The almost sure invariance principle for unbounded functions of expanding maps

J. Dedecker*, S. Gouëzel†, and F. Merlevède‡

Abstract

We consider two classes of piecewise expanding maps T of $[0,1]$: a class of uniformly expanding maps for which the Perron-Frobenius operator has a spectral gap in the space of bounded variation functions, and a class of expanding maps with a neutral fixed point at zero. In both cases, we give a large class of unbounded functions f for which the partial sums of $f \circ T^i$ satisfy an almost sure invariance principle. This class contains piecewise monotonic functions (with a finite number of branches) such that:

- For uniformly expanding maps, they are square integrable with respect to the absolutely continuous invariant probability measure.
- For maps having a neutral fixed point at zero, they satisfy an (optimal) tail condition with respect to the absolutely continuous invariant probability measure.

Mathematics Subject Classifications (2000): 37E05, 37C30, 60F15.
Key words: Expanding maps, intermittency, strong invariance principle.

1 Introduction and main results

Our goal in this article is to prove the almost sure invariance principle with error rate $o(\sqrt{n \ln \ln n})$ for several classes of one-dimensional dynamical systems, under very weak integrability or regularity assumptions. We will consider uniformly expanding maps, and maps with an indifferent fixed point, as defined below.

Several classes of uniformly expanding maps of the interval are considered in the literature. We will use the very general definition of Rychlik (1983) to allow infinitely many branches. For notational simplicity, we will assume that there is a single absolutely invariant measure and that it is mixing (the general case can be reduced to this one by looking at subintervals and

*Université Paris Descartes, Laboratoire MAP5 and CNRS UMR 8145, 45 rue des Saints Pères, 75270 Paris Cedex 06, France. E-mail: jerome.dedecker@parisdescartes.fr
†Université Rennes 1, IRMAR and CNRS UMR 6625. E-mail: sebastien.gouezel@univ-rennes1.fr
‡Université Paris Est-Marne la Vallée, LAMA and CNRS UMR 8050. Florence.Merlevede@univ-mlv.fr

at an iterate of the map). We will also need to impose a nontrivial restriction on the density of the measure: it should be bounded away from 0 on its support. This is not always the case, but it is true if there are only finitely many different images (see Zweimüller (1998) for a neat introduction to such classes of maps, or Broise (1996)).

Definition 1.1. *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.*

From this point on, we will simply refer to such maps as *uniformly expanding*. This definition encompasses for instance piecewise C^2 maps with finitely many branches which are all onto, and with derivative everywhere strictly larger than 1 in absolute values.

We consider now a class of expanding maps with a neutral fixed point at zero, as defined below.

Definition 1.2. *A map $T : [0,1] \to [0,1]$ is a generalized Pomeau-Manneville map (or GPM map) of parameter $\gamma \in (0,1)$ if there exist $0 = y_0 < y_1 < \cdots < y_d = 1$ such that, writing $I_k = (y_k, y_{k+1})$,*

1. *The restriction of T to I_k admits a C^1 extension $T_{(k)}$ to $\overline{I_k}$.*

2. *For $k \geq 1$, $T_{(k)}$ is C^2 on $\overline{I_k}$, and $|T'_{(k)}| > 1$.*

3. *$T_{(0)}$ is C^2 on $(0, y_1]$, with $T'_{(0)}(x) > 1$ for $x \in (0, y_1]$, $T'_{(0)}(0) = 1$ and $T''_{(0)}(x) \sim cx^{\gamma-1}$ when $x \to 0$, for some $c > 0$.*

4. *T is topologically transitive.*

For such maps, almost sure invariance principles with good remainder estimates (of the form $O(n^{1/2-\alpha})$ for some $\alpha > 0$) have been established by Melbourne and Nicol (2005) for Hölder observables, and by Merlevède and Rio (2012) under rather mild integrability assumptions. For instance, for uniformly expanding maps, Merlevède and Rio (2012) obtain such a result for a class of observables f in $\mathbb{L}^p(\nu)$ for $p > 2$. This leaves open the question of the boundary case $f \in \mathbb{L}^2(\nu)$. In this case, just like in the i.i.d. case, one can not hope for a remainder

$O(n^{1/2-\alpha})$ with $\alpha > 0$, but one might expect to get $o(\sqrt{n \ln \ln n})$. This would for instance be sufficient to deduce the functional law of the iterated logarithm from the corresponding result for the Brownian motion. The corresponding boundary case for GPM maps has been studied in Dedecker, Gouëzel and Merlevède (2010): we proved a bounded law of the iterated logarithm (i.e., almost surely, $\limsup \sum_{i=0}^{n-1} f \circ T^i / \sqrt{n \log \log n} \leq A < +\infty$), but we were not able to obtain the almost sure invariance principle.

Our goal in the present article is to solve this issue by combining the arguments of the two above papers: we will approximate a function in the boundary case by a function with better integrability properties, use the almost sure invariance principle of Merlevède and Rio (2011) for this better function, and show that the bounded law of the iterated logarithm makes it possible to pass the results from the better function to the original function. This is an illustration of a general method in mathematics: to prove results for a wide class of systems, it is often sufficient to prove results for a smaller (but dense) class of systems, and to prove uniform (maximal) inequalities. This strategy gives the almost sure invariance principle in the boundary case for GPM maps (see Theorem 1.6 below). In the case of uniformly expanding maps the almost sure invariance principle for a dense set of functions has been proved by Hofbauer and Keller (1982) for a smaller class than that given in Definition 1.1, and follows from Merlevède and Rio (2012) for the class of uniformly expanding maps considered in the present paper. However, the bounded law of the iterated logarithm for the boundary case is not available in the literature: we will prove it in Proposition 5.3.

We now turn to the functions for which we can prove the almost sure invariance principle. The main feature of our arguments is that they work with the weakest possible integrability condition (merely $\mathbb{L}^2(\nu)$ for uniformly expanding maps), and without any condition on the modulus of continuity: we only need the functions to be piecewise monotonic. More precisely, the results are mainly proved for functions which are monotonic on a single interval, and they are then extended by linearity to convex combinations of such functions. Such classes are described in the following definition.

Definition 1.3. *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 $\mathbb{L}^1(\mu)$ 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 $f_\ell \in \mathrm{Mon}_p(M, \mu)$.*

The above definition deals with $\mathbb{L}^p$-like spaces, with an additional monotonicity condition. In some cases, it is also important to deal with spaces similar to weak $\mathbb{L}^p$, where one only requires a uniform bound on the tails of the functions. Such spaces are described in the following definition.

Definition 1.4. *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 μ is a probability measure on $\mathbb{R}$ and H is a tail function, let $\mathrm{Mon}(H, \mu)$ denote the set of functions $f : \mathbb{R} \to \mathbb{R}$*

which are monotonic on some interval and null elsewhere and such that $\mu(|f| > t) \le H(t)$*. Let* $\mathrm{Mon}^c(H,\mu)$ *be the closure in* $\mathbb{L}^1(\mu)$ *of the set of functions which can be written as* $\sum_{\ell=1}^L a_\ell f_\ell$*, where* $\sum_{\ell=1}^L |a_\ell| \le 1$ *and* $f_\ell \in \mathrm{Mon}(H,\mu)$*.*

Our main theorems follow. For uniformly expanding maps, it involves an $\mathbb{L}^2$-integrability condition, while for GPM maps the boundary case is formulated in terms of tails.

Theorem 1.5. *Let* T *be a uniformly expanding map with absolutely continuous invariant measure* ν*. Then, for any* $M > 0$ *and any* $f \in \mathrm{Mon}_2^c(M,\nu)$*, the series*

$$\sigma^2 = \sigma^2(f) = \nu((f-\nu(f))^2) + 2\sum_{k>0} \nu((f-\nu(f))f\circ T^k) \tag{1.1}$$

converges absolutely to some nonnegative number. Moreover,

1. *On the probability space* $([0,1],\nu)$*, the process*

$$\Big\{\frac{1}{\sqrt{n}} \sum_{i=0}^{[(n-1)t]} (f\circ T^i - \nu(f)),\ t\in[0,1]\Big\}$$

converges in distribution in the Skorokhod topology to σW*, where* W *is a standard Wiener process.*

2. *There exists a nonnegative constant* A *such that*

$$\sum_{n=1}^{\infty} \frac{1}{n}\nu\Big(\max_{1\le k\le n}\Big|\sum_{i=0}^{k-1}(f\circ T^i - \nu(f))\Big| \ge A\sqrt{n\log\log n}\Big) < \infty\,.$$

3. *Enlarging* $([0,1],\nu)$ *if necessary, there exists a sequence* $(Z_i)_{i\ge0}$ *of i.i.d. Gaussian random variables with mean zero and variance* σ^2 *defined by* (1.1)*, such that*

$$\Big|\sum_{i=0}^{n-1}(f\circ T^i - \nu(f) - Z_i)\Big| = o(\sqrt{n\log\log n})\,, \text{ almost surely.} \tag{1.2}$$

Theorem 1.6. *Let* T *be a GPM map with parameter* $\gamma \in (0,1/2)$ *and invariant measure* ν*. Let* H *be a tail function with*

$$\int_0^\infty x(H(x))^{\frac{1-2\gamma}{1-\gamma}}\,dx < \infty\,. \tag{1.3}$$

Then, for any $f \in \mathrm{Mon}^c(H,\nu)$*, the series* σ^2 *defined in* (1.1) *converges absolutely to some nonnegative number, and the asymptotic results 1., 2. and 3. of Theorem 1.5 hold.*

In particular, it follows from Theorem 1.6 that, if T is a GPM map with parameter $\gamma \in (0,1/2)$, then the almost sure invariance principle (1.2) holds for any positive and nonincreasing function f on (0,1) such that

$$f(x) \leq \frac{C}{x^{(1-2\gamma)/2}|\ln(x)|^b} \quad \text{near 0, for some } b > 1/2.$$

Note that (1.2) cannot be true if f is exactly of the form $f(x) = x^{-(1-2\gamma)/2}$. Indeed, in that case, Gouëzel (2004) proved that the central limit theorem holds with the normalization $\sqrt{n\ln(n)}$, and the corresponding almost sure result is

$$\lim_{n\to 0} \frac{1}{\sqrt{n}(\ln(n))^b} \sum_{i=0}^{n-1} (f \circ T^i - \nu(f)) = 0 \quad \text{almost everywhere, for any } b > 1/2.$$

We refer to the paper by Dedecker, Gouëzel and Merlevède (2010) for a deeper discussion on the optimality of the conditions.

The plan of the paper is as follows. In Section 2, we explain how functions in $\mathrm{Mon}_p^c(M,\mu)$ or $\mathrm{Mon}^c(H,\mu)$ can be approximated by bounded variation functions (to which the results of Merlevède and Rio (2012) regarding the almost sure invariance principle apply). In Section 3, we show how an almost sure invariance principle for a sequence of approximating processes implies an almost sure invariance principle for a given process, if one also has uniform estimates (for instance, a bounded law of the iterated logarithm). Those two results together with the bounded law of the iterated logarithm of Dedecker, Gouëzel and Merlevède (2010) readily give the almost sure invariance principle in the boundary case for GPM maps, as we explain in Section 4. In Section 5, we prove a bounded law of the iterated logarithm under a polynomial assumption on mixing coefficients, and we use this estimate in Section 6 to obtain the almost sure invariance principle in the boundary case for uniformly expanding maps, following the same strategy as above.

2 Approximation by bounded variation functions

Let us define the variation $\|f\|_v$ of a function $f : \mathbb{R} \to \mathbb{R}$ as the supremum of the quantities $|f(a_0)| + \sum_{i=0}^{k-1} |f(a_{i+1} - f(a_i)| + |f(a_k)|$ over all finite sequences $a_0 < \cdots < a_k$. A function f has bounded variation if $\|f\|_v < \infty$.

In this section, we want to approximate a function in $\mathrm{Mon}_2^c(M,\mu)$ or $\mathrm{Mon}^c(H,\mu)$ in a suitable way. For $\mathrm{Mon}^c(H,\mu)$, we shall use the following compactness lemma. It is mainly classical (compare for instance Hofbauer and Keller (1982) Lemma 5), but since we have not been able to locate a reference with this precise statement we will give a complete proof.

Lemma 2.1. *Let μ be a probability measure on $\mathbb{R}$. Let f_n be a sequence of functions on $\mathbb{R}$ with $\|f_n\|_v \leq C$. Then there exists $f : \mathbb{R} \to \mathbb{R}$ with $\|f\|_v \leq C$ such that a subsequence $f_{\varphi(n)}$ tends to f in $\mathbb{L}^1(\mu)$.*

Proof. We will first prove that f_n admits a convergent subsequence in $\mathbb{L}^1(\mu)$. By a classical diagonal argument, it suffices to show that, for any $\epsilon > 0$, one can find a subsequence with $\limsup_{n\to\infty} \sup_{m\geq n} \left\|f_{\varphi(n)} - f_{\varphi(m)}\right\|_{\mathbb{L}^1(\mu)} \leq D\epsilon$, for some $D > 0$ not dependending on ϵ.

We consider a finite number of points $a_0 < \cdots < a_k$ such that (letting $a_{-1} = -\infty$ and $a_{n+1} = +\infty$), the measure of every interval (a_i, a_{i+1}) is at most ϵ. One can find a subsequence of f_n such that each $f_{\varphi(n)}(a_i)$ converges, we claim that it satisfies the desired property. It suffices to show that a function g with $|g(a_i)| \leq \epsilon$ for all i and $\|g\|_v \leq 2C$ satisfies

$$\|g\|_{\mathbb{L}^1(\mu)} \leq D\epsilon. \tag{2.1}$$

Consider in each interval (a_i, a_{i+1}) a point b_i such that $\sup_{(a_i,a_{i+1})} |g| \leq 2|g(b_i)|$. We have

$$\begin{aligned}\|g\|_{\mathbb{L}^1(\mu)} &\leq \sum \mu(a_i, a_{i+1}) \sup_{(a_i,a_{i+1})} |g| + \sum \mu\{a_i\}|g(a_i)| \\ &\leq 2\sum \mu(a_i, a_{i+1})(|g(b_i) - g(a_i)| + |g(a_i)|) + \sum \mu\{a_i\}|g(a_i)|.\end{aligned}$$

Since $|g(a_i)| \leq \epsilon$ and μ is a probability measure, the contribution of the terms $|g(a_i)|$ to this expression is at most 2ϵ. Moreover, $\sum \mu(a_i, a_{i+1})|g(b_i) - g(a_i)| \leq \epsilon \sum |g(b_i) - g(a_i)| \leq \epsilon \|g\|_v$. This proves (2.1).

We have proved that f_n admits a subsequence (that we still denote f_n) that converges in $\mathbb{L}^1(\mu)$ to a function f. Extracting further if necessary, we may also assume that it converges to f on a set Ω with full measure. On $\overline{\Omega} - \Omega$, we define $f(x)$ to be $\limsup f(y)$ where y tends to x in Ω. Finally, on the open set $\mathbb{R} - \overline{\Omega}$ (which may be nonempty if μ does not have full support), we define $f(x)$ to be $\max(f(a), f(b))$ where a and b are the endpoints of the connected component of x in $\mathbb{R} - \overline{\Omega}$ (if one of those endpoints is $-\infty$ or $+\infty$, we only use the other endpoint). Then f_n converges to f in $\mathbb{L}^1(\mu)$, and we claim that f has variation at most C.

Indeed, consider a sequence $a_0 < \cdots < a_k$, we want to estimate $|f(a_0)| + \sum |f(a_{i+1}) - f(a_i)| + |f(a_k)|$. Let $b_i = a_i$ if $a_i \in \Omega$. By construction of f, for all $a_i \notin \Omega$, one may find a point b_i in Ω such that $|f(a_i) - f(b_i)|$ is small, say $< \epsilon/(k+1)$, and we may ensure that $b_0 \leq \cdots \leq b_k$. Then

$$\begin{aligned}|f(a_0)| + \sum |f(a_{i+1}) - f(a_i)| + |f(a_n)| &\leq 4\epsilon + |f(b_0)| + \sum |f(b_{i+1}) - f(b_i)| + |f(b_k)| \\ &= 4\epsilon + \lim\Big(|f_n(b_0)| + \sum |f_n(b_{i+1}) - f_n(b_i)| + |f_n(b_k)|\Big).\end{aligned}$$

Since the variation of f_n is at most C, this is bounded by $4\epsilon + C$. Letting ϵ tend to 0, we get $\|f\|_v \leq C$. □

Lemma 2.2. *Let H be a tail function, and consider $f \in \mathrm{Mon}^c(H, \mu)$. For any $m > 0$, one can write $f = \bar{f}_m + g_m$ where $\bar{f}_m$ has bounded variation and $g_m \in \mathrm{Mon}^c(H_m, \mu)$ where $H_m(x) = \min(H(m), H(x))$.*

Proof. Consider $f \in \mathrm{Mon}^c(H, \mu)$. By definition, there exists a sequence of functions $f_L =$

$\sum_{\ell=1}^{L} a_{\ell,L} g_{\ell,L}$ with $g_{\ell,L}$ belonging to $\mathrm{Mon}(H,\mu)$ and $\sum_{\ell=1}^{L} |a_{\ell,L}| \leq 1$, such that f_L converges in $\mathbb{L}^1(\mu)$ to f. Define then

$$f_{L,m} = \sum_{\ell=1}^{L} a_{\ell,L} g_{\ell,L} \mathbf{1}_{|g_{\ell,L}| \leq m} .$$

Note that $f_{L,m}$ is such that $\|f_{L,m}\|_v \leq 3m$. Applying Lemma 2.1, there exists a subsequence $f_{\varphi(L),m}$ converging in $\mathbb{L}^1(\mu)$ to a limit $\bar{f}_m$ such that $\|\bar{f}_m\|_v \leq 3m$. Hence $f - \bar{f}_m$ is the limit in $\mathbb{L}^1(\mu)$ of

$$f_{\varphi(L)} - f_{\varphi(L),m} = \sum_{\ell=1}^{\varphi(L)} a_{\ell,\varphi(L)} g_{\ell,\varphi(L)} \mathbf{1}_{|g_{\ell,\varphi(L)}| > m} .$$

Now $g_{\ell,\varphi(L)} \mathbf{1}_{|g_{\ell,\varphi(L)}|>m}$ belongs to $\mathrm{Mon}(\min(H(m), H), \mu)$. It follows that $f - \bar{f}_m$ belongs to the class $\mathrm{Mon}^c(H_m, \mu)$. □

A similar result holds for the space $\mathrm{Mon}_2^c(M,\mu)$:

Lemma 2.3. *Consider $f \in \mathrm{Mon}_2^c(M,\mu)$. For any $m > 0$, one can write $f = \bar{f}_m + g_m$, where $\bar{f}_m$ has bounded variation and $g_m \in \mathrm{Mon}_2^c(1/m,\mu)$.*

The above proof does not work to obtain this result (the problem is that the function $g_\ell \mathbf{1}_{|g_\ell|>m}$ usually does not satisfy better $\mathbb{L}^2$ bounds than the function g_ℓ, at least not uniformly in g_ℓ). To prove this lemma, we will therefore need to understand more precisely the structure of elements of $\mathrm{Mon}_2^c(M,\mu)$. We will show that they are extended convex combinations of elements of $\mathrm{Mon}_2(M,\mu)$, i.e., they can be written as $\int g d\beta(g)$ for some probability measure β on $\mathrm{Mon}_2(M,\mu)$ (the case $\sum a_\ell g_\ell$ corresponds to the case where β is an atomic measure).

To justify this assertion, the first step is to be able to speak of measures on $\mathrm{Mon}_2(M,\mu)$. We need to specify a topology on $\mathrm{Mon}_2(M,\mu)$. We use the weak topology (inherited from the space $\mathbb{L}^2(\mu)$, that contains $\mathrm{Mon}_2(M,\mu)$): a sequence $f_n \in \mathrm{Mon}_2(M,\mu)$ converges to f if, for any continuous compactly supported function $u : \mathbb{R} \to \mathbb{R}$ (or, equivalently, for any $\mathbb{L}^2(\mu)$ function u), $\int f_n(x)u(x)d\mu(x) \to \int f(x)u(x)d\mu(x)$.

Lemma 2.4. *The space $\mathrm{Mon}_2(M,\mu)$, with the topology of weak convergence, is a compact metrizable space.*

Proof. Consider a countable sequence of continuous compactly supported functions $u_k : \mathbb{R} \to \mathbb{R}$, which is dense in this space for the topology of uniform convergence. We define a distance on $\mathbb{L}^2(\mu)$ by

$$d(f_1, f_2) = \sum 2^{-k} \min\left(1, \left|\int (f_1 - f_2) u_k d\mu\right|\right) .$$

Convergence for this distance is clearly equivalent to weak convergence.

Let us now prove that $\mathrm{Mon}_2(M,\mu)$ is compact. Consider a sequence f_n in this space. In particular, it is bounded in $\mathbb{L}^2(\mu)$. By weak compactness of the unit ball of a Hilbert space, we can find a subsequence (still denoted by f_n) which converges weakly in $\mathbb{L}^2(\mu)$, to a function f.

In particular, $\int f_n u d\mu$ converges to $\int f u d\mu$ for any continuous compactly supported function u. Moreover, f is bounded by M in $\mathbb{L}^2(\mu)$. To conclude, it remains to show that f has a version which is monotonic on an interval, and vanishes elsewhere.

A function in $\text{Mon}_2(M, \mu)$ can be either nonincreasing or nondecreasing, on an interval which is half-open or half-closed to the left and to the right, there are therefore eight possible combinatorial types. Extracting a further subsequence if necessary, we may assume that all the functions f_n have the same combinatorial type. For simplicity, we will describe what happens for one of those types, the other ones are handled similarly. We will assume that all the functions f_n are nondecreasing on an interval $(a_n, b_n]$. We may also assume that a_n and b_n are either constant, or increasing, or decreasing (since any sequence in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ admits a subsequence with this property). In particular, those sequences converge in $\overline{\mathbb{R}}$ to limits a and b. Let I be the interval with endpoints a and b, where we include a in I if a_n is increasing and exclude it otherwise, and where we include b if b_n is decreasing or constant and exclude it otherwise. The Banach-Saks theorem shows that (extracting further if necessary) we may ensure that the sequence of functions $g_N = \frac{1}{N} \sum_{n=1}^{N} f_n$ converges to f in $\mathbb{L}^2(\mu)$ and on a set A of full measure. It readily follows that f is nondecreasing on $A \cap I$ and vanishes on $A \cap (\mathbb{R} - I)$. Modifying f on the zero measure set $\mathbb{R} - A$, we get a function in $\text{Mon}_2(M, \mu)$ as claimed. □

The Borel structure coming from the weak topology on $\mathbb{L}^2(\mu)$ coincides with the Borel structure coming from the norm topology (since an open ball for the norm topology can be written as a countable intersection of open sets for the weak topology, by the Hahn-Banach theorem). Therefore, all the usual functions on $\text{Mon}_2(M, \mu)$ are measurable.

If β is a probability measure on $\text{Mon}_2(M, \mu)$, we can define a function $f \in \mathbb{L}^2(\mu)$ by $f(x) = \int g(x) d\beta(g)$. We claim that the elements of $\text{Mon}_2^c(M, \mu)$ are exactly such functions:

Proposition 2.5. *We have*

$$\text{Mon}_2^c(M, \mu) = \left\{ \int_{\text{Mon}_2(M,\mu)} g d\beta(g) \,:\, \beta \textit{ probability measure on } \text{Mon}_2(M, \mu) \right\}.$$

Proof. We have two inclusions to prove.

Consider first $f \in \text{Mon}_2^c(M, \mu)$, we will show that it can be written as $\int g d\beta(g)$ for some measure β. By definition of $\text{Mon}_2^c(M, \mu)$, there exists a sequence of atomic probability measures β_n on $\text{Mon}_2(M, \mu)$ such that $f_n = \int g d\beta_n(g)$ converges in $\mathbb{L}^1(\mu)$ to f. Since the space $\text{Mon}_2(M, \mu)$ is compact, the sequence of measures β_n admits a convergent subsequence (that we still denote by β_n), to a measure β. By definition of vague convergence, for any continuous function Ψ on $\text{Mon}_2(M, \mu)$, $\int \Psi(g) d\beta_n(g)$ tends to $\int \Psi(g) d\beta(g)$. Fix a continuous compactly supported function u on $\mathbb{R}$. By definition of the topology on $\text{Mon}_2(M, \mu)$, the map $\Psi_u : g \mapsto \int u(x) g(x) d\mu(x)$ is continuous. Therefore, $\int \Psi_u(g) d\beta_n(g)$ tends to $\int \Psi_u(g) d\beta(g)$, i.e., $\int u(x) f_n(x) d\mu(x)$ tends to $\int u(x) f_\beta(x) d\mu(x)$, where $f_\beta = \int g d\beta(g)$. This shows that f_n converges weakly to f_β. However, by assumption, f_n converges in $\mathbb{L}^1(\mu)$ to f. We deduce that $f = f_\beta$, as desired.

Conversely, consider a function f_β for some probability measure β on $\mathrm{Mon}_2(M,\mu)$, let us show that it belongs to $\mathrm{Mon}_2^c(M,\mu)$. Let us consider a sequence of atomic probability measures β_n converging vaguely to β. The arguments in the previous paragraph show that the functions f_{β_n} converge weakly to f_β. By Banach-Saks theorem, extracting a subsequence if necessary, we can ensure that $f_N = N^{-1}\sum_{n=1}^N f_{\beta_n}$ converges almost everywhere and in $\mathbb{L}^2(\mu)$ to f_β. In particular, it converges to f_β in $\mathbb{L}^1(\mu)$. Since f_N can be written as $\sum a_{\ell,N} f_{\ell,N}$ for some functions $f_{\ell,N} \in \mathrm{Mon}_2(M,\mu)$ and some coefficients $a_{\ell,N}$ with sum bounded by 1, this shows that f_β belongs to $\mathrm{Mon}_2^c(M,\mu)$. □

Proof of Lemma 2.3. Consider $f \in \mathrm{Mon}_2^c(M,\mu)$, and $\epsilon > 0$. By Proposition 2.5, there exists a measure β on $\mathrm{Mon}_2(M,\mu)$ such that $f = \int g \, d\beta(g)$. For each $g \in \mathrm{Mon}_2(M,\mu)$, let $K(g)$ be the smallest number such that $\int g^2 1_{|g|\geq K(g)} \leq \epsilon^2$. Fix some $K > 0$. We have

$$\begin{aligned} f(x) &= \int_{K(g)<K} g(x)d\beta(g) + \int_{K(g)\geq K} g(x)d\beta(g) \\ &= \int_{K(g)<K} g(x)\mathbf{1}_{|g(x)|\leq K(g)}d\beta(g) + \int_{K(g)<K} g(x)\mathbf{1}_{|g(x)|>K(g)}d\beta(g) + \int_{K(g)\geq K} g(x)d\beta(g). \end{aligned}$$

The first term has variation bounded by $3K$. In the second term, each function $g\mathbf{1}_{|g|>K(g)}$ is monotonic on an interval and null elsewhere, with $\mathbb{L}^2(\mu)$ norm bounded by ϵ. Therefore, the second term belongs to $\mathrm{Mon}_2^c(\epsilon,\mu)$. Writing $A(K) = \{g : K(g) \geq K\}$ and $\alpha(K) = \beta(A(K))$, the third term is the average over $A(K)$ of the functions $\alpha(K)g \in \mathrm{Mon}_2(\alpha(K)M,\mu)$ with respect to the probability measure $\mathbf{1}_{A(K)}d\beta(g)/\alpha(K)$. Therefore, it belongs to $\mathrm{Mon}_2^c(\alpha(K)M,\mu)$. Taking K large enough so that $\alpha(K)M \leq \epsilon$, we infer that f is the sum of a function of bounded variation and a function in $\mathrm{Mon}_2^c(2\epsilon,\mu)$. □

3 Strong invariance principle by approximation

Let $(X_i)_{i\geq 1}$ be a sequence of random variables. Assume that

1. For each $m \in \mathbb{N}$ there exists a sequence $(X_{i,m})_{i\geq 1}$ such that

$$\limsup_{n\to\infty} \left| \frac{\sum_{i=1}^n X_i - X_{i,m}}{\sqrt{n \log\log n}} \right| \leq \epsilon(m) \quad \text{almost surely,}$$

   where $\epsilon(m)$ tends to 0 as m tends to infinity.

2. For each $m \in \mathbb{N}$, the sequence $(X_{i,m})_{i\geq 1}$ satisfies a strong invariance principle: there exists a sequence $(Z_{i,m})_{i\geq 1}$ of i.i.d. Gaussian random variables with mean 0 and variance σ_m^2 such that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n X_{i,m} - Z_{i,m}}{\sqrt{n \log\log n}} = 0 \quad \text{almost surely.}$$

   We also assume that σ_m^2 converges as $m \to \infty$ to a limit σ^2.

3. There exists an infinite subset $\mathcal{A}$ of $\mathbb{N}$ such that, for any $A \in \mathcal{A}$, the σ-algebras $\sigma(Z_{i,m})_{i<A,\, m\in\mathbb{N}}$ and $\sigma(Z_{i,m})_{i\geq A,\, m\in\mathbb{N}}$ are independent.

Proposition 3.1. *Under the assumptions 1, 2 and 3, there exists a sequence $(Z_i)_{i\geq 1}$ of i.i.d. Gaussian random variables with mean zero and variance σ^2 such that*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n X_i - Z_i}{\sqrt{n \log\log n}} = 0 \quad \textit{almost surely.} \tag{3.1}$$

Proof. The idea of the proof is to use a diagonal argument: we will use the $Z_{i,0}$ for some time, then the $Z_{i,1}$ for a longer time, and so on, to construct the Z_i.

Let A_m be a sequence of elements of $\mathcal{A}$ tending to infinity fast enough. More precisely, we choose A_m in such a way that there exists a set Ω_m with probability greater than $1 - 2^{-m}$ on which, for any $n \geq A_m$,

$$\left| \frac{\sum_{i=1}^n X_{i,m} - Z_{i,m}}{\sqrt{n \log\log n}} \right| \leq \epsilon(m) \quad \text{and} \quad \left| \frac{\sum_{i=1}^n X_i - X_{i,m}}{\sqrt{n \log\log n}} \right| \leq 2\epsilon(m).$$

The assumptions 1 and 2 ensure that these two properties are satisfied provided A_m is large enough. We also choose A_m in a such a way that, for $j < m - 1$,

$$\epsilon(j)\sqrt{A_{j+1} \log\log A_{j+1}} < 2^{-(m-j)}\epsilon(m)\sqrt{A_m \log\log A_m}. \tag{3.2}$$

Indeed, if the A_j's have been defined for $j < m$, it suffices to take A_m large enough for (3.2) to hold.

With this choice of A_m, we infer that for any $\omega \in \Omega_m$ and any $n \geq A_m$,

$$\left| \sum_{i=1}^n X_i - Z_{i,m} \right| \leq 3\epsilon(m)\sqrt{n \log\log n}.$$

Hence, for any $\omega \in \Omega_m$ and any $n \geq A_m$,

$$\left| \sum_{i=A_m}^n X_i - Z_{i,m} \right| \leq 6\epsilon(m)\sqrt{n \log\log n}. \tag{3.3}$$

For $i \in [A_m, A_{m+1} - 1]$, let $m(i) = m$. Let $(\delta_k)_{k\geq 1}$ be a sequence of i.i.d. Gaussian random variables with mean zero and variance σ^2, independent of the array $(Z_{i,m})_{i\geq 1, m\geq 1}$. We now construct the sequence Z_i as follows: if $\sigma_{m(i)} = 0$, then $Z_i = \delta_i$, else $Z_i = (\sigma/\sigma_{m(i)})Z_{i,m(i)}$. By construction, thanks to the assumption 3, the Z_i's are i.i.d. Gaussian random variables with mean zero and variance σ^2. Let us show that they satisfy (3.1).

Let $D_i = Z_i - Z_{i,m(i)}$ and note that $(D_i)_{i\geq 1}$ is a sequence of independent Gaussian random variables with mean zero and variances $\mathrm{Var}(D_i) = (\sigma - \sigma_{m(i)})^2$. Since $\sigma_{m(i)}$ converges to σ as i

tends to infinity, it follows that

$$\text{letting} \quad v_n = \frac{1}{n} \operatorname{Var}\Big(\sum_{i=1}^{n} D_i\Big), \quad \text{then} \quad \lim_{n\to\infty} v_n = 0.$$

From the basic inequality

$$\mathbb{P}\Big(\max_{1\le k\le n}\Big|\sum_{i=1}^{k} D_i\Big| > x\Big) \le 2\exp\Big(-\frac{x^2}{2nv_n}\Big),$$

it follows that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} Z_{i,m(i)} - Z_i}{\sqrt{n\log\log n}} = 0 \quad \text{almost surely.}$$

To conclude the proof, it remains to prove that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i - Z_{i,m(i)}}{\sqrt{n\log\log n}} = 0 \quad \text{almost surely.} \tag{3.4}$$

Let $B = \{\omega : \omega \in \liminf \Omega_m\}$. By Borel-Cantelli, $\mathbb{P}(B) = 1$. For $\omega \in B$, there exists $m_0(\omega)$ such that ω belongs to all the Ω_m for $m \ge m_0(\omega)$. For $n \ge A_{m_0(\omega)}$, we have (denoting by M the greater integer such that $A_M \le n$)

$$\Big|\sum_{i=1}^{n} X_i - Z_{i,m(i)}\Big| \le \sum_{i=1}^{A_{m_0(\omega)}-1} |X_i - Z_{i,m(i)}| + \sum_{m=m_0(\omega)}^{M-1} \Big|\sum_{i=A_m}^{A_{m+1}-1} X_i - Z_{i,m}\Big| + \Big|\sum_{i=A_M}^{n} X_i - Z_{i,M}\Big|.$$

Taking into account (3.2) and (3.3), we obtain

$$\begin{aligned}\Big|\sum_{i=1}^{n} X_i - Z_{i,m(i)}\Big| &\le C(\omega) + \sum_{m=1}^{M-1} 6\epsilon(m)\sqrt{A_{m+1}\log\log A_{m+1}} + 6\epsilon(M)\sqrt{n\log\log n} \\ &\le C(\omega) + \sum_{m=1}^{M-2} 6\epsilon(M)\sqrt{A_M\log\log A_M}2^{-(M-m)} \\ &\quad + 6\epsilon(M-1)\sqrt{A_M\log\log A_M} + 6\epsilon(M)\sqrt{n\log\log n} \\ &\le C(\omega) + 9(\epsilon(M-1)+\epsilon(M))\sqrt{n\log\log n}.\end{aligned}$$

Since $\epsilon(M-1)+\epsilon(M)$ tends to zero as n tends to infinity, this proves (3.4) and completes the proof of Proposition 3.1. □

Remark 3.2. *The proposition would also apply to random variables taking values in $\mathbb{R}^d$ or in Banach spaces (with the same proof), but we have formulated it only for real-valued random variables in view of our applications. Indeed, the class of functions we consider relies on monotonicity which is a purely one-dimensional notion.*

4 Proof of Theorem 1.6 on GPM maps

To prove Theorem 1.6, we should establish the convergence of the series (1.1) as well as the asymptotic results 1., 2. and 3. described in Theorem 1.5. The convergence of (1.1) and the asymptotics 1. and 2. have been proved in Dedecker, Gouëzel and Merlevède (2010). Therefore it only remains to prove the almost sure invariance principle.

To do this, we apply Proposition 3.1 to the sequences $X_i = f \circ T^i - \nu(f)$ and $X_{i,m} = \bar{f}_m \circ T^i - \nu(\bar{f}_m)$, where the function $\bar{f}_m$ has been constructed in Lemma 2.2. Let us denote by $S_n(f) = \sum_{i=0}^{n-1}(f \circ T^i - \nu(f))$. To apply Proposition 3.1, we have to check the assumptions 1., 2. and 3. of Section 3.

The function $g_m = f - \bar{f}_m$ belongs to $\text{Mon}^c(H_m, \nu)$ where $H_m = \min(H(m), H)$, by Lemma 2.2. Therefore, it belongs to the class of functions to which the results of Dedecker, Gouëzel and Merlevède (2010) apply: $S_n(g_m)$ satisfies a central limit theorem and a bounded law of the iterated logarithm. In particular, applying Theorem 1.5 of this article (and Section 4.5 there to compute the constant $M(m)$) we get that, almost surely,

$$\limsup \frac{1}{\sqrt{n \log\log n}} \left| \sum_{i=0}^{n-1} (g_m \circ T^i - \nu(g_m)) \right| \leq M(m),$$

where $M(m) = C \int_0^\infty x(H_m(x))^{\frac{1-2\gamma}{1-\gamma}} dx$, C being some positive constant. Since $M(m)$ tends to zero as m tends to infinity, the assumption 1. of Section 3 follows by choosing $\epsilon(m) = 2M(m)$.

Since the function $\bar{f}_m$ has bounded variation we can apply Item 2 of Theorem 3.1 of Merlevède and Rio (2012) to the sequence $(X_{i,m})$ (see their Remark 3.1 for the case of GPM maps). Hence there exists a sequence $(Z_{i,m})_{i \geq 1}$ of i.i.d. Gaussian random variables with mean 0 and variance $\sigma_m^2 = \sigma^2(\bar{f}_m)$ such that

$$\lim_{n \to \infty} \frac{\sum_{i=1}^n X_{i,m} - Z_{i,m}}{\sqrt{n \log\log n}} = 0 \quad \text{almost surely.}$$

More precisely, it follows from their construction (see the definition of the variables $V_{k,L}^*$ in Section 4.2 of Merlevède and Rio (2010)) that the assumption 3. of Section 3 is satisfied with $\mathcal{A} = \{2^L, L \in \mathbb{N}^*\}$.

To check the assumption 2. of Section 3, it remains only to prove that σ_m^2 converges to σ^2 as m tends to infinity. We have $f = \bar{f}_m + g_m$, therefore

$$\frac{S_n(f)}{\sqrt{n}} = \frac{S_n(\bar{f}_m)}{\sqrt{n}} + \frac{S_n(g_m)}{\sqrt{n}}.$$

The term on the left converges in distribution to a Gaussian with variance σ^2, and the terms on the right converge to

(non-independent) Gaussians with respective variances σ_m^2 and $\sigma^2(g_m)$. To conclude, it suffices to show that $\sigma^2(g_m)$ converges to 0 when m tends to infinity.

As we have explained above, the results of Dedecker, Gouëzel and Merlevède (2010) apply, and show that $S_n(g_m)$ satisfies a central limit theorem. From the same paper (see Sections 2.2 and 4.1 there), we get the following estimate on the asymptotic variance $\sigma^2(g_m)$ of $n^{-1/2}S_n(g_m)$: there exists a positive constant C such that

$$\sigma^2(g_m) \le C \int_0^\infty x(H_m(x))^{\frac{1-2\gamma}{1-\gamma}}\, dx \,,$$

and the second term on right hand tends to zero as m tends to infinity by using (1.3) and the dominated convergence theorem. The result follows.

Hence, we have checked that the assumptions 1., 2. and 3. of Section 3 are satisfied. This completes the proof of the almost sure invariance principle. □

5 A bounded LIL for ϕ-dependent sequences

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and let $\theta : \Omega \mapsto \Omega$ be a bijective bimeasurable transformation preserving the probability $\mathbb{P}$. Let $\mathcal{F}_0$ be a sub-σ-algebra of $\mathcal{A}$ satisfying $\mathcal{F}_0 \subseteq \theta^{-1}(\mathcal{F}_0)$.

Definition 5.1. *For any integrable random variable X, let us write $X^{(0)} = X - \mathbb{E}(X)$. For any random variable $Y = (Y_1, \cdots, Y_k)$ with values in $\mathbb{R}^k$ and any σ-algebra $\mathcal{F}$, let*

$$\phi(\mathcal{F}, Y) = \sup_{(x_1,\ldots,x_k)\in\mathbb{R}^k} \left\| \mathbb{E}\Big(\prod_{j=1}^k (\mathbb{1}_{Y_j \le x_j})^{(0)} \Big| \mathcal{F}\Big)^{(0)} \right\|_\infty .$$

For a sequence $\mathbf{Y} = (Y_i)_{i\in\mathbb{Z}}$, where $Y_i = Y_0 \circ \theta^i$ and Y_0 is an $\mathcal{F}_0$-measurable and real-valued random variable, let

$$\phi_{k,\mathbf{Y}}(n) = \max_{1\le l\le k} \sup_{n\le i_1\le \ldots \le i_l} \phi(\mathcal{F}_0, (Y_{i_1}, \ldots, Y_{i_l})).$$

The interest of those mixing coefficients is that they are not too restrictive, so they can be used to study several classes of dynamical systems, and that on the other hand they are strong enough to yield correlation bounds for piecewise monotonic functions (or, more generally, functions in $\mathrm{Mon}_p^c(M, \mu)$). In particular, we have the following:

Lemma 5.2. *Let $\mathbf{Y} = (Y_i)_{i\in\mathbb{Z}}$, where $Y_i = Y_0 \circ \theta^i$ and Y_0 is an $\mathcal{F}_0$-measurable random variable. 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 $\|f(Y_0)\|_p < \infty$, then, for any positive integer k,*

$$\|\mathbb{E}(f(Y_k)|\mathcal{F}_0) - \mathbb{E}(f(Y_k))\|_p \le 2(2\phi_{1,\mathbf{Y}}(k))^{(p-1)/p}\|f(Y_0)\|_p \,.$$

If moreover $p \ge 2$ and $\|g(Y_0)\|_p < \infty$, then for any positive integers $i \ge j \ge k$,

$$\|\mathbb{E}(f(Y_i)^{(0)}g(Y_j)^{(0)}|\mathcal{F}_0) - \mathbb{E}(f(Y_i)^{(0)}g(Y_j)^{(0)})\|_{p/2} \le 8(4\phi_{2,\mathbf{Y}}(k))^{(p-2)/p}\|f(Y_0)\|_p\|g(Y_0)\|_p \,.$$

Proof. Note first that, for any positive integers $i \geq j \geq k$,

$$\phi(\mathcal{F}_0, f(Y_k)) \leq 2\phi(\mathcal{F}_0, Y_k) \leq 2\phi_{1,\mathbf{Y}}(k),$$
$$\phi(\mathcal{F}_0, (f(Y_j), g(Y_i))) \leq 4\phi(\mathcal{F}_0, (Y_j, Y_i)) \leq 4\phi_{2,\mathbf{Y}}(k)\,.$$

This follows from definition (5.1), by noting that $\{f \leq t\}$ (and also $\{g \leq s\}$) is either an interval or the complement of an interval.

To prove the first inequality of the lemma, let us note that

$$\|\mathbb{E}(f(Y_k)|\mathcal{F}_0) - \mathbb{E}(f(Y_k))\|_p = \sup_{Z \in B_{p/(p-1)}(\mathcal{F}_0)} \mathrm{Cov}(Z, f(Y_k))\,,$$

where $B_q(\mathcal{F}_0)$ is the set of $\mathcal{F}_0$-measurable random variables Z such that $\|Z\|_q \leq 1$. Proposition 2.1 of Dedecker (2004) states that $|\mathrm{Cov}(Z,Y)| \leq 2\phi(\sigma(Z), Y)^{(p-1)/p}\|Y\|_p\|Z\|_{p/(p-1)}$. Since $\phi(\sigma(Z), f(Y_k)) \leq \phi(\mathcal{F}_0, f(Y_k)) \leq 2\phi_{1,\mathbf{Y}}(k)$, we obtain the first inequality of Lemma 5.2 as desired.

For the second inequality, we note in the same way that

$$\|\mathbb{E}(f(Y_i)^{(0)}g(Y_j)^{(0)}|\mathcal{F}_0) - \mathbb{E}(f(Y_i)^{(0)}g(Y_j)^{(0)})\|_{p/2} = \sup_{Z \in B_{p/(p-2)}(\mathcal{F}_0)} \mathrm{Cov}(Z, f(Y_i)^{(0)}g(Y_j)^{(0)})\,.$$

Proposition 6.1 of Dedecker, Merlevède and Rio (2009) gives a control of the covariance in terms of $\phi(\mathcal{F}_0, (f(Y_j), g(Y_i)))$. Since this quantity is bounded by $4\phi_{2,\mathbf{Y}}(k)$, the result follows. □

The main result of this section is the following proposition, showing that a suitable polynomial assumption on mixing coefficients implies a bounded law of the iterated logarithm for piecewise monotonic $\mathbb{L}^2$ functions.

Proposition 5.3. *Let $X_i = f(Y_i) - \mathbb{E}(f(Y_i))$, where $Y_i = Y_0 \circ \theta^i$ and Y_0 is an $\mathcal{F}_0$-measurable random variable. Let*

$$S_n = S_n(f) = \sum_{k=1}^{n} X_k\,,$$

and let P_{Y_0} be the distribution of Y_0. Assume that

$$\sum_{k\geq 1} k^{1/\sqrt{3}-1/2}\phi_{2,\mathbf{Y}}^{1/2}(k) < \infty\,. \tag{5.1}$$

If f belongs to $\mathrm{Mon}_2^c(M, P_{Y_0})$ for some $M > 0$, then

$$\sum_{n>0} \frac{1}{n}\mathbb{P}\Big(\max_{1\leq k\leq n} |S_k| > 3CM\sqrt{n\log\log n}\Big) < \infty\,, \tag{5.2}$$

where $C = 16\sum_{k\geq 0}\phi_{1,\mathbf{Y}}^{1/2}(k)$.

Proof. Let $f \in \mathrm{Mon}_2^c(M, P_{Y_0})$. By definition of $\mathrm{Mon}_2^c(M, P_{Y_0})$, there exists $f_L = \sum_{\ell=1}^L a_{\ell,L} g_{\ell,L}$ with $g_{\ell,L}$ belonging to $\mathrm{Mon}_2(M, P_{Y_0})$ and $\sum_{\ell=1}^L |a_{\ell,L}| \leq 1$, and such that f_L converges in $\mathbb{L}^1(P_{Y_0})$ to f. It follows that $X_{i,L} = f_L(Y_i) - \mathbb{E}(f_L(Y_i))$ converges in $\mathbb{L}^1$ to X_i as L tends to infinity. Extracting a subsequence if necessary, one may also assume that the convergence holds almost surely.

Hence, for any fixed n, $S_n(f_L) = \sum_{k=1}^n X_{k,L}$ converges almost surely and in $\mathbb{L}^1$ to $S_n(f)$. Assume that one can prove that, for any positive integer L,

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} |S_k(f_L)| > 3CM\sqrt{n \log\log n}\Big) < K\,, \tag{5.3}$$

for some positive constant K not depending on L. Let us explain why (5.3) implies (5.2). Let $Z_n = \max_{1\leq k\leq n} |S_k(f)|/\sqrt{M^2 n \log\log n}$. By Beppo-Levi,

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} |S_k(f)| > 3CM\sqrt{n \log\log n}\Big) = \lim_{k\to\infty} \mathbb{E}\Big(\sum_{n>0} \frac{1}{n} \mathbf{1}_{Z_n > 3C + k^{-1}}\Big). \tag{5.4}$$

Let h_k be a continuous function from $\mathbb{R}$ to $[0, 1]$, such that $h_k(x) = 1$ if $x > 3C + k^{-1}$ and $h_k(x) = 0$ if $x < 3C$. Let $Z_{n,L} = \max_{1\leq k\leq n} |S_{k,L}|/\sqrt{M^2 n \log\log n}$. By Fatou's lemma,

$$\begin{aligned}\mathbb{E}\Big(\sum_{n>0} \frac{1}{n} \mathbf{1}_{Z_n > 3C+k^{-1}}\Big) &\leq \mathbb{E}\Big(\sum_{n>0} \frac{1}{n} h_k(Z_n)\Big) \\ &\leq \liminf_{L\to\infty} \mathbb{E}\Big(\sum_{n>0} \frac{1}{n} h_k(Z_{n,L})\Big) \leq \liminf_{L\to\infty} \mathbb{E}\Big(\sum_{n>0} \frac{1}{n} \mathbf{1}_{Z_{n,L} > 3C}\Big). \end{aligned} \tag{5.5}$$

From (5.3), (5.4) and (5.5), we infer that

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} |S_k(f)| > 3C\sqrt{M(f) n \log\log n}\Big) \leq \liminf_{L\to\infty} \mathbb{E}\Big(\sum_{n>0} \frac{1}{n} \mathbf{1}_{Z_{n,L} > 3C}\Big) \leq K\,,$$

and (5.2) follows.

Hence, it remains to prove (5.3), or more generally that: if $f = \sum_{\ell=1}^L a_\ell f_\ell$ with f_ℓ belonging to $\mathrm{Mon}_2(M, P_{Y_0})$ and $\sum_{\ell=1}^L |a_\ell| \leq 1$, then

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} |S_k(f)| > 3CM\sqrt{n \log\log n}\Big) < K\,, \tag{5.6}$$

for some positive constant K not depending on f.

We now prove (5.6). We will need to truncate the functions. It turns out that the optimal truncation level is at $\sqrt{n}/\sqrt{\log\log n}$: the large part can then be controlled by a simple $\mathbb{L}^1$ estimate, while the truncated part can be estimated thanks to a maximal inequality of Pinelis (1994) (after a reduction to a martingale). Let $g_n(x) = x \mathbf{1}_{|x| \leq M n^{1/2}/\sqrt{\log\log n}}$. For any $i \geq 0$, we first define

$$X'_{i,n} = \sum_{\ell=1}^{L} a_\ell \, g_n \circ f_\ell(Y_i) - \sum_{\ell=1}^{L} a_\ell \mathbb{E}(g_n \circ f_\ell(Y_i)) \quad \text{and} \quad X''_{i,n} = X_i - X'_{i,n}\,.$$

Let

$$d_{i,n} = \sum_{j \geq i} \mathbb{E}(X'_{j,n}|\mathcal{F}_i) - \mathbb{E}(X'_{j,n}|\mathcal{F}_{i-1}) \quad \text{and} \quad M_{k,n} = \sum_{i=1}^{k} d_{i,n}\,.$$

The following decomposition holds

$$X_0 = d_{0,n} + \sum_{k\geq 0} \mathbb{E}(X'_{k,n}|\mathcal{F}_{-1}) - \sum_{k \geq 0} \mathbb{E}(X'_{k+1,n}|\mathcal{F}_0) + X''_{0,n}\,.$$

Let $h_n = \sum_{k\geq 0} \mathbb{E}(X'_{k,n}|\mathcal{F}_{-1})$. One can write

$$X_i = d_{0,n} \circ \theta^i + h_n \circ \theta^i - h_n \circ \theta^{i+1} + X''_{0,n} \circ \theta^i\,,$$

and consequently

$$S_k = M_{k,n} + h_n \circ \theta - h_n \circ \theta^{k+1} + S''_{k,n}\,,$$

with $S''_{k,n} = \sum_{i=1}^{k} X''_{0,n} \circ \theta^i$. Hence, for any $x > 0$,

$$\mathbb{P}(\max_{1\leq k\leq n} |S_k| \geq 3x) \leq \mathbb{P}(\max_{1\leq k\leq n} |M_{k,n}| \geq x) \\ + \mathbb{P}(\max_{1\leq k\leq n} |h_n \circ \theta - h_n \circ \theta^{k+1}| \geq x) + \mathbb{P}(\max_{1\leq k\leq n} |S''_{k,n}| \geq x)\,. \quad (5.7)$$

Let us first control the coboundary term. We have

$$\|\mathbb{E}(X'_{k,n}|\mathcal{F}_0)\|_\infty \leq \sum_{\ell=1}^{L} |a_\ell| \|\mathbb{E}(g_n \circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n \circ f_\ell(Y_k))\|_\infty\,.$$

Applying Lemma 5.2, $\|\mathbb{E}(g_n \circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n \circ f_\ell(Y_k))\|_\infty \leq 4M\phi_{1,\mathbf{Y}}(k)\sqrt{n}/\sqrt{\log\log n}$. It follows that

$$\|h_n\|_\infty \leq 4M\Big(\sum_{k=1}^{\infty} \phi_{1,\mathbf{Y}}(k)\Big)\frac{\sqrt{n}}{\sqrt{\log\log n}}\,.$$

Hence, there exists a positive constant K_1 such that

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} |h_n \circ \theta - h_n \circ \theta^{k+1}| \geq CM\sqrt{n \log\log n}\Big) < K_1\,. \quad (5.8)$$

Let us now control the large part X''. We will prove the existence of a positive constant K_2 such that

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} |S''_{k,n}| \geq CM\sqrt{n \log\log n}\Big) < K_2\,. \quad (5.9)$$

We shall use the following lemma, whose proof is straightforward:

Lemma 5.4.

$$\mathbb{P}\Big(\max_{1\le k\le n}|S''_{k,n}|\ge x\Big)\le \frac{2n}{x}\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(|f_\ell(Y_0)|\mathbf{1}_{|f_\ell(Y_0)|>Mn^{1/2}/\sqrt{\log\log n}})\,.$$

Applying Lemma 5.4 with $x=CM\sqrt{n\log\log n}$, we obtain that

$$\mathbb{P}\Big(\max_{1\le k\le n}|S''_{k,n}|\ge CM\sqrt{n\log\log n}\Big)$$
$$\le \frac{2n}{CM\sqrt{n\log\log n}}\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(|f_\ell(Y_0)|\mathbf{1}_{|f_\ell(Y_0)|>Mn^{1/2}/\sqrt{\log\log n}}).$$

Now, via Fubini, there exists a positive constant A_1 such that

$$\sum_{n>0}\frac{1}{n}\frac{n}{\sqrt{n\log\log n}}\mathbb{E}(|f_\ell(Y_0)|\mathbf{1}_{|f_\ell(Y_0)|>Mn^{1/2}/\sqrt{\log\log n}})<A_1\|f_\ell(Y_0)\|_2^2\le A_1M^2\,,$$

and (5.9) follows with $K_2=(2A_1M)/C$.

Next, we turn to the main term, that is the martingale term. We will prove that there exists a positive constant K_3 such that

$$\sum_{n>0}\frac{1}{n}\mathbb{P}\Big(\sup_{1\le j\le n}|M_{j,n}|\ge CM\sqrt{n\log\log n}\Big)<K_3\,. \tag{5.10}$$

The main contribution will be controlled through the following maximal inequality.

Lemma 5.5. *Let*

$$c_n=\frac{8M\sqrt{n}}{\sqrt{\log\log n}}\sum_{k\ge0}\phi^{1/2}_{1,\mathbf{Y}}(k)\,.$$

The following upper bound holds: for any positive reals x and y,

$$\mathbb{P}\Big(\sup_{1\le j\le n}|M_{j,n}|\ge x,\sum_{j=1}^{n}\mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1})\le 2y\Big)\le 2\exp\left(-\frac{2y}{c_n^2}\,h\Big(\frac{xc_n}{2y}\Big)\right)\,,$$

where $h(u)=(1+u)\ln(1+u)-u\ge u\ln(1+u)/2$.

Proof. Note first that

$$\|d_{0,n}\|_\infty\le 2\sum_{k\ge0}\|\mathbb{E}(X'_{k,n}|\mathcal{F}_0)\|_\infty\le 2\sum_{k\ge0}\sum_{\ell=1}^{L}|a_\ell|\|\mathbb{E}(g_n\circ f_\ell(Y_k)|\mathcal{F}_0)-\mathbb{E}(g_n\circ f_\ell(Y_k))\|_\infty\,.$$

Now, applying Lemma 5.2,

$$\|\mathbb{E}(g_n \circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n \circ f_\ell(Y_k))\|_\infty \le \frac{4M\sqrt{n}}{\sqrt{\log\log n}}\phi_{1,\mathbf{Y}}(k)\,,$$

so that

$$\|d_{0,n}\|_\infty \le \frac{8M\sqrt{n}}{\sqrt{\log\log n}}\Big(\sum_{k\ge 0}\phi_{1,\mathbf{Y}}(k)\Big) \le c_n\,.$$

Proposition A.1 in Dedecker, Gouëzel and Merlevède (2010) shows that any sequence of martingale differences d_j which is bounded by a constant c satisfies

$$\mathbb{P}\Big(\sup_{1\le j\le n}|M_j| \ge x, \sum_{j=1}^n \mathbb{E}(d_j^2|\mathcal{F}_{j-1}) \le 2y\Big) \le 2\exp\Big(-\frac{2y}{c^2}\,h\Big(\frac{xc}{2y}\Big)\Big)\,.$$

The sequence $d_j = d_{j,n}$ satisfies the assumptions of this proposition for $c = c_n$. Therefore, Lemma 5.5 follows. □

Notice that

$$\sum_{j=1}^n \mathbb{E}(d_{j,n}^2) = n\mathbb{E}(d_{1,n}^2) \le 4n\Big\|\sum_{j\ge 0}\mathbb{E}(X'_{j,n}|\mathcal{F}_0)\Big\|_2^2.$$

Now,

$$\|\mathbb{E}(X'_{k,n}|\mathcal{F}_0)\|_2 \le \sum_{\ell=1}^L |a_\ell|\|\mathbb{E}(g_n \circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n \circ f_\ell(Y_k))\|_2\,.$$

Applying Lemma 5.2, $\|\mathbb{E}(g_n \circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n \circ f_\ell(Y_k))\|_2 \le 2\sqrt{2}\phi_{1,\mathbf{Y}}^{1/2}(k)\|f_\ell(Y_0)\|_2$. It follows that

$$\|\mathbb{E}(X'_{k,n}|\mathcal{F}_0)\|_2 \le 2\sqrt{2}\phi_{1,\mathbf{Y}}^{1/2}(k)M$$

and consequently

$$\sum_{j=1}^n \mathbb{E}(d_{j,n}^2) \le 32\,n\Big(\sum_{k\ge 0}\phi_{1,\mathbf{Y}}^{1/2}(k)\Big)^2 M^2\,.$$

We apply Lemma 5.5 with

$$y = y_n = 32\,n\Big(\sum_{k\ge 0}\phi_{1,\mathbf{Y}}^{1/2}(k)\Big)^2 M^2\,. \tag{5.11}$$

Letting $x_n = CM\sqrt{n\log\log n}$, we have

$$\sum_{n>0}\frac{1}{n}\mathbb{P}\Big(\sup_{1\le j\le n}|M_{j,n}| \ge x_n, \sum_{j=1}^n \mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) \le 2y_n\Big)$$
$$\le 2\sum_{n>0}\frac{1}{n}\exp\Big(-\frac{x_n}{2c_n}\ln(1 + x_n c_n/(2y_n))\Big)\,.$$

Now, the choice of C imply that $x_n = 4y_n/c_n$ and $2y_n = c_n^2(\log\log n)$. It follows that

$$\sum_{n>0} \frac{1}{n} \exp\Big(-\frac{x_n}{2c_n} \ln(1 + x_n c_n/(2y_n))\Big) = \sum_{n>0} \frac{1}{n} \exp\big(-(\log\log n)\log 3\big) < \infty\,.$$

To prove (5.10), it remains to prove that there exists a positive constant K_4 such that

$$\sum_{n\geq 1} \frac{1}{n}\mathbb{P}\Big(\sum_{j=1}^{n} \mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) \geq 2y_n\Big) < K_4\,.$$

Since $\sum_{j=1}^n \mathbb{E}(d_{j,n}^2) \leq y_n$, it suffices to prove that

$$\sum_{n\geq 1} \frac{1}{n}\mathbb{P}\Big(\Big|\sum_{j=1}^{n} (\mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) - \mathbb{E}(d_{j,n}^2))\Big| \geq y_n\Big) < K_4\,. \tag{5.12}$$

To prove (5.12), we shall use the following lemma:

Lemma 5.6. *If* (5.1) *holds, there exists a positive constant* $C_2(\phi)$ *such that for any* $y > 0$*,*

$$\mathbb{P}\Big(\Big|\sum_{j=1}^{n} (\mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) - \mathbb{E}(d_{j,n}^2))\Big| \geq y\Big) \leq \frac{nC_2(\phi)}{y^2} \sum_{\ell=1}^{L} |a_\ell| \mathbb{E}(f_\ell(Y_0)^4 \mathbf{1}_{|f_\ell(Y_0)|\leq Mn^{1/2}})\,.$$

Before proving Lemma 5.6, let us complete the proof of (5.12), (5.10) and (5.2). Since y_n is given by (5.11), we infer from Lemma 5.6 that there exists a positive constant $C_3(\phi)$ such that

$$\sum_{n>0} \frac{1}{n}\mathbb{P}\Big(\Big|\sum_{j=1}^{n} (\mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) - \mathbb{E}(d_{j,n}^2))\Big| \geq y_n\Big) \leq \frac{C_3(\phi)}{M^4} \sum_{\ell=1}^{L} |a_\ell| \sum_{n>0} \frac{1}{n^2}\mathbb{E}(f_\ell(Y_0)^4 \mathbf{1}_{|f_\ell(Y_0)|\leq Mn^{1/2}})\,.$$

By Fubini, the last sum in this equation is bounded by $4M^2\|f_\ell(Y_0)\|_2^2 \leq 4M^4$. Therefore, (5.12) follows with $K_4 = 4C_3(\phi)$. This completes the proof of (5.10). Now, the proof of (5.6) follows from (5.7), (5.8), (5.9) and (5.10). The inequality (5.2) of Proposition 5.3 is proved. □

It remains to prove Lemma 5.6.

Proof of Lemma 5.6. In a sense, the contribution coming from Lemma 5.6 is less essential than the contribution we estimated thanks to the maximal inequality. However, it is rather technical to estimate. To handle this term, we will argue in the other direction, and go from the martingale to the partial sums of the original random variables.

We apply Theorem 3 in Wu and Zhao (2008): for any $q \in (1,2]$ there exists a positive constant C_q such that

$$\mathbb{E}\Big(\Big|\sum_{j=1}^{n}(\mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) - \mathbb{E}(d_{j,n}^2))\Big|^q\Big) \le C_q n \mathbb{E}(|d_{1,n}|^{2q}) + C_q n \Delta_{n,q}^*$$

where

$$\Delta_{n,q}^* = \Big(\sum_{k=1}^{n}\frac{1}{k^{1+1/q}}\|\mathbb{E}(M_{k,n}^2|\mathcal{F}_0) - \mathbb{E}(M_{k,n}^2)\|_q\Big)^q .$$

Hence, by Markov's inequality with $q=2$, one has

$$\mathbb{P}\Big(\Big|\sum_{j=1}^{n}(\mathbb{E}(d_{j,n}^2|\mathcal{F}_{j-1}) - \mathbb{E}(d_{j,n}^2))\Big| \ge y\Big) \le \frac{C_2 n}{y^2}\Big(\mathbb{E}(|d_{1,n}|^4) + \Delta_{n,2}^*\Big) .$$

Note first that

$$\mathbb{E}(|d_{1,n}|^4) \le 16\Big(\sum_{j\ge 0}\|\mathbb{E}(X'_{j,n}|\mathcal{F}_0)\|_4\Big)^4 .$$

Now

$$\|\mathbb{E}(X'_{k,n}|\mathcal{F}_0)\|_4 \le \sum_{\ell=1}^{L}|a_\ell|\|\mathbb{E}(g_n\circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n\circ f_\ell(Y_k))\|_4 .$$

Applying Lemma 5.2, $\|\mathbb{E}(g_n\circ f_\ell(Y_k)|\mathcal{F}_0) - \mathbb{E}(g_n\circ f_\ell(Y_k))\|_4 \le 2(2\phi_{1,\mathbf{Y}}(k))^{3/4}\|g_n\circ f_\ell(Y_0)\|_4$. It follows that

$$\mathbb{E}(|d_{1,n}|^4) \le 2^{11}\Big(\sum_{k>0}\phi_{1,\mathbf{Y}}(k)^{3/4}\Big)^4\Big(\sum_{\ell=1}^{L}|a_\ell|\|g_n\circ f_\ell(Y_0)\|_4\Big)^4 .$$

Applying Jensen's inequality,

$$\mathbb{E}(|d_{1,n}|^4) \le 2^{11}\Big(\sum_{k>0}\phi_{1,\mathbf{Y}}(k)^{3/4}\Big)^4\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\le Mn^{1/2}}) . \tag{5.13}$$

Now, letting $S'_{k,n} = \sum_{i=1}^{k} X'_{i,n}$, one has $M_{k,n} = S'_{k,n} - R_{k,n}$, with

$$R_{k,n} = \sum_{i\ge 1}\mathbb{E}(X'_{i,n}|\mathcal{F}_0) - \sum_{i\ge k+1}\mathbb{E}(X'_{i,n}|\mathcal{F}_k) .$$

Hence

$$\Delta_{n,2}^* \le 3\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\|\mathbb{E}(S_{k,n}'^2|\mathcal{F}_0) - \mathbb{E}(S_{k,n}'^2)\|_2\Big)^2 + 3\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\|R_{k,n}^2\|_2\Big)^2$$
$$+ 12\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\|\mathbb{E}(S'_{k,n}R_{k,n}|\mathcal{F}_0) - \mathbb{E}(S'_{k,n}R_{k,n})\|_2\Big)^2 .$$

Arguing as for the proof of (5.13), we obtain that

$$\begin{aligned}\|R_{k,n}^2\|_2 &\leq 4\Big\|\sum_{i\geq 1}\mathbb{E}(X'_{i,n}|\mathcal{F}_0)\Big\|_4^2\\ &\leq 32\sqrt{2}\Big(\sum_{k>0}\phi_{1,\mathbf{Y}}(k)^{3/4}\Big)^2\Big(\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\leq Mn^{1/2}})\Big)^{1/2}.\end{aligned}$$

From the proof of Corollary 2.1 in Dedecker, Doukhan and Merlevède (2011), for any $\gamma \in (0,1]$ (to be chosen later), there exists a positive constant B such that

$$\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\|\mathbb{E}(S_{k,n}'^2|\mathcal{F}_0)-\mathbb{E}(S_{k,n}'^2)\|_2\Big)^2 \leq BI_1^2 + BI_2^2 \tag{5.14}$$

where

$$\begin{aligned}I_1 &= \sum_{m>0}\frac{m^\gamma}{m^{1/2}}\sup_{i\geq j\geq m}\|\mathbb{E}(X'_{i,n}X'_{j,n}|\mathcal{F}_0)-\mathbb{E}(X'_{i,n}X'_{j,n})\|_2\\ I_2 &= \Big(\sum_{k>0}\frac{k^{1/(2\gamma)}}{k^{1/4}}\|\mathbb{E}(X'_{k,n}|\mathcal{F}_0)\|_4\Big)^2.\end{aligned}$$

Arguing as for the proof of (5.13), we obtain that

$$I_2 \leq 8\sqrt{2}\Big(\sum_{k>0}\frac{k^{1/(2\gamma)}}{k^{1/4}}\phi_{1,\mathbf{Y}}(k)^{3/4}\Big)^2\Big(\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\leq Mn^{1/2}})\Big)^{1/2}. \tag{5.15}$$

To bound I_1, note that

$$\begin{aligned}&\|\mathbb{E}(X'_{i,n}X'_{j,n}|\mathcal{F}_0)-\mathbb{E}(X'_{i,n}X'_{j,n})\|_2\\ &\leq \sum_{k=1}^{L}\sum_{\ell=1}^{L}|a_k||a_\ell|\|\mathbb{E}((g_n\circ f_k(Y_i))^{(0)}(g_n\circ f_\ell(Y_j))^{(0)}|\mathcal{F}_0)-\mathbb{E}((g_n\circ f_k(Y_i))^{(0)}(g_n\circ f_\ell(Y_j))^{(0)})\|_2.\end{aligned}$$

Applying Lemma 5.2, for $i\geq j\geq m$,

$$\begin{aligned}&\|\mathbb{E}((g_n\circ f_k(Y_i))^{(0)}(g_n\circ f_\ell(Y_j))^{(0)}|\mathcal{F}_0)-\mathbb{E}((g_n\circ f_k(Y_i))^{(0)}(g_n\circ f_\ell(Y_j))^{(0)})\|_2\\ &\qquad\leq 16\phi_{2,\mathbf{Y}}(m)^{1/2}\|g_n\circ f_k(Y_0)\|_4\|g_n\circ f_\ell(Y_0)\|_4.\end{aligned}$$

It follows that

$$I_1 \leq \Big(16\sum_{m>0}\frac{m^\gamma}{m^{1/2}}\phi_{2,\mathbf{Y}}(m)^{1/2}\Big)\Big(\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\leq Mn^{1/2}})\Big)^{1/2}. \tag{5.16}$$

Let $\gamma = 1/\sqrt{3}$. If the condition (5.1) holds, then

$$\sum_{k>0} \frac{k^{\sqrt{3}/2}}{k^{1/4}} \phi_{1,\mathbf{Y}}(k)^{3/4} < \infty \quad \text{and} \quad \sum_{m>0} \frac{m^{1/\sqrt{3}}}{m^{1/2}} \phi_{2,\mathbf{Y}}(m)^{1/2} < \infty.$$

To see that the convergence of the second series implies the convergence of the first series, it suffices to note that $\phi_{1,\mathbf{Y}}(n) \le \phi_{2,\mathbf{Y}}(n)$ and that, since $\phi_{2,\mathbf{Y}}(n)$ is nonincreasing, $\phi_{2,\mathbf{Y}}(n) = o(n^{-(2+\sqrt{3})/\sqrt{3}})$.

We infer from (5.14), (5.15) and (5.16) that, if (5.1) holds, there exists a positive constant $C_4(\phi)$ such that

$$\Big(\sum_{k=1}^{n} \frac{1}{k^{3/2}} \|\mathbb{E}(S'^2_{k,n}|\mathcal{F}_0) - \mathbb{E}(S'^2_{k,n})\|_2\Big)^2 \le C_4(\phi) \sum_{\ell=1}^{L} |a_\ell| \mathbb{E}(f_\ell(Y_0)^4 \mathbf{1}_{|f_\ell(Y_0)| \le M n^{1/2}}) . \tag{5.17}$$

Let us consider now the term

$$\Big(\sum_{k=1}^{n} \frac{1}{k^{3/2}} \|\mathbb{E}(S'_{k,n} R_{k,n}|\mathcal{F}_0) - \mathbb{E}(S'_{k,n} R_{k,n})\|_2\Big)^2 .$$

As for the proof of (5.13), one has

$$\begin{aligned}\Big\|\mathbb{E}(S'_{k,n}|\mathcal{F}_0) \sum_{i\ge1} \mathbb{E}(X'_{i,n}|\mathcal{F}_0)\Big\|_2^2 &\le \Big(\sum_{i\ge1} \|\mathbb{E}(X'_{i,n}|\mathcal{F}_0)\|_4\Big)^2 \\ &\le 8\sqrt{2}\Big(\sum_{i\ge1} \phi^{3/4}_{1,\mathbf{Y}}(i)\Big)^2 \Big(\sum_{\ell=1}^{L} |a_\ell| \mathbb{E}(f_\ell(Y_0)^4 \mathbf{1}_{|f_\ell(Y_0)| \le M n^{1/2}})\Big)^{1/2} .\end{aligned}$$

Next, we need to bound

$$\Big(\sum_{k=1}^{n} \frac{1}{k^{3/2}} \Big\|\mathbb{E}(S'_{k,n} \sum_{i\ge k+1} \mathbb{E}(X'_{i,n}|\mathcal{F}_k)|\mathcal{F}_0) - \mathbb{E}(S'_{k,n} \sum_{i\ge k+1} \mathbb{E}(X'_{i,n}|\mathcal{F}_k))\Big\|_2\Big)^2 .$$

First, we see that

$$\sum_{i\ge k+1} \mathbb{E}(X'_{i,n}|\mathcal{F}_k) = \mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k) + \sum_{j\ge 2k+1} \mathbb{E}(X'_{j,n}|\mathcal{F}_k) .$$

Since $S'_{k,n}$ is $\mathcal{F}_k$-measurable, we get that

$$\begin{aligned}&\|\mathbb{E}(S'_{k,n}\mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k)|\mathcal{F}_0) - \mathbb{E}(S'_{k,n}\mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k))\|_2 \\ &\qquad = \|\mathbb{E}(S'_{k,n}(S'_{2k,n} - S'_{k,n})|\mathcal{F}_0) - \mathbb{E}(S'_{k,n}(S'_{2k,n} - S'_{k,n}))\|_2 .\end{aligned}$$

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

$$2\|\mathbb{E}(S'_{k,n}\mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k)|\mathcal{F}_0) - \mathbb{E}(S'_{k,n}\mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k))\|_2 \\ \le \|\mathbb{E}(S'^2_{2k,n}|\mathcal{F}_0) - \mathbb{E}(S'^2_{2k,n})\|_2 + 2\|\mathbb{E}(S'^2_{k,n}|\mathcal{F}_0) - \mathbb{E}(S'^2_{k,n})\|_2\,,$$

which combined with (5.17) implies that

$$\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\|\mathbb{E}(S'_{k,n}\mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k)|\mathcal{F}_0) - \mathbb{E}(S'_{k,n}\mathbb{E}(S'_{2k,n} - S'_{k,n}|\mathcal{F}_k))\|_2\Big)^2 \\ \le 6\,C_4(\phi)\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\le bn^{1/2}})\,.$$

It remains to bound

$$\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\Big\|\mathbb{E}\Big(S'_{k,n}\sum_{j\ge 2k+1}\mathbb{E}(X'_{j,n}|\mathcal{F}_k)\Big|\mathcal{F}_0\Big)\Big\|_2\Big)^2.$$

By stationarity,

$$\sum_{j\ge 2k+1}\|\mathbb{E}(S'_{k,n}\mathbb{E}(X'_{j,n}|\mathcal{F}_k))|\mathcal{F}_0)\|_2 \le k\sum_{j\ge k+1}\|X'_{0,n}\mathbb{E}(X'_{j,n}|\mathcal{F}_0)\|_2\,.$$

Now, as for the proof of (5.13),

$$\sum_{j\ge k+1}\|X'_{0,n}\mathbb{E}(X'_{j,n}|\mathcal{F}_0)\|_2 \le \|X'_{0,n}\|_4\sum_{j\ge k+1}\|\mathbb{E}(X'_{j,n}|\mathcal{F}_0)\|_4 \\ \le 2\Big(2\sum_{j\ge k+1}(2\phi_{1,\mathbf{Y}}(j))^{3/4}\Big)\Big(\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\le bn^{1/2}})\Big)^{1/2},$$

and consequently, there exists a positive constant D such that

$$\Big(\sum_{k=1}^{n}\frac{1}{k^{3/2}}\Big\|\mathbb{E}\Big(S'_{k,n}\sum_{j\ge 2k+1}\mathbb{E}(X'_{j,n}|\mathcal{F}_k)\Big|\mathcal{F}_0\Big)\Big\|_2\Big)^2 \\ \le D\Big(\sum_{j\ge 2}j^{1/2}\phi_{1,\mathbf{Y}}(j)^{3/4}\Big)^2\sum_{\ell=1}^{L}|a_\ell|\mathbb{E}(f_\ell(Y_0)^4\mathbf{1}_{|f_\ell(Y_0)|\le Mn^{1/2}})\,.$$

The lemma is proved. □

6 Proof of Theorem 1.5 on uniformly expanding maps

Let $(Y_i)_{i\geq 0}$ be the stationary Markov chain with transition kernel K corresponding to the iteration of the inverse branches of T, and let $X_n = f(Y_n) - \nu(f)$. Concerning Item 1 in Theorem 1.5, it is well known that it is equivalent to prove it for the iteration of the map or of the Markov chain, since the distributions are the same (see for instance the proof of Theorem 2.1 in Dedecker and Merlevède (2009)). Therefore, it is enough to show that the process

$$\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} X_i,\ t\in[0,1]\Big\}$$

converges in distribution in the Skorokhod topology to σW, where W is a standard Wiener process. Now, as shown by Heyde (1975), this property as well as the absolute convergence of the series (1.1) will be true provided that Gordin's condition (1969) holds, that is

$$\sum_{n=0}^{\infty} \|K^n(f) - \nu(f)\|_{\mathbb{L}^2(\nu)} < \infty\,. \tag{6.1}$$

By definition of $\mathrm{Mon}_2^c(M,\nu)$, there exists a sequence of functions $f_L = \sum_{\ell=1}^L a_{\ell,L} g_{\ell,L}$ with $g_{\ell,L}$ belonging to $\mathrm{Mon}_2(M,\nu)$ and $\sum_{\ell=1}^L |a_{\ell,L}| \leq 1$, such that f_L converges in $\mathbb{L}^1(\nu)$ to f. It follows that, for any nonnegative integer n, $K^n(f_L) - \nu(f_L)$ converges to $K^n(f) - \nu(f)$ in $\mathbb{L}^1(\nu)$. Hence, there exists a subsequence $K^n(f_{\varphi(L)}) - \nu(f_{\varphi(L)})$ converging to $K^n(f) - \nu(f)$ almost surely and in $\mathbb{L}^1(\nu)$. Applying Fatou's lemma, we infer that

$$\|K^n(f) - \nu(f)\|_{\mathbb{L}^2(\nu)} \leq \liminf_{L\to\infty} \|K^n(f_{\varphi(L)}) - \nu(f_{\varphi(L)})\|_{\mathbb{L}^2(\nu)}\,. \tag{6.2}$$

Applying Lemma 5.2, for any g in $\mathrm{Mon}_2(M,\nu)$, $\|K^n(g) - \nu(g)\|_{\mathbb{L}^2(\nu)} \leq 2\sqrt{2}\phi_{1,\mathbf{Y}}^{1/2}(n)M$. Hence

$$\|K^n(f_{\varphi(L)}) - \nu(f_{\varphi(L)})\|_{\mathbb{L}^2(\nu)} \leq \sum_{\ell=1}^L |a_{\ell,L}|\|K^n(g_{\ell,\varphi(L)}) - \nu(g_{\ell,\varphi(L)})\|_{\mathbb{L}^2(\nu)} \leq 2\sqrt{2}\phi_{1,\mathbf{Y}}^{1/2}(n)M\,.$$

From (6.2), it follows that $\|K^n(f) - \nu(f)\|_{\mathbb{L}^2(\nu)} \leq 2\sqrt{2}\phi_{1,\mathbf{Y}}^{1/2}(n)M$, and (6.1) holds provided that $\sum_{n>0}\phi_{1,\mathbf{Y}}^{1/2}(n) < \infty$. Now, if T is uniformly expanding, it follows from Section 6.3 in Dedecker and Prieur (2007) that $\phi_{2,\mathbf{Y}}(n) = O(\rho^n)$ for some $\rho\in(0,1)$, and Item 1 is proved.

According to the inequality (4.1) in Dedecker, Gouëzel and Merlevède (2010), we have

$$\nu\Big(\max_{1\leq k\leq n}\Big|\sum_{i=0}^{k-1}(f\circ T^i - \nu(f))\Big| > x\Big) \leq \nu\Big(2\max_{1\leq k\leq n}\Big|\sum_{i=1}^{k} X_i\Big| > x\Big)\,.$$

Therefore, Item 2 follows from Proposition 5.3 applied to the sequences $(X_i)_{i\geq 1}$ as soon as (5.1) holds, which is clearly true.

For Item 3, we proceed exactly as in the case of GPM maps, relying on the approximation $f = \bar{f}_m + g_m$ given by Lemma 2.3 to apply Proposition 3.1. Since $g_m \in \mathrm{Mon}_2^c(1/m, \nu)$, Proposition 5.3 shows that almost surely

$$\limsup \frac{1}{\sqrt{n \log \log n}} \left| \sum_{i=0}^{n-1} (g_m \circ T^i - \nu(g_m)) \right| \leq C/m \,,$$

for some constant C. Moreover, the proof of Theorem 3.1 in Merlevède and Rio (2012) shows that the sequence $\bar{f}_m \circ T^i - \nu(f_m)$ satisfies an almost sure principle, towards a Gaussian with variance σ_m^2. It only remains to show that σ_m^2 converges to σ^2. We start from the basic inequality

$$\sigma^2(g_m) \leq 2 \|g_m\|_{\mathbb{L}^2(\nu)} \sum_{n=0}^{\infty} \|K^n(g_m) - \nu(g_m)\|_{\mathbb{L}^2(\nu)} \,.$$

Arguing as in (6.2), we infer that

$$\sigma^2(g_m) \leq 16 m^{-2} \sum_{k=0}^{\infty} \phi_{1,\mathbf{Y}}^{1/2}(k)$$

and the series on the right hand side is finite since $\phi_{1,\mathbf{Y}}(n) = O(\rho^n)$ for some $\rho \in (0,1)$. Therefore, $\sigma^2(g_m)$ converges to 0. □

References

[1] Broise, A. (1996). Transformations dilatantes de l'intervalle et théorèmes limites. Études spectrales d'opérateurs de transfert et applications. Astérisque **238** 1-109.

[2] Dedecker, J. (2004). Inégalités de covariance. *C. R. Math. Acad. Sci. Paris* **339** 503-506.

[3] Dedecker, J., Doukhan, P. and Merlevède, F. (2011). Rates of convergence in the strong invariance principle under projective criteria. Preprint. Available at `http://perso-math.univ-mlv.fr/users/merlevede.florence/`

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

[5] Dedecker, J. and Merlevède, F. (2009). Weak invariance principle and exponential bounds for some special functions of intermittent maps. *High dimensional probability V: the Luminy volume. Institute of mathematical Statistics* 60-72.

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

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

[8] Gouëzel, S. (2004). Central limit theorem and stable laws for intermittent maps. *Probab. Theory Relat. Fields* **128** 82-122.

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

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

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

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

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

[14] Pinelis, I. (1994). Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.* **22** 1679-1706.

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

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

[17] Zweimüller, R. (1998). Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points. *Nonlinearity* **11** 1263-1276.