Superdiffusive limits beyond the Marcus regime for deterministic fast-slow systems

Ilya Chevyrev*   Alexey Korepanov†   Ian Melbourne‡

23 December 2023

Abstract

We consider deterministic fast-slow dynamical systems of the form

$$x_{k+1}^{(n)} = x_k^{(n)} + n^{-1}A(x_k^{(n)}) + n^{-1/\alpha}B(x_k^{(n)})v(y_k), \quad y_{k+1} = Ty_k,$$

where $\alpha \in (1,2)$ and $x_k^{(n)} \in \mathbb{R}^m$. Here, T is a slowly mixing nonuniformly hyperbolic dynamical system and the process $W_n(t) = n^{-1/\alpha}\sum_{k=1}^{[nt]} v(y_k)$ converges weakly to a d-dimensional α-stable Lévy process L_α.

We are interested in convergence of the m-dimensional process $X_n(t) = x_{[nt]}^{(n)}$ to the solution of a stochastic differential equation (SDE)

$$dX = A(X)\,dt + B(X)\,dL_\alpha.$$

In the simplest cases considered in previous work, the limiting SDE has the Marcus interpretation. In particular, the SDE is Marcus if the noise coefficient B is exact or if the excursions for W_n converge to straight lines as $n \to \infty$.

Outside these simplest situations, it turns out that typically the Marcus interpretation fails. We develop a general theory that does not rely on exactness or linearity of excursions. To achieve this, it is necessary to consider suitable spaces of "decorated" càdlàg paths and to interpret the limiting decorated SDE. In this way, we are able to cover more complicated examples such as billiards with flat cusps where the limiting SDE is typically non-Marcus for $m \geq 2$.

Contents

*School of Mathematics, The University of Edinburgh, Edinburgh, EH9 3FD, UK. ichevyrev@gmail.com

†Department of Mathematical Sciences, Loughborough University, Loughborough, Leicestershire, LE11 3TU, UK. khumarahn@gmail.com

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

1   Introduction

Homogenisation for systems with multiple timescales is a longstanding and very active area of research [36]. Recently there has been considerable interest in the case where the underlying multiscale system is deterministic, see [13, 15, 16, 17, 21, 26, 27, 28, 34, 37] as well as our survey paper [14]. Although the system is deterministic, chaotic dynamics often leads to a limiting stochastic differential equation (SDE) driven by Brownian motion.

We are interested in the case where the limiting SDE is driven by a superdiffusive α-stable Lévy process. This situation was considered previously in [13, 21] though, as we shall see, under restrictions that gave a misleadingly simple picture.

Let $\alpha \in (1, 2)$. The multiscale equations that we consider have the form

$$\begin{cases} x_{k+1}^{(n)} = x_k^{(n)} + n^{-1}A(x_k^{(n)}) + n^{-1/\alpha}B(x_k^{(n)})v(y_k), \\ y_{k+1} = Ty_k \end{cases} \tag{1.1}$$

defined on $\mathbb{R}^m \times \Lambda$ where Λ is a bounded metric space. Here,

$$A : \mathbb{R}^m \to \mathbb{R}^m, \quad B : \mathbb{R}^m \to \mathbb{R}^{m \times d}, \quad v : \Lambda \to \mathbb{R}^d, \quad T : \Lambda \to \Lambda.$$

It is assumed that the fast dynamical system $T : \Lambda \to \Lambda$ has an ergodic invariant probability measure μ_Λ. Also, $v : \Lambda \to \mathbb{R}^d$ is Hölder with $\int v \, d\mu_\Lambda = 0$. Define

$$W_n(t) = n^{-1/\alpha} \sum_{j=0}^{[nt]-1} v \circ T^j. \tag{1.2}$$

Then W_n belongs to $D([0, 1], \mathbb{R}^d)$, the Skorokhod space of càdlàg functions, and can be viewed as a random process on the probability space (Λ, μ_Λ) depending on the initial condition $y_0 \in \Lambda$. We assume that T and v exhibit superdiffusive behaviour, so that the process W_n converges weakly in $D([0, 1], \mathbb{R}^d)$ to an α-stable Lévy process L_α.

Now fix $x_0^{(n)} = \xi \in \mathbb{R}^m$, and solve (1.1) to obtain $(x_k^{(n)}, y_k)$ depending on the initial condition $y_0 \in (\Lambda, \mu_\Lambda)$. Define the càdlàg process $X_n \in D([0, 1], \mathbb{R}^m)$ given by $X_n(t) = x_{[nt]}^{(n)}$; again we view this as a process on (Λ, μ_Λ). Our aim is to show, under mild regularity assumptions on

the functions $A : \mathbb{R}^m \to \mathbb{R}^m$ and $B : \mathbb{R}^m \to \mathbb{R}^{m \times d}$, that $X_n \to_{\mu_\Lambda} X$ [1] where X is the solution of the SDE

$$dX = A(X)\, dt + B(X)\, dL_\alpha, \qquad X(0) = \xi.$$

A key part of the theory is to explain the sense in which the limiting SDE is to be interpreted.

In previous work [21, 13], the limiting SDE had the Marcus interpretation [31] (see [29, 3, 7] for the general theory of Marcus SDEs and their applications). The Marcus integral is the analogue of Stratonovich in the Lévy case: it is the integral that transforms under the standard laws of calculus. Gottwald and Melbourne [21, Section 5] deals with the special case $d = m = 1$ and more generally the case of exact noise: $d = m$ and $B = (Dr)^{-1}$ for some diffeomorphism $r : \mathbb{R}^m \to \mathbb{R}^m$. However, exactness typically fails outside the case $d = m = 1$ and then [21] does not apply.

In [13], we gave general conditions under which there is convergence to a Marcus SDE, but a crucial assumption was that W_n converges to L_α in the $\mathcal{M}_1$ Skorokhod topology. To circumvent nonexactness, we used topologies inspired by rough paths: Weak convergence of W_n no longer suffices and we required in addition that W_n is tight in p-variation. Moreover, although convergence of W_n was assumed in the $\mathcal{M}_1$ topology, X_n need not converge in any of the Skorokhod topologies [41] even in the simplest situations [13]. Hence we needed to consider convergence of X_n in generalised Skorokhod topologies as introduced recently in Chevyrev and Friz [12]. In Theorem 3.7, we extend the result in [13] to include general "linear excursions" for W_n (by which we mean that each excursion is asymptotically a line segment, possibly with overshoots). This includes the situations where W_n converges weakly in the $\mathcal{M}_1$ or $\mathcal{M}_2$ topologies.

When $d \geq 2$, linear excursions for W_n are highly unlikely unless there is a very simple mechanism underlying the excursions, such as a neutral fixed point for an intermittent map. Moreover, as shown in Example 1.1 below, typically the limiting SDE is not Marcus when B is not exact and the excursions are not linear. In this paper, we present a general theory under which the slow dynamics converges to an SDE driven by L_α regardless of the nature of the convergence of W_n. Moreover, we give a full description of how to interpret the limiting SDE for both the Marcus and non-Marcus cases.

Example 1.1. We show that in the simplest nonexact examples, convergence to a Marcus integral typically fails unless the excursions of W_n are linear. Take $d = m = 2$, $A \equiv 0$, $\xi = 0$ and

$$B : \mathbb{R}^2 \to \mathbb{R}^{2 \times 2}, \qquad B(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ 0 & x_1 \end{pmatrix}.$$

Then $dX = A(X)\, dt + B(X)\, dL$ becomes

$$dX^1 = dL^1, \qquad dX^2 = dX^1 dL^2, \qquad X(0) = 0.$$

In particular, $X^1(t) = L^1(t)$ and $X^2(t) = \int_0^t L^1(t)\, dL^2(t)$.

Choose $h = (h_1, h_2) \in C([0,1], \mathbb{R}^2)$ where $h(0) = (0,0)$, $h(1) = (1,1)$ and h_2 is C^1. We define $W_n \in D([0,1], \mathbb{R}^2)$ by

$$W_n(t) = \begin{cases} (0,0) & \text{if } 0 \leq t < \frac{1}{2} - \frac{1}{n} \\ h\big(n(t - \frac{1}{2} + \frac{1}{n})\big) & \text{if } \frac{1}{2} - \frac{1}{n} \leq t < \frac{1}{2} \\ (1,1) & \text{if } \frac{1}{2} \leq t \leq 1 \end{cases}.$$

---

[1]We write $\to_{\mu_\Lambda}$ to denote weak convergence, emphasising the probability space on which X_n are defined. As in [13], μ_Λ can be replaced by any probability measure absolutely continuous with respect to μ_Λ.

Then W_n converges pointwise to $L \in D([0,1], \mathbb{R}^2)$ given by $L(t) = \mathbf{1}_{t \geq \frac{1}{2}}(1,1)$. Moreover, a direct calculation shows that the solutions X_n to the ODE $dX_n = B(X_n)\,dW_n$ converge pointwise to $X \in D([0,1], \mathbb{R}^2)$ where $X(t) = \mathbf{1}_{t \geq \frac{1}{2}}\left(1, \int_0^1 h_1(s)\,h_2'(s)\,ds\right)$.

We distinguish the case where the excursion of W_n is linear, so $\{h(s); 0 \leq s \leq 1\}$ is contained in a line. In other words, $h_1 \equiv h_2$. This includes $\mathcal{M}_1$ and $\mathcal{M}_2$ convergence as special cases. Indeed W_n converges to L in the $\mathcal{M}_2$ topology if and only if $\{h(s); 0 \leq s \leq 1\}$ is the line segment joining $(0,0)$ to $(1,1)$. For convergence in the $\mathcal{M}_1$ topology, it is required moreover that h_1 is non-decreasing.

In the linear case $h_1 \equiv h_2$, we obtain by a change of variables that $X(t) = \mathbf{1}_{t \geq \frac{1}{2}}\left(1, \frac{1}{2}\right)$ and the stochastic integral $dX = B(L)\,dL$ has the Marcus interpretation. Otherwise, even though the endpoints of the excursion for W_n on $[\frac{1}{2} - \frac{1}{n}, \frac{1}{2}]$ are fixed and the limiting path L is fixed, the values of the limiting integral X on $[\frac{1}{2}, 1]$ depend sensitively on the excursion h.

Decorated paths Example 1.1 shows that it is necessary to include excursions in the description of the limiting path L and to consider spaces of decorated càdlàg paths. Such spaces were introduced by Whitt [44] and studied further by Freitas *et al.* [19, 18]. Here, we build on and generalise the spaces used by Chevyrev *et al.* [11, 12, 13] which sufficed for the Marcus limits in [13]. As in [13], weak convergence of W_n is insufficient to deduce convergence of X_n and we also require tightness of W_n in p-variation.

Remark 1.2. The specific decorated càdlàg space $\mathscr{D}$ that we introduce in Section 2.1 is related to but different from the spaces F in [44, Sec. 15.7] and F' in [19, 18]. For our purposes, the main advantage is that it is well-suited for studying fast-slow systems. But it also distinguishes more decorated paths than F' (see Example 2.3) which is already finer than F (see [19, Section 1.1]). In addition, our metric α_∞ on $\mathscr{D}$ is complete and separable whereas the metrics considered in [19, 44] are not simultaneously complete and separable.

Interestingly, there are natural embeddings $\iota, \jmath : D \hookrightarrow \mathscr{D}$ under which the metric α_∞ induces the $\mathcal{J}_1$ and $\mathcal{M}_1$ topologies on D respectively (see Remark 2.8). This shows that our construction simultaneously generalises the Skorokhod $\mathcal{J}_1$ and $\mathcal{M}_1$ topologies. Moreover, the first embedding ι is dense, showing that $(\mathscr{D}, \alpha_\infty)$ is isometric to the completion of D under the standard $\mathcal{J}_1$-metric (Corollary 8.15).

Solving decorated SDEs A key part of our paper is to show how to solve equations driven by decorated processes. In the framework of Example 1.1, we argue that W_n converges to a decorated process $\hat{L}$ consisting of L together with an excursion $\{h(s); 0 \leq s \leq 1\}$ at the discontinuity point $t = \frac{1}{2}$. Solving the equation $d\widehat{X} = B(\widehat{X})\,d\hat{L}$ yields a decorated process $\widehat{X}$, and we show that X_n converges weakly to $\widehat{X}$.

To solve $d\widehat{X} = B(\widehat{X})\,d\hat{L}$ we proceed as follows: First, we insert an interval $[\frac{1}{2}, \frac{1}{2} + \delta]$ into the interval $[0,1]$ and define a δ-*extension* $L^\delta \in D([0, 1+\delta], \mathbb{R}^2)$ coinciding with L on $[0, \frac{1}{2}] \cup [\frac{1}{2} + \delta, 1 + \delta]$ (suitably displaced) and given by h (suitably scaled) on $[\frac{1}{2}, \frac{1}{2} + \delta]$. Then we solve the Young equation $X^\delta = B(X^\delta)\,dL^\delta$ for $X^\delta \in D([0, 1+\delta], \mathbb{R}^2)$. Finally, we collapse the inserted interval $[\frac{1}{2}, \frac{1}{2} + \delta]$ back to a point $\frac{1}{2}$ to obtain the desired decorated limit $\widehat{X}$ consisting of a càdlàg path on $[0,1]$ and an excursion at $t = \frac{1}{2}$. Precise definitions are given in Section 2.

Example 1.3. The motivating example for the results in this paper is a class of dispersing billiards with flat cusps introduced by [25]. The billiard table $Q \subset \mathbb{R}^2$ has a boundary consisting of C^3 curves with at least one flat cusp. In local coordinates $(s, z) \in \mathbb{R}^2$, the i'th cusp lies at $(0,0)$ and the bounding curves are given by $\Gamma_\pm = \{(s, \pm s^{\beta_i})\}$ close to $(0,0)$, where $\beta_i > 2$. We let $\beta = \max \beta_i$ corresponding to the flattest cusp(s).

The phase space of the billiard map (collision map) T is given by $\Lambda = \partial Q \times [-\pi/2, \pi/2]$, with coordinates (r, θ) where r denotes arc length along ∂Q and θ is the angle between the normal to the boundary and the collision vector in the clockwise direction. (A standard reference for billiard maps is [9].) There is a natural ergodic T-invariant probability measure $d\mu_\Lambda = (2|\partial Q|)^{-1} \cos\theta \, dr \, d\theta$ on Λ, where $|\partial Q|$ is the length of ∂Q.

Let $v : \Lambda \to \mathbb{R}$ be a Hölder observable with $\int_\Lambda v \, d\mu_\Lambda = 0$ and define W_n as in (1.2). By [25, 24], $W_n(1)$ converges weakly to an α-stable law with $\alpha = \beta/(\beta-1) \in (1, 2)$. Convergence to the corresponding Lévy process is considered in [33, 24, 23]. Such convergence is impossible in the standard Skorokhod $\mathcal{J}_1$ topology since the jumps are bounded. If v has constant sign on each cusp, then convergence holds in the $\mathcal{M}_1$ topology. For more general v, convergence of W_n might hold only in the $\mathcal{M}_2$ topology [33] or may fail in any Skorokhod topology [23, 33]. Indeed, the latter is typically the case when $d \geq 2$. Nevertheless, convergence always holds in a suitable decorated path space [18].

For purposes of illustration, consider the case of a billiard with a single flat cusp as shown below.
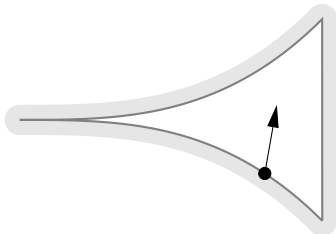
Take $\beta = 3$, $\alpha = \frac{3}{2}$ and $v : \Lambda \to \mathbb{R}^2$ to be the mean zero observable $v(r, \theta) = (\cos 3\theta, \cos 5\theta)$. We show a parametric plot of a computer-simulated $W_n(t)$, $0 \leq t \leq 1$, with separate plots for its coordinates $W_n^1(t)$, $W_n^2(t)$.[2]
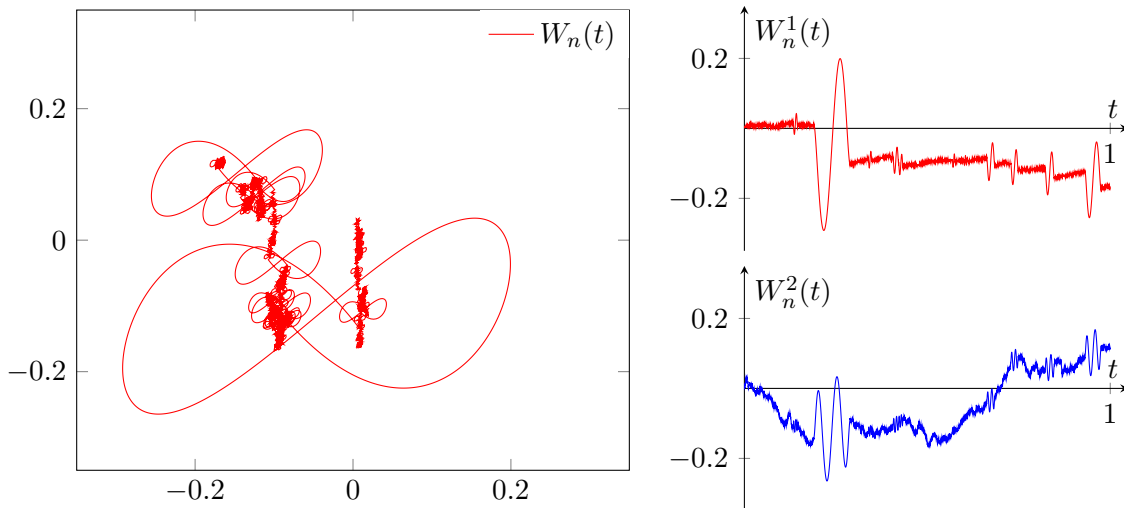
Every deep excursion into the cusp resembles the figure ∞ (or a butterfly) with the start and end points close to the centre. Despite the geometric simplicity of the example, such excursions are clearly incompatible with the Skorokhod topologies. Moreover, the excursions are not linear and hence outside the regime where one should expect a Marcus SDE.

Our main results incorporate a large class of nonuniformly hyperbolic systems with superdiffusive behaviour, including billiards with flat cusps. For such systems, we recover the

---

[2]The data can be accessed at `https://khumarahn.github.io/CKM23`.

result of [18] that W_n converges weakly to a decorated α-stable Lévy process. Moreover, we relax certain assumptions in [18]: we do not require that endpoints of excursions are in distinct directions, nor that they are nonzero. In addition, we verify tightness of W_n in p-variation (extending the corresponding result in [13] to a larger class of dynamical systems). Theorem 3.6 guarantees that these are the ingredients required to show weak convergence of X_n to solutions of a correspondingly decorated SDE.

The remainder of this paper is organised as follows. In Section 2, we define our space of decorated càdlàg paths and explain rigorously how to solve SDEs driven by such decorated paths. Then in Section 3, we state our main results. Sections 4 to 7 are devoted to the proofs of these main results. In Section 8, we prove the results from Section 2 for the decorated path space.

Notation We use "big O" and $\ll$ notation interchangeably, writing $a_n = O(b_n)$ or $a_n \ll b_n$ if there are constants $C > 0$, $n_0 \geq 1$ such that $a_n \leq C b_n$ for all $n \geq n_0$. As usual, $a_n = o(b_n)$ means that $a_n/b_n \to 0$ and $a_n \sim b_n$ means that $a_n/b_n \to 1$.

For a function $h : [a, b] \to \mathbb{R}^d$, we denote $h(t^-) = \lim_{s \uparrow t} h(s)$ whenever this limit exists. We let $\to_w$ denote weak convergence of random variables. We write $\mathbb{Z}^+ = \{1, 2, \ldots\}$.

2 Decorated path space and p-variation

In this section we define the space $\mathscr{D}$ of decorated càdlàg paths with metric α_∞. We also define the enhanced space $\mathscr{D}^{p\text{-var}}$ with metric $\alpha_{p\text{-var}}$ based on p-variation. We then show how to construct solutions X to differential equations driven by elements $\phi \in \mathscr{D}^{p\text{-var}}$, and end with Theorem 2.10 which states that the solution map $\phi \mapsto X = \mathcal{S}(\phi)$ preserves weak convergence.

Proofs of the results in this section are delayed to Section 8 where independence of definitions on various choices is also verified.

2.1 Decorated path space

Fix $d \geq 1$. Let $I = [a, b] \subset \mathbb{R}$ be a bounded closed interval. We denote by $C(I, \mathbb{R}^d)$ and $D(I, \mathbb{R}^d)$ the spaces of continuous and càdlàg functions equipped respectively with the uniform norm $|h|_{\infty;I} = \sup_{x \in I} |h(x)|$ and the Skorokhod $\mathcal{J}_1$-metric

$$\sigma_{\infty;I}(h_1, h_2) = \inf_{\lambda \in \Lambda} \max\{|\lambda - \mathrm{Id}|_\infty, |h_1 \circ \lambda - h_2|_\infty\}.$$

Here, Λ is the set of continuous increasing bijections $\lambda : I \to I$. We will drop I from our notation in norms and metrics whenever it is clear from the context.

In the remainder of this subsection, we suppress the dependence on $\mathbb{R}^d$, writing $C(I)$, $D(I)$ and so on. Also, we write for instance $D[a, b]$ instead of $D([a, b])$

Consider a function $\phi : I \to D[0, 1]$. If $\phi(t)$ is a constant path in $D[0, 1]$, then we say that $t \in I$ is a *stationary* point for ϕ.

Definition 2.1. We let $\bar{\mathscr{D}}(I)$ denote the space of pairs (ϕ, Π) where $\phi : I \to D[0, 1]$ and

(a) the function $t \mapsto \phi(t)(1)$ lies in $D(I)$,

(b) $\Pi \subset I$ is an at most countable set (possibly empty) containing the non-stationary points of ϕ,

(c) for all $\varepsilon > 0$, there exist only finitely many $t \in \Pi$ such that $\sup_{s \in [0,1]} |\phi(t)(s) - \phi(t)(0)| > \varepsilon$.

Let $(\phi, \Pi) \in \bar{\mathscr{D}}[a, b]$ and write $\Pi = \{t_j : 1 \le j \le \kappa\}$ where $\kappa \in \{0, 1, \dots\} \cup \{\infty\}$. Given $\delta > 0$, we define the δ-extension $\phi^\delta \in D[a, b+\delta]$ as follows. When $\kappa = 0$, we define $\phi^\delta(t) = \phi(t)$ for $t \in [a, b]$ and $\phi^\delta(t) = \phi(b)$ for $t \in [b, b+\delta]$. Otherwise, define $r = \sum_{j=1}^\kappa 2^{-j} > 0$. We insert in $[a, b]$ *fictitious* intervals I_j of length $2^{-j}\delta/r$ after each t_j to obtain the interval $[a, b+\delta]$. Define $\phi^\delta : [a, b+\delta] \to \mathbb{R}^d$ to coincide with ϕ on $[a, b+\delta] \setminus \bigcup_j I_j$ and to coincide with the appropriate time-scaled version of $\phi(t_j)$ on I_j. (So if $I_j = [c_j, c_j + 2^{-j}\delta/r] \subset [a, b+\delta]$, then $\phi^\delta(c_j + s) = \phi(t_j)(s2^j r/\delta)$ for $0 \le s \le 2^{-j}\delta/r$.) See Section 8.1 for a more precise definition. By Lemma 8.4, $\phi^\delta \in D[a, b+\delta]$ for all $(\phi, \Pi) \in \bar{\mathscr{D}}[a, b]$.

For $(\phi_1, \Pi_1), (\phi_2, \Pi_2) \in \bar{\mathscr{D}}[a, b]$ we set

$$\alpha_{\infty;[a,b]}(\phi_1, \phi_2) = \lim_{\delta \to 0} \sigma_{\infty;[a,b+\delta]}(\phi_1^\delta, \phi_2^\delta).$$

By Lemma 8.12(i), this limit exists and is independent of Π_1, Π_2 and the various other choices above. From now on we speak of $\phi \in \bar{\mathscr{D}}[a, b]$ and suppress the set Π.

We are now in a position to define our metric space of decorated càdlàg paths. Given $h_1, h_2 : [0, 1] \to \mathbb{R}^d$, we say that h_1 is a *reparametrisation* of h_2 if $\inf_{\lambda \in \Lambda} |h_1 \circ \lambda - h_2|_\infty = 0$. Say that $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a, b]$ are *equivalent* if the paths $s \mapsto \phi_i(t^-)(1)\mathbf{1}_{s=0} + \phi_i(t)(s)\mathbf{1}_{s>0}$ for $i = 1, 2$ are reparametrisations of each other for all $t \in [a, b]$.

Definition 2.2. We let $\mathscr{D}[a, b] = \mathscr{D}([a, b], \mathbb{R}^d)$ be the space of equivalence classes in $\bar{\mathscr{D}}[a, b]$.

By Lemma 8.12(i) and Theorem 8.14, α_∞ is constant on equivalence classes and defines a complete separable metric on $\mathscr{D}[a, b]$. Elements of $\mathscr{D}[a, b]$ are called *decorated paths*, generalising a similar notion introduced in [11, 12] called path functions.

Example 2.3. The following are examples of decorated paths that are distinct in $\mathscr{D}$ but identified in the space F' considered in [19, 18]:

$$\phi(t)(s) = \begin{cases} \sin(2\pi s) & \text{if } t = 1/3 \\ \sin(4\pi s) & \text{if } t = 2/3 \\ 0 & \text{else} \end{cases} \quad \text{and} \quad \psi(t)(s) = \begin{cases} \sin(4\pi s) & \text{if } t = 1/3 \\ \sin(2\pi s) & \text{if } t = 2/3 \\ 0 & \text{else} \end{cases}$$
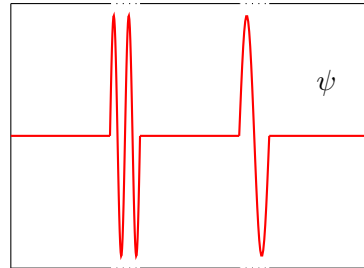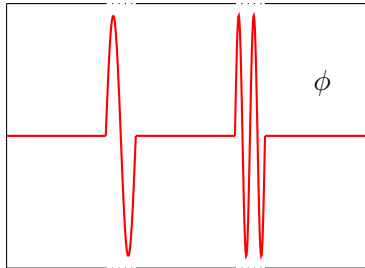
We end this subsection by recording some elementary properties of $\mathscr{D}$ which are used in Section 6. The following result is immediate from the definitions.

Proposition 2.4. *Let $a < c < b$. Then*

$$\alpha_{\infty;[a,b]}(\phi_1, \phi_2) \le \max\{\alpha_{\infty;[a,c]}(\phi_1, \phi_2), \alpha_{\infty;(c,b]}(\phi_1, \phi_2)\}$$

for all $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a, b]$.[3] ∎

Definition 2.5. Define the embedding $\iota : D(I) \hookrightarrow \bar{\mathscr{D}}(I)$ by setting $\iota h(t)(s) = h(t)$. We call ιh the *trivial lift* of h.

---

[3] We can define α_∞ on $\bar{\mathscr{D}}(c, b]$ with the obvious modifications.

Remark 2.6. Let $h_1, h_2 \in D(I)$. Then $\alpha_\infty(\iota h_1, \iota h_2) = \sigma_\infty(h_1, h_2) \leq |h_1 - h_2|_\infty$.

Proposition 2.7. *Let $\phi \in \bar{\mathscr{D}}(I)$ such that $\phi(t)(s) \equiv \phi(a)(0)$ for all $t \in [a, b)$, $s \in [0, 1]$, and such that $\phi(b) \in D[0, 1]$. Define $h \in D[a, b]$, $h(s) = \phi(b)(\frac{s-a}{b-a})$. Then $\alpha_\infty(\phi, \iota h) \leq b - a$.*

Proof. We take $\Pi = \{b\}$. Note that $(\iota h)^\delta$ is a reparametrisation of ϕ^δ on $[a, b + \delta]$ so

$$\sigma_\infty(\phi^\delta, (\iota h)^\delta) \leq b - a + \delta.$$

The result follows by definition of α_∞. ∎

Remark 2.8. There is another embedding $\jmath : D[a, b] \hookrightarrow \mathscr{D}[a, b]$ for which $(\jmath h)(t)(s) = (1 - s)h(t^-) + sh(t)$, the linear path from $h(t^-)$ to $h(t)$. Then the topology on $D[a, b]$ induced by α_∞ via the embedding $\jmath$ is the $\mathcal{M}_1$ topology, see [12, Proposition 2.9]. Together with Remark 2.6, this shows that $(\mathscr{D}[a, b], \alpha_\infty)$ generalises simultaneously the $\mathcal{J}_1$ and the $\mathcal{M}_1$ topologies on $D[a, b]$ via the embeddings $\iota, \jmath : D[a, b] \hookrightarrow \mathscr{D}[a, b]$ respectively.

2.2 p-variation

Fix $p \in [1, \infty]$. The p-variation semi-norm of a function $h : [a, b] \to \mathbb{R}^d$ is given by

$$|h|_{p\text{-var};[a,b]} = \sup_{a=t_0 < t_1 < \cdots < t_k = b} \Big(\sum_{j=1}^k |h(t_j) - h(t_{j-1})|^p \Big)^{1/p}, \tag{2.1}$$

understood as $\sup_{s,t \in [a,b]} |h_s - h_t|$ if $p = \infty$. Define also the norm $\|h\|_{p\text{-var}} = |h(a)| + |h|_{p\text{-var}}$.

Now let $\phi \in \bar{\mathscr{D}}[a, b]$. Recalling the definition of $\phi^\delta \in D[a, b + \delta]$, we set $|\phi|_{p\text{-var}} = |\phi^\delta|_{p\text{-var}}$. This is well-defined, independent of δ, and constant on equivalence classes by Lemma 8.11. We can therefore define $|\phi|_{p\text{-var}}$ for $\phi \in \mathscr{D}[a, b]$. Let

$$\mathscr{D}^{p\text{-var}}[a, b] = \{\phi \in \mathscr{D}[a, b] : |\phi|_{p\text{-var}} < \infty\}. \tag{2.2}$$

Similarly, define $\bar{\mathscr{D}}^{p\text{-var}}[a, b]$.

Remark 2.9. Let $h \in D[a, b]$ with trivial lift $\iota h \in \mathscr{D}[a, b]$. Then $|\iota h|_{p\text{-var}} = |h|_{p\text{-var}}$.

Next, we introduce a Skorokhod version of p-variation. For $h_1, h_2 \in D[a, b]$, define the metric

$$\sigma_{p\text{-var}}(h_1, h_2) = \inf_{\lambda \in \Lambda} \max \big\{|\lambda - \mathrm{Id}|_\infty, \|h_1 \circ \lambda - h_2\|_{p\text{-var}}\big\}.$$

For $\phi_1, \phi_2 \in \bar{\mathscr{D}}^{p\text{-var}}[a, b]$, we set

$$\alpha_{p\text{-var}}(\phi_1, \phi_2) = \lim_{\delta \to 0} \sigma_{p\text{-var}}(\phi_1^\delta, \phi_2^\delta).$$

Again, by Lemma 8.12(ii) this limit exists and is constant on equivalence classes. In this way, we define the metric $\alpha_{p\text{-var}}$ on $\mathscr{D}^{p\text{-var}}[a, b]$.

2.3 Differential equations driven by decorated paths

Fix $m, d \geq 1$ and $p \in [1, 2)$. For $\gamma > 0$, let $C^\gamma(\mathbb{R}^m, \mathbb{R}^d)$ denote the space of functions $f : \mathbb{R}^m \to \mathbb{R}^d$ such that

$$\|f\|_{C^\gamma} = \max_{|\alpha|=0,\ldots,[\gamma]} \|D^\alpha f\|_\infty + \sup_{x \neq y \in \mathbb{R}^m} \max_{|\alpha|=[\gamma]} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x - y|^{\gamma - [\gamma]}} < \infty.$$

(Note $f \in C^N$ for $N \in \mathbb{Z}^+$ implies only that the $(N-1)$-th derivative of f is Lipschitz.) Assume that $A \in C^\beta(\mathbb{R}^m, \mathbb{R}^m)$ for $\beta > 1$ and $B \in C^\gamma(\mathbb{R}^m, \mathbb{R}^{m \times d})$ for $\gamma > p$. Let $\xi \in \mathbb{R}^m$. We consider initial value problems of the form

$$dX = A(X)\, dt + B(X)\, d\phi, \quad X(0) = \xi, \tag{2.3}$$

where $\phi \in \mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^d)$, $X \in \mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^m)$.

To interpret and solve (2.3), consider the decorated path $(\omega, \Pi) = ((\iota\mathrm{Id}, \phi), \Pi) \in \bar{\mathscr{D}}^{p\text{-var}}([0,1], \mathbb{R}^{1+d})$, where $\mathrm{Id} : [0,1] \to [0,1]$ is the identity map. Consider also a δ-extension $\omega^\delta \in D^{p\text{-var}}([0, 1+\delta], \mathbb{R}^{1+d})$, $\delta > 0$, as described in Section 2.1. There is a unique solution $X^\delta \in D^{p\text{-var}}([0, 1+\delta], \mathbb{R}^m)$ to the initial value problem

$$dX^\delta = (A, B)(X^\delta)\, d\omega^\delta, \quad X^\delta(0) = \xi$$

defined as a Young ODE, see Section 8.3. Define $(X, \Pi) \in \bar{\mathscr{D}}^{p\text{-var}}([0,1], \mathbb{R}^m)$ with $X(t_j) \in D^{p\text{-var}}([0,1], \mathbb{R}^m)$ for $t_j \in \Pi$ given by the suitably rescaled version of X^δ restricted to the interval $I_j = [c_j, c_j + 2^{-j}\delta/r]$ (see Section 8.3 for details). By construction, X^δ is precisely the δ-extension of (X, Π).

As shown in Proposition 8.16, the resulting map

$$\mathcal{S} : \mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^d) \to \mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^m), \qquad \phi \mapsto X,$$

is well-defined and continuous (with respect to the $\alpha_{p\text{-var}}$ topology on domain and range). We call X the *solution* of (2.3) and $\mathcal{S}$ the *solution map*.

The following is our main convergence criterion for ODEs driven by random decorated paths, the proof of which we give at the end of Section 8.3.

Theorem 2.10. *Suppose ϕ_n is a sequence of random variables in $\mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^d)$ and ϕ is a random variable in $\mathscr{D}([0,1], \mathbb{R}^d)$, such that $\phi_n \to_w \phi$ in the α_∞ topology. Suppose further that $|\phi_n|_{p\text{-var}}$ is tight.*

Then $|\phi|_{p\text{-var}} < \infty$ a.s. and the solutions $X_n = \mathcal{S}(\phi_n)$, $X = \mathcal{S}(\phi) \in \mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^m)$ satisfy $X_n \to_w X$ in the $\alpha_{q\text{-var}}$ topology for all $q > p$.

3 Dynamical setup and main results

In this section, we state the main result in this paper.

3.1 Assumptions

Let Λ be a compact metric space with Borel probability measure μ_Λ, and let $T : \Lambda \to \Lambda$ be an ergodic and mixing measure-preserving dynamical system. Fix $\Sigma \subset \Lambda$ with $\mu_\Lambda(\Sigma) > 0$. We define the *first return time*

$$R : \Sigma \to \mathbb{Z}^+, \qquad R(x) = \inf\{n \geq 1 : T^n x \in \Sigma\},$$

and the *first return map*

$$f : \Sigma \to \Sigma, \qquad fx = T^{R(x)}x.$$

The map $f : \Sigma \to \Sigma$ preserves the probability measure $\mu = \mu_\Lambda|_\Sigma / \mu_\Lambda(\Sigma)$.

As described below, we assume that T falls within the so-called Chernov-Markarian-Zhang framework [32, 10]. Informally this means that (i) we have good control on the first return time R, and (ii) f is modelled by a Young tower with exponential tails.

This framework incorporates uniformly expanding/hyperbolic systems such as the intermittent maps [40] as analysed in [13], but is flexible enough to include nonuniformly hyperbolic systems such as the billiards with flat cusps [25] from Example 1.3.

Assumptions on the first return time. Our first assumption is that R is regularly varying with exponent $\alpha \in (1, 2)$. That is

$$\mu(R > t) \sim \ell(t)t^{-\alpha} \quad \text{as } t \to \infty \tag{3.1}$$

where $\ell : (0, \infty) \to (0, \infty)$ is a slowly varying function.

Introduce $b_n > 0$ such that $b_n^\alpha \sim n\ell(b_n)$ as $n \to \infty$. We require that for every $s > 0$, there exist $\theta_1 > 1$ and $C > 0$ such that

$$\mu(R > sb_n \text{ and } R \circ f^j > sb_n) \leq Cn^{-\theta_1} \quad \text{for all } 1 \leq j \leq n. \tag{3.2}$$

This assumption means that large first returns are not too clustered.

Assumptions on the first return map. We suppose that (f, Σ, μ) is ergodic and mixing and modelled by a Young tower with exponential tails [46] as discussed further in Section 4. In particular,

- There is an inducing set $Y \subset \Sigma$ with $\mu(Y) > 0$ and a return time $\tau : Y \to \mathbb{Z}^+$ (not necessarily the first return time) such that $F = f^\tau$ maps Y into Y;

- The tail probabilities $\mu(\tau > n)$ decay exponentially;

- Associated to F, there is a countable partition $\{Y_j : j \geq 1\}$ of Y and τ is constant on elements Y_j.

We require that R is constant on elements of

$$\mathcal{C} = \{f^\ell Y_j : 0 \leq \ell < \tau|_{Y_j}, j \geq 1\}.$$

Assumptions on excursions. Let Σ_i, $i = 1, \ldots, e$, be a finite collection of disjoint subsets of Σ such that the Σ_i are unions of elements of $\mathcal{C}$. (When $T : \Lambda \to \Lambda$ is a billiard with flat cusps, the sets Σ_i correspond to trajectories entering the i'th cusp, restricting to flattest cusps.)

We suppose that

$$\mu(R\mathbf{1}_{\Sigma_i} > t) \sim c_i\ell(t)t^{-\alpha} \quad \text{as } t \to \infty \tag{3.3}$$

for each i, where $c_i > 0$ and $\sum_{i=1}^e c_i = 1$. Setting $\Sigma_0 = \Sigma \setminus \bigcup_{i=1,\ldots,e} \Sigma_i$, we require that there exist $C > 0$ and $\eta > 0$ such that

$$\mu(R\mathbf{1}_{\Sigma_0} > t) \leq Ct^{-(\alpha+\eta)} \quad \text{for all } t > 0. \tag{3.4}$$

(Together, (3.3) and (3.4) imply assumption (3.1).)

Let $v : \Lambda \to \mathbb{R}^d$ be a Hölder observable with $\int_\Lambda v \, d\mu_\Lambda = 0$. We consider a set of *profiles* $P = \{P_i, i = 1, \ldots, e\} \subset C([0,1], \mathbb{R}^d)$ with $P_i(0) = 0$ for each i. Our final assumption is that there exist $C > 0$ and $\eta \in (0, 1)$ such that

$$\mathbf{1}_{\Sigma_i} \Big| \sum_{\ell=0}^k v \circ T^\ell - P_i(k/R)R \Big| \leq CR^{1-\eta} \quad \text{for all } 0 \leq k \leq R-1, i = 1, \ldots, e. \tag{3.5}$$

Remark 3.1. By (3.5) and continuity of P_i, each profile P_i is Lipschitz. It is equivalent to (3.5) that

$$\mathbf{1}_{\Sigma_i} \Big| \sum_{\ell=0}^{[Rt]-1} v \circ T^\ell - P_i(t)R \Big| \leq CR^{1-\eta} \quad \text{for all } t \in [0,1], i = 1, \ldots, e. \tag{3.6}$$

Remark 3.2. Unlike in [18], we do not require that the endpoints $P_i(1) \in \mathbb{R}^d$ are distinct and nonzero.

3.2 The decorated Lévy process L_α^P

Let $\omega_1 = (1, 0, \ldots, 0), \ldots, \omega_e = (0, \ldots, 0, 1)$ be the standard basis of $\mathbb{R}^e$. Define

$$Z : \Sigma \to \mathbb{R}^e, \qquad Z = \sum_{i=1}^{e} \omega_i R \mathbf{1}_{\Sigma_i}. \tag{3.7}$$

Then Z is regularly varying with spectral measure ν on $\mathbb{S}^{e-1}$ given by $\nu = \sum_{i=1}^{e} c_i \delta_{\omega_i}$. This means that

$$\lim_{t \to \infty} \frac{\mu(|Z| > rt, \ Z/|Z| \in E)}{\mu(|Z| > t)} = r^{-\alpha} \nu(E) \tag{3.8}$$

for all $r > 0$ and all Borel sets $E \subset \mathbb{S}^{e-1}$ with $\nu(\partial E) = 0$.

Let G_α be the corresponding e-dimensional α-stable law with characteristic function

$$\mathbb{E} \, e^{is \cdot G_\alpha} = \exp \left\{ - \int_{\mathbb{S}^{e-1}} |s \cdot x|^\alpha \left(1 - i \operatorname{sgn}(s \cdot x) \tan \frac{\pi\alpha}{2} \right) \cos \frac{\pi\alpha}{2} \, \Gamma(1 - \alpha) \, d\nu(x) \right\}$$

for $s \in \mathbb{R}^e$. Then Z is in the domain of attraction of G_α. That is, if $Z_1, Z_2, \ldots$ are i.i.d. copies of Z, then $b_n^{-1} \big(\sum_{j=1}^{n} Z_j - n \int_\Sigma Z \, d\mu \big) \to_w G_\alpha$.

Let $\bar{R} = \int_\Sigma R \, d\mu$ and let $L_\alpha \in D([0, 1], \mathbb{R}^e)$ be the e-dimensional α-stable Lévy process corresponding to the stable law $\bar{R}^{-1/\alpha} G_\alpha$.

Let $\Gamma = (P_1(1), \ldots, P_e(1)) \in \mathbb{R}^{d \times e}$, and choose $\Phi : \mathbb{S}^{e-1} \to \operatorname{Lip}([0, 1], \mathbb{R}^d)$ such that

- $(x, s) \mapsto \Phi(x)(s)$ is Lipschitz continuous;

- Φ is constant in a neighbourhood of each ω_i;

- $\Phi(\omega_i) = P_i$ for each i.

We define the decorated Lévy process $L_\alpha^P \in \mathscr{D}([0, 1], \mathbb{R}^d)$:

$$L_\alpha^P(t)(s) = \Gamma L_\alpha(t^-) + |L_\alpha(t) - L_\alpha(t^-)| \Phi \left(\frac{L_\alpha(t) - L_\alpha(t^-)}{|L_\alpha(t) - L_\alpha(t^-)|} \right)(s),$$

with the convention that $L_\alpha^P(t) \equiv \Gamma L_\alpha(t)$ at continuity points t of L_α. Although there is some flexibility in the choice of Φ, the decorated process L_α^P depends only on L_α and the set of profiles P.

Remark 3.3. (a) It is easily checked that $L_\alpha^P \in \mathscr{D}([0, 1], \mathbb{R}^d)$. We need to verify that $L_\alpha^P(t)(1) = \Gamma L_\alpha(t)$ for discontinuity points t of L_α. To this end, $L_\alpha(t) - L_\alpha(t^-) = \omega_i |L_\alpha(t) - L_\alpha(t^-)|$ for some i and $\Phi(\omega_i)(1) = P_i(1) = \Gamma \omega_i$. Hence

$$\begin{aligned} L_\alpha^P(t)(1) &= \Gamma L_\alpha(t^-) + |L_\alpha(t) - L_\alpha(t^-)| \Phi(\omega_i)(1) \\ &= \Gamma L_\alpha(t^-) + |L_\alpha(t) - L_\alpha(t^-)| \Gamma \omega_i \\ &= \Gamma L_\alpha(t^-) + \Gamma(L_\alpha(t) - L_\alpha(t^-)) = \Gamma L_\alpha(t). \end{aligned}$$

(b) Each δ-extension $(L_\alpha^P)^\delta$ is an element of $C([0, 1 + \delta], \mathbb{R}^d)$.

3.3 Main results

Let $v : \Lambda \to \mathbb{R}^d$ be a Hölder observable with $\int_\Lambda v \, d\mu_\Lambda = 0$. Define the càdlàg processes on (Λ, μ_Λ),

$$W_n \in D([0,1], \mathbb{R}^d), \qquad W_n(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} v \circ T^j.$$

Using the embedding $\iota : D([a,b], \mathbb{R}^d) \hookrightarrow \mathscr{D}([a,b], \mathbb{R}^d)$ from Definition 2.5, we obtain the trivial lift

$$\widehat{W}_n = \iota W_n \in \mathscr{D}([0,1], \mathbb{R}^d), \qquad \widehat{W}_n(t)(s) = W_n(t).$$

We prove:

Theorem 3.4. $\widehat{W}_n \to_{\mu_\Lambda} L_\alpha^P$ in the α_∞ topology as $n \to \infty$.

Theorem 3.5. The sequence $|W_n|_{p\text{-var}}$ is tight on (Λ, μ_Λ) for every $p > \alpha$.

The proofs are postponed to Sections 4, 5, 6 and 7. As a consequence, we obtain:

Theorem 3.6. Let $d, m \geq 1$, $\xi \in \mathbb{R}^m$. Assume that $A : \mathbb{R}^m \to \mathbb{R}^m$ and $B : \mathbb{R}^m \to \mathbb{R}^{m \times d}$ are sufficiently smooth as in Section 2.3. Let $X_n, X \in \mathscr{D}([0,1], \mathbb{R}^m)$ be solutions for

$$dX_n = A(X_n) \, dt + B(X_n) \, d\widehat{W}_n, \qquad dX = A(X) \, dt + B(X) \, dL_\alpha^P$$

with $X_n(0) = X(0) = \xi$. Then

$$X_n \to_{\mu_\Lambda} X \quad \text{in the } \alpha_{q\text{-var}} \text{ topology as } n \to \infty$$

for every $q > \alpha$.

Proof. By Theorem 3.5 and Remark 2.9, $|\widehat{W}_n|_{p\text{-var}}$ is tight for $p > \alpha$. Hence it follows from Theorems 3.4 and 2.10 that $X_n = \mathcal{S}(\widehat{W}_n) \to_{\mu_\Lambda} \mathcal{S}(L_\alpha^P) = X$. ∎

We say that the profile P_i is *linear* if $\{P_i(s) : 0 \leq s \leq 1\}$ is contained in a line. Under the extra assumption that excursions are linear, i.e. each profile P_i is linear, we can identify our solution with a Marcus SDE as follows.

Theorem 3.7. In the setting of Theorem 3.6, assume additionally that P_i is linear for each $i = 1, \dots, e$. Then the path $t \mapsto X(t)(1)$ solves the Marcus SDE $dY = A(Y) \, dt + B(Y) \diamond dL_\alpha$.

Proof. To show the first claim, consider the decorated process $L_\alpha^\ell \in \mathscr{D}([0,1], \mathbb{R}^d)$ defined by $L_\alpha^\ell(t)(s) = (1-s)L_\alpha(t^-) + sL_\alpha(t)$ and the solution $X^\ell \in \mathscr{D}([0,1], \mathbb{R}^m)$ to

$$dX^\ell = A(X^\ell) \, dt + B(X^\ell) \, dL_\alpha^\ell, \quad X(0) = \xi.$$

By [12, Proposition 4.16],[4] the path $t \mapsto X^\ell(t)(1)$ is almost surely equal to the solution Y of the stated Marcus SDE. On the other hand, for any $i = 1, \dots, e$ and $\lambda \in \mathbb{R}$, since P_i takes values on a line through $P_i(0) = 0$, one has $Z(1) = Z^\ell(1)$ for the solutions $Z, Z^\ell : [0,1] \to \mathbb{R}^m$ to the ODEs $dZ = B(Z)\lambda \, dP_i$ and $dZ^\ell = B(Z^\ell)\lambda P_i(1) \, ds$ with same initial value $Z(0) = Z^\ell(0)$. Since every excursion of L_α^P is of the form $s \mapsto \Gamma L_\alpha(t-) + \lambda P_i(s)$ for some $i = 1, \dots, e$ and $\lambda \in \mathbb{R}$, it readily follows from the construction of the solutions X, X^ℓ in Section 2.3 that $X(t)(1) = X^\ell(t)(1)$. ∎

---

[4][12, Proposition 4.16] assumes that the functions A, B are C^γ for $\gamma > 2$, but the same (even simpler) proof applies in our setting since we are in the Young regime $p \in [1, 2)$.

4 Preliminaries about Gibbs-Markov maps and Young towers

In this section, we elaborate on the structure of the dynamical systems (f, Σ, μ) and (F, Y, μ_Y) from Section 3.1 and we collect together some ergodic-theoretic properties of such dynamical systems for easy reference during Sections 5, 6 and 7.

Let $(\bar{Y}, \mu_{\bar{Y}})$ be a probability space with an at most countable measurable partition $\{\bar{Y}_j : j \geq 1\}$ and let $\bar{F} : \bar{Y} \to \bar{Y}$ be an ergodic measure-preserving map. Define the separation time $s(y, y')$ to be the least integer $n \geq 0$ such that $\bar{F}^n y$ and $\bar{F}^n y'$ lie in distinct partition elements. We assume that $s(y, y') = \infty$ if and only if $y = y'$; then $d_\theta(y, y') = \theta^{s(y,y')}$ is a metric for $\theta \in (0, 1)$.

We say that $\bar{F} : \bar{Y} \to \bar{Y}$ is a *(full-branch) Gibbs-Markov map* if (i) $\bar{F} : \bar{Y}_j \to \bar{Y}$ is a measure-theoretic bijection for each $j \geq 1$; (ii) there exists $\theta \in (0, 1)$ such that $\log \xi$ is d_θ-Lipschitz, where $\xi = d\mu_{\bar{Y}}/d\mu_{\bar{Y}} \circ \bar{F}$. For standard facts about Gibbs-Markov maps, we refer to [1, 2].

We say that a function g defined on $\bar{Y}$ is *piecewise constant* if g is constant on partition elements.

Proposition 4.1. *Let $g \in L^2(\bar{Y})$ be piecewise constant with $\int_{\bar{Y}} g \, d\mu_{\bar{Y}} = 0$. Then*

$$\left\| \max_{1 \leq k \leq n} \left| \sum_{j=0}^{k-1} g \circ \bar{F}^j \right| \right\|_{L^2(\bar{Y})} \leq C n^{1/2} |g|_{L^2(\bar{Y})}$$

where $C > 0$ is a constant independent of g and n.

Proof. We suppress the bars. Also, $|\cdot|_p$ denotes $|\cdot|_{L^p(Y)}$. Given $v : Y \to \mathbb{R}$ continuous, we define $\|v\|_\theta = \|v\|_\infty + \text{Lip}_\theta v$ where $\text{Lip}_\theta v = \sup_{y \neq y'} |v(y) - v(y')|/d_\theta(y, y')$. In particular, we can fix $\theta \in (0, 1)$ so that $\text{Lip}_\theta \log \xi < \infty$.

Define the transfer operator $P : L^1(Y) \to L^1(Y)$ (so $\int_Y Pv \, w \, d\mu_Y = \int_Y v \, w \circ F \, d\mu_Y$ for $v \in L^1(Y)$, $w \in L^\infty(Y)$). Then $(Pv)(y) = \sum_j \xi(y_j) v(y_j)$ where y_j is the unique preimage of y under $F|_{Y_j}$. There exists $C > 0$ such that $\xi(y_j) \leq C\mu(Y_j)$ and $|\xi(y_j) - \xi(y_j')| \leq C\mu(Y_j) d_\theta(y, y')$ for all $y, y' \in Y_j$, $j \geq 1$. It follows easily that $\|Pg\|_\theta \ll |g|_1 \leq |g|_2$.

There exist constants $\gamma \in (0, 1)$, $C > 0$ such that $\|P^n v\|_\theta \leq C\gamma^n \|v\|_\theta$ for all continuous $v : Y \to \mathbb{R}$ with $\int_Y v \, d\mu_Y = 0$ and all $n \geq 1$. Define $\chi = \sum_{n=1}^\infty P^n g$. Then $|\chi|_\infty \leq \sum_{n=1}^\infty \|P^{n-1} Pg\|_\theta \ll \|Pg\|_\theta \ll |g|_2$.

Now define the "martingale-coboundary decomposition"

$$g = m + \chi \circ F - \chi$$

where $|m|_2 \ll |g|_2$. Since $m \in \ker P$, it follows easily that $\left| \sum_{j<n} m \circ F^j \right|_2 = n^{1/2} |m|_2$. Moreover, $\{m \circ F^j, j \geq 0\}$ is a sequence of reverse martingale differences so, by Doob's inequality, $\left| \max_{k \leq n} \left| \sum_{j<k} m \circ F^j \right| \right|_2 \leq 2n^{1/2} |m|_2$. Hence

$$\left| \max_{k \leq n} \left| \sum_{j<k} g \circ F^j \right| \right|_2 \leq 2n^{1/2} |m|_2 + 2|\chi|_\infty \ll n^{1/2} |g|_2$$

as required. ∎

We require that the map (F, Y, μ_Y) from Section 3.1 quotients to a Gibbs-Markov map $(\bar{F}, \bar{Y}, \mu_{\bar{Y}})$. Specifically, there is a measure-preserving semiconjugacy $\bar{\pi} : Y \to \bar{Y}$ between F and $\bar{F}$. We say that a function g defined on Y is *piecewise constant* if g is constant on $Y_j = \bar{\pi}^{-1} \bar{Y}_j$ for each j.

Suppose as in Section 3.1 that $b_n^\alpha \sim n\ell(b_n)$ where $\alpha \in (1, 2)$ and ℓ is slowly varying.

Proposition 4.2. *Suppose that (F, Y, μ_Y) quotients to a Gibbs-Markov map. Let $g : Y \to [0, \infty)$ be piecewise constant and integrable with $\int_Y g \, d\mu_Y = 0$. Then the following are equivalent:*

(a) $b_n^{-1} \sum_{j=0}^{n-1} g \circ F^j \to_{\mu_Y} G$ *where G is a nondegenerate α-stable law.*

(b) $\mu_Y(g > t) \sim c\ell(t)t^{-\alpha}$ *as $t \to \infty$ for some $c > 0$.*

Proof. Since $\bar{\pi} : Y \to \bar{Y}$ is a measure-preserving semiconjugacy, it suffices to prove this result at the level of $(\bar{F}, \bar{Y}, \bar{\mu}_Y)$. Suppose that $\mu_{\bar{Y}}(g > t) \sim c\ell(t)t^{-\alpha}$ as $t \to \infty$ for some $c > 0$. By [1], $b_n^{-1} \sum_{j=0}^{n-1} g \circ \bar{F}^j \to_{\mu_{\bar{Y}}} G$ where G is a nondegenerate α-stable law. The converse holds by [22]. ∎

Exponential Young towers   Let (Σ, d) be a metric space and $f : \Sigma \to \Sigma$ a measurable map. Let $Y \subset \Sigma$ be a measurable subset, Let $\tau : Y \to \mathbb{Z}^+$ be a return time (so $f^{\tau(y)}y \in Y$ for $y \in Y$). Define $F = f^\tau : Y \to Y$ and let μ_Y be an F-invariant ergodic Borel probability measure on Y.

We suppose as above that (F, Y, μ_Y) quotients to a Gibbs-Markov map $(\bar{F}, \bar{Y}, \mu_{\bar{Y}})$. Moreover, we require that $\tau : Y \to \mathbb{Z}^+$ is piecewise constant. (In particular, τ is well-defined on $\bar{Y}$.) We suppose further that $\mu_Y(\tau > n) = O(e^{-cn})$ for some $c > 0$.

Define the *Young tower* $f_\Delta : \Delta \to \Delta$,

$$\Delta = \{(y, \ell) \in Y \times \mathbb{Z} : 0 \le \ell < \tau(y)\}, \qquad f_\Delta(y, \ell) = \begin{cases} (y, \ell+1) & \text{if } \ell \le \tau(y) - 2 \\ (Fy, 0) & \text{if } \ell = \tau(y) - 1 \end{cases}$$

with ergodic f_Δ-invariant probability measure $\mu_\Delta = (\mu_Y \times \text{counting})/\bar{\tau}$ where $\bar{\tau} = \int_Y \tau \, d\mu_Y$.

Define the semiconjugacy

$$\pi_\Delta : \Delta \to \Sigma, \qquad \pi_\Delta(y, \ell) = f^\ell(y),$$

between f_Δ and f. Then $\mu = (\pi_\Delta)_* \mu_\Delta$ is an f-invariant ergodic probability measure on Σ. Under certain further technical assumptions as in Young [46], we say that (f, Σ, μ) is *modelled by a Young tower with exponential tails*. (These technical assumptions ensure the validity of Theorem 4.4 and Lemma 4.5 below, and are not used further in this paper.)

For $n \ge 1$, define the *lap number* $N_n : \Delta \to \mathbb{Z}^+$ to be the integer satisfying

$$\sum_{j=0}^{N_n(x)-1} \tau(F^j y) \le n + \ell < \sum_{j=0}^{N_n(x)} \tau(F^j y) \tag{4.1}$$

for $x = (y, \ell) \in \Delta$.

Proposition 4.3. *There exists $C > 0$ such that $\big| \max_{1 \le k \le n} |N_k - k\bar{\tau}^{-1}| \big|_{L^1(\Delta)} \le Cn^{1/2}$ for all $n \ge 1$.*

Proof. Write $\tau_k = \sum_{j=0}^{k-1} \tau \circ F^j$. Then, for $x = (y, \ell) \in \Delta$,

$$\tau_{N_n(x)}(y) - \tau(y) \le \sum_{j=0}^{N_n(x)-1} \tau(F^j y) - \ell \le n \le \sum_{j=0}^{N_n(x)} \tau(F^j y) - \ell \le \tau_{N_n(x)}(y) + \tau(F^{N_n(x)}y).$$

Hence,
$$|n - N_n(x)\bar{\tau}| \le |\tau_{N_n(x)}(y) - N_n(x)\bar{\tau}| + \max\{\tau(y), \tau(F^{N_n(x)}y)\}.$$

14

Since $N_n \leq n$,

$$\max_{1 \leq k \leq n} |k\bar{\tau}^{-1} - N_k(x)| \leq \bar{\tau}^{-1} H_n(y)$$

where $H_n(y) = \max_{0 \leq j \leq n} |\tau_j(y) - j\bar{\tau}| + \max_{0 \leq j \leq n} \tau(F^j y)$.

Since $\tau \in L^2(Y)$ and H_n is independent of ℓ,

$$\big| \max_{1 \leq k \leq n} |k\bar{\tau}^{-1} - N_k|\big|_{L^1(\Delta)} \leq \bar{\tau}^{-1} \int_Y \tau\, \bar{\tau}^{-1} H_n\, d\mu_Y \ll |H_n|_{L^2(Y)}.$$

Applying Proposition 4.1 with $g = \tau - \bar{\tau}$,

$$\big| \max_{0 \leq j \leq n} |\tau_j - j\bar{\tau}|\big|_{L^2(Y)} = \big| \max_{0 \leq j \leq n} |\tau_j - j\bar{\tau}|\big|_{L^2(\bar{Y})} \ll n^{1/2}.$$

Also, $|\max_{0 \leq j \leq n} \tau \circ F^j|_{L^2(Y)} \leq n^{1/2} |\tau|_{L^2(Y)}$. Hence $|H_n|_{L^2(Y)} \ll n^{1/2}$ as required. ∎

Let $\mathcal{A}$ consist of unions of elements of $\mathcal{C}$. Let $\mathcal{B}$ consist of subsets $B \subset \Sigma$ that are unions of sets of the form $f^{-j}A$ where $A \in \mathcal{A}$, $j \geq 0$. Clearly, $\mathcal{A} \subset \mathcal{B}$.

Theorem 4.4. *Suppose that $f : \Sigma \to \Sigma$ is mixing and modelled by a Young tower with exponential tails. Then there exist constants $C > 0$, $\gamma \in (0,1)$ such that*

$$\big|\mu\big(A \cap f^{-n}B\big) - \mu(A)\mu(B)\big| \leq C\gamma^n$$

for all $A \in \mathcal{A}$, $B \in \mathcal{B}$, $n \geq 1$.

Proof. In general, the Young towers associated to f are mixing only up to a finite cycle. However, since f is mixing, we can reduce by [8, Section 10] (see also [6, Section 4.1]) to the case when $f_\Delta : \Delta \to \Delta$ is mixing.

The observables $\mathbf{1}_A$ and $\mathbf{1}_B$ on Σ lift to observables $v = \mathbf{1}_A \circ \pi_\Delta$, $w = \mathbf{1}_B \circ \pi_\Delta$ on the tower Δ. Since $A, B \in \mathcal{B}$, it follows from Young [46] that,

$$\big|\mu\big(A \cap f^{-n}B\big) - \mu(A)\mu(B)\big| \ll (|v|_\infty + \mathrm{Lip}_\theta\, v)\, |w|_\infty\, \gamma^n,$$

where $\mathrm{Lip}_\theta\, v = \sup_{(y,\ell) \neq (y',\ell)} |v(y,\ell) - v(y',\ell)|/d_\theta(y,y')$. The result follows since $|v|_\infty \leq 1$, $\mathrm{Lip}_\theta\, v \leq 1$ and $|w|_\infty \leq 1$. ∎

Lemma 4.5. *Suppose that $f : \Sigma \to \Sigma$ is mixing and modelled by a Young tower with exponential tails and fix $q > 1$. Let $H : \Sigma \to \mathbb{R}^d$ be such that $H \in L^q$ and $\int_\Sigma H\, d\mu = 0$. Suppose that there exists $C > 0$, $\eta \in (0,1)$ such that $|H(x) - H(x')| \leq CR(x)d(x,x')^\eta$ for all x, x' in the same element of $\mathcal{C}$.*

Define $H_Y = \sum_{\ell=0}^{\tau-1} H \circ f^\ell : Y \to \mathbb{R}^d$. Then

$$H^Y = m + \chi \circ F - \chi$$

where $m, \chi \in L^p$ for all $p < q$ and $\{m \circ F^{n-j} : 0 \leq j \leq n\}$ is a martingale difference sequence for each $n \geq 1$.

Proof. This is part of the proof of [33, Lemma 6.2]. ∎

Corollary 4.6. *Assume the setup of Lemma 4.5 with $q > \alpha$. Then*

$$b_n^{-1} \max_{k \leq n} \big| \sum_{j=0}^{k-1} H \circ f^j \big| \to_\mu 0.$$

15

Proof. Choose $p \in (\alpha, q)$. Following the proof of [33, Lemma 6.2], we can apply inequalities of Doob and Burkholder to deduce from Lemma 4.5 that

$$\left| \max_{k \le n} \left| \sum_{j=0}^{k-1} H^Y \circ F^j \right| \right|_p \le \left| \max_{k \le n} \left| \sum_{j=0}^{k-1} m \circ F^j \right| \right|_p + 2 \left| \max_{k \le n} |\chi \circ F^k| \right|_p \ll n^{1/p} = o(b_n).$$

In particular, $b_n^{-1} \max_{k \le n} \left| \sum_{j=0}^{k-1} H^Y \circ F^j \right| \to_{\mu_Y} 0$. The result follows by a standard inducing argument (e.g. a very simplified special case of [35, Theorem 2.2]). ∎

Remark 4.7. Let $(T, \Lambda, \mu_\Lambda)$ and R be as in Section 3.1. Suppose that $v : \Lambda \to \mathbb{R}^d$ is a Hölder function with $\int_\Lambda v \, d\mu_\Lambda = 0$ and let $H = \sum_{\ell=0}^{R-1} v \circ T^\ell$. Then $\int_\Sigma H \, d\mu = 0$ and H satisfies the estimate on $H(x) - H(x')$ in Lemma 4.5.

Lemma 4.8. *Let $V \in L^1(\Sigma)$ with $\int_\Sigma V \, d\mu = 0$. Define the induced observable $V^Y : Y \to \mathbb{R}$, $V^Y = \sum_{j=0}^{\tau-1} V \circ f^j$. Let G be a random variable. Then the following are equivalent:*

(a) $b_n^{-1} \sum_{j=0}^{n-1} V \circ f^j \to_\mu G$.

(b) $b_n^{-1} \sum_{j=0}^{n-1} V^Y \circ F^j \to_{\mu_Y} \bar{\tau}^{1/\alpha} G$.

Proof. Since Y quotients to a Gibbs-Markov map and τ is piecewise constant and square integrable, it follows for example from [1] that $n^{-1/2}(\sum_{j=0}^{n-1} \tau \circ F^j - n\bar\tau) \to_{\mu_Y} \mathcal{N}$ where $\mathcal{N}$ is a possibly degenerate Gaussian. In particular, $b_n^{-1}(\sum_{j=0}^{n-1} \tau \circ F^j - n\bar\tau) \to_{\mu_Y} 0$. Hence, the result follows from [33, Theorem A.1 and Remark A.3]. ∎

5   Convergence in $\mathcal{J}_1$ for the first return dynamics

Let (f, Σ, μ) and $R : \Sigma \to \mathbb{Z}^+$ be as in Section 3. In particular, the map f is mixing and modelled by a Young tower with exponential tails. We assume conditions (3.2) and (3.3) and make use of Propositions 4.1 and 4.3 and Theorem 4.4. The underlying system $(T, \Lambda, \mu_\Lambda)$ plays no role in this section.

In Section 3.2, we defined

$$Z : \Sigma \to \mathbb{R}^e, \qquad Z = \sum_{i=1}^e \omega_i R \mathbf{1}_{\Sigma_i} : \Sigma \to \mathbb{R}^e.$$

Recall also $b_n^\alpha \sim n\ell(b_n)$ as $n \to \infty$ from Section 3.1. Let $\widetilde{Z} = Z - \int_\Sigma Z \, d\mu$ and define the sequence of processes

$$W_n^Z \in D([0,1], \mathbb{R}^e), \qquad W_n^Z(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} \widetilde{Z} \circ f^j,$$

on the probability space (Σ, μ). Let $\tilde{L}_\alpha$ be the e-dimensional α-stable Lévy process corresponding to the stable law G_α described in Section 3.2 (so $\tilde{L}_\alpha = \bar{R}^{1/\alpha} L_\alpha$).

In this section, we prove

Theorem 5.1. *Assume that $f : \Sigma \to \Sigma$ is mixing and modelled by a Young tower with exponential tails. Suppose that conditions (3.2) and (3.3) hold. Then $W_n^Z \to_\mu \tilde{L}_\alpha$ in the $\mathcal{J}_1$ topology.*

Our proof of Theorem 5.1 largely follows the approach in [24, Section 4] which was written specifically for billiards with flat cusps in the case $d = 1$.

On the probability space (Σ, μ), we define the sequence of random point processes $\mathcal{N}_n = \sum_{j=1}^{n} \delta_{(\frac{j}{n}, b_n^{-1} \widetilde{Z} \circ f^{j-1})}$ on $(0, \infty) \times (\mathbb{R}^e \setminus \{0\})$. The Lévy measure Π corresponding to the Lévy process $\widetilde{L}_\alpha$ is given by

$$\Pi(B) = \alpha \int_{\mathbb{S}^{e-1}} \int_0^\infty \mathbf{1}_B(rx) r^{-\alpha-1} \, dr \, d\nu(x)$$

where $\nu = \sum_{i=1}^e c_i \delta_{\omega_i}$ is the spectral measure from Section 3.2. Let $\mathcal{N}$ be the Poisson point process on $(0, \infty) \times (\mathbb{R}^e \setminus \{0\})$ with mean measure $\mathrm{Leb} \times \Pi$.

By [42, Theorem 4.1] (see also [43, Theorem 1.2] and [24, Proposition 4.4]), to prove Theorem 5.1 it is enough to verify two conditions:

Condition I (Point process convergence). $\mathcal{N}_n \to_\mu \mathcal{N}$ as $n \to \infty$ in the space of point measures defined on $(0, \infty) \times (\mathbb{R}^e \setminus \{0\})$.

Condition II (Vanishing small values). For every $\gamma > 0$,

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \mu \left(\max_{1 \leq k \leq n} \Big| \sum_{j=0}^{k-1} \big(\widetilde{Z} \mathbf{1}_{\{|\widetilde{Z}| < b_n \varepsilon\}} \big) \circ f^j - k \int_\Sigma \widetilde{Z} \mathbf{1}_{\{|\widetilde{Z}| < b_n \varepsilon\}} \, d\mu \Big| > b_n \gamma \right) = 0.$$

These conditions are verified in the next two subsections. We denote by $B_a(c) \subset \mathbb{R}^e$ the open ball of radius a centred at c.

5.1 Point process convergence

In this subsection, we verify Condition I. We follow [24, Section 4.5] using [38, Theorem 2.1]. It is enough to prove convergence of $\mathcal{N}_n$ to $\mathcal{N}$ on

$$(0, \infty) \times U, \qquad U = \mathbb{R}^e \setminus \overline{B_{a_0}(0)}$$

for each fixed $a_0 > 0$.

Fix a_0 and let

$$A_n = \{|\widetilde{Z}| > a_0 b_n\}.$$

Let $\mathcal{W}$ be the ring of subsets of U generated by sets of the type $\{x \in \mathbb{R}^e : a < |x| < a', \ x/|x| \in E\}$, where $a_0 < a < a'$ and $E \subset \mathbb{S}^{e-1}$ is open with $\nu(\partial E) = 0$. Note that $\mathcal{W}$ generates the Borel sigma-algebra on U and that $\Pi(\partial W) = 0$ for all $W \in \mathcal{W}$.

For a collection $\mathcal{F}$ of measurable subsets of Σ, define

$$\mathcal{Q}_p(\mathcal{F}) = \sup_{\substack{A \in \mathcal{F} \\ B \in \sigma(\bigcup_{j \geq p} f^{-j} \mathcal{F})}} \big| \mu(A \cap B) - \mu(A)\mu(B) \big|.$$

Let $J_n = b_n^{-1} \widetilde{Z}$. By [38, Theorem 2.1], to prove Condition I it suffices to prove the following:

Lemma 5.2. (a) $\lim_{n \to \infty} \mu \big(J_n^{-1} W | A_n \big) = \Pi(W|U)$ for all $W \in \mathcal{W}$.

(b) $\mathcal{Q}_1(J_n^{-1} \mathcal{W}_0) = o(\mu(A_n))$ for every finite subset $\mathcal{W}_0$ of $\mathcal{W}$.

Proof of Lemma 5.2(a). Without loss, we may suppose that $W = \{x \in \mathbb{R}^e : a < |x| < a', x/|x| \in E\}$ where $a_0 < a < a'$ and $E \subset \mathbb{S}^{e-1}$ is open with $\nu(\partial E) = 0$. Since $Z - \widetilde{Z}$ is constant, it follows from (3.8) that

$$
\lim_{n \to \infty} \mu\big(J_n^{-1}W | A_n\big) = \lim_{n \to \infty} \frac{\mu\big(|\widetilde{Z}| \in (ab_n, a'b_n), \, \widetilde{Z}/|\widetilde{Z}| \in E\big)}{\mu(|\widetilde{Z}| > a_0 b_n)}
$$
$$
= \lim_{n \to \infty} \frac{\mu\big(|Z| \in (ab_n, a'b_n), \, Z/|Z| \in E\big)}{\mu(|Z| > a_0 b_n)} = a_0^\alpha (a^{-\alpha} - (a')^{-\alpha})\nu(E).
$$

On the other hand, by the definition of Π,

$$
\Pi(W|U) = \frac{\Pi(W)}{\Pi(U)} = \frac{(a^{-\alpha} - (a')^{-\alpha})\nu(E)}{a_0^{-\alpha}},
$$

so the result is proved. ∎

Next, we prove Lemma 5.2(b).

Proposition 5.3. $\mu(A_n) \sim a_0^{-\alpha} n^{-1}$ as $n \to \infty$.

Proof. Since $Z - \widetilde{Z}$ is constant, $\mu(A_n) \sim \mu(A_n')$ where $A_n' = \{|Z| > a_0 b_n\}$.
By (3.3),
$$
\mu(|Z| > t) = \mu(R\mathbf{1}_{\bigcup_{i=1,\dots,e} \Sigma_i} > t) \sim \ell(t)t^{-\alpha}.
$$

Using that ℓ is slowly varying and that $b_n^\alpha \sim n\ell(b_n)$,

$$
\mu(A_n) \sim \mu(A_n') \sim \ell(a_0 b_n)(a_0 b_n)^{-\alpha} \sim a_0^{-\alpha}\ell(a_0 b_n)\ell(b_n)^{-1}n^{-1} \sim a_0^{-\alpha}n^{-1},
$$

as required. ∎

Let $\theta_1 > 1$ be as in (3.2) for $s = a_0/2$. Without loss, we can suppose that $\theta_1 \in (1,2)$. Let $\gamma \in (0,1)$ be as in Theorem 4.4.

Proposition 5.4. There exists $C > 0$ such that $\mu(A_n \cap f^{-k}A_n) \leq Cn^{-\theta_1}$ for all $k, n \geq 1$.

Proof. By Theorem 4.4, $\big|\mu(A_n \cap f^{-k}A_n) - \mu(A_n)^2\big| \ll \gamma^k$ for all $k \geq 1$. By Proposition 5.3, $\mu(A_n)^2 \ll n^{-2} \leq n^{-\theta_1}$. Hence $\mu(A_n \cap f^{-k}A_n) \ll n^{-\theta_1}$ uniformly in $k \geq n$.
Let $D = \{R > a_0 b_n/2\}$. Then

$$
A_n \subset \{2|Z| > a_0 b_n\} \subset \{2R > a_0 b_n\} = D.
$$

Hence $A_n \cap f^{-k}A_n \subset D \cap f^{-k}D$ and, by (3.2), $\mu(A_n \cap f^{-k}A_n) \ll n^{-\theta_1}$ uniformly in $1 \leq k \leq n$.
Combining the estimates for $k \leq n$ and $k \geq n$ yields the desired result. ∎

Now fix $\mathcal{W}_0 \subset \mathcal{W}$ finite, and let $\mathcal{Q}_{n,p} = \mathcal{Q}_p(J_n^{-1}\mathcal{W}_0)$.

Proposition 5.5. $\mathcal{Q}_{n,p} \leq C\gamma^p$ for all $n, p \geq 1$.

Proof. Let $A \in J_n^{-1}\mathcal{W}_0$, $B \in \sigma\big(\bigcup_{j \geq p} f^{-j}J_n^{-1}\mathcal{W}_0\big)$. Write $B = f^{-p}B'$ where $B' \in \sigma\big(\bigcup_{j \geq 0} f^{-j}J_n^{-1}\mathcal{W}_0\big)$. Then $|\mu(A \cap B) - \mu(A)\mu(B)| \ll \gamma^p$ by Theorem 4.4. ∎

Let $\tau_{A_n}(x) = \min\{n \geq 1 : f^n x \in A_n\}$ for $x \in \Sigma$.

Lemma 5.6. $\mathcal{Q}_{n,1} \leq \mathcal{Q}_{n,p+1} + \mu(A_n \cap \{\tau_{A_n} \leq p\}) + \mu(A_n)\mu(\tau_{A_n} \leq p)$ for all $n, p \geq 1$.

Proof. This is identical to [24, Lemma 4.9]. We give the short argument for completeness.

Let $A \in J_n^{-1}\mathcal{W}_0$ and $B \in \sigma\big(\bigcup_{j \geq 1} f^{-j} J_n^{-1}\mathcal{W}_0\big)$. Note that $A \subset A_n$ by definition of $\mathcal{W}$.

Suppose that $\mathcal{W}_0 = \{W_1, \ldots, W_K\}$. Observe that there exists a function $g : (\{0,1\}^K)^{\mathbb{Z}^+} \to \{0,1\}$ such that $\mathbf{1}_B = g(Y_1, Y_2, \ldots)$, where $Y_i = \big(\mathbf{1}_{J_n^{-1}W_1}, \ldots, \mathbf{1}_{J_n^{-1}W_K}\big) \circ f^i$. Define $B' \in \sigma\big(\bigcup_{j \geq p+1} f^{-j}(J_n^{-1}\mathcal{W}_0)\big)$ by

$$\mathbf{1}_{B'} = g(0, \ldots, 0, Y_{p+1}, Y_{p+2}, \ldots).$$

Then $|\operatorname{Cov}(\mathbf{1}_A, \mathbf{1}_{B'})| \leq \mathcal{Q}_{n,p+1}$. Moreover, $|\mathbf{1}_B - \mathbf{1}_{B'}| \leq \mathbf{1}_{\{\tau_{A_n} \leq p\}}$, so

$$\big|\operatorname{Cov}(\mathbf{1}_A, \mathbf{1}_B) - \operatorname{Cov}(\mathbf{1}_A, \mathbf{1}_{B'})\big| \leq \mu\big(A \cap \{\tau_{A_n} \leq p\}\big) + \mu(A)\mu(\tau_{A_n} \leq p)$$
$$\leq \mu(A_n \cap \{\tau_{A_n} \leq p\}) + \mu(A_n)\mu(\tau_{A_n} \leq p).$$

The result follows. ∎

Proof of Lemma 5.2(b). Let $p_n = [n^{\theta_2}]$ where $0 < \theta_2 < \theta_1 - 1 < 1$. Applying Proposition 5.4,

$$\mu(A_n \cap \{\tau_{A_n} \leq p_n\}) = \mu\bigg(A_n \cap \bigcup_{k=1}^{p_n} f^{-k} A_n\bigg)$$
$$\leq \sum_{k=1}^{p_n} \mu(A_n \cap f^{-k} A_n) = O(p_n n^{-\theta_1}) = o(n^{-1}).$$

Using Proposition 5.3 and invariance of μ under f,

$$\mu(\tau_{A_n} \leq p_n) = \mu\bigg(\bigcup_{k=1}^{p_n} f^{-k} A_n\bigg) \leq \sum_{k=1}^{p_n} \mu(f^{-k} A_n) = p_n \mu(A_n) = O(p_n n^{-1}) = o(1).$$

Hence it follows from Proposition 5.3 and Lemma 5.6 that $\mathcal{Q}_{n,1} \leq \mathcal{Q}_{n,p_n+1} + o(\mu(A_n))$. By Propositions 5.3 and 5.5, $\mathcal{Q}_{n,p_n+1} = o(n^{-1}) = o(\mu(A_n))$, so $\mathcal{Q}_1(J_n^{-1}\mathcal{W}_0) = \mathcal{Q}_{n,1} = o(\mu(A_n))$. ∎

This concludes the verification of Condition I.

5.2 Vanishing small values

In this subsection, we verify Condition II. The regularly varying sequence b_n is asymptotically equivalent to a strictly increasing sequence, so for the purposes of proving Theorem 5.1 we may suppose without loss of generality that b_n strictly increases and $b_1 = 1$.

It is convenient to work at the level of the Young tower $f_\Delta : \Delta \to \Delta$. Recall from Section 4 that $\pi_\Delta : \Delta \to \Sigma$, $\pi(y, \ell) = f^\ell y$ is a measure-preserving semiconjugacy from $(f_\Delta, \Delta, \mu_\Delta)$ to (f, Σ, μ). Let $\Delta_i = \pi_\Delta^{-1} \Sigma_i$, $i = 1, \ldots, e$. Abusing notation, we denote lifted observables $R \circ \pi_\Delta : \Delta \to \mathbb{Z}^+$, $\widetilde{Z} \circ \pi_\Delta : \Delta \to \mathbb{R}^e$ and so on simply by R, $\widetilde{Z}$ and so on.

Fix $i = 1, \ldots, e$ and set $R^{(i)} = R\mathbf{1}_{\Delta_i}$. Let $\bar{Z} = \int_\Delta Z \, d\mu_\Delta$ and

$$\widetilde{R}_\xi^{(i)} = R^{(i)} \mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}} - \int_\Delta R^{(i)} \mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}} \, d\mu_\Delta, \quad \xi > 0.$$

The main step in the proof is:

Lemma 5.7. *There is a constant $C > 0$ such that*

$$\limsup_{n \to \infty} b_n^{-1} \bigg\| \max_{1 \leq k \leq n} \bigg| \sum_{j=0}^{k-1} \widetilde{R}_{\varepsilon b_n}^{(i)} \circ f_\Delta^j \bigg| \bigg\|_{L^1(\Delta)} \leq C \varepsilon^{1-\alpha/2} \quad \text{for all } \varepsilon > 0.$$

Most of the remainder of this subsection is concerned with proving Lemma 5.7. At the end of the subsection, we show that Condition II follows from the lemma.

We begin with the following consequence of condition (3.2). Let $\theta_1 > 1$ be as in (3.2) with $s = 1$.

Proposition 5.8. *There exists $a_1 > 0$ such that*

$$\mu_\Delta(R > \eta \text{ and } R \circ f^j > \eta) \ll \left(\frac{\eta^\alpha}{\ell(\eta)}\right)^{-\theta_1} \quad \text{for all } 1 \le j \le a_1 \eta^\alpha/\ell(\eta), \ \eta > 1.$$

Proof. Since b_n is strictly increasing, we can write $b_n = \tilde{b}(n)$ where $\tilde{b} : (0, \infty) \to (0, \infty)$ is strictly increasing and $\tilde{b}(x)^\alpha \sim x\ell(\tilde{b}(x))$ as $x \to \infty$. Note that $\tilde{b}^{-1}(x) \sim x^\alpha/\ell(x)$ as $x \to \infty$. Choose $a_1 > 0$ such that $a_1 x^\alpha/\ell(x) \le \frac{1}{2}\tilde{b}^{-1}(x)$ for all $x \ge 1$.

Given $\eta > 1 = b_1$, there exists $n \in \mathbb{Z}^+$ such that $b_n < \eta \le b_{n+1}$. By (3.2),

$$\mu_\Delta(R > \eta \text{ and } R \circ f^j > \eta) \le \mu_\Delta(R > b_n \text{ and } R \circ f^j > b_n)$$

$$\ll n^{-\theta_1} \ll (n+1)^{-\theta_1} \le (\tilde{b}^{-1}(\eta))^{-\theta_1} \ll \left(\frac{\eta^\alpha}{\ell(\eta)}\right)^{-\theta_1}$$

for all $1 \le j \le n$. Finally, if $j \le a_1\eta^\alpha/\ell(\eta)$, then $j \le \frac{1}{2}\tilde{b}^{-1}(\eta) \le \frac{1}{2}(n+1) \le n$ so the estimate holds for all such j. $\blacksquare$

From now on, we fix θ_1 and a_1 as in Proposition 5.8.

Let $1 \le \eta \le \xi$. Define $H = H_\xi : Y \to [0, \infty)$, $s = s_\eta : Y \to \mathbb{Z}$,

$$H(y) = \sum_{0 \le j \le \tau(y)-1} (R^{(i)}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}})(y, j), \quad s(y) = \#\{0 \le j \le \tau(y) - 1 : R(y, j) > \eta\}.$$

Note that $H = H' + H'' + H'''$, where

$$H'(y) = \mathbf{1}_{\{s(y)=1\}} \sum_{j < \tau(y)} (R^{(i)}\mathbf{1}_{\{R > \eta\}}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}})(y, j),$$

$$H''(y) = \sum_{j < \tau(y)} (R^{(i)}\mathbf{1}_{\{R \le \eta\}}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}})(y, j),$$

$$H'''(y) = \mathbf{1}_{\{s(y) \ge 2\}} \sum_{j < \tau(y)} (R^{(i)}\mathbf{1}_{\{R > \eta\}}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}})(y, j).$$

Proposition 5.9. *Let $\delta > 0$. There exist $M, C > 0$ such that*

(a) $|H'|_{L^2(Y)} \le C\ell(\xi)^{1/2}\xi^{1-\alpha/2}$,

(b) $|H''|_{L^2(Y)} \le C\eta^{1-\alpha/2+\delta}$,

(c) $|H'''|_{L^2(Y)} \le C\xi^{1+\delta}\left(\frac{\eta^\alpha}{\ell(\eta)}\right)^{-\theta_1/2}$,

for all $1 < \eta \le \xi$ satisfying $M \log \xi \le a_1\eta^\alpha/\ell(\eta)$ and $\xi > |\bar{Z}|$.

Proof. We use throughout that $R^{(i)} \le R$ for all i.
(a) Observe that $H' \le \xi'$ where $\xi' = \xi + |\bar{Z}|$. Also for $0 \le n \le \xi'$,

$$\mu_Y(H' = n) \le \mu_Y(y \in Y : R(y, j) = n \text{ for some } 0 \le j < \tau(y)\})$$

$$\le \int_Y \sum_{j=0}^{\tau(y)-1} \mathbf{1}_{\{y \in Y : R(y,j)=n\}} \, d\mu_Y(y) = \bar{\tau}\mu_\Delta(R = n).$$

Hence
$$|H'|^2_{L^2(Y)} = \sum_{n \le \xi'} n^2 \mu_Y(H' = n) \le \bar{\tau} \sum_{n \le \xi'} n^2 \mu_\Delta(R = n) \ll \sum_{n \le \xi'} n \mu_\Delta(R \ge n).$$

Using (3.3) and applying Karamata's inequality, we conclude that $|H'|^2_{L^2(Y)} \ll \ell(\xi') \xi'^{2-\alpha} \ll \ell(\xi) \xi^{2-\alpha}$ as required.

(b) Let $q > 2$, $\varepsilon > 0$ and observe that $|R\mathbf{1}_{\{R \le \eta\}}|^q_{L^q(\Delta)} \ll \sum_{j \le \eta} j^{q-1} \mu_\Delta(R \ge j) \ll \eta^{q+\varepsilon-\alpha}$. Let $g_j(y) = (R\mathbf{1}_{\{R \le \eta\}}) \circ f^j_\Delta(y, 0)$. Since f_Δ is measure-preserving,

$$\left| \sum_{j<k} g_j \right|^q_{L^q(Y)} = \int_Y \left| \sum_{j<k} g_j \right|^q d\mu_Y \le \int_Y \sum_{\ell=0}^{\tau(y)-1} \left| \sum_{j<k} (R\mathbf{1}_{\{R \le \eta\}}) \circ f^j_\Delta(y, \ell) \right|^q d\mu_Y(y)$$

$$= \bar{\tau} \int_\Delta \left| \sum_{j<k} (R\mathbf{1}_{\{R \le \eta\}}) \circ f^j_\Delta \right|^q d\mu_\Delta \le \bar{\tau} k^q |R\mathbf{1}_{\{R \le \eta\}}|^q_{L^q(\Delta)} \ll k^q \eta^{q+\varepsilon-\alpha}.$$

Note also that $H'' \le \sum_{j<\tau} g_j$. Since τ has exponential tails, there exists $c_0 > 0$ such that

$$|H''|^2_{L^2(Y)} \le \sum_{k=1}^\infty \left| \mathbf{1}_{\{\tau=k\}} \sum_{j<k} g_j \right|^2_{L^2(Y)} \ll \sum_{k=1}^\infty e^{-c_0 k} \left| \sum_{j<k} g_j \right|^2_{L^q(Y)}$$

$$\ll \sum_{k=1}^\infty e^{-c_0 k} (k \eta^{(q+\varepsilon-\alpha)/q})^2 \ll \eta^{2(q+\varepsilon-\alpha)/q}.$$

The desired estimate follows for q and ε sufficiently close to 2 and 0.

(c) Since τ has exponential tails, there exists $c_1 > 0$ such that $k^2 \mu_Y(\tau = k) \ll e^{-c_1 k}$. Set $M = 2\alpha\theta_1/c_1$.

Suppose that $1 < \eta \le \xi$ satisfies $M \log \xi \le a_1 \eta^\alpha/\ell(\eta)$. For any $y \in \{\tau = k, s \ge 2\}$, there exist $0 \le j_1 < j_2 < k$ such that $R(y, j_1) > \eta$ and $R(y, j_2) > \eta$. By Proposition 5.8,

$$\mu_Y(\tau = k, s \ge 2) \le \bar{\tau} \sum_{0 \le j_1 < j_2 < k} \mu_\Delta(R \circ f^{j_1}_\Delta > \eta \text{ and } R \circ f^{j_2}_\Delta > \eta)$$

$$\le k\bar{\tau} \sum_{1 \le j < k} \mu_\Delta(R > \eta \text{ and } R \circ f^j_\Delta > \eta) \ll k^2 \left(\frac{\eta^\alpha}{\ell(\eta)} \right)^{-\theta_1}$$

for all $k \le a_1 \eta^\alpha/\ell(\eta)$ and hence for $k \le M \log \xi$.

Noting that $H'''(y) \ll \mathbf{1}_{\{s(y) \ge 2\}} \xi \tau(y)$,

$$|H'''|^2_{L^2(Y)} \ll \xi^2 \int_Y \mathbf{1}_{\{s \ge 2\}} \tau^2 d\mu_Y$$

$$\ll \xi^2 \sum_{k \le M \log \xi} k^2 \mu_Y(\tau = k, s \ge 2) + \xi^2 \sum_{k > M \log \xi} k^2 \mu_Y(\tau = k).$$

Now,

$$\sum_{k \le M \log \xi} k^2 \mu_Y(\tau = k, s \ge 2) \ll \sum_{k \le M \log \xi} k^4 \left(\frac{\eta^\alpha}{\ell(\eta)} \right)^{-\theta_1} \ll (\log \xi)^5 \left(\frac{\eta^\alpha}{\ell(\eta)} \right)^{-\theta_1},$$

$$\sum_{k > M \log \xi} k^2 \mu_Y(\tau = k) \ll \sum_{k > M \log \xi} e^{-c_1 k} \ll \xi^{-c_1 M} \le \eta^{-c_1 M} = \eta^{-2\alpha\theta_1}.$$

Hence, $|H'''|^2_{L^2(Y)} \ll \xi^2 (\log \xi + 1)^5 \left(\frac{\eta^\alpha}{\ell(\eta)} \right)^{-\theta_1}$ and the result follows. ∎

Corollary 5.10. *There is a constant $C > 0$ such that*

$$\limsup_{n \to \infty} b_n^{-1} n^{1/2} |H_{\varepsilon b_n}|_{L^2(Y)} \leq C\varepsilon^{1-\alpha/2} \quad \text{for all } \varepsilon > 0.$$

Proof. Recalling that $\theta_1 > 1$, we choose $\gamma \in (0, 1/\alpha)$ sufficiently close to $1/\alpha$ that $\gamma\alpha\theta_1 > 1$. We apply Proposition 5.9 with $\xi = \varepsilon b_n$, $\eta = n^\gamma$. (Note that for all $\varepsilon > 0$, $|\bar{Z}| \in \mathbb{R}^e$, the constraints $1 < \eta \leq \xi$, $M \log \xi \leq a_1 \eta^\alpha / \ell(\alpha)$, $\xi > |\bar{Z}|$ are satisfied for n sufficiently large.) Then

$$|H'|_{L^2(Y)} \ll \varepsilon^{1-\alpha/2} b_n \Big(\frac{\ell(\varepsilon b_n)}{\ell(b_n)} \Big)^{1/2} \big(\ell(b_n) b_n^{-\alpha} \big)^{1/2} \sim \varepsilon^{1-\alpha/2} b_n n^{-1/2}.$$

Also, $|H''|_{L^2(Y)} \ll n^{\gamma(1-\alpha/2+\delta)} = o(b_n n^{-1/2})$ for $\delta > 0$ sufficiently small. Finally, $|H'''|_{L^2(Y)} \ll b_n^{1+\delta} n^{-\gamma\alpha\theta_1/2} \tilde{\ell}(n)$ where $\tilde{\ell}$ is slowly varying. Hence, $|H'''|_{L^2(Y)} \ll b_n^{1+\delta} n^{-\gamma\alpha\theta_1/2} n^\delta$. Since $\gamma\alpha\theta_1 > 1$, we can shrink δ if necessary so that $b_n^\delta n^\delta n^{-\gamma\alpha\theta_1/2} = o(n^{-1/2})$. Hence, $|H'''|_{L^2(Y)} = o(b_n n^{-1/2})$. ∎

Define $\psi_\xi : \Delta \to [0, \infty)$ by $\psi_\xi(y, \ell) = \sum_{0 \leq j < \ell} (R^{(i)} \mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}})(y, j)$.

Corollary 5.11. *There exists $C > 0$ such that*

$$\limsup_{n \to \infty} b_n^{-1} \big| \max_{k \leq n} \psi_{\varepsilon b_n} \circ f_\Delta^k \big|_{L^1(\Delta)} \leq C\varepsilon^{1-\alpha/2} \quad \text{for all } \varepsilon > 0.$$

Proof. Define $h_\xi : \Delta \to [0, \infty)$ by $h_\xi(y, \ell) = H_\xi(y)$. Then $\max_{k \leq n} h_\xi \circ f_\Delta^k(y, \ell) \leq \max_{k \leq n} H_\xi(F^k y)$. Since τ has exponential tails,

$$\big| \max_{k \leq n} h_\xi \circ f_\Delta^k \big|_{L^1(\Delta)} \leq \bar{\tau}^{-1} \int_Y \tau \max_{k \leq n} H_\xi \circ F^k \, d\mu_Y$$
$$\ll \big| \max_{k \leq n} H_\xi \circ F^k \big|_{L^2(Y)} \ll n^{1/2} |H_\xi|_{L^2(Y)}.$$

Hence it follows from Corollary 5.10 that $\limsup_{n \to \infty} b_n^{-1} |\max_{k \leq n} h_{\varepsilon b_n} \circ f_\Delta^k|_{L^1(\Delta)} \ll \varepsilon^{1-\alpha/2}$. Since $0 \leq \psi_{\varepsilon b_n} \leq h_{\varepsilon b_n}$, we obtain the desired estimate for $\max_{k \leq n} \psi_{\varepsilon b_n} \circ f_\Delta^k$. ∎

Define $Q_{n,\xi} : \Delta \to \mathbb{R}$ by $Q_{n,\xi} = N_n \int_Y H_\xi \, d\mu_Y - n \int_\Delta R^{(i)} \mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}} \, d\mu_\Delta$ where N_n is the lap number defined in (4.1).

Proposition 5.12. *There exists $C > 0$ such that*

$$\limsup_{n \to \infty} b_n^{-1} \big| \max_{k \leq n} |Q_{k,\varepsilon b_n}| \big|_{L^1(\Delta)} \leq C\varepsilon^{1-\alpha/2} \quad \text{for all } \varepsilon > 0.$$

Proof. Observe that $Q_{n,\xi} = (N_n - n\bar{\tau}^{-1}) \int_Y H_\xi \, d\mu_Y$. Hence

$$|Q_{n,\xi}| \leq |N_n - n\bar{\tau}^{-1}| \int_Y H_\xi \, d\mu_Y \leq |N_n - n\bar{\tau}^{-1}| \, |H_\xi|_{L^2(Y)}.$$

The result follows from Proposition 4.3 and Corollary 5.10. ∎

Proof of Lemma 5.7. Let $\widetilde{H}_\xi = H_\xi - \int_Y H_\xi \, d\mu_Y$. Define $G_{n,\xi} : Y \to \mathbb{R}$, $g_{n,\xi} : \Delta \to \mathbb{R}$,

$$G_{n,\xi} = \max_{k \leq n} \sum_{j < k} \widetilde{H}_\xi \circ F^j, \quad g_{n,\xi}(y, \ell) = G_{n,\xi}(y).$$

By Proposition 4.1 and Corollary 5.10,

$$\limsup_{n\to\infty} b_n^{-1}|G_{n,\varepsilon b_n}|_{L^2(Y)} \ll \limsup_{n\to\infty} b_n^{-1}n^{1/2}|H_{\varepsilon b_n}|_{L^2(Y)} \ll \varepsilon^{1-\alpha/2}.$$

Since τ has exponential tails,

$$\int_\Delta |g_{n,\varepsilon b_n}|\,d\mu_\Delta = \bar{\tau}^{-1}\int_Y \tau|G_{n,\varepsilon b_n}|\,d\mu_Y \ll |G_{n,\varepsilon b_n}|_{L^2(Y)},$$

so

$$\limsup_{n\to\infty} b_n^{-1}\int_\Delta |g_{n,\varepsilon b_n}|\,d\mu_\Delta \ll \varepsilon^{1-\alpha/2}. \tag{5.1}$$

Next, by the definition of the lap number N_n,

$$\sum_{j=0}^{n-1}(R^{(i)}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}})(f_\Delta^j(y,\ell)) = \sum_{j=0}^{N_n(y,\ell)-1} H_\xi(F^j y) + \psi_\xi(f_\Delta^n(y,\ell)) - \psi_\xi(y,\ell),$$

and so

$$\sum_{j=0}^{n-1}\widetilde{R}_\xi^{(i)}(f_\Delta^j(y,\ell)) = \sum_{j=0}^{N_n(y,\ell)-1} \widetilde{H}_\xi(F^j y) + \psi_\xi(f_\Delta^n(y,\ell)) - \psi_\xi(y,\ell) + Q_{n,\xi}(y,\ell).$$

Since $N_n \le n$ for all n,

$$\max_{k\le n}\Big|\sum_{j=0}^{k-1}\widetilde{R}_\xi^{(i)}\circ f_\Delta^j(y,\ell)\Big| \ll \max_{k\le n}\Big|\sum_{j<k}\widetilde{H}_\xi(F^j y)\Big| + \max_{k\le n}\psi_\xi(f_\Delta^k(y,\ell)) + \max_{k\le n}|Q_{k,\xi}(y,\ell)|.$$

In other words,

$$\max_{k\le n}\Big|\sum_{j=0}^{k-1}\widetilde{R}_\xi^{(i)}\circ f_\Delta^j\Big| \ll g_{n,\xi} + \max_{k\le n}\psi_\xi\circ f_\Delta^k + \max_{k\le n}|Q_{k,\xi}|.$$

Setting $\xi = \varepsilon b_n$, the result follows from (5.1), Corollary 5.11 and Proposition 5.12. ∎

We can now complete the verification of Condition II. First notice that

$$R^{(i)}\mathbf{1}_{\{|\widetilde{Z}|<\xi\}} = R^{(i)}\mathbf{1}_{\{|\omega_i R^{(i)} - \int_\Delta Z\,d\mu|<\xi\}} = R^{(i)}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}}.$$

Hence

$$Z\mathbf{1}_{\{|\widetilde{Z}|<\xi\}} = \sum_i \omega_i R^{(i)}\mathbf{1}_{\{\omega_i R \in B_\xi(\bar{Z})\}}$$

and so

$$Z\mathbf{1}_{\{|\widetilde{Z}|<\xi\}} - \int_\Delta Z\mathbf{1}_{\{|\widetilde{Z}|<\xi\}}\,d\mu_\Delta = \sum_i \omega_i \widetilde{R}_\xi^{(i)}.$$

By Lemma 5.7,

$$\limsup_{n\to\infty} b_n^{-1}\Big|\max_{1\le k\le n}\Big|\sum_{j=0}^{k-1}\big(Z\mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}}\big)\circ f_\Delta^j - k\int_\Delta Z\mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}}\,d\mu_\Delta\Big|\Big|_{L^1(\Delta)} \ll \varepsilon^{1-\alpha/2}.$$

Also, by Proposition 4.1 with $g = \mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}} - \int_\Delta \mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}}\,d\mu_\Delta$,

$$\Big|\max_{1\le k\le n}\Big|\sum_{j=0}^{k-1}\mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}}\circ f_\Delta^j - k\int_\Delta \mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}}\,d\mu_\Delta\Big|\Big|_{L^2(\Delta)} \ll n^{1/2}.$$

23

Combining these two estimates yields

$$\limsup_{n\to\infty} b_n^{-1} \left\| \max_{1\le k\le n} \left| \sum_{j=0}^{k-1} \left(\widetilde{Z}\mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}}\right) \circ f_\Delta^j - k \int_\Delta \widetilde{Z}\mathbf{1}_{\{|\widetilde{Z}|<\varepsilon b_n\}} d\mu_\Delta \right| \right\|_{L^1(\Delta)} \ll \varepsilon^{1-\alpha/2}.$$

Hence Condition II follows from Markov's inequality and the fact that $\pi_\Delta : \Delta \to \Sigma$ is a measure-preserving semiconjugacy. This completes the proof of Theorem 5.1.

6 Proof of Theorem 3.4

In this section, we complete the proof of Theorem 3.4, using the results of Section 4 (specifically Lemma 4.5 and Remark 4.7) and the $\mathcal{J}_1$ convergence from Section 5. The structure of the proof broadly follows [18, 35].

In Section 5, we defined processes

$$W_n^Z \in D([0,1], \mathbb{R}^e), \qquad W_n^Z(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} \widetilde{Z} \circ f^j,$$

where $\widetilde{Z} = Z - \int_\Sigma Z \, d\mu$. By Theorem 5.1, $W_n^Z \to_\mu \tilde{L}_\alpha$ in the $\mathcal{J}_1$ topology, where $\tilde{L}_\alpha$ is an e-dimensional α-stable Lévy process. Recall that $L_\alpha = \bar{R}^{-1/\alpha}\tilde{L}_\alpha$ is the α-stable Lévy process corresponding to the stable law $\bar{R}^{-1/\alpha}G_\alpha$ and that $L_\alpha^P \in \mathscr{D}([0,1], \mathbb{R}^d)$ is the decorated Lévy process.

Step 1: Define processes on (Σ, μ),

$$U_n \in D([0,1], \mathbb{R}^e), \qquad U_n(t) = b_n^{-1} \sum_{j=0}^{N_{[nt]}^R-1} \widetilde{Z} \circ f^j,$$

where N_k^R is the largest integer n satisfying $\sum_{j=0}^{n-1} R \circ f^j \le k$. (This is analogous to, but different from, the lap number N_n defined in Section 4.) Then U_n is a time-changed version of W_n^Z, and using Theorem 5.1 we show that

$$U_n \to_\mu L_\alpha \quad \text{in the } \mathcal{J}_1 \text{ topology.}$$

Step 2: We define decorated processes $G(U_n) \in \mathscr{D}([0,1], \mathbb{R}^d)$ and show that

$$G(U_n) \to_\mu L_\alpha^P \quad \text{in the } \alpha_\infty \text{ topology.}$$

Step 3: Recall that $\widehat{W}_n \in \mathscr{D}([0,1], \mathbb{R}^d)$ is the trivial lift of W_n. We show that $\alpha_\infty(G(U_n), \widehat{W}_n) \to_\mu 0$. From this and Step 2,

$$\widehat{W}_n \to_\mu L_\alpha^P \quad \text{in the } \alpha_\infty \text{ topology.}$$

Step 4: We apply [47, Theorem 1] to show that $\widehat{W}_n \to_{\mu_\Lambda} L_\alpha^P$ in the α_∞ topology.

In the remainder of this section we implement the steps above. To carry out Step 1, we apply the argument in [18, Lemma 5.7]. In this way we obtain the following consequence of Theorem 5.1:

Corollary 6.1. $U_n \to_\mu L_\alpha$ in the $\mathcal{J}_1$ topology. ∎

Turning to Step 2, recall $\Gamma \in \mathbb{R}^{d \times e}$ and the map $\Phi : \mathbb{S}^{e-1} \to \mathrm{Lip}([0,1], \mathbb{R}^d)$ from Section 3.2. Define

$$G : D([0,1], \mathbb{R}^e) \to \mathscr{D}([0,1], \mathbb{R}^d), \qquad h \mapsto G(h)$$

where

$$G(h)(t)(s) = \begin{cases} \Gamma h(t^-) + |h(t) - h(t^-)|\Phi\big(\frac{h(t)-h(t^-)}{|h(t)-h(t^-)|}\big)(s) & \text{if } s \in [0,1) \text{ and } h(t^-) \neq h(t) \\ \Gamma h(t) & \text{if } s = 1 \text{ or } h(t^-) = h(t) \end{cases}$$

Lemma 6.2. *The map G is continuous from $(D([0,1], \mathbb{R}^e), \sigma_\infty)$ to $(\mathscr{D}([0,1], \mathbb{R}^d), \alpha_\infty)$.*

Proof. It is clear that $G(h) \in \mathscr{D}([0,1], \mathbb{R}^d)$ for all $h \in D([0,1], \mathbb{R}^e)$ since $G(h)(\cdot)(1) = \Gamma h$ and h is càdlàg by assumption.

To prove continuity of G, let $\varepsilon > 0$ and $g, h \in D([0,1], \mathbb{R}^e)$ with $\sigma_\infty(g, h) < \varepsilon$. Choose $\lambda \in \Lambda$ such that $|g \circ \lambda - h|_\infty < \varepsilon$ and $|\lambda - \mathrm{Id}|_\infty < \varepsilon$. Let $\Pi = \{t_1, t_2 \ldots\} \subset [0,1]$ be the set of discontinuities of the triple $(g, g \circ \lambda, h)$ ordered arbitrarily. All δ-extensions taken below are with respect to the set Π.

Let $|\Gamma|$ denote the operator norm of Γ and let $|\Phi|_\infty = \sup_{x \in \mathbb{S}^{e-1}, s \in [0,1]} |\Phi(x)(s)|$. Choose $L > 0$ such that $|\Phi(x)(s) - \Phi(y)(s)| \leq L|x - y|$. For $\delta > 0$, we show that

$$|G(g \circ \lambda)^\delta - G(h)^\delta|_\infty \leq |\Gamma|\varepsilon + 8\varepsilon|\Phi|_\infty + 8\varepsilon L. \tag{6.1}$$

Note also that $G(g \circ \lambda)^\delta$ is a reparametrisation of $G(g)^\delta$, so $\sigma_\infty(G(g)^\delta, G(g \circ \lambda)^\delta) \leq \varepsilon + \delta$. Hence $\sigma_\infty(G(g)^\delta, G(h)^\delta) \leq |\Gamma|\varepsilon + 8\varepsilon|\Phi|_\infty + 8\varepsilon L + \varepsilon + \delta$ and the result follows.

Turning to the verification of (6.1), let $\tilde{g} = g \circ \lambda$ and set $a = G(\tilde{g})$, $b = G(h)$. Let $I_j = [c_j, d_j] \subset [0, 1+\delta]$ be the fictitious time intervals appearing in the construction of a^δ and b^δ (see Section 2.1) and denote $J = \bigcup_j [c_j, d_j]$. Since $|\tilde{g} - h|_\infty < \varepsilon$,

$$\sup_{t \in [0,1+\delta] \setminus J} |a^\delta(t) - b^\delta(t)| \leq \sup_{t \in [0,1]} |\Gamma(\tilde{g}(t) - h(t))| \leq |\Gamma|\varepsilon. \tag{6.2}$$

Consider now $t_j \in \Pi$. Let $\Delta_g = \tilde{g}(t_j^-) - \tilde{g}(t_j)$, $\Delta_h = h(t_j^-) - h(t_j)$ and $\Delta = \max\{|\Delta_g|, |\Delta_h|\}$. We have the basic estimate

$$\sup_{t \in [c_j, d_j)} |a^\delta(t) - b^\delta(t)| \leq |\Gamma(\tilde{g}(t_j) - h(t_j))| + 2\Delta|\Phi|_\infty \leq |\Gamma|\varepsilon + 2\Delta|\Phi|_\infty. \tag{6.3}$$

In addition, we note that $|\Delta_g - \Delta_h| < 2\varepsilon$, so if $\Delta > 4\varepsilon$ then

$$\left| \frac{\Delta_g}{|\Delta_g|} - \frac{\Delta_h}{|\Delta_h|} \right| \leq 2|\Delta_g - \Delta_h|/|\Delta_g| < 4\varepsilon/(\Delta - 2\varepsilon) < 8\varepsilon/\Delta.$$

Again in the case $\Delta > 4\varepsilon$, note that t_j is necessarily a discontinuity of both $\tilde{g}$ and h. Hence

$$\sup_{t \in [c_j, d_j)} |a^\delta(t) - b^\delta(t)| \leq |\Gamma(\tilde{g}(t_j^-) - h(t_j^-))| + |\Delta_g - \Delta_h||\Phi|_\infty$$

$$+ |\Delta_h|\sup_{s \in [0,1]} |\Phi(\Delta_g/|\Delta_g|)(s) - \Phi(\Delta_h/|\Delta_h|)(s)|$$

$$\leq |\Gamma|\varepsilon + 2\varepsilon|\Phi|_\infty + 8\varepsilon L. \tag{6.4}$$

Applying (6.3) for $\Delta \leq 4\varepsilon$ and (6.4) for $\Delta > 4\varepsilon$, we obtain

$$\sup_{t \in I_j} |a^\delta(t) - b^\delta(t)| \leq |\Gamma|\varepsilon + 8\varepsilon|\Phi|_\infty + 8\varepsilon L.$$

This combined with (6.2) yields the required estimate (6.1) ∎

25

We now complete Step 2, noting that $L_\alpha^P = G(L_\alpha)$.

Corollary 6.3. $G(U_n) \to_\mu L_\alpha^P$ in the α_∞ topology.

Proof. This follows from Corollary 6.1, Lemma 6.2 and the continuous mapping theorem. ∎

Next, we carry out Step 3. Recall that $v : \Lambda \to \mathbb{R}^d$ is Hölder with $\int_\Lambda v \, d\mu_\Lambda = 0$. Define $R' : \Sigma \to \mathbb{Z}$ and $V : \Sigma \to \mathbb{R}^d$,

$$R' = R^{1-\eta} + R\mathbf{1}_{\Sigma_0}, \qquad V = \sum_{j=0}^{R-1} v \circ T^j,$$

where $\eta > 0$ is as in (3.5).

Proposition 6.4. There exists $p > \alpha$ such that $R' \in L^p$. Moreover, $V = \Gamma Z + H$ where $H \in L^p$.

Proof. Since $R \in L^q$ for all $q < \alpha$, it follows that $R^{1-\eta} \in L^p$ for some $p > \alpha$. Also, $R\mathbf{1}_{\Sigma_0} \in L^p$ for some $p > \alpha$ by (3.4).

By (3.5),
$$|H| = |V - \Gamma Z| = |V - P_i(1)R| \ll R^{1-\eta}$$

on Σ_i, $i = 1, \ldots, e$. On Σ_0, we have $|H| = |V| \ll R$. Hence $|H| \ll R' \in L^p$. ∎

Note that $\int_\Sigma V \, d\mu = 0$. Hence $V = \Gamma \widetilde{Z} + \widetilde{H}$ where $\widetilde{H} = H - \int_\Sigma H \, d\mu$. Define $\widetilde{H}_k = \sum_{j=0}^{k-1} \widetilde{H} \circ f^j$.

Using that R is the first return time to Σ, define $R_\Lambda : \Lambda \to \mathbb{Z}$ so that $R_\Lambda(T^\ell x) = R(x)$ for every $x \in \Sigma$ and $0 \leq \ell < R(x)$.

Lemma 6.5. There exists $C > 0$ such that

$$\alpha_\infty(G(U_n), \widehat{W}_n) \leq \max_{k < n}\{b_n^{-1}|\widetilde{H}_k| + Cb_n^{-1}R' \circ f^k + n^{-1}R \circ f^k\} + |v|_\infty b_n^{-1}R_\Lambda \circ T^n$$

for all $n \geq 1$.

Proof. Set $t_0 = 0$ and $t_{k+1} = t_k + n^{-1}R \circ f^k$, so that the time intervals $[t_k, t_{k+1}]$ correspond to excursions from Σ. In particular, $W_n(t_k) = b_n^{-1}\sum_{j=0}^{k-1} V \circ f^j$ and $U_n(t_k) = b_n^{-1}\sum_{j=0}^{k-1} \widetilde{Z} \circ f^j$. Hence

$$W_n(t_k) - \Gamma U_n(t_k) = b_n^{-1}\sum_{j=0}^{k-1}(V - \Gamma\widetilde{Z}) \circ f^j = b_n^{-1}\widetilde{H}_k. \qquad (6.5)$$

Let $k_* = \max\{k : t_k \leq 1\}$. Note that t_k and k_* are random variables depending also on n, and that $k_* \leq n$.

We are going to bound $\alpha_{\infty;(t_k,t_{k+1}]}(G(U_n), \widehat{W}_n)$ with $0 \leq k < k_*$ and $\alpha_{\infty;(t_{k_*},1]}(G(U_n), \widehat{W}_n)$ separately. (The last term is absent if $t_{k_*} = 1$.) By Proposition 2.4,

$$\alpha_\infty(G(U_n), \widehat{W}_n) \leq \max\Big\{\max_{0 \leq k < k_*} \alpha_{\infty;(t_k,t_{k+1}]}(G(U_n), \widehat{W}_n) \,,\; \alpha_{\infty;(t_{k_*},1]}(G(U_n), \widehat{W}_n)\Big\}. \qquad (6.6)$$

Let $0 \leq k < k_*$. Define $E \in D([t_k, t_{k+1}], \mathbb{R}^d)$,

$$E(t) = G(U_n)(t_{k+1})(s) \qquad s = \frac{t - t_k}{t_{k+1} - t_k} = \frac{(t - t_k)n}{R \circ f^k}.$$

Let $\widehat{E} = \iota E \in \mathscr{D}((t_k, t_{k+1}], \mathbb{R}^d)$ denote the trivial lift of E from Definition 2.5. Recall that U_n is constant on $[t_k, t_{k+1})$. By Proposition 2.7,

$$\alpha_{\infty;(t_k,t_{k+1}]}(G(U_n), \widehat{E}) \le t_{k+1} - t_k = n^{-1} R \circ f^k.$$

Next, we estimate $|W_n - E|_{\infty,(t_k,t_{k+1}]}$. By (6.5),

$$W_n(t_k) - E(t_k) = W_n(t_k) - \Gamma U_n(t_k) = b_n^{-1} \widetilde{H}_k. \tag{6.7}$$

Note that $U_n(t_{k+1}) - U_n(t_k) = b_n^{-1} \widetilde{Z} \circ f^k$. Hence, for $t \in (t_k, t_{k+1})$,

$$E(t) - E(t_k) = b_n^{-1} \big\{ |\widetilde{Z}| \, \Phi(\widetilde{Z}/|\widetilde{Z}|)(s) \big\} \circ f^k.$$

Recall that $\widetilde{Z} = Z - \int_\Sigma Z \, d\mu$, $Z = \sum_{i=1}^e \omega_i R \mathbf{1}_{\Sigma_i}$. Since $\Phi : \mathbb{S}^{e-1} \to \mathrm{Lip}([0,1], \mathbb{R}^d)$ is constant on a neighbourhood of each ω_i with value P_i, there exists $R_0 > 0$ such that

$$\mathbf{1}_{\Sigma_i} \Phi(\widetilde{Z}/|\widetilde{Z}|) = \mathbf{1}_{\Sigma_i} \Phi(Z/|Z|) = \Phi(\omega_i) = P_i$$

for $R \ge R_0$ for each $i = 1, \ldots, e$. Also, $\mathbf{1}_{\Sigma_0} |\widetilde{Z}| \, \Phi(\widetilde{Z}/|\widetilde{Z}|) \le |\int_\Sigma Z \, d\mu|_\infty |\Phi|_\infty < \infty$. Hence, for $t \in (t_k, t_{k+1})$,

$$E(t) - E(t_k) = b_n^{-1} \sum_{i=1}^e \Big\{ R \mathbf{1}_{\Sigma_i} P_i \big(\tfrac{(t-t_k)n}{R} \big) \Big\} \circ f^k + O(b_n^{-1}). \tag{6.8}$$

On the other hand, for $t \in (t_k, t_{k+1}]$, it follows from (3.6) that

$$W_n(t) - W_n(t_k) = b_n^{-1} \sum_{\ell=0}^{[(t-t_k)n]-1} v \circ T^\ell \circ f^k \tag{6.9}$$

$$= b_n^{-1} \Big\{ \sum_{i=1}^e R \mathbf{1}_{\Sigma_i} P_i \big(\tfrac{(t-t_k)n}{R} \big) + O(R^{1-\eta}) + O(R \mathbf{1}_{\Sigma_0}) \Big\} \circ f^k.$$

Combining (6.7), (6.8) and (6.9), there exists $C > 0$ such that

$$|W_n - E|_{\infty;(t_k,t_{k+1}]} \le b_n^{-1} \big\{ \max\{ |\widetilde{H}_k|, |\widetilde{H}_{k+1}| \} + C R' \circ f^k \big\}.$$

Since $\widehat{W}_n$ and $\widehat{E}$ are trivial lifts, it follows from Remark 2.6 that

$$\alpha_{\infty;(t_k,t_{k+1}]}(\widehat{W}_n, \widehat{E}) \le |W_n - E|_{\infty;(t_k,t_{k+1}]} \le b_n^{-1} \big\{ \max\{ |\widetilde{H}_k|, |\widetilde{H}_{k+1}| \} + C R' \circ f^k \big\}.$$

Combining the estimates for $\alpha_\infty(G(U_n), \widehat{E})$ and $\alpha_\infty(\widehat{W}_n, \widehat{E})$, we obtain

$$\alpha_{\infty;(t_k,t_{k+1}]}(G(U_n), \widehat{W}_n) \le \alpha_{\infty;(t_k,t_{k+1}]}(G(U_n), \widehat{E}) + \alpha_{\infty;(t_k,t_{k+1}]}(\widehat{W}_n, \widehat{E})$$
$$\le \max_{k \le n} \{ b_n^{-1} |\widetilde{H}_k| + C b_n^{-1} R' \circ f^k + n^{-1} R \circ f^k \}. \tag{6.10}$$

Finally, we treat the interval $(t_{k_*}, 1]$ for $t_{k_*} < 1$. On this interval, $G(U_n) = \Gamma U_n(t_{k_*})$ is constant. Note that $R \circ f^{k_*} = R_\Lambda \circ T^n$ and it follows that $|W_n(t_{k_*}) - W_n|_{\infty;(t_{k_*},1]} \le b_n^{-1} |v|_\infty R_\Lambda \circ T^n$. By Remark 2.6 and (6.5),

$$\alpha_{\infty;(t_{k_*},1]}(G(U_n), \widehat{W}_n) \le |\Gamma U_n(t_{k_*}) - W_n|_{\infty;(t_{k_*},1]}$$
$$\le |\Gamma U_n(t_{k_*}) - W_n(t_{k_*})| + |W_n(t_{k_*}) - W_n|_{\infty;(t_{k_*},1]} \tag{6.11}$$
$$\le b_n^{-1} \big\{ |\widetilde{H}_{k_*}| + |v|_\infty R_\Lambda \circ T^n \big\}.$$

The result follows from (6.6), (6.10) and (6.11). $\blacksquare$

27

We can now complete Step 3.

Corollary 6.6. $\alpha_\infty(G(U_n), \widehat{W}_n) \to_\mu 0$.

Proof. We show that all the terms on the right-hand side in Lemma 6.5 converge to 0 in probability on (Σ, μ).

Let $p > \alpha$ be as in Proposition 6.4. By Proposition 6.4, $\widetilde{H} = V - \Gamma \widetilde{Z} \in L^p$. Also $\int_\Sigma \widetilde{H} \, d\mu = 0$. For x, x' in the same element of $\mathcal{C}$, we have $\widetilde{H}(x) - \widetilde{H}(x') = V(x) - V(x')$. Hence it follows from Corollary 4.6 and Remark 4.7 that $b_n^{-1} |\max_{k \le n} |\widetilde{H}_k|| \to_\mu 0$.

By the pointwise ergodic theorem $n^{-1/q} \max_{k \le n} g \circ f^k \to 0$ a.e. for all L^q functions $g : \Sigma \to \mathbb{R}$, $1 \le q < \infty$. Taking $q = 1$, it follows that $n^{-1} \max_{k \le n} R \circ f^k \to_\mu 0$. Taking $q = p$, we have $b_n^{-1} \max_{k \le n} R' \circ f^k \ll n^{-1/p} \max_{k \le n} R' \circ f^k \to_\mu 0$.

Finally, $b_n^{-1} R_\Lambda \to_\mu 0$ a.e. on (Λ, μ_Λ). Since μ_Λ is T-invariant, $b_n^{-1} R_\Lambda \circ T^n \to_{\mu_\Lambda} 0$. But μ is absolutely continuous with respect to μ_Λ, so $b_n^{-1} R_\Lambda \circ T^n \to_\mu 0$ too. ∎

The key ingredient of Step 4 is:

Proposition 6.7. $\alpha_\infty(\widehat{W}_n \circ T, \widehat{W}_n) \le 2|v|_\infty b_n^{-1}$.

Proof. Since $\widehat{W}_n$ (and hence $\widehat{W}_n \circ T$) is a trivial lift, it follows from Remark 2.6 that

$$\alpha_\infty(\widehat{W}_n \circ T, \widehat{W}_n) \le |W_n \circ T - W_n|_\infty \le 2|v|_\infty b_n^{-1},$$

as required. ∎

In particular, $\alpha_\infty(\widehat{W}_n \circ T, \widehat{W}_n) \to_\mu 0$. Hence we have verified [47, Condition (1)] in Proposition 6.7. Since μ is absolutely continuous with respect to μ_Λ, it follows from [47, Theorem 1] that we can replace μ in Step 3 by μ_Λ, thereby completing the proof of Theorem 3.4.

7 Tightness of W_n in p-variation

In this section, we prove Theorem 3.5. Recall that we have mixing ergodic dynamical systems $(T, \Lambda, \mu_\Lambda)$, (f, Σ, μ) and (F, Y, μ_Y) where $f = T^R$ and $F = f^\tau$, Here, $R : Y \to \mathbb{Z}^+$ is a first return time and $\tau : Y \to \mathbb{Z}^+$ is a return time with exponential tails. Moreover (F, Y, μ_Y) quotients to a Gibbs-Markov map $(\bar{F}, \bar{Y}, \mu_{\bar{Y}})$.

Define the induced return time function

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

Proposition 7.1. $\mu_Y(R^Y > t) \sim c\ell(t)t^{-\alpha}$ as $t \to \infty$ for some $c > 0$.

Proof. In this proof, we denote Birkhoff sums under f by $R_n = \sum_{j=0}^{n-1} R \circ f^j$ and so on. Let M denote the $1 \times e$ matrix all of whose entries are 1. Then MG_α is a nondegenerate totally skewed 1-dimensional α-stable law. Let $\widetilde{R} = R - \int_\Sigma R \, d\mu$. We claim that $b_n^{-1} \widetilde{R}_n \to_\mu MG_\alpha$.

The result follows from the claim by an argument in [33, Lemma 5.1(c)]. Indeed, define the piecewise constant function $\widetilde{R}^Y = \sum_{j=0}^{\tau-1} \widetilde{R} \circ f^j$. By Lemma 4.8,

$$b_n^{-1} \sum_{j=0}^{n-1} \widetilde{R}^Y \circ F^j \to_{\mu_Y} \bar{\tau}^{1/\alpha} MG_\alpha.$$

By Proposition 4.2, $\mu_Y(\widetilde{R}^Y > t) \sim c\ell(t)t^{-\alpha}$. Finally, $R^Y = \widetilde{R}^Y + \tau \int_\Sigma R\,d\mu$ and τ has exponential tails, so $\mu_Y(R^Y > t) \sim c\ell(t)t^{-\alpha}$.

It remains to prove the claim. Write $R = Q + Q'$ where $Q = R\mathbf{1}_{\Sigma\setminus\Sigma_0}$ and $Q' = R\mathbf{1}_{\Sigma_0}$. Define $\widetilde{Q} = Q - \int_\Sigma Q\,d\mu$, $\widetilde{Q}' = Q' - \int_\Sigma Q'\,d\mu$.

Since $M\omega_i = 1$ for all i, it follows that $\widetilde{Q} = M\widetilde{Z}$. Evaluating at $t = 1$ in Theorem 5.1, $b_n^{-1}\widetilde{Z}_n \to_\mu \tilde{L}_\alpha(1) = G_\alpha$. By the continuous mapping theorem, $b_n^{-1}\widetilde{Q}_n = b_n^{-1}M\widetilde{Z} \to_\mu MG_\alpha$.

By Proposition 6.4, $R\mathbf{1}_{\Sigma_0} \in L^q$ for some $q > \alpha$ and hence $\widetilde{Q}' \in L^q$. Also, $\int_\Sigma \widetilde{Q}'\,d\mu = 0$ and $\widetilde{Q}'$ is constant on elements of $\mathcal{C}$. By Corollary 4.6, $b_n^{-1}\widetilde{Q}'_n \to_\mu 0$.

The claim for $\widetilde{R} = \widetilde{Q} + \widetilde{Q}'$ now follows by combining the results for $\widetilde{Q}$ and $\widetilde{Q}'$. ∎

Recall that $V : \Sigma \to \mathbb{R}^d$ is given by $V = \sum_{j=0}^{R-1} v \circ T^j$. Define

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

Define the corresponding processes on (Y, μ_Y),

$$W_n^Y \in D([0,1], \mathbb{R}^d), \qquad W_n^Y(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} V^Y \circ F^j.$$

Lemma 7.2. $\sup_n \int_Y |W_n^Y|_{p\text{-var}}\,d\mu_Y < \infty$ for all $p > \alpha$.

Proof. Let $H = V - \sum_{i=1}^e P_i(1)\widetilde{R}_i$ where $R_i = R\mathbf{1}_{\Sigma_i}$, $\widetilde{R}_i = R_i - \int_\Sigma R_i\,d\mu$. Then $\int_\Sigma H\,d\mu = 0$. Also, $H = V - \sum_{i=1}^e P_i(1)R_i + C$ where $C = \sum_{i=1}^e P_i(1)\int_\Sigma R_i\,d\mu \in \mathbb{R}^d$ is a constant. In particular, $|H| \le |v|_\infty R + |C|$ on Σ_0 and $H = V - P_i(1)R + |C|$ on Σ_i, $i = 1,\ldots,e$. Hence by assumptions (3.4) and (3.5), $H \in L^q(\Sigma, \mu)$ for some $q > \alpha$. Let $p \in (\alpha, q)$. It suffices to show that $\sup_n \int_Y |W_n^Y|_{p\text{-var}}\,d\mu_Y < \infty$.

Write $V^Y = \sum_{i=1}^e P_i(1)\widetilde{R}_i^Y + H^Y : Y \to \mathbb{R}^d$, where

$$\widetilde{R}_i^Y = \sum_{j=0}^{\tau-1} \widetilde{R}_i \circ f^j, \qquad H^Y = \sum_{j=0}^{\tau-1} H \circ f^j.$$

Correspondingly, write $W_n^Y = \sum_{i=1}^e P_i(1)A_{i,n} + B_n$ on Y where

$$A_{i,n}(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} \widetilde{R}_i^Y \circ F^j, \qquad B_n(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} H^Y \circ F^j.$$

Now $\widetilde{R}_i^Y$ is piecewise constant on Y and hence well-defined as a piecewise constant function $\bar{Y}$. Let

$$\bar{A}_{i,n}(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} \widetilde{R}_i^{\bar{Y}} \circ \bar{F}^j : \bar{Y} \to \mathbb{R}.$$

Let $\sigma = R^{\bar{Y}} + (\int_\Sigma R\,d\mu)\tau$. We are going to apply [13, Theorem 4.4] (with $\bar{Y}$, $\widetilde{R}_i^{\bar{Y}}$ and σ playing the roles of Z, V and τ in [13]). By Proposition 7.1, $R^{\bar{Y}}$ is regularly varying with exponent α. Since τ has exponential tails, σ is regularly varying with exponent α. For all $y, y' \in \bar{Y}_j$ and all partition elements $\bar{Y}_j$,

$$|\widetilde{R}_i^{\bar{Y}}(y)| \le \sigma(y), \qquad \widetilde{R}_i^{\bar{Y}}(y) - \widetilde{R}_i^{\bar{Y}}(y') = 0,$$

29

verifying the condition [13, eq. (4.2)]. Hence it follows from [13, Theorem 4.4] that

$$\sup_n \int_Y |A_{i,n}|_{p\text{-var}} \, d\mu_Y = \sup_n \int_{\bar{Y}} |\bar{A}_{i,n}|_{p\text{-var}} \, d\mu_{\bar{Y}} < \infty \quad \text{for each } i = 1, \ldots, e. \qquad (7.1)$$

Next, by Lemma 4.5,

$$H^Y = m + \chi \circ F - \chi,$$

where $m, \chi \in L^p(\mu_Y)$ and $\{m \circ F^{n-j} : 0 \leq j \leq n\}$ is a martingale difference sequence for each $n \geq 1$. By a simpler argument than the one in the proof of [13, Theorem 4.4], we show that $\lim_{n\to\infty} \int_Y |B_n|_{p\text{-var}}^p \, d\mu_Y = 0$. Indeed, we can write $B_n = M_n + D_n$ where

$$M_n(t) = b_n^{-1} \sum_{j=0}^{[nt]-1} m \circ F^j, \qquad D_n(t) = b_n^{-1}(\chi \circ F^{[nt]} - \chi).$$

Define also $M_n^-(t) = b_n^{-1} \sum_{j=1}^{[nt]} m \circ F^{n-j}$, so that $M_n^-(t)$ is a martingale. By [39, Theorem 2.1], there is $C_p > 0$ depending only on p so that

$$\int |M_n^-|_{p\text{-var}}^p \, d\mu_Y \leq C_p b_n^{-p} \sum_{j=1}^n \int |m \circ F^{n-j}|^p \, d\mu_Y = C_p n b_n^{-p} \int |m|^p \, d\mu_Y.$$

Hence

$$\int |M_n|_{p\text{-var}}^p \, d\mu_Y = \int |M_n^-|_{p\text{-var}}^p \, d\mu_Y \leq C_p n b_n^{-p} \int |m|^p \, d\mu_Y. \qquad (7.2)$$

Further, for $1 \leq j_1 < j_2 < \cdots < j_k \leq n-1$,

$$\sum_{i=1}^k |\chi \circ F^{j_i} - \chi \circ F^{j_{i-1}}|^p \leq \sum_{i=1}^k \left(|\chi| \circ F^{j_i} + |\chi| \circ F^{j_{i-1}}\right)^p$$

$$\leq 2^{p-1} \sum_{i=1}^k \left(|\chi|^p \circ F^{j_i} + |\chi|^p \circ F^{j_{i-1}}\right) \leq 2^p \sum_{j=0}^{n-1} |\chi|^p \circ F^j.$$

We deduce that $|D_n|_{p\text{-var}}^p \leq 2^p b_n^{-p} \sum_{j=0}^{n-1} |\chi|^p \circ F^j$, and hence

$$\int_Y |D_n|_{p\text{-var}}^p \, d\mu_Y \leq 2^p n b_n^{-p} \int |\chi|^p \, d\mu_Y. \qquad (7.3)$$

Combining the estimates (7.2) and (7.3) and using $\lim_{n\to\infty} n b_n^{-p} = 0$, we see that $\lim_{n\to\infty} \int_Y |B_n|_{p\text{-var}}^p \, d\mu_Y = 0$. This, together with (7.1), shows that $\sup_n \int_Y |W_n^Y|_{p\text{-var}} \, d\mu_Y < \infty$, as required. ∎

Proof of Theorem 3.5. Note that $V^Y = \sum_{j=0}^{R^Y-1} v \circ T^j$. By Proposition 7.1, R^Y is regularly varying. Applying [13, Theorem 5.2] (with Λ, Y, W_n^Y and R^Y playing the roles of Y, Z, $\widetilde{W}_n$ and τ in [13]) and Lemma 7.2, we obtain that $|W_n|_{p\text{-var}}$ is tight on (Λ, μ_Λ). ∎

8 Path space – Proofs

In this section we define the space $\mathscr{D}$ in more detail, collect its important properties, construct solutions of differential equations driven by elements of $\mathscr{D}$, and prove that the solution map is continuous (Theorem 2.10).

We follow the notation and terminology introduced in Section 2 and often drop reference to the target space $\mathbb{R}^d$ and write, e.g. $D[a, b]$ for $D([a, b], \mathbb{R}^d)$. We also drop reference to the interval $[a, b]$ when it is clear from the context. For intervals $I, J \subset \mathbb{R}$, let $\Lambda_{I,J}$ denote the set of all continuous increasing bijections $\lambda : I \to J$. We use the shorthand $\Lambda_I = \Lambda_{I,I}$.

8.1 The decorated path space $\mathscr{D}$

Definition 8.1. Define the space $\mathcal{D} = \bigcup_{a<b} D[a,b]$ of càdlàg paths parametrised by compact intervals. For $h_1 \in D[a_1, b_1]$, $h_2 \in D[a_2, b_2]$, define the *Fréchet distance* by

$$d_F(h_1, h_2) = \inf_{\lambda \in \Lambda_{[a_1,b_1],[a_2,b_2]}} |h_1 \circ \lambda - h_2|_\infty. \tag{8.1}$$

We say that h_1 is a *reparametrisation* of h_2, and write $h_1 \sim_F h_2$, if $d_F(h_1, h_2) = 0$. Denote $[h] = \{h' \in \mathcal{D} : h \sim_F h'\}$.

Remark 8.2. d_F defines a metric on $\mathcal{D}/\sim_F$. Indeed, this follows from the fact that h_1, h_2 are reparametrisations if and only if there exist non-decreasing surjections $\lambda_1 : [0,1] \to [a_1, b_1]$ and $\lambda_2 : [0,1] \to [a_2, b_2]$ such that $h_1 \circ \lambda_1 = h_2 \circ \lambda_2$. See for example [4, Exercise 2.5.3] and [5, Proposition 5.3].

Lemma 8.3. *$\mathcal{D}/\sim_F$ is complete under the metric d_F.*

Proof. We can extract from a Cauchy sequence $[X_n]$ a Cauchy sequence $Y_n \in [X_n]$ in the space $D[0,1]$ with the uniform norm, and apply the fact that this space is complete. ∎

For the rest of this subsection , consider $\delta > 0$, a function $\phi : [a,b] \to D([0,1], \mathbb{R}^d)$, and a countable set $\Pi = \{t_1, t_2, \ldots\} \subset [a,b]$ such that Π contains the non-stationary points of ϕ. We construct functions $\phi^\delta : [a, b+\delta] \to \mathbb{R}^d$ and $\psi^\delta : [a, b+\delta] \to \mathbb{R}^d \times [a,b]$ as follows.

First we define $\tau_\delta : [a,b] \to [a, b+\delta]$. Let $0 \le \kappa \le \infty$ be the cardinality of Π. If $\kappa = 0$, we set $\tau_\delta(t) = t$. Otherwise, let $r = \sum_{j=1}^\kappa 2^{-j} > 0$ and

$$\tau_\delta(t) = t + \sum_{j=1}^\kappa \frac{\delta 2^{-j}}{r} \mathbf{1}_{t_j \le t}.$$

Note that τ_δ is a strictly increasing càdlàg function with $\tau_\delta(t^-) < \tau_\delta(t)$ if and only if $t = t_j$ for some $1 \le j < \kappa + 1$. Moreover, the interval $[\tau_\delta(t_j^-), \tau_\delta(t_j))$ is of length $\delta 2^{-j}/r$.

Let $\phi^\delta(t) = \phi(\min\{t, b\})$ if $\kappa = 0$ and otherwise

$$\phi^\delta(u) = \begin{cases} \phi(t)(1) & \text{if } u = \tau_\delta(t) \text{ for some } t \in [a,b], \\ \phi(t_j)\left(\frac{u - \tau_\delta(t_j^-)}{\delta 2^{-j}/r}\right) & \text{if } u \in [\tau_\delta(t_j^-), \tau_\delta(t_j)) \text{ for some } 1 \le j < \kappa + 1. \end{cases} \tag{8.2}$$

We call ϕ^δ the *δ-extension* of (ϕ, Π).

Furthermore, let $\tau_\delta^{-1} : [a, b+\delta] \to [a,b]$ be a left inverse of τ_δ defined by

$$\tau_\delta^{-1}(t) = \inf\{s \in [a,b] : \tau_\delta(s) \ge t\}.$$

This way, τ_δ is a non-decreasing surjection which maps each interval $[\tau_\delta(t_j^-), \tau_\delta(t_j)]$ to t_j, and $\tau_\delta^{-1} \circ \tau_\delta = \mathrm{Id}$. Finally, define

$$\psi^\delta(u) = (\phi^\delta(u), \tau_\delta^{-1}(u)). \tag{8.3}$$

We call ψ^δ the *δ-parametric representation* of (ϕ, Π).

Recall the spaces $\bar{\mathscr{D}}[a,b]$ and $\mathscr{D}[a,b]$ as in Definitions 2.1 and 2.2. We write $\phi_1 \sim \phi_2$ if ϕ_1 and ϕ_2 are equivalent in $\bar{\mathscr{D}}[a,b]$, so that $\mathscr{D}[a,b] = \bar{\mathscr{D}}[a,b]/\sim$.

Lemma 8.4. *The following statements are equivalent.*

(i) At least one of ψ^δ or ϕ^δ is càdlàg.

(ii) Both ψ^δ and ϕ^δ are càdlàg.

(iii) $(\phi, \Pi) \in \bar{\mathscr{D}}$.

Proof. (ii) $\Rightarrow$ (i) is trivial while (i) $\Rightarrow$ (ii) follows from the fact that τ_δ^{-1} is continuous. To prove (iii) $\Rightarrow$ (i), suppose that $(\phi, \Pi) \in \bar{\mathscr{D}}$. We will prove that ϕ^δ is càdlàg. If $t \notin \Pi$, then Lemma 8.5 below implies that ϕ^δ is continuous at $\tau_\delta(t)$. If $t \in \Pi$, then Lemma 8.5 implies that ϕ^δ has a left limit at $\tau_\delta(t^-)$ while the fact that $\phi(t) \in D[0,1]$ implies that ϕ^δ is càdlàg on $[\tau_\delta(t^-), \tau_\delta(t)]$. Hence (iii) $\Rightarrow$ (i).

Conversely, suppose that $\phi \notin \bar{\mathscr{D}}$, in which case either condition (c) in Definition 2.1 does not hold or $h : t \mapsto \phi(t)(1)$ is not càdlàg. If h is not càdlàg, then clearly ϕ^δ is also not càdlàg. If Definition 2.1(c) does not hold, then there exists $\varepsilon > 0$ and a subsequence $n(k)$ such that $\sup_{t \in [0,1]} |\phi(t_{n(k)})(t) - \phi(t_n)(0)| > \varepsilon$ for all $k \geq 1$. Furthermore, there exists $t \in [a,b]$ such that either (a) $t_{n(k)} \uparrow t$ as $k \to \infty$ or (b) $t_{n(k)} \downarrow t$ as $k \to \infty$. In case (a), ϕ^δ does not have a left limit at $\tau_\delta(t^-)$, and in case (b), ϕ^δ does not have a right limit at $\tau_\delta(t)$. Therefore (ii) $\Rightarrow$ (iii). ■

Lemma 8.5. *For all $\phi \in \bar{\mathscr{D}}[a,b]$, $t \in [a,b]$ and $\varepsilon > 0$, there exists $\gamma > 0$ such that*

$$\sup_{u \in (t-\gamma,t)} \sup_{s \in [0,1]} |\phi(u)(s) - \phi(t^-)(1)| < \varepsilon, \qquad \sup_{u \in (t,t+\gamma)} \sup_{s \in [0,1]} |\phi(u)(s) - \phi(t)(1)| < \varepsilon.$$

Proof. This follows from conditions (a) and (c) in Definition 2.1. ■

Remark 8.6. For $\phi \in \bar{\mathscr{D}}[a,b]$ and $\delta_1, \delta_2 > 0$, we have $\psi^{\delta_1} \sim_F \psi^{\delta_2}$. Furthermore, ϕ and ψ^δ are related via $\psi^\delta = (\phi^\delta, \tau_\delta^{-1})$ and $\phi(t) \sim_F \phi^\delta |_{[\tau_\delta(t^-),\tau_\delta(t)]}$ for each $t \in [a,b]$. Note that the interval $[\tau_\delta(t^-), \tau_\delta(t)]$ is precisely the set of all $y \in [a, b+\delta]$ such that $\tau_\delta^{-1}(y) = t$. In particular, if $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a,b]$, then $\psi_1^\delta \sim_F \psi_2^\delta$ for some $\delta > 0$ if and only if $\phi_1 \sim \phi_2$.

Remark 8.7. Lemma 8.4 remains true (with essentially the same proof) when one replaces "càdlàg" by "continuous" in (i)–(ii) and $\bar{\mathscr{D}}$ by $\bar{\mathscr{D}}_c$ defined as the space of all pairs $(\phi, \Pi) \in \mathscr{D}$ such that $\phi : [a,b] \to C[0,1]$ and $\phi(t)(0) = \phi(t^-)(1)$ in (iii).

We point out a minor error in [11, 12, 13]: the constructions therein implicitly claimed that ϕ^δ is in $C[a, b+\delta]$ whenever $\phi \in \bar{\mathscr{D}}_c$ without assuming condition (c) in Definition 2.1, which is false. However, the results in those articles always deal with decorated paths (called path functions therein) having further nice properties, such as *endpoint continuity* or *finiteness of p-variation*, for which Definition 2.1(c) holds and thus $\phi^\delta \in C[a, b+\delta]$.

Remark 8.8 (Intrinsic definition of ϕ^δ)**.** The construction of ϕ^δ involves several arbitrary choices: the set $\Pi = \{t_1, t_2, \ldots\}$ with the enumeration of its points and the sequence 2^{-j}. If $\bar{\phi}^\delta$ is constructed similarly, but using another Π and another summable sequence, then ϕ^δ and $\bar{\phi}^\delta$ are reparametrisations of each other. See Lemma 8.12 where these arbitrary choices are in essence removed. The same remarks apply to ψ^δ.

8.2 A generalisation of Skorokhod's $\mathcal{M}_1$ topology

We fix $p \in [1, \infty]$. Recall the semi-norm $|\cdot|_{p\text{-var}}$ and space $D^{p\text{-var}}[a,b]$ from Section 2.2. For later use, we record the interpolation estimate

$$|f|_{q\text{-var}} \leq |f|_{\infty\text{-var}}^{1-p/q} |f|_{p\text{-var}}^{p/q} \quad \text{for} \quad 1 \leq p \leq q \leq \infty. \tag{8.4}$$

For $f : I \to \mathbb{R}$ denote

$$[\![f]\!] = \sup_{x \in I} |f(x) - x|.$$

Remark that if $g : J \to I$, then $[\![f \circ g]\!] \leq [\![f]\!] + [\![g]\!]$.

Lemma 8.9. *Consider $\delta > 0$, $g_i \in D[a,b]$ and $h_i \in D[a, b+\delta]$ for $i = 1, 2$ such that $h_i = g_i \circ \lambda_i$ for some $\lambda_i \in \Lambda_{[a,b+\delta],[a,b]}$. Then*

$$|\sigma_\infty(g_1, g_2) - \sigma_\infty(h_1, h_2)| \leq [\![\lambda_1]\!] + [\![\lambda_2]\!]. \tag{8.5}$$

If $g_i, h_i \in D^{p\text{-var}}[a,b]$, then (8.5) holds with σ_∞ replaced by $\sigma_{p\text{-var}}$.

Proof. By definition of σ_∞,

$$|\sigma_\infty(g_1, g_2) - \sigma_\infty(h_1, h_2)| = \Big| \inf_{\omega_1 \in \Lambda_{[a,b]}} \max\{[\![\omega_1]\!], |g_1 \circ \omega_1 - g_2|_\infty\}$$
$$- \inf_{\omega_2 \in \Lambda_{[a,b+\delta]}} \max\{[\![\omega_2]\!], |g_1 \circ \lambda_1 \circ \omega_2 \circ \lambda_2^{-1} - g_2|_\infty\} \Big|$$

from which (8.5) follows upon making the substitution $\omega_1 = \lambda_1 \circ \omega_2 \circ \lambda_2^{-1}$ and using $[\![\lambda_1 \circ \omega_2 \circ \lambda_2^{-1}]\!] \leq [\![\lambda_1]\!] + [\![\omega_2]\!] + [\![\lambda_2]\!]$. An identical argument shows the final claim. ∎

Lemma 8.10. *For all $\delta_2 > \delta_1 > 0$ and $(\phi, \Pi) \in \bar{\mathscr{D}}[a,b]$, there exists $\lambda \in \Lambda_{[a,b+\delta_1],[a,b+\delta_2]}$ such that $\phi^{\delta_2} \circ \lambda = \phi^{\delta_1}$ and $[\![\lambda]\!] = \delta_2 - \delta_1$.*

Proof. Obvious from the construction of ϕ^δ. ∎

Lemma 8.11. *Consider $\delta_1, \delta_2 > 0$ and $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a,b]$ with $\phi_1 \sim \phi_2$. Then $|\phi_1^{\delta_1}|_{p\text{-var};[a,b+\delta_1]} = |\phi_2^{\delta_2}|_{p\text{-var};[a,b+\delta_2]}$.*

Proof. This is immediate from the fact that $\phi_1^{\delta_1}$ is a reparametrisation of $\phi_2^{\delta_2}$. ∎

Recall the space $\mathscr{D}^{p\text{-var}}[a,b]$ from (2.2).

Lemma 8.12. (i) *The limit $\alpha_\infty(\phi_1, \phi_2) = \lim_{\delta \to 0} \sigma_\infty(\phi_1^\delta, \phi_2^\delta)$ exists and is well-defined for all $(\phi_1, \Pi_1), (\phi_2, \Pi_2) \in \mathscr{D}[a,b]$, i.e. is independent of the choice of the series $\sum_{j=1}^\infty 2^{-j}$ used in Section 8.1, the choice of Π_1 and Π_2, the respective enumeration of points, or the parametrisations of $\phi_1(t)$, $\phi_2(t)$ for each t.*

(ii) *The limit $\alpha_{p\text{-var}}(\phi_1, \phi_2) = \lim_{\delta \to 0} \sigma_{p\text{-var}}(\phi_1^\delta, \phi_2^\delta)$ exists and is well-defined for $\phi_1, \phi_2 \in \mathscr{D}^{p\text{-var}}[a,b]$ in the same sense as in (i).*

Proof. (i) Lemmas 8.9 and 8.10 imply that, for $\delta_2 > \delta_1 > 0$ and $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a,b]$,

$$|\sigma_\infty(\phi_1^{\delta_2}, \phi_2^{\delta_2}) - \sigma_\infty(\phi_1^{\delta_1}, \phi_2^{\delta_1})| \leq 2(\delta_2 - \delta_1),$$

from which the existence of the limit follows. The fact that the limit is independent of the series $\sum_j 2^{-j}$, the enumeration of non-stationary points, and representatives of equivalence classes is clear. (ii) follows in an identical manner. ∎

Lemma 8.13. *Let $\delta > 0$ and let $\psi_1^\delta, \psi_2^\delta : [0,1] \to \mathbb{R}^d \times [a,b]$ be the δ-parametric representations of $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a,b]$ as in (8.3). Then*

$$\alpha_\infty(\phi_1, \phi_2) = d_F(\psi_1^\delta, \psi_2^\delta),$$

where we equip $\mathbb{R}^d \times \mathbb{R}$ with the norm $\max\{|h|, |t|\}$ for $(h,t) \in \mathbb{R}^d \times \mathbb{R}$.

Proof. Observe that $d_F(\psi_1^\delta, \psi_2^\delta)$ does not depend on δ, hence, where convenient, we can suppose that δ is arbitrarily small.

Suppose $\lambda \in \Lambda_{[a,b+\delta]}$. Let $\tau_{1,\delta}, \tau_{2,\delta}$ and $\tau_{1,\delta}^{-1}, \tau_{2,\delta}^{-1}$ be the time-changes associated to ϕ_1, ϕ_2 respectively as in Section 8.1. Then

$$\llbracket \tau_{1,\delta}^{-1} \circ \lambda \circ \tau_{2,\delta} \rrbracket \leq \sup_{t \in [a,b+\delta]} |\tau_{1,\delta}^{-1}(\lambda(t)) - \tau_{2,\delta}^{-1}(t)| \leq \llbracket \tau_{1,\delta}^{-1} \rrbracket + \llbracket \tau_{2,\delta}^{-1} \rrbracket + \llbracket \lambda \rrbracket,$$

where in the first bound we used that $\tau_{2,\delta}^{-1}$ is surjective. Therefore

$$\begin{aligned} |\psi_1^\delta \circ \lambda - \psi_2^\delta|_\infty &= \sup_{t \in [a,b+\delta]} \max\{|\phi_1^\delta(\lambda(t)) - \phi_2^\delta(t)|, |\tau_{1,\delta}^{-1}(\lambda(t)) - \tau_{2,\delta}^{-1}(t)|\} \\ &\geq \max\{|\phi_1^\delta \circ \lambda - \phi_2^\delta|_\infty, \llbracket \tau_{1,\delta}^{-1} \circ \lambda \circ \tau_{2,\delta} \rrbracket\}. \end{aligned}$$

Furthermore, $\llbracket \tau_{1,\delta}^{-1} \circ \lambda \circ \tau_{2,\delta} \rrbracket \to \llbracket \lambda \rrbracket$ as $\delta \to 0$, and thus $d_F(\psi_1^\delta, \psi_2^\delta) \geq \alpha_\infty(\phi_1, \phi_2)$. Conversely,

$$|\psi_1^\delta \circ \lambda - \psi_2^\delta|_\infty \leq \max\{|\phi_1^\delta \circ \lambda - \phi_2^\delta|_\infty, \llbracket \lambda \rrbracket + \llbracket \tau_{1,\delta}^{-1} \rrbracket + \llbracket \tau_{2,\delta}^{-1} \rrbracket\}$$

which proves $d_F(\psi_1^\delta, \psi_2^\delta) \leq \alpha_\infty(\phi_1, \phi_2)$ since $\llbracket \tau_{1,\delta}^{-1} \rrbracket + \llbracket \tau_{2,\delta}^{-1} \rrbracket \to 0$ as $\delta \to 0$. ∎

Theorem 8.14. *$\mathscr{D}[a,b]$ equipped with α_∞ is a complete separable metric space.*

Proof. We first prove that α_∞ is a metric. It is clear that α_∞ satisfies the triangle inequality, so following Lemma 8.12 it remains only to show that if $\alpha_\infty(\phi_1, \phi_2) = 0$ for some $\phi_1, \phi_2 \in \bar{\mathscr{D}}[a,b]$, then $\phi_1 \sim \phi_2$. This in turn follows from Lemma 8.13 and Remarks 8.6 and 8.2.

Next, if $\phi_n \in \mathscr{D}$ is a Cauchy sequence for α_∞, then $\psi_n^\delta \in \mathcal{D}$ are a Cauchy sequence for d_F by Lemma 8.13. By Lemma 8.3, $\mathcal{D}/\sim_F$ is complete under d_F, therefore $\lim_{n\to\infty} d_F(\psi_n^\delta, \psi) \to 0$ for some $\psi \in \mathcal{D}$. Since the second component of each ψ_n^δ is a non-decreasing surjection, the same holds for ψ. We can then readily define $\phi \in \mathscr{D}$ with δ-parametric representation ψ^δ such that $\psi \sim_F \psi^\delta$, and thus $\phi_n \to \phi$ in α_∞ due to Lemma 8.13. Hence $(\mathscr{D}, \alpha_\infty)$ is complete.

Finally, note that $(\mathcal{D}, d_F)$ is a separable metric space (this follows, e.g. from separability of the Skorokhod space). Again by Lemma 8.13, we can identify $(\mathscr{D}, \alpha_\infty)$ with a subspace of $(\mathcal{D}, d_F)$, and thus $(\mathscr{D}, \alpha_\infty)$ is separable. ∎

Corollary 8.15. *$\mathscr{D}[a,b]$ is isometric to the completion of $D[a,b]$ under the Skorokhod $\mathcal{J}_1$-metric.*

Proof. By Theorem 8.14, $(\mathscr{D}[a,b], \alpha_\infty)$ is complete. Furthermore, recall the embedding $\iota : D[a,b] \to \mathscr{D}[a,b]$ from Definition 2.5. It is easy to see that $\alpha_\infty = \sigma_\infty$ on $\iota D[a,b]$, thus $(\iota D[a,b], \alpha_\infty)$ is isometric to $(D[a,b], \sigma_\infty)$ (Remark 2.6). Finally, $\iota D[a,b]$ is dense in $\mathscr{D}[a,b]$ since $\alpha_\infty(\phi, \iota \phi^\delta \circ \lambda) \leq 2\delta$ where λ is the linear reparametrisation $\lambda : [a,b+\delta] \to [a,b]$. ∎

8.3 Differential equations driven by decorated paths

Consider $p \in [1,2)$, $\xi \in \mathbb{R}^m$, and A, B as in Section 2.3. In this section we make precise the meaning of and study the differential equation

$$dX = A(X)\,dt + B(X)\,d\phi, \quad X(a) = \xi,$$

where $\phi \in \mathscr{D}^{p\text{-var}}([a,b], \mathbb{R}^d)$ and which we now pose on a general interval $[a,b]$.

We first recall and make precise the construction of X as an element of $\mathscr{D}^{p\text{-var}}([a,b], \mathbb{R}^m)$. Recall $(\omega, \Pi) = ((\iota \mathrm{Id}, \phi), \Pi) \in \bar{\mathscr{D}}^{p\text{-var}}([a,b], \mathbb{R}^{1+d})$ from Section 2.3 where $\mathrm{Id} : [a,b] \to [a,b]$ is

the identity map. By [15, Theorem 3.6],[5] the Young ODE $dX^\delta = (A, B)(X^\delta)\, d\omega^\delta$ admits a unique solution $X^\delta \in D^{p\text{-var}}([a, b], \mathbb{R}^m)$ since $p \in [1, 2)$. We define $X \colon [a, b] \to D^{p\text{-var}}([0, 1], \mathbb{R}^m)$ by taking $X(t)$ as the linear reparametrisation of $X^\delta|_{[\tau_\delta(t^-), \tau_\delta(t)]}$, where τ_δ is associated to (ω, Π) as in Section 8.1.

The main continuity result on ODEs driven by decorated paths in the Young regime is the following.

Proposition 8.16. *The map* $\mathcal{S} \colon (\xi, \phi) \mapsto X$ *is well-defined as a map* $\mathbb{R}^m \times \mathscr{D}^{p\text{-var}}([a, b], \mathbb{R}^d) \to \mathscr{D}^{p\text{-var}}([a, b], \mathbb{R}^m)$ *and is locally Hölder continuous when the spaces* $\mathscr{D}^{p\text{-var}}$ *on both sides are equipped with* $\alpha_{p\text{-var}}$.

Proof. For ξ, ϕ as above, we first verify that $(X, \Pi) \in \mathscr{D}^{p\text{-var}}([a, b], \mathbb{R}^m)$. If $t \notin \Pi$, then t is a continuity point of τ_δ, hence $[\tau_\delta(t^-), \tau_\delta(t)]$ is a singleton and $X(t) \equiv X^\delta(\tau_\delta(t))$ is constant. Therefore Π contains all the non-stationary points of X. Furthermore $X(t)(1) = X^\delta \circ \tau_\delta(t)$ which is càdlàg since X^δ is càdlàg and τ_δ is càdlàg and increasing. The fact that $\sup_{s \in [0,1]} |X(t)(s) - X(t)(0)| > \varepsilon > 0$ for at most finitely many $t \in \Pi$ follows from $|X^\delta|_{p\text{-var}} < \infty$. Therefore $(X, \Pi) \in \mathscr{D}([a, b], \mathbb{R}^m)$ and thus $(X, \Pi) \in \mathscr{D}^{p\text{-var}}([a, b], \mathbb{R}^m)$.

Next, we recall that ODEs commute with reparametrisations: if $Y_i \in D^{p\text{-var}}([a, b], \mathbb{R}^m)$ for $i = 1, 2$ solves $dY_i = (A, B)(Y_i)\, du_i$ for $u_i \in D^{p\text{-var}}([a, b], \mathbb{R}^{1+d})$ with first component of bounded variation, and $u_1 \circ \lambda = u_2$ for some $\lambda \in \Lambda_{[a,b]}$, then $Y_1 \circ \lambda = Y_2$. It readily follows that if X' is a solution associated to (ξ, ϕ') with $\omega' = (\iota\mathrm{Id}, \phi') \sim \omega$, then $X \sim X'$. Moreover, $(\iota\mathrm{Id}, \phi') \sim (\iota\mathrm{Id}, \phi)$ if and only if $\phi' \sim \phi$, therefore $\mathcal{S}$ is well-defined.

It remains to prove that $\mathcal{S}$ is locally Hölder continuous. Consider $\xi_i, \phi_i, \omega_i, X_i$ as above for $i = 1, 2$ and take Π to contain all non-stationary points of ϕ_1, ϕ_2. Remark that $\omega_i^\delta = (g^\delta, \phi_i^\delta) \in D^{p\text{-var}}([a, b + \delta], \mathbb{R}^{1+d})$ for a non-decreasing surjection $g^\delta \colon [a, b + \delta] \to [a, b]$ with $[\![g^\delta]\!] \leq \delta$.

Recalling that $A \in C^\beta(\mathbb{R}^m, \mathbb{R}^m)$ for $\beta > 1$, take $q \in (1, \beta)$. Since ODEs commute with reparameterisations, $X_1^\delta \circ \lambda$ solves the ODE $dX = (A, B)(X)\, d(\omega_1^\delta \circ \lambda)$ for any $\lambda \in \Lambda_{[a,b+\delta]}$. Therefore, by [15, Theorem 3.6],

$$|X_1^\delta \circ \lambda - X_2^\delta|_{p\text{-var}} \leq C(|g^\delta \circ \lambda - g^\delta|_{q\text{-var}} + |\phi_1^\delta \circ \lambda - \phi_2^\delta|_{p\text{-var}} + |\xi_1 - \xi_2|),$$

where C here and below depends on ϕ_i, ξ_i only through $|\xi_1 - \xi_2| + \sum_{i=1}^2 |\phi_i|_{p\text{-var}}$. By (8.4),

$$|g^\delta \circ \lambda - g^\delta|_{q\text{-var}} \leq |g^\delta \circ \lambda - g^\delta|_{\infty\text{-var}}^{1-1/q}\, |g^\delta \circ \lambda - g^\delta|_{1\text{-var}}^{1/q}.$$

Using $|\cdot|_{\infty\text{-var}} \leq 2|\cdot|_\infty$ and $[\![g^\delta]\!] \leq \delta$, we obtain $|g^\delta \circ \lambda - g^\delta|_{\infty\text{-var}} \leq 4\delta + 2[\![\lambda]\!]$. Furthermore, $|g^\delta \circ \lambda - g^\delta|_{1\text{-var}} \leq |g^\delta \circ \lambda|_{1\text{-var}} + |g^\delta|_{1\text{-var}} \leq 2(b - a)$, and thus

$$\alpha_{p\text{-var}}(X_1, X_2) \leq C \liminf_{\delta \to 0} \min_{\lambda \in \Lambda} \max\{[\![\lambda]\!], [\![\lambda]\!]^{1-1/q} + |\phi_1^\delta \circ \lambda - \phi_2^\delta|_{p\text{-var}} + |\xi_1 - \xi_2|\}$$

$$\leq C\big(\alpha_{p\text{-var}}(\phi_1, \phi_2) + \alpha_{p\text{-var}}(\phi_1, \phi_2)^{1-1/q} + |\xi_1 - \xi_2|\big). \qquad \blacksquare$$

To get control on $\alpha_{p\text{-var}}$, there is a useful interpolation inequality.

Lemma 8.17. *Let* $g, h \colon [a, b] \to \mathbb{R}^d$ *(not necessarily càdlàg) and* $1 \leq p \leq q \leq \infty$. *Then*

$$\sigma_{q\text{-var}}(g, h) \leq \{1 + (|g|_{p\text{-var}} + |h|_{p\text{-var}})^{p/q}\}\{\sigma_{\infty\text{-var}}(g, h)^{1-p/q} + \sigma_{\infty\text{-var}}(g, h)\}.$$

---

[5]There is an assumption in [15, Theorem 3.6] that the drift has finite $p/2$-variation; we can drop this assumption here as we are working in the Young regime $p \in [1, 2)$, see also [20, Theorem 3.2 & Remark 6.1].

Proof. It follows from (8.4) that

$$\sigma_{q\text{-var}}(g, h) = \inf_{\lambda} \max\{\llbracket\lambda\rrbracket, |g \circ \lambda - h|_{q\text{-var}} + |g(a) - h(a)|\}$$

$$\leq \inf_{\lambda} \max\{\llbracket\lambda\rrbracket, |g \circ \lambda - h|_{p\text{-var}}^{p/q} \, |g \circ \lambda - h|_{\infty\text{-var}}^{1-p/q} + |g(a) - h(a)|\}$$

$$\leq \inf_{\lambda} \max\{\llbracket\lambda\rrbracket, (|g|_{p\text{-var}} + |h|_{p\text{-var}})^{p/q} \, |g \circ \lambda - h|_{\infty\text{-var}}^{1-p/q} + |g(a) - h(a)|\}$$

$$\leq \{1 + (|g|_{p\text{-var}} + |h|_{p\text{-var}})^{p/q}\}\{\sigma_{\infty\text{-var}}(g, h)^{1-p/q} + \sigma_{\infty\text{-var}}(g, h)\}. \qquad \blacksquare$$

Lemma 8.18. *For $\phi_1, \phi_2 \in \mathscr{D}^{q\text{-var}}$ and $1 \leq p \leq q \leq \infty$,*

$$\alpha_{q\text{-var}}(\phi_1, \phi_2) \leq \{1 + (|\phi_1|_{p\text{-var}} + |\phi_2|_{p\text{-var}})^{p/q}\}\{\alpha_{\infty\text{-var}}(\phi_1, \phi_2)^{1-p/q} + \alpha_{\infty\text{-var}}(\phi_1, \phi_2)\}.$$

Proof. We apply Lemma 8.17 to ϕ_1^δ and ϕ_2^δ for each $\delta > 0$ and use the definition of $\alpha_{q\text{-var}}$ and $\alpha_{\infty\text{-var}}$. $\qquad \blacksquare$

The following is a minor extension of the continuous mapping theorem.

Lemma 8.19. *Let E, F be metric spaces equipped with their Borel σ-algebras and let $H : E \to F$ be a mapping. Suppose that there exist closed sets $D_k \subset E$ for $k \geq 1$ such that $D_k \subset D_{k+1}$ and $\bigcup_{k\geq 1} D_k = E$, and that H is continuous on every D_k. Suppose that X_n are E-valued random variables such that $\lim_{k\to\infty} \sup_n \mathbb{P}[X_n \notin D_k] = 0$. Finally, suppose that $X_n \to_w X$ in E as $n \to \infty$. Then $H(X_n) \to_w H(X)$ in F as $n \to \infty$.*

Proof. Let $f : F \to \mathbb{R}$ be a lower semicontinuous function such that $|f| \leq M$. We claim that the function $H_k = \mathbf{1}_{D_k} f \circ H + \mathbf{1}_{D_k^c} M$ is lower semicontinuous on E. We suppose that $x_n \to x$ in E and we show that $\liminf_{n\to\infty} H_k(x_n) \geq H_k(x)$. The verification is straightforward if $\liminf_{n\to\infty} H_k(x_n) = M$. Suppose now $\liminf_{n\to\infty} H_k(x_n) < M$. Let x_{n_m} be a subsequence such that $\liminf_{n\to\infty} H_k(x_n) = \lim_{m\to\infty} H_k(x_{n_m})$. Note that x_{n_m} is eventually in D_k and thus $x \in D_k$ by the assumption that D_k is closed. Then by continuity of H on D_k and lower semicontinuity of f,

$$\liminf_{n\to\infty} H_k(x_n) = \lim_{m\to\infty} H_k(x_{n_m}) = \lim_{m\to\infty} f \circ H(x_{n_m}) \geq f \circ H(x) = H_k(x).$$

The claim is proved.

Next, by the Portmanteau lemma,

$$\liminf_{n\to\infty} \mathbb{E}[H_k(X_n)] \geq \mathbb{E}[H_k(X)].$$

Moreover, H_k is decreasing in k with $H_k \to f \circ H$ pointwise, and therefore

$$\lim_{k\to\infty} \mathbb{E}[H_k(X)] = \mathbb{E}[f \circ H(X)].$$

On the other hand,

$$\lim_{k\to\infty} \sup_n \big(\mathbb{E}[H_k(X_n)] - \mathbb{E}[f \circ H(X_n)]\big) = 0$$

by the assumption that $\lim_{k\to\infty} \sup_n \mathbb{P}(X_n \notin D_k) = 0$. Combining these facts, we obtain

$$\liminf_{n\to\infty} \mathbb{E}[f \circ H(X_n)] \geq \mathbb{E}[f \circ H(X)].$$

Since this holds for all bounded lower semicontinuous functions f, it follows that $H(X_n) \to_w H(X)$ in F by the Portmanteau lemma. $\qquad \blacksquare$

We are now ready to prove Theorem 2.10.

Proof of Theorem 2.10. The fact that $|\phi|_{p\text{-var}} < \infty$ follows from lower semicontinuity of p-variation and the Portmanteau lemma (more specifically, we use that $\mathbf{1}_{|\phi|_{p\text{-var}}>n}$ is lower semicontinuous). Consider now the metric spaces $E = (\mathscr{D}^{p\text{-var}}([a,b], \mathbb{R}^d), \alpha_\infty)$ and $F = (\mathscr{D}^{p\text{-var}}([a,b], \mathbb{R}^m), \alpha_{q\text{-var}})$ with $q \in (p, \gamma)$. Recall that $B \in C^\gamma(\mathbb{R}^m, \mathbb{R}^{m \times d})$ for $\gamma > p$. By stability of Young ODEs (Proposition 8.16), interpolation (Lemma 8.18), and the elementary inequality $|f(a)| + |f|_{\infty\text{-var}} \le 3|f|_\infty$ and thus $\alpha_{\infty\text{-var}} \le 3\alpha_\infty$, the solution map $\mathcal{S}$ maps every subset $D_k = \{\phi \in E : |\phi|_{p\text{-var}} \le k\}$ continuously into F. By lower semicontinuity of p-variation, each set D_k is closed in E, and, by the assumption that $|\phi_n|_{p\text{-var}}$ is tight, $\lim_{k\to\infty} \sup_n \mathbb{P}[\phi_n \notin D_k] = 0$. Finally $\phi_n \to_w \phi$ in E by assumption. It follows from Lemma 8.19 that $\mathcal{S}(\phi_n) \to_w \mathcal{S}(\phi)$ in F as $n \to \infty$ as claimed. ∎

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] J. Aaronson, M. Denker and M. Urbański. Ergodic theory for Markov fibred systems and parabolic rational maps. *Trans. Amer. Math. Soc.* **337** (1993) 495–548.

[3] D. Applebaum. *Lévy processes and stochastic calculus.* Second edition. Cambridge Stud. Adv. Math. **116** (2009).

[4] D. Burago, Y. Burago, S. Ivanov. *A course in metric geometry.* Graduate Studies in Mathematics **33**, American Mathematical Society, Providence, RI, 2001.

[5] H. Boedihardjo, X. Geng. The uniqueness of signature problem in the non-Markov setting. *Stochastic Process. Appl.* **125** (2015) 4674–4701.

[6] H. Bruin, I. Melbourne and D. Terhesiu. Sharp polynomial bounds on decay of correlations for multidimensional nonuniformly hyperbolic systems and billiards. *Annales Henri Lebesgue* **4** (2021) 407–451.

[7] A. Chechkin and I. Pavlyukevich. Marcus versus Stratonovich for systems with jump noise. *J. Phys. A* **47** (2014) 342001.

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

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

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

[11] I. Chevyrev. Random walks and Lévy processes as rough paths. *Probab. Theory Related Fields* **170** (2018) 891–932.

[12] I. Chevyrev and P. K. Friz. Canonical RDEs and general semimartingales as rough paths. *Ann. Probab.* **47** (2019) 420–463.

[13] I. Chevyrev, P. K. Friz, A. Korepanov and I. Melbourne. Superdiffusive limits for deterministic fast-slow dynamical systems. *Probab. Theory Related Fields* **178** (2020) 735–770.

[14] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang, Multiscale systems, homogenization, and rough paths, *Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016* (P. Friz et al., ed.), Springer Proceedings in Mathematics & Statistics **283** (2019) 17–42.

[15] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 1328–1350.

[16] D. Dolgopyat. Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* **356** (2004) 1637–1689.

[17] D. Dolgopyat. Averaging and invariant measures. *Mosc. Math. J.* **5** (2005) 537–576.

[18] A. C. M. Freitas, J. M. Freitas, I. Melbourne and M. Todd. Convergence to decorated Lévy processes in non-Skorohod topologies for dynamical systems. arXiv:2310.00978.

[19] A. C. M. Freitas, J. M. Freitas and M. Todd. Enriched functional limit theorems for dynamical systems. arXiv:2011.10153.

[20] P. K. Friz, H. Zhang. Differential equations driven by rough paths with jumps. *J. Differential Equations* **264** (2018) 6226–6301.

[21] G. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. Roy. Soc. London A* **469** (2013) 20130201.

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

[23] 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. *Stoch. Dyn.* **21** (2021) 2150024, 8 pages.

[24] P. Jung, F. Pène and H.-K. Zhang. Convergence to α-stable Lévy motion for chaotic billiards with several cusps at flat points. *Nonlinearity* **33** (2020) 807–839.

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

[26] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2016) 479–520.

[27] D. Kelly and I. Melbourne. Homogenization for deterministic fast-slow systems with multidimensional multiplicative noise. *J. Funct. Anal.* **272** (2017) 4063–4102.

[28] A. Korepanov, Z. Kosloff, and I. Melbourne. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 1305–1327.

[29] T.G. Kurtz, E. Pardoux and P. Protter. Stratonovich stochastic differential equations driven by general semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.* **31** (1995) 351–377.

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

[31] S. I. Marcus. Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics* **4** (1980/81) 223–245.

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

[33] I. Melbourne and P. Varandas. Convergence to a Lévy process in the Skorohod $\mathcal{M}_1$ and $\mathcal{M}_2$ topologies for nonuniformly hyperbolic systems, including billiards with cusps. *Comm. Math. Phys.* **375** (2020) 653–678.

[34] I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew product flows. *Nonlinearity* **24** (2011) 1361–1367.

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

[36] G. A. Pavliotis and A. M. Stuart. *Multiscale methods.* Texts in Applied Mathematics **53**, Springer, New York, 2008. Averaging and homogenization.

[37] F. Pène. Averaging method for differential equations perturbed by dynamical systems. *ESAIM Probab. Statist.* **6** (2002) 33–88.

[38] F. Pène and B. Saussol. Spatio-temporal Poisson processes for visits to small sets. *Israel J. Math.* **240** (2020) 625–665.

[39] G. Pisier and Q. H. Xu. The strong p-variation of martingales and orthogonal series. *Probab. Theory Related Fields* **77** (1988) 497–514.

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

[41] A. V. Skorohod. Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** (1956) 261–290.

[42] M. Tyran-Kamińska. Convergence to Lévy stable processes under some weak dependence conditions. *Stochastic Process. Appl.* **120** (2010) 1629–1650.

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

[44] 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.

[45] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.* **36** (1965) 1560–1564.

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

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