{n\geq 1}$, we infer that (5.15) will hold (and then (5.12) also) if we can prove that under the conditions (3.2) and (3.3),

$$\sum_{n\geq 2}\frac{1}{n^2(\log n)^{(t-1)p/2}}\|\mathbb{E}(V_n^2|\mathcal{F}_0)-\mathbb{E}(V_n^2)\|_{p/2}^{p/2}<\infty\,. \tag{5.16}$$

But, recalling that by definition,

$$\varphi(Y_k, Y_{k-1}) := \sum_{n\geq 0}(K^n g(Y_k) - K^{n+1}g(Y_{k-1})) = \sum_{n\geq k}\big(\mathbb{E}(g(Y_n)|\mathcal{F}_k) - \mathbb{E}(g(Y_n)|\mathcal{F}_{k-1})\big)\,,$$

we notice that (5.16) is exactly the second part of condition (2.5) in [4]. Therefore by the proof of their Theorem 2.3 and their Proposition 2.2, we infer that (5.16) holds provided that

$$\sum_{n\geq 2}\frac{n^{p-1}}{n^{2/p}(\log n)^{\frac{(t-1)p}{2}}}\|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_p^p < \infty \text{ and } \sum_{n\geq 2}\frac{n^{3p/4}}{n^2(\log n)^{\frac{(t-1)p}{2}}}\|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_2^{p/2} < \infty\,, \tag{5.17}$$

and

$$\sum_{n\geq 2}\frac{1}{n^2(\log n)^{(t-1)p/2}}\|\mathbb{E}(\widetilde{S}_n^2(f)|\mathcal{F}_0) - \mathbb{E}(\widetilde{S}_n^2(f))\|_{p/2}^{p/2} < \infty\,, \tag{5.18}$$

where

$$\widetilde{S}_n(f) = \sum_{k=1}^{n}(f(Y_k) - \nu(f)) := \sum_{k=1}^{n} g(Y_k)\,.$$

Since $K^n g(Y_0) = \mathbb{E}(g(Y_n)|\mathcal{F}_0)$ for any non-negative integer n, the conditions (5.17) can be rewritten as (3.2), whereas condition (5.18) is implied by (3.3) by noticing that

$$\|\mathbb{E}(\widetilde{S}_n^2(f)|\mathcal{F}_0) - \mathbb{E}(\widetilde{S}_n^2(f))\|_{p/2} \leq 2\sum_{i=1}^{n}\sum_{j=0}^{n-i}\|\mathbb{E}(g(Y_i)g(Y_{i+j})|\mathcal{F}_0) - \mathbb{E}(g(Y_i)g(Y_{i+j}))\|_{\frac{p}{2}}\,. \tag{5.19}$$

and that, for any non-negative integers i and j,

$$\|\mathbb{E}(g(Y_i)g(Y_{i+j})|\mathcal{F}_0) - \mathbb{E}(g(Y_i)g(Y_{i+j}))\|_{\frac{p}{2}} = \|K^i(gK^j(g)) - \nu(gK^j(g))\|_{\frac{p}{2},\nu}\,. \tag{5.20}$$

This ends the proof of (5.16) and then of the theorem. □

5.2 Proof of Theorem 3.2

Notice that since $(\|K^n g\|_4)_{n\geq 0}$ is a non-increasing sequence, the first part of condition (3.6) implies (5.8) with $p = 4$. Therefore the sequence $(D_\ell)_{\ell\geq 0}$ of reverse martingale differences defined in Proposition 5.1 is well defined in $\mathbb{L}^4(X, \Sigma, \nu)$ and σ^2 is absolutely convergent. Setting $M_n(f) = \sum_{\ell=0}^{n-1} D_\ell$, the theorem will then follow if we can prove that

$$\sup_{1\leq k\leq n}\big|S_k(f) - M_k(f)\big| = O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})\ \ \nu\text{-a.s.}, \tag{5.21}$$

and, enlarging the probability space (X, Σ, ν) if necessary, there exists a sequence $(Z_i)_{i\geq 0}$ of iid Gaussian random variables with mean zero and variance σ^2 such that

$$\sup_{1\leq k\leq n}\Big|\sum_{i=1}^{k}(D_i - Z_i)\Big| = O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})\ \ \nu\text{-a.s.} \tag{5.22}$$

To prove (5.21), it suffices to apply Proposition 5.2 by recalling that the first part of condition (3.6) implies (5.8) with $p = 4$.

We prove now that (5.22) holds. According to Corollary 2.7, the almost sure approximation (5.22) will follow if we can prove that

$$\sum_{k=0}^{n-1}(\mathbb{E}_\nu(D^2\circ T^k|\mathcal{G}_{k+1}) - \nu(D^2)) = O(n^{1/2}(\log\log n)^{1/2}) \qquad \nu\text{-a.s.} \tag{5.23}$$

Setting $h := \int_X \varphi^2(y,\cdot)K(\cdot, dy)$, recall that $\mathbb{E}(D^2\circ T^k|\mathcal{G}_{k+1}) = \mathbb{E}(D^2\circ T^k|T^{k+1}) = h\circ T^k$ ν-a.s. Notice that that $\nu(h^2) < \infty$. Let $\widetilde{h} = h - \nu(h)$. Assume that we can prove that

$$\sum_{n\geq 1}(\log n)^3\frac{\|\widetilde{h} + K\widetilde{h} + \cdots + K^{n-1}\widetilde{h}\|_{2,\nu}^2}{n^2} < \infty\,. \tag{5.24}$$

This condition implies in particular that $\sum_{n\geq 1} n^{-3/2}\|\widetilde{h}+K\widetilde{h}+\cdots+K^{n-1}\widetilde{h}\|_{2,\nu}<\infty$. By Lemma 2 of [27] (and its proof), it follows that $m(Y_1,Y_0):=\lim_n n^{-1}\sum_{j=1}^{n}\sum_{k=0}^{j-1}(K^k\widetilde{h}(Y_1)-K^{k+1}\widetilde{h}(Y_0))$ exists in $\mathbb{L}^2$ and $V_n:=\sum_{k=1}^{n} m(Y_1,Y_0)$ is a stationary martingale with respect to $(\mathcal{F}_k)$ such that

$$\|\widetilde{h}(Y_1)+\cdots+\widetilde{h}(Y_n)-V_n\|_2 \ll \sqrt{n}\sum_{k\geq n}\frac{\|\widetilde{h}+K\widetilde{h}+\cdots+K^{k-1}\widetilde{h}\|_{2,\nu}}{k^{3/2}}\,. \tag{5.25}$$

Define now $V_n^*:=\sum_{k=0}^{n-1} m(T^k,T^{k+1})$. Then V_n^* is a sum associated with a stationary sequence of reverse martingale differences in $\mathbb{L}^2_\nu$, with respect to $(\mathcal{G}_k)_{k\geq 0}$. By Corollary 4.2 of [3] with $b(n)=\log n$, if

$$\sum_n \frac{\log n}{n^2}\|\widetilde{h}+\cdots+\widetilde{h}\circ T^{n-1}-V_n^*\|_{2,\nu}^2<\infty\,, \tag{5.26}$$

then

$$\frac{\widetilde{h}+\cdots+\widetilde{h}\circ T^{n-1}-V_n^*}{\sqrt{n\log\log n}}\to 0 \qquad \nu\text{-a.s.}$$

But, by using (5.1) and (5.25), we have that

$$\|\widetilde{h}+\cdots+\widetilde{h}\circ T^{n-1}-V_n^*\|_{2,\nu} \ll \sqrt{n}\sum_{k\geq n}\frac{\|\widetilde{h}+K\widetilde{h}+\cdots+K^{k-1}\widetilde{h}\|_{2,\nu}}{k^{3/2}}\,. \tag{5.27}$$

By taking into account (5.27), it is easy to see that (5.26) holds as soon as (5.24) is satisfied. Using then Corollary 2.5 to observe that $V_n^*=O(\sqrt{n\log\log n})$ ν-a.s., we conclude that (5.23) (and then (5.22)) holds as soon as (5.24) does. Notice now that (5.24) can be rewritten as

$$\sum_{n\geq 1}(\log n)^3\frac{\|\mathbb{E}(V_n^2|\mathcal{F}_0)-\mathbb{E}(V_n^2)\|_2^2}{n^2}<\infty\,. \tag{5.28}$$

According to the proofs of Theorem 2.3 and Proposition 2.2 in [4], this will hold true provided that

$$\sum_{n\geq 2} n^{5/2}(\log n)^3\|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_4^4<\infty \text{ and } \sum_{n\geq 2} n(\log n)^3\|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_2^2<\infty\,, \tag{5.29}$$

and

$$\sum_{n\geq 2}\frac{(\log n)^3}{n^2}\|\mathbb{E}(\widetilde{S}_n^2(f)|\mathcal{F}_0)-\mathbb{E}(\widetilde{S}_n^2(f))\|_2^2<\infty\,, \tag{5.30}$$

where $\widetilde{S}_n(f)=\sum_{k=1}^{n}(f(Y_k)-\nu(Y_k))$. To end the proof, it suffices to notice that, since $K^n g(Y_0)=\mathbb{E}(g(Y_n)|\mathcal{F}_0)$ for any non negative integer n, the conditions (5.29) can be rewritten as (3.6), whereas condition (5.30) is implied by (3.7) by using (5.19) and (5.20). This ends the proof of (5.28) and then of the theorem. □

6 Proof of the results of Section 3.2

All along this section, K will be the Perron-Frobenius operator associated with the dynamical system T under consideration. Recall that it is defined by (3.1).

6.1 Proof of Corollary 3.4

It suffices to check that the conditions of Theorem 3.1 are satisfied. With this aim, it is convenient to consider $(Y_n)_{n\in\mathbb{Z}}$ a stationary Markov chain with transition kernel K and invariant distribution ν. Denoting $\mathcal{F}_k=\sigma(Y_j, j\leq k)$, we recall that the conditions (3.2) can be rewritten as (5.17), whereas according to (5.20), (3.3) can be rewritten as

$$\sum_{n\geq 2}\frac{1}{n^2(\log n)^{(t-1)p/2}}\Big(\sum_{i=1}^{n}\sum_{j=0}^{n-i}\|\mathbb{E}(g(Y_i)g(Y_{i+j})|\mathcal{F}_0)-\mathbb{E}(g(Y_i)g(Y_{i+j}))\|_{\frac{p}{2}}\Big)^{p/2}<\infty\,. \tag{6.1}$$

To end the proof it suffices to notice that, according to the proof of Corollary 3.8 in [4], the conditions (5.17) and (6.1) are satisfied with $t=1$ under (3.10). □

6.2 Proof of Corollary 3.7

The result will follow if we can prove that the conditions (3.6) and (3.7) of Theorem 3.2 are satisfied. With this aim, we start with a lemma whose proof is given later and that is a key tool also in the proof of Theorem 3.8.

Lemma 6.1 *Let T be a uniformly expanding map as defined in Definition 3.5, with absolutely continuous invariant measure ν. Let K be its associated Perron-Frobenius operator defined by (3.1). Let f and g be two functions from $\mathbb{R}$ to $\mathbb{R}$ which are monotonic on some interval and null elsewhere. Let $p \in [1, \infty]$. If $\|f\|_{p,\nu} < \infty$, then, for any positive integer, there exists $\rho \in (0, 1)$ such that for any positive integer n,*

$$\|K^n f - \nu(f)\|_{p,\nu} \ll \rho^{n(p-1)/p} \|f\|_{p,\nu} .$$

If moreover $p \geq 2$ and $\|g\|_{p,\nu} < \infty$, then setting $f^{(0)} = f - \nu(f)$ and $g^{(0)} = g - \nu(g)$, we have for any positive integers $j \geq i \geq n$,

$$\|K^i(f^{(0)} K^{j-i}(g^{(0)})) - \nu(f^{(0)} K^{j-i}(g^{(0)}))\|_{\frac{p}{2},\nu} \ll \rho^{n(p-2)/p} \|f\|_{p,\nu} \|g\|_{p,\nu} .$$

To make use of this lemma to prove that (3.6) and (3.7) are satisfied when $f \in \mathrm{Mon}_4^c(M, \nu)$, we continue with some preliminary observations. Since $f \in \mathrm{Mon}_4^c(M, \nu)$, by definition, there exists a sequence of functions

$$f_L = \sum_{k=1}^{L} a_{k,L} g_{k,L} , \tag{6.2}$$

with $g_{k,L}$ belonging to $\mathrm{Mon}_4(M, \nu)$ and $\sum_{k=1}^{L} |a_{k,L}| \leq 1$, such that f_L converges in $\mathbb{L}^4(\nu)$ to f.

Therefore

$$\|K^n f - \nu(f)\|_{4,\nu} \leq \liminf_{L\to\infty} \sum_{k=1}^{L} |a_{k,L}| \, \|K^n g_{k,L} - \nu(g_{k,L})\|_{4,\nu} .$$

Next, by Lemma 6.1, $\|K^n g_{k,L} - \nu(g_{k,L})\|_{4,\nu} \ll M\rho^{3n/4}$. So overall, since $\sum_{k=1}^{L} |a_{k,L}| \leq 1$, we get that

$$\|K^n f - \nu(f)\|_{4,\nu} \ll M\rho^{3n/4} , \tag{6.3}$$

which clearly shows that (3.6) is satisfied.

On the other hand, for any non-negative integers i and j,

$$\|K^i(f^{(0)} K^j(f^{(0)})) - \nu(f^{(0)} K^j(f^{(0)}))\|_{2,\nu} \\ \leq \liminf_{L\to\infty} \sum_{k=1}^{L} \sum_{\ell=1}^{L} |a_{k,L}||a_{\ell,L}| \, \|K^i(g^{(0)}_{k,L} K^j(g^{(0)}_{\ell,L})) - \nu(g^{(0)}_{k,L} K^j(g^{(0)}_{\ell,L}))\|_{2,\nu} . \tag{6.4}$$

Applying Lemma 6.1 and using the fact that $\sum_{k=1}^{L} |a_{k,L}| \leq 1$, it follows that, for any non-negative integers i and j,

$$\|K^i(f^{(0)} K^j(f^{(0)})) - \nu(f^{(0)} K^j(f^{(0)}))\|_{2,\nu} \ll M^2 \rho^{i/2} . \tag{6.5}$$

On another hand, starting from (6.4) and applying Lemma 6.1, we also have that for any non-negative integers i and j,

$$\|K^i(f^{(0)} K^j(f^{(0)})) - \nu(f^{(0)} K^j(f^{(0)}))\|_{2,\nu} \leq 2\|f - \nu(f)\|_{4,\nu} \|K^j f - \nu(f)\|_{4,\nu} \ll M^2 \rho^{3j/4} . \tag{6.6}$$

Hence, by using (6.5) and (6.6), we get that, for any non-negative integers i and j,

$$\|K^i(f^{(0)} K^j(f^{(0)})) - \nu(f^{(0)} K^j(f^{(0)}))\|_{2,\nu} \ll \rho^{(i \vee j)/2} . \tag{6.7}$$

This last estimate implies that

$$\sum_{i=1}^{n} \sum_{j=0}^{n-i} \|K^i(f^{(0)} K^j(f^{(0)})) - \nu(f^{(0)} K^j(f^{(0)}))\|_{2,\nu} = 0(1) ,$$

which clearly proves (3.7). To end the proof of the corollary, it remains to prove Lemma 6.1.

Proof of Lemma 6.1. This lemma comes from Lemma 5.2 in [6]. Indeed if $\mathbf{Y} = (Y_n)_{n\in\mathbb{Z}}$ is a stationary Markov chain with transition kernel K and invariant distribution ν, and if $\mathcal{F}_k = \sigma(Y_j, j \le k)$, then we have

$$\|K^n f - \nu(f)\|_{p,\nu} = \|\mathbb{E}(f(Y_k)|\mathcal{F}_0) - \mathbb{E}(f(Y_k))\|_p .$$

and for any $j \ge i$,

$$\|K^i(f^{(0)}K^{j-i}(g^{(0)})) - \nu(f^{(0)}K^{j-i}(g^{(0)}))\|_{\frac{p}{2},\nu} = \|\mathbb{E}(f(Y_i)^{(0)}g(Y_j)^{(0)}|\mathcal{F}_0) - \mathbb{E}(f(Y_i)^{(0)}g(Y_j)^{(0)})\|_{p/2} .$$

The lemma follows directly by using the estimates given in Lemma 5.2 of [6] by taking into account that since K is the Perron-Frobenius operator associated with (T,ν) defined in Definition 3.5, their $\phi_{1,\mathbf{Y}}(k)$ and $\phi_{2,\mathbf{Y}}(k)$ coefficients can be bounded as follows: there exist a positive constant C and a real $\rho \in (0,1)$, such that for any non-negative integer k,

$$\phi_{1,\mathbf{Y}}(k)) \le \phi_{2,\mathbf{Y}}(k)) \le C\rho^k$$

(see Section 6.3 in [10]). □

6.3 Proof of Theorem 3.8

Let $p \in]2,4[$. The strategy of proof is the same as the one developed in the proof of Theorem 3.1 and we shall use the same definitions and notations as those introduced at the beginning of Section 5. So, we shall define a sequence of reverse martingale differences $(D_\ell)_{\ell\ge 0}$ with respect to the non-increasing filtration $(\mathcal{G}_\ell)_{\ell\ge 0}$ that is in $\mathbb{L}^p(\nu)$, such that, setting $M_n(f) = \sum_{\ell=0}^{n-1} D_\ell$,

$$\sup_{1\le k\le n} \big|S_k(f) - M_k(f)\big| = o(n^{1/p}(\log n)^{1-2/p}) \ \nu\text{-a.s.}\,, \tag{6.8}$$

and satisfying the following strong invariance principle: enlarging our probability space if necessary, there exists a sequence $(Z_i)_{i\ge 0}$ of iid Gaussian random variables with mean zero and variance σ^2 such that

$$\sup_{1\le k\le n} \Big|\sum_{\ell=0}^{k-1}(D_\ell - Z_\ell)\Big| = o(n^{1/p}(\log n)^{1-2/p}) \ \nu\text{-a.s.} \tag{6.9}$$

As we shall see a suitable sequence of reverse martingale differences that satisfies the above convergence conditions, can be defined with the help of truncation arguments leading to a non-stationary sequence.

6.3.1 Step 1. Truncation arguments and preliminary facts

Since $f \in \mathrm{Mon}_p^c(M,\nu)$, by definition, there exists a sequence of functions

$$f_L = \sum_{k=1}^{L} a_{k,L} g_{k,L}\,, \tag{6.10}$$

with $g_{k,L}$ belonging to $\mathrm{Mon}_p(M,\nu)$ and $\sum_{k=1}^{L}|a_{k,L}| \le 1$, such that f_L converges in $\mathbb{L}^p(\nu)$ to f.

Let j be a non-negative integer and

$$c(0) = 1 \ \text{ and } c(j) = 2^{j/p} j^{-2/p}\,, \ \text{ if } j > 0\,.$$

We define in what follows some functions $(f_{L,j})_{L\ge 1}$ by truncating the functions $g_{k,L}$ defined in (6.10) at the level $c(j)$. So, setting $t_j(x) = x\mathbf{1}_{|x|\le c(j)}$, we then define

$$f_{L,j} = \sum_{k=1}^{L} a_{k,L}\, t_j \circ g_{k,L}\,. \tag{6.11}$$

By assumption and by construction, for any non-negative integers j and L, the function $f_{L,j}$ has a variation bounded by $4c(j)$. Hence, by Lemma 2.1 of [6], $(f_{L,j})_L$ admits a subsequence, say $(f_{\varphi(L),j})_L$ converging in $\mathbb{L}^1(\nu)$, hence in $\mathbb{L}^p(\nu)$, say to $\bar f_j$. Moreover the function $f - \bar f_j$ is the limit in $\mathbb{L}^p(\nu)$ of

$$f_{\varphi(L)} - f_{\varphi(L),j} = \sum_{k=1}^{\varphi(L)} a_{k,\varphi(L)} \widetilde{t}_j \circ g_{k,\varphi(L)} \ \text{where } \widetilde{t}_j = x\mathbf{1}_{|x|>c(j)}\,. \tag{6.12}$$

6.3.2 Step 2. Construction of the approximating sequence of reverse martingale differences

For any non-negative integer j, we consider the function $\bar{f}_j$ defined in Step 1 above. Let us prove that the series

$$m_j := \sum_{n\geq 0}(K^n \bar{f}_j - UK^{n+1}\bar{f}_j) \tag{6.13}$$

converges in $\mathbb{L}^q(\nu)$ for any $q \in [1,\infty]$. According to Proposition 5.1, this holds if

$$\sum_{n\geq 1} \frac{1}{n^{1/q}} \|K^n \bar{f}_j - \nu(\bar{f}_j)\|_{q,\nu} < \infty\,. \tag{6.14}$$

But,

$$\|K^n \bar{f}_j - \nu(\bar{f}_j)\|_{q,\nu} \leq \liminf_{L\to\infty} \sum_{k=1}^{\varphi(L)} |a_{k,\varphi(L)}| \|K^n t_j \circ g_{k,\varphi(L)} - \nu(t_j \circ g_{k,\varphi(L)})\|_{q,\nu}\,.$$

Hence, by using Lemma 6.1 and the fact that $\sum_{k=1}^{\varphi(L)} |a_{k,\varphi(L)}| \leq 1$, we get that

$$\|K^n \bar{f}_j - \nu(\bar{f}_j)\|_{q,\nu} \ll \rho^{n(q-1)/q} c(j)\,,$$

which clearly proves (6.14). Notice also that m_j is $\mathcal{G}_0$-measurable and is such that $\mathbb{E}_\nu(m_j|\mathcal{G}_1) = 0$ ν-a.s.

We define now the sequence of reverse martingale differences $(D_\ell)_{\ell\geq 0}$ for which the convergences (6.8) and (6.9) will hold. Let $D_0 := m_0$ and, for any non-negative integer j and any $\ell \in \{2^j, ..., 2^{j+1}-1\}$, let

$$D_\ell := d_{j,\ell} \quad \text{where} \quad d_{j,\ell} = m_j \circ T^\ell\,. \tag{6.15}$$

By the previous considerations, $(D_\ell)_{\ell\geq 0}$ forms a sequence of reverse martingale differences in $\mathbb{L}^\infty(\nu)$, with respect to the non-increasing filtration $(\mathcal{G}_\ell)_{\ell\geq 0}$. Note also that although the sequence $(D_\ell)_{\ell\geq 0}$ is by construction non-stationary, it becomes stationary on each dyadic interval $\{2^j, ..., 2^{j+1}-1\}$.

6.3.3 Step 3. Proof of the almost sure approximation (6.8)

We consider the sequence $(D_\ell)_{\ell\geq 0}$ defined in the previous step and we set $M_n(f) = \sum_{\ell=0}^{n-1} D_\ell$. To prove that the almost sure approximation (6.8) holds, we shall then prove that

$$\sup_{1\leq k\leq 2^N} \big|S_k(f) - M_k(f)\big| = o(2^{N/p} N^{1-2/p}) \ \nu\text{-a.s.} \tag{6.16}$$

Let $N \in \mathbb{N}^*$ and let $k \in]1, 2^N]$. We first notice that if L is the integer such that $2^{L-1} < k \leq 2^L$, then

$$|S_k(f) - M_k(f)| \leq |f - \nu(f) - D_0| + \sum_{j=0}^{L-1} \sup_{0\leq r\leq 2^j-1} \Big|\sum_{\ell=2^j}^{r+2^j} (f\circ T^\ell - \nu(f) - D_\ell)\Big|\,.$$

Consequently, since $L \leq N$ and $D_\ell = d_{j,\ell}$ if $\ell \in \{2^j, ..., 2^{j+1}-1\}$, we derive that

$$\begin{aligned}\sup_{1\leq k\leq 2^N} |S_k(f) - M_k(f)| &\leq |f - \nu(f) - D_0| + \sum_{j=0}^{N-1} \sup_{0\leq r\leq 2^j-1} \Big|\sum_{\ell=2^j}^{r+2^j} (f\circ T^\ell - \nu(f) - d_{j,\ell})\Big| \\ &\leq |f - \nu(f) - D_0| + \sum_{j=0}^{N-1} \sup_{0\leq r\leq 2^j-1} \Big|\sum_{\ell=2^j}^{r+2^j} ((f-\bar{f}_j)\circ T^\ell - \nu(f-\bar{f}_j))\Big| \\ &\qquad + \sum_{j=0}^{N-1} \sup_{0\leq r\leq 2^j-1} \Big|\sum_{\ell=2^j}^{r+2^j} (\bar{f}_j\circ T^\ell - \nu(\bar{f}_j) - d_{j,\ell})\Big|\,.\end{aligned}$$

Therefore (6.16) (and then (6.8)) will follow if we can prove that

$$\sup_{0\leq r\leq 2^j-1} \Big|\sum_{\ell=2^j}^{r+2^j} ((f-\bar{f}_j)\circ T^\ell - \nu(f-\bar{f}_j))\Big| = o\big(j^{1-2/p}\, 2^{j/p}\big) \ \nu\text{-a.s.} \tag{6.17}$$

and that

$$\sup_{0\le r\le 2^j-1}\Big|\sum_{\ell=2^j}^{r+2^j}(\bar f_j\circ T^\ell-\nu(\bar f_j)-d_{j,\ell})\Big| = o\big(j^{1-2/p}\,2^{j/p}\big)\ \nu\text{-a.s.} \tag{6.18}$$

We first show the almost sure convergence (6.17). By stationarity, this will hold true if we can show that

$$\sum_{j\ge1}\frac{\|\sup_{1\le k\le 2^j}|\sum_{\ell=0}^{k-1}((f-\bar f_j)\circ T^\ell-\nu(f-\bar f_j))|\|_{2,\nu}^2}{2^{2j/p}j^{2-4/p}}<\infty\,. \tag{6.19}$$

Recall that $f-\bar f_j$ is the limit in $\mathbb{L}^p(\nu)$ of the function $f_{\varphi(L)}-f_{\varphi(L),j}$ defined in (6.12). Therefore, to prove (6.19), it suffices to prove that there exists some positive constant C, such that, for any positive integer L,

$$\sum_{j\ge1}\frac{\|\sup_{1\le k\le 2^j}|\sum_{\ell=0}^{k-1}((f_{\varphi(L)}-f_{\varphi(L),j})\circ T^\ell-\nu(f_{\varphi(L)}-f_{\varphi(L),j}))|\|_{2,\nu}^2}{2^{2j/p}j^{2-4/p}}<C\,. \tag{6.20}$$

To prove (6.20), we use Proposition 1 in [22] (that is the corresponding maximal inequality of [31] for non-invertible maps). Using stationarity, this leads to

$$\begin{aligned}&\|\sup_{1\le k\le 2^j}|\sum_{\ell=0}^{k-1}((f_{\varphi(L)}-f_{\varphi(L),j})\circ T^\ell-\nu(f_{\varphi(L)}-f_{\varphi(L),j}))|\|_{2,\nu}^2\\&\qquad\ll 2^j\|f_{\varphi(L)}-f_{\varphi(L),j}\|_{2,\nu}^2+2^j\Big(\sum_{n=1}^{2^j}n^{-1/2}\|K^n(f_{\varphi(L)}-f_{\varphi(L),j})-\nu(f_{\varphi(L)}-f_{\varphi(L),j})\|_{2,\nu}\Big)^2.\end{aligned} \tag{6.21}$$

Notice that by (6.12), $\|f_{\varphi(L)}-f_{\varphi(L),j}\|_{2,\nu}\le\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|\widetilde t_j\circ g_{\ell,\varphi(L)}\|_{2,\nu}$. Therefore, since $\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\le 1$, by Jensen's inequality,

$$\|f_{\varphi(L)}-f_{\varphi(L),j}\|_{2,\nu}^2\le\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|\widetilde t_j\circ g_{\ell,\varphi(L)}\|_{2,\nu}^2\,. \tag{6.22}$$

On the other hand, by (6.12),

$$\|K^n(f_{\varphi(L)}-f_{\varphi(L),j})-\nu(f_{\varphi(L)}-f_{\varphi(L),j})\|_{2,\nu}\le\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|K^n(\widetilde t_j\circ g_{\ell,\varphi(L)})-\nu(\widetilde t_j\circ g_{\ell,\varphi(L)})\|_{2,\nu}\,.$$

Applying Lemma 6.1,

$$\|K^n(\widetilde t_j\circ g_{\ell,\varphi(L)})-\nu(\widetilde t_j\circ g_{k,\varphi(L)})\|_{2,\nu}\ll\rho^{n/2}\|\widetilde t_j\circ g_{\ell,\varphi(L)}\|_{2,\nu}\,. \tag{6.23}$$

Hence by Jensen's inequality,

$$\begin{aligned}&\Big(\sum_{n=1}^{2^j}n^{-1/2}\|K^n(f_{\varphi(L)}-f_{\varphi(L),j})-\nu(f_{\varphi(L)}-f_{\varphi(L),j})\|_{2,\nu}\Big)^2\\&\qquad\ll\Big(\sum_{n\ge1}n^{-1/2}\rho^{n/2}\Big)^2\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|\widetilde t_j\circ g_{\ell,\varphi(L)}\|_{2,\nu}^2\,.\end{aligned} \tag{6.24}$$

Therefore, using (6.21) together with the upper bounds (6.22) and (6.24), we derive that, for any positive integer j,

$$\begin{aligned}&\frac{\|\sup_{1\le k\le 2^j}|\sum_{\ell=0}^{k-1}((f_{\varphi(L)}-f_{\varphi(L),j})\circ T^\ell-\nu(f_{\varphi(L)}-f_{\varphi(L),j}))|\|_{2,\nu}^2}{2^{2j/p}j^{2-4/p}}\\&\qquad\ll\frac{2^{j(p-2)/p}}{j^{2(p-2)/p}}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|\widetilde t_j\circ g_{\ell,\varphi(L)}\|_{2,\nu}^2\,.\end{aligned}$$

Now, via Fubini's theorem, there exists a positive constant C not depending on L such that

$$\sum_{j\geq 1}\frac{2^{j(p-2)/p}}{j^{2(p-2)/p}}\|\widetilde{t}_j\circ g_{\ell,\varphi(L)}\|_{2,\nu}^2=\sum_{j\geq 1}\frac{2^{j(p-2)/p}}{j^{2(p-2)/p}}\nu\Big(g_{\ell,\varphi(L)}^2\mathbf{1}_{|g_{\ell,\varphi(L)}|>2^{j/p}j^{-2/p}}\Big)$$
$$<C\|g_{\ell,\varphi(L)}\|_{p,\nu}^p\leq CM^p\,.$$

Using the fact that $\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\leq 1$, (6.20) follows. This ends the proof of the (6.17).

We turn now to the proof of (6.18). By stationarity, this will hold true if we can show that

$$\sum_{j\geq 1}\frac{\big\|\sup_{1\leq k\leq 2^j}\big|\sum_{\ell=0}^{k-1}(\bar f_j\circ T^\ell-\nu(\bar f_j)-m_j\circ T^\ell)\big|\big\|_{4,\nu}^4}{2^{4j/p}j^{4(p-2)/p}}<\infty\,. \tag{6.25}$$

To prove it we shall apply Proposition 5.1 with $f=\bar f_j$ (recall that (6.14) holds). This leads to

$$\Big\|\sup_{1\leq k\leq 2^j}\Big|\sum_{\ell=0}^{k-1}(\bar f_j\circ T^\ell-\nu(\bar f_j)-m_j\circ T^\ell)\Big|\Big\|_{4,\nu}$$
$$\ll\sum_{k=1}^{2^{2j}}\|K^k\bar f_j-\nu(\bar f_j)\|_{4,\nu}+2^{j/2}\sum_{k\geq 2^{2j}}\frac{\|K^k\bar f_j-\nu(\bar f_j)\|_{4,\nu}}{k^{1/4}}+2^{j/4}\sum_{k=1}^{2^j}\frac{\|K^k\bar f_j-\nu(\bar f_j)\|_{4,\nu}}{\ell^{1/4}}\,. \tag{6.26}$$

But,

$$\|K^k\bar f_j-\nu(\bar f_j)\|_{4,\nu}\leq\liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|K^kt_j\circ g_{\ell,\varphi(L)}-\nu(t_j\circ g_{\ell,\varphi(L)})\|_{4,\nu}\,.$$

Hence, by using Lemma 6.1, we get that

$$\|K^k\bar f_j-\nu(\bar f_j)\|_{4,\nu}\ll\rho^{3k/4}\liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|t_j\circ g_{\ell,\varphi(L)})\|_{4,\nu}\,,$$

Taking into account this last estimate in (6.26), we derive that

$$\Big\|\sup_{1\leq k\leq 2^j}\Big|\sum_{\ell=0}^{k-1}(\bar f_j\circ T^\ell-\nu(\bar f_j)-m_j\circ T^\ell)\Big|\Big\|_{4,\nu}\ll 2^{j/4}\liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|t_j\circ g_{\ell,\varphi(L)})\|_{4,\nu}\,.$$

Next, since $\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\leq 1$, by using Jensen's inequality,

$$\Big\|\sup_{1\leq k\leq 2^j}\Big|\sum_{\ell=0}^{k-1}(\bar f_j\circ T^\ell-\nu(\bar f_j)-m_j\circ T^\ell)\Big|\Big\|_{4,\nu}^4\ll 2^{j}\liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|t_j\circ g_{\ell,\varphi(L)})\|_{4,\nu}^4\,.$$

Hence, by Fatou's lemma, we get that

$$\sum_{j\geq 1}\frac{\big\|\sup_{1\leq k\leq 2^j}\big|\sum_{\ell=0}^{k-1}(\bar f_j\circ T^\ell-\nu(\bar f_j)-m_j\circ T^\ell)\big|\big\|_{4,\nu}^4}{2^{4j/p}j^{4(p-2)/p}}$$
$$\ll\liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\sum_{j\geq 1}\frac{2^j\|t_j\circ g_{\ell,\varphi(L)})\|_{4,\nu}^4}{2^{4j/p}j^{4(p-2)/p}}\,. \tag{6.27}$$

Notice now that, by applying Fubini's theorem, there exists a positive constant C not depending on L such that

$$\sum_{j\geq 1}\frac{2^j}{2^{4j/p}j^{4(p-2)/p}}\|t_j\circ g_{\ell,\varphi(L)}\|_{4,\nu}^4\leq\sum_{j\geq 1}\Big(\frac{2^{j/p}}{j^{2/p}}\Big)^{p-4}\frac{2^j}{2^{4j/p}j^{4(p-2)/p}}\nu\Big(g_{\ell,\varphi(L)}^4\mathbf{1}_{|g_{\ell,\varphi(L)}|\leq 2^{j/p}j^{-2/p}}\Big)$$
$$\leq C\|g_{\ell,\varphi(L)}\|_{p,\nu}^p\leq CM^p\,,$$

Therefore, by taking into account this last estimate in (6.27) together with the fact that $\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\leq 1$, the convergence (6.25) follows. This ends the proof of (6.16) and then of the almost sure convergence (6.8).

□

6.3.4 Step 4. Proof of the almost sure approximation (6.9)

We shall first prove that

$$\sigma^2 = \lim_{n\to\infty} n^{-1} \sum_{\ell=0}^{n-1} \mathbb{E}_\nu(D_\ell^2)\,, \tag{6.28}$$

where σ^2 is defined by (3.4). With this aim, we first notice that, since $f \in \mathrm{Mon}_p^c(M,\nu)$, the condition (5.8) is satisfied. Indeed

$$\|K^n f - \nu(f)\|_{p,\nu} \le \liminf_{L\to\infty} \sum_{k=1}^{\varphi(L)} |a_{k,\varphi(L)}| \|K^n g_{k,\varphi(L)} - \nu(g_{k,\varphi(L)})\|_{p,\nu}\,.$$

Hence, by using Lemma 6.1 and the fact that $\sum_{k=1}^{\varphi(L)} |a_{k,\varphi(L)}| \le 1$, we get that

$$\|K^n f - \nu(f)\|_{p,\nu} \ll \rho^{n(p-1)/p} M\,, \tag{6.29}$$

which clearly shows that (5.8) holds. Therefore, according to Proposition 5.1, the series $D := \sum_{n\ge0}(K^n f - UK^{n+1}f)$ converges in $\mathbb{L}^p(\nu)$, and $\mathbb{E}_\nu(D|\mathcal{G}_1) = 0$ ν-a.s. Moreover $\nu(D^2) = \sigma^2$. Therefore (6.28) will follow if we can prove that

$$\Big\| \sum_{\ell=0}^{n-1} (D_\ell - D\circ T^\ell) \Big\|_{2,\nu} = o(n^{1/2})\,. \tag{6.30}$$

Let N be the positive integer such that $2^{N-1} < n \le 2^N$. Since $(D_\ell - D\circ T^\ell)_{\ell\ge0}$ is a sequence of reverse martingale differences, we get that

$$\Big\| \sum_{\ell=0}^{n-1} (D_\ell - D\circ T^\ell) \Big\|_{2,\nu}^2 = \sum_{\ell=0}^{n-1} \big\| D_\ell - D\circ T^\ell \big\|_{2,\nu}^2 \le 2\|D - m_0\|_{2,\nu}^2 + \sum_{j=1}^{N-1} 2^j \|D - m_j\|_{2,\nu}^2\,. \tag{6.31}$$

But, for any non-negative integer j,

$$\|D - m_j\|_{2,\nu} \le 2 \sum_{k\ge0} \|K^k(f - \bar f_j) - \nu(f - \bar f_j)\|_{2,\nu}\,,$$

and, by (6.12),

$$\|K^k(f - \bar f_j) - \nu(f - \bar f_j)\|_{2,\nu} \le \liminf_{L\to\infty} \sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \|K^k(\widetilde{t}_j \circ g_{\ell,\varphi(L)}) - \nu(\widetilde{t}_j \circ g_{\ell,\varphi(L)})\|_{2,\nu}\,.$$

Therefore by taking into account (6.23),

$$\|D - m_j\|_{2,\nu} \ll \liminf_{L\to\infty} \sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \|g_{\ell,\varphi(L)} \mathbf{1}_{|g_{\ell,\varphi(L)}| > 2^{j/p}(j+1)^{-2/p}}\|_{2,\nu} \tag{6.32}$$

Since $\|g_{\ell,\varphi(L)} \mathbf{1}_{|g_{\ell,\varphi(L)}| > 2^{j/p}(j+1)^{-2/p}}\|_{2,\nu}^2 \le 2^{j(2-p)/p}(j+1)^{2(p-2)/p} \|g_{\ell,\varphi(L)}\|_{p,\nu}^p$ and $\sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \le 1$, we get that

$$\sum_{j=0}^{N-1} 2^j \|D - m_j\|_{2,\nu}^2 \ll M^p \sum_{j=1}^{N} 2^{2j/p} j^{2(p-2)/p} \ll M^p 2^{2N/p} N^{2-4/p}\,,$$

which combined with (6.31) proves (6.30). This ends the proof of (6.28). Notice also that (6.29) implies (5.7) that is a sufficient condition for the absolute convergence of the series σ^2.

In the trivial case where $\sigma = 0$, the theorem is true by (6.16) (with an even better rate). Hence we shall assume that $\sigma > 0$.

Let now $a_n = n^{2/p}(\log n)^{1-4/p}$ and assume, for a while, that we can prove that

$$\sum_{\ell\ge2} a_\ell^{-\alpha} \mathbb{E}_\nu(|D_\ell|^{2\alpha}) < \infty \quad \text{for some } 1 \le \alpha \le 2\,, \tag{6.33}$$

and that

$$\sum_{\ell=0}^{n-1}\big(\mathbb{E}_\nu(D_\ell^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(D_\ell^2)\big)=o(a_n)\ \ \nu\text{-a.s.} \tag{6.34}$$

Then, according to Theorem 2.3, we see that, enlarging our probability space if necessary, one may find a sequence $(\bar{Z}_\ell)_{\ell\geq 1}$ of independent Gaussian random variables with zero mean and variance $\mathbb{E}(\bar{Z}_\ell)^2 = \mathbb{E}_\nu(D_\ell^2)=\sigma_\ell^2$ such that

$$\sup_{1\leq k\leq n}\big|M_k(f)-\sum_{\ell=1}^{k}\bar{Z}_\ell\big|=o\big(a_n^{1/2}(\log n)^{1/2}\big)\ \ \nu\text{-a.s.} \tag{6.35}$$

Let $(\delta_k)_{k\geq1}$ be a sequence of iid Gaussian random variables with mean zero and variance σ^2, independent of the sequence $(\bar{Z}_\ell)_{\ell\geq1}$. We now construct a sequence $(Z_\ell)_{\ell\geq1}$ as follows. If $\bar{\sigma}_\ell=0$, then $Z_\ell=\delta_\ell$, else $Z_\ell=(\sigma/\bar{\sigma}_\ell)\bar{Z}_\ell$. By construction, the Z_ℓ's are iid Gaussian random variables with mean zero and variance σ^2. Let $G_\ell=Z_\ell-\bar{Z}_\ell$ and note that $(G_\ell)_{\ell\geq1}$ is a sequence of independent Gaussian random variables with mean zero and variances $\mathbb{E}_\nu(G_\ell^2)=(\sigma-\bar{\sigma}_\ell)^2$. Assume that we can prove that

$$\sum_{n\geq3}\frac{\mathbb{E}_\nu(G_n^2)}{a_n\log n}<\infty\,. \tag{6.36}$$

Then by the Kolmogorov theorem (or Lemma 4.2), it will follow that the series $\sum_{n\geq3}\frac{G_n}{(a_n\log n)^{1/2}}$ converges ν-a.s. Hence, Kronecker lemma will imply that $\sum_{\ell=1}^n G_\ell=o((a_n\log n)^{1/2})$ ν-a.s. Therefore starting from (6.35), we will conclude that if (6.33) and (6.34) hold then (6.9) does. So the proof of the theorem will be complete if we can prove (6.33), (6.34) and (6.36).

We start by proving (6.36). With this aim, we first notice that

$$\mathbb{E}_\nu(G_n^2)=\big(\|D\circ T^n\|_{2,\nu}-\|D_n\|_{2,\nu}\big)^2\leq\|D\circ T^n-D_n\|_{2,\nu}^2\,.$$

Next

$$\sum_{n\geq3}\frac{\mathbb{E}_\nu(G_n^2)}{a_n\log n}\leq\sum_{j\geq1}\frac{1}{2^{2j/p}j^{2-4/p}}\sum_{\ell=2^j}^{2^{j+1}-1}\mathbb{E}_\nu((D\circ T^\ell-d_{j,\ell})^2)=\sum_{j\geq1}\frac{2^j}{2^{2j/p}j^{2-4/p}}\mathbb{E}_\nu((D-m_j)^2)\,.$$

Taking into account (6.32), and using Fatou's lemma and Jensen's inequality (recall that $\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\leq 1$), we then get that

$$\sum_{n\geq3}\frac{\mathbb{E}_\nu(G_n^2)}{a_n\log n}\ll\liminf_{L\to\infty}\sum_{j\geq1}\frac{2^j}{2^{2j/p}j^{2-4/p}}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|g_{\ell,\varphi(L)}\mathbf{1}_{|g_{\ell,\varphi(L)}|>2^{j/p}j^{-2/p}}\|_{2,\nu}^2\,.$$

Hence, by Fubini's theorem, there exists a positive constant C not depending on L such that

$$\begin{aligned}\sum_{n\geq3}\frac{\mathbb{E}_\nu(G_n^2)}{a_n\log n}&\ll\liminf_{L\to\infty}\sum_{j\geq1}\frac{2^j}{2^{2j/p}j^{2-4/p}}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\nu(g_{\ell,\varphi(L)}^2\mathbf{1}_{|g_{\ell,\varphi(L)}|>2^{j/p}j^{-2/p}})\\&\ll\liminf_{L\to\infty}\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\sum_{j\geq1}\Big(\frac{2^{j/p}}{j^{2/p}}\Big)^{p-2}\nu(g_{\ell,\varphi(L)}^2\mathbf{1}_{|g_{\ell,\varphi(L)}|>2^{j/p}j^{-2/p}})\\&<C\liminf_{L\to\infty}\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\|g_{\ell,\varphi(L)}\|_{p,\nu}^p\leq CM^p\,.\end{aligned}$$

This ends the proof of (6.36).

We prove now that (6.33) holds for $\alpha=2$. With this aim, we first notice that

$$\sum_{\ell\geq2}a_\ell^{-2}\mathbb{E}_\nu(D_\ell^4)=\sum_{j\geq1}\sum_{\ell=2^j}^{2^{j+1}-1}a_\ell^{-2}\mathbb{E}_\nu(d_{j,\ell}^4)\ll\sum_{j\geq1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}\mathbb{E}_\nu(m_j^4)\,.$$

But $\|m_j\|_{4,\nu} \le 2\sum_{k\ge 0}\|K^k(\bar f_j)-\nu(\bar f_j)\|_{4,\nu}$ and, by (6.12),

$$\|K^k(\bar f_j)-\nu(\bar f_j)\|_{4,\nu} \le \liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|K^k(t_j\circ g_{\ell,\varphi(L)})-\nu(t_j\circ g_{\ell,\varphi(L)})\|_{4,\nu}\,. \tag{6.37}$$

Applying Lemma 6.1 and Jensen's inequality (recall that $\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\le 1$), it follows that

$$\|m_j\|_{4,\nu}^4 \ll \liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|g_{\ell,\varphi(L)}\mathbf{1}_{|g_{\ell,\varphi(L)}|\le 2^{j/p}j^{-2/p}}\|_{4,\nu}^4\,. \tag{6.38}$$

Therefore, by Fatou's lemma followed by Fubini's theorem, we infer that there exists a positive constant C not depending on L such that

$$\begin{aligned}\sum_{\ell\ge 2}a_\ell^{-2}\mathbb{E}_\nu(D_\ell^4) &\ll \liminf_{L\to\infty}\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\sum_{j\ge 1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}|\nu(g^4_{k,\varphi(L)}\mathbf{1}_{|g_{k,\varphi(L)}|\le 2^{j/p}j^{-2/p}})\\ &\ll \liminf_{L\to\infty}\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\sum_{j\ge 1}\Big(\frac{2^{j/p}}{j^{2/p}}\Big)^{p-4}\nu(g^4_{k,\varphi(L)}\mathbf{1}_{|g_{k,\varphi(L)}|\le 2^{j/p}j^{-2/p}})\\ &< C\liminf_{L\to\infty}\sum_{k=1}^{\varphi(L)}|a_{k,\varphi(L)}|\|g_{k,\varphi(L)}\|_{p,\nu}^p \le CM^p\,.\end{aligned} \tag{6.39}$$

This ends the proof of (6.33).

We prove now that (6.34) is satisfied. With this aim, we first notice that, according to (6.15), for any non-negative integer j,

$$A_j := \sup_{1\le r\le 2^j}\Big|\sum_{\ell=2^j}^{2^j+r-1}\big(\mathbb{E}_\nu(D_\ell^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(D_\ell^2)\big)\Big| = \sup_{1\le r\le 2^j}\Big|\sum_{\ell=2^j}^{2^j+r-1}\big(\mathbb{E}_\nu(d_{j,\ell}^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(d_{j,\ell}^2)\big)\Big|\,.$$

Let $N\in\mathbb{N}^*$ and let $k\in\{1,\dots,2^N\}$. Note that $A_j\ge|\sum_{\ell=2^j}^{2^{j+1}-1}(\mathbb{E}_\nu(d_{j,\ell}^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(d_{j,\ell}^2))|$. Hence, if L is the integer such that $2^{L-1}<k\le 2^L$, then

$$\Big|\sum_{\ell=0}^{k-1}\big(\mathbb{E}_\nu(D_\ell^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(D_\ell^2)\big)\Big| \le |\mathbb{E}_\nu(D_0^2|\mathcal{G}_1)-\mathbb{E}_\nu(D_0^2)|+\sum_{j=0}^{L-1}A_j\,.$$

Consequently since $L\le N$,

$$\sup_{1\le k\le 2^N}\Big|\sum_{\ell=0}^{k-1}\big(\mathbb{E}_\nu(D_\ell^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(D_\ell^2)\big)\Big| \le |\mathbb{E}_\nu(D_0^2|\mathcal{G}_1)-\mathbb{E}_\nu(D_0^2)|+\sum_{j=0}^{N-1}A_j\,.$$

Therefore to prove (6.34), it is enough to show that, for any non-negative integer j, the following almost sure convergence holds: $A_j=o(2^{2j/p}(j+1)^{1-4/p})$ ν-a.s. In particular, it suffices to prove that

$$\sum_{j\ge 0}\frac{\big\|\sup_{1\le r\le 2^j}\big|\sum_{\ell=0}^{r-1}\big(\mathbb{E}_\nu(d_{j,\ell}^2|\mathcal{G}_{\ell+1})-\mathbb{E}_\nu(d_{j,\ell}^2)\big)\big|\big\|_{2,\nu}^2}{2^{4j/p}(j+1)^{2(1-4/p)}}<\infty\,. \tag{6.40}$$

For any non-negative integer j, define

$$\varphi_j(x,y) := \sum_{n\ge 0}(K^n\bar f_j(x)-K^{n+1}\bar f_j(y))\,.$$

Recall that it has been shown in Step 2 that this function is well defined in $\mathbb{L}^q(\nu)$ for any $q\in[1,\infty]$. Consider now a stationary Markov chain $(Y_n)_{n\in\mathbb{Z}}$ with transition kernel K and invariant distribution ν. Let $\mathcal{F}_k=\sigma(Y_\ell,\ell\le k)$. Define

$$V_{j,n} := \sum_{k=1}^{n}\varphi_j(Y_k,Y_{k-1})\,.$$

Arguing as in Step 5 of the proof of Theorem 3.1, we infer that

$$\Big\| \sup_{1\le r\le 2^j} \Big| \sum_{\ell=0}^{r-1} \big((\mathbb{E}_\nu(d_{j,\ell}^2|\mathcal{G}_{\ell+1}) - \mathbb{E}_\nu(d_{j,\ell}^2) \big) \Big| \Big\|_{2,\nu}$$
$$\le 2^{j/2}\|\varphi_j(Y_1,Y_0)\|_4^2 + 2^{j/2}\sum_{r=0}^{j} 2^{-r/2}\|\mathbb{E}(V_{j,2^r}^2|\mathcal{F}_0) - \mathbb{E}(V_{j,2^r}^2)\|_2 .$$

Therefore, (6.40) will hold if we can prove that

$$\sum_{j\ge 0} \frac{2^j\|\varphi_j(Y_1,Y_0)\|_4^4}{2^{4j/p}(j+1)^{2(1-4/p)}} < \infty \quad \text{and} \quad \sum_{j\ge 0} \frac{2^j\big(\sum_{r=0}^{j} 2^{-r/2}\|\mathbb{E}(V_{j,2^r}^2|\mathcal{F}_0) - \mathbb{E}(V_{j,2^r}^2)\|_2\big)^2}{2^{4j/p}(j+1)^{2(1-4/p)}} < \infty . \tag{6.41}$$

By (5.1), $\|\varphi_j(Y_1,Y_0)\|_4 = \|m_j\|_{4,\nu}$. The first part of (6.41) then follows from (6.38) and the computations used to get (6.39). On another hand, a careful analysis of the proof of Theorem 2.3 in [4] reveals that

$$\|\mathbb{E}(V_{j,2^r}^2|\mathcal{F}_0) - \mathbb{E}(V_{j,2^r}^2)\|_2 \ll \|\mathbb{E}(\tilde S_{j,2^r}^2|\mathcal{F}_0) - \mathbb{E}(\tilde S_{j,2^r}^2)\|_2$$
$$+ \Big(\sum_{n\ge 0}\|K^n(\bar f_j) - \nu(\bar f_j)\|_{4,\nu}\Big)^2 + 2^r\sum_{n\ge 2^r}\|K^n(\bar f_j) - \nu(\bar f_j)\|_{2,\nu} ,$$

where $\tilde S_{j,n} = \sum_{\ell=1}^{n}(\bar f_j(Y_\ell) - \nu(\bar f_j))$. Denoting $\bar f_j^{(0)} = \bar f_j - \nu(\bar f_j)$, we have that

$$\|\mathbb{E}(\tilde S_{j,2^r}^2|\mathcal{F}_0) - \mathbb{E}(\tilde S_{j,2^r}^2)\|_2 \le 2\sum_{i=1}^{2^r}\sum_{k=0}^{2^r-i}\|K^i(\bar f_j^{(0)}K^k(\bar f_j^{(0)})) - \nu(\bar f_j^{(0)}K^k(\bar f_j^{(0)}))\|_{2,\nu} .$$

So, overall, the second part of (6.41) will hold provided we can show that

$$\sum_{j\ge 1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}\Big(\sum_{r=0}^{j}2^{-r/2}\sum_{i=1}^{2^r}\sum_{k=0}^{2^r-i}\|K^i(\bar f_j^{(0)}K^k(\bar f_j^{(0)})) - \nu(\bar f_j^{(0)}K^k(\bar f_j^{(0)}))\|_{2,\nu}\Big)^2 < \infty , \tag{6.42}$$

$$\sum_{j\ge 1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}\Big(\sum_{n\ge 0}\|K^n(\bar f_j) - \nu(\bar f_j)\|_{4,\nu}\Big)^4 < \infty , \tag{6.43}$$

and

$$\sum_{j\ge 1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}\Big(\sum_{r=0}^{j-1}2^{r/2}\sum_{k\ge 2^r}\|K^k(\bar f_j) - \nu(\bar f_j)\|_{2,\nu}\Big)^2 < \infty . \tag{6.44}$$

By using (6.37), Lemma 6.1, Fatou's lemma and Jensen's inequality (since $\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}| \le 1$), we derive that

$$\sum_{j\ge 1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}\Big(\sum_{n\ge 0}\|K^n(\bar f_j) - \nu(\bar f_j)\|_{4,\nu}\Big)^4$$
$$\ll \liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\sum_{j\ge 1}\frac{2^j}{2^{4j/p}j^{2(1-4/p)}}\nu\Big(g_{\ell,\varphi(L)}^4\mathbf{1}_{|g_{\ell,\varphi(L)}|\le 2^{j/p}j^{-2/p}}\Big) ,$$

which together with an application of Fubini's theorem show (6.43). We turn to the proof of (6.44). We first write that

$$\|K^k(\bar f_j) - \nu(\bar f_j)\|_{2,\nu} \le \liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\|K^k(t_j\circ g_{\ell,\varphi(L)}) - \nu(t_j\circ g_{\ell,\varphi(L)})\|_{2,\nu} .$$

Therefore, applying Lemma 6.1, Fatou's lemma and Jensen's inequality, we get that

$$\Big(\sum_{r=0}^{j-1}2^{r/2}\sum_{k\ge 2^r}\|K^k(\bar f_j) - \nu(\bar f_j)\|_{2,\nu}\Big)^2 \ll \liminf_{L\to\infty}\sum_{\ell=1}^{\varphi(L)}|a_{\ell,\varphi(L)}|\nu\Big(g_{\ell,\varphi(L)}^2\mathbf{1}_{|g_{\ell,\varphi(L)}|\le 2^{j/p}j^{-2/p}}\Big) \ll M^2 ,$$

which implies (6.44) since $p < 4$.

It remains to show that (6.42) holds true. Denoting $u_{j,\ell,L} = t_j \circ g_{\ell,\varphi(L)} - \nu(t_j \circ g_{\ell,\varphi(L)})$, we notice that, for any non-negative integers i and k,

$$\|K^i(\bar f_j^{(0)} K^k(\bar f_j^{(0)})) - \nu(\bar f_j^{(0)} K^k(\bar f_j^{(0)}))\|_{2,\nu} \\ \leq \liminf_{L\to\infty} \sum_{\ell=1}^{\varphi(L)} \sum_{s=1}^{\varphi(L)} |a_{\ell,\varphi(L)}||a_{s,\varphi(L)}| \, \|K^i(u_{j,\ell,L}K^k(u_{j,s,L})) - \nu(u_{j,\ell,L}K^k(u_{j,s,L}))\|_{2,\nu} \,. \quad (6.45)$$

Applying Lemma 6.1, we get that, for any non-negative integers i and k,

$$\|K^i(u_{j,\ell,L}K^k(u_{j,s,L})) - \nu(u_{j,\ell,L}K^k(u_{j,s,L}))\|_{2,\nu} \ll \rho^{i/2} \|t_j \circ g_{\ell,\varphi(L)}\|_{4,\nu} \|t_j \circ g_{s,\varphi(L)}\|_{4,\nu} \,. \quad (6.46)$$

On another hand, applying Lemma 6.1, we also have that for any non-negative integers i and k,

$$\|K^i(u_{j,\ell,L}K^k(u_{j,s,L})) - \nu(u_{j,\ell,L}K^k(u_{j,s,L}))\|_{2,\nu} \\ \leq 2\|u_{j,\ell,L}\|_{4,\nu} \|K^k(u_{j,s,L}))\|_{4,\nu} \ll \rho^{3k/4} \|t_j \circ g_{\ell,\varphi(L)}\|_{4,\nu} \|t_j \circ g_{s,\varphi(L)}\|_{4,\nu} \,. \quad (6.47)$$

Starting from (6.45) and using (6.46), (6.47) and Fatou's lemma, we get that

$$\Big(\sum_{r=0}^{j} 2^{-r/2} \sum_{i=1}^{2^r} \sum_{k=0}^{2^r - i} \|K^i(\bar f_j^{(0)} K^k(\bar f_j^{(0)})) - \nu(\bar f_j^{(0)} K^k(\bar f_j^{(0)}))\|_{2,\nu} \Big)^2 \ll \liminf_{L\to\infty} \Big(\sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \|t_j \circ g_{\ell,\varphi(L)}\|_{4,\nu} \Big)^4 .$$

Therefore, by Fatou's lemma and Jensen's inequality (since $\sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \leq 1$), we derive that

$$\sum_{j\geq 1} \frac{2^j}{2^{4j/p} j^{2(1-4/p)}} \Big(\sum_{r=0}^{j} 2^{-r/2} \sum_{i=1}^{2^r} \sum_{k=0}^{2^r - i} \|K^i(\bar f_j^{(0)} K^k(\bar f_j^{(0)})) - \nu(\bar f_j^{(0)} K^k(\bar f_j^{(0)}))\|_{2,\nu} \Big)^2 \\ \ll \liminf_{L\to\infty} \sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \sum_{j\geq 1} \Big(\frac{2^{j/p}}{j^{2/p}} \Big)^{p-4} \nu\big(g^4_{\ell,\varphi(L)} \mathbf{1}_{|g_{\ell,\varphi(L)}| \leq 2^{j/p} j^{-2/p}} \big) .$$

Whence, the convergence (6.42) follows by applying Fubini's theorem and taking into account the fact that $\sum_{\ell=1}^{\varphi(L)} |a_{\ell,\varphi(L)}| \leq 1$. This ends the proof of the second part of (6.41) and then of (6.34). The proof of the theorem is therefore complete. □

6.4 Proof of Corollary 3.9

We first recall the spectral gap property of K on the space of Hölder functions.

Let $\alpha \in (0,1]$ and denote by $C^\alpha(X)$ the Banach space of real-valued α-Hölder functions, with norm $\|\cdot\|_\alpha$ given by $\|f\|_\alpha = \|f\|_\infty + \sup_{x\neq y} \frac{|f(x)-f(y)|}{d(x,y)^\alpha}$.

It follows from Theorem 2.4 of [1] that K is quasi-compact on $C^\alpha(X)$. More precisely (see the first equation page 85 of [1] and use that φ is bounded from below), there exist $C > 0$ and $0 < \tau < 1$ such that

$$\|K^n f - \nu(f)\|_\alpha \leq C\tau^n \|f\|_\alpha \qquad \forall f \in C^\alpha(X) \,. \quad (6.48)$$

Using Lemma 15 in [9] (their proof works in our setting, see also the proof of Lemma 6.2 below), we derive that, for every $f \in \Gamma_c$,

$$\|K^n(f) - \nu(f)\|_{\infty,\nu} \leq M_f c(C\tau^n) \,, \quad (6.49)$$

where $M_f > 0$.

We shall now obtain a similar bound for $\|K^n(fK^m f) - \nu(K^n(fK^m f))\|_{\infty,\nu}$ when $f \in \Gamma_c$, thanks to a lemma similar to Lemma 15 of [9].

Define a metric d_1 on X^2, by $d_1((x_1,x_2),(y_1,y_2)) := \frac{1}{2}(d(x_1,y_1) + d(x_2,y_2))$ and denote by $\Lambda(X^2)$ the set of Lipschitz function on X^2 with respect to d_1 with Lipschitz norm $\|\cdot\|_{\Lambda(X^2)}$.

For every $h \in \Lambda(X^2)$ and every $\ell \geq 0$, write $h_\ell := \int_X h(\cdot,y) K^\ell(\cdot,dy)$.

Let us admit the following lemma for a while. Let us also admit that condition (6.50) holds true with $\rho = \tau$.

Lemma 6.2 *Assume that there exists $C > 0$ and $\rho \in (0,1)$ such that for every $h \in \Lambda(X^2)$ and every integers $n, \ell \geq 0$,*

$$\|K^n h_\ell - \nu(K^n h_\ell)\|_{\infty,\nu} \leq C\rho^n \|h\|_{\Lambda(X^2)} \,. \tag{6.50}$$

Then, for every non-decreasing concave function c with $c(0) = 0$, and every $f \in \Gamma_c$, the following inequality holds: for any non-negative integers n and m,

$$\|K^n(fK^m f) - \nu(K^n(fK^m f))\|_{\infty,\nu} \leq 2\|f\|_\infty c(C\rho^n) \,. \tag{6.51}$$

Let $2 < p \leq 4$. We shall apply Theorem 3.1 with $t = 1$. Let $f \in \Gamma_c$ and write $g := f - \nu(f)$.

Let $(Y_n)_{n\in\mathbb{Z}}$ be a stationary Markov chain on X with transition probability given by K and stationary distribution ν. Then, conditions (3.2) and (3.3) (with $t = 1$) read

$$\sum_{n\geq 2} \frac{n^{p-1}}{n^{2/p}} \|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_p^p < \infty \text{ and } \sum_{n\geq 2} \frac{n^{3p/4}}{n^2} \|\mathbb{E}(g(Y_n)|\mathcal{F}_0)\|_2^{p/2} < \infty \,, \tag{6.52}$$

and

$$\sum_{n\geq 2} \frac{1}{n^2} \Big(\sum_{i=1}^{n} \sum_{j=0}^{n-i} \|\mathbb{E}(g(Y_i)g(Y_{i+j})|\mathcal{F}_0) - \mathbb{E}(g(Y_i)g(Y_{i+j}))\|_{\frac{p}{2}} \Big)^{p/2} < \infty \,. \tag{6.53}$$

Hence, by the proof of Corollary 2.1 of [4], we see that conditions (3.2) and (3.3) (with $t = 1$) will hold as soon as there exists $\gamma \in]0,1]$ such that

$$\sum_{n\geq 2} \frac{n^{(\frac{p}{2}-1)(\frac{1}{\gamma}+1)}}{n^{1/2}} \|K^n g\|_{p,\nu}^{p/2} < \infty \,,$$

and

$$\sum_{n\geq 2} \frac{n^{(\gamma+1)p/2}}{n^2} \sup_{j\geq 0} \|K^n(gK^j(g)) - \nu(gK^j(g))\|_{\frac{p}{2},\nu}^{p/2} < \infty \,.$$

Taking $\gamma = \frac{1+\sqrt{1+4p(p-2)}}{2p}$ and using (6.49) and (6.51), we then see that (3.2) and (3.3) hold (with $t = 1$) as soon as

$$\sum_{n\geq 1} n^\delta \big(c(C\tau^n)\big)^{p/2} < \infty \,,$$

where $\delta = \frac{2p+\sqrt{1+4p(p-2)}}{4} - \frac{7}{4}$. By using standard comparison between series with integrals, this condition holds as soon as (3.11) does. □

To end the proof of Corollary 3.9 it suffices to prove (6.50) and to prove Lemma 6.2.

Proof of (6.50)**.** By (6.48), it suffices to prove that there exists $C > 0$, such that for every $h \in \Lambda(X^2)$ and every $\ell \geq 0$,

$$\|h_\ell\|_\alpha \leq C\|h\|_{\Lambda(X^2)} \,.$$

Let $x, x' \in X$. Write $\varphi_x(\cdot) = h(x,\cdot)$ and notice that φ_x is α-Hölder with $\|\varphi_x\|_\alpha \leq C\|h\|_{\Lambda(X^2)}$, for some C depending on the diameter of X. We have, using (6.48),

$$\begin{aligned} |h_\ell(x) - h_\ell(x')| &\leq \int_X |h(x,y) - h(x',y)| K^\ell(x',dy) \;+\; |K^\ell\varphi_x(x) - K^\ell\varphi_x(x')| \\ &\leq \frac{1}{2} d(x,x')\|h\|_{\Lambda(X^2)} + C\tau^\ell \|h\|_{\Lambda(X^2)} d(x,x')^\alpha \leq \tilde{C}\|h\|_{\Lambda(X^2)} d(x,x')^\alpha \,. \end{aligned}$$

Proof of Lemma 6.2. Let $f \in \Gamma_c$.

From Lemma 1 in Dedecker and Merlevède [7], we know that for i and j non-negative integers, there exists (Y_i^*, Y_j^*) distributed as (Y_i, Y_j) and independent of Y_0 such that

$$\big\|\mathbb{E}(d_1((Y_i,Y_j),(Y_i^*,Y_j^*))|Y_0)\big\|_\infty = \sup_{h\in\Lambda(X^2), \|h\|_{\Lambda(X^2)}\leq 1} \big\|\mathbb{E}(h(Y_i,Y_j)|Y_0) - \mathbb{E}(h(Y_i,Y_j))\big\|_\infty \,.$$

Now for every $i \geq j \geq 0$ and every $h \in \Lambda(X^2)$,

$$\big\|\mathbb{E}(h(Y_i, Y_j)|Y_0) - \mathbb{E}(h(Y_i, Y_j))\big\|_\infty = \|K^j(h_{i-j}) - \nu\big((h_{i-j})\big)\|_{\infty,\nu} \leq C\rho^j \|h\|_{\Lambda(X^2)} .$$

On another hand, we clearly have that

$$\begin{aligned}\|K^j(fK^{i-j}(f)) - \nu(fK^{i-j}(f))\|_{\infty,\nu} &= \big\|\mathbb{E}(f(Y_i)f(Y_j) - f(Y_i^*)f(Y_j^*)|Y_0)\big\|_\infty \\ &= \big\|\mathbb{E}((f(Y_i) - f(Y_i^*))f(Y_j)|Y_0) - \mathbb{E}(f(Y_i^*)(f(Y_j^*) - f(Y_j))|Y_0)\big\|_\infty .\end{aligned}$$

Hence,

$$\|K^j(fK^{i-j}(f)) - \nu(fK^{i-j}(f))\|_{\infty,\nu} \leq \|f\|_\infty \big\|\mathbb{E}(c(d(Y_i, Y_i^*))|Y_0)\big\|_\infty + \|f\|_\infty \big\|\mathbb{E}(c(d(Y_j, Y_j^*))|Y_0)\big\|_\infty .$$

Since c is concave and non-decreasing, it follows that

$$\begin{aligned}\|K^j(fK^{i-j}(f)) - \nu(fK^{i-j}(f))\|_{\infty,\nu} &\leq \|f\|_\infty \big\|c\big(\mathbb{E}(d(Y_i, Y_i^*)|Y_0)\big)\big\|_\infty + \|f\|_\infty \big\|c\big(\mathbb{E}(d(Y_j, Y_j^*)|Y_0)\big)\big\|_\infty \\ &\leq \|f\|_\infty c\big(\big\|\mathbb{E}(d(Y_i, Y_i^*)|Y_0)\ \big\|_\infty\big) + \|f\|_\infty c\big(\big\|\mathbb{E}(d(Y_j, Y_j^*)|Y_0)\big\|_\infty\big) \\ &\leq 2\|f\|_\infty c\Big(2\big\|\mathbb{E}(d_1((Y_i, Y_j), (Y_i^*, Y_j^*))|Y_0)\big\|_\infty\Big) .\end{aligned}$$

So overall,

$$\sup_{i \geq j \geq n} \|K^j(fK^{i-j}(f)) - \nu(fK^{i-j}(f))\|_{\infty,\nu} \leq 2\|f\|_\infty c(C\rho^n) ,$$

This ends the proof of the lemma. □

Acknowledgements. The authors would like to thank the two referees for carefully reading the manuscript and for numerous suggestions that improved the presentation of this paper.

References

[1] Baladi, V. *Positive transfer operators and decay of correlations.* (English summary) Advanced Series in Nonlinear Dynamics, 16. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[2] Conze, J.P. and Le Borgne, S. *Méthode de martingales et flot géodésique sur une surface de courbure constante négative.* Ergodic Theory Dynam. Systems 21 (2001), no. 2, 421–441.

[3] Cuny, C. *Pointwise ergodic theorems with rate and application to limit theorems for stationary processes.* Stoch. Dyn. 11 (2011), no. 1, 135–155.

[4] Dedecker, J., Doukhan, P. and Merlevède, F. *Rates of convergence in the strong invariance principle under projective criteria.* Electron. J. Probab. 17 (2012), no. 16, 1–31.

[5] Dedecker, J., Gouëzel, S. and Merlevède, F. *Some almost sure results for intermittent maps and their associated Markov chains.* Ann. Inst. H. Poincaré Probab. Statist. 46 (2010), 796–821.

[6] Dedecker, J., Gouëzel, S. and Merlevède F. *The almost sure invariance principle for unbounded functions of expanding maps.* ALEA, Lat. Am. J. Probab. Math. Stat. 9 (2012), 141–163.

[7] Dedecker, J. and Merlevède, F. *Inequalities for partial sums of Hilbert-valued dependent sequences and applications.* Math. Methods Statist. 15 (2006), 176–206.

[8] Dedecker, J., Merlevède, F. and Peligrad, M. *Invariance principles for linear processes with application to isotonic regression.* Bernoulli 17 (2011), 88–113.

[9] Dedecker, J., Merlevède, F., Peligrad, M. and Utev, S. *Moderate deviations for stationary sequences of bounded random variables.* Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009), 453–476.

[10] Dedecker, J. and Prieur, C. *An empirical central limit theorem for dependent sequences.* Stochastic Process. Appl. 117 (2007), 121–142.

[11] Field, M.J., Melbourne, I. and Török, A. *Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions.* Ergod. Th. & Dyn. Sys. 23 (2003), 87–110.

[12] Gouëzel, S. *Almost sure invariance principle for dynamical systems by spectral methods.* Ann. Probab. 38 (2010), no. 4, 1639–1671.

[13] Gordin, M.I. *The central limit theorem for stationary processes.* Soviet Math. Dokl. 10 (1969), 1174–1176.

[14] Hanson, D.L. and Russo, R. P. *Some results on increments of the Wiener process with applications to lag sums of i.i.d. random variables.* Ann. Probab. 11 (1983), no. 3, 609–623.

[15] Hennion, H. and Hervé. L. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness.* Lecture Notes in Mathematics **1766** Springer. (2001).

[16] Heyde, C.C. *On central limit and iterated logarithm supplements to the martingale convergence theorem.* J. Appl. Probability 14 (1977), no. 4, 758–75.

[17] Hofbauer, F. and Keller, G. *Ergodic properties of invariant measures for piecewise monotonic transformations.* Math. Z. 180 (1982), 119–140.

[18] Huggins, R.M. *A strong invariance principle for reverse martingales.* Acta Math. Hungar. 46 (1985), no. 1-2, 93–100.

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

[20] Ledoux, M. and Talagrand, M. *Probability in Banach spaces. Isoperimetry and processes.* Springer-Verlag, Berlin. 1991.

[21] Liverani, C., Saussol, B. and Vaienti S. *A probabilistic approach to intermittency.* Ergodic Theory Dynam. Systems 19 (1999) 671–685.

[22] Mackey, M. C. and Tyran-Kamińska, M. *Central limit theorems for non-invertible measure preserving maps.* Colloq. Math. 110 (2008), 167–191.

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

[24] Melbourne, I. and Nicol, M. *Almost sure invariance principle for nonuniformly hyperbolic systems.* Commun. Math. Phys. 260 (2005), 131–146.

[25] Melbourne, I. and Nicol, M. *A vector-valued almost sure invariance principle for hyperbolic dynamical systems.* Ann. Probab. (2009), 478–505.

[26] Melbourne, I and Török, A. *Central limit theorems and invariance principles for time-one maps of hyperbolic flows.* Comm. Math. Phys. 229 (2002), 57–71.

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

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

[29] Merlevède, F. and Rio, E. *Strong approximation of partial sums under dependence conditions with application to dynamical systems.* Stochastic Process. Appl. 122 (2012), 386–417.

[30] Neveu, J. *Bases mathmatiques du calcul des probabilits.* (French) Prface de R. Fortet. Deuxime dition, revue et corrige Masson et Cie, diteurs, Paris 1970.

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

[32] Philipp, W. and Stout, W.F. *Almost sure invariance principle for partial sums of weakly dependent random variables.* Mem. of the Amer. Math. Soc. 161 (1975), Providence, RI: Amer. Math. Soc.

[33] Rychlik, M. *Bounded variation and invariant measures.* Studia Math. 76 (1983) 69–80.

[34] Scott, D.J. and Huggins, R.M. *On the embedding ofprocesses in Brownian motion and the law of the iterated logarithm for reverse martingales.* Bull. Austral. Math. Soc. 27 (1983), no. 3, 443–459

[35] Shao, Q.M. *Almost sure invariance principles for mixing sequences of random variables.* Stochastic Process. Appl. 48 (1993), 319–334.

[36] Strassen, V. *An invariance principle for the law of the iterated logarithm.* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 3 (1964), 211–226.

[37] Wu, L. *Functional law of iterated logarithm for additive functionals of reversible Markov processes.* Acta Math. Appl. Sinica, 16 (2000), no. 2, 149–161.

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

[39] Wu, W.B. and Zhao, Z. *Moderate deviations for stationary processes.* Statist. Sinica 18 (2008), 769–782.

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