# Weak Convergence to Stable Lévy Processes for Nonuniformly Hyperbolic Dynamical Systems

Ian Melbourne* Roland Zweimüller†

11 January 2013; final version 18 September 2013

**Abstract**

We consider weak invariance principles (functional limit theorems) in the domain of a stable law. A general result is obtained on lifting such limit laws from an induced dynamical system to the original system. An important class of examples covered by our result are Pomeau-Manneville intermittency maps, where convergence for the induced system is in the standard Skorohod $\mathcal{J}_1$ topology. For the full system, convergence in the $\mathcal{J}_1$ topology fails, but we prove convergence in the $\mathcal{M}_1$ topology.

**Keywords:** Nonuniformly hyperbolic systems, functional limit theorems, Lévy processes, induced dynamical systems

**MSC codes:** 37D25; 28D05, 37A50, 60F17

## 1 Introduction

For large classes of dynamical systems with good mixing properties, it is possible to obtain strong statistical limit laws such as the central limit theorems and its refinements including the almost sure invariance principle (ASIP) [21, 12, 10, 15, 27, 25, 26, 5, 20]. An immediate consequence of the ASIP is the weak invariance principle (WIP) which is the focus of this paper.

*Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK i.melbourne@warwick.ac.uk

†Faculty of Mathematics, University of Vienna, Vienna, Austria

Thus the standard WIP (weak convergence to Brownian motion) holds for general Axiom A diffeomorphisms and flows, and also for nonuniformly hyperbolic maps and flows modelled by Young towers [35, 36] with square integrable return time function (including Hénon-like attractors [7], finite horizon Lorentz gases [25], and the Lorenz attractor [22]).

Recently, there has been interest in statistical limit laws for dynamical systems with weaker mixing properties such as those modelled by a Young tower where the return time function is not square integrable. In the borderline case where the return time lies in $L^p$ for all $p < 2$, it is often possible to prove a central limit theorem with nonstandard norming (nonstandard domain of attraction of the normal distribution). This includes important examples such as the infinite horizon Lorentz gas [32], the Bunimovich stadium [4] and billiards with cusps [3]. In such cases, it is also possible to obtain the corresponding WIP (see for example [3, 11]).

For Young towers with non-square-integrable return time function, the central limit theorem generally fails. Gouëzel [18] (see also Zweimüller [37]) obtained definitive results on convergence in distribution to stable laws. The only available results on the corresponding WIP are due to Tyran-Kamińska [33] who gives necessary and sufficient conditions for weak convergence to the appropriate stable Lévy process in the standard Skorohod $\mathcal{J}_1$ topology [31]. However in the situations we are interested in, the $\mathcal{J}_1$ topology is too strong and the results in [33] prove that weak convergence fails in this topology.

In this paper, we repair the situation by working with the $\mathcal{M}_1$ topology (also introduced by Skorohod [31]). In particular, we give general conditions for systems modelled by a Young tower, whereby convergence in distribution to a stable law can be improved to weak convergence in the $\mathcal{M}_1$ topology to the corresponding Lévy process.

The proof is by inducing (see [30, 28, 19] for proofs by inducing of convergence in distribution). Young towers by definition have a good inducing system, namely a Gibbs-Markov map (a Markov map with bounded distortion and big images [1]). The results of Tyran-Kamińska [33] often apply positively for such induced maps (see for example the proof of Theorem 4.1 below) and yield weak convergence in the $\mathcal{J}_1$ topology, and hence the $\mathcal{M}_1$ topology, for the induced system. The main theoretical result of the present paper discusses how $\mathcal{M}_1$ convergence in an induced system lifts to the original system (even when convergence in the $\mathcal{J}_1$ topology does not lift).

As a special case, we recover the aforementioned results [3, 11] on the WIP in the nonstandard domain of attraction of the normal distribution.

In the remainder of the introduction, we describe how our results apply to Pomeau-Manneville intermittency maps [29]. In particular, we consider the family of maps $f : X \to X$, $X = [0, 1]$, studied by [23], given by

$$f(x) = \begin{cases} x(1 + 2^\gamma x^\gamma), & x \in [0, \frac{1}{2}] \\ 2x - 1, & x \in (\frac{1}{2}, 1] \end{cases} \tag{1.1}$$

For $\gamma \in [0, 1)$, there is a unique absolutely continuous ergodic invariant probability measure $\mu$. Suppose that $\phi : X \to \mathbb{R}$ is a Hölder observable with $\int_X \phi \, d\mu = 0$. Let $\phi_n = \sum_{j=0}^{n-1} \phi \circ f^j$. For the map in (1.1), our main result implies the following:

**Theorem 1.1** *Let $f : [0, 1] \to [0, 1]$ be the map (1.1) with $\gamma \in (\frac{1}{2}, 1)$ and set $\alpha = 1/\gamma$. Let $\phi : [0, 1] \to \mathbb{R}$ be a mean zero Hölder observable and suppose that $\phi(0) \neq 0$. Define $W_n(t) = n^{-1/\alpha}\phi_{\lfloor nt \rfloor}$. Then $W_n$ converges weakly in the Skorohod $\mathcal{M}_1$ topology to an $\alpha$-stable Lévy process. (The specific Lévy process is described below.)*

**Remark 1.2** *The $\mathcal{J}_1$ and $\mathcal{M}_1$ topologies are reviewed in Section 2.1. Roughly speaking, the difference is that the $\mathcal{M}_1$ topology allows numerous small jumps for $W_n$ to accumulate into a large jump for $W$, whereas the $\mathcal{J}_1$ topology would require a large jump for $W$ to be approximated by a single large jump for $W_n$. Since the jumps in $W_n$ are bounded by $n^{-1/\alpha}|\phi|_\infty$, it is evident that in Theorem 1.1 convergence cannot hold in the $\mathcal{J}_1$ topology.*

*Situations in the probability theory literature where convergence holds in the $\mathcal{M}_1$ topology but not the $\mathcal{J}_1$ topology include [2, 6].*

Theorem 1.1 completes the study of weak convergence for the intermittency map (1.1) with $\gamma \in [0, 1)$ and typical Hölder observables. We recall the previous results in this direction. If $\gamma \in [0, \frac{1}{2})$ then it is well-known that $\phi$ satisfies a central limit theorem, so $n^{-\frac{1}{2}}\phi_n$ converges in distribution to a normal distribution with mean zero and variance $\sigma^2$, where $\sigma^2$ is typically positive. Moreover, [25] proved the ASIP. An immediate consequence is the WIP: $W_n(t) = n^{-\frac{1}{2}}\phi_{\lfloor nt \rfloor}$ converges weakly to Brownian motion.

If $\gamma = \frac{1}{2}$ and $\phi(0) \neq 0$, then Gouëzel [18] proved that $\phi$ is in the nonstandard domain of attraction of the normal distribution: $(n \log n)^{-\frac{1}{2}}\phi_n$ converges in distribution to a normal distribution with mean zero and variance $\sigma^2 > 0$. Dedecker & Merlevede [11] obtained the corresponding WIP in this situation (with $W_n(t) = (n \log n)^{-\frac{1}{2}}\phi_{\lfloor nt \rfloor}$).

Finally, if $\gamma \in (\frac{1}{2}, 1)$ and $\phi(0) \neq 0$, then Gouëzel [18] proved that $n^{-1/\alpha}\phi_n$ converges in distribution to a one-sided stable law $G$ with exponent $\alpha = \gamma^{-1}$.

The stable law in question has characteristic function

$$E(e^{itG}) = \exp\big\{-c|t|^\alpha(1 - i\,\mathrm{sgn}(\phi(0)t)\tan(\alpha\pi/2)\big\},$$

where $c = \frac{1}{4}h(\frac{1}{2})(\alpha|\phi(0)|)^\alpha\Gamma(1-\alpha)\cos(\alpha\pi/2)$ and $h = \frac{d\mu}{dx}$ is the invariant density. Let $\{W(t);\, t \geq 0\}$ denote the corresponding $\alpha$-stable Lévy process (so $\{W(t)\}$ has independent and stationary increments with cadlag sample paths and $W(t) =_d t^{1/\alpha}G$). Tyran-Kamińska [33] verified that $W_n(t) = n^{-1/\alpha}\phi_{\lfloor nt\rfloor}$ does not converge weakly to $W$ in the $\mathcal{J}_1$ topology. In contrast, Theorem 1.1 shows that $W_n$ converges weakly to $W$ in the $\mathcal{M}_1$ topology.

The remainder of this paper is organised as follows. In Section 2 we state our main abstract result, Theorem 2.2, on inducing the WIP. In Section 3 we prove Theorem 2.2. In Section 4 we consider some examples which include Theorem 1.1 as a special case.

# 2 Inducing a weak invariance principle

In this section, we formulate our main abstract result Theorem 2.2. The result is stated in Subsection 2.2 after some preliminaries in Subsection 2.1.

## 2.1 Preliminaries

**Distributional convergence.** To fix notations, let $(X, P)$ be a probability space and $(R_n)_{n\geq 1}$ a sequence of Borel measurable maps $R_n : X \to S$, where $(S, d)$ is a separable metric space. Then distributional convergence of $(R_n)_{n\geq 1}$ w.r.t. $P$ to some random element $R$ of $S$ will be denoted by $R_n \overset{P}{\Longrightarrow} R$. *Strong distributional convergence* $R_n \overset{\mathcal{L}(\mu)}{\Longrightarrow} R$ on a measure space $(X, \mu)$ means that $R_n \overset{P}{\Longrightarrow} R$ for all probability measures $P \ll \mu$.

**Skorohod spaces.** We briefly review the required background material on the Skorohod $\mathcal{J}_1$ and $\mathcal{M}_1$ topologies [31] on the linear spaces $\mathcal{D}[0,T]$, $\mathcal{D}[T_1,T_2]$, and $\mathcal{D}[0,\infty)$ of real-valued *cadlag* functions (right-continuous $g(t^+) = g(t)$ with left-hand limits $g(t^-)$) on the respective interval, referring to [34] for proofs and further information. Both topologies are Polish, with $\mathcal{J}_1$ stronger than $\mathcal{M}_1$.

It is customary to first deal with bounded time intervals. We thus fix some $T > 0$ and focus on $\mathcal{D} = \mathcal{D}[0,T]$. (Everything carries over to $\mathcal{D}[T_1,T_2]$ in an obvious fashion.) Throughout, $\|.\|$ will denote the uniform norm. Two functions $g_1, g_2 \in \mathcal{D}$ are close in the $\mathcal{J}_1$-*topology* if they are uniformly close after a small distortion of the domain. Formally, let $\Lambda$ be the set of increasing homeomorphisms $\lambda : [0,T] \to [0,T]$, and let $\lambda_{id} \in \Lambda$ denote the identity. Then $d_{\mathcal{J}_1,T}(g_1,g_2) = \inf_{\lambda\in\Lambda}\{\|g_1 \circ \lambda - g_2\| \vee \|\lambda - \lambda_{id}\|\}$ defines a metric on $\mathcal{D}$ which induces the $\mathcal{J}_1$-topology. While its restriction to $\mathcal{C} = \mathcal{C}[0,T]$ coincides with the uniform topology, discontinuous functions are $\mathcal{J}_1$-close to each other if they have jumps of similar size at similar positions.

In contrast, the $\mathcal{M}_1$-*topology* allows a function $g_1$ with a jump at $t$ to be approximated arbitrarily well by some continuous $g_2$ (with large slope near $t$). For convenience, we let $[a,b]$ denote the (possibly degenerate) closed interval with endpoints $a, b \in \mathbb{R}$, irrespective of their order. Let $\Gamma(g) := \{(t,x) \in [0,T] \times \mathbb{R} : x \in [g(t^-), g(t)]\}$ denote the completed graph of $g$, and let $\Lambda^*(g)$ be the set of all its parametrizations, that is, all continuous $G = (\lambda, \gamma) : [0,T] \to \Gamma(g)$ such that $t' < t$ implies either $\lambda(t') < \lambda(t)$ or $\lambda(t') = \lambda(t)$ plus $|\gamma(t) - g(\lambda(t))| \le |\gamma(t') - g(\lambda(t))|$. Then $d_{\mathcal{M}_1,T}(g_1,g_2) = \inf_{G_i=(\lambda_i,\gamma_i)\in\Lambda^*(g_i)}\{\|\lambda_1 - \lambda_2\| \vee \|\gamma_1 - \gamma_2\|\}$ gives a metric inducing $\mathcal{M}_1$.

On the space $\mathcal{D}[0,\infty)$ the $\tau$-topology, $\tau \in \{\mathcal{J}_1, \mathcal{M}_1\}$, is defined by the metric $d_{\tau,\infty}(g_1,g_2) := \int_0^\infty e^{-t}(1 \wedge d_{\tau,t}(g_1,g_2))\,dt$. Convergence $g_n \to g$ in $(\mathcal{D}[0,\infty),\tau)$ means that $d_{\tau,T}(g_n,g) \to 0$ for every continuity point $T$ of $g$.

For either topology, the corresponding Borel $\sigma$-field $\mathcal{B}_{\mathcal{D},\tau}$ on $\mathcal{D}$, generated by the $\tau$-open sets, coincides with the usual $\sigma$-field $\mathcal{B}_{\mathcal{D}}$ generated by the canonical projections $\pi_t(g) := g(t)$. Therefore, any family $W = (W_t)_{t\in[0,T]}$ or $(W_t)_{t\in[0,\infty)}$ of real random variables $W_t$ such that each path $t \mapsto W_t$ is cadlag, can be regarded as a random element of $\mathcal{D}$, equipped with $\tau = \mathcal{J}_1$ or $\mathcal{M}_1$.

## 2.2 Statement of the main result

Recall that for any ergodic measure preserving transformation (*m.p.t.*) $f$ on a probability space $(X,\mu)$, and any $Y \subset X$ with $\mu(Y) > 0$, the *return time* function $r : Y \to \mathbb{N} \cup \{\infty\}$ given by $r(y) := \inf\{k \ge 1 : f^k(y) \in Y\}$ is integrable with mean $\int_Y r\,d\mu_Y = \mu(Y)^{-1}$ (Kac' formula), where $\mu_Y(A) := \mu(Y \cap A)/\mu(Y)$. Moreover, the *first return map* or *induced map* $F := f^r : Y \to Y$ is an ergodic m.p.t. on the probability space $(Y, \mu_Y)$. This is widely used as a tool in the study of complicated systems, where $Y$ is chosen in such a way that $F$ is more convenient than $f$. In particular, given an *observable* (i.e. a measurable function) $\phi : X \to \mathbb{R}$, it may be easier to first consider its *induced version* $\Phi : Y \to \mathbb{R}$ on $Y$, given by $\Phi := \sum_{\ell=0}^{r-1} \phi \circ f^\ell$. By standard arguments, if $\phi \in L^1(X,\mu)$ then $\Phi \in L^1(Y,\mu_Y)$ and $\int_Y \Phi\,d\mu_Y = \mu(Y)^{-1}\int_X \phi\,d\mu$. In this setup, we will denote the corresponding ergodic sums by $\phi_k := \sum_{\ell=0}^{k-1} \phi \circ f^\ell$ and $\Phi_n := \sum_{j=0}^{n-1} \Phi \circ F^j$, respectively.

Our core result allows us to pass from a weak invariance principle for the induced version to one for the original observable. Such a step requires some a priori control of the behaviour of ergodic sums $\phi_k$ during an excursion from $Y$. We shall express this in terms of the function $\Phi^* : Y \to [0, \infty]$ given by

$$\Phi^*(y) := \Big( \max_{0 \le \ell' \le \ell \le r(y)} (\phi_{\ell'}(y) - \phi_\ell(y)) \Big) \wedge \Big( \max_{0 \le \ell' \le \ell \le r(y)} (\phi_\ell(y) - \phi_{\ell'}(y)) \Big).$$

Note that $\Phi^*$ vanishes if and only if the ergodic sums $\phi_k$ grow monotonically (nonincreasing or nondecreasing) during each excursion. Hence bounding $\Phi^*$ means limiting the growth of $\phi_\ell$ until the first return to $Y$ in at least one direction. The expression $\Phi^*$ can be understood also in terms of the maximal and minimal processes $\phi_\ell^\uparrow, \phi_\ell^\downarrow$ defined during each excursion $0 \le \ell \le r(y)$ by

$$\phi_\ell^\uparrow(y) = \max_{0 \le \ell' \le \ell} \phi_{\ell'}(y), \quad \phi_\ell^\downarrow(y) = \min_{0 \le \ell' \le \ell} \phi_{\ell'}(y).$$

**Proposition 2.1**
*(i) In the "predominantly increasing" case* $\Phi^*(y) = \max_{0 \le \ell' \le \ell \le r(y)} (\phi_{\ell'}(y) - \phi_\ell(y))$, *we have* $\Phi^*(y) = \max_{0 \le \ell \le r(y)} \Big( \phi_\ell^\uparrow(y) - \phi_\ell(y) \Big)$.

*(ii) In the "predominantly decreasing" case* $\Phi^*(y) = \max_{0 \le \ell' \le \ell \le r(y)} (\phi_\ell(y) - \phi_{\ell'}(y))$, *we have* $\Phi^*(y) = \max_{0 \le \ell \le r(y)} \Big( \phi_\ell(y) - \phi_\ell^\downarrow(y) \Big)$.

**Proof.** This is immediate from the definition of $\phi_\ell^\uparrow$ and $\phi_\ell^\downarrow$. ■

We use $\Phi^*$ to impose a weak monotonicity condition for $\phi_\ell$ during excursions.

**Theorem 2.2 (Inducing a weak invariance principle)** *Let $f$ be an ergodic m.p.t. on the probability space $(X, \mu)$, and let $Y \subset X$ be a subset of positive measure with return time $r$ and first return map $F$. Suppose that the observable $\phi : X \to \mathbb{R}$ is such that its induced version $\Phi$ satisfies a WIP on $(Y, \mu_Y)$ in that*

$$(P_n(t))_{t \ge 0} := \left( \frac{\Phi_{\lfloor tn \rfloor}}{B(n)} \right)_{t \ge 0} \overset{\mathcal{L}(\mu_Y)}{\Longrightarrow} (W(t))_{t \ge 0} \quad \text{in } (\mathcal{D}[0, \infty), \mathcal{M}_1), \tag{2.1}$$

*where $B$ is regularly varying of index $\gamma > 0$, and $(W(t))_{t \ge 0}$ is a process with cadlag paths. Moreover, assume that*

$$\frac{1}{B(n)} \left( \max_{0 \le j \le n} \Phi^* \circ F^j \right) \overset{\mu_Y}{\Longrightarrow} 0. \tag{2.2}$$

*Then $\phi$ satisfies a WIP on $(X, \mu)$ in that*

$$(W_n(s))_{s\geq 0} := \left(\frac{\phi_{\lfloor sn \rfloor}}{B(n)}\right)_{s\geq 0} \overset{\mathcal{L}(\mu)}{\Longrightarrow} (W(s\mu(Y)))_{s\geq 0} \text{ in } (\mathcal{D}[0,\infty), \mathcal{M}_1). \quad (2.3)$$

**Remark 2.3 ($\alpha$-stable processes)** *If the process $W$ in* (2.1) *for the induced system is an $\alpha$-stable Lévy process, then the limiting process in* (2.3) *is $(\int_Y r\, d\mu_Y)^{-1/\alpha} W$.*

**Remark 2.4** *In general, the convergence from (2.3) fails in $(\mathcal{D}[0,\infty), \mathcal{J}_1)$, even if (2.1) holds in the $\mathcal{J}_1$-topology. That this is the case for the intermittent maps* (1.1) *was pointed out in [33, Example 2.1].*

**Remark 2.5 (Continuous sample paths)** *If the process $W$ in* (2.1) *for the induced system has continuous sample paths, then the statement and proof of Theorem 2.2 is greatly simplified and the uniform topology (corresponding to the uniform norm $\| \; \|$) can be used throughout. In particular, the function $\Phi^*$ is replaced by $\Phi^*(y) = \max_{0\leq \ell < r(y)} |\phi_\ell(y)|$. In the case of normal diffusion $B(n) = n^{\frac{1}{2}}$, condition* (2.2) *is then satisfied if $\Phi^* \in L^2$.*

*A simplified proof based on the one presented here is written out in [17, Appendix].*

**Remark 2.6 (Centering)** *In the applications that we principally have in mind (including the maps* (1.1)*), the observable $\phi : X \to \mathbb{R}$ is integrable, and hence so is its induced version $\Phi : Y \to \mathbb{R}$. In particular, if $\phi$ has mean zero, then $\Phi$ has mean zero and we are in a situation to apply Theorem 2.2. From this, it follows easily that if condition* (2.1) *holds with*

$$P_n(t) = \frac{\Phi_{\lfloor tn \rfloor} - tn \int_Y \Phi \, d\mu_Y}{B(n)},$$

*and condition* (2.2) *holds with $\phi$ replaced throughout by $\phi - \int_X \phi \, d\mu$ in the definition of $\Phi^*$, then conclusion* (2.3) *is valid with*

$$W_n(s) = \frac{\phi_{\lfloor sn \rfloor} - sn \int_X \phi \, d\mu}{B(n)}.$$

*With a little more effort it is also possible to handle more general centering sequences where the process $(P_n)$ in condition* (2.1) *takes the form*

$$P_n(t) = \frac{\Phi_{\lfloor tn \rfloor} - tA(n)}{B(n)},$$

*for real sequences $A(n)$, $B(n)$ with $B(n) \to \infty$.*

The monotonicity condition (2.2) will be shown to hold, for example, if we have sufficiently good pointwise control for single excursions:

**Proposition 2.7 (Pointwise weak monotonicity)** *Let $f$ be an ergodic m.p.t. on the probability space $(X, \mu)$, and let $Y \subset X$ be a subset of positive measure with return time $r$. Let $B$ be regularly varying of index $\gamma > 0$. Suppose that for the observable $\phi : X \to \mathbb{R}$ there is some $\eta \in (0, \infty)$ such that for a.e. $y \in Y$,*

$$\Phi^*(y) \leq \eta \, B(r(y)). \tag{2.4}$$

*Then the weak monotonicity condition (2.2) holds.*

The proofs of Theorem 2.2 and Proposition 2.7 are given in Section 3.

Assuming strong distributional convergence $\overset{\mathcal{L}(\mu_Y)}{\Longrightarrow}$ in (2.1), rather than $\overset{\mu_Y}{\Longrightarrow}$, is not a restriction, as an application of the following result to the induced system $(Y, \mu_Y, F)$ shows.

**Proposition 2.8 (Automatic strong distributional convergence)** *Let $f$ be an ergodic m.p.t. on a $\sigma$-finite space $(X, \mu)$. Let $\tau = \mathcal{J}_1$ or $\mathcal{M}_1$ and let $A(n), B(n)$ be real sequences with $B(n) \to \infty$, Assume that $\phi : X \to \mathbb{R}$ is measurable, and that there is some probability measure $P \ll \mu$ and some random element $R$ of $\mathcal{D}[0, \infty)$ such that*

$$R_n := \left( \frac{\phi_{\lfloor tn \rfloor} - tA(n)}{B(n)} \right)_{t \geq 0} \overset{P}{\Longrightarrow} R \quad \textit{in } (\mathcal{D}[0, \infty), \tau), \tag{2.5}$$

*Then, $R_n \overset{\mathcal{L}(\mu)}{\Longrightarrow} R$ in $(\mathcal{D}[0, \infty), \tau)$.*

**Proof.** This is based on ideas in [14]. According to Zweimüller [38, Theorem 1], it suffices to check that $d_{\tau,\infty}(R_n \circ f, R_n) \overset{\mu}{\longrightarrow} 0$. The proof of [38, Corollary 3] shows that $B(n) \to \infty$ alone (that is, even without (2.5)) implies $d_{\mathcal{J}_1,\infty}(R_n \circ f, R_n) \overset{\mu}{\longrightarrow} 0$. Since $d_{\mathcal{M}_1,\infty} \leq d_{\mathcal{J}_1,\infty}$ (see [34, Theorem 12.3.2]), the case $\tau = \mathcal{M}_1$ then is a trivial consequence. ∎

**Remark 2.9** *There is a systematic typographical error in [38] in that the factor $t$ in the centering process $tA(n)/B(n)$ is missing, but the arguments there work, without any change, for the correct centering.*

# 3 Proof of Theorem 2.2.

In this section, we give the proof of Theorem 2.2 and also Proposition 2.7. Throughout, we assume the setting of Theorem 2.2. In particular, we suppose that $f$ is an ergodic m.p.t. on the probability space $(X, \mu)$, and that $Y \subset X$ is a subset of positive measure with return time $r$ and first return map $F$.

## 3.1 Decomposing the processes.

When $Y$ is chosen appropriately, many features of $f$ are reflected in the behaviour of the ergodic sums $r_n = \sum_{j=0}^{n-1} r \circ F^j$, i.e. the times at which orbits return to $Y$. These are intimately related to the *occupation times* or *lap numbers*

$$N_k := \sum_{\ell=1}^{k} 1_Y \circ f^\ell = \max\{n \geq 0 : r_n \leq k\} \leq k, \;\; k \geq 0.$$

The visits to $Y$ counted by the $N_k$ separate the consecutive *excursions from* $Y$, that is, the intervals $\{r_j, \dots, r_{j+1} - 1\}$, $j \geq 0$. Decomposing the $f$-orbit of $y$ into these excursions, we can represent the ergodic sums of $\phi$ as

$$\phi_k = \Phi_{N_k} + R_k \quad \text{on } Y$$

with remainder term $R_k = \sum_{\ell=r_{N_k}}^{k-1} \phi \circ f^\ell = \phi_{k-r_{N_k}} \circ F^{N_k}$ encoding the contribution of the incomplete last excursion (if any). Next, decompose the rescaled processes accordingly, writing

$$W_k(s) = U_k(s) + V_k(s),$$

with $U_k(s) := B(k)^{-1}\Phi_{N_{\lfloor sk \rfloor}}$, and $V_k(s) := B(k)^{-1}R_{\lfloor sk \rfloor}$. On the time scale of $U_n$, the excursions correspond to the intervals $[t_{n,j}, t_{n,j+1})$, $j \geq 0$, where $t_{n,j} : Y \to [0, \infty)$ is given by $t_{n,j} := r_j/n$. Note that the interval containing a given point $t > 0$ is that with $j = N_{\lfloor tn \rfloor}$. Hence

$$t \in [t_{n,N_{\lfloor tn \rfloor}}, t_{n,N_{\lfloor tn \rfloor}+1}) \quad \text{for } t > 0 \text{ and } n \geq 1. \tag{3.1}$$

## 3.2 Some almost sure results

In this subsection, we record some consequences of the ergodic theorem which we will use below. But first an elementary observation, the proof of which we omit.

**Lemma 3.1** *Let $(c_n)_{n\geq 0}$ be a sequence in $\mathbb{R}$ such that $n^{-1}c_n \to c \in \mathbb{R}$. Define a sequence of functions $C_n : [0,\infty) \to \mathbb{R}$ by letting $C_n(t) := n^{-1}c_{\lfloor tn \rfloor} - tc$. Then, for any $T > 0$, $(C_n)_{n\geq 1}$ converges to $0$ uniformly on $[0,T]$.*

For the occupation times of $Y$, we then obtain:

**Lemma 3.2 (Strong law of large numbers for occupation times)** *The occupation times $N_k$ satisfy*
***a)*** *$k^{-1}N_k \longrightarrow \mu(Y)$ a.e. on $X$ as $k \to \infty$.*
***b)*** *Moreover, for any $T > 0$,*

$$\sup_{t\in[0,T]} \left| k^{-1}N_{\lfloor tk \rfloor} - t\mu(Y) \right| \longrightarrow 0 \quad \text{a.e. on } X \quad \text{as } k \to \infty.$$

**Proof.** The first statement is immediate from the ergodic theorem. The second then follows by the preceding lemma. ■

**Lemma 3.3** *For any $T > 0$, $\lim_{n\to\infty} n^{-1} \max_{0\leq j\leq \lfloor Tn \rfloor +1}(r\circ F^j) = 0$ a.e. on $Y$.*

**Proof.** Applying the ergodic theorem to $F$ and the integrable function $r$, we get $n^{-1}\sum_{j=0}^{n-1} r \circ F^j \to \mu(Y)^{-1}$, and hence also $n^{-1}(r \circ F^n) \to 0$ a.e. on $Y$. The result follows from Lemma 3.1. ■

**Pointwise control of monotonicity behaviour.** We conclude this subsection by establishing Proposition 2.7.

**Proof of Proposition 2.7.** We may suppose without loss that the sequence $B(n)$ is nondecreasing. Since this sequence is regularly varying, $B(\widehat{\delta}n)/B(n) \to \widehat{\delta}^\gamma$ for all $\widehat{\delta} > 0$. Hence for $\delta > 0$ fixed, there are $\widehat{\delta} > 0$ and $\widehat{n} \geq 1$ s.t. $\eta\, B(h)/B(n) < \delta$ whenever $n \geq \widehat{n}$ and $h \leq \widehat{\delta}n$.

As a consequence of Lemma 3.3, there is some $\widetilde{n} \geq 1$ such that $Y_n := \{n^{-1} \max_{0\leq j\leq n}(r \circ F^j) < \widehat{\delta}\}$ satisfies $\mu_Y(Y_n^c) < \varepsilon$ for $n \geq \widetilde{n}$. In view of (2.4) we then see (using monotonicity of $B$ again) that

$$\begin{aligned} \frac{1}{B(n)}\left(\max_{0\leq j\leq n} \Phi^* \circ F^j\right) &\leq \frac{1}{B(n)}\left(\max_{0\leq j\leq n} \eta\, B(r \circ F^j)\right) \\ &\leq \frac{\eta\, B(\max_{0\leq j\leq n}(r \circ F^j))}{B(n)} < \delta \quad \text{on } Y_n \quad \text{for } n \geq \widehat{n}, \end{aligned}$$

which proves (2.2). ■

## 3.3 Convergence of $(U_n)$.

As a first step towards Theorem 2.2, we prove that switching from $\Phi_{\lfloor tn \rfloor}$ to $\Phi_{N_{\lfloor sk \rfloor}}$ preserves convergence in the Skorohod space.

**Lemma 3.4 (Convergence of $(U_n)$)** *Under the assumptions of Theorem 2.2,*

$$(U_n(s))_{s\geq 0} \overset{\mu_Y}{\Longrightarrow} (W(s\,\mu(Y)))_{s\geq 0} \quad \text{in } (\mathcal{D}[0,\infty), \mathcal{M}_1).$$

**Proof.** For $n \geq 1$ and $s \in [0,\infty)$, we let $u_n(s) := n^{-1} N_{\lfloor sn \rfloor}$. Since $\lfloor u_n(s)\, n \rfloor = N_{\lfloor sn \rfloor}$, we have

$$U_n(s) = P_n(u_n(s)) \quad \text{on } Y \text{ for } n \geq 1 \text{ and } s \geq 0. \tag{3.2}$$

We regard $U_n, P_n, W$, and $u_n$ as random elements of $(\mathcal{D}, \mathcal{M}_1) = (\mathcal{D}[0,\infty), \mathcal{M}_1)$. Note that $u_n \in \mathcal{D}_\uparrow := \{g \in \mathcal{D} : g(0) \geq 0 \text{ and } g \text{ non-decreasing}\}$. Let $u$ denote the constant random element of $\mathcal{D}$ given by $u(s)(y) := s\mu(Y)$, $s \geq 0$.

Recalling Lemma 3.2 b), we see that for $\mu_Y$-a.e. $y \in Y$ we have $u_n(.)(y) \to u(.)(y)$ uniformly on compact subsets of $[0,\infty)$. Hence, $u_n \to u$ in $(\mathcal{D}, \mathcal{M}_1)$ holds $\mu_Y$-a.e. In particular,

$$u_n \overset{\mu_Y}{\Longrightarrow} u \quad \text{in } (\mathcal{D}, \mathcal{M}_1).$$

By assumption (2.1) we also have $P_n \overset{\mathcal{L}(\mu_Y)}{\Longrightarrow} W$ in $(\mathcal{D}, \mathcal{M}_1)$. But then we automatically get

$$(P_n, u_n) \overset{\mu_Y}{\Longrightarrow} (W, u) \quad \text{in } (\mathcal{D}, \mathcal{M}_1)^2, \tag{3.3}$$

since the limit $u$ of the second component is deterministic.

The composition map $(\mathcal{D}, \mathcal{M}_1) \times (\mathcal{D}_\uparrow, \mathcal{M}_1) \to (\mathcal{D}, \mathcal{M}_1)$, $(g, v) \mapsto g \circ v$, is continuous at every pair $(g, v)$ with $v \in \mathcal{C}_{\uparrow\uparrow} := \{g \in \mathcal{D} : g(0) \geq 0$ and $g$ strictly increasing and continuous$\}$, cf. [34, Theorem 13.2.3]. As the limit $(W, u)$ in (3.3) satisfies $\Pr((W, u) \in \mathcal{D} \times \mathcal{C}_{\uparrow\uparrow}) = 1$, the standard mapping theorem for distributional convergence (cf. [34, Theorem 3.4.3]) applies to $(P_n, u_n)$, showing that

$$P_n \circ u_n \overset{\mu_Y}{\Longrightarrow} W \circ u \quad \text{in } (\mathcal{D}, \mathcal{M}_1).$$

In view of (3.2), this is what was to be proved. ■

## 3.4 Control of excursions

Passing from convergence of $(U_n)$ to convergence of $(W_n)$ requires a little preparation.

**Lemma 3.5** *(i) Let $g, g' \in \mathcal{D}[0,T]$ and $0 = T_0 < \ldots < T_m = T$. Then*

$$d_{\mathcal{M}_1,T}(g,g') \le \max_{1\le j\le m} d_{\mathcal{M}_1,[T_{j-1},T_j]}(g|_{[T_{j-1},T_j]}, g'|_{[T_{j-1},T_j]}).$$

*(ii) Let $g_j \in \mathcal{D}[T_{j-1},T_j]$ and $\bar{g}_j := 1_{[T_{j-1},T_j)}\, g_j(T_{j-1}) + 1_{\{T_j\}}\, g_j(T_j)$. Then*

$$d_{\mathcal{M}_1,[T_{j-1},T_j]}(g_j,\bar{g}_j) \le 2g_j^* + (T_j - T_{j-1}),$$

*where $g_j^* := (\sup_{T_{j-1}\le s\le t\le T_j}(g_j(s) - g_j(t))) \wedge (\sup_{T_{j-1}\le s\le t\le T_j}(g_j(t) - g_j(s)))$.*

**Proof.** The first assertion is obvious. To validate the second, assume without loss that $j = 1$ and that $g_1$ is predominantly increasing in that $g_1^* = \sup_{T_0\le s\le t\le T_1}(g_1(s) - g_1(t))$. In this case, $g_1^* = \sup_{T_0\le t\le T_1}(g_1^{\uparrow}(t) - g_1(t))$ for the nondecreasing function $g_1^{\uparrow}(t) := \sup_{T_0\le s\le t} g_1(s)$. Therefore,

$$d_{\mathcal{M}_1,[T_0,T_1]}(g_1, g_1^{\uparrow}) \le \|g_1 - g_1^{\uparrow}\| = g_1^*.$$

Letting $\bar{g_1}^{\uparrow} := 1_{[T_0,T_1)}\, g_1^{\uparrow}(T_0) + 1_{\{T_1\}}\, g_1^{\uparrow}(T_1)$, it is clear that

$$d_{\mathcal{M}_1,[T_0,T_1]}(\bar{g}_1^{\uparrow}, \bar{g}_1) \le \|\bar{g}_1^{\uparrow} - \bar{g}_1\| = |g_1^{\uparrow}(T_1) - g_1(T_1)| \le g_1^*.$$

Finally, we check that

$$d_{\mathcal{M}_1,[T_0,T_1]}(g_1^{\uparrow}, \bar{g}_1^{\uparrow}) \le T_1 - T_0.$$

To this end, we refer to Figure 1 where $\Gamma_1 = \Gamma(g_1^{\uparrow})$ and $\Gamma_2 = \Gamma(\bar{g}_1^{\uparrow})$ represent the completed graphs of $g_1^{\uparrow}$ and $\bar{g}_1^{\uparrow}$ respectively. Here $\Gamma_2$ consists of one horizontal line segment followed by one vertical segment. The picture of $\Gamma_1$ is schematic, it may also contain horizontal and vertical line segments.

Choose $C$ on the graph of $\Gamma_1$ that is equidistant from $AD$ and $DB$ and let $E$ be the point on $DB$ that is the same height as $C$. Choose parametrizations $G_i = (\lambda_i, \gamma_i)$ of $\Gamma_i$, $i = 1, 2$, satisfying

(i) $G_1(0) = G_2(0) = A$, $\quad G_1(1) = G_2(1) = B$,

(ii) $G_1(\frac{1}{2}) = C$, $\quad G_2(\frac{1}{2}) = E$,

(iii) $\gamma_1(t) = \gamma_2(t)$ for all $t \in [\frac{1}{2}, 1]$.

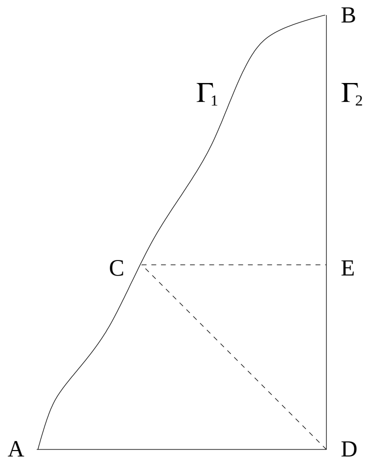


Figure 1: A monotone excursion

Automatically $\|\lambda_1 - \lambda_2\| \leq |AD| = T_1 - T_0$ and by construction $\|\gamma_1 - \gamma_2\| \leq |DE| \leq |AD|$, as required. ■

As a consequence, we obtain:

**Lemma 3.6** $d_{\mathcal{M}_1,T}(W_n, U_n) \leq \max_{0\leq j\leq \lfloor Tn\rfloor+1}(n^{-1}r + 2B(n)^{-1}\Phi^*) \circ F^j$.

**Proof.** Let $y \in Y$ and decompose $[0,T]$ according to the consecutive excursions, letting $T_j := t_{n,j}(y) \wedge T$, $j \leq m := \lfloor Tn\rfloor + 1$. Consider $g(t) := W_n(t)(y)$, $t \in [0,T]$. If we set $g_j := g|_{[T_{j-1},T_j]}$, then $\bar{g}_j$ as defined in Lemma 3.5 coincides with $U_n(.)(y)|_{[T_{j-1},T_j]}$, so that

$$d_{\mathcal{M}_1,T}(W_n(.)(y), U_n(.)(y)) \leq \max_{0\leq j\leq m} d_{\mathcal{M}_1,[T_{j-1},T_j]}(g_j, \bar{g}_j).$$

But $T_j - T_{j-1} = n^{-1}r \circ F^j$, and since $g_j(s) - g_j(t) = B(n)^{-1}(\phi_{\ell'} - \phi_\ell) \circ F^j(y)$, for suitable $0 \leq \ell' \leq \ell \leq r$, we see that Lemma 3.5 gives

$$d_{\mathcal{M}_1,[T_{j-1},T_j]}(g_j, \bar{g}_j) \leq (n^{-1}r + 2B(n)^{-1}\Phi^*) \circ F^j,$$

as required. ■

**Proof of Theorem 2.2.** Fix any $T > 0$. By Lemma 3.3, $n^{-1}\max_{0\leq j\leq \lfloor Tn\rfloor+1}(r \circ F^j) \to 0$ a.e. and by assumption (2.2) $B(n)^{-1}\max_{0\leq j\leq \lfloor Tn\rfloor+1}(\Phi^* \circ F^j) \overset{\mu_Y}{\Longrightarrow} 0$. Hence Lemma 3.6 guarantees that

$$d_{\mathcal{M}_1,T}(W_n, U_n) \overset{\mu_Y}{\Longrightarrow} 0. \tag{3.4}$$

Recall also from Lemma 3.4 that

$$(U_n(s))_{0\le s\le T} \overset{\mu_Y}{\Longrightarrow} (W(s\,\mu(Y)))_{0\le s\le T} \quad \text{in } (\mathcal{D}[0,T],\mathcal{M}_1). \tag{3.5}$$

It follows (see [8, Theorem 3.1]) from (3.4) and (3.5) that

$$(W_n(s))_{0\le s\le T} \overset{\mu_Y}{\Longrightarrow} (W(s\,\mu(Y)))_{0\le s\le T} \quad \text{in } (\mathcal{D}[0,T],\mathcal{M}_1).$$

This immediately gives $(W_n(s))_{t\ge 0} \overset{\mu_Y}{\Longrightarrow} (W(s\,\mu(Y)))_{s\ge 0}$ in $(\mathcal{D}[0,\infty),\mathcal{M}_1)$. Strong distributional convergence as asserted in (2.3) follows via Proposition 2.8. ■

# 4 Examples

We continue to suppose that $f$ is an ergodic m.p.t. on a probability space $(X,\mu)$ with first return map $F = f^r : Y \to Y$ where $\mu(Y) > 0$. Suppose further that the induced map $F : Y \to Y$ is Gibbs-Markov with ergodic invariant probability measure $\mu_Y$ and partition $\beta$, and that $r|_a$ is constant for each $a \in \beta$. Let $\phi : X \to \mathbb{R}$ be an $L^\infty$ mean zero observable, with induced observable $\Phi : Y \to \mathbb{R}$.

**Theorem 4.1** *Suppose that $\phi$ is constant on $f^\ell a$ for every $a \in \beta$ and $\ell \in \{0,\dots,r|_a - 1\}$. If $\Phi$ lies in the domain of an $\alpha$-stable law, then $\Phi$ satisfies the WIP in $(D,\mathcal{J}_1)$ with $B(n) = n^{1/\alpha}$. If in addition condition* (2.2) *holds, then $\phi$ satisfies the WIP in $(D,\mathcal{M}_1)$.*

**Proof.** By [1], $n^{-1/\alpha}\Phi_n$ converges in distribution to the given stable law. The assumptions guarantee that the induced observable $\Phi$ is constant on each $Y_j$. Hence we can apply Tyran-Kamińska [33, Corollary 4.1] to deduce that $\Phi$ satisfies the corresponding $\alpha$-stable WIP in $(D,\mathcal{J}_1)$. In particular, condition (2.1) is satisfied. The final statement follows from Theorem 2.2. ■

For certain examples, including Pomeau-Manneville intermittency maps, we can work with general Hölder observables, thus improving upon [33, Example 4.1]. The idea is to decompose the observable $\phi$ into a piecewise constant observable $\phi_0$ and a Hölder observable $\tilde{\phi}$ in such a way that only $\phi_0$ "sees" the source of the anomalous behaviour.

In the remainder of this section, we carry out this procedure for the maps (1.1) and thereby prove Theorem 1.1. (Lemma 4.2 and Proposition 4.3 below hold in the general context of induced Gibbs-Markov maps.)

Fix $\theta \in (0,1)$ and let $d_\theta$ denote the symbolic metric on $Y$, so $d_\theta(x,y) = \theta^{s(x,y)}$ where $s(x,y)$ is the least integer $n \ge 0$ such that $F^n x, F^n y$ lie in distinct elements of $\beta$. An observable $\Phi : Y \to \mathbb{R}$ is *piecewise Lipschitz* if $D_a(\Phi) := \sup_{x,y\in a, x\neq y} |\Phi(x) - \Phi(y)|/d_\theta(x,y) < \infty$ for each $a \in \beta$, and *Lipschitz* if $\|\Phi\|_\theta = |\Phi|_\infty + \sup_{a\in\beta} D_a(\Phi) < \infty$. The space Lip of Lipschitz observables $\Phi : Y \to \mathbb{R}$ is a Banach space. Note that $\Phi$ is integrable with $\sum_{a\in\beta} \mu_Y(a) D_a(\Phi) < \infty$ if and only if $\sum_{a\in\beta} \mu_Y(a)\|1_a\Phi\|_\theta < \infty$.

Let $L$ denote the transfer operator for $F : Y \to Y$.

**Lemma 4.2** ***a)*** *The essential spectral radius of* $L : \mathrm{Lip} \to \mathrm{Lip}$ *is at most* $\theta$.
***b)*** *Suppose that* $\Phi : Y \to \mathbb{R}$ *is a piecewise Lipschitz observable satisfying* $\sum_{a\in\beta} \mu_Y(a)\|1_a\Phi\|_\theta < \infty$. *Then* $L\Phi \in \mathrm{Lip}$.

**Proof.** This is standard. See for example [1, Theorem 1.6] for part (a) and [25, Lemma 2.2] for part (b). ■

**Proposition 4.3** *Let* $\Phi : Y \to \mathbb{R}$ *be a piecewise Lipschitz mean zero observable lying in* $L^p$*, for some* $p \in (1,2)$. *Assume that* $\sum_{a\in\beta} \mu_Y(a)\|1_a\Phi\|_\theta < \infty$. *Then* $\max_{j=0,\dots,n-1} n^{-\gamma}\Phi_j \to_d 0$ *for all* $\gamma > 1/p$.

**Proof.** Suppose first that $F$ is weak mixing (this assumption is removed below). Then $L : \mathrm{Lip} \to \mathrm{Lip}$ has no eigenvalues on the unit circle except for the simple eigenvalue at 1 (corresponding to constant functions). By Lemma 4.2(a), there exists $\tau < 1$ such that the remainder of the spectrum of $L$ lies strictly inside the ball of radius $\tau$. In particular, there is a constant $C > 0$ such that $\|L^n v - \int v\, d\mu_Y\| \le C\tau^n\|v\|$ for all $v \in \mathrm{Lip}$, $n \ge 1$.

By Lemma 4.2(b), $L\Phi \in \mathrm{Lip}$. Hence $\chi = \sum_{j=1}^\infty L^j\Phi \in \mathrm{Lip}$. Following Gordin [16], write $\Phi = \hat\Phi + \chi \circ F - \chi$. Then $\hat\Phi \in L^p$ (since $\chi \in \mathrm{Lip}$ and $\Phi \in L^p$). Applying $L$ to both sides and noting that $L(\chi\circ F) = \chi$, we obtain that $L\hat\Phi = 0$. It follows that the sequence $\{\hat\Phi_n;\, n \ge 1\}$ defines a reverse martingale sequence.

By Burkholder's inequality [9, Theorem 3.2],

$$
\begin{aligned}
|\hat\Phi_n|_p &\ll \Big\| \Big(\sum_{j=1}^n \hat\Phi^2 \circ F^j\Big)^{1/2} \Big\|_p = \Big( \int \Big(\sum_{j=1}^n \hat\Phi^2\circ F^j\Big)^{p/2}\Big)^{1/p} \\
&\le \Big( \int \sum_{j=1}^n |\hat\Phi|^p \circ F^j \Big)^{1/p} = |\hat\Phi|_p n^{1/p}.
\end{aligned}
$$

By Doob's inequality [13] (see also [9, Equation (1.4), p. 20], $|\max_{j=0,\dots,n-1}\hat\Phi_j|_p \ll |\hat\Phi_n|_p \ll n^{1/p}$. By Markov's inequality, for $\epsilon > 0$ fixed, $\mu_Y(|\max_{j=0,\dots,n-1}\hat\Phi_j| \ge$

$\epsilon n^\gamma) \le |\max_{j=0,\dots,n-1}\hat\Phi_j|_p^p/(\epsilon^p n^{\gamma p}) \ll n^{-(\gamma p-1)} \to 0$ as $n\to\infty$. Hence $\max_{j=0,\dots,n-1} n^{-\gamma}\hat\Phi_j \to_d 0$. Since $\Phi$ and $\hat\Phi$ differ by a bounded coboundary, $\max_{j=0,\dots,n-1} n^{-\gamma}\Phi_j \to_d 0$ as required.

It remains to remove the assumption about eigenvalues (other than 1) for $L$ on the unit circle. Suppose that there are $k$ such eigenvalues $e^{i\omega_\ell}$, $\omega_\ell \in (0,2\pi)$, $\ell = 1,\dots,k$ (including multiplicities). Then we can write $\Phi = \Psi_0 + \sum_{\ell=1}^k \Psi_\ell$ where $\|L^n\Psi_0\|_\theta \le C\tau^n\|\Psi_0\|_\theta$ and $L\Psi_\ell = e^{i\omega_\ell}\Psi_\ell$. In particular, the above argument applies to $\Psi_0$, while $L\Psi_\ell = e^{i\omega_\ell}\Psi_\ell$, $\ell = 1,\dots k$.

A simple argument (see [24]) shows that $\Psi_\ell\circ F = e^{-i\omega_\ell}\Psi_\ell$ for $\ell = 1,\dots,k$, so that $|\sum_{j=1}^n \Psi_\ell\circ F^j|_\infty \le 2|e^{i\omega_\ell}-1|^{-1}|\Psi_\ell|_\infty$ which is bounded in $n$. Hence the estimate for $\Phi$ follows from the one for $\Psi_0$. ■

**Proof of Theorem 1.1.** We verify the hypotheses of Theorem 2.2. A convenient inducing set for the maps (1.1) is $Y = [\frac12, 1]$. Let $\phi_0 = \phi(0) - \mu(Y)^{-1}\phi(0)1_Y$. (The first term is the important one, and the second term is simply an arbitrary choice that ensures that $\phi_0$ has mean zero while preserving the piecewise constant requirement in Theorem 4.1.) Write $\phi = \phi_0 + \tilde\phi$ and note that $\tilde\phi$ is a mean zero piecewise Hölder observable vanishing at 0. We have the corresponding decomposition $\Phi = \Phi_0 + \tilde\Phi$ for the induced observables. By Theorem 4.1, $\Phi_0$ satisfies the WIP (in the $\mathcal{J}_1$ topology).

Let $\eta$ denote the Hölder exponent of $\phi$. By the proof of [18, Theorem 1.3], $\tilde\phi$ induces to a piecewise Lipschitz mean zero observable $\tilde\Phi$ satisfying $\sum_{a\in\beta}\mu_Y(a)\|1_a\tilde\Phi\|_\theta < \infty$ for suitably chosen $\theta$. Moreover [18] shows that $\tilde\Phi$ lies in $L^2$ provided that $\eta > \gamma - \frac12$. Exactly the same argument shows that $\tilde\Phi$ lies in $L^p$ provided $\eta > \gamma - \frac1p$. In particular, for any $\eta > 0$, there exists $p > 1/\gamma$ such that $\tilde\Phi \in L^p$. Since we are normalising by $B(n) = n^{1/\alpha} = n^\gamma$, it follows from Proposition 4.3 that $\tilde\Phi$ does not contribute to the WIP.

Combining the results for $\Phi_0$ and $\tilde\Phi$, we deduce that $\Phi$ satisfies the WIP (in the $\mathcal{J}_1$ topology). In particular, condition (2.1) is satisfied.

It remains to verify condition (2.4). In fact, we show that $\Phi^*$ is bounded. Suppose that $\phi(0) > 0$ (the case $\phi(0) < 0$ is treated similarly). Choose $\epsilon > 0$ such that $\phi > 0$ on $[0,\epsilon]$. Define the decreasing sequence $x_n \in (0,\frac12)$ where $f(x_n) = x_{n-1}$, $x_1 = \frac12$, and let $k$ be such that $x_n \in (0,\epsilon)$ for all $n \ge k$. Then for $y\in Y$, $f^\ell y \in [0,\epsilon]$ for $0 \le \ell \le r(y) - k$.

Now observe that

(i) $\phi_{\ell'}(y) - \phi_\ell(y) \le 0$ for $1 \le \ell' \le \ell \le r(y) - k$,

(ii) $\phi_{\ell'}(y) - \phi_\ell(y) \le k|\phi|_\infty$ for $r(y) - k \le \ell' \le \ell \le r(y)$,

(iii) $\phi_{\ell'}(y) - \phi_\ell(y) \le \phi_{r(y)-k}(y) - \phi_\ell(y) \le k|\phi|_\infty$ for $1 \le \ell' \le r(y) - k \le \ell \le r(y)$,

Hence

$$\begin{aligned}\Phi^*(y) &\le \max_{0\le\ell'\le\ell\le r(y)} (\phi_{\ell'}(y) - \phi_\ell(y)) \\ &\le |\phi|_\infty + \max_{1\le\ell'\le\ell\le r(y)} (\phi_{\ell'}(y) - \phi_\ell(y)) \le (k+1)|\phi|_\infty,\end{aligned}$$

as required. ■

**Remark 4.4** *The arguments in the proof of Theorem 1.1 apply to a much wider class of examples, including intermittent maps with neutral periodic points or with multiple neutral fixed/periodic points. In such cases, condition* (2.2) *is again automatically satisfied.*

**Acknowledgements** The research of IM was supported in part by EPSRC Grant EP/F031807/1 held at the University of Surrey. We are very grateful to the referee for very helpful suggestions that led to a significantly simplified proof of the main result in this paper.

# References

[1] J. Aaronson and M. Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.* **1** (2001) 193–237.

[2] F. Avram and M. S. Taqqu. Weak convergence of sums of moving averages in the $\alpha$-stable domain of attraction. *Ann. Probab.* **20** (1992) 483–503.

[3] P. Bálint, N. Chernov and D. Dolgopyat. Limit theorems for dispersing billiards with cusps. *Comm. Math. Phys.* **308** (2011) 479–510.

[4] P. Bálint and S. Gouëzel. Limit theorems in the stadium billiard. *Comm. Math. Phys.* **263** (2006) 461–512.

[5] P. Bálint and I. Melbourne. Decay of correlations and invariance principles for dispersing billiards with cusps, and related planar billiard flows. *J. Stat. Phys.* **133** (2008) 435–447.

[6] G. Ben Arous and J. Černý. Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.* **35** (2007) 2356–2384.

[7] M. Benedicks and L.-S. Young. Markov extensions and decay of correlations for certain Hénon maps. *Astérisque* (2000) no. 261, 13–56.

[8] P. Billingsley. *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999.

[9] D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Probability* **1** (1973) 19–42.

[10] J.-P. Conze and S. Le Borgne. Méthode de martingales et flow géodésique sur une surface de courbure constante négative. *Ergodic Theory Dynam. Systems* **21** (2001) 421–441.

[11] J. Dedecker and F. Merlevède. Weak invariance principle and exponential bounds for some special functions of intermittent maps. *High dimensional probability* **5** (2009) 60–72.

[12] M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory Dynam. Systems* **4** (1984) 541–552.

[13] J. L. Doob. *Stochastic processes.* Wiley, New York, 1953.

[14] G. K. Eagleson. Some simple conditions for limit theorems to be mixing. *Teor. Verojatnost. i Primenen* **21** (1976) 653–660.

[15] M. J. Field, I. Melbourne and A. Török. Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergodic Theory Dynam. Systems* **23** (2003) 87–110.

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

[17] G. A. Gottwald and I. Melbourne. Central limit theorems and suppression of anomalous diffusion for systems with symmetry. Preprint.

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

[19] S. Gouëzel. Statistical properties of a skew product with a curve of neutral points. *Ergodic Theory Dynam. Systems* **27** (2007) 123–151.

[20] S. Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.* **38** (2010) 1639–1671.

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

[22] M. Holland and I. Melbourne. Central limit theorems and invariance principles for Lorenz attractors. *J. London Math. Soc.* **76** (2007) 345–364.

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

[24] I. Melbourne and M. Nicol. Statistical properties of endomorphisms and compact group extensions. *J. London Math. Soc.* **70** (2004) 427–446.

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

[26] I. Melbourne and M. Nicol. A vector-valued almost sure invariance principle for hyperbolic dynamical systems. *Ann. Probability* **37** (2009) 478–505.

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

[28] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.* **144** (2004) 191–209.

[29] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980) 189–197.

[30] M. Ratner. The central limit theorem for geodesic flows on $n$-dimensional manifolds of negative curvature. *Israel J. Math.* **16** (1973) 181–197.

[31] A. V. Skorohod. Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen.* **1** (1956) 289–319.

[32] D. Szász and T. Varjú. Limit laws and recurrence for the planar Lorentz process with infinite horizon. *J. Stat. Phys.* **129** (2007) 59–80.

[33] M. Tyran-Kamińska. Weak convergence to Lévy stable processes in dynamical systems. *Stoch. Dyn.* **10** (2010) 263–289.

[34] W. Whitt. *Stochastic-process limits.* Springer Series in Operations Research, Springer-Verlag, New York, 2002.

[35] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.

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

[37] R. Zweimüller. Stable limits for probability preserving maps with indifferent fixed points. *Stoch. Dyn.* **3** (2003) 83–99.

[38] R. Zweimüller. Mixing limit theorems for ergodic transformations. *J. Theoret. Probab.* **20** (2007) 1059–1071.