Convergence to a Lévy process in the Skorohod $\mathcal{M}_1$ and $\mathcal{M}_2$ topologies for nonuniformly hyperbolic systems, including billiards with cusps

Ian Melbourne[*] Paulo Varandas[†]

21 September 2018. Updated 15 May 2019

Abstract

We prove convergence to a Lévy process for a class of dispersing billiards with cusps. For such examples, convergence to a stable law was proved by Jung & Zhang. For the corresponding functional limit law, convergence is not possible in the usual Skorohod $\mathcal{J}_1$ topology. Our main results yield elementary geometric conditions for convergence (i) in $\mathcal{M}_1$, (ii) in $\mathcal{M}_2$ but not $\mathcal{M}_1$.

In general, we show for a large class of nonuniformly hyperbolic systems how to deduce functional limit laws once convergence to the corresponding stable law is known.

1 Introduction

It is by now well-known that deterministic dynamical systems often satisfy statistical limit theorems from classical probability theory. Following Sinai [42], a rich source of examples is provided by dispersing billiards [15] which are based on deterministic Lorentz gas models [33]. By [11, 12], the central limit theorem (CLT) and functional central limit theorem or weak invariance principle (WIP) hold for planar periodic dispersing billiards. The CLT asserts convergence to a normal distribution and the WIP deals with convergence to the corresponding Brownian notion. These limit laws also hold for Sinai billiards where the boundary of the table is a simple closed curve consisting of finitely many C^3 convex inwards curves with nonvanishing curvature and nonzero angles at corner points [19]. For billiards with cusps (corner points with zero

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

[†]Departamento de Matemática, Universidade Federal da Bahia, 40170-110 Salvador, Brazil. pcvarand@gmail.com

angle), the CLT and WIP were obtained by [3] but with the weakly superdiffusive normalization $(n \log n)^{1/2}$ instead of the standard diffusion rate $n^{1/2}$.

Stronger superdiffusion rates $n^{1/\alpha}$, $\alpha < 2$, with limiting fluctuations governed by an α-stable Lévy process rather than a Brownian motion, have been the focus of much attention across the physical sciences. See for example [5, 22, 24, 31, 37, 38, 39, 41, 44, 47] and references therein. Whereas Brownian motions are continuous processes with finite variance, Lévy processes exhibit jumps of all sizes and have infinite variance.

In this paper, we show for the first time that convergence to a Lévy process occurs in dispersing billiards. The example is elementary to write down and the mechanism for superdiffusion is intuitively transparent. Moreover, our analysis casts light on the mode of convergence, an aspect which has received little attention previously.

Recently, Jung & Zhang [30] considered a class of billiards with cusps where there is vanishing curvature at the cusp and proved convergence to totally skewed α-stable laws with $\alpha \in (1, 2)$. However, they were unable to prove the functional WIP version of their limit law (i.e. weak convergence to the corresponding α-stable Lévy process).

In this paper, as part of a general framework including [30], we show how to pass from the stable law to the WIP. The standard $\mathcal{J}_1$ Skorohod topology [43, 47] is always too strong for these examples, but we obtain convergence in the $\mathcal{M}_1$ and $\mathcal{M}_2$ topologies. The definition of these Skorokhod topologies is recalled in Appendix B.

It is well-known that the $\mathcal{J}_1$ topology is often too strong, and there are many natural examples where the $\mathcal{M}_1$ topology is the appropriate one, see for example [2, 7, 37, 47]. Indeed, Whitt [47, p. xii] writes

> Thus, while the $\mathcal{J}_1$ topology sometimes cannot be used, the $\mathcal{M}_1$ topology can almost always be used. Moreover, the extra strength of the $\mathcal{J}_1$ topology is rarely exploited. Thus, we would be so bold as to suggest that, *if only one topology on the function space D is to be considered, then it should be the $\mathcal{M}_1$ topology.*

Jakubowski [27] writes

> All these reasons bring interest also to the weaker Skorokhod's topologies $\mathcal{J}_2$, $\mathcal{M}_1$ and $\mathcal{M}_2$. Among them practically only the topology $\mathcal{M}_1$ proved to be useful.

Nevertheless, in this paper we provide natural examples where the $\mathcal{M}_1$ topology is too strong and the $\mathcal{M}_2$ topology is the appropriate one. The only previous such example that we know of can be found in [6].

Example 1.1 We consider the Jung & Zhang example [30] consisting of a planar dispersing billiard with a cusp at a flat point. A standard reference for background material on billiards is [15].

The billiard table $Q \subset \mathbb{R}^2$ has a boundary consisting of a finite number of C^3 curves Γ_i, $i = 1, \dots, n_0$, where $n_0 \geq 3$ with a cusp formed by two of these curves Γ_1, Γ_2. In coordinates $(s, z) \in \mathbb{R}^2$, the cusp lies at $(0, 0)$ and Γ_1, Γ_2 are tangent to the s-axis at $(0,0)$. Moreover, close to $(0,0)$, we have $\Gamma_1 = \{(s, \beta^{-1}s^\beta)\}$, $\Gamma_2 = \{(s, -\beta^{-1}s^\beta)\}$, where $\beta > 2$. See Figure 1. [1]

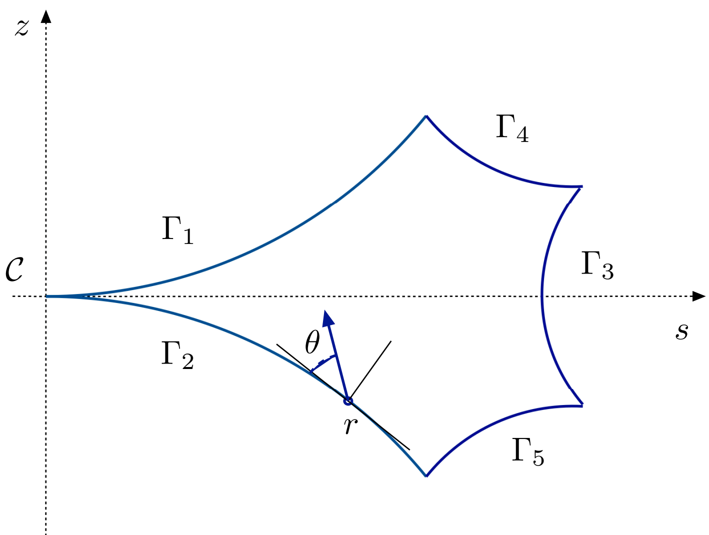

Figure 1: Billiard with a cusp at a flat point as studied by Jung & Zhang.

The phase space of the billiard map (or collision map) T is given by $\Lambda = \partial Q \times [0, \pi]$, with coordinates (r, θ) where r denotes arc length along ∂Q and θ is the angle between the tangent line of the boundary and the collision vector in the clockwise direction. There is a natural ergodic invariant probability measure $d\mu = (2|\partial Q|)^{-1} \sin\theta \, dr \, d\theta$ on Λ, where $|\partial Q|$ is the length of ∂Q.

In configuration space, the cusp is a single point $(0,0) = \Gamma_1 \cap \Gamma_2$. Let $r' \in \Gamma_1$ and $r'' \in \Gamma_2$ be the arc length coordinates of $(0,0)$. Then in phase space Λ, the cusp is the union of two line segments

$$\mathcal{C} = \{(r', \theta) : 0 \le \theta \le \pi\} \cup \{(r'', \theta) : 0 \le \theta \le \pi\}.$$

Let $v : \Lambda \to \mathbb{R}$ be a Hölder continuous observable with $\int_\Lambda v \, d\mu = 0$ and define[2]

$$I_v(s) = \frac{1}{2} \int_0^s \{v(r', \theta) + v(r'', \pi - \theta)\} (\sin\theta)^{1/\alpha} \, d\theta, \quad s \in [0, \pi]. \tag{1.1}$$

where $\alpha = \frac{\beta}{\beta - 1} \in (1, 2)$. Suppose that $I_v(\pi) > 0$ (the case $I_v(\pi) < 0$ is identical with the obvious modifications). Let G be the totally skewed α-stable law with

[1] In [30], it is assumed in addition that the trajectory running out of the cusp along the s-axis hits Γ_3 perpendicularly, but this was only done for convenience and is not present in [29].

[2] Our definitions differ from those in [30] by constant factors, leading to simpler formulas in Section 8.

characteristic function

$$\mathbb{E}(e^{iuG}) = \exp\{-|u|^\alpha\sigma^\alpha(1-i\,\mathrm{sgn}\,u\tan\tfrac{\pi\alpha}{2})\}, \quad \sigma^\alpha = (\beta|\partial Q|2^{\alpha-1})^{-1}I_v(\pi)^\alpha\Gamma(1-\alpha)\cos\tfrac{\pi\alpha}{2}.$$

Jung & Zhang [30, Theorem 1.1] prove:

Theorem 1.2 $n^{-1/\alpha}\sum_{j=0}^{n-1} v\circ T^j \to_d G$. ∎

Let $D[0,\infty)$ denote the set of real-valued càdlàg functions (right-continuous with left-hand limits) on $[0,\infty)$, and let $W \in D[0,\infty)$ be the α-stable Lévy process with $W(1) =_d G$. Define

$$W_n : \Lambda \to D[0,\infty), \qquad W_n(t) = n^{-1/\alpha}\textstyle\sum_{j=0}^{[nt]-1} v\circ T^j.$$

Since the increments of W_n are bounded by $n^{-1/\alpha}|v|_\infty$ and W has jumps with probability one, W_n does not converge to W in the $\mathcal{J}_1$ topology. However, the weaker $\mathcal{M}_1$ topology allows an amalgamation of numerous small increments for W_n to approximate a single jump for W. This is analogous to the situation for intermittent maps of Pomeau-Manneville type [40] studied in [37]. In contrast to [37], convergence in $\mathcal{M}_1$ is not automatic. Instead, there is a simple geometric condition on $v|_\mathcal{C}$ which characterizes convergence in $\mathcal{M}_1$:

Theorem 1.3 *$W_n \to_w W$ in $(D[0,\infty),\mathcal{M}_1)$ if and only if $v(r',\theta) + v(r'',\pi-\theta) \geq 0$ for all $\theta \in [0,\pi]$. (Equivalently, $s \mapsto I_v(s)$ is nondecreasing on $[0,\pi]$.)*

We also have a sufficient condition for convergence in the even weaker $\mathcal{M}_2$ topology.

Theorem 1.4 *If $I_v(s) \in [0, I_v(\pi)]$ for all $s \in [0,\pi]$, then $W_n \to_w W$ in $(D[0,\infty),\mathcal{M}_2)$.*

It is now easy to construct a Hölder continuous mean zero observable $v : \Lambda \to \mathbb{R}$ so that convergence holds in $\mathcal{M}_2$ but not in $\mathcal{M}_1$. For example, choose v so that $v(r',\theta) + v(r'',\pi-\theta)$ is positive on $[0,\frac{\pi}{3})\cup(\frac{2\pi}{3},\pi]$ and negative on $(\frac{\pi}{3},\frac{2\pi}{3})$. See Figure 2(b). The change of sign violates the condition for $\mathcal{M}_1$-convergence in Theorem 1.3, while it is clear that if v is small enough on $(\frac{\pi}{3},\frac{2\pi}{3})$ comparable to its values on $[0,\frac{\pi}{3})\cup(\frac{2\pi}{3},\pi]$, then the condition for $\mathcal{M}_2$-convergence in Theorem 1.4 is satisfied.

Remark 1.5 (a) After writing this paper, we learned of independent work of [29] on billiards with several cusps at flat points. They considered the case where v has constant sign near each cusp and proved convergence to a Lévy process in the $\mathcal{M}_1$ topology.

(b) In a previous version of this paper, we conjectured that the condition in Theorem 1.4 for convergence in the $\mathcal{M}_2$ topology is necessary and sufficient. This has

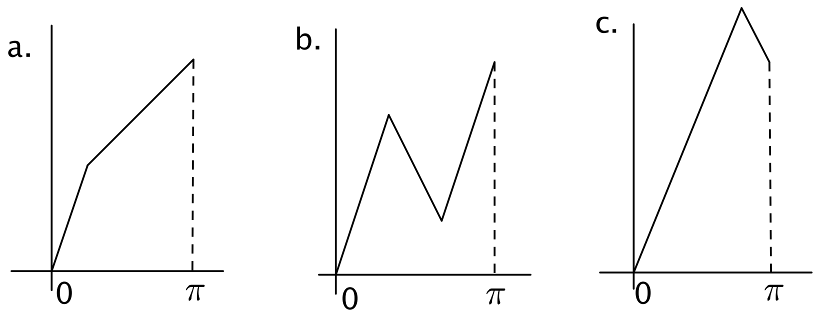

Figure 2: Different possible shapes of the function I_v for the Jung & Zhang example: (a) WIP holds in the $\mathcal{M}_1$ (hence also in the $\mathcal{M}_2$) topology; (b) WIP holds in the $\mathcal{M}_2$ topology but not in the $\mathcal{M}_1$ topology; (c) the WIP does not hold even in the $\mathcal{M}_2$ topology.

now been shown to be the case in [28]. An interesting open question is to consider alternative weaker modes of convergence in situations such as Figure 2(c) where $\mathcal{M}_2$-convergence fails. (Such a weakening entails diminishing the class of continuous functionals under which weak convergence is preserved. For example, weak convergence in any of the Skorokhod topologies mentioned above implies weak convergence of the supremum process, i.e. $\sup_{[0,t]} W_n \to_w \sup_{[0,t]} W$, see [47, Section 13.4], but this appears unlikely in the situation of Figure 2(c).)

Strategy of proof The proof of Theorems 1.3 and 1.4 fits into a general framework [18, 34] which has been used to study large classes of examples from billiards specifically and nonuniformly hyperbolic dynamical systems in general. This framework is described in Section 2. (It includes the setting of intermittent maps as a very special case, see Remark 3.7.) Let $X \subset \Lambda$ be a cross-section with first return time $\varphi : X \to \mathbb{Z}^+$ and first return map $f = T^\varphi : X \to X$ as in (2.4). In Example 1.1, $X = (\Gamma_3 \cup \cdots \cup \Gamma_{n_0}) \times [0, \pi]$. We require that f is modelled by a Young tower with exponential tails [48] over a "uniformly hyperbolic" subset $Y \subset X \subset \Lambda$. Associated to the observable $v : \Lambda \to \mathbb{R}$, we have the induced observable $V = \sum_{\ell=0}^{\varphi-1} v \circ T^\ell : X \to \mathbb{R}$. Also, associated to φ, V on X there are induced versions φ^Y, V^Y on Y.

The key argument of [30, Theorem 3.1] proves a stable law for $\varphi : X \to \mathbb{Z}^+$. Our approach deduces the WIP for v on Λ from the stable law for φ on X. The idea is to first induce the stable law for φ to a stable law for φ^Y on Y. Since the dynamics on Y is very well-understood, this leads via results of Gouëzel [26] and Tyran-Kamińska [46] to convergence to a Lévy process in the $\mathcal{J}_1$ topology for φ^Y and thereby V^Y. The WIP for V^Y uninduces to convergence in the $\mathcal{M}_1$ topology for V

on X. Under certain conditions, this uninduces to convergence in the $\mathcal{M}_1$ or $\mathcal{M}_2$ topology for v. The strategy can be represented diagrammatically as follows:

Λ		X		Y
		stable law for φ	$\Longrightarrow$	stable law for φ^Y
				$\Downarrow$
				WIP in $\mathcal{J}_1$ for φ^Y
				$\Downarrow$
WIP in $\mathcal{M}_1$ / $\mathcal{M}_2$ for v	$\Longleftarrow$	WIP in $\mathcal{M}_1$ for V	$\Longleftarrow$	WIP in $\mathcal{J}_1$ for V^Y

The remainder of the paper is organized as follows. In Section 2, we consider the Chernov-Markarian-Zhang framework where the underlying system has a first return map modelled by a Young tower with exponential tails. In Section 3, we state our main results on stable laws and WIPs for systems with a Chernov-Markarian-Zhang structure. In Section 4, we state and prove a purely probabilistic result on uninducing WIPs in the $\mathcal{M}_1$ or $\mathcal{M}_2$ topology, extending a result of [37]. Section 5 contains limit laws for the return times φ and φ^Y, and Section 6 contains some estimates for induced Hölder observables. These are combined in Section 7 to prove our main results from Section 3. In Section 8, we return to Example 1.1, proving Theorems 1.3 and 1.4 as well as giving a streamlined proof of Theorem 1.2.

Notation We use the "big O" and $\ll$ notation interchangeably, writing $a_n = O(b_n)$ or $a_n \ll b_n$ if there is a constant $C > 0$ such that $a_n \le Cb_n$ for all $n \ge 1$. Also, we write $a_n \approx b_n$ if $a_n \ll b_n \ll a_n$. As usual, $a_n \sim b_n$ as $n \to \infty$ means that $\lim_{n\to\infty} a_n/b_n = 1$.

For $a, b \in \mathbb{R}$, we write $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$.

Recall that a sequence $b_n \in (0, \infty)$ is *regularly varying of index* $p > 0$ if $b_{\lambda n}/b_n \to \lambda^p$ as $n \to \infty$ for all $\lambda \ge 1$.

2 Preliminaries

In this section, we recall the Chernov-Markarian-Zhang framework [18, 34]. Roughly speaking, this means that there is a convenient first return map that is modelled by a Young tower with exponential tails [48]. The full details from Young [48] are not required for our main theorems, so we recall here only those aspects that are needed.

2.1 Towers and return maps

In this subsection, we review a purely measure-theoretic framework of tower maps and return maps that arises throughout this paper.

Let $F : Y \to Y$ be a measure-preserving transformation on a probability space (Y, μ_Y), and let $\tau : Y \to \mathbb{Z}^+$ be integrable. The tower $\Delta = Y^\tau$ and tower map $\hat{f} : \Delta \to \Delta$ are given by

$$\Delta = \{(y, \ell) \in Y \times \mathbb{Z} : 0 \le \ell < \tau(y)\}, \quad \hat{f}(y, \ell) = \begin{cases} (y, \ell + 1) & \ell \le \tau(y) - 2 \\ (Fy, 0) & \ell = \tau(y) - 1 \end{cases}. \tag{2.1}$$

Define $\bar{\tau} = \int_Y \tau \, d\mu_Y$. Then $\mu_\Delta = (\mu_Y \times \text{counting})/\bar{\tau}$ is an $\hat{f}$-invariant probability measure on Δ. We call $\hat{f} : \Delta \to \Delta$ the *tower with base map F and return time τ*.

Next, let $f : X \to X$ be a measure-preserving transformation on a probability space (X, μ_X), and $Y \subset X$ a positive measure subset. Let $\tau : Y \to \mathbb{Z}^+$ be measurable such that $f^{\tau(y)} y \in Y$ for a.e. $y \in Y$; define $F = f^\tau : Y \to Y$. Suppose that μ_Y is an F-invariant probability measure on Y and that τ is integrable with respect to μ_Y. Let $\hat{f} : \Delta \to \Delta$ denote the tower with base map F and return time τ, and let $\pi : \Delta \to X$ be the semiconjugacy $\pi(y, \ell) = f^\ell y$. Assume that $\mu_X = \pi_* \mu_\Delta$. If all these assumptions are satisfied, we call τ a *return time* and F a *return map*.

2.2 Young towers with exponential tails

Let $f : X \to X$ be a measure-preserving transformation defined on a metric space (X, d) with Borel probability measure μ_X. Suppose that Y is a positive measure subset of X and that $\tau : Y \to \mathbb{Z}^+$ is a return time with return map $F = f^\tau : Y \to Y$. In particular, there is an F-invariant probability measure μ_Y on Y such that τ is μ_Y-integrable. Let $\Delta = Y^\tau$ and $\hat{f} : \Delta \to \Delta$ be the tower with base map F and return time τ as in Subsection 2.1 with $\hat{f}$-invariant probability measure μ_Δ and semiconjugacy $\pi : \Delta \to X$ such that $\mu_X = \pi_* \mu_\Delta$. In addition, we assume that μ_Y and μ_Δ (and hence μ_X) are ergodic. Moreover, we assume the exponential tails condition

$$\mu_Y(y \in Y : \tau(y) > n) = O(e^{-cn}) \quad \text{for some } c > 0. \tag{2.2}$$

Let $\mathcal{W}^s$ be a cover of Y by disjoint measurable subsets (called "local stable leaves") and let W^s_y denote the local stable leaf containing y. We require that $F(W^s_y) \subset W^s_{Fy}$ for all $y \in Y$. Let $\bar{Y}$ be the quotient space obtained from Y by quotienting along local stable manifolds and denote by $\bar{\pi} : Y \to \bar{Y}$ the corresponding projection. The probability measure $\bar{\mu}_Y = \bar{\pi}_* \mu_Y$ is ergodic and invariant under the quotient map $\bar{F} : \bar{Y} \to \bar{Y}$, and $\bar{\pi}$ defines a measure-preserving semiconjugacy between F and $\bar{F}$.

Let $\{a\}$ be an at most countable measurable partition of $\bar{Y}$. Define $s(y, y')$ to be the least integer $n \ge 0$ such that $F^n y$, $F^n y'$ lie in distinct partition elements. It is assumed that $s(y, y') = \infty$ if and only if $y = y'$. We require that $\bar{F}|_a : a \to \bar{Y}$ is a measurable bijection for all a and that there are constants $C > 0$, $\theta \in (0, 1)$ such that

$$\left| \log \tfrac{d\bar{\mu}_Y}{d\bar{\mu}_Y \circ \bar{F}}(y) - \log \tfrac{d\bar{\mu}_Y}{d\bar{\mu}_Y \circ \bar{F}}(y') \right| \le C \theta^{s(y, y')} \quad \text{for all } y, y' \in a \text{ and all } a.$$

Under these conditions, $\bar{F} : \bar{Y} \to \bar{Y}$ is called a *(full branch) Gibbs-Markov map* [1].

We require that $\tau : Y \to \mathbb{Z}^+$ is constant on $\bar{\pi}^{-1}a$ for all a. Hence τ is well-defined on $\bar{Y}$ and constant on partition elements.

Finally, assume that there are constants $C > 0$, $\gamma_0 \in (0,1)$ such that

$$d(F^n y, F^n y') \le C\gamma_0^n \quad \text{for all } y, y' \in Y,\ y' \in W_y^s,\ n \ge 0. \tag{2.3}$$

Under these assumptions, we say that $f : X \to X$ is modelled by a *Young tower $\Delta = Y^\tau$ with exponential tails.*

2.3 Chernov-Markarian-Zhang framework

Let $T : \Lambda \to \Lambda$ be an ergodic measure-preserving transformation defined on a metric space (Λ, d) with Borel probability measure μ. Let $X \subset \Lambda$ be a Borel subset of positive measure and define the *first return time* $\varphi : X \to \mathbb{Z}^+$ and *first return map* $f = T^\varphi : X \to X$,

$$\varphi(x) = \inf\{n \ge 1 : T^n x \in X\}, \qquad f(x) = T^{\varphi(x)}x. \tag{2.4}$$

Then φ is integrable and $\mu_X = \mu|_X/\mu(X)$ is an ergodic f-invariant probability measure on X. Define $\bar{\varphi} = \int_X \varphi\, d\mu_X$.

Next, we suppose that $f : X \to X$ is modelled by a Young tower $\Delta = Y^\tau$ with exponential tails as in Subsection 2.2. Define the *induced return time function*

$$\varphi^Y : Y \to \mathbb{Z}^+, \qquad \varphi^Y = \textstyle\sum_{\ell=0}^{\tau-1} \varphi \circ f^\ell.$$

Assume that $\varphi : X \to \mathbb{Z}^+$ is constant on $f^\ell \bar{\pi}^{-1}a$ for all $0 \le \ell < \tau(a)$ and all a. Then φ^Y is well-defined on $\bar{Y}$ and constant on partition elements.

The final condition is somewhat technical and is based on [4, Lemma 5.4] which is itself based on [48, Sublemma, p. 612]. Given $h \in C^\eta(\Lambda)$, define $H^Y = \sum_{\ell=0}^{\varphi^Y-1} h \circ T^\ell : Y \to \mathbb{R}$. Let $\mathcal{B}$ be the σ-algebra generated by $\mathcal{W}^s$. Then $\mathbb{E}(H^Y|\mathcal{B}) = \zeta \circ \bar{\pi}$ where $\zeta \in L^1(\bar{Y})$. It is immediate that

$$|\zeta(y)| \le |h|_\infty \varphi^Y(a) \quad \text{for all } y \in a \text{ and all } a. \tag{2.5}$$

We require that there are constants $C > 0$, $\gamma_0 \in (0,1)$ such that

$$|\zeta(y) - \zeta(y')| \le C\varphi^Y(a)\gamma_0^{s(y,y')} \quad \text{for all } y, y' \in a \text{ and all } a. \tag{2.6}$$

Under these assumptions, we say that $T : \Lambda \to \Lambda$ possesses a *Chernov-Markarian-Zhang structure.*

Remark 2.1 The exponential tail condition for τ is assumed for convenience, but the abstract results require only that $\mu_Y(\tau > n) = O(n^{-q})$ for q sufficiently large.

Remark 2.2 The method of choosing a first return map modelled by a Young tower with exponential tails arises in various contexts in the literature, see for example [9, 10] in the noninvertible context. However, the method plays a special role in the context of billiards as we now briefly recall.

Young [48] introduced Young towers with exponential tails as a general method for dealing with diffeomorphisms with singularities; the initial landmark application was to prove exponential decay of correlations for planar finite horizon dispersing billiards. Chernov [14] simplified the construction of exponential Young towers and used this to prove exponential decay of correlations for planar dispersing billiards with infinite horizon. Then Young [49] studied examples with subexponential decay of correlations using Young towers with subexponential tails. Markarian [34], noting that Chernov's simplification no longer applies in the subexponential case, devised the method outlined in this section: namely to construct a first return map for which Chernov [14] applies. This was used to prove the decay of correlations bound $O(1/n)$ for Bunimovich stadia. The method was extended and simplified by Chernov & Zhang [18] who applied it to a large class of billiard examples. Subsequent applications of the method include [16, 17] as well as Zhang [50] who analysed the examples discussed in this paper.

3 Statement of main results

Throughout this section, we suppose that $T:\Lambda\to\Lambda$ possesses a Chernov-Markarian-Zhang structure as in Section 2.3, with first return map $f=T^\varphi:X\to X$ modelled by a Young tower with exponential tails.

For random elements X_n, X taking values in a metric space, we write $X_n\to_w X$ if $\mathbb{P}(X_n\in B)\to\mathbb{P}(X\in B)$ for all Borel sets B with $\mathbb{P}(X\in \partial B)=0$. When the metric space is $\mathbb{R}$, we write $\to_d$ instead of $\to_w$. When the metric space is $\mathbb{R}$ and $X=0$, then this is equivalent to the simpler concept, convergence in probability, denoted $X_n\to_p 0$. For background on stable laws and Lévy processes, we refer to [41].

We assume that there exists $\alpha\in(1,2)$ such that the first return time $\varphi:X\to\mathbb{Z}^+$ satisfies the limit law

$$\frac{1}{n^{1/\alpha}}\Big(\sum_{j=0}^{n-1}\varphi\circ f^j-n\int_X\varphi\,d\mu_X\Big)\to_d G\quad\text{on } (X,\mu_X),\tag{3.1}$$

where G is an α-stable law. Since $\varphi\ge 1$, this stable law is totally skewed to the right.

Let $v:\Lambda\to\mathbb{R}$ be a Hölder observable with $\int_\Lambda v\,d\mu=0$. Define the associated induced observable $V:X\to\mathbb{R}$ given by $V(x)=\sum_{\ell=0}^{\varphi(x)-1}v(T^\ell x)$. We assume that

$$V-I\varphi\in L^p(X)\quad\text{for some } I\neq 0,\ p>\alpha.\tag{3.2}$$

Define $v_n=\sum_{j=0}^{n-1}v\circ T^j:\Lambda\to\mathbb{R}$.

Theorem 3.1 (Stable law) *Suppose that $T : \Lambda \to \Lambda$ possesses a Chernov-Markarian-Zhang structure and that $v : \Lambda \to \mathbb{R}$ is a Hölder mean zero observable. Assume (3.1) and (3.2). Then $n^{-1/\alpha} v_n \to_d (\bar\varphi)^{-1/\alpha} IG$.*

Next, let $W \in D[0,\infty)$ be the α-stable Lévy process with $W(1) =_d G$. Define $W_n : \Lambda \to D[0,\infty)$ by $W_n(t) = n^{-1/\alpha} v_{[nt]}$.

Define $M_1,\ M_2 : X \to [0,\infty)$,

$$M_1 = \max_{1\le \ell' \le \ell \le \varphi} (v_{\ell'} - v_\ell) \wedge \max_{1\le \ell' \le \ell \le \varphi} (v_\ell - v_{\ell'}),$$
$$M_2 = \Big\{ \max_{0\le \ell \le \varphi} (-v_\ell) + \max_{0\le \ell \le \varphi} (v_\ell - V) \Big\} \wedge \Big\{ \max_{0\le \ell \le \varphi} v_\ell + \max_{0\le \ell \le \varphi} (V - v_\ell) \Big\}.$$

Note that $M_1 = 0$ if and only if excursions between returns to X are monotone [37], and $M_2 = 0$ if and only if excursions starting at $x \in X$ remain between 0 and $V(x)$.

Theorem 3.2 (WIP) *Suppose that $T : \Lambda \to \Lambda$ possesses a Chernov-Markarian-Zhang structure and that $v : \Lambda \to \mathbb{R}$ is a Hölder mean zero observable. Assume (3.1) and (3.2).*

(a) If $n^{-1/\alpha} \max_{j\le n} M_1 \circ f^j \to_p 0$ on (X, μ_X), then $W_n \to_w (\bar\varphi)^{-1/\alpha} IW$ on (Λ, μ) in $(D[0,\infty), \mathcal{M}_1)$.

(b) If $n^{-1/\alpha} \max_{j\le n} M_2 \circ f^j \to_p 0$ on (X, μ_X), then $W_n \to_w (\bar\varphi)^{-1/\alpha} IW$ on (Λ, μ) in $(D[0,\infty), \mathcal{M}_2)$.

The theorem asserts that whenever excursions satisfy a mild monotonicity condition ($n^{-1/\alpha} \max_{j\le n} M_1 \circ f^j \to_p 0$), or lie within a controlled distance from its endpoints ($n^{-1/\alpha} \max_{j\le n} M_2 \circ f^j \to_p 0$), then we obtain the WIP in the $\mathcal{M}_1$ or $\mathcal{M}_2$ topology respectively.

Remark 3.3 Let $(\Omega, \mathbb{P})$ be a probability space and $R_n : \Omega \to S$ a sequence of Borel measurable maps where S is a metric space. *Strong distributional convergence* of R_n to a random element R on $(\Omega, \mathbb{P})$ means that $R_n \to_w R$ in S on the probability space $(\Omega, \mathbb{P}')$ for all probability measures $\mathbb{P}' \ll \mathbb{P}$.

In the context of Theorem 3.2, strong distributional convergence on (Λ, μ) is automatic. Let T be an ergodic measure-preserving transformation on a probability space (Λ, μ) and let μ' be an absolutely continuous probability measure. Based on ideas of [21], it was shown in [52, Theorem 1 and Corollary 3] that distributional convergence in $(D[0,\infty), \mathcal{J}_1)$ holds on (Λ, μ) if and only if it holds on (Λ, μ'). Hence distributional convergence in $D[0,\infty)$ with the $\mathcal{J}_1$ topology on (Λ, μ) is equivalent to strong distributional convergence. As pointed out in [37, Proposition 2.8], this carries over immediately to weaker topologies on $D[0,\infty)$ such as $\mathcal{M}_1$ and $\mathcal{M}_2$.

Remark 3.4 A more concise formula for M_2 can be obtained by noting that

$$\begin{aligned} M_2 &= \Big\{ -\min_{0\le\ell\le\varphi} v_\ell + \max_{0\le\ell\le\varphi} v_\ell - V \Big\} \wedge \Big\{ \max_{0\le\ell\le\varphi} v_\ell + V - \min_{0\le\ell\le\varphi} v_\ell \Big\} \\ &= \max_{0\le\ell\le\varphi} v_\ell - \min_{0\le\ell\le\varphi} v_\ell - |V|. \end{aligned}$$

The next result, proved in Section 5, extends and significantly improves [37, Proposition 2.7].

Proposition 3.5 *Let $i \in \{1,2\}$. Suppose that there are constants $C>0$, $\delta\in(0,1)$ such that $M_i \le C\varphi^\delta$ almost everywhere. Then the assumption on M_i in Theorem 3.2 is satisfied.*

In Section 8, we require the following converse result for the $\mathcal{M}_1$ topology. (There is no such converse result for $\mathcal{M}_2$.)

Proposition 3.6 *If $W_n \to_w IW$ in $(D[0,\infty),\mathcal{M}_1)$ for some constant $I \ne 0$, then $n^{-1/\alpha}\max_{0\le j\le n} M_1\circ f^j \to_p 0$.*

Proof Without loss, $I=1$. Fix $c>0$. Define $\Delta W(t) = W(t)-W(t_-)$. The stable law G is totally skewed with Lévy measure supported in $(0,\infty)$, so $\mathbb{P}\{\Delta W(t) < -c \text{ for some } 0\le t\le 2\bar\varphi\} = 0$.

For $\delta>0$, define

$$E_\delta = \{u\in D[0,2\bar\varphi] : u(t)-u(t') < -c \quad \text{for some } 0\le t'<t<(t'+\delta)\wedge 2\bar\varphi\}.$$

Since $W_n \to_w W$ in $\mathcal{M}_1$, for any $\epsilon>0$ there exists $\delta>0$, $n_0\ge 1$ such that

$$\mu(W_n\in E_\delta) < \epsilon \quad \text{for } n\ge n_0.$$

Let $\varphi_n = \sum_{j=0}^{n-1}\varphi\circ f^j$. Since φ is integrable, it follows from the ergodic theorem that $n^{-1}\varphi_n \to \bar\varphi$ a.e. and so $n^{-1}\varphi\circ f^n \to 0$ a.e. It follows easily that $n^{-1}\max_{j\le n}\varphi\circ f^j \to 0$ a.e. Hence there exists $n_1\ge n_0$ such that

$$\mu\big(n^{-1}\max_{j\le n}\varphi\circ f^j \ge \delta\big) + \mu(n^{-1}\varphi_n \ge 2\bar\varphi) < \epsilon \quad \text{for } n\ge n_1.$$

Now,

$$n^{-1/\alpha}\max_{j\le n} M_1\circ f^j \le n^{-1/\alpha}\max_{j\le n}\max_{0\le\ell'<\ell<\varphi\circ f^j}(v_{\ell'}-v_\ell)\circ f^j \le \max{}^*(W_n(t')-W_n(t)),$$

where $\max^*$ is the maximum over $0\le t'<t<(t'+n^{-1}\max_{j\le n}\varphi\circ f^j)\wedge n^{-1}\varphi_n$.

It follows that for $n\ge n_1$,

$$\begin{aligned} &\mu\big\{n^{-1/\alpha}\max_{j\le n} M_1\circ f^j > c\big\} \le \mu\big\{\max{}^*(W_n(t')-W_n(t)) > c\big\} \\ &\le \mu(n^{-1}\varphi_n\ge 2\bar\varphi) + \mu\big(n^{-1}\max_{j\le n}\varphi\circ f^j\ge\delta\big) + \mu\Big(\max_{0\le t'<t<(t'+\delta)\wedge 2\bar\varphi}(W_n(t')-W_n(t)) > c\Big) \\ &< \epsilon + \mu(W_n\in E_\delta) < 2\epsilon. \end{aligned}$$

Hence $n^{-1/\alpha}\max_{j\le n} M_1\circ f^j \to_p 0$. ∎

Remark 3.7 Convergence results in the $\mathcal{M}_1$ topology for nonuniformly hyperbolic maps were considered previously by [37] with applications to Markov Pomeau-Manneville intermittent maps [40]. Such maps fall into a greatly simplified version of the Chernov-Markov-Zhang framework. Fix $\alpha\in(1,2)$ and set $\Lambda=[0,1]$. A prototypical example [32] is the map $T:\Lambda\to\Lambda$ given by $Tx=\begin{cases} x(1+2^{1/\alpha}x^{1/\alpha}) & x<\frac12 \\ 2x-1 & x>\frac12\end{cases}$, but the method applies equally to the general class of intermittent Markov maps considered by [45]. Taking $X=[\frac12,1]$, the first return map $f=T^\varphi:X\to X$ is already Gibbs-Markov, so there is no need to consider an induced return map $F=T^{\varphi^Y}$, nor to quotient along stable leaves. In other words, $X=Y=\bar Y$. For these examples, condition (3.1) holds by [24]. Condition (3.2) and condition (a) in Theorem 3.2 were verified in [37, Section 4].

Theorem 3.2(a) also applies to non-Markovian intermittent maps $T:\Lambda\to\Lambda$: the so-called AFN maps studied by [51]. A specific example is given by $Tx=x(1+bx^{1/\alpha}) \bmod 1$ which is not Markov when the positive constant b is not an integer. As far as we know, the WIP for stable laws has not been previously studied for such maps. Since this is a much simpler situation than for our main billiard example, we just sketch the details. (In fact, the situation lies in between those for Markov intermittent maps and billiards: quotienting along stable leaves is not required, but we do need to consider an induced map $F=T^{\varphi^Y}$.)

Take X to be the interval of domain of the rightmost branch of T. Let $v:\Lambda\to\mathbb{R}$ be Hölder with $v(0)\neq0$ and define $V=\sum_{\ell=0}^{\varphi-1}v\circ T^\ell$ where $\varphi:X\to\mathbb{Z}^+$ is the first return time. The same calculations as in the Markov case show that $\mu_X(\varphi>n)\sim cn^{-\alpha}$ for some $c>0$ and that $V-v(0)\varphi\in L^p(X)$ for some $p>\alpha$. Hence (3.2) is satisfied. Also condition (a) of Theorem 3.2 holds as in the Markov case. By [10, Section 9], $f=T^\varphi$ is modelled by a Young tower with exponential tails so these maps fall into the Chernov-Markarian-Zhang framework.

It remains to verify the stable law (3.1). One method is to proceed as in [30, Section 3], but alternatively we can make use of the fact proved in [10] that φ^Y inherits the tail asymptotic satisfied by φ. Since $F:Y\to Y$ is Gibbs-Markov and φ^Y is constant on partition elements, a stable law for φ^Y is immediate by [1, Theorem 6.1]. This yields the desired stable law for φ by Theorem A.1.

4 Inducing functional limit laws

The proof of Theorem 3.2(a) makes use of a purely probabilistic result [37, Theorem 2.2] on inducing functional limit laws on $D[0,\infty)$ with the $\mathcal{M}_1$ topology. The result in [37] is stated in a slightly generalised form in Theorem 4.1 below. The proof of Theorem 3.2(b) makes use of the corresponding result in the $\mathcal{M}_2$ topology. In this section, it is not required that W is a Lévy process.

We assume the set up in Section 2.1 but with different notation (this simplifies the application of Theorem 4.1 in Sections 5 and 7). Let $S:\Omega\to\Omega$ be an ergodic measure-preserving transformation on a probability space (Ω,μ_Ω) and fix a positive measure subset $\Omega_0\subset\Omega$. Let μ_{Ω_0} be a probability measure on Ω_0 and let $r:\Omega_0\to\mathbb{Z}^+$ be an integrable return time such that the return map (not necessarily the first return) $S_0=S^r:\Omega_0\to\Omega_0$ is measure-preserving and ergodic. Define $\bar r=\int_{\Omega_0} r\,d\mu_{\Omega_0}$. Let $\widehat S:\Delta\to\Delta$ denote the tower with base map S_0 and return time r, and let $\pi:\Delta\to\Omega$ be the semiconjugacy $\pi(y,\ell)=S^\ell y$. We assume that $\mu_\Delta=(\mu_{\Omega_0}\times\text{counting})/\bar r$ is ergodic and that $\pi_*\mu_\Delta=\mu_\Omega$.

Let $\phi:\Omega\to\mathbb{R}$ be measurable, with induced observable $\Phi:\Omega_0\to\mathbb{R}$ given by $\Phi=\sum_{\ell=0}^{r-1}\phi\circ S^\ell$. Let b_n be a sequence of positive numbers. Define càdlàg processes ψ_n on Ω and Ψ_n on Ω_0:

$$\psi_n(t)=\frac{1}{b_n}\sum_{j=0}^{[nt]-1}\phi\circ S^j,\qquad \Psi_n(t)=\frac{1}{b_n}\sum_{j=0}^{[nt]-1}\Phi\circ S_0^j.$$

Let $W\in D[0,\infty)$ and define $\widetilde W(t)=W(\bar r t)$. (If W is an α-stable Lévy process, $\alpha\in(0,2]$, then $\widetilde W=\bar r^{1/\alpha}W$.) Also, define $\phi_\ell=\sum_{j=0}^{\ell-1}\phi\circ S^j$ and

$$\begin{aligned} M_1&=\max_{1\le\ell'\le\ell\le r}(\phi_{\ell'}-\phi_\ell)\wedge\max_{1\le\ell'\le\ell\le r}(\phi_\ell-\phi_{\ell'}),\\ M_2&=\Big\{\max_{0\le\ell\le r}(-\phi_\ell)+\max_{0\le\ell\le r}(\phi_\ell-\Phi)\Big\}\wedge\Big\{\max_{0\le\ell\le r}\phi_\ell+\max_{0\le\ell\le r}(\Phi-\phi_\ell)\Big\}.\end{aligned}$$

Theorem 4.1 *Let $i\in\{1,2\}$. Suppose that on $(\Omega_0,\mu_{\Omega_0})$*

1. *$\Psi_n\to_w\widetilde W$ in $(D[0,\infty),\mathcal{M}_i)$ and*
2. *$\frac{1}{b_n}\max_{0\le j\le n}M_i\circ S_0^j\to_p 0$.*

Then $\psi_n\to_w W$ in $(D[0,\infty),\mathcal{M}_i)$ on (Ω,μ_Ω).

Proof of Theorem 4.1 for $i=1$. Under the additional assumptions that b_n is regularly varying and r is the first return time, this is is precisely [37, Theorem 2.2]. (The conclusion in [37, Theorem 2.2] is stated slightly differently using that $\bar r^{-1}=\mu_\Omega(\Omega_0)$ for first return times.) It is easily checked that the proof in [37] does not use any properties of the sequence b_n.

It remains to drop the assumption that r is the first return time to Ω_0. Note that $\Omega_0\subset\Omega$ is naturally identified with $\Delta_0=\{(y,0):y\in\Omega_0\}\subset\Delta$ and $r:\Delta_0\to\mathbb{R}$ is now the first return to Δ_0 for the dynamics on Δ. Define $\widehat S_0:\Delta_0\to\Delta_0$, $\widehat S_0(y,0)=(S_0y,0)$.

The observable $\phi : \Lambda \to \mathbb{R}$ lifts to an observable $\hat{\phi} = \phi \circ \pi : \Delta \to \mathbb{R}$. Define the corresponding càdlàg process $\widehat{\psi}_n(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} \hat{\phi} \circ \widehat{S}^j$ on Δ. Also, we define $\widehat{\Phi}$, $\widehat{\Psi}_n$, $\widehat{M_1}$ on Δ_0 (corresponding to Φ, Ψ_n, M_1 on Ω_0) using $\hat{\phi}$, $\widehat{S}$, $\widehat{S}_0$ instead of ϕ, S, S_0, so

$$\widehat{\Phi} = \sum_{\ell=0}^{\hat{r}-1} \hat{\phi} \circ \widehat{S}^\ell, \quad \widehat{\Psi}_n(t) = \frac{1}{b_n} \sum_{j=0}^{[nt]-1} \widehat{\Phi} \circ \widehat{S}_0^j, \quad \widehat{M_1} = \max_{1 \le \ell' \le \ell \le \hat{r}} (\hat{\phi}_{\ell'} - \hat{\phi}_\ell) \wedge \max_{1 \le \ell' \le \ell \le \hat{r}} (\hat{\phi}_\ell - \hat{\phi}_{\ell'}),$$

where $\hat{\phi}_\ell = \sum_{j=0}^{\ell-1} \hat{\phi} \circ \widehat{S}^j$ and $\hat{r}(y,0) = r(y)$.

Note that

$$\widehat{\Phi}(y,0) = \Phi(y), \quad \widehat{\Psi}_n(t)(y,0) = \Psi_n(t)(y), \quad \widehat{M_1}(y,0) = M_1(y).$$

In particular, the assumptions 1 and 2 for Ψ_n and M_1 on Ω_0 imply the corresponding assumptions for $\widehat{\Psi}_n$ and $\widehat{M_1}$ on Δ_0. Since $\hat{r} : \Delta_0 \to \mathbb{Z}^+$ is the first return time, $\widehat{\psi}_n \to_w W$ on (Δ, μ_Δ) in $(D[0,\infty), \mathcal{M}_1)$ by [37, Theorem 2.2]. The result follows since π is a measure-preserving semiconjugacy. ∎

Proof of Theorem 4.1 for $i = 2$. The strategy here is similar to the one of [37, Theorem 2.2]. As in the proof for $i = 1$, by considering the associated tower we may suppose without loss that $r : \Omega_0 \to \mathbb{Z}^+$ is the first return time.

Write $\psi_n = U_n + R_n$, where

$$U_n(t) = \frac{1}{b_n} \sum_{\ell=0}^{N_{[nt]}-1} \Phi \circ S_0^\ell \quad \text{and} \quad R_n(t) = \frac{1}{b_n} \left(\sum_{\ell=0}^{[nt]-r_{N_{[nt]}}-1} \phi \circ S^\ell \right) \circ S_0^{N_{[nt]}}.$$

Here, $r_k = \sum_{j=0}^{k-1} r \circ S^j$ and $N_k(x) = \max\{\ell \ge 1 : r_\ell(x) \le k\}$ is the number of returns of x to the set Ω_0, under iteration by S, up to time k.

By [37, Lemma 3.4], $U_n \to_w (\widetilde{W}(\bar{r}^{-1}t))_{t \ge 0} = W$ in $(D[0,\infty), \mathcal{M}_2)$. (The hypotheses of [37, Lemma 3.4] are with respect to the $\mathcal{M}_1$ topology. However, most of the proof holds in any separable metric space and the only ingredient that relies on the specific topology is [47, Theorem 13.2.3] which is formulated for both $\mathcal{M}_1$ and $\mathcal{M}_2$.)

We claim that $d_{\mathcal{M}_2,[0,K]}(\psi_n, U_n) \to 0$ as $n \to \infty$ for each $K \in \mathbb{N}$. Then by [8, Theorem 3.1], $\psi_n \to_w W$ in $(D[0,K], \mathcal{M}_2)$ for each $K \in \mathbb{N}$, and the result follows.

It remains to verify the claim. This means taking into account the contribution of the final incomplete excursion from Ω_0 (if any) encoded by R_n. Following [37, Lemmas 3.5 and 3.6], given $x \in \Omega_0$, $n \ge 1$, write $g_j(t) = \psi_n(t)(x)|_{[t_j,t_{j+1}]}$ for every $0 \le j \le Kn+1$, where $t_j = \frac{1}{n} r_j \wedge K$. Then

$$d_{\mathcal{M}_2,[0,K]}(\psi_n(\cdot)(x), U_n(\cdot)(x)) \le \max_{0 \le j \le Kn+1} d_{\mathcal{M}_2,[t_j,t_{j+1}]}(g_j, \bar{g}_j),$$

where $\bar{g}_j = U_n|_{[t_j,t_{j+1}]} = g_j|_{[t_j,t_{j+1})} + 1_{\{t_{j+1}\}} g_j(t_{j+1})$.

By Lemma B.1,

$$d_{\mathcal{M}_2,[t_j,t_{j+1}]}(g_j,\bar g_j) \le t_{j+1}-t_j + A_j\wedge B_j \le \frac{1}{n}r(S_0^jx) + A_j\wedge B_j,$$

where

$$\begin{aligned}
A_j &= \sup_{t\in[t_j,t_{j+1}]}(g_j(t_j)-g_j(t)) + \sup_{t\in[t_j,t_{j+1}]}(g_j(t)-g_j(t_{j+1})) \\
&= \sup_{t\in[t_j,t_{j+1}]}(\psi_n(t_j)(x)-\psi_n(t)(x)) + \sup_{t\in[t_j,t_{j+1}]}(\psi_n(t)(x)-\psi_n(t_{j+1})(x)) \\
&= \frac{1}{b_n}\max_{0\le\ell\le r(S_0^jx)}(-\phi_\ell(S_0^jx)) + \frac{1}{b_n}\max_{0\le\ell\le r(S_0^jx)}(\phi_\ell(S_0^jx)-\Phi(S_0^jx)),
\end{aligned}$$

and similarly

$$\begin{aligned}
B_j &= \sup_{t\in[t_j,t_{j+1}]}(g_j(t)-g_j(t_j)) + \sup_{t\in[t_j,t_{j+1}]}(g_j(t_{j+1})-g_j(t)) \\
&= \frac{1}{b_n}\max_{0\le\ell\le r(S_0^jx)}\phi_\ell(S_0^jx) + \frac{1}{b_n}\max_{0\le\ell\le r(S_0^jx)}(\Phi(S_0^jx)-\phi_\ell(S_0^jx)).
\end{aligned}$$

In particular, $A_j\wedge B_j \le \frac{1}{b_n}M_2(S_0^jx)$. Hence we have shown that

$$d_{\mathcal{M}_2,[0,K]}(\psi_n,U_n) \le \frac{1}{n}\max_{0\le j\le Kn+1} r\circ S_0^j + \frac{1}{b_n}\max_{0\le j\le Kn+1} M_2\circ S_0^j.$$

The first term converges to zero a.e. by ergodicity, and the second term converges to zero in probability by the assumption on M_2. ∎

5 Limit laws for φ and φ^Y

Recall that $T:\Lambda\to\Lambda$ is assumed to possess a Chernov-Markarian-Zhang structure, with first return map $f=T^\varphi:X\to X$ modelled by a Young tower $\Delta=Y^\tau$ with exponential tails and induced return time $\varphi^Y=\sum_{\ell=0}^{\tau-1}\varphi\circ f^\ell : Y\to\mathbb{Z}^+$.

In this section, we show how to pass from the stable law (3.1) for φ to a stable law for φ^Y and WIPs for φ and φ^Y. We also prove Proposition 3.5.

Note that $\int_Y \varphi^Y\,d\mu_Y=\bar\varphi\bar\tau$. Define the centered return times

$$\widetilde\varphi=\varphi-\bar\varphi, \qquad \widetilde\varphi^Y=\varphi^Y-\tau\bar\varphi, \qquad \widetilde{\varphi^Y}=\varphi^Y-\bar\tau\bar\varphi.$$

Define càdlàg processes A_n and A_n^Y on X and Y,

$$A_n(t)=n^{-1/\alpha}\textstyle\sum_{j=0}^{[nt]-1}\widetilde\varphi\circ f^j, \qquad A_n^Y(t)=n^{-1/\alpha}\sum_{j=0}^{[nt]-1}\widetilde\varphi^Y\circ F^j. \tag{5.1}$$

Lemma 5.1 *Assume that (3.1) holds and let W be the α-stable Lévy process corresponding to the totally skewed α-stable law G in (3.1). Then*

(a) $n^{-1/\alpha}\sum_{j=0}^{n-1}\widetilde{\varphi}^Y\circ F^j\to(\bar\tau)^{1/\alpha}G$ *on* (Y,μ_Y).

(b) $n^{-1/\alpha}\sum_{j=0}^{n-1}\widetilde{\varphi^Y}\circ F^j\to(\bar\tau)^{1/\alpha}G$ *on* (Y,μ_Y).

(c) $A_n^Y\to_w(\bar\tau)^{1/\alpha}W$ *on* (Y,μ_Y) *in* $(D[0,\infty),\mathcal{J}_1)$.

(d) $A_n\to_w W$ *on* (X,μ_X) *in* $(D[0,\infty),\mathcal{M}_1)$.

Proof (a) Since $\tau:\bar Y\to\mathbb{Z}^+$ has exponential tails, we certainly have that $\tau\in L^2$. Also τ is constant on partition elements and $\bar F:\bar Y\to\bar Y$ is Gibbs-Markov, so it is standard (see for example [26, Theorem 1.5]) that $n^{-1/2}(\sum_{j=0}^{n-1}\tau\circ\bar F^j-n\bar\tau)$ converges in distribution (to a possibly degenerate normal distribution). Since $\alpha\in(1,2)$,

$$n^{-1/\alpha}\Big(\sum_{j=0}^{n-1}\tau\circ F^j-n\bar\tau\Big)=_d n^{-1/\alpha}\Big(\sum_{j=0}^{n-1}\tau\circ\bar F^j-n\bar\tau\Big)\to_p 0. \tag{5.2}$$

By assumption (3.1), the centered return time function $\widetilde{\varphi}$ satisfies a stable law on X. Hence condition (a) in Theorem A.1 is satisfied with $b_n=n^{1/\alpha}$ and it follows from Theorem A.1 and Remark A.3 that $\widetilde{\varphi}^Y$ satisfies the required stable law on Y.

(b) By (5.2) and part (a),

$$\begin{aligned}n^{-1/\alpha}\sum_{j=0}^{n-1}\widetilde{\varphi^Y}\circ F^j&=n^{-1/\alpha}\Big(\sum_{j=0}^{n-1}\varphi^Y\circ F^j-n\bar\tau\bar\varphi\Big)\\&=n^{-1/\alpha}\sum_{j=0}^{n-1}\widetilde{\varphi}^Y\circ F^j+\bar\varphi n^{-1/\alpha}\Big(\sum_{j=0}^{n-1}\tau\circ F^j-n\bar\tau\Big)\to_d(\bar\tau)^{1/\alpha}G.\end{aligned}$$

(c) Recall that $\widetilde{\varphi}^Y$ is constant on partition elements of the Gibbs-Markov map $\bar F:\bar Y\to\bar Y$. By part (a), $n^{-1/\alpha}\sum_{j=0}^{n-1}\widetilde{\varphi}^Y\circ\bar F^j\to_d(\bar\tau)^{1/\alpha}G$. By Gouëzel [26, Theorem 1.5], $\widetilde{\varphi}^Y$ lies in the domain of attraction of the stable law $(\bar\tau)^{1/\alpha}G$ and hence has tails that are regularly varying with index α. We have verified the hypotheses of Tyran-Kamińska [46, Corollary 4.1][3], and so deduce that $A_n^Y\to_w(\bar\tau)^{1/\alpha}W$ in the $\mathcal{J}_1$ topology.

(d) We apply Theorem 4.1 with $i=1$ to pass from $A_n^Y:Y\to\mathbb{R}$ to $A_n:X\to\mathbb{R}$ via the inducing time $\tau:Y\to\mathbb{Z}^+$. (The spaces $\Omega_0\subset\Omega$ in Theorem 4.1 correspond to the spaces $Y\subset X$ here. Similarly ϕ, Φ, ψ_n, Ψ_n, r are called $\widetilde{\varphi}$, $\widetilde{\varphi}^Y$, A_n, A_n^Y, τ, and the maps $S:\Omega\to\Omega$, $S_0:\Omega_0\to\Omega_0$ are called $f:X\to X$, $F:Y\to Y$.)

Condition 1 of Theorem 4.1 is immediate from part (c). Define $M_1:Y\to\mathbb{R}$,

$$M_1=\max_{1\le\ell'\le\ell\le\tau}(\widetilde{\varphi}_{\ell'}-\widetilde{\varphi}_\ell)\wedge\max_{1\le\ell'\le\ell\le\tau}(\widetilde{\varphi}_\ell-\widetilde{\varphi}_{\ell'}),$$

[3] The hypothesis "exponentially continued fraction mixing" in [46, Corollary 4.1] is automatic for full-branch Gibbs-Markov maps (see the discussion immediately after [46, Example 4.1]).

where $\widetilde{\varphi}_\ell = \sum_{j=0}^{\ell-1} \widetilde{\varphi} \circ f^j$. We claim that $n^{-1/\alpha} \max_{0\le j\le n} M_1 \circ F^j \to 0$ a.e. This implies condition 2 of Theorem 4.1 and the result follows.

By positivity of φ,

$$M_1 \le \max_{1\le \ell' \le \ell \le \tau} (\widetilde{\varphi}_{\ell'} - \widetilde{\varphi}_\ell) = \max_{1\le \ell' \le \ell \le \tau} \big\{ (\varphi_{\ell'} - \varphi_\ell) - (\ell' - \ell)\bar{\varphi} \big\} \le \tau \bar{\varphi}.$$

Since τ has exponential tails, it is certainly the case that $\tau \in L^\alpha(Y)$. By the ergodic theorem, $n^{-1} \sum_{j=0}^{n-1} \tau^\alpha \circ F^j \to \int_Y \tau^\alpha \, d\mu_Y$ a.e. and so $\tau \circ F^n = o(n^{1/\alpha})$ a.e. It follows easily that $\max_{j\le n} \tau \circ F^j = o(n^{1/\alpha})$ a.e. Hence $n^{-1/\alpha} \max_{j\le n} M_1 \circ F^j \to 0$ a.e. as required. ∎

Corollary 5.2 *Under the assumptions of Lemma 5.1, if $\delta \in (0,1)$ then $n^{-1/\alpha} \max_{0\le j\le n} |\widetilde{\varphi}|^\delta \circ f^j \to_p 0$ on (X, μ_X).*

Proof The functional $\chi : (D[0,\infty), \mathcal{M}_1) \to \mathbb{R}$, $\chi(g) = \sup_{[0,1]} |g|$ is continuous so, by the continuous mapping theorem applied to Lemma 5.1(d), we have $\chi(A_n) \to_w \chi(W)$ on (X, μ_X). Hence $n^{-1/\alpha} \max_{0\le j\le n} |\widetilde{\varphi}| \circ f^j = \chi(A_n)$ converges in distribution and so $n^{-1/(\delta\alpha)} \max_{0\le j\le n} |\widetilde{\varphi}| \circ f^j \to_p 0$. The result follows. ∎

Proof of Proposition 3.5 We have $\varphi = \bar{\varphi} + \widetilde{\varphi} \ll 1 + |\widetilde{\varphi}|$, so $M_i \ll \varphi^\delta \ll 1 + |\widetilde{\varphi}|^\delta$. Hence $n^{-1/\alpha} \max_{j\le n} M_i \circ f^j \ll n^{-1/\alpha}(1 + \max_{j\le n} |\widetilde{\varphi}|^\delta \circ f^j) \to_p 0$ by Corollary 5.2. ∎

Remark 5.3 As seen in the proof of Lemma 5.1(c), $\widetilde{\varphi}^Y$ lies in the domain of attraction of an α-stable law, so $\varphi^Y \in L^q(Y)$ for all $q < \alpha$. It follows easily that $\varphi \in L^q(X)$ for all $q < \alpha$.

6 Moment estimates for induced observables

In this section, we consider estimates for certain induced observables. We continue to assume that $T : \Lambda \to \Lambda$ possesses a Chernov-Markarian-Zhang structure. Our method follows [4, Section 5].

Proposition 6.1 *Let $H : X \to \mathbb{R}$ and suppose that $H \in L^q(X)$ for some $q > 1$. Define $H^Y = \sum_{\ell=0}^{\tau-1} H \circ f^\ell$. Then $H^Y \in L^p(Y)$ for all $p < q$.*

Proof Let $a > 1$ with $1/a + 1/q = 1/p$. Let $c' = c/a$, where $c > 0$ is given by (2.2). By Hölder's inequality,

$$\begin{aligned} |H^Y|_{L^p(Y)} &\ll \sum_{n\ge 1} \Big| 1_{\{\tau=n\}} \sum_{\ell=0}^{n-1} H \circ f^\ell \Big|_{L^p(Y)} \le \sum_{n\ge 1} \mu_Y(\tau = n)^{1/a} \Big| 1_{\{\tau=n\}} \sum_{\ell=0}^{n-1} H \circ f^\ell \Big|_{L^q(Y)} \\ &\le \sum_{n\ge 1} e^{-c'n} \sum_{\ell=0}^{n-1} \Big| 1_{\{\tau=n\}} H \circ f^\ell \Big|_{L^q(Y)} \ll \sum_{n\ge 1} e^{-c'n} n |H|_{L^q(X)} \ll |H|_{L^q(X)} < \infty, \end{aligned}$$

as required. ∎

Lemma 6.2 *Let $p \in (1,2]$. Let $h \in C^\eta(\Lambda)$ with induced observables*

$$H = \sum_{\ell=0}^{\varphi-1} h \circ T^\ell : X \to \mathbb{R}, \qquad \widetilde{H}^Y = \sum_{\ell=0}^{\tau-1} \widetilde{H} \circ f^\ell : Y \to \mathbb{R},$$

where $\widetilde{H} = H - \int_X H \, d\mu_X$. Suppose that $H \in L^q(X)$ for some $q > p$. Then $\big| \max_{j \le n} | \sum_{i=0}^{j-1} \widetilde{H}^Y \circ F^i | \big|_p \ll n^{1/p}$.

Proof Following [4], we apply a Gordin type argument [23] to obtain an L^p martingale-coboundary decomposition.

First, by Proposition 6.1 we may suppose that $\widetilde{H}^Y \in L^q(Y)$ for some (smaller) $q > p$. Let $\overline{\mathcal{B}}$ denote the underlying σ-algebra on $\bar{Y}$ and let $\mathcal{B} = \bar{\pi}^{-1}\overline{\mathcal{B}}$. Then $\{F^n\mathcal{B},\, n \in \mathbb{Z}\}$ defines an increasing sequence of σ-algebras on Y. We claim that

$$\sum_{n=0}^{\infty} \big|\mathbb{E}(\widetilde{H}^Y | F^n\mathcal{B}) - \widetilde{H}^Y\big|_p < \infty, \qquad \sum_{n=1}^{\infty} \big|\mathbb{E}(\widetilde{H}^Y | F^{-n}\mathcal{B})\big|_p < \infty. \tag{6.1}$$

Suppose that the claim is true. Then equivalently,

$$\sum_{n=0}^{\infty} \big|\mathbb{E}(\widetilde{H}^Y \circ F^n | \mathcal{B}) - \widetilde{H}^Y \circ F^n\big|_p < \infty, \qquad \sum_{n=1}^{\infty} \big|\mathbb{E}(\widetilde{H}^Y \circ F^{-n} | \mathcal{B})\big|_p < \infty,$$

so the series

$$\chi = \sum_{n=0}^{\infty} (\mathbb{E}(\widetilde{H}^Y \circ F^n | \mathcal{B}) - \widetilde{H}^Y \circ F^n) + \sum_{n=1}^{\infty} \mathbb{E}(\widetilde{H}^Y \circ F^{-n} | \mathcal{B}),$$

converges in $L^p(Y)$. Define

$$m = \widetilde{H}^Y + \chi - \chi \circ F \in L^p(Y). \tag{6.2}$$

Then

$$m = \textstyle\sum_{n=-\infty}^{\infty} (g_n - g_n \circ F) = \sum_{n=-\infty}^{\infty} (g_{n+1} - g_n \circ F), \tag{6.3}$$

where $g_n = \mathbb{E}[\widetilde{H}^Y \circ F^n | \mathcal{B}]$.

Now, g_n is $\mathcal{B}$-measurable, while $g_n \circ F$ is measurable with respect to $F^{-1}\mathcal{B} \subset \mathcal{B}$. Hence m is $\mathcal{B}$-measurable. Next, $g_n \circ F = \mathbb{E}[\widetilde{H}^Y \circ F^n | \mathcal{B}] \circ F = \mathbb{E}[\widetilde{H}^Y \circ F^{n+1} | F^{-1}\mathcal{B}]$. It follows that

$$\mathbb{E}[g_n \circ F | F^{-1}\mathcal{B}] = \mathbb{E}[\widetilde{H}^Y \circ F^{n+1} | F^{-1}\mathcal{B}] = \mathbb{E}[\mathbb{E}[\widetilde{H}^Y \circ F^{n+1} | \mathcal{B}] | F^{-1}\mathcal{B}] = \mathbb{E}[g_{n+1} | F^{-1}\mathcal{B}],$$

where we used again that $F^{-1}\mathcal{B} \subset \mathcal{B}$. Substituting into (6.3), we obtain $\mathbb{E}[m|F^{-1}\mathcal{B}] = 0$. Hence $\{m \circ F^{-n}; n \in \mathbb{Z}\}$ is a martingale difference sequence with respect to the filtration $\{F^n\mathcal{B}; n \in \mathbb{Z}\}$.

By Burkholder's inequality [13, Theorem 3.2],

$$\Big|\sum_{j=1}^{n} m \circ F^{-j}\Big|_p \ll \Big|\Big(\sum_{j=1}^{n} m^2 \circ F^{-j}\Big)^{1/2}\Big|_p = \Big(\int \Big(\sum_{j=1}^{n} m^2 \circ F^{-j}\Big)^{p/2}\Big)^{1/p}$$
$$\le \Big(\int \sum_{j=1}^{n} |m|^p \circ F^{-j}\Big)^{1/p} = |m|_p\, n^{1/p}.$$

By Doob's inequality [20] (see also [13, Equation (1.4), p. 20]),

$$\Big|\max_{j\le n}\Big|\sum_{i=0}^{j-1} m \circ F^{i}\Big|\Big|_p \le 2\Big|\max_{j\le n}\Big|\sum_{i=1}^{j} m \circ F^{-i}\Big|\Big|_p \ll n^{1/p}.$$

Also,

$$\int_Y \max_{j\le n} |\chi \circ F^j - \chi|^p \le 2\sum_{j=0}^{n} |\chi \circ F^j|_p^p = 2(n+1)|\chi|_p^p,$$

so $\big|\max_{j\le n}|\chi \circ F^j - \chi|\big|_p \ll n^{1/p}$.

By (6.2), $\sum_{j=0}^{n-1} \widetilde{H}^Y \circ F^j = \sum_{j=0}^{n-1} m \circ F^j + \chi \circ F^n - \chi$, so the desired estimate for $\widetilde{H}^Y$ follows from the estimates for m and χ.

It remains to verify the claim. The argument is identical to the one in [4, Lemma 5.3] except for the order of integrability. Hence we only sketch the argument referring to [4] for the details (especially the prerequisite estimates for systems modelled by Young towers).

If $y, y' \in Y$ lie in the same stable leaf, then $|\widetilde{H}^Y(y) - \widetilde{H}^Y(y')| \ll \varphi^Y(y) d(y,y')^\eta$. By (2.3), the atoms of $F^n\mathcal{B}$ have diameter at most $C\gamma_0^n$ for some $C > 0$, $\gamma_0 \in (0,1)$. Hence setting $\gamma = \gamma_0^\eta$, we have (cf. [4, Estimate (54)])

$$|\widetilde{H}^Y - \mathbb{E}(\widetilde{H}^Y|F^n\mathcal{B})| \ll \varphi^Y \gamma^n.$$

Choose $r > 1$ with $1/r + 1/q = 1/p$. In the case that the inducing time is large,

$$\big|1_{\{\varphi^Y > n^{2r}\}}\widetilde{H}^Y\big|_p \le \mu(\varphi^Y > n^{2r})^{1/r}|\widetilde{H}^Y|_q \le n^{-2}|\varphi^Y|_1^{1/r}|\widetilde{H}^Y|_q,$$

and similarly,

$$\big|1_{\{\varphi^Y > n^{2r}\}}\mathbb{E}(\widetilde{H}^Y|F^n\mathcal{B})\big|_p \le n^{-2}|\varphi^Y|_1^{1/r}|\mathbb{E}(\widetilde{H}^Y|F^n\mathcal{B})|_q \le n^{-2}|\varphi^Y|_1^{1/r}|\widetilde{H}^Y|_q.$$

Hence $\big|1_{\{\varphi^Y > n^{2r}\}}\{\widetilde{H}^Y - \mathbb{E}(\widetilde{H}^Y|F^n\mathcal{B})\}\big|_p \le 2n^{-2}|\varphi^Y|_1^{1/r}|\widetilde{H}^Y|_q$. On the other hand,

$$\big|1_{\{\varphi^Y \le n^{2r}\}}\{\widetilde{H}^Y - \mathbb{E}(\widetilde{H}^Y|F^n\mathcal{B})\}\big|_\infty \ll n^{2r}\gamma^n.$$

Combining the last two estimates, we obtain the first part of (6.1).

Next, write $\mathbb{E}(\widetilde{H}^Y|\mathcal{B}) = \zeta \circ \bar{\pi}$ where $\zeta \in L^1(\bar{Y})$. Let $P : L^1(\bar{Y}) \to L^1(\bar{Y})$ be the transfer operator associated to $\bar{F}$ (so $\int_{\bar{Y}} Pv\, w\, d\bar{\mu}_Y = \int_{\bar{Y}} v\, w \circ \bar{F}\, d\bar{\mu}_Y$ for all $w \in L^\infty(\bar{Y})$). By standard methods (see for example [35, Corollary 2.3(a)]), it follows from integrability of φ^Y and the estimates (2.5) and (2.6) that there exist constants $C > 0$, $\gamma \in (0,1)$ such that $|P^n\zeta|_\infty \le C\gamma^n$. Moreover $\mathbb{E}(\cdot|\bar{F}^{-n}\overline{\mathcal{B}}) = (UP)^n = U^n P^n$ where $Uv = v \circ \bar{F}$. Hence

$$\begin{aligned}\mathbb{E}(\widetilde{H}^Y|F^{-n}\mathcal{B}) &= \mathbb{E}(\mathbb{E}(\widetilde{H}^Y|\mathcal{B})|F^{-n}\mathcal{B}) = \mathbb{E}(\zeta \circ \bar{\pi}|F^{-n}\mathcal{B}) \\ &= \mathbb{E}(\zeta \circ \bar{\pi}|\bar{\pi}^{-1}\bar{F}^{-n}\overline{\mathcal{B}}) = \mathbb{E}(\zeta|\bar{F}^{-n}\overline{\mathcal{B}}) \circ \bar{\pi} = (U^n P^n \zeta) \circ \bar{\pi},\end{aligned}$$

and so

$$\big|\mathbb{E}(\widetilde{H}^Y|F^{-n}\mathcal{B})\big|_{L^p(Y)} = |U^n P^n \zeta|_{L^p(\bar{Y})} = |P^n\zeta|_{L^p(\bar{Y})} \le |P^n\zeta|_\infty \ll \gamma^n.$$

Hence $\big|\mathbb{E}(\widetilde{H}^Y|F^{-n}\mathcal{B})\big|_{L^p(Y)}$ is summable, completing the proof of (6.1). ∎

7 Proof of Theorems 3.1 and 3.2

In this section, we complete the proof of the main results in Section 3.

Proof of Theorem 3.1 Define the Hölder observable $h = v - I : \Lambda \to \mathbb{R}$. As in the statement of Lemma 6.2, define $H = V - I\varphi$, $\widetilde{H} = V - I\widetilde{\varphi}$, $\widetilde{H}^Y = V^Y - I\widetilde{\varphi}^Y$. By (3.2), $H \in L^p(X)$ for some $p > \alpha$ and hence by Lemma 6.2,

$$n^{-1/\alpha} \max_{j \le n} \Big| \sum_{i=0}^{j-1} \widetilde{H}^Y \circ F^i \Big| \to_p 0. \tag{7.1}$$

Hence by Lemma 5.1(a),

$$n^{-1/\alpha} \sum_{j=0}^{n-1} V^Y \circ F^j = n^{-1/\alpha} \sum_{j=0}^{n-1} (I\widetilde{\varphi}^Y + \widetilde{H}^Y) \circ F^j \to_d (\bar{\tau})^{1/\alpha} IG. \tag{7.2}$$

As in the proof of Theorem 4.1, we can suppose without loss that $F = T^{\varphi^Y} : Y \to Y$ is a first return map. As a consequence of (7.2) and Lemma 5.1(b) we can apply [25, Theorem A.1] (see Remark A.2) and it follows that $n^{-1/\alpha} \sum_{j=0}^{n-1} v \circ T^j \to_d (\int_Y \varphi^Y\, d\mu_Y)^{-1/\alpha} (\bar{\tau})^{1/\alpha} IG = (\bar{\varphi})^{-1/\alpha} IG$. (In applying Remark A.2, it should be noted that $T : \Lambda \to \Lambda$, v, φ^Y are called $f : X \to X$, V, τ in Appendix A.) ∎

Recall that $W_n^Y(t) = n^{-1/\alpha} \sum_{j=0}^{[nt]-1} V^Y \circ F^j$ is a càdlàg process on Y.

Lemma 7.1 (WIP on Y) *Under the assumptions of Theorem 3.1, $W_n^Y \to_w (\bar{\tau})^{1/\alpha} IW$ on (Y, μ_Y) in $(D[0,\infty), \mathcal{J}_1)$.*

Proof Write $V^Y = I\widetilde{\varphi}^Y + \widetilde{H}^Y$ as in the proof of Theorem 3.1. Then $W_n^Y = IA_n^Y + B_n^Y$ where A_n^Y is as in (5.1) and $B_n^Y(t) = n^{-1/\alpha}\sum_{j=0}^{[nt]-1}\widetilde{H}^Y \circ F^j$. By (7.1), for every $K > 0$ one has that $\sup_{[0,K]}|B_n^Y| \ll n^{-1/\alpha}\max_{j\le Kn}|\sum_{i=0}^{j-1}\widetilde{H}^Y \circ F^i| \to_p 0$. Hence the result follows from Lemma 5.1(c). ∎

Next, recall that $W_n^X(t) = n^{-1/\alpha}\sum_{j=0}^{[nt]-1} V \circ f^j$ is a càdlàg process on X.

Lemma 7.2 (WIP on X) *Under the assumptions of Theorem 3.1, $W_n^X \to_w IW$ on (X, μ_X) in $(D[0,\infty), \mathcal{M}_1)$.*

Proof We apply Theorem 4.1 (with $i = 1$) to pass from $V^Y : Y \to \mathbb{R}$ to $V : X \to \mathbb{R}$ via the inducing time $\tau : Y \to \mathbb{Z}^+$. (The spaces $\Omega_0 \subset \Omega$ in Theorem 4.1 correspond to the spaces $Y \subset X$ here. Similarly ϕ, Φ, ψ_n, Ψ_n, r are called V, V^Y, W_n^X, W_n^Y, τ, and the maps $S : \Omega \to \Omega$, $S_0 : \Omega_0 \to \Omega_0$ are called $f : X \to X$, $F : Y \to Y$.)

Condition 1 of Theorem 4.1 is immediate from Lemma 7.1. Define $M_1 : Y \to \mathbb{R}$,

$$M_1 = \max_{1\le \ell' \le \ell \le \tau}(V_{\ell'} - V_\ell) \wedge \max_{1\le \ell' \le \ell \le \tau}(V_\ell - V_{\ell'}),$$

where $V_\ell = \sum_{j=0}^{\ell-1} V \circ f^j$. We claim that $n^{-1/\alpha}\max_{0\le j\le n} M_1 \circ F^j \to 0$ in $L^1(Y)$. This implies condition 2 of Theorem 4.1 and the result follows.

It remains to verify the claim. Recall that $V = I\varphi + H$ and correspondingly $V^Y = I\varphi^Y + H^Y$. Define $H_\ell = \sum_{j=0}^{\ell-1} H \circ f^j$ and $H^* = |H|^Y = \sum_{\ell=0}^{\tau-1}|H| \circ f^\ell$. By assumption (3.2) and Proposition 6.1, $H^* \in L^p(Y)$ for some $p > \alpha$.

Suppose that $I > 0$ (the case $I < 0$ is similar). Then $I\varphi > 0$ and

$$M_1 \le \max_{1\le \ell' \le \ell \le \tau}(V_{\ell'} - V_\ell) \le \max_{1\le \ell' \le \ell \le \tau}(H_{\ell'} - H_\ell) \le H^*.$$

Hence

$$\int_Y (\max_{j\le n} M_1 \circ F^j)^p \, d\mu \le \int_Y \sum_{j=0}^{n}(H^* \circ F^j)^p \, d\mu = (n+1)\int_Y H^{*p} \, d\mu \ll n,$$

and so $n^{-1/p}\max_{j\le n} M_1 \circ F^j$ is bounded in $L^p(Y)$ proving the claim. ∎

Proof of Theorem 3.2 We apply Theorem 4.1 to pass from $V : X \to \mathbb{R}$ to $v : \Lambda \to \mathbb{R}$ via the return time $\varphi : X \to \mathbb{Z}^+$. (This time, the spaces $\Omega_0 \subset \Omega$ in Theorem 4.1 correspond to the spaces $X \subset \Lambda$ here. Similarly ϕ, Φ, ψ_n, Ψ_n, r are called v, V, W_n, W_n^X, φ, and the maps $S : \Omega \to \Omega$, $S_0 : \Omega_0 \to \Omega_0$ are called $T : \Lambda \to \Lambda$, $f : X \to X$. Also, M_1 and M_2 are defined as in Section 3.)

(a) Conditions 1 and 2 of Theorem 4.1 with $i = 1$ correspond to Lemma 7.2 and the assumption on M_1 respectively.

(b) Lemma 7.2 asserts convergence in the $\mathcal{M}_1$ topology and hence in the $\mathcal{M}_2$ topology, so condition 1 of Theorem 4.1 ($i = 2$) is satisfied. Condition 2 of Theorem 4.1 corresponds to the assumption on M_2. ∎

8 Billiards with cusps at flat points

We consider the Jung & Zhang example [30] described in Example 1.1. Zhang [50] showed that such billiard maps $T:\Lambda\to\Lambda$ fit in the Chernov-Markarian-Zhang framework with first return map $f=T^\varphi:X\to X$ where $X=(\Gamma_3\cup\cdots\cup\Gamma_{n_0})\times[0,\pi]$. Recall that $\alpha=\frac{\beta}{\beta-1}\in(1,2)$. Define $I_v(s)$ as in (1.1) for continuous functions $v:\Lambda\to\mathbb{R}$.

In the remainder of this section, we fix $v:\Lambda\to\mathbb{R}$ Hölder continuous with mean zero such that $I_v(\pi)>0$. Define the strictly increasing, hence invertible, function $\Psi(s)=I_1(\pi)^{-1}I_1(s)$, $s\in[0,\pi]$.

Proposition 8.1 *Let $\delta=\eta/(\beta-1)$ where η is the Hölder exponent of v. There is a constant $C>0$ such that for all $0\le\ell\le\varphi(x)$, $x\in X$,*

$$v_\ell(x)=\varphi(x)I_1(\pi)^{-1}I_v\circ\Psi^{-1}(\ell/\varphi(x))+E_\ell(x),\qquad |E_\ell(x)|\le C\varphi(x)^{1-\delta}.$$

Proof Let $\widetilde{v}(\theta)=\frac12\{v(r',\theta)+v(r'',\pi-\theta)\}$. Proceeding as in [30, Section 6: Proof of Lemma 4.4], for $0\le\ell\le\varphi/2$,

$$\begin{aligned}v_\ell=\sum_{j=0}^{\ell-1}v\circ T^j&=\sum_{j=0}^{\ell-1}\widetilde{v}\circ\Psi^{-1}(j/\varphi)+O(\varphi^{1-\delta})=\varphi\int_0^{\ell/\varphi}\widetilde{v}\circ\Psi^{-1}\,d\theta+O(\varphi^{1-\delta})\\&=\varphi I_1(\pi)^{-1}\int_0^{\Psi^{-1}(\ell/\varphi)}\widetilde{v}(\theta)(\sin\theta)^{1/\alpha}\,d\theta+O(\varphi^{1-\delta})\\&=\varphi I_1(\pi)^{-1}I_v\circ\Psi^{-1}(\ell/\varphi)+O(\varphi^{1-\delta}).\end{aligned}$$

In particular, $v_{\varphi/2}=\varphi I_1(\pi)^{-1}I_v(\pi/2)+O(\varphi^{1-\delta})$.

For $\varphi/2\le\ell\le\varphi$, using time reversibility and the estimates in [30],

$$\begin{aligned}v_\varphi-v_\ell=\sum_{j=\ell}^{\varphi-1}v\circ T^j&=\varphi\int_{\ell/\varphi}^1\widetilde{v}\circ\Psi^{-1}\,d\theta+O(\varphi^{1-\delta})\\&=\varphi I_1(\pi)^{-1}\int_{\Psi^{-1}(\ell/\varphi)}^{\Psi^{-1}(1)}\widetilde{v}(\theta)(\sin\theta)^{1/\alpha}\,d\theta+O(\varphi^{1-\delta})\qquad(8.1)\\&=\varphi I_1(\pi)^{-1}\{I_v(\pi)-I_v\circ\Psi^{-1}(\ell/\varphi)\}+O(\varphi^{1-\delta}).\end{aligned}$$

In particular, $v_\varphi-v_{\varphi/2}=\varphi I_1(\pi)^{-1}\{I_v(\pi)-I_v(\pi/2)\}+O(\varphi^{1-\delta})$, so $v_\varphi=\varphi I_1(\pi)^{-1}I_v(\pi)+O(\varphi^{1-\delta})$. Substituting the final estimate into (8.1) completes the proof. ∎

Lemma 8.2 *Conditions (3.1) and (3.2) hold.*

Proof Condition (3.1) holds by [30, Theorem 3.1]. Taking $\ell = \varphi(x)$ in Proposition 8.1, $V = I\varphi + E$ where $I = I_v(\pi)/I_1(\pi)$ and $|E| \ll \varphi^{1-\delta}$ for some $\delta > 0$. Choose $p \in (\alpha, \alpha/(1-\delta))$. Then $\varphi^{(1-\delta)p}$ is integrable by Remark 5.3 and $\int_X |E|^p \, d\mu_X \ll \int_X \varphi^{(1-\delta)p} \, d\mu_X$, so (3.2) is satisfied. ∎

Note that Theorem 1.2 is an immediate consequence of Theorem 3.1 and Lemma 8.2.

Corollary 8.3 *The conditions on $I_v(s)$ in Theorems 1.3 and 1.4 are sufficient for the WIP. In particular, Theorem 1.4 holds.*

Proof We verify the assumptions of Theorem 3.2. Conditions (3.1) and (3.2) hold by Lemma 8.2. Hence it suffices to prove that $n^{-1/\alpha} \max_{j \le n} M_i \circ f^j \to_p 0$ where $i \in \{1, 2\}$ respectively.

First suppose that $s \mapsto I_v(s)$ is nondecreasing. Note that Ψ is increasing, so $v_\ell - E_\ell$ is a nondecreasing function of ℓ. Hence by Proposition 8.1,

$$M_1 \le \max_{1 \le \ell' \le \ell \le \varphi} (v_{\ell'} - v_\ell) \le \max_{1 \le \ell' \le \ell \le \varphi} (E_{\ell'} - E_\ell) \le 2C\varphi^{1-\delta}.$$

By Proposition 3.5, $n^{-1/\alpha} \max_{j \le n} M_1 \circ f^j \to_p 0$.

Next suppose that $I_v(s) \in [0, I_v(\pi)]$ for all s. Then $v_\ell \ge E_\ell$ and $v_\varphi - v_\ell \ge E_\varphi - E_\ell$ for $0 \le \ell \le \varphi$. By Proposition 8.1,

$$M_2 \le \max_{0 \le \ell \le \varphi} (-v_\ell) + \max_{0 \le \ell \le \varphi} (v_\ell - v_\varphi) \le \max_{0 \le \ell \le \varphi} (-E_\ell) + \max_{0 \le \ell \le \varphi} (E_\ell - E_\varphi) \le 3C\varphi^{1-\delta}.$$

Again, it follows from Proposition 3.5 that $n^{-1/\alpha} \max_{j \le n} M_2 \circ f^j \to_p 0$. ∎

It remains to prove necessity of the conditions for the WIP in Theorem 1.3. We require one further result from [30].

Proposition 8.4 $n^{-1/\alpha} \max_{0 \le j \le n} \varphi \circ f^j \not\to_p 0$.

Proof Define $\nu_n = \sum_{j=1}^n \delta_{n^{-1/\alpha}\varphi \circ f^j}$. This is the expression in [30, Eq. (3.15)]. By [30, Eq. (3.7) and Section 3.2], $\mu(\nu_n((1, \infty)) = 0) \to c < 1$. In particular, $\mu(n^{-1/\alpha} \max_{j \le n} \varphi \circ f^j > 1) = \mu(\nu_n(1, \infty) \ge 1) \not\to 0$. ∎

Proof of Theorem 1.3 Suppose that $I_v(s)$ is not monotone. Then there exists $0 < s_1 < s_2 < \pi$ such that $I_v(s_2) < I_v(s_1)$.

For each $x \in X$, set $\ell_r(x) = [\varphi(x)\Psi(s_r)]$ for $r = 1, 2$. Then $0 \le \ell_1 \le \ell_2 \le \varphi$. By Proposition 8.1, $v_{\ell_r} = \varphi I_v(s_r) + O(\varphi^{1-\delta})$, so

$$v_{\ell_1} - v_{\ell_2} = c_1\varphi + O(\varphi^{1-\delta}), \qquad V = c_2\varphi + O(\varphi^{1-\delta}),$$

where $c_1, c_2 > 0$. Hence $M_1 \ge c\varphi + O(\varphi^{1-\delta})$ where $c = c_1 \wedge c_2 > 0$. By Proposition 8.4, $n^{-1/\alpha} \max_{j \le n} M_1 \circ f^j \not\to_p 0$. By Proposition 3.6, $W_n \not\to_w W$ in $\mathcal{M}_1$. The other direction was proved in Corollary 8.3 so this completes the proof. ∎

A Inducing stable laws in both directions

We assume the set up from Section 2.1 with measure-preserving transformations f, $F = f^\tau$ and $\hat{f}$ on probability spaces (X, μ_X), (Y, μ_Y) and (Δ, μ_Δ) respectively. Let $\pi : \Delta \to X$ be the measure-preserving semiconjugacy $\pi(y, \ell) = f^\ell y$ and set $\bar{\tau} = \int_Y \tau \, d\mu_Y$. We assume in addition that the probability measures μ_X, μ_Y, μ_Δ are ergodic.

In the following result, based on [25, 36], we relate limit theorems on X and Y.

Theorem A.1 *Let $V \in L^1(X)$ with $\int_X V \, d\mu_X = 0$. Define the induced observable*

$$V^Y : Y \to \mathbb{R}, \qquad V^Y = \textstyle\sum_{\ell=0}^{\tau-1} V \circ f^\ell,$$

and the Birkhoff sums

$$V_n = \textstyle\sum_{j=0}^{n-1} V \circ f^j, \qquad V_n^Y = \textstyle\sum_{j=0}^{n-1} V^Y \circ F^j, \qquad \tau_n = \textstyle\sum_{j=0}^{n-1} \tau \circ F^j, \quad n \geq 1.$$

Let G be a random variable. Let $b_n > 0$ be a sequence with $b_n \to \infty$, such that $\inf_{n\geq 1} b_n / b_{[\bar{\tau}^{-1} n + c b_n]} > 0$ for each $c > 0$. Assume that $b_n^{-1}(\tau_n - n\bar{\tau}) \to_p 0$ as $n \to \infty$. Then the following are equivalent:

(a) $b_n^{-1} V_n \to_d G$ on (X, μ_X) as $n \to \infty$.

(b) $b_n^{-1} V^Y_{[n/\bar{\tau}]} \to_d G$ on (Y, μ_Y) as $n \to \infty$.

Remark A.2 It is a special case of [25, Theorem A.1] that (b) implies (a). Moreover, instead of condition $b_n^{-1}(\tau_n - n\bar{\tau}) \to_p 0$ it suffices that $b_n^{-1}(\tau_n - n\bar{\tau})$ is tight in [25, Theorem A.1].

Proof Note that

$$\begin{aligned} \int_Y |V^Y| \, d\mu_Y \leq \int_Y \sum_{\ell=0}^{\tau-1} |V \circ f^\ell| \, d\mu_Y &= \int_Y \sum_{\ell=0}^{\tau(y)-1} |V \circ \pi(y,\ell)| \, d\mu_Y(y) \\ &= \bar{\tau} \int_\Delta |V| \circ \pi \, d\mu_\Delta = \bar{\tau} \int_X |V| \, d\mu_X < \infty. \end{aligned}$$

So $V^Y \in L^1(Y)$ and similarly $\int_Y V^Y \, d\mu_Y = 0$.

Define $\widehat{V} = V \circ \pi : \Delta \to \mathbb{R}$ and $\widehat{V}_n = \sum_{j=0}^{n-1} \widehat{V} \circ \hat{f}^j$. Since π is a measure-preserving semiconjugacy, condition (a) is equivalent to

(a′) $b_n^{-1} \widehat{V}_n \to_d G$ on (Δ, μ_Δ) as $n \to \infty$.

Note that μ_Y can be viewed as a probability measure on Δ supported on Y. As such, $\mu_Y \ll \mu_\Delta$. By Remark 3.3, we obtain that condition (a′) is equivalent to (a$''$) $b_n^{-1}\widehat{V}_n \to_d G$ on (Y,μ_Y) as $n\to\infty$.

The lap number $N_n : Y \to \mathbb{Z}^+$ is defined by the relation

$$\tau_{N_n(y)}(y) \le n < \tau_{N_n(y)+1}(y).$$

For initial conditions $y \in Y$, we write

$$\widehat{V}_n(y) = V^Y_{N_n(y)}(y) + H(\hat{f}^n y),$$

where $H : \Delta \to \mathbb{R}$ is given by $H(y,\ell) = \sum_{\ell'=0}^{\ell-1} \widehat{V}(y,\ell')$. Now

$$\begin{aligned}\mu_Y(y\in Y : b_n^{-1}|H(\hat{f}^n y)| \ge a) &= \bar\tau\mu_\Delta(y\in Y : b_n^{-1}|H(\hat{f}^n y)|\ge a)\\ &\le \bar\tau\mu_\Delta(x\in\Delta : b_n^{-1}|H(\hat{f}^n x)|\ge a) = \bar\tau\mu_\Delta(x\in\Delta : b_n^{-1}|H(x)|\ge a)\to 0\end{aligned}$$

as $n\to\infty$ since H is measurable. Hence condition (a$''$) is equivalent to

(a$'''$) $b_n^{-1}V^Y_{N_n} \to_d G$ on (Y,μ_Y) as $n\to\infty$.

It remains to prove that conditions (a$'''$) and (b) are equivalent. In other words, we must show that $b_n^{-1}(V^Y_{N_n} - V^Y_{[n/\bar\tau]}) \to_p 0$ on (Y,μ_Y).

We recall some properties of the lap number. By the ergodic theorem, $\lim_{n\to\infty} n^{-1}N_n = \bar\tau^{-1}$ a.e. Also, $\tau_k \le n$ if and only if $N_n \ge k$. Let $c>0$ and set $k = k(n) = [n/\bar\tau + cb_n]$. A calculation shows that if $b_n^{-1}|N_n - n/\bar\tau| > c$ then $b_n^{-1}|\tau_k - k\bar\tau| \ge c\bar\tau + O(b_n^{-1})$, so $b_k^{-1}|\tau_k - k\bar\tau| \ge cb_k^{-1}b_n\bar\tau + O(b_k^{-1})$. It follows from the assumptions on τ_n and b_n that

$$\mu_Y\big(b_k^{-1}|\tau_k - k\bar\tau| \ge cb_k^{-1}b_n\bar\tau + O(b_k^{-1})\big)\to 0 \quad \text{as } n\to\infty,$$

and hence that

$$b_n^{-1}(N_n - [n/\bar\tau]) \to_p 0 \quad \text{as } n\to\infty. \tag{A.1}$$

Passing to the natural extension, we can suppose without loss that F is invertible. For $n\le -1$, we write $V^Y_n = \sum_{j=n}^{-1} V^Y\circ F^j$. Then

$$V^Y_{N_n(y)}(y) - V^Y_{[n/\bar\tau]}(y) = V^Y_{\widetilde{N}_n(y)}(F^{[n/\bar\tau]}y) \quad \text{where} \quad \widetilde{N}_n(y) = N_n(y) - [n/\bar\tau].$$

Since F is measure-preserving, it suffices to show that $b_n^{-1}V^Y_{\widetilde{N}_n} \to_p 0$.

By the ergodic theorem, $n^{-1}V^Y_n \to 0$ a.e. and hence in probability as $n\to\pm\infty$. Let $\epsilon>0$. We can choose $\widetilde{Y}\subset Y$ with $\mu_Y(\widetilde{Y}) > 1-\epsilon$ and $N_0\ge 1$ such that $|n^{-1}V^Y_n| < \epsilon$ on $\widetilde{Y}$ for all $|n|\ge N_0$.

For each $n\ge 1$, define

$$Y'_n = \{y\in Y : |\widetilde{N}_n(y)| \le N_0\}, \qquad Y''_n = \{y\in Y : |\widetilde{N}_n(y)| > N_0\}.$$

For $y \in Y_n'$, we have $|V^Y_{\widetilde{N}_n(y)}| \le \Psi$, where $\Psi(y) = \sum_{j=-N_0}^{N_0-1} |V^Y(F^j y)|$. Note that $|\Psi|_1 \le 2N_0|V^Y|_1 < \infty$, so

$$\mu_Y(y \in Y_n' : |b_n^{-1} V^Y_{\widetilde{N}_n(y)}(y)| > \epsilon) \le \mu_Y(b_n^{-1}\Psi > \epsilon) < \epsilon$$

for n sufficiently large.

For $y \in Y_n'' \cap \widetilde{Y}$, we have $\left|\frac{1}{|\widetilde{N}_n|} V^Y_{\widetilde{N}_n}\right| < \epsilon$, and hence $|b_n^{-1} V^Y_{\widetilde{N}_n(y)}| < \epsilon b_n^{-1}|\widetilde{N}_n|$, so that

$$\mu_Y(y \in Y_n'' : |b_n^{-1} V^Y_{\widetilde{N}_n(y)}(y)| \ge \epsilon) \le \mu_Y(b_n^{-1}|\widetilde{N}_n| \ge 1) + \epsilon.$$

By (A.1), $b_n^{-1}|\widetilde{N}_n| \to_p 0$. Hence $b_n^{-1} V^Y_{\widetilde{N}_n} \to_p 0$, completing the proof. ∎

Remark A.3 Suppose that b_n is regularly varying of index $1/\alpha$ with $\alpha > 1$. Then the assumptions on b_n in Theorem A.1 are satisfied, and condition (b) can be restated as $b_n^{-1} V_n^Y \to_d \bar{\tau}^{1/\alpha} G$.

B The Skorohod topologies on $D[0,1]$

Let $D[0,1]$ denote the càdlàg space of right-continuous functions $g : [0,1] \to \mathbb{R}$ with left limits. The uniform topology on $D[0,1]$ is not suitable for many purposes; on the theoretical side it is not separable, and for applications it is too strong since functions must have jumps in exactly the same place in order to be close to each other.

To circumvent these issues, Skorohod [43] introduced four topologies on $D[0,1]$ that are separable and sufficiently strong for theoretical purposes, whilst being sufficiently weak to allow the flexibility for functions to be close to each other in reasonable situations. The four topologies are ordered by

$$\mathcal{J}_1 > \mathcal{J}_2 > \mathcal{M}_2 \quad \text{and} \quad \mathcal{J}_1 > \mathcal{M}_1 > \mathcal{M}_2$$

where $>$ means stronger than. The $\mathcal{M}_1$ and $\mathcal{J}_2$ topologies are not comparable. All these topologies are weaker than the uniform topology. The $\mathcal{J}_2$ topology plays no role in this paper; we define the remaining topologies below. For simplicity, we restrict to the interval $[0,1]$. (The differences between $D[0,1]$ and $D[0,\infty)$ are of a purely technical nature.) We refer the reader to [43, 47] for more details and proofs.

The Skorohod $\mathcal{J}_1$ topology The first Skorohod topology, the $\mathcal{J}_1$ topology, is metrizable and is defined through the metric d_{J_1} given by

$$d_{J_1}(g_1, g_2) = \inf_{\lambda \in \Lambda} \left\{ \|g_2 \circ \lambda - g_1\| \vee \|\lambda - id\| \right\}$$

for $g_1, g_2 \in D[0,1]$, where Λ denotes the space of strictly increasing reparametrizations mapping $[0,1]$ onto itself and $\|\cdot\|$ denotes the uniform norm. This strong topology, which coincides with the uniform topology on the subspace of continuous functions $C[0,1] \subset D[0,1]$, is suitable to define convergence of discontinuous functions when discontinuities and magnitudes of the jumps are close. For instance, if $a_n \to 1$ then the function $g_n = a_n 1_{[\frac{1}{2}-\frac{1}{n},1]}$ converges to the function $g = 1_{[\frac{1}{2},1]}$ in the $\mathcal{J}_1$ topology as $n \to \infty$. (Note that $\|g_n - g\| = |a_n|$ for all n, so there is no convergence in the uniform topology.)

The Skorohod $\mathcal{M}_1$ topology In many situations, a single jump in the limit function g corresponds to multiple smaller jumps in the functions g_n. In this paper, as in [37], the jumps of g_n are $o(1)$ and the limit function g has jumps, so a more flexible topology on $D[0,1]$ is required.

The $\mathcal{M}_1$ topology on $D[0,1]$ is again metrizable and is defined in terms of the Hausdorff distance between completed graphs of elements of $D[0,1]$. Given $g \in D[0,1]$, the *completed graph* of g is the set

$$\Gamma(g) = \{(t,s) \in [0,1] \times \mathbb{R} : s = \alpha g(t^-) + (1-\alpha) g(t),\ \alpha \in [0,1]\}.$$

Let $\Lambda^*(g)$ denotes the space of parameterizations $G = (\lambda, \gamma)\colon [0,1] \to \Gamma(g)$ such that $t' < t$ implies either $\lambda(t') < \lambda(t)$, or $\lambda(t') = \lambda(t)$ and $|\gamma(t) - g(\lambda(t))| \le |\gamma(t') - g(\lambda(t'))|$. Then the $\mathcal{M}_1$-metric is defined by

$$d_{\mathcal{M}_1}(g_1, g_2) = \inf_{G_i = (\lambda_i, \gamma_i) \in \Lambda^*(g_i)} \left\{ \|\gamma_1 - \gamma_2\| \vee \|\lambda_1 - \lambda_2\| \right\}$$

An example in the spirit of Figure 2(a) is obtained by defining $g_n = \frac{3}{4} 1_{[\frac{1}{2}-\frac{1}{n},\frac{1}{2})} + a_n 1_{[\frac{1}{2},1]}$. If $a_n \to 1$ then g_n converges to $g = 1_{[\frac{1}{2},1]}$ in the $\mathcal{M}_1$ topology as $n \to \infty$, but not in the $\mathcal{J}_1$ topology.

The Skorohod $\mathcal{M}_2$ topology The $\mathcal{M}_2$ topology on $D[0,1]$ is also defined in terms of the Hausdorff distance between completed graphs of elements of $D[0,1]$, namely $d_{\mathcal{M}_2}(g_1, g_2) = \rho(\Gamma(g_1), \Gamma(g_2)) \vee \rho(\Gamma(g_2), \Gamma(g_1))$ where

$$\rho(\Gamma(g_1), \Gamma(g_2)) = \sup_{(t_1,s_1) \in \Gamma(g_1)} \inf_{(t_2,s_2) \in \Gamma(g_2)} \|(t_1, s_1) - (t_2, s_2)\|.$$

(Here, $\|(t_1, s_1) - (t_2, s_2)\| = |t_1 - t_2| + |s_1 - s_2|$.) An example in the spirit of Figure 2(b) is obtained by defining $g_n = \frac{3}{4} 1_{[\frac{1}{2}-\frac{1}{n},\frac{1}{2})} + \frac{1}{3} 1_{[\frac{1}{2},\frac{1}{2}+\frac{1}{n})} + a_n 1_{[\frac{1}{2}+\frac{1}{n},1]}$. If $a_n \to 1$ then g_n converges to $g = 1_{[\frac{1}{2},1]}$ in the $\mathcal{M}_2$ topology as $n \to \infty$, but not in the $\mathcal{M}_1$ topology.

An example in the spirit of Figure 2(c) is obtained by defining $g_n = \frac{5}{4} 1_{[\frac{1}{2}-\frac{1}{n},\frac{1}{2})} + a_n 1_{[\frac{1}{2},1]}$, where $a_n \to 1$. Then g_n fails to converge in any of the Skorokhod topologies.

We end this appendix with the following instrumental lemma.

Lemma B.1 *Given $g \in D[a,b]$ take $\bar{g} \in D[a,b]$ given by $\bar{g} = 1_{[a,b)} g(a) + 1_{\{b\}} g(b)$. Then*

$$d_{\mathcal{M}_2,[a,b]}(g, \bar{g}) \le b - a + A \wedge B,$$

where

$$A = \sup_{t\in[a,b]} (g(a) - g(t)) + \sup_{t\in[a,b]} (g(t) - g(b)),$$
$$B = \sup_{t\in[a,b]} (g(t) - g(a)) + \sup_{t\in[a,b]} (g(b) - g(t)).$$

Proof We assume that $g(b) \geq g(a)$ (the case $g(b) < g(a)$ is entirely analogous). Then $\Gamma(\bar{g}) = \{(t, g(a)) : a \leq t \leq b\} \cup \{(b, s) : g(a) \leq s \leq g(b)\}$. Also $\Gamma(g) \subset [a, b] \times \mathbb{R}$ and intersects every horizontal line between $s = g(a)$ and $s = g(b)$.

For every $(t, s) \in \Gamma(\bar{g})$, there exists $t' \in [a, b]$ such that $(t', s) \in \Gamma(g)$. Then $\|(t, s) - (t', s)\| \leq b - a$ and hence $\rho(\Gamma(\bar{g}), \Gamma(g)) \leq b - a$.

It remains to estimate $\rho(\Gamma(g), \Gamma(\bar{g}))$. Let $(t, s) \in \Gamma(g)$.

- If $s \in [g(a), g(b)]$, then $(b, s) \in \Gamma(\bar{g})$ and $\|(t, s) - (b, s)\| \leq b - a$.
- If $s < g(a)$, then $(t, g(a)) \in \Gamma(\bar{g})$ and $g(t) \leq s < g(a)$, so $\|(t, s) - (t, g(a))\| = g(a) - s \leq g(a) - g(t) = (g(a) - g(t)) \wedge (g(b) - g(t)) \leq A \wedge B$.
- If $s > g(b)$, then $(b, g(b)) \in \Gamma(\bar{g})$ and there exists $t' \in [a, b]$ such that $g(t') \geq s > g(b)$. Hence $\|(t, s) - (b, g(b))\| \leq b - a + s - g(b) \leq b - a + g(t') - g(b) = b - a + (g(t') - g(b)) \wedge (g(t') - g(a)) \leq b - a + A \wedge B$.

In all cases, $\inf_{(\bar{t},\bar{s})\in\Gamma(\bar{g})} \|(t, s) - (\bar{t}, \bar{s})\| \leq b - a + A \wedge B$ so $\rho(\Gamma(g), \Gamma(\bar{g})) \leq b - a + A \wedge B$ completing the proof. ∎

Acknowledgements The research of IM was supported in part by a European Advanced Grant *StochExtHomog* (ERC AdG 320977) and by CNPq (Brazil) through PVE grant number 313759/2014-6. We are grateful to Adam Jakubowski for pointing out reference [6].

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 α-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] F. Bartumeus and S. A. Levin. Fractal reorientation clocks: Linking animal behavior to statistical patterns of search. *Proc. Natl. Acad. Sci. USA* **105** (2008) 19072–19077.

[6] B. Basrak and D. Krizmanić. A limit theorem for moving averages in the α-stable domain of attraction. *Stochastic Process. Appl.* **124** (2014) 1070–1083.

[7] B. Basrak, D. Krizmanić and J. Segers. A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40** (2012) 2008–2033.

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

[9] H. Bruin, S. Luzzatto and S. van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup.* **36** (2003) 621–646.

[10] H. Bruin and D. Terhesiu. Upper and lower bounds for the correlation function via inducing with general return times. *Ergodic Theory Dynam. Systems* **38** (2018) 34–62.

[11] L. A. Bunimovich and Y. G. Sinaĭ. Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.* **78** (1980/81) 479–497.

[12] L. A. Bunimovich, Y. G. Sinaĭ and N. I. Chernov. Statistical properties of two-dimensional hyperbolic billiards. *Uspekhi Mat. Nauk* **46** (1991) 43–92.

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

[14] N. Chernov. Decay of correlations and dispersing billiards. *J. Statist. Phys.* **94** (1999) 513–556.

[15] N. Chernov and R. Markarian. *Chaotic billiards*. Mathematical Surveys and Monographs **127**, American Mathematical Society, Providence, RI, 2006.

[16] N. Chernov and R. Markarian. Dispersing billiards with cusps: slow decay of correlations. *Comm. Math. Phys.* **270** (2007) 727–758.

[17] N. Chernov and H.-K. Zhang. A family of chaotic billiards with variable mixing rates. *Stoch. Dyn.* **5** (2005) 535–553.

[18] N. I. Chernov and H.-K. Zhang. Billiards with polynomial mixing rates. *Nonlinearity* **18** (2005) 1527–1553.

[19] J. De Simoi and I. P. Tóth. An expansion estimate for dispersing planar billiards with corner points. *Ann. Henri Poincaré* **15** (2014) 1223–1243.

[20] J. L. Doob. *Stochastic processes.* Wiley Classics Library, John Wiley & Sons, Inc., New York, 1990. Reprint of the 1953 original. A Wiley-Interscience Publication.

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

[22] P. Gaspard and X.-J. Wang. Sporadicity: Between periodic and chaotic dynamical behaviors. *Proc. Natl. Acad. Sci. USA* **85** (1988) 4591–4595.

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

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

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

[26] S. Gouëzel. Characterization of weak convergence of Birkhoff sums for Gibbs-Markov maps. *Israel J. Math.* **180** (2010) 1–41.

[27] A. Jakubowski. The Skorokhod Space in functional convergence: a short introduction. *International conference: Skorokhod Space. 50 years on, 17-23 June 2007, Kyiv, Ukraine, Part I, s. 11-18*, 2007.

[28] P. Jung, I. Melbourne, F. Pène, P. Varandas and H.-K. Zhang. Necessary and sufficient condition for $\mathcal{M}_2$-convergence to a Lévy process for billiards with cusps at flat points. Preprint, 2019. arXiv:1902.08958.

[29] P. Jung, F. Pène and H.-K. Zhang. Convergence to α-stable Lévy motion for chaotic billiards with several cusps at flat points. Preprint, 2018. arXiv:1809.08021.

[30] P. Jung and H.-K. Zhang. Stable laws for chaotic billiards with cusps at flat points. *Annales Henri Poincaré* **19** (2018) 3815–3853.

[31] R. Klages, S. Gallegos, J. Solanpää, M. Sarvilahti and E Räsänen. Normal and anomalous diffusion in soft Lorentz gases. *Phys. Rev. Lett.* **122** (2019) 064102.

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

[33] H. Lorentz. The motion of electrons in metallic bodies. *Proc. Amst. Acad.* **7** (1905) 438–453.

[34] R. Markarian. Billiards with polynomial decay of correlations. *Ergodic Theory Dynam. Systems* **24** (2004) 177–197.

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

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

[37] I. Melbourne and R. Zweimüller. Weak convergence to stable Lévy processes for nonuniformly hyperbolic dynamical systems. *Ann Inst. H. Poincaré (B) Probab. Statist.* **51** (2015) 545–556.

[38] R. Metzler, J-H. Jeon, A. G. Cherstvy and E. Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.* **16** (2014) 24128–24164.

[39] B. Podobnik, A. Valentinčič, D. Horvatić and H. E. Stanley. Asymmetric Lévy flight in financial ratios. *Proc. Natl. Acad. Sci. USA* **108** (2011) 17883–17888.

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

[41] G. Samorodnitsky and M. S. Taqqu. *Stable non-Gaussian random processes: Stochastic models with infinite variance.* Chapman & Hall, New York, 1994.

[42] Y. G. Sinaĭ. Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Uspehi Mat. Nauk* **25** (1970) 141–192.

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

[44] T. H. Solomon, E. R. Weeks and H. L. Swinney. Chaotic advection in a two-dimensional flow: Lévy flights and anomalous diffusion. *Phys. D* **76** (1994) 70–84.

[45] M. Thaler. A limit theorem for the Perron-Frobenius operator of transformations on $[0, 1]$ with indifferent fixed points. *Israel J. Math.* **91** (1995) 111–127.

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

[47] W. Whitt. *Stochastic-process limits.* Springer Series in Operations Research, Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.

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

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

[50] H.-K. Zhang. Decay of correlations for billiards with flat points II: cusps effect. *Dynamical systems, ergodic theory, and probability: in memory of Kolya Chernov*. Contemp. Math. **698**, Amer. Math. Soc., Providence, RI, 2017, pp. 287–316.

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

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