Sharp Statistical Properties for a Family of Multidimensional NonMarkovian Nonconformal Intermittent Maps.

Peyman Eslami * Ian Melbourne † Sandro Vaienti ‡

24 March 2019. Updated 13 December 2020.

Abstract

Intermittent maps of Pomeau-Manneville type are well-studied in one-dimension, and also in higher dimensions if the map happens to be Markov. In general, the nonconformality of multidimensional intermittent maps represents a challenge that up to now is only partially addressed. We show how to prove sharp polynomial bounds on decay of correlations for a class of multidimensional intermittent maps. In addition we show that the optimal results on statistical limit laws for one-dimensional intermittent maps hold also for the maps considered here. This includes the (functional) central limit theorem and local limit theorem, Berry-Esseen estimates, large deviation estimates, convergence to stable laws and Lévy processes, and infinite measure mixing.

1 Introduction

Intermittent maps were introduced by Pomeau & Manneville [56] as a model for turbulence. These are maps that are uniformly expanding except for the presence of neutral fixed points. In the smooth ergodic theory literature, they have provided the archetypal examples of nonuniformly expanding dynamical systems. For one-dimensional intermittent maps, [62] studied the invariant densities in the case when the map is Markov with respect to a suitable partition, and the nonMarkovian case was analysed in [66].

*Dipartimento di Matematica, II Università di Roma (Tor Vergata), 00133 Roma, Italy. peslami7@gmail.com)

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

‡Aix Marseille Université, Université de Toulon, CNRS, CPT, 13009 Marseille, France. vaienti@cpt.univ-mrs.fr

The paper of Liverani, Saussol & Vaienti [47] set out to study the statistical properties of one-dimensional intermittent maps by considering the simplest possible example $f : [0,1] \to [0,1]$, namely

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

Here $\gamma > 0$ is a real parameter. For $\gamma \in (0,1)$, there is a unique absolutely continuous probability measure μ. Let

$$\rho_{v,w}(n) = \int v\, w \circ f^n \, d\mu - \int v \, d\mu \int w \, d\mu. \tag{1.2}$$

By [37, 65], $\rho_{v,w}(n) = O(n^{-(\frac{1}{\gamma}-1)})$ for v Hölder and $w \in L^\infty$ and this decay rate is optimal [28, 58]. For $\gamma \in (0, \frac{1}{2})$, the central limit theorem (CLT) holds for Hölder observables by [47, 65] as does the weak invariance principle (WIP) [50]. Berry-Esseen estimates and local limit theorems were obtained in [30]. When $\gamma \in [\frac{1}{2}, 1)$, the CLT fails for Hölder observables that are nonzero at $x = 0$; stable laws were proved in this situation by [29] and the corresponding WIP holds by [55].

In addition, for $\gamma \in (0,1)$, sharp results on large deviations and convergence of moments were obtained in [21, 34, 49, 51, 53].

For $\gamma \ge 1$, there is a unique absolutely continuous invariant σ-finite measure up to scaling, but the measure is infinite. Results on mixing for infinite measure systems were obtained in [33, 52]

Although [47] initially focused on the specific maps (1.1), the results described above have by now been shown to hold for very general classes of one-dimensional intermittent maps and extend to many multi-dimensional examples in cases when the map f is Markov. In such cases, the standard approach is to construct an induced map F with infinitely many branches and to deduce quasicompactness properties of the transfer operator for F acting on a suitable function space. In the Markov case, it is natural to consider observables that are Hölder with respect to a symbolic metric; in the one-dimensional case, one can consider observables of bounded variation.

Currently, multidimensional intermittent maps are poorly understood in general. The aim of this paper is to approach the problem of multidimensional intermittent maps in the same spirit that [47] approached one-dimensional intermittent maps, focusing on some simple examples that exhibit all the problematic features: multidimensional, intermittent, nonconformal, nonMarkovian.

1.1 Statement of the main results

Let $M = [0,1] \times \mathbb{T}$ where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Our counterpart of the family (1.1) is the family of maps $f : M \to M$ with $f(x,\theta) = (f_1(x,\theta), f_2(\theta))$, where $f_1 : M \to [0,1]$ is a (not necessarily Markov) nonuniformly expanding map for each θ. Specifically, we assume that

$$f_1(x,\theta)=\begin{cases} x(1+x^\gamma u(x,\theta)), & 0\le x\le \frac34 \\ 4x-3, & \frac34<x\le 1\end{cases}, \qquad f_2(\theta)=4\theta \bmod 1, \tag{1.3}$$

where $\gamma>0$, and $u:[0,\frac34]\times\mathbb{T}\to(0,\infty)$ is a positive C^2 function satisfying $u(0,\theta)\equiv c_0>0$. (Implicitly, it is assumed that $x(1+x^\gamma u(x,\theta))\le 1$ for all $x\in[0,\frac34]$, $\theta\in\mathbb{T}$.) In addition, we assume that $|(Df)_{(x,\theta)}v|\ge |v|$ for all $(x,\theta)\in[0,\frac34]\times\mathbb{T}$, $v\in\mathbb{R}^2$.

In particular, as in [47], f is an everywhere expanding map with a neutral invariant circle $\{x=0\}$, and f is uniformly expanding on $[\delta,1]\times\mathbb{T}$ for all $\delta>0$. Also, $f_1(\frac34,\theta)>\frac34$ for $\theta\in\mathbb{T}$. For definiteness, we suppose that $f_1(\frac34,\theta)>\frac{15}{16}$ for $\theta\in\mathbb{T}$. Our final assumption is that u is sufficiently close to constant, in the sense that $|x\frac{\partial u}{\partial x}|_\infty$ and $|\frac{\partial u}{\partial\theta}|_\infty$ are sufficiently small. (See Remarks 2.6 and 4.8.)

Note that f has 8 branches; again as in [47] half of the branches are linear. However, typically the remaining branches are not full and there is no Markovian structure. Moreover, the maps expand polynomially in x and exponentially in θ and hence are highly nonconformal.

For the maps (1.3), we obtain almost identical results to the ones described above for the one-dimensional maps (1.1). Recall that $v:M\to\mathbb{R}$ is Hölder with exponent $\eta\in(0,1)$, denoted $v\in C^\eta(M)$, if $\|v\|_\eta=|v|_\infty+\sup_{x\ne y}|v(x)-v(y)|/|x-y|^\eta$ is finite. Our results are formulated mainly for Hölder observables, but occasionally for observables in $\mathrm{BV}_\infty(M)=\mathrm{BV}(M)\cap L^\infty(M)$. (The definition of bounded variation on M is recalled in Section 4.) In particular, all of the results hold for C^1 observables.

Lemma 3.4 states that for $\gamma<1$, there is a unique absolutely continuous f-invariant probability measure, denoted μ, and this measure is mixing. Our main result gives sharp polynomial upper and lower bounds on the rate of mixing. We set $\alpha=1/\gamma$ throughout. Define the correlation function $\rho_{v,w}$ as in (1.2).

Theorem 1.1 *Suppose that $\gamma<1$.*

(a) Let $\eta\in(0,1)$. There exists $C>0$ such that

$$|\rho_{v,w}(n)|\le C\|v\|_\eta|w|_\infty\, n^{-(\alpha-1)} \quad \textit{for all } n\ge 1,$$

for all $v\in C^\eta(M)$, $w\in L^\infty(M)$.

(b) Define $E(n)=\begin{cases} n^{-\alpha} & \alpha>2\\ n^{-2}\log n & \alpha=2 \\ n^{-2(\alpha-1)} & 1<\alpha<2\end{cases}$. *There exists $C>0$, $c>0$ such that*

$$\Big|\rho_{v,w}(n)-cn^{-(\alpha-1)}\int v\,d\mu\int w\,d\mu\Big|\le CE(n)(\|v\|_{\mathrm{BV}}+|v|_\infty)|w|_1 \quad \textit{for all } n\ge 1,$$

for all $v\in\mathrm{BV}_\infty(M)$, $w\in L^1(M)$ supported in $[\frac34,1]\times\mathbb{T}$. In particular, $\rho_{v,w}(n)\sim cn^{-(\alpha-1)}\int v\,d\mu\int w\,d\mu$ as $n\to\infty$.

(c) There exists $C > 0$ such that

$$|\rho_{v,w}(n)| \le C(\|v\|_{\mathrm{BV}} + |v|_\infty)|w|_1 n^{-\alpha} \quad \textit{for all } n \ge 1,$$

for all $v \in \mathrm{BV}_\infty(M)$, $w \in L^1(M)$ supported in $[\frac{3}{4}, 1] \times \mathbb{T}$ with $\int v\, d\mu = 0$.

For $\gamma < \frac{1}{2}$, corresponding to summable decay of correlations in Theorem 1.1(a), we obtain the CLT and related results. Define $v_n = \sum_{j=0}^{n-1} v \circ f^j$. Also, define $W_n(t) = n^{-1/2} v_{nt}$ for $t = 0, \frac{1}{n}, \frac{2}{n}, \dots, 1$ and linearly interpolate to obtain $W_n \in C[0, 1]$.

Theorem 1.2 *Suppose that $\gamma < \frac{1}{2}$. Let $v : M \to \mathbb{R}$ be Hölder with $\int v\, d\mu = 0$.*

(a) **CLT** *$n^{-1/2} v_n$ converges in distribution*[1] *to a normal distribution $G =_d N(0, \sigma^2)$. The variance σ^2 is zero if and only if $v = \chi \circ f - \chi$ for some χ measurable.*

(b) **Berry-Esseen** *There exists $C > 0$ such that*

$$|\mu(n^{-1/2} v_n \le a) - \mathbb{P}(G \le a)| \le C n^{-q} \quad \textit{for all } a \in \mathbb{R},\ n \ge 1,$$

where $q = \frac{1}{2}$ for $\alpha > 3$ and $q = (\alpha - 2)/2$ for $\alpha \in (2, 3)$ (Any $q < \frac{1}{2}$ works for $\alpha = 3$.)

(c) **Local limit theorem** *Suppose that v is aperiodic*[2]*. For all $a, b, \kappa \in \mathbb{R}$ with $a < b$, all $k_n \in \mathbb{R}$ with $k_n \sim \kappa n^{1/2}$, and all $u \in C^\eta(M)$, $w : M \to \mathbb{R}$ measurable,*

$$\lim_{n\to\infty} n^{1/2} \mu\big\{x \in M : v_n(x) - \kappa_n - u(x) - w(f^n x)\} \in [a, b]\big\} = (b - a)\frac{e^{-\kappa^2/(2\sigma^2)}}{(2\pi\sigma^2)^{1/2}}.$$

(d) **WIP** *W_n converges weakly in $C[0, 1]$ to Brownian motion W with $W(1) =_d G$.*

(e) **Error rate in WIP** *For any $q < (\alpha - 2/(4\alpha)$, there exists $C > 0$ such that $\pi_1(W_n, W) \le C n^{-q}$ for all $n \ge 1$.* [3]

(f) **Almost sure invariance principle** *For any $\epsilon > 0$, there is a probability space supporting W and a sequence of random variables $\{\tilde{v}_n;\ n \ge 1\}$ with the same joint distributions as $\{v_n\}$ such that $\tilde{v}_n = W(n) + O(n^\gamma (\log n)^{\gamma+\epsilon})$ a.e.*

For $\gamma \in (\frac{1}{2}, 1)$, the CLT with normalization $n^{-1/2}$ fails for general Hölder observables, and we obtain results on anomalous diffusion. Let G_α denote the totally skewed α-stable law with characteristic function $\mathbb{E}(e^{itG_\alpha}) = \exp\{-|t|^\alpha (1 - i \operatorname{sgn} t \tan \frac{\alpha\pi}{2})\}$.

---

[1] Here and elsewhere, convergence in distribution (or weak convergence) holds on the probability space (M, μ) and equivalently [67] on the probability space (M, Leb_M) where Leb_M denotes normalised Lebesgue measure on the support of μ.

[2] Aperiodic means that it is not possible to write $v \equiv \chi - \chi \circ T +$ constant mod $\lambda\mathbb{Z}$ for some χ measurable and $\lambda > 0$.

[3] Let A^ϵ denote the ϵ-neighborhood of A. The Prokhorov metric π_1 is given by
$\pi_1(X, Y) = \inf\{\epsilon > 0 : \mathbb{P}(X \in A) \le \mathbb{P}(Y \in A^\epsilon) + \epsilon$ for all closed sets $A \subset C[0, 1]\}$.

Theorem 1.3 *Let $v : M \to \mathbb{R}$ be Hölder with $\int v\,d\mu = 0$. Suppose that $\int_{\mathbb{T}} v(0,\theta)\,d\theta \neq 0$. Then there exists $c > 0$ such that $n^{-1/\alpha}v_n$ converges in distribution to cG_α.*

Moreover the process defined by $W_n(t) = n^{-1/\alpha}v_{[nt]}$ converges weakly in $D[0,1]$ with the $\mathcal{M}_1$ topology[4] *to the α-stable Lévy process W with $W(1) =_d cG_\alpha$.*

Next, we consider large deviation estimates and moment estimates.

Theorem 1.4 *Suppose that $\gamma < 1$ and let $v : M \to \mathbb{R}$ be Hölder.*

(a) **Large deviation estimates** *For any $a > 0$, there exists $C > 0$ such that*

$$\mu\Big\{\Big|\frac{1}{n}v_n - \int v\,d\mu\Big| > a\Big\} \le Cn^{-(\alpha-1)} \quad \text{for all } n \ge 1.$$

(b) **Moment estimates** *For any $p \ge 1$, there exists $C > 0$ such that for all $n \ge 1$*

$$\int |v_n|^p\,d\mu \le C\max\{g(n), n^{p-\alpha+1}\} \quad \text{where} \quad g(n) = \begin{cases} n^{p/2} & \alpha > 2 \\ (n\log n)^{p/2} & \alpha = 2 \\ n^{p/\alpha} & 1 < \alpha < 2,\ p \neq \alpha \\ n\log n & 1 < \alpha < 2,\ p = \alpha \end{cases}.$$

(c) **Convergence of moments** *If $\gamma < \frac{1}{2}$, then $\int |n^{-1/2}v_n|^p\,d\mu \to \mathbb{E}|G|^p$ for all $p < 2(\alpha-1)$.*

If $\gamma \in (\frac{1}{2}, 1)$, then $\int |n^{-1/\alpha}v_n|^p\,d\mu \to \mathbb{E}|cG_\alpha|^p$ for all $p < \alpha$ where c is the constant in Theorem 1.3.

For $\gamma \ge 1$, Lemma 3.4 states that up to scaling there is a unique absolutely continuous f-invariant σ-finite measure μ, but now $\mu(M) = \infty$. We prove the following mixing property for infinite measure systems.

Theorem 1.5 *(a) Suppose that $\gamma > 1$. There exists $c > 0$ such that*

$$\lim_{n\to\infty} n^{1-\alpha}\int v\,w\circ f^n\,d\mu = c\int v\,d\mu\int w\,d\mu,$$

for all $v \in \mathrm{BV}_\infty(M)$, $w \in L^1(M)$ supported in $[\frac{3}{4}, 1] \times \mathbb{T}$.

For $\gamma = 1$, the same result holds with $n^{1-\alpha}$ replaced by $\log n$.

(b) For $\gamma > 1$, there exists $C > 0$ such that

$$\Big|\int v\,w\circ f^n\,d\mu\Big| \le C(\|v\|_{\mathrm{BV}} + |v|_\infty)|w|_1 n^{-\alpha} \quad \text{for all } n \ge 1,$$

for all $v \in \mathrm{BV}_\infty(M)$, $w \in L^1(M)$ supported in $[\frac{3}{4}, 1] \times \mathbb{T}$ with $\int v\,d\mu = 0$.

[4] We refer to [60, 64] for background information on $D[0,1]$ and the Skorohod $\mathcal{M}_1$ topology.

Remark 1.6 The constants c in Theorems 1.1(b), 1.3 and 1.5(a) are given explicitly in Sections 5 and 6.

Remark 1.7 It is an easy but tedious exercise to extend to cases where θ is of general dimension and $f_2 : \mathbb{T}^{d-1} \to \mathbb{T}^{d-1}$ is a general smooth uniformly expanding map with worst expansion sufficiently large (strictly larger than 3 suffices when $d = 2$), but we restrict to the current situation for readability.

A notationally simpler example would have $f_2(\theta) = 2\theta \bmod 1$, but it is well-known that the extra expansion is useful in higher dimensions. Our assumption that u is sufficiently close to constant is of the same flavour and can be relaxed by assuming sufficient expansivity of f, see Remark 4.8.

1.2 Comparison with other results and methods

There is a considerable amount of work on uniformly expanding maps in higher dimensions. In the analytic setting, see [11, 63]. For C^2 maps, still with finitely many branches, see [19, 20, 26]. The paper [46] sets out a general approach to multidimensional uniformly expanding maps with infinitely many branches. This method could in principle be applied to the first return maps F mentioned in Subsection 1.3. However, the assumptions therein do not hold for the examples in [38, 39] nor the examples (1.3). (Condition 4 in [46] fails due to the lack of conformality; the condition has the form $\lim_{\epsilon\to 0} A_\epsilon = 0$ but in our examples $A_\epsilon = \infty$ for $\epsilon > 0$.) Another approach in [59] uses quasi-Hölder spaces, but these also have drawbacks as discussed below.

Turning to multidimensional intermittent maps, we mention the work of [5, 6, 7, 31] which treats examples like those in (1.3) with γ depending on θ. However, these papers require that f is Markovian and hence do not encounter the issues treated here.

In contrast, there are has been very little work on multidimensional nonMarkovian nonuniformly expanding maps. We now list all papers on this topic that we know of. A large class of multidimensional intermittent maps was considered in [38, 39] using the quasi-Hölder spaces from [59]. In particular, [38] obtained results on existence of absolutely continuous invariant measures, but it was convenient to consider maps that were close to conformal. For statistical limit laws, it seems that quasi-Hölder spaces handle nonconformality of multidimensional maps quite poorly. A more recent paper [10] obtains almost optimal, but still nonoptimal, results on decay of correlations for the maps in [38, 39]. Moreover, the methods in [10] do not seem to apply to the maps (1.3) considered here.

1.3 Structure of the paper

The method in this paper starts off, as usual, by constructing a convenient first return map $F : Y \to Y$, and from then on is a hybrid of two standard methods.

Reinducing enables us to model f by a Young tower with polynomial tails, leading to existence of absolutely continuous invariant measures and a spectral decomposition. For $\gamma \in (0, \frac{1}{2})$, this already yields sharp upper bounds on decay of correlations as well as a number of statistical limit laws. Combining the information on invariant measures with bounded variation methods for F, we obtain sharp lower bounds on decay of correlations, as well as convergence to stable laws and Lévy processes, and results on infinite measure mixing. This hybrid method bypasses many of the problems associated with multi-dimensional bounded variation (namely, that the function space is not contained in L^∞; supports of invariant densities need not *a priori* have nonempty interior; certain aperiodicity assumptions are hard to verify).

The reinducing step, Lemma 3.1 below, makes use of recent work [23] based on the method of standard pairs [12, 22], and gives precise joint control on the first return time to Y and the reinducing return time (denoted respectively as φ and ρ below). As already noted, the reinducing approach adopted in [10] seems not applicable for the examples in this paper and in any case gives much less control on return times.

The remainder of this paper is organised as follows. In Section 2, we construct a convenient first return map F and obtain estimates for the first return time and distortion bounds for F. In Section 3, we derive mixing properties of f and F and results on aperiodicity. In the process of doing this, we show that f can be modelled by a Young tower with polynomial decay of correlations. We use this to prove Theorems 1.1(a), 1.2 and 1.4.

Section 4 contains functional analytic estimates in bounded variation. In Section 5, we prove Theorems 1.1(b,c) and 1.5. Finally, we prove Theorem 1.3 in Section 6.

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

We set $\overline{\mathbb{D}} = \{\omega \in \mathbb{C} : |\omega| \le 1\}$. Throughout, $|\cdot|$ denotes Euclidean distance.

2 Estimates for the first return map F

2.1 Construction of F

Let $f : M \to M$, $M = [0,1] \times \mathbb{T}$, belong to the class of maps (1.3).

Define

$$X_i = \{(x,\theta) \in M : 0 \le x \le f_1(\tfrac{3}{4}, \tfrac{i+\theta}{4})\} = f([0, \tfrac{3}{4}] \times [\tfrac{i}{4}, \tfrac{i+1}{4}]), \quad i = 0,1,2,3.$$

Then $X = \bigcup_{i=0}^{3} X_i$ is an invariant set for f and $f(X) = X$.

We induce on the set $Y = ([\frac{3}{4}, 1] \times \mathbb{T}) \cap X$. Let $\varphi : Y \to \mathbb{Z}^+$ be the first return time, with first return map $F = f^\varphi : Y \to Y$. The sets

$$Y_{n,j} = \{(y,\theta) \in Y : \varphi(y,\theta) = n : (j-1)/4^n < \theta < j/4^n\}, \quad n \ge 1,\, 1 \le j \le 4^n,$$

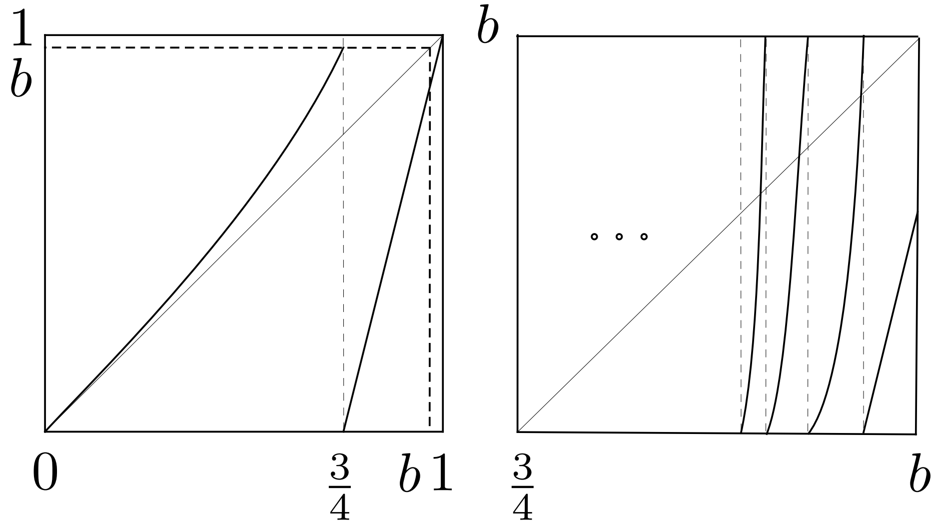

Figure 1: The maps f and F for fixed θ. The letter b denotes $f_1(\frac{3}{4}, \theta)$

form a (mod 0) partition α^Y of Y. Note that $F : a \to Fa$ is a diffeomorphism for each $a \in \alpha^Y$. We have

$$Y_{1,j} = \{(y, \theta) \in Y : \tfrac{15}{16} < y < f_1(\tfrac{3}{4}),\ (j-1)/4 < \theta < j/4\}. \tag{2.1}$$

Also, $FY_{n,j} \in \{([\frac{3}{4}, 1] \times \mathbb{T}) \cap X_i\}_{i=0}^{3}$ for $n \geq 2$. In particular, F has finitely many images, i.e. $\{Fa : a \in \alpha^Y\}$ is finite.

Figure 1 is a sketch of $f(\cdot, \theta)$ and $F(\cdot, \theta)$ for θ fixed. Figure 2 is a schematic picture of the partition $\alpha^Y = \{Y_{n,j} : n \geq 1,\ 1 \leq j \leq 4^n\}$.

Proposition 2.1 *$|(DF)_{(y,\theta)}v| \geq 4|v|$ for all $(y, \theta) \in Y$, $v \in \mathbb{R}^2$.*

Proof We have $Df = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$ on Y and $|(Df)v| \geq |v|$ on X. ∎

Proposition 2.2 *$f : X \to X$ is topologically exact: for any nonempty open subset $U \subset X$, there exists $n \geq 0$ such that $f^n U \supset X \setminus \{x = 0\}$.*

Proof It suffices to consider rectangles $U = U_1 \times U_2$ where $U_1 \subset [0, 1]$, $U_2 \subset \mathbb{T}$ are intervals. Let $\pi : M \to [0, 1]$ be projection onto the first coordinate. Note that

$$\text{For any } (x, \theta) \in (0, \tfrac{3}{4}) \times \mathbb{T}, \text{ there exists } n \geq 1 \text{ such that } \pi f^n(x, \theta) > \tfrac{3}{4}. \tag{2.2}$$

If $\frac{3}{4} \in \pi U$, then $0 \in \overline{\pi f U}$. Since 0 is a fixed point and f_1 is continuous on $[0, \frac{3}{4}] \times \mathbb{T}$, it follows from (2.2) that $(0, \frac{3}{4}] \subset \pi f^n U$ for all n sufficiently large. Also $f_2 : \mathbb{T} \to \mathbb{T}$ is continuous and uniformly expanding, so it is immediate that $(0, \frac{3}{4}] \times \mathbb{T} \subset f^n U$ for some n and hence that $f^{n+1} U \supset X \setminus \{x = 0\}$.

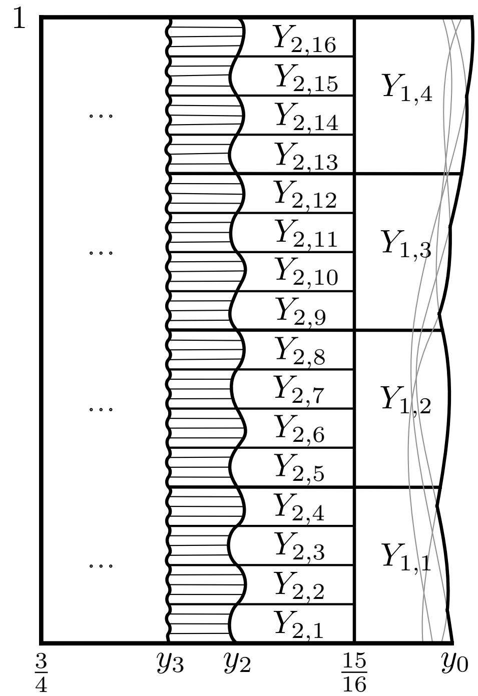

Figure 2: The partition $\alpha^Y = \{Y_{n,j},\, n \geq 1,\, 1 \leq j \leq 4^n\}$

For a general rectangle U, it remains to show that $\frac{3}{4} \in \pi f^n U$ for some $n \geq 0$. Suppose this is not the case. Since f_1 is continuous on $[0, \frac{3}{4}) \times \mathbb{T}$ and $(\frac{3}{4}, 1] \times \mathbb{T}$, it follows that $\pi f^n U$ is an interval for all n. By (2.2), $\pi f^n U \subset (\frac{3}{4}, 1]$ infinitely often. On each such occasion, $\operatorname{diam} \pi f^{n+1} U = 4 \operatorname{diam} \pi f^n U$, so $\operatorname{diam} \pi f^n U \to \infty$ which is impossible. ∎

Proposition 2.3 *$F : Y \to Y$ is topologically exact.*

Proof Let $U \subset Y$ be a nonempty open rectangle. If U intersects the boundary of the strip $\{\varphi = n\}$ for some n, then $\overline{FU} \cap (\{\frac{3}{4}\} \times \mathbb{T}) \neq \emptyset$. But then FU contains a partition element $Y_{n,j}$. It follows that $F^2 U \supset FY_{n,j} = Y \cap X_i$ for some i and hence that $F^3 U = Y$.

Again, let π denote projection onto the first coordinate. If $F^k U$ does not intersect the boundary of $\{\varphi = n\}$ for all n and all $k \geq 0$, then $\operatorname{diam} \pi F^k U \geq 4^k \operatorname{diam} \pi U$ for all k, which is impossible. ∎

2.2 Estimates for the partition

Recall that $u : [0, \frac{3}{4}] \times \mathbb{T} \to (0, \infty)$ is C^2 and $u(0, \theta) \equiv c_0 > 0$. In fact, we only use the following consequences of this property:

$$u(x, \theta) - c_0 = o(1), \quad x\frac{\partial u}{\partial x}(x, \theta) = o(1), \quad \text{and} \quad x^2\frac{\partial^2 u}{\partial x^2}(x, \theta) = o(1),$$

as $x \to 0$ uniformly in θ.

Proposition 2.4 *Suppose that (x_n, θ_n), $n \geq 1$, is a sequence in $[0, \frac{3}{4}] \times \mathbb{T}$ such that $f^n(x_n, \theta_n) = (\frac{3}{4}, \theta)$ where $\theta \in \mathbb{T}$. Then*

$$x_n \sim c_1 n^{-\alpha} \ \text{ and } \ x_n - x_{n+1} \sim c' n^{-(1+\alpha)} \ \text{ as } n \to \infty,$$

uniformly in θ, where $c_1 = (c_0\gamma)^{-\alpha}$ and $c' = c_1^{1+\gamma}c_0$.

In addition, the curve $\theta \to x_n(\theta)$ is C^1 and there exists a constant $C > 0$ independent of n, θ such that $|x_n'(\theta)| \leq Cn^{-(1+\alpha)}$.

Proof By construction, for each choice of inverse sequence θ_n, the sequence x_n is unique and monotonically decreasing to zero. We do the computation for $\theta_n = \theta/4^n$, but the result is independent of this choice. Write $\theta_0 = \theta$.

The inverse branch $\psi : M \to [0, \frac{3}{4}] \times [0, \frac{1}{4}]$ has the form $\psi(x, \theta) = (\psi_1(x, \theta), \frac{1}{4}\theta)$. Compute that

$$\begin{aligned}\psi_1(x, \theta) &= x(1 - x^\gamma \tilde{u}(x, \theta)) = [x^{-\gamma}(1 - x^\gamma \tilde{u}(x, \theta))^{-\gamma}]^{-\alpha} \\ &= [x^{-\gamma} + \gamma \hat{u}(x, \theta)]^{-\alpha},\end{aligned}$$

where $\hat{u}(x, \theta) = c_0 + o(1)$ as $x \to 0$ uniformly in θ. Inductively,

$$[\psi^n]_1(x, \theta) = \Big[x^{-\gamma} + \gamma \sum_{j=0}^{n-1} \hat{u}(\psi^j(x, \theta))\Big]^{-\alpha}.$$

In particular,

$$x_n = [\psi^n]_1(\tfrac{3}{4}, \theta_0) = \Big[(\tfrac{4}{3})^\gamma + \gamma \sum_{j=0}^{n-1} \hat{u}(x_j, \theta_j)\Big]^{-\alpha} = \Big[\gamma \sum_{j=0}^{n-1} \hat{u}(x_j, \theta_j) + O(1)\Big]^{-\alpha}. \tag{2.3}$$

Since $x_n \to 0$, we have $\hat{u}(x_j, \theta_j) \to c_0$ and hence $\sum_{j=0}^{n-1} \hat{u}(\psi^j(x, \theta)) = nc_0 + o(n)$. Substituting this into (2.3) yields the desired expression for x_n.

Since $(x_n, \theta_n) = f(x_{n+1}, \theta_{n+1})$, we have $x_n = x_{n+1}(1 + x_{n+1}^\gamma u(x_{n+1}, \theta_{n+1}))$ and so

$$x_n - x_{n+1} = x_{n+1}^{1+\gamma} u(x_{n+1}, \theta_{n+1}) \sim c_1^{1+\gamma} n^{-(1+\alpha)} c_0 = c' n^{-(1+\alpha)}.$$

Differentiating the formula for x_n in (2.3),

$$x_n'(\theta) = -\Big[(\tfrac{4}{3})^\gamma + \gamma \sum_{j=0}^{n-1} \hat{u}(x_j,\theta_j)\Big]^{-(1+\alpha)} \Big(\sum_{j=0}^{n-1} \frac{\partial \hat{u}}{\partial x}(x_j,\theta_j)x_j'(\theta) + \sum_{j=0}^{n-1} \frac{\partial \hat{u}}{\partial \theta}(x_j,\theta_j)4^{-j}\Big).$$

It is easy to verify that $\hat{u}$ inherits the properties $x\dfrac{\partial \hat{u}}{\partial x} = o(1)$ and $\dfrac{\partial \hat{u}}{\partial \theta} = O(1)$ imposed on u. Hence, there is a constant $K > 0$ such that

$$|x_n'(\theta)| \le Kn^{-(1+\alpha)} + Kn^{-(1+\alpha)} \sum_{j=0}^{n-1} j^{1/\gamma}|x_j'(\theta)| \le Kn^{-(1+\alpha)} + Kn^{-1} \sum_{j=0}^{n-1} |x_j'(\theta)|.$$

The discrete version of Gronwall's inequality states that if $|b_n| \le C_1 + C_2 \sum_{j=0}^{n-1} |b_j|$, then $|b_n| \le C_1(1+C_2)^n$. Hence $|x_n'(\theta)| \le Kn^{-(1+\alpha)}(1+Kn^{-1})^n \ll n^{-(1+\alpha)}$.

For each $n \ge 1$, we have established smoothness of the curve $x_n(\theta)$ and the estimate $|x_n'(\theta)| \le Cn^{-(1+\alpha)}$ except at finitely many points related to the partition into inverse branches of f^n. Since $x_n(\theta) = \{(x,\theta) \in [0,\frac{3}{4}] \times \mathbb{T}^1 : f^n(x,\theta) \in \{\frac{3}{4}\} \times \mathbb{T}\}$ is defined intrinsically on the cylindrical domain $[0,\frac{3}{4}] \times \mathbb{T}^1$, independent of any choice of partition, these smoothness properties are uniform in $\theta \in \mathbb{T}$. ∎

Remark 2.5 It follows from Proposition 2.4 that if $\theta, \theta' \in [(j-1)4^{-n}, j4^{-n}]$ for some $j = 1,\dots,4^n$, then $x_n(\theta) - x_{n+1}(\theta') \sim c'n^{-(1+\alpha)}$ uniformly in j.

The partition elements $Y_{1,j}$, $1 \le j \le 4$, are as in (2.1). The remaining partition elements $Y_{n,j}$, $n \ge 2$, $1 \le j \le 4^n$, are given by

$$Y_{n,j} = \{(y,\theta) : y \in (y_n(\theta), y_{n-1}(\theta)),\, \theta \in ((j-1)/4^n, j/4^n)\},$$

where $y_n(\theta) = \frac{1}{4}(x_{n-1}(f_2(\theta)) + 3)$. Note that $y_n'(\theta) = x_{n-1}'(\theta)$.

By Remark 2.5, $Y_{n,j}$ is almost rectangular for n large, and $y_n(\theta) - y_{n+1}(\theta) \sim \frac{1}{4}c'n^{-(1+\alpha)}$ uniformly in θ.

Remark 2.6 By choosing u sufficiently close to constant, in the sense that $|x\frac{\partial u}{\partial x}|_\infty$ and $|\frac{\partial u}{\partial \theta}|_\infty$ are sufficiently small, we can arrange that $|x_n'(\theta)|$ is uniformly small in n and θ. In fact, we require that $|x_n'(\theta)| < 7/\sqrt{72}$ for all n and θ. This turns out to be convenient for technical reasons, see Remark 4.8.

Corollary 2.7 *$\mathrm{Leb}(\varphi > n) \sim \frac{1}{4}c_1 n^{-\alpha}$ as $n \to \infty$. In particular, $\varphi \in L^1$ if and only if $\gamma < 1$. In addition, $\mathrm{Leb}(\varphi = n) \sim \frac{1}{4}c'(n^{-(1+\alpha)})$.*

Proof Let $C_n = \{x_n(\theta)\}$, $n \ge 0$, be the smooth curve defined in Proposition 2.4 and let X_{n+1} be the region in $[0,\frac{3}{4}] \times \mathbb{T}$ to the left of C_n. Then X_{n+1} consists of precisely those points in $[0,\frac{3}{4}] \times \mathbb{T}$ that require at least $n+1$ iterates of f to enter Y for the first time. By Proposition 2.4, $\mathrm{Leb}(X_n) \sim c_1 n^{-\alpha}$ and $\mathrm{Leb}(X_{n-1} \setminus X_n) \sim c'n^{-(1+\alpha)}$.

Now, f maps $\{\varphi = n\}$ onto $X_{n-1} \setminus X_n$ for $n \geq 2$, and the mapping is 4 to 1. Each branch is linear and scales areas by a factor of 4^2, so $\mathrm{Leb}(\{\varphi = n\}) = \frac{1}{4}\,\mathrm{Leb}(X_{n-1} \setminus X_n)$. Hence $\mathrm{Leb}(\varphi = n) \sim \frac{1}{4} c' n^{-(1+\alpha)}$ and $\mathrm{Leb}(\varphi > n) = \frac{1}{4}\,\mathrm{Leb}(X_n) \sim \frac{1}{4} c_1 n^{-\alpha}$. ∎

Remark 2.8 Suppose further that $u(x,\theta) = c_0 + O(x^\gamma)$ uniformly in θ. This property is inherited by $\tilde{u}$ and $\hat{u}$ in the above calculations and we obtain that $x_n = c_1 n^{-\alpha}(1 + O(n^{-1}\log n))$ uniformly in θ.

2.3 Distortion estimates

Lemma 2.9 *Let $(y,\theta) \in Y_{n,j}$. Then*

$$\prod_{k=\ell}^{m} \frac{\partial f_1}{\partial x}(f^k(y,\theta)) \sim \Big(\frac{n-\ell}{n-m}\Big)^{1+\alpha} \quad \textit{as } n \to \infty.$$

This estimate holds uniformly in $(y,\theta) \in Y_{n,j}$, in j, and in $1 \leq \ell \leq m \leq n-1$.

Proof Let $(y_k, \theta_k) = f^k(y,\theta)$, $k = 0, \dots, n-1$ and recall from Proposition 2.4 that $y_{n-k} \sim (c_0\gamma k)^{-\alpha}$ as $k \to \infty$ uniformly in the initial choice of θ. Now,

$$\frac{\partial f_1}{\partial x}(x,\theta) = 1 + (1+\gamma)x^\gamma u(x,\theta) + x^{1+\gamma}\frac{\partial u}{\partial x}(x,\theta) = 1 + c_0(1+\gamma)x^\gamma(1+o(1)), \quad (2.4)$$

as $x \to 0$ uniformly in θ, so

$$\log \frac{\partial f_1}{\partial x}(f^{n-k}(y,\theta)) \sim c_0(1+\gamma)y_{n-k}^\gamma \sim (1+\alpha)k^{-1}.$$

It follows that

$$\log \prod_{k=\ell}^{m} \frac{\partial f_1}{\partial x}(f^k(y,\theta)) = \sum_{k=n-m}^{n-\ell} \log \frac{\partial f_1}{\partial x}(f^{n-k}(y,\theta)) = (1+\alpha)\sum_{k=n-m}^{n-\ell} b_k(\theta),$$

where $b_k(\theta) \sim k^{-1}$ as $k \to \infty$ uniformly in θ. Hence $\log \prod_{k=\ell}^{m} \frac{\partial f_1}{\partial x}(f^k(y,\theta)) \sim (1+\alpha)(\log(n-\ell) - \log(n-m))$ and the result follows. ∎

Corollary 2.10 *Let $(y,\theta) \in Y_{n,j}$. Then*

$$(DF)_{(y,\theta)} = \begin{pmatrix} A(y,\theta) & B(y,\theta) \\ 0 & 4^n \end{pmatrix},$$

where $A(y,\theta) \sim 4n^{1+\alpha}$ as $n \to \infty$ and $B(y,\theta) = O(4^n)$. These estimates hold uniformly in $(y,\theta) \in Y_{n,j}$ and in j.

Proof Let $z_k = f^k(y,\theta)$, $k = 0,\dots,n-1$ and write

$$(DF)_{(y,\theta)} = (Df)_{z_{n-1}} \cdots (Df)_{z_1}(Df)_{z_0}.$$

Clearly, $(Df)_{z_k}$ is upper triangular with diagonal entries $\frac{\partial f_1}{\partial x}(z_k)$ and 4. Hence $(DF)_{(y,\theta)}$ has the required form with

$$A(y,\theta) = \prod_{k=0}^{n-1} \frac{\partial f_1}{\partial x}(z_k) = 4\prod_{k=1}^{n-1} \frac{\partial f_1}{\partial x}(z_k)$$

The required asymptotics for $A(y,\theta)$ follows from Lemma 2.9.

Next,

$$\frac{\partial [f^{k+1}]_1}{\partial\theta}(z_0) = \frac{\partial f_1}{\partial x}(z_k)\frac{[\partial f^k]_1}{\partial\theta}(z_0) + \frac{\partial f_1}{\partial\theta}(z_k)4^k,$$

so by induction,

$$\frac{\partial [f^k]_1}{\partial\theta}(y,\theta) = \sum_{\ell=0}^{k-1}\Big[\prod_{m=\ell+1}^{k-1}\frac{\partial f_1}{\partial x}(z_m)\Big]\frac{\partial f_1}{\partial\theta}(z_\ell)4^\ell.$$

By Lemma 2.9, $\prod_{m=\ell}^{k-1}\frac{\partial f_1}{\partial x}(z_m) \ll [(n-\ell)/(n-k)]^{1+\alpha}$ and so

$$\begin{aligned}\frac{\partial [f^k]_1}{\partial\theta}(y,\theta) &\ll (n-k)^{-(1+\alpha)}\sum_{\ell=1}^{k}(n-\ell)^{1+\alpha}4^\ell \\ &\ll (n-k)^{-(1+\alpha)}\sum_{\ell=n-k}^{n}\ell^{1+\alpha}4^{n-\ell} \ll (n-k)^{-(1+\alpha)}4^n,\end{aligned}$$

yielding the required estimate for $B(y,\theta)$. ∎

Let $JF = \det DF$. By Corollary 2.10, $JF \sim 4^{n+1}n^{1+\alpha}$ on $Y_{n,j}$.

Lemma 2.11 *There is a constant $C > 0$ such that*

$$\Big|\frac{\partial}{\partial y}(JF)^{-1}_{(y,\theta)}\Big| \le C4^{-n}, \quad \Big|\frac{\partial}{\partial\theta}(JF)^{-1}_{(y,\theta)}\Big| \le Cn^{-(1+\alpha)},$$

for all $(y,\theta)\in Y_{n,j}$ and all n, j.

Proof By Corollary 2.10, $JF(y,\theta) = 4^nA(y,\theta)$ where $A(y,\theta)\sim 4n^{1+\alpha}$ uniformly on $Y_{n,j}$. We have

$$\begin{aligned}\frac{\partial}{\partial y}(JF(y,\theta))^{-1} &= 4^{-n}(A(y,\theta))^{-2}\frac{\partial}{\partial y}A(y,\theta) \\ &= 4^{-n}A(y,\theta)^{-1}\frac{\partial}{\partial y}\log A(y,\theta) \sim 4^{-(n+1)}n^{-(1+\alpha)}\frac{\partial}{\partial y}\log A(y,\theta).\end{aligned}$$

Similarly, $\frac{\partial}{\partial\theta}(JF(y,\theta))^{-1} \sim 4^{-(n+1)}n^{-(1+\alpha)}\frac{\partial}{\partial\theta}\log A(y,\theta)$. Hence, it suffices to show that

$$\frac{\partial}{\partial y}\log A(y,\theta) \ll n^{1+\alpha}, \quad \frac{\partial}{\partial\theta}\log A(y,\theta) \ll 4^n. \tag{2.5}$$

Writing $z_k = f^k(y,\theta)$,

$$\frac{\partial}{\partial y}\log A(y,\theta) = \sum_{k=1}^{n-1}\frac{\partial}{\partial y}\log\frac{\partial f_1}{\partial x}(z_k) = \sum_{k=1}^{n-1}\Big(\frac{\partial f_1}{\partial x}(z_k)\Big)^{-1}\frac{\partial^2 f_1}{\partial x^2}(z_k)\frac{\partial[f^k]_1}{\partial y}(y,\theta).$$

By (2.4) and Proposition 2.4, $\frac{\partial f_1}{\partial x}(z_k) \sim 1$ and $\frac{\partial^2 f_1}{\partial x^2}(z_k) \ll z_k^{\gamma-1} \ll (n-k)^{-(1-\alpha)}$. Moreover, it follows from Lemma 2.9 that

$$\frac{\partial[f^k]_1}{\partial y}(y,\theta) \ll (n/(n-k))^{1+\alpha}.$$

Hence

$$\frac{\partial}{\partial y}\log A(y,\theta) \ll \sum_{k=1}^{n-1}(n-k)^{-(1-\alpha)}n^{1+\alpha}(n-k)^{-(1+\alpha)} = n^{1+\alpha}\sum_{k=1}^{n}k^{-2} \ll n^{1+\alpha},$$

establishing the first estimate in (2.5).

Proceeding similarly for the second estimate

$$\frac{\partial}{\partial\theta}\log A(y,\theta) = \sum_{k=1}^{n-1}\Big(\frac{\partial f_1}{\partial x}(z_k)\Big)^{-1}\Big[\frac{\partial^2 f_1}{\partial x^2}(z_k)\frac{\partial[f^k]_1}{\partial\theta}(y,\theta) + \frac{\partial^2 f_1}{\partial x\partial\theta}(z_k)\frac{\partial[f^k]_2}{\partial\theta}(y,\theta)\Big]. \tag{2.6}$$

Our assumptions on f_1 imply in particular that $\frac{\partial^2 f_1}{\partial x\partial\theta}$ is bounded, so the second term in (2.6) is $O(4^n)$. The calculation at the end of the proof of Corollary 2.10 shows that $\frac{\partial[f^k]_1}{\partial\theta}(y,\theta) \ll (n-k)^{-(1+\alpha)}4^n$. Hence the first term in (2.6) is bounded up to a constant by

$$\sum_{k=1}^{n-1}(n-k)^{-(1-\alpha)}(n-k)^{-(1+\alpha)}4^n = 4^n\sum_{k=1}^{n-1}k^{-2} \ll 4^n,$$

establishing the second estimate in (2.5). ∎

Corollary 2.12 *There is a constant $C > 0$ such that*

$$|1/JF(y,\theta) - 1/JF(y',\theta')| \le C\inf_a(1/JF)|F(y,\theta) - F(y',\theta')|,$$

for all (y,θ), $(y',\theta') \in a = Y_{n,j}$ and all n, j.

Proof First, we prove the result under the simplifying assumption that a is a rectangle. In particular, the line segments $[(y,\theta),(y',\theta)]$ and $[(y',\theta),(y',\theta')]$ lie in a. By Lemma 2.11 and the mean value theorem, $|1/JF(y,\theta)-1/JF(y',\theta)| \ll 4^{-n}|y-y'|$. By Corollary 2.10, $4^{-n}|y-y'| \ll 4^{-n}n^{-(1+\alpha)}|F(y,\theta)-F(y',\theta)| \ll \inf_a(1/JF)|F(y,\theta)-F(y',\theta)|$. Hence $|1/JF(y,\theta)-1/JF(y',\theta)| \ll \inf_a(1/JF)|F(y,\theta)-F(y',\theta)|$. Similarly, $|1/JF(y',\theta)-1/JF(y',\theta')| \ll \inf_a(1/JF)|F(y',\theta)-F(y',\theta')|$. The desired estimate follows.

In general, Proposition 2.4 ensures that there is a constant $c_2>0$ such that the line segments lie in the union of partition elements $Y_{m,j}$ with $m \ge c_2 n$, and the argument above is unaffected. ∎

Let α_k^Y denote the refinement of α^Y into k-cylinders.

Corollary 2.13 *There exists a constant $C>0$ such that $\sup_a JF^k \le C\inf_a JF^k$ for all $a\in\alpha_k^Y$, $k\ge 1$.*

Proof Write $x=(y,\theta)$, $x'=(y',\theta')$. First suppose that $x,x'\in a$, $a\in\alpha^Y$. By Corollary 2.12, there is a constant $C_1>0$ such that

$$\frac{JF(x)}{JF(x')} = 1+\frac{1/JF(x')-1/JF(x)}{1/JF(x)} \le 1+\frac{|1/JF(x)-1/JF(x')|}{\inf_a 1/JF}$$
$$\le 1+C_1|Fx-Fx'| \le e^{C_1|Fx-Fx'|}.$$

Hence $|\log JF(x)-\log JF(x')| \le C_1|Fx-Fx'|$.

Now suppose that $x,x'\in a$, $a\in\alpha_k^Y$ for some $k\ge 1$. Then

$$|\log JF^k(x) - \log JF^k(x')| \le \sum_{j=0}^{k-1}|\log JF(F^jx)-\log JF(F^jx')|$$
$$\le C_1\sum_{j=0}^{k-1}|F^{j+1}x-F^{j+1}x'| \le C_1\sum_{j=0}^{k-1}4^{-j}|F^kx-F^kx'| \le \frac{4}{3}C_1 \operatorname{diam} Y.$$

The result follows with $C=e^{\frac{4}{3}C_1\operatorname{diam} Y}$. ∎

Corollary 2.14 (Bounded distortion) *There is a constant $C>0$ such that*

$$\sup_a |\nabla(JF)^{-1})(DF)^{-1}|JF \le C \quad \text{for all } a\in\alpha^Y.$$

Proof Let $a=Y_{n,j}$. By Corollary 2.10,

$$(DF)^{-1} = 4^{-n}A^{-1}\begin{pmatrix}4^n & -B\\ 0 & A\end{pmatrix} \ll \begin{pmatrix} n^{-(1+\alpha)} & n^{-(1+\alpha)} \\ 0 & 4^{-n}\end{pmatrix} \tag{2.7}$$

uniformly on a. By Lemma 2.11,

$$|(\nabla(JF)^{-1})(DF^{-1})| \ll \begin{pmatrix}4^{-n} & n^{-(1+\alpha)}\end{pmatrix}\begin{pmatrix}n^{-(1+\alpha)} & n^{-(1+\alpha)} \\ 0 & 4^{-n}\end{pmatrix} \ll 4^{-n}n^{-(1+\alpha)}$$

on a. Finally, apply Corollary 2.10. ∎

3 Mixing properties of f and F

In this section, we show that the first return map $F : Y \to Y$ has a unique absolutely continuous invariant probability measure μ_Y and that F is mixing. We also show that the underlying map $f : X \to X$ has a unique (up to scaling) absolutely continuous invariant σ-finite measure μ. When $\gamma < 1$, this is a finite measure and it is mixing. These results are obtained in Subsection 3.1. In the process of obtaining these results we show that f is modelled by a Young tower with polynomial tails. A logarithmic factor in this tail rate is removed in Subsection 3.2. This is already sufficient to obtain many of the results announced in the Introduction, as explained in Subsection 3.3. In Subsection 3.4, we obtain an aperiodicity property for F.

Recall that the first return map $F : Y \to Y$ is topologically mixing with finite images, and has bounded distortion. If in addition F were Markov, then the results in this section would be easier to deduce from standard results. Our strategy is to further induce F, with exponential tails, to a full-branched Gibbs-Markov map $G : Z \to Z$ as follows:

Lemma 3.1 *There exists a refinement α_1^Y of the partition α^Y for $F : Y \to Y$, an open set $Z \subset Y$ consisting of a union of elements of α_1^Y and a map $\rho : Z \to \mathbb{Z}^+$ constant on elements of $\alpha^Z = \{a \in \alpha_1^Y : a \subset Z\}$ such that*

(a) $G = F^\rho : Z \to Z$ *is a full-branched Gibbs-Markov map with partition* α^Z.

(b) $\mathrm{Leb}(\rho > k) = O(\delta^k)$ *for some* $\delta \in (0, 1)$.

(c) $\gcd\{n \geq 1 : \{\varphi = n,\, \rho = 1\} \neq \emptyset\} = 1$.

(d) There exists $n \geq 1$ *such that* $F\big(Z \cap \mathrm{Int}\{\varphi = n\}\big) = Y$.

We postpone the proof of Lemma 3.1 to Appendix A. Parts (a) and (b) can be proven in the general setting of piecewise expanding maps, but parts (c) and (d) are specific to our map F. Part (c) is used to prove that F and f are mixing in Lemmas 3.2 and 3.4 respectively. Part (d) is used in the proof of Lemma 3.2 to prove that the invariant density for F is bounded below.

By [1, Theorem 4.7.4], there exists a unique absolutely continuous G-invariant probability measure μ_Z on Z. Moreover, μ_Z is mixing and the density $h_Z = d\mu_Z/d\,\mathrm{Leb}$ is bounded above and below on Z.

3.1 Densities and mixing

In this subsection, we study the mixing properties of f and F, the existence and uniqueness of absolutely continuous invariant measures, and the boundedness properties of the corresponding densities.

Lemma 3.2 *There exists a unique absolutely continuous F-invariant probability measure μ_Y. The density $h_Y = d\mu_Y/d\,\mathrm{Leb}$ is bounded above and below and F is mixing.*

Proof Let $G = F^\rho : Z \to Z$ be the full-branched Gibbs-Markov map on $Z \subset Y$ obtained in Lemma 3.1, with ergodic invariant probability measure μ_Z. By Lemma 3.1(b), $\bar\rho = \int_Z \rho\, d\mu_Z < \infty$.

Form the Young tower $g : \Delta \to \Delta$ where

$$\Delta = \{(z,\ell) \in Z \times \mathbb{Z} : 0 \le \ell \le \rho(z) - 1\}, \qquad g(z,\ell) = \begin{cases} (z,\ell+1) & \ell \le \rho(z) - 2 \\ (Gz, 0) & \ell = \rho(z) - 1 \end{cases}.$$

The measure $\mu_\Delta = (\mu_Z \times \text{counting})/\bar\rho$ is an ergodic g-invariant probability measure on Δ. The projection $\pi : \Delta \to Y$, $\pi(z,\ell) = F^\ell z$ defines a semiconjugacy between g and F, and $\mu_Y = \pi_*\mu_\Delta$ is an absolutely continuous F-invariant probability measure on Y. Since G is full-branch and $\gcd(\rho(a) : a \in \alpha^Z) = 1$ by Lemma 3.1(c), it follows from [65, Theorem 1] that μ_Y is mixing.

Next, for $E \subset Z$ measurable,

$$\begin{aligned} \mu_Y(E) = \mu_\Delta(\pi^{-1}E) &= (1/\bar\rho) \int_Z \sum_{\ell=0}^{\rho-1} 1_E \circ F^\ell \, d\mu_Z \\ &\ge (1/\bar\rho) \int_Z 1_E \, d\mu_Z = (1/\bar\rho)\mu_Z(E). \end{aligned} \tag{3.1}$$

It follows that $h_Y \ge (1/\bar\rho) h_Z$ on Z. Moreover, letting $n \ge 1$ be as in Lemma 3.1(d), for any $y \in Y$ there exists $z \in Z \cap \mathrm{Int}\{\varphi = n\}$ with $Fz = y$. Since f^n has finitely many continuous branches, $M = |Jf^n|_\infty < \infty$. We obtain

$$\begin{aligned} h_Y(y) = \sum_{Fy'=y} JF(y')^{-1} h_Y(y') &\ge JF(z)^{-1} h_Y(z) \\ &= Jf^n(z)^{-1} h_Y(z) \ge \bar\rho^{-1} M^{-1} \inf h_Z > 0. \end{aligned}$$

Hence h_Y is bounded below. Uniqueness of h_Y follows.

It remains to show that h_Y is bounded above. This follows from a result of Rychlik [57, Theorem 1] once we check three conditions:

1. There exists a constant $C > 0$ such that $\sup_a JF^k \le C \inf_a JF^k$ for all $a \in \alpha_k^Y$, $k \ge 1$.

2. There exists $\epsilon > 0$, $r \in (0,1)$ such that if $a \in \alpha_k^Y$ for some $k \geq 1$ and $\mathrm{Leb}(F^k a) < \epsilon$, then $\sum_{\{a' \in \alpha^Y : \mathrm{Leb}(a' \cap F^k a) > 0\}} \sup_{a'} 1/JF \leq r$.

3. $\sum_{a \in \alpha^Y} \sup_a 1/JF < \infty$.

By [57, Theorem 1], there exists an F-invariant density $h_1 \in L^\infty(Y)$. Since h_Y is the unique F-invariant density, we have $h_Y = h_1$ bounded.

Now, condition 1 holds by Corollary 2.13. Condition 2 is trivially satisfied since the set $\{F^k a : a \in \alpha_k^Y,\ k \geq 1\}$ is finite. By Corollary 2.10, $1/JF \sim 4^{-n+1} n^{-(1+\alpha)}$ uniformly on $a = Y_{n,j}$, $j = 1, \dots, 4^n$ as $n \to \infty$, and the third condition follows. ∎

Define $\tau = \varphi_\rho = \sum_{\ell=0}^{\rho-1} \varphi \circ F^\ell : Z \to \mathbb{Z}^+$.

Proposition 3.3 *τ is Lebesgue integrable (equivalently $\int_Z \tau \, d\mu_Z < \infty$) if and only if $\gamma < 1$.*

Proof A standard argument, see for instance [13, 48], shows that $\tau = \varphi_\rho$ satisfies $\mu_Z(\tau > n) = O((\log n)^\alpha n^{-\alpha})$. Integrability for $\gamma < 1$ follows.

Similarly, $\mu_Z(\tau > n) \gg (\log n)^{-1} n^{-\alpha}$ (see for example [10, Proposition 5.1(b)]) proving nonintegrability for $\gamma \geq 1$. ∎

Lemma 3.4 *There exists a unique (up to scaling) absolutely continuous f-invariant σ-finite measure μ. Moreover, the density $h_X = d\mu/d\,\mathrm{Leb}$ is bounded below.*

The measure μ is finite if and only if $\gamma < 1$, in which case f is mixing.

Proof Since $F = f^\varphi : Y \to Y$ and $G = F^\rho : Z \to Z$, it follows that $G = f^\tau : Z \to Z$.

We proceed similarly to the proof of Lemma 3.2 but with ρ replaced by τ and F replaced by f. Form the new Young tower $\tilde{g} : \tilde{\Delta} \to \tilde{\Delta}$,

$$\tilde{\Delta} = \{(y, \ell) \in Z \times \mathbb{Z} : 0 \leq \ell \leq \tau(y) - 1\}, \qquad \tilde{g}(y, \ell) = \begin{cases} (y, \ell + 1) & \ell \leq \tau(y) - 2 \\ (Gy, 0) & \ell = \tau(y) - 1 \end{cases}.$$

The ergodic $\tilde{g}$-invariant measure $\mu_Z \times$ counting is finite if and only if $\bar{\tau} = \int_Z \tau \, d\mu_Z < \infty$. Equivalently $\int_Z \tau \, d\,\mathrm{Leb} < \infty$, and by Proposition 3.3, this holds if and only if $\gamma < 1$.

When $\gamma < 1$, the measure $\tilde{\mu}_\Delta = (\mu_Z \times \text{counting})/\bar{\tau}$ is an ergodic $\tilde{g}$-invariant probability measure on $\tilde{\Delta}$. The projection $\tilde{\pi} : \tilde{\Delta} \to X$, $\tilde{\pi}(y, \ell) = f^\ell y$ defines a semiconjugacy between $\tilde{g}$ and f, and $\mu = \tilde{\pi}_* \tilde{\mu}_\Delta$ is an absolutely continuous f-invariant probability measure. Lemma 3.1(c) implies that $\gcd\{\tau(a),\, a \in \alpha^Z\} = 1$. Since G is a full-branch Gibbs-Markov map, it follows from [65, Theorem 1] that $\tilde{\mu}_\Delta$, and hence μ, is mixing.

Again, as in the proof of Lemma 3.2, $h_X \geq (1/\bar{\tau}) h_Z$ on Z. By Lemma 3.1, Z is open, so by Proposition 2.2 there exists $n \geq 1$ such that $f^n Z = X$. Since f^n has finitely many branches, $M = |Jf^n|_\infty < \infty$. Given $x \in X$, choose $z \in Z$ such that $f^n z = x$. Then

$$h_X(x) = \sum_{f^n x' = x} Jf^n(x')^{-1} h_X(x') \geq Jf^n(z)^{-1} h_X(z) \geq (1/\bar{\tau}) M^{-1} \inf h_Z > 0.$$

Hence h_X is bounded below, and uniqueness of μ follows.

When $\gamma \geq 1$, we proceed in the same way but without normalising by $\bar{\tau}$. ∎

3.2 Tail estimate for $\tau = \varphi_\rho$

As noted in the proof of Proposition 3.3, the induced return time $\tau = \varphi_\rho$ satisfies the tail estimate $\mu(\tau > n) = O((\log n)^\alpha n^{-\alpha})$. In this subsection, we show how to remove the logarithmic factor using ideas from Szász and Varjú [61].

Lemma 3.5 *$\mu_Z(\tau > n) = O(n^{-\alpha})$.*

Following [14, Lemma 5.1] and [61, Lemma 16], the crucial ingredient for proving Lemma 3.5 is the following estimate. Fix $p, q \in (0, 1)$ satisfying $p < (1 - q)\alpha$. Let

$$Y(k, n) = \{\varphi = n \text{ and } \varphi \circ F^k > n^{1-q}\} \subset Y.$$

Proposition 3.6 *There exists $C > 0$ such that*

$$\mu_Y(Y(k, n)) \leq C n^{-(1+\alpha+p)} \quad \textit{for all } k, n \geq 1.$$

Proof For $k \geq 1$, denote by $\mathcal{H}^k$ the set of inverse branches $h : F^k a^h \to a^h$ of F^k. Define $S_n = \{\varphi > n^{1-q}\}$, and notice that

$$Y(k, n) = \left(\bigcup_{j=1}^{4^n} Y_{n,j} \right) \cap F^{-k}(S_n) = \bigcup_{j=1}^{4^n} \bigcup_{h \in \mathcal{H}^k} h(S_n) \cap Y_{n,j}.$$

But $h(S_n)$ is contained in the k-cylinder $a^h \in \alpha_k^Y$ while $Y_{n,j} \in \alpha^Y$. Therefore, if $h(S_n) \cap Y_{n,j} \neq \emptyset$ then $a^h \subset Y_{n,j}$. It follows that $\bigcup_{h \in \mathcal{H}^k} h(S_n) \cap Y_{n,j} \subset \bigcup_{h \in \mathcal{H}^k : a^h \subset Y_{n,j}} h(S_n)$ and so

$$Y(k, n) \subset \bigcup_{j=1}^{4^n} \bigcup_{h \in \mathcal{H}^k \,:\, a^h \subset Y_{n,j}} h(S_n).$$

Hence

$$\mu_Y(Y(k, n)) \leq \sum_{\{a \in \alpha^Y \,:\, \varphi(a) = n\}} \sum_{\{h \in \mathcal{H}^k \,:\, a^h \subset a\}} \mu_Y(h(S_n)).$$

If $a^h \subset a$, then $h = h_a \circ \tilde{h}$, where $h_a \in \mathcal{H}^1$ and $\tilde{h} \in \mathcal{H}^{k-1}$ are inverse branches with $h_a : Fa \to a$. By Lemma 3.2, the density $d\mu_Y/d\,\mathrm{Leb}$ is bounded above and below. Using this and F-invariance of μ_Y,

$$\sum_{\{h\in\mathcal{H}^k\,:\,a^h\subset a\}} \mu_Y(h(S_n)) \ll |Jh_a|_\infty \sum_{\tilde{h}\in\mathcal{H}^{k-1}} \mu_Y(\tilde{h}(S_n))$$
$$= |Jh_a|_\infty\, \mu_Y(F^{-(k-1)}(S_n)) = |Jh_a|_\infty\, \mu_Y(S_n).$$

By Corollaries 2.7 and 2.10,

$$\mu_Y(Y(k,n)) \ll \mu_Y(S_n) \sum_{\{a\in\alpha^Y\,:\,\varphi(a)=n\}} |Jh_a|_\infty \ll n^{-(1-q)\alpha} n^{-(1+\alpha)} \ll n^{-(1+\alpha+p)}$$

by the choice of p and q. ∎

Lemma 3.5 now holds by standard arguments. We follow the exposition in [9]. Define

$$Z_b(n) = \big\{\rho \le b\log n \text{ and } \max_{0\le\ell<\rho} \varphi\circ F^\ell \le \tfrac{1}{2}n \text{ and } \tau \ge n\big\} \subset Z.$$

Corollary 3.7 *Let $b > 0$. Then $\mu_Z(Z_b(n)) = o(n^{-\alpha})$.*

Proof Define

$$Y_b(n) = \{\varphi = n \text{ and } \varphi\circ F^k > n^{1-q} \text{ for some } 1 \le k \le 2b\log n\} \subset Y.$$

By Proposition 3.6,

$$\mu_Y(Y_b(n)) \le C(2b\log n)n^{-(1+\alpha+p)} \ll n^{-(1+\alpha+p/2)}.$$

Let $z \in Z_b(n)$. Define $\varphi_1(z) = \max_{0\le\ell<\rho(z)} \varphi(F^\ell z)$ and choose $\ell_1(z) \in \{0,\dots,\rho(z)-1\}$ such that $\varphi_1(z) = \varphi(F^{\ell_1(z)}z)$. Define $\varphi_2(z) = \max_{0\le\ell<\rho(z)\,,\,\ell\ne\ell_1(z)} \varphi(F^\ell z)$.

Now, $n \le \tau \le \varphi_1 + (\rho-1)\varphi_2 \le \frac{1}{2}n + (b\log n)\varphi_2$. Hence

$$\frac{n}{2b\log n} \le \varphi_2 \le \varphi_1 \le \frac{n}{2}.$$

In particular, $\varphi_1 > \varphi_2^{1-q}$ and $\varphi_2 > \varphi_1^{1-q}$ for n large.

Choose $\ell_2(z) \in \{0,\dots,\rho(z)-1\}$ such that $\ell_2(z) \ne \ell_1(z)$ and $\varphi_2(z) = \varphi(F^{\ell_2(z)})$. Suppose for definiteness that $\ell_1(z) < \ell_2(z)$ (the other case is similar). Let $m = \varphi_1(z)$, $k = \ell_2(z) - \ell_1(z)$. Then

- $\varphi(F^{\ell_1(z)}z) = \varphi_1(z) = m$;
- $\varphi\circ F^k(F^{\ell_1(z)}z) = \varphi(F^{\ell_2(z)}z) = \varphi_2(z) > \varphi_1(z)^{1-q} = m^{1-q}$;
- $1 \le k \le \ell_2(z) \le b\log n \le 2b\log\varphi_1(z) = 2b\log m$ for n large.

Hence, $F^{\ell_1(z)}z \in Y_b(m)$ for n large. We have shown that

$$Z_b(n) \subset F^{-\ell}Y_b(m) \quad \text{for some } \ell < b\log n,\ m \geq n/(2b\log n),$$

and so

$$\begin{aligned}\mu_Z(Z_b(n)) \ll \mu_Y(Z_b(n)) \ll \log n \sum_{m\geq n/(2b\log n)} \mu_Y(Y_b(m)) \\ \ll \log n \sum_{m\geq n/(2b\log n)} m^{-(1+\alpha+p/2)} \ll \log n(n/\log n)^{-(\alpha+p/2)} = o(n^{-\alpha})\end{aligned}$$

as required. ∎

Proof of Lemma 3.5 Let $\Delta = \{(z,\ell) \in Z \times \mathbb{Z} : 0 \leq \ell < \rho(z)\}$ be the Young tower from the proof of Lemma 3.2 with probability measure $\mu_\Delta = (\mu_Z \times \text{counting})/\bar{\rho}$. Recall that $\mu_Y = \pi_*\mu_\Delta$ where $\pi(z,\ell) = F^\ell z$.

Write $\max_{0\leq\ell<\rho(z)} \varphi(F^\ell z) = \varphi(F^{\ell_1(z)}z)$ where $\ell_1(z) \in \{0,\dots,\rho(z)-1\}$. Then

$$\begin{aligned}\mu_Z\big(z \in Z : \max_{0\leq\ell<\rho(z)} \varphi(F^\ell z) > n/2\big) &= \bar{\rho}\mu_\Delta\{(z,0) \in \Delta : \varphi(F^{\ell_1(z)}z) > n/2\} \\ &= \bar{\rho}\mu_\Delta\{(z,\ell_1(z)) : \varphi(F^{\ell_1(z)}z) > n/2\} = \bar{\rho}\mu_\Delta\{(z,\ell_1(z)) : \varphi\circ\pi(z,\ell_1(z)) > n/2\} \\ &\leq \bar{\rho}\mu_\Delta\{p \in \Delta : \varphi\circ\pi(p) > n/2\} = \bar{\rho}\mu_Y\{y \in Y : \varphi > n/2\} = O(n^{-\alpha}).\end{aligned}$$

Hence, by Corollary 3.7,

$$\mu_Z(\rho \leq b\log n \text{ and } \tau > n) = O(n^{-\alpha}).$$

Finally, by Lemma 3.1(a), $\mu_Z(\rho > b\log n) = O(\delta^{b\log n}) = O(n^{b\log\delta}) = o(n^{-\alpha})$ for any b fixed sufficiently large. Hence $\mu_Z(\tau > n) = O(n^{-\alpha})$ as required. ∎

3.3 Proof of upper bounds for decay of correlations, and various statistical properties

We suppose throughout this subsection that $\gamma < 1$, and set $\alpha = 1/\gamma$. In the proof of Lemma 3.4, we showed that the intermittent map $f : X \to X$ is modelled by a Young tower $\tilde{g} : \tilde{\Delta} \to \tilde{\Delta}$ with first return $G = f^\tau : Z \to Z$. By Lemma 3.5, the return time tails satisfy $\mu_Z(\tau > n) = O(n^{-\alpha})$. Accordingly we can read off numerous statistical properties that hold for all Hölder (and dynamically Hölder) observables $v : X \to \mathbb{R}$.

Recall that $\tilde{\pi} : \tilde{\Delta} \to X$, given by $\tilde{\pi}(z,\ell) = f^\ell z$ is a semiconjugacy between $\tilde{g}$ and f. Moreover, we have invariant ergodic probability measures $\tilde{\mu}_\Delta$ on $\tilde{\Delta}$ and μ on X where $\tilde{\mu}_\Delta = (\mu_Z \times \text{counting})/\bar{\tau}$ and $\mu = \tilde{\pi}_*\mu_\Delta$.

Recall from Lemma 3.1 that $G : Z \to Z$ is a full-branched Gibbs-Markov map with partition α^Z. Define the separation time $s(z,z')$ on Z to be the least integer

$n \ge 0$ such that $G^n z, G^n z'$ lie in distinct partition elements in α^Z. For $\theta \in (0,1)$ and $v : X \to \mathbb{R}$, define

$$\|v\|_{\mathcal{H}_\theta} = |v|_\infty + |v|_{\mathcal{H}_\theta}, \quad |v|_{\mathcal{H}_\theta} = \sup_{z,z' \in Z: z \ne z'} \sup_{0 \le \ell < \tau(z)} \frac{|v(f^\ell z) - v(f^\ell z')|}{\theta^{s(z,z')}}.$$

An observable $v : X \to \mathbb{R}$ is said to be *dynamically-Hölder* if $\|v\|_{\mathcal{H}_\theta} < \infty$ for some choice of θ.

Proposition 3.8 *Hölder observables are dynamically Hölder. Moreover, for $v \in C^\eta(X)$, $\eta \in (0,1)$, we have $\|v\|_{\mathcal{H}_\theta} \le 2^{\eta/2}\theta^{-1}\|v\|_\eta$ where $\theta = 4^{-\eta}$.*

Proof Let $z, z' \in Z$ with $s(z,z') = n$. Then

$$\sqrt{2} = \operatorname{diam} M \ge |G^n z - G^n z'| \ge 4^n |z - z'|,$$

so $|z - z'| \le \sqrt{2}\, 4^{-s(z,z')}$ for all $z, z' \in Z$.

Now, let $z, z' \in Z$, $0 \le \ell < \tau(z)$. Then

$$\begin{aligned} |v(f^\ell z - v(f^\ell z')| &\le \|v\|_\eta \, |f^\ell z - f^\ell z'|^\eta \le \|v\|_\eta \, |Gz - Gz'|^\eta \\ &\le \|v\|_\eta 2^{\eta/2}\, \theta^{s(Gz,Gz')} \le \|v\|_\eta 2^{\eta/2}\theta^{-1}\, \theta^{s(z,z')} \end{aligned}$$

yielding the desired estimate. ∎

Proof of Theorems 1.1(a), 1.2 and 1.4 (The proof of Theorem 1.4(c) for $\gamma \in (\frac{1}{2}, 1)$ is momentarily contingent on Theorem 1.3.)

Given $v : X \to \mathbb{R}$, we define the lifted observable $\tilde{v} = v \circ \tilde{\pi} : \tilde{\Delta} \to \mathbb{R}$. Since $\tilde{\pi}$ is a measure-preserving semiconjugacy, the desired statistical properties for v follow from those for $\tilde{v}$. Also, $|v|_{\mathcal{H}_\theta} = \sup_{z,z' \in Z: z \ne z'} \sup_{0 \le \ell < \tau(z)} \frac{|\tilde{v}(z,\ell) - \tilde{v}(z',\ell)|}{\theta^{s(z,z')}}$, so dynamically Hölder observables lie in the standard function space $C_\theta(\tilde{\Delta})$ considered on one-sided Young towers [65]. The upper bound on decay of correlations in Theorem 1.1(a) now follows from [65, Theorem 3].

Next, we turn to Theorem 1.2. Part (a) holds by [65, Theorem 4]. Parts (b) and (c) follow respectively from [30, Theorems 1.3 and 1.2]. Part (f) is proved in [17, Theorem 5.3] and part (d) is a standard consequence. Part (e) is proved in [4, Theorem 2.2].

Finally, we consider Theorem 1.4. Part (a) follows from Theorem 1.1(a) by [49, Theorem 1.2]. Part (b) is proved in [34, Theorem 1.4]. Part (c) is an immediate consequence of part (b) together with the CLT for $\gamma < \frac{1}{2}$ and Theorem 1.3 for $\gamma \in (\frac{1}{2}, 1)$. ∎

Remark 3.9 Alternative references for some of the results in the above proof include [18, 21, 43, 44, 50, 51, 53].

For brevity, we have omitted various other statistical properties that follow from the existence of the Young tower $\tilde{\Delta}$ such as concentration inequalities [34]. Also, for $\gamma < \frac{1}{2}$, homogenization (convergence of fast-slow systems to a stochastic differential equation) holds when the fast dynamics is given by f, see [15, 16, 27, 41, 42, 45].

3.4 Aperiodicity

Let $S^1 = \{\omega \in \mathbb{C} : |\omega| = 1\}$ and consider the cohomological equation

$$v \circ F = \omega^\varphi v, \tag{3.2}$$

where $v : Y \to S^1$ is measurable and $\omega \in S^1$. If $\omega = 1$, then since F is ergodic, the measurable solutions to equation (3.2) are precisely the constant solutions. Absence of solutions for $\omega \neq 1$ is called *aperiodicity*. In this subsection, we prove:

Lemma 3.10 *For each $\omega \in S^1 \setminus \{1\}$ there are no measurable solutions $v : Y \to S^1$ to equation (3.2).*

Aperiodicity is useful for ruling out peripheral spectra for certain twisted transfer operators. Instances of this are seen in Corollaries 4.10(e) and 5.2(ii) below.

For the moment, consider an arbitrary ergodic measure-preserving transformations $F : Y \to Y$ defined on a probability space (Y, μ_Y). Let $U : L^1(Y) \to L^1(Y)$ denote the Koopman operator $Uv = v \circ F$ and define the transfer operator $R : L^1(Y) \to L^1(Y)$, where $\int_Y Rv\, w\, d\mu_Y = \int_Y v\, Uw\, d\mu_Y$ for all $v \in L^1(Y)$, $w \in L^\infty(Y)$.

For $\omega \in S^1$, we define the twisted Koopman and transfer operators $U(\omega)v = \bar{\omega}^\varphi Uv = \bar{\omega}^\varphi v \circ F$ and $R(\omega)v = R(\omega^\varphi v)$. Note that $R(\omega)$ is the L^2 adjoint of $U(\omega)$ but that $R(\omega) : L^1 \to L^1$ is the dual of $U(\bar{\omega}) : L^\infty \to L^\infty$. (This discrepancy between adjoints and duals over the complex numbers is standard.)

Proposition 3.11 *Suppose that $F : Y \to Y$ is ergodic. Let $\omega \in S^1$ and let $v : Y \to \mathbb{C}$ be L^1. Then $U(\omega)v = v$ if and only if $R(\omega)v = v$, in which case $|v|$ is constant.*

Proof First, note that if $U(\omega)v = v$, then $|v| \circ F = |v|$ and so $|v|$ is constant by ergodicity.

Next, recall that $RU = I$ and hence $R(\omega)U(\omega) = I$ for all ω. If $U(\omega)v = v$, then $v = R(\omega)U(\omega)v = R(\omega)v$, proving one direction.

Conversely, suppose that $R(\omega)v = v$. By duality, $\int_Y v\, U(\bar{\omega})^n w\, d\mu_Y = \int_Y vw\, d\mu_Y$ for every $w \in L^\infty$ and $n \geq 1$. We claim that v is bounded and $|v|_\infty \leq |v|_1$. Suppose the claim is false. Then there is a set E of positive measure and $c > |v|_1$ such that

$|v| \ge c$ on E. Choose $w = 1_E\bar v/|v|$ on $\{v \ne 0\}$ and $w = 1_E$ elsewhere. Then

$$\begin{aligned} c\mu_Y(E) &\le \int_E |v|\,d\mu_Y = \int_Y vw\,d\mu_Y = \int_Y v\frac{1}{n}\sum_{j=0}^{n-1} U(\bar\omega)^j w\,d\mu_Y \\ &= \int_Y v\frac{1}{n}\sum_{j=0}^{n-1} \omega^{\varphi_j}\, w\circ F^j\,d\mu_Y \le \int_Y |v|\frac{1}{n}\sum_{j=0}^{n-1} |w|\circ F^j\,d\mu_Y. \end{aligned}$$

The last integrand is dominated by $|w|_\infty|v| \in L^1$ and converges a.e. to $|v|\int_Y |w|\,d\mu_Y$ by the pointwise ergodic theorem. By the dominated convergence theorem,

$$\lim_{n\to\infty}\int_Y |v|\frac{1}{n}\sum_{j=0}^{n-1} |w|\circ F^j\,d\mu_Y = \int_Y |v|\,d\mu_Y\ \mu_Y(E).$$

Hence $c \le \int_Y |v|\,d\mu_Y$ which is a contradiction.

This proves the claim, so v is bounded. In particular, $v \in L^2$ and a computation using that $R(\omega) = U(\omega)^*$ and $R(\omega)v = v$ shows that $\langle U(\omega)v - v, U(\omega)v - v\rangle = 0$ so that $U(\omega)v = v$ as required. ∎

Returning to the intermittent maps (1.3), we obtain

Corollary 3.12 *For each $\omega \in S^1 \setminus \{1\}$ there are no L^1 functions $v : Y \to S^1$ such that $R(\omega)v = v$.*

Proof This is immediate from Lemma 3.10 and Proposition 3.11. ∎

To prove Lemma 3.10, we make use of two Young towers $g : \Delta \to \Delta$ and $\tilde g : \tilde\Delta \to \tilde\Delta$. The second of these coincides with the tower in the proof of Lemma 3.4. The first tower is different from those considered so far in this paper (in particular, that of Lemma 3.2), and is defined as follows:

$$\Delta = \{(y,\ell) \in Y \times \mathbb{Z} : 0 \le \ell \le \varphi(y) - 1\}, \qquad g(y,\ell) = \begin{cases} (y,\ell+1) & \ell \le \varphi(y) - 2 \\ (Fy, 0) & \ell = \varphi(y) - 1 \end{cases}.$$

As in Lemma 3.4, we have ergodic g-invariant and $\tilde g$-invariant measures $\mu_\Delta = \mu_Y \times \text{counting}$ and $\tilde\mu_\Delta = (\mu_Z \times \text{counting})/\bar\sigma$ on Δ and $\tilde\Delta$. (When $\gamma < 1$, it follows from Corollary 2.7 and Lemma 3.2 that $\bar\varphi = \int_Y \varphi\,d\mu_Y < \infty$ and hence we can normalize further by $\bar\varphi$ to obtain probability measures μ_Δ and $\tilde\mu_\Delta$.)

Remark 3.13 A standard strategy, used below, to establish aperiodicity is to show that g is weak mixing. This is made complicated by that fact that g is nonMarkov, so we pass to the Markov extension $\tilde g$. (This is similar in spirit, though the notation is more complicated, to the derivation of mixing properties for F from mixing properties for G in Subsection 3.1.)

Recall that $\tau : Z \to \mathbb{Z}^+$, $\tau = \varphi_\rho = \sum_{i=0}^{\rho-1} \varphi \circ F^i$. Write $\varphi_j = \sum_{i=0}^{j-1} \varphi \circ F^i$. Any element of $\tilde\Delta$ can be written uniquely as $(z, \varphi_j(z) + \ell)$ where $0 \le j \le \rho(z) - 1$ and $0 \le \ell \le \varphi(F^j z) - 1$, valid for $z \in Z$.

Define

$$\pi_\Delta : \tilde\Delta \to \Delta, \qquad \pi_\Delta(z, \varphi_j(z) + \ell) = (F^j z, \ell).$$

Proposition 3.14 *π_Δ is a measure-preserving semiconjugacy from $\tilde g$ to g.*

Proof To verify that π_Δ is a semiconjugacy ($\pi_\Delta \circ \tilde g = g \circ \pi_\Delta$), we show that $\pi_\Delta \circ \tilde g(z, \varphi_j(z) + \ell) = g \circ \pi_\Delta(z, \varphi_j(z) + \ell)$ for all z, j, ℓ.

Now,

$$g \circ \pi_\Delta(z, \varphi_j(z) + \ell) = g(F^j z, \ell) = \begin{cases} (F^j z, \ell + 1) & \ell \le \varphi(F^j z) - 2 \\ (F^{j+1} z, 0) & \ell = \varphi(F^j z) - 1. \end{cases}$$

Also,

$$\begin{aligned} \pi_\Delta \circ \tilde g(z, \varphi_j(z) + \ell) &= \begin{cases} \pi_\Delta(z, \varphi_j(z) + \ell + 1) & \ell \le \varphi(F^j z) - 2 \\ \pi_\Delta(z, \varphi_{j+1}(z)) & \ell = \varphi(F^j z) - 1,\ j \le \rho(z) - 2 \\ \pi_\Delta(Gz, 0) & \ell = \varphi(F^j z) - 1,\ j = \rho(z) - 1 \end{cases} \\ &= \begin{cases} (F^j z, \ell + 1) & \ell \le \varphi(F^j z) - 2 \\ (F^{j+1} z, 0) & \ell = \varphi(F^j z) - 1. \end{cases} \end{aligned}$$

Hence π_Δ is a semiconjugacy.

It remains to show that $(\pi_\Delta)_* \tilde\mu_\Delta = \mu_\Delta$. It suffices to test this for sets $E \times \{\ell\}$ where E is a measurable subset of a partition element $a \subset Z$, $a \in \alpha^Z$, and $0 \le \ell \le \varphi(a) - 1$. By (3.1), $\mu_\Delta(E \times \{\ell\}) = \mu_Y(E) = (1/\bar\rho) \int_Z \sum_{j=0}^{\rho-1} 1_E \circ F^j \, d\mu_Z$. On the other hand

$$\begin{aligned} (\pi_\Delta)_* \tilde\mu_\Delta(E \times \{\ell\}) &= \tilde\mu_\Delta(\pi_\Delta^{-1}(E \times \{\ell\})) = \tilde\mu_\Delta\{(x, \varphi_j(z) + \ell) : F^j z \in E,\ j < \rho(a)\} \\ &= (1/\bar\rho) \int_Z \sum_{j=0}^{\rho-1} 1_E \circ F^j \, d\mu_Z. \end{aligned}$$

This completes the proof. ∎

Proof of Lemma 3.10 Suppose that $u : \tilde\Delta \to S^1$ is measurable and $u \circ \tilde g = \omega u$ for some $\omega \in S^1$. Define $V : Z \to S^1$, $V(z) = u(z, 0)$. Then $V(Gz) = u(Gz, 0) = u \circ \tilde g^{\tau(z)}(z, 0) = \omega^{\tau(z)} V(z)$. Since G is a full branch Gibbs-Markov map, for every $a \in \alpha^Z$ there exists $z_a \in a$ with $Gz_a = z_a$ and so $V(z_a) = \omega^{\tau(a)} V(z_a)$. Hence $\omega^{\tau(a)} = 1$ for all a. By Lemma 3.1(c), $\gcd\{\tau(a),\ a \in \alpha^Z\} = 1$ and it follows that $\omega = 1$. In other words, $\tilde g : \tilde\Delta \to \tilde\Delta$ is weak mixing.

By Proposition 3.14, $g : \Delta \to \Delta$ is weak mixing. Again this means that the equation $u \circ g = \omega u$ has no measurable solutions $u : \Delta \to S^1$ for each $\omega \in S^1 \setminus \{1\}$.

Let $\omega \in S^1 \setminus \{1\}$ and suppose that $v : Y \to S^1$ is a measurable solution to (3.2). Define $u(y, \ell) = \omega^\ell v(y)$. Then $u : \Delta \to S^1$ is measurable. If $\ell \le \varphi(y) - 2$, then $u \circ g(y, \ell) = u(y, \ell + 1) = \omega^{\ell+1} v(y) = \omega u(y, \ell)$. If $\ell = \varphi(y) - 1$, then $u \circ g(y, \ell) = u(Fy, 0) = v(Fy) = \omega^{\varphi(y)} v(y) = \omega^{\ell+1} v(y) = \omega u(y, \ell)$. This shows that $u \circ g = \omega u$ which is impossible since g is weak mixing. Hence there are no such measurable solutions to (3.2). ∎

4 Estimates in two-dimensional BV

Let λ_m denote m-dimensional Lebesgue measure. For $v \in L^1(Y)$, define the variation

$$\operatorname{Var} v = \sup_\omega \int_{\mathbb{R}^2} v \operatorname{div} \omega \, d\lambda_2,$$

where the supremum is taken over all compactly supported C^1 test functions $\omega : \mathbb{R}^2 \to \mathbb{R}^2$ such that $|\omega|_\infty \le 1$. Let $\mathrm{BV}(Y)$ consist of those functions $v \in L^1(Y)$ such that $\operatorname{Var} v < \infty$. This is a Banach space with norm $\|v\|_{\mathrm{BV}} = \int_Y |v| \, d\lambda_2 + \operatorname{Var} v$. Recall that C^1 functions lie in $\mathrm{BV}(Y)$ and $\operatorname{Var} v = \int_Y |\nabla v| \, d\lambda_2$ for such functions, where $|\nabla v| = (|\partial v / \partial y|^2 + |\partial v / \partial \theta|^2)^{\frac{1}{2}}$.

We use the fact [25, Remark 2.14] that if w is continuous on a set U with Lipschitz boundary and w is C^1 on $\operatorname{Int} U$, then $\operatorname{Var}(1_U w) = \int_U |\nabla w| \, d\lambda_2 + \int_{\partial U} |w| \, d\lambda_1$. (The measure will often be suppressed when the meaning is clear.)

The following standard result [25, Theorem 1.17] allows us to reduce to considering C^1 functions $v : \mathbb{R}^2 \to \mathbb{R}$ in many estimates.

Proposition 4.1 *Let $v \in \mathrm{BV}(Y)$. There exists a sequence of C^1 functions $v_n : \mathbb{R}^2 \to \mathbb{R}$ such that $v_n \to v$ in $L^1(Y)$ and $\lim_{n\to\infty} \operatorname{Var} v_n = \operatorname{Var} v$.* ∎

Corollary 4.2 *Let $A : L^1(Y) \to L^1(Y)$ be a bounded linear operator. If $\operatorname{Var}(Av) \le C_1 \int_Y |v| + C_2 \operatorname{Var} v$ for all $v \in C^1$, then $\operatorname{Var}(Av) \le C_1 \int_Y |v| + C_2 \operatorname{Var} v$ for all $v \in \mathrm{BV}(Y)$.*

Proof Let $v \in \mathrm{BV}(Y)$ and choose a sequence v_n as in Proposition 4.1. Let ω be a C^1 test function. Since $Av_n \to Av$ in $L^1(Y)$,

$$\begin{aligned}\int_{\mathbb{R}^2} Av \operatorname{div} \omega = \lim_{n\to\infty} \int_{\mathbb{R}^2} Av_n \operatorname{div} \omega = \limsup_{n\to\infty} \int_{\mathbb{R}^2} Av_n \operatorname{div} \omega \le \limsup_{n\to\infty} \operatorname{Var}(Av_n) \\ \le \limsup_{n\to\infty} \Big(C_1 \int_Y |v_n| + C_2 \operatorname{Var} v_n \Big) = C_1 \int_Y |v| + C_2 \operatorname{Var} v.\end{aligned}$$

Taking the supremum over ω yields the desired result. ∎

We make some additional observations that are used in Section 5.

Remark 4.3 Returning to Proposition 4.1, suppose in addition that $v \in L^\infty(Y)$. Then the sequence v_n can be chosen to have the additional property that $\sup_Y |v_n| \le 3|v|_\infty$. To see this we use the notation from the proof of [25, Theorem 1.17] where v is denoted by f and the approximating sequence v_n is denoted by $f_\epsilon = \sum_i \eta_{\epsilon_i} \star (f\varphi_i)$. It is immediate from the definitions in [25] that $\sup_Y |f_\epsilon| \le 3\max_i \sup_Y |\eta_{\epsilon_i} \star (f\varphi_i)| \le 3\max_i(\int_Y |\eta_{\epsilon_i}|)\sup_Y |f\varphi_i| \le 3|f|_\infty$.

Let $\mathrm{BV}_\infty(Y) = \mathrm{BV}(Y) \cap L^\infty(Y)$ with norm $\|v\|_{\mathrm{BV}_\infty} = |v|_\infty + \operatorname{Var} v$.

Corollary 4.4 *$\mathrm{BV}_\infty(Y)$ is a Banach algebra.*

Proof Let $v, w \in \mathrm{BV}_\infty(Y)$. By Remark 4.3, there exists a sequence of C^1 functions v_n such that $v_n \to v$ in $L^1(Y)$, $\sup_Y |v_n| \le 3|v|_\infty$ and $\operatorname{Var} v_n \to \operatorname{Var} v$. Let w_n be a similar approximating sequence for w.

Note that $v_n w_n$ is C^1 and hence lies in BV, while $\int_Y |v_n w_n - vw| \le \sup_Y |v_n| \int_Y |w_n - w| + (\int_Y |v_n - v|)|w|_\infty \le 3|v|_\infty \int |w_n - w| + |w|_\infty \int_Y |v_n - v| \to 0$ as $n \to \infty$. Hence it follows from [25, Theorem 1.9] that $\operatorname{Var}(vw) \le \liminf_{n\to\infty} \operatorname{Var}(v_n w_n)$.

Since v_n and w_n are C^1, we have $\operatorname{Var}(v_n w_n) \le \sup_Y |v_n| \operatorname{Var} w_n + \sup_Y |w_n| \operatorname{Var} w_n$. Hence

$$\operatorname{Var}(vw) \le \liminf_{n\to\infty} 3(|v|_\infty \operatorname{Var} w_n + |w|_\infty \operatorname{Var} v_n) = 3(|v|_\infty \operatorname{Var} w + |w|_\infty \operatorname{Var} v).$$

It follows that

$$\|vw\|_{BV_\infty} \le |vw|_\infty + 3(|v|_\infty \operatorname{Var} w + |w|_\infty \operatorname{Var} v) \le 3\|v\|_{BV_\infty}\|w\|_{BV_\infty}$$

as required. ∎

Throughout the remainder of this section, $|v|_1$ denotes $\int_Y |v|\, d\lambda_2$.

4.1 Boundary terms

The primary difficulty in dealing with multidimensional BV is the occurrence of certain boundary terms. Let $a \in \alpha^Y$ denote a partition element and consider the branch $F_a : a \to Fa$. Let ∂F_a denote $F|_{\partial a} : \partial a \to \partial Fa$ with 1-dimensional derivative $D\partial F_a$. For the Lasota-Yorke inequality (Subsection 4.2 below) given $v \in C^1$, we are required to estimate terms of the form

$$\int_{\partial a} |v||D\partial F_a / JF_a|,$$

relative to the BV norm $\|v\|_{\mathrm{BV}(Y)}$. In one dimension, $\mathrm{BV}(Y)$ is embedded in L^∞ which simplifies the estimates considerably. For higher dimensions, much more work is required, see [19, 20, 26] and references therein.

Our main results in this subsection are:

Lemma 4.5 *Let $v : \mathbb{R}^2 \to \mathbb{R}$ be C^1. There is a constant $C_1 > 0$ such that*

$$\int_{\partial a} |v||D\partial F_a/JF_a| \le C_1(|1_a v|_1 + n^{-(1+\alpha)}|1_a\nabla v|_1) \quad \textit{for all } a \in \alpha^Y \textit{ with } \varphi(a) = n.$$

Lemma 4.6 *Let $v : \mathbb{R}^2 \to \mathbb{R}$ be C^1. Suppose that u is sufficiently close to constant as in Remark 2.6. Then there exists $\kappa_0 \in (0, \frac{3}{4})$, and for any $N_0 \ge 1$, there exists a constant $C_2 > 0$ such that*

$$\int_{\partial a} |v||D\partial F_a/JF_a| \le C_2|1_a v|_1 + \kappa_0|1_a\nabla v|_1$$

for all $a \in \alpha^Y$ with $\varphi(a) \le N_0$.

An immediate consequence is:

Corollary 4.7 *Let $v : \mathbb{R}^2 \to \mathbb{R}$ be C^1. There exists $\kappa_0 \in (0, \frac{3}{4})$ and $C_3 > 0$ such that*

$$\sum_a \int_{\partial a} |v||D\partial F_a/JF_a| \le C_3 \sum_a |1_a v|_1 + \kappa_0 \sum_a |1_a\nabla v|_1 = C_3|v|_1 + \kappa_0 \operatorname{Var} v. \qquad \blacksquare$$

In the remainder of this subsection, we prove Lemmas 4.5 and 4.6.

Recall from Section 2.2 that the partition elements form a 'rectangular' grid $\{Y_{n,j},\ n \ge 1,\ j = 1 \dots, 4^n\}$ where there are infinitely many columns $\mathcal{C}_n$, $n \ge 1$, bounded by 'vertical' curves $\xi_n(\theta)$, $0 \le \theta \le 1$. The column $\mathcal{C}_n$ is divided into 4^n partition elements $\{Y_{n,j}\}$ bounded by horizontal lines $\theta = j4^{-n}$, $j = 0, \dots, 4^n$. In particular, the partition element $a = Y_{n,j}$ is given by

$$Y_{n,j} = \{(y, \theta) : \xi_n(\theta) \le y \le \xi_{n-1}(\theta),\ (j-1)4^{-n} \le \theta \le j4^{-n}\}.$$

By Proposition 2.4,

$$(\xi_{n-1}(\theta) - \xi_n(\theta))^{-1} \ll n^{1+\alpha}, \tag{4.1}$$

uniformly in θ. Also, by Proposition 2.4,

$$M_n = \max\{|\xi'_{n-1}|_\infty, |\xi'_n|_\infty\} \ll n^{-(1+\alpha)}. \tag{4.2}$$

We write $\partial a = H_a \cup V_a$ where H_a is the union of the two horizontal edges and V_a is the union of the two 'vertical' edges.

Proof of Lemma 4.5 By Lemma B.5 and (4.1), (4.2),

$$\begin{aligned}\int_{H_a} |v| &\ll (4^n + M_n(\xi_{n-1} - \xi_n)^{-1})|1_a v|_1 + |1_a\partial_\theta v|_1 + M_n|1_a\partial_y v|_1 \\ &\ll 4^n|1_a v|_1 + |1_a\partial_\theta v|_1 + |1_a\partial_y v|_1 \ll 4^n|1_a v|_1 + |1_a\nabla v|_1.\end{aligned}$$

On the horizontal edges, $\partial F_a(y, \theta_0) = F_1(y, \theta_0)$ since horizontal lines are mapped to horizontal lines. By Corollary 2.10, $D\partial F_a = A$ and $|D\partial F_a/JF_a| = 4^{-n}$. Hence

$$\int_{H_a} |v|D\partial F_a/JF_a \ll |1_a v|_1 + 4^{-n}|1_a\nabla v|_1. \tag{4.3}$$

Similarly, it follows from Lemma B.4 that

$$\int_{V_a} |v| \ll n^{1+\alpha}|1_a v|_1 + |1_a \nabla v|_1.$$

On the 'vertical' edges, by Corollary 2.10, $|D\partial F_a| \ll 4^n$ and $|D\partial F_a/JF_a| \ll n^{-(1+\alpha)}$. Hence

$$\int_{V_a} |v||D\partial F_a/JF_a| \ll |1_a v|_1 + n^{-(1+\alpha)}|1_a \nabla v|_1. \tag{4.4}$$

Combining (4.3) and (4.4), we obtain the result. ∎

Proof of Lemma 4.6 We apply Theorem B.1. Since u is nearly constant as in Remark 2.6, we have $|M|_\infty < 7/\sqrt{72}$ by Remark 2.6 and hence

$$\sqrt{2}((1+|M|_\infty^2)^{1/2} + |M|_\infty) = 4\kappa_0,$$

where $\kappa_0 < \frac{3}{4}$. Hence $\int_{\partial a} |v| \le K(a)|1_a v|_1 + 4\kappa_0|1_a \nabla v|_1$ where $K(a)$ is a constant. The result follows since $|D\partial F_a/JF_a| \le \frac{1}{4}$. ∎

Remark 4.8 We can relax the artificial condition in Remark 2.6 by increasing the expansivity of f so that $D\partial F_a/JF_a$ is sufficiently small.

Alternatively, we could consider higher iterates. But now we have to check that the calculations in Lemma 4.5 remain intact. Note that Lemma 3.1 also requires a certain amount of expansion for F to overcome the complexity growth of discontinuities of F.

4.2 Lasota-Yorke inequality

Let $\widehat{F} : L^1(Y) \to L^1(Y)$ be the transfer operator corresponding to F relative to Lebesgue measure. (So $\int_Y \widehat{F}v\, w\, d\lambda_2 = \int_Y v\, w \circ F\, d\lambda_2$ for all $v \in L^1$, $w \in L^\infty$.) Then $\widehat{F}v = \sum_a 1_{Fa}(gv) \circ F_a^{-1}$ where $g = 1/|\det(DF)| = (JF)^{-1}$.

Also, for $\omega \in \overline{\mathbb{D}}$, we consider the twisted transfer operator $\widehat{F}(\omega)$ given by $\widehat{F}(\omega)v = \widehat{F}(\omega^\varphi v)$, so

$$\widehat{F}(\omega)v = \sum_a 1_{Fa}(g\omega^\varphi v) \circ F_a^{-1} = \sum_a \omega^{\varphi(a)} 1_{Fa}(gv) \circ F_a^{-1}.$$

We note that

$$|\widehat{F}(\omega)v|_1 = |\widehat{F}(\omega^\varphi v)|_1 \le |\omega^\varphi v|_1 \le |\omega||v|_1. \tag{4.5}$$

Lemma 4.9 *There exist constants $C > 0$ and $\kappa_1 \in (0,1)$ such that*

$$\operatorname{Var}(\widehat{F}(\omega)v) \le |\omega|(C|v|_1 + \kappa_1 \operatorname{Var} v) \quad \textit{for all } v \in \mathrm{BV}(Y),\ \omega \in \overline{\mathbb{D}}.$$

Proof By Corollary 4.2, it suffices to prove this for $v \in C^1$.

First we consider the case $\omega = 1$. Note that $(gv) \circ F_a^{-1}$ is C^1 on Fa, and so

$$\begin{aligned}\operatorname{Var}(\widehat{F}v) &\leq \sum_a \operatorname{Var}(1_{Fa}(gv) \circ F_a^{-1}) \\ &= \sum_a \int_{Fa} |\nabla[(gv) \circ F_a^{-1}]| \, d\lambda_2 + \sum_a \int_{\partial Fa} |(gv) \circ F_a^{-1}| \, d\lambda_1.\end{aligned}$$

Now,

$$\begin{aligned}\int_{Fa} |\nabla[(gv) \circ F_a^{-1}]| &= \int_{Fa} |\nabla(gv) \circ F_a^{-1} \cdot D(F_a^{-1})| \\ &= \int_{Fa} |\nabla(gv) \circ F_a^{-1} \cdot (DF_a)^{-1} \circ F_a^{-1}| = \int_a |\nabla(gv)(DF_a)^{-1}| JF_a \\ &\leq \int_a |v||(\nabla g)(DF_a)^{-1}| JF_a + \int_a |g||\nabla v||(DF_a)^{-1}| JF_a \\ &= \int_a |v||(\nabla g)(DF_a)^{-1}| JF_a + \int_a |\nabla v||(DF_a)^{-1}| \leq C_1 |1_a v|_1 + \tfrac{1}{4} |1_a \nabla v|_1,\end{aligned}$$

where $C_1 = \sup_a |(\nabla g)(DF_a)^{-1}| JF_a < \infty$ by Corollary 2.14 and we have used the fact that $|DF| \geq 4$. Also,

$$\int_{\partial Fa} |(gv) \circ F_a^{-1}| = \int_{\partial a} |gv||D(\partial F_a)| = \int_{\partial a} |v||D\partial F_a / JF_a|.$$

We have shown that

$$\operatorname{Var}(\widehat{F}v) \leq C_1 |v|_1 + \frac{1}{4} |\nabla v|_1 + \sum_a \int_{\partial a} |v||D\partial F_a / JF_a|.$$

Applying Corollary 4.7, we obtain $\operatorname{Var}(\widehat{F}v) \leq C|v|_1 + \kappa_1 \operatorname{Var} v$, with $\kappa_1 = \frac{1}{4} + \kappa_0$.

For general ω, we have an extra factor of $|\omega|^{\varphi(a)}$ throughout. Since $|\omega| \leq 1$ and $\varphi \geq 1$, this is bounded by $|\omega|$. ∎

Corollary 4.10 *(a) 1 is a simple eigenvalue for $\widehat{F} : L^1(Y) \to L^1(Y)$ with eigenfunction $h_Y = d\mu_Y / d\,\mathrm{Leb}$.*

(b) $\widehat{F}(\omega) : L^1(Y) \to L^1(Y)$ has spectral radius at most $|\omega|$ for all $\omega \in \overline{\mathbb{D}}$.

(c) Let $\kappa_1 \in (0,1)$ be as in Lemma 4.9. Then $\widehat{F}(\omega) : \mathrm{BV}(Y) \to \mathrm{BV}(Y)$ has essential spectral radius at most $\kappa_1 |\omega|$ for all $\omega \in \overline{\mathbb{D}}$.

(d) $h_Y \in \mathrm{BV}_\infty(Y)$.

(e) Regarding $\widehat{F}(\omega)$ as operators on $\mathrm{BV}(Y)$, it holds that 1 is a simple isolated eigenvalue in $\operatorname{spec} \widehat{F}$ and $1 \notin \operatorname{spec} \widehat{F}(\omega)$ for all $\omega \in \overline{\mathbb{D}} \setminus \{1\}$.

Proof (a) We have $\widehat{F}h_Y = h_Y$ so 1 is an eigenvalue for $\widehat{F}$. Simplicity follows from the fact that h_Y is the unique invariant density (Lemma 3.2).

(b) This is immediate by (4.5).

(c,d) By (4.5) and Lemma 4.9, $\|\widehat{F}(\omega)v\|_{\mathrm{BV}} \le |\omega|\{(C+1)|v|_1 + \kappa_1\|v\|_{\mathrm{BV}}\}$. Since the unit ball in $\mathrm{BV}(Y)$ is compact in $L^1(Y)$, the estimate on the essential spectrum radius follows from [35]. Moreover, 1 is an eigenvalue for $\widehat{F} : \mathrm{BV}(Y) \to \mathrm{BV}(Y)$ by [35]. By Lemma 3.2, the corresponding density coincides with h_Y, so $h_Y \in \mathrm{BV}(Y)$. Also, $h_Y \in L^\infty(Y)$ by Lemma 3.2.

(e) By part (c), it suffices to consider the multiplicity of 1 as an eigenvalue for $\widehat{F}(\omega)$ acting on $L^1(Y)$. Hence the result follows from part (a) for $\omega = 1$ and part (b) for $|\omega| < 1$. Finally, we note that $\widehat{F}(\omega) = h_Y R(\omega) h_Y^{-1}$ where $R(\omega)$ is the normalized transfer operator corresponding to the invariant measure μ_Y. By Corollary 3.12, 1 is not an eigenvalue for $R(\omega)$ when $\omega \in S^1 \setminus \{1\}$. Hence the same holds for $\widehat{F}(\omega)$. ∎

4.3 Tail of the return time function

Let $c' > 0$ be as in Proposition 2.4.

Proposition 4.11 *There exists a constant $C > 0$ such that $\int_{\{\varphi=n\}} |v|\, d\,\mathrm{Leb} \le Cn^{-(1+\alpha)}\|v\|_{\mathrm{BV}}$ for all $v \in \mathrm{BV}(Y)$. Moreover, for $v \in \mathrm{BV}(Y)$,*

$$\int_{\{\varphi=n\}} v\, d\,\mathrm{Leb} \sim \tfrac{1}{4}c' \int_{\mathbb{T}} v(\tfrac{3}{4}+, \theta)\, d\theta\, n^{-(1+\alpha)}, \tag{4.6}$$

where the one-sided limit $v(\frac{3}{4}+,\theta) = \lim_{y\to\frac{3}{4}+} v(y,\theta)$ exists for almost every θ and is integrable.

Taking v to be the density h_Y, we obtain

$$\mu_Y(\varphi = n) \sim c_2 n^{-(1+\alpha)} \quad \text{where } c_2 = \tfrac{1}{4}c' \int_{\mathbb{T}} h_Y(\tfrac{3}{4}+,\theta)\, d\theta.$$

Proof By [25, p. 29], $\mathrm{Var}\, v = \int \mathrm{Var}^y\, v\, d\theta = \int \mathrm{Var}^\theta\, v\, dy$ where $(\mathrm{Var}^y\, v)(\theta)$ denotes the one-dimensional variation of $v(\cdot,\theta)$ in the y-variable, and similarly for $(\mathrm{Var}^\theta\, v)(y)$.

Recall from Section 2.2 that for $n \ge 1$, $\{\varphi = n\} = \{(y,\theta) : y \in [y_n(\theta), y_{n-1}(\theta)], \theta \in \mathbb{T}\}$ and that $y_{n-1} - y_n \sim \frac{1}{4}c' n^{-(1+\alpha)}$ uniformly in θ. Hence for a.e. θ,

$$\begin{aligned} n^{1+\alpha} \int_{\{\varphi(\cdot,\theta)=n\}} |v(y,\theta)|\, dy &= n^{1+\alpha} \int_{y_n(\theta)}^{y_{n-1}(\theta)} |v(y,\theta)|\, dy \\ &\le n^{1+\alpha}(y_{n-1}(\theta) - y_n(\theta)) \sup\nolimits_y |v(y,\theta)| \ll \int |v(y,\theta)|\, dy + (\mathrm{Var}^y\, v)(\theta), \end{aligned}$$

so

$$n^{1+\alpha}\int_{\{\varphi=n\}}|v(y,\theta)|\,d\,\mathrm{Leb}\ll\int_{\mathbb{T}}\Big(\int|v(y,\theta)|\,dy+(\mathrm{Var}^y\,v)(\theta)\Big)\,d\theta=\|v\|_{\mathrm{BV}}.$$

This completes the proof of the first statement.

Next, note that BV functions restrict to BV functions on almost all one-dimensional slices (see [46, Lemma A.1, third statement] which is based on [24, Section 5.10.2, Theorem 2]. Moreover, one-dimensional BV functions have one-sided limits. Hence $J(\theta)=\lim_{y\to0+}v(\frac34+y,\theta)$ exists a.e. and is measurable (being a limit of measurable functions by Fubini's theorem). For a.e. θ,

$$|J(\theta)|\le\sup_y|v(y,\theta)|\le\int|v(y,\theta)|\,dy+(\mathrm{Var}^y\,v)(\theta).$$

Hence J is integrable and both sides of (4.6) are well-defined.

Let $A_n=n^{1+\alpha}\int_{\{\varphi=n\}}v\,d\,\mathrm{Leb}-\frac14c'\int_{\mathbb{T}}J(\theta)\,d\theta$. To prove validity of (4.6), we must show that $\lim_{n\to\infty}A_n=0$. Write

$$A_n=\int_{\mathbb{T}}B_n(\theta)\,d\theta,\qquad B_n(\theta)=n^{1+\alpha}\int_{y_n(\theta)}^{y_{n-1}(\theta)}v(y,\theta)\,dy-\tfrac14c'J(\theta).$$

We apply the dominated convergence theorem.

We have already seen that B_n is dominated by the L^1 function $\int|v(y,\cdot)|\,dy+\mathrm{Var}^y+\frac14c'|J|$. Next,

$$B_n(\theta)=\big\{n^{1+\alpha}(y_{n-1}(\theta)-y_n(\theta))-\tfrac14c'\big\}J(\theta)+n^{1+\alpha}\int_{y_n(\theta)}^{y_{n-1}(\theta)}\big\{v(y,\theta)-J(\theta)\big\}\,dy.$$

The first term converges to zero a.e. by the estimate for $y_{n-1}-y_n$. Also, $y_n\to\frac34+$, so

$$\begin{aligned}\Big|n^{1+\alpha}\int_{y_{n+1}(\theta)}^{y_n(\theta)}(v(y,\theta)-J(\theta))\,dy\Big|&\le n^{1+\alpha}(y_{n-1}(\theta)-y_n(\theta))\sup_{y\in[\frac34,y_{n-1}(\theta)]}|v(y,\theta)-J(\theta)|\\&\ll\sup_{y\in[\frac34,y_{n-1}(\theta)]}|v(y,\theta)-J(\theta)|\to0\ a.e.\end{aligned}$$

by the definition of $J(\theta)$. Hence $B_n(\theta)\to0$ a.e. completing the proof of (4.6).

The estimate for $\mu_Y(\varphi>n)$ follows immediately. ∎

4.4 Estimates for $\|F_n\|_{\mathrm{BV}}$

Define the family of operators $\widehat{F}_n:\mathrm{BV}(Y)\to\mathrm{BV}(Y)$, $n\ge1$, given by $\widehat{F}_nv=\widehat{F}(1_{\{\varphi=n\}}v)$.

Lemma 4.12 $\|\widehat{F}_n\|_{\mathrm{BV}} = O(n^{-(1+\alpha)})$.

Proof Recall that $\|\widehat{F}_n v\|_{\mathrm{BV}} = |\widehat{F}_n v|_1 + \mathrm{Var}(\widehat{F}_n v)$. Now,

$$|\widehat{F}_n v|_1 = |\widehat{F}(1_{\{\varphi=n\}}v)|_1 \le |1_{\{\varphi=n\}}v|_1 \ll n^{-(1+\alpha)}\|v\|_{\mathrm{BV}}$$

by Proposition 4.11.

Next, we estimate $\mathrm{Var}(\widehat{F}_n v)$. By Corollary 4.2, it suffices to do this for $v \in C^1$. Adapting the calculations in the proof of Lemma 4.9, we have $\mathrm{Var}(\widehat{F}_n v) \le I_1 + I_2 + I_3$, where

$$I_1 = \sum_{\varphi(a)=n} |1_a(\nabla g)(DF_a)^{-1}JF|_\infty \int_a |v|,$$

$$I_2 = \sum_{\varphi(a)=n} |(DF_a)^{-1}|_\infty \int_a |\nabla v|, \qquad I_3 = \sum_{\varphi(a)=n} \int_{\partial a} |v||D\partial F_a/JF_a|.$$

We consider partition elements of the form $a = Y_{n,j}$. By Corollary 2.14, $|1_a(\nabla g)(DF_a)^{-1}JF|_\infty = O(1)$. Hence $I_1 \ll \int_{\{\varphi=n\}} |v| \ll n^{-(1+\alpha)}\|v\|_{\mathrm{BV}}$ by Proposition 4.11.

By (2.7), $|(DF_a)^{-1}|_\infty \ll n^{-(1+\alpha)}$. Hence $I_2 \ll n^{-(1+\alpha)} \int_{\{\varphi=n\}} |\nabla v| \le n^{-(1+\alpha)}\|v\|_{\mathrm{BV}}$.

Finally, by Lemma 4.5 and Proposition 4.11, $I_3 \ll \int_{\{\varphi=n\}} |v| + n^{-(1+\alpha)} \int_{\{\varphi=n\}} |\nabla v| \ll n^{-(1+\alpha)}\|v\|_{\mathrm{BV}}$. ∎

5 Lower bounds on decay of correlations and infinite measure mixing

By Proposition 4.11, the return time φ is integrable if and only if $\gamma < 1$. In this section, we establish lower bounds on decay of correlations for a class of observables supported on Y when $\gamma < 1$. Also, for $\gamma \ge 1$, we obtain results on mixing for the infinite measure μ_Y for the same class of observables.

By Lemma 3.2, the invariant density $h_Y = d\mu_Y/d\,\mathrm{Leb}$ is bounded above and below, so the L^p spaces with respect to μ_Y and Leb are identical and we can just write $L^p(Y)$. We have $\mathrm{BV}(Y) \subset L^2(Y)$ since the domain Y is two-dimensional.

The transfer operator R corresponding to the F-invariant measure μ_Y is given by $Rv = h_Y^{-1}\widehat{F}(h_Y v)$. Again, we consider the twisted transfer operators $R(\omega)v = R(\omega^\varphi v) = h_Y^{-1}\widehat{F}(\omega)(h_Y v)$. These act naturally on the Banach space $\mathcal{B}(Y) = h_Y^{-1}\mathrm{BV}(Y)$ which consists of functions $v : Y \to \mathbb{R}$ such that $h_Y v \in \mathrm{BV}(Y)$ with norm $\|v\|_{\mathcal{B}} = \|h_Y v\|_{\mathrm{BV}}$. Similarly we define $R_n : \mathcal{B} \to \mathcal{B}$, $n \ge 1$, given by $R_n v = R(1_{\{\varphi=n\}}v) = h_Y^{-1}\widehat{F}_n(h_Y v)$.

Recall that $h_Y \in \mathrm{BV}_\infty(Y) = \mathrm{BV}(Y) \cap L^\infty(Y)$.

Proposition 5.1 $\mathrm{BV}_\infty(Y) \subset \mathcal{B}(Y) \subset L^2(Y)$.

Proof Let $v \in \mathcal{B}(Y)$. Then $\int v^2 \, d\,\mathrm{Leb} \le |h_Y^{-2}|_\infty \int (h_Y v)^2 \, d\,\mathrm{Leb} \ll \|h_Y v\|_{\mathrm{BV}}^2 = \|v\|_{\mathcal{B}}^2$ establishing the second inclusion.

For the first inclusion, let $v \in \mathrm{BV}_\infty(Y)$. By Corollary 4.4, $h_Y v \in \mathrm{BV}_\infty(Y) \subset \mathrm{BV}(Y)$, so $v \in \mathcal{B}(Y)$. ∎

Corollary 5.2 *Consider the operators* $R(\omega) : \mathcal{B}(Y) \to \mathcal{B}(Y)$.

(i) 1 *is a simple isolated eigenvalue in the spectrum of* R.

(ii) $1 \notin \operatorname{spec} R(\omega)$ *for all* $\omega \in \overline{\mathbb{D}} \setminus \{1\}$.

(iii) $\|R_n\|_{\mathcal{B}} = O(n^{-(1+\alpha)})$.

Proof Multiplication by h_Y^{-1} is an isomorphism from $\mathrm{BV}(Y) \to \mathcal{B}(Y)$ that conjugates $\widehat{F}(\omega)$ to $R(\omega)$. Hence $R(\omega)$ inherits properties of $\widehat{F}(\omega)$ in Corollary 4.10. Similarly, R_n inherits properties of $\widehat{F}_n$ in Lemma 4.12. ∎

For observables v, w supported in Y, we can write $\int v\, w \circ f^n \, d\mu = \int T_n v \, w \, d\mu$ where $T_n = 1_Y L^n 1_Y$ and L is the transfer operator for f. Defining $T(\omega) = \sum_{n=0}^{\infty} T_n \omega^n$, we recall from [58, Proposition 1] the operator renewal equation $T(\omega) = (I - R(\omega))^{-1}$.

Proof of Theorem 1.1(b,c) Recall that F is the first return map to Y so $\mu_Y = \mu|_Y/\mu(Y)$. By Corollary 5.2, we have verified the assumptions of Gouëzel [28]. Let $v \in \mathcal{B}(Y)$ and $w \in L^2(Y)$ [5]. By [28, Theorem 1.1],

$$\Big|\rho(n) - \mu(Y) \sum_{j>n} \mu_Y(\varphi > j) \int v \, d\mu \int w \, d\mu\Big| \ll E(n) \|v\|_{\mathcal{B}} |w|_2$$

But $\sum_{j>n} \mu_Y(\varphi > j) \sim \gamma(\alpha-1)^{-1} c_2 n^{-(\alpha-1)}$ with c_2 as given in Proposition 4.11. Part (b) follows with $c = \mu(Y)\gamma(\alpha-1)^{-1}c_2$.

Part (c) is a consequence of [28, Theorem 1.2]. ∎

Proof of Theorem 1.5 By Corollary 5.2 we have verified the assumptions in Gouëzel [33, Theorem 1.4] and Melbourne & Terhesiu [52, Theorem 2.1]. Let $d_\gamma = \frac{1}{\pi} \sin \alpha\pi$. Let c_2 be the constant in Proposition 4.11 and define $c_4 = \mu(Y)\gamma c_2$. For $\gamma > 1$, we obtain

$$c_4 n^{1-\alpha} \int v \, w \circ f^n \, d\mu \sim d_\gamma \int v \, d\mu \int w \, d\mu,$$

for all $v \in \mathcal{B}(Y)$ and $w \in L^2(Y)$. For $\gamma = 1$ the asymptotic holds with $d_\gamma = 1$ and $n^{1-\alpha}$ replaced by $\log n$. Part (a) follows with $c = d_\gamma / c_4$.

Part (b) is a consequence of [52, Theorem 2.2(c)]. ∎

[5] In general, we require w in L^2 since Proposition 5.1 only gives $\mathcal{B} \subset L^2$. For $v \in \mathrm{BV}_\infty$, we can take $w \in L^1$.

6 Convergence to stable laws and Lévy processes

In this section, we prove Theorem 1.3. Set $\alpha = \frac{1}{\gamma} \in (1,2)$ and let G_α denote the totally skewed α-stable law in Theorem 1.3. Define $\sigma = \frac{1}{4}c'\gamma\int_{\mathbb{T}} h_Y(\frac{3}{4}+,\theta)\,d\theta\,\Gamma(1-\alpha)\cos\frac{\alpha\pi}{2}$ where c' is as in Proposition 2.4.

Proposition 6.1 $n^{-1/\alpha}\sum_{j=0}^{n-1}(\varphi\circ f^j - \int_Y \varphi\,d\mu_Y) \to_d \sigma^{1/\alpha}G_\alpha$.

Proof We verify the conditions stated in Appendix C. Taking $\omega = 1$ in Corollary 5.2, we see that $R:\mathcal{B}(Y)\to\mathcal{B}(Y)$ satisfies the required spectral gap condition.

Let $\psi = \varphi - \int_Y \varphi\,d\mu_Y$. This is an L^1 function with mean zero. Clearly ψ is bounded below. By Proposition 4.11, $\mu_Y(\psi > x) \sim \sigma_1 x^{-\alpha}$ where $\sigma_1 = \frac{1}{4}c'\gamma\int_{\mathbb{T}} h_Y(\frac{3}{4}+,\theta)\,d\theta$. Hence condition (C.1) is satisfied (with $\sigma_2 = 0$).

Define $R_t = R(e^{it})$ for $t \in \mathbb{R}$. By Section 5, R_t is a bounded linear operator on $\mathcal{B}(Y)$ for all t. Note that $R_t = \sum_{n=1}^\infty R_n e^{int}$. It follows from Corollary 5.2(iii) that $\sum_{n=1}^\infty n\|R_n\|_{\mathcal{B}} < \infty$ so $t\mapsto R_t$ is C^1. In particular, $\|R_t\|_{\mathcal{B}} = O(|t|)$. The result now follows from Theorem C.1 (with $\beta = 1$). ∎

Proposition 6.2 *Let $v : X \to \mathbb{R}$ be Hölder. Suppose that $v(0,\theta) \equiv I$ for some $I \in \mathbb{R}$. Define $V = \sum_{\ell=0}^{\varphi-1} v\circ f^\ell$. Then $V - I\varphi \in L^p(Y)$ for some $p > \alpha$.*

Proof Let $\eta \in (0,1)$ be the Hölder exponent for v and suppose without loss that $\eta < \gamma$. Set $\delta = \eta\alpha \in (0,1)$. Since $\varphi \in L^q(Y)$ for all $q < \alpha$, it suffices to show that $V - I\varphi = O(\varphi^{1-\delta})$.

Let $(y,\theta) \in Y_{n,j}$. Then

$$V(y,\theta) - I\varphi(y,\theta) = \sum_{\ell=0}^{n-1} v(f^\ell(y,\theta)) - nI = \sum_{\ell=0}^{n-1}(v(f^\ell(y,\theta) - v(f^\ell(0,\theta)).$$

Write $f^\ell(y,\theta) = (y_\ell,\theta_\ell)$. Then $f^\ell(0,\theta) = (0,\theta_\ell)$, so $|V(y,\theta) - I\varphi(y,\theta)| \le |v|_\eta \sum_{\ell=0}^{n-1} y_\ell^\eta$. By Proposition 2.4, $y_\ell \ll (n-\ell)^{-\alpha}$ for $\ell = 0,\dots,n-1$, so $|V - I\varphi| \ll |v|_\eta n^{1-\eta\alpha} \ll \varphi^{1-\delta}$ as required. ∎

Proof of Theorem 1.3 First we prove the result under the additional assumption that $v(0,\theta)$ is independent of θ. Evidently this constant value is I_v, so Proposition 6.2 implies that $V - I_v\varphi \in L^p(Y)$ for some $p > \alpha$. This is [54, condition (3.2)]. Also, [54, condition (3.1)] follows from Proposition 6.1. Hence convergence to the desired stable law follows from [54, Theorem 3.1] with $c = \bar{\varphi}^{-1/\alpha} I_v\sigma$.

Define $M_1 = \max_{1\le\ell'\le\ell\le\varphi}(v_{\ell'} - v_\ell) \wedge \max_{1\le\ell'\le\ell\le\varphi}(v_\ell - v_{\ell'})$ where $v_\ell = \sum_{j=0}^{\ell-1} v\circ f^j$. Suppose for definiteness that $I_v > 0$ (the case I_v is treated similarly). The calculation in Proposition 6.2 shows that $v_\ell = I_v\ell + O(\varphi^{1-\delta})$ for all $0 \le \ell \le \varphi$. Hence

$$0 \le M_1 \le \max_{1\le\ell'\le\ell\le\varphi}(v_{\ell'} - v_\ell) = \max_{1\le\ell'\le\ell\le\varphi} I_v(\ell' - \ell) + O(\varphi^{1-\delta}).$$

Since $I_v > 0$ it follows that $M_1 \ll \varphi^{1-\delta}$. By [54, Proposition 3.5], $n^{-1/\alpha} \max_{j\le n} M_1 \circ F^j \to_p 0$ on (Y, μ_Y). Hence convergence to the desired Lévy process follows from [54, Theorem 3.2(a)].

Finally, we relax the additional assumption on v. Write $v = v' + v''$ where $v''(y,\theta) = v(0,\theta) - I_v$. We have $W_n = W_n' + W_n''$ where

$$W_n'(t) = n^{-1/\alpha} \textstyle\sum_{j=0}^{[nt]-1} v' \circ f^j, \qquad W_n''(t) = n^{-1/\alpha} \textstyle\sum_{j=0}^{[nt]-1} v'' \circ f^j.$$

Note that v'', and hence v', is Hölder and mean zero. Moreover, $v'(0,\theta) \equiv I_v$, so $W_n' \to_w W$ in $(D[0,\infty), \mathcal{M}_1)$. Also, $u(\theta) = v''(y,\theta)$ is a Hölder mean zero observable for the uniformly expanding map $f_2 : \mathbb{T} \to \mathbb{T}$, so $n^{-1/2} \sum_{j=0}^{[nt]-1} u \circ f_2^j$ converges weakly to Brownian motion in the uniform topology (see for example [36, Theorem 5] which establishes the ASIP and hence the weak convergence). Hence $n^{-1/2} \sum_{j=0}^{[nt]-1} v'' \circ f^j$ converges weakly, so $W_n'' \to_w 0$. The result follows. ∎

A Construction of the Gibbs-Markov map G

This section is devoted to the proof of Lemma 3.1. The main step is to verify the hypotheses of Theorem 3 of [23]. This is done using Theorem A.1 below.

Recall that $Y \subset \mathbb{R}^2$ is endowed with the Euclidean metric $|(y_1,\theta_1) - (y_2,\theta_2)| = ((y_1-y_2)^2 + (\theta_1-\theta_2)^2)^{1/2}$. For $x \in \mathbb{R}^2$ and $A \subset \mathbb{R}^2$, let $\mathbf{d}(x,A) = \inf_{y\in A} |x-y|$ (with $\mathbf{d}(x,A) = \infty$ if $A = \emptyset$). Given $A \subset \mathbb{R}^2$, $\varepsilon > 0$, define

$$\partial_\varepsilon A = \{x \in A : \mathbf{d}(x, \partial A) \le \varepsilon\} \subset A,$$

where ∂A is the boundary of A as a subset of $\mathbb{R}^2$.

We prove that the first return map $F : Y \to Y$ satisfies the following properties:

Theorem A.1 *Let $\Lambda \in (\frac14, \frac13)$. There exists $\varepsilon_0 \in (0,1)$ and $C > 0$ such that the following hold:*

Uniform expansion: $|F_a^{-1} z_1 - F_a^{-1} z_2| \le \Lambda |z_1 - z_2|$ *for all $z_1, z_2 \in a$ with $|z_1 - z_2| < \varepsilon_0$ and all $a \in \alpha^Y$.*

Bounded distortion: $(JF_a^{-1})(z_1) \le e^{C|z_1-z_2|}(JF_a^{-1})(z_2)$ *for all $z_1, z_2 \in a$ and all $a \in \alpha^Y$.*

Controlled complexity: *For every open set $I \subset Y$ with* $\operatorname{diam} I \le \varepsilon_0$ *and all $\varepsilon < \varepsilon_0$,*

$$\sum_{a\in\alpha^Y} \frac{\operatorname{Leb}(F_a^{-1}(\partial_\varepsilon F(I\cap a)) \setminus \partial_{\varepsilon\Lambda} I)}{\operatorname{Leb} \partial_{\varepsilon\Lambda} I} < \Lambda^{-1} - 1.$$

Set Z: *For all $\delta > 0$ sufficiently small, there exist rectangles $Z, Z' \subset Y$ with* $\operatorname{Leb} Z < \operatorname{Leb} Z'$ *and* $\operatorname{diam} Z' \le \delta$ *such that*

$$F(Z \cap \mathrm{Int}\{\varphi = n\}) = Y \quad \textit{for sufficiently large } n; \tag{A.1}$$

and, there exists $a_1, a_2 \in \alpha^Y$ *such that* $a_i \subset Z$ *and* $Fa_i \supset Z'$ *for* $i = 1, 2$ *and*

$$\varphi|_{a_2} - \varphi|_{a_1} = 1. \tag{A.2}$$

Proof of Lemma 3.1 Theorem A.1 implies in particular that we have verified the hypotheses of [23, Theorem 3]. This guarantees the existence of the desired refinement α_1^Y, the subset Z (as given in Theorem A.1), the return time $\rho : Z \to \mathbb{Z}^+$ constant on elements of $\alpha^Z = \{a' \in \alpha_1^Y : a' \subset Z\}$ and the induced map $G = F^\rho : Z \to Z$. Moreover, conditions (a) and (b) of Lemma 3.1 follow directly from [23, Theorem 3](a),(c).

In addition, [23, Theorem 3](b) states that $\mathrm{Leb}(\{\rho = 1\} \cap a_i) > 0$ for $i = 1, 2$. This combined with (A.2) guarantees that Lemma 3.1(c) holds. Finally, Lemma 3.1(d) follows from (A.1). ∎

In the next four subsections, we verify the four properties listed in Theorem A.1.

A.1 Uniform expansion

By Proposition 2.1, $|DF_a^{-1}| \le \frac{1}{4}$ on Fa for all $a \in \alpha^Y$.

Lemma A.2 *For every* $\delta > 0$ *there exists* $\varepsilon_0 > 0$ *such that for all* $z_1, z_2 \in Fa$ *with* $|z_1 - z_2| < \varepsilon_0$ *and all* $a \in \alpha^Y$*, there exists a path* $\gamma : [0, 1] \to \mathbb{R}^2$ *contained in* Fa*, joining* z_1 *and* z_2*, and having length bounded by* $(1 + \delta)|z_1 - z_2|$.

Proof The boundary of Fa is a rectangle except that its right boundary is a C^1 curve which we denote by ψ. Denote the line segment joining z_1 and z_2 by S. If S lies in Fa, then take γ to be the path corresponding to this line segment. If not, then S intersects the boundary of Fa. Let p_1, p_2 be the points of intersection closest to z_1, z_2, respectively. Define γ to be the path corresponding to starting at z_1, travelling on S until p_1, then travelling on the boundary of Fa until p_2 and then continuing on S to z_2. Since ψ is smooth, the length of γ can be made arbitrarily close to the length of S by choosing ε_0 sufficiently small. (The path γ may not be entirely contained in Fa, but a small translation of it will be entirely inside Fa.) ∎

Choose δ so that $\frac{1}{4}(1 + \delta) < \Lambda$, and fix ε_0 as in Lemma A.2. Let $z_1, z_2 \in Fa$ with $|z_1 - z_2| < \varepsilon_0$ and choose γ as in Lemma A.2.

Now, $F_a^{-1} z_2 - F_a^{-1} z_1 = (F_a^{-1} \circ \gamma)(1) - (F_a^{-1} \circ \gamma)(0) = \int_0^1 D(F_a^{-1} \circ \gamma)(t) dt$, so by Lemma A.2,

$$\begin{aligned} |F_a^{-1} z_1 - F_a^{-1} z_2| &\le \int_0^1 |(DF_a^{-1})(\gamma(t))\, \gamma'(t)|\, dt \le \sup_{t \in [0,1]} |DF_a^{-1}(\gamma(t))| \int_0^1 |\gamma'(t)|\, dt \\ &\le \tfrac{1}{4}(1 + \delta)|z_1 - z_2| < \Lambda |z_1 - z_2|, \end{aligned}$$

as required.

A.2 Bounded distortion

By Corollary 2.12, there exists $C > 0$ such that

$$|1/JF(y_1,\theta_1) - 1/JF(y_2,\theta_2)| \le C \inf_a (1/JF)|F(y_1,\theta_1) - F(y_2,\theta_2)|$$

for all (y_1,θ_1), $(y_2,\theta_2) \in a$ and all a. Writing $z_1 = F(y_1,\theta_1)$, $z_2 = F(y_2,\theta_2)$, it follows that

$$\frac{(JF_a^{-1})(z_1)}{(JF_a^{-1})(z_2)} \le 1 + C|z_1 - z_2| \le e^{C|z_1-z_2|},$$

yielding the desired distortion condition.

A.3 Controlled Complexity

For the proof of this property we need a generalization of [8, Sublemma C.1], which appears below as Proposition A.5.

Recall that $\partial_\varepsilon A$ is defined as a subset of A. We also define $\tilde{\partial}_\varepsilon A = \{x \in \mathbb{R}^2 : \mathbf{d}(x, \partial A) \le \varepsilon\}$. (Hence $\partial_\varepsilon A = \tilde{\partial}_\varepsilon A \cap A$).

Let us recall [8, Sublemma C.1] in a form that suffices for our purposes. We refer to its proof briefly at the end of the proof of Lemma A.4.

Lemma A.3 (Sublemma C.1 of [8]) *Suppose I is a non-empty measurable bounded subset of the plane and E is a straight line cutting I into left and right parts I_l and I_r. Then for all $\varepsilon \ge 0$, $0 \le \xi \le 1$,*

$$\begin{aligned}\mathrm{Leb}(\{x \in I_l : \mathbf{d}(x,E) \le \varepsilon\xi\} \setminus \{x \in I : \mathbf{d}(x,\partial I) \le \varepsilon\}) \le \\ \xi\, \mathrm{Leb}\{x \in I_r : \mathbf{d}(x,\partial I) \le \varepsilon\}.\end{aligned} \tag{A.3}$$

In Proposition A.5 we generalize to the case where E is the graph of a Lipschitz function, but first we prove a lemma which is similar in flavour but applies to segments which may or may not intersect I. This lemma would follow from the one above if $0 \le \xi \le 1$ and S (taking the place of E) were a hyperplane in $\mathbb{R}^2$ cutting through I.

Given a straight line segment $S \in \mathbb{R}^2$ and $x \in \mathbb{R}^2$, define $\mathbf{d}^\perp$ as follows. Suppose $x \in \mathbb{R}^2$. If there exists a line that passes through x, intersects S and is perpendicular to S, then $\mathbf{d}^\perp(x,S) = \mathbf{d}(x,S)$. If not, then $\mathbf{d}^\perp(x,S) = \infty$. If S is the graph of a piecewise constant function, then one can define $\mathbf{d}^\perp(x,S)$ similarly.

Lemma A.4 *Suppose I is a measurable bounded subset of the plane and S is a straight line segment in the plane. Then for all $\varepsilon \ge 0$, $\xi \ge 0$,*

$$\begin{aligned}\mathrm{Leb}(\{x \in I : \mathbf{d}^\perp(x,S) \le \varepsilon\xi\} \setminus \{x \in I : \mathbf{d}(x,\partial I) \le \varepsilon\}) \le \\ \xi\, \mathrm{Leb}\{x \in I : \mathbf{d}(x,\partial I) \le \varepsilon\}.\end{aligned} \tag{A.4}$$

Proof Fix $\varepsilon \geq 0$, $\xi \geq 0$. Let

$$A = \{x \in I : \mathbf{d}^{\perp}(x, S) \leq \varepsilon\xi\} \setminus B, \text{ where } B = \{x \in I : \mathbf{d}(x, \partial I) \leq \varepsilon\}.$$

We show that $\operatorname{Leb} A \leq \xi \operatorname{Leb} B$.

Let e_z be the line perpendicular to S at the point $z \in S$ and let $A_z = A \cap e_z$ and $B_z = B \cap e_z$. Given $\varepsilon' > 0$ and an interval $J_{\varepsilon'}$ of length ε' inside e_z, denote $A_z(\varepsilon') = A_z \cap J_{\varepsilon'}$. Points of $A_z(\varepsilon)$ are by definition at least distance ε from ∂I so there exists a translate of $A_z(\varepsilon)$ along e_z that lies in B_z. It follows from translation invariance of Lebesgue measure Leb_z on e_z that $\operatorname{Leb}_z A_z(\varepsilon) \leq \operatorname{Leb}_z B_z$.

Let us write $\xi = \lfloor \xi \rfloor + \{\xi\}$, where $\{\xi\}$ denotes the fractional part of ξ. Since A_z can be partitioned by $\lfloor \xi \rfloor$ sets of the form $A_z(\varepsilon)$ plus one remainder set of the form $A_z(\{\xi\}\varepsilon)$, it follows that

$$\operatorname{Leb}_z A_z \leq \lfloor \xi \rfloor \operatorname{Leb}_z(B_z) + \operatorname{Leb}_z A_z(\varepsilon\{\xi\}). \tag{A.5}$$

Now we show that $\operatorname{Leb}_z A_z(\varepsilon\{\xi\}) \leq \{\xi\} \operatorname{Leb}_z B_z$ bounding the second term of (A.5). If $z \in S \setminus I$, then $A_z(\varepsilon\{\xi\}) = \emptyset$ because $\{\xi\} < 1$, so we are done. Otherwise, if $z \in S \cap I$, the claim follows directly from the proof of Lemma A.3 given in [8, p.1364] because $0 \leq \{\xi\} < 1$.

We have proved that $\operatorname{Leb}_z A_z \leq \xi \operatorname{Leb}_z B_z$. Integrating over $z \in S$ with respect to Lebesgue measure on S, we obtain $\operatorname{Leb} A \leq \xi \operatorname{Leb} B$ as required. ∎

Proposition A.5 *Suppose I is a measurable bounded subset of the plane and E is the graph of an L-Lipschitz function in the plane. Then for all $\varepsilon \geq 0$, $0 \leq \bar{\xi} \leq 1$*

$$\begin{aligned}\operatorname{Leb}(\{x \in I : \mathbf{d}(x, E) \leq \varepsilon\bar{\xi}\} \setminus \{x \in I : \mathbf{d}(x, \partial I) \leq \varepsilon\}) \leq \\ \bar{\xi}(1 + L) \operatorname{Leb}\{x \in I : \mathbf{d}(x, \partial I) \leq \varepsilon\}.\end{aligned}$$

In other words, $\operatorname{Leb}((I \cap \tilde{\partial}_{\varepsilon\bar{\xi}} E) \setminus \partial_\varepsilon I) \leq \bar{\xi}(1 + L) \operatorname{Leb} \partial_\varepsilon I$.

Proof Suppose E is the graph of the Lipschitz function $\psi : \mathbb{R} \to \mathbb{R}$. By a rotation of I and E, we can suppose that the domain of ψ is the horizontal axis.

Fix $\varepsilon \geq 0$, $0 \leq \bar{\xi} \leq 1$. For $t > 0$, let $\{A_j\}$ be a partition of the horizontal axis $\mathbb{R}$ into intervals of length $t\varepsilon$. Define $g_t : \mathbb{R} \to \mathbb{R}$, $g_t|_{A_j} \equiv (\operatorname{Leb} A_j)^{-1} \int_{A_j} \psi$, and denote $S = \operatorname{graph} g_t$. Note that $|\psi - g_t|_\infty \leq Lt\varepsilon$.

We claim that

$$\text{If } \mathbf{d}(x, E) \leq \varepsilon\bar{\xi}, \text{ then } \mathbf{d}^{\perp}(x, E) \leq \varepsilon\bar{\xi}(1 + L).$$

Here $\mathbf{d}^{\perp}(x, E)$ means the vertical distance from x to E. Since $|\psi - g_t|_\infty \leq Lt\varepsilon$, it follows from the claim that

$$\{x \in I : \mathbf{d}(x, E) \leq \varepsilon\bar{\xi}\} \subset \{x \in I : \mathbf{d}^{\perp}(x, S) \leq \varepsilon\xi_t\},$$

where $\xi_t = \bar{\xi}(1+L) + Lt$.

Now applying Lemma A.4 with $\xi = \xi_t$ on constant pieces graph$(g_t|_{A_j})$ of S separately and adding the contributions, we get

$$\begin{aligned}\operatorname{Leb}(\{x \in I : \mathbf{d}(x,E) \le \varepsilon\bar{\xi}\} \setminus \partial_\varepsilon I) &\le \operatorname{Leb}(\{x \in I : \mathbf{d}^\perp(x,S) \le \varepsilon\xi_t\} \setminus \partial_\varepsilon I) \\ &\le \xi_t \operatorname{Leb} \partial_\varepsilon I = (\bar{\xi}(1+L) + Lt) \operatorname{Leb} \partial_\varepsilon I.\end{aligned}$$

Since $t > 0$ is arbitrary, we obtain the desired result.

It remains to prove the claim. Write $x = (x_1, x_2)$ and choose $z = (z_1, z_2) \in E$ with $|x - z| \le \varepsilon\bar{\xi}$. Let $v = (v_1, v_2) \in E$ with $v_1 = x_1$. Then

$$\begin{aligned}\mathbf{d}^\perp(x,E) = |x_2 - v_2| &\le |x_2 - z_2| + |v_2 - z_2| \le |x_2 - z_2| + L|v_1 - z_1| \\ &= |x_2 - z_2| + L|x_1 - z_1| \le (1+L)|x-z| \le \varepsilon\bar{\xi}(1+L),\end{aligned}$$

as required. ∎

Verification of controlled complexity Set $\Lambda_n = \sup_j |DF_{n,j}^{-1}|$ and note that $\Lambda_n \le \frac{1}{4} < \Lambda$. Also, $\Lambda_n = O(n^{-(1+\alpha)})$ by (2.7). Recall that $Y_n = \bigcup_{j=1}^{4^n} Y_{n,j}$ is bounded by two flat horizontal sides and two smooth vertical curves. By Proposition 2.4, the Lipschitz constants corresponding to the vertical curves are bounded by some $L_0 > 0$.

We choose $n_0 \ge 1$ sufficiently large so that $\sum_{n=n_0}^\infty \frac{\Lambda_n}{\Lambda}(1+L_0) \le \frac{1}{8}$ and then shrink ε_0 if needed so that if $I \subset Y$ has $\operatorname{diam} I \le \varepsilon_0$ then at least one of the following holds:

(i) $I \subset \bigcup_{n=1}^{n_0} Y_n$ and $I \cap \bigcup_{n=1}^{n_0} \sum_{j=1}^{4^n} \partial Y_{n,j}$ consists of at most one horizontal curve H and one vertical curve V.

(ii) $I \subset \bigcup_{n=n_0}^\infty Y_n$.

In case (i), H is flat and V is smooth. Recall that $\Lambda < \frac{1}{3}$. Shrinking ε_0 further, we can suppose that $I \cap V$ is the graph of a function with Lipschitz constant L_1 satisfying $3 + L_1 < \Lambda^{-1}$.

Let $I \subset Y$ be an open subset with $\operatorname{diam} I \le \varepsilon_0$. Note that

$$F_{n,j}^{-1}(\partial_\varepsilon F(I \cap Y_{n,j})) \setminus \partial_{\varepsilon\Lambda} I \subset (I \cap \partial_{\varepsilon\Lambda_n} Y_{n,j}) \setminus \partial_{\varepsilon\Lambda} I.$$

Recall that $\partial Y_{n,j} = H_{n,j} \cup V_{n,j}$ where $H_{n,j}$ consists of two flat horizontal edges and $V_{n,j}$ consists of two vertical curves. Hence

$$F_{n,j}^{-1}(\partial_\varepsilon F(I \cap Y_{n,j})) \setminus \partial_{\varepsilon\Lambda} I \subset \{(I \cap \tilde{\partial}_{\varepsilon\Lambda_n} H_{n,j}) \setminus \partial_{\varepsilon\Lambda} I\} \cup \{(I \cap \tilde{\partial}_{\varepsilon\Lambda_n} V_{n,j}) \setminus \partial_{\varepsilon\Lambda} I\}. \tag{A.6}$$

Case (i). Since the only intersections are with H and V, (A.6) simplifies to

$$\bigcup_a F_a^{-1}(\partial_\varepsilon F(I \cap a)) \setminus \partial_{\varepsilon\Lambda} I \subset \{(I \cap \tilde{\partial}_{\varepsilon\Lambda} H) \setminus \partial_{\varepsilon\Lambda} I\} \cup \{(I \cap \tilde{\partial}_{\varepsilon\Lambda} V) \setminus \partial_{\varepsilon\Lambda} I\}.$$

By Proposition A.5 (taking $\bar{\xi} = 1$ and replacing ε by $\varepsilon\Lambda$),

$$\mathrm{Leb}((I \cap \tilde{\partial}_{\varepsilon\Lambda} V) \setminus \partial_{\varepsilon\Lambda} I) \leq (1 + L_1)\,\mathrm{Leb}\,\partial_{\varepsilon\Lambda} I.$$

Similarly, $\mathrm{Leb}((I \cap \tilde{\partial}_{\varepsilon\Lambda} H) \setminus \partial_{\varepsilon\Lambda} I) \leq \mathrm{Leb}\,\partial_{\varepsilon\Lambda} I$. Hence

$$\sum_a \frac{\mathrm{Leb}(F_a^{-1}(\partial_\varepsilon F(I \cap a)) \setminus \partial_{\varepsilon\Lambda} I)}{\mathrm{Leb}\,\partial_{\varepsilon\Lambda} I} \leq 2 + L_1 < \Lambda^{-1} - 1.$$

Case (ii). By (A.6).

$$\bigcup_a F_a^{-1}(\partial_\varepsilon F(I \cap a)) \setminus \partial_{\varepsilon\Lambda} I \subset S_H + S_V,$$

where

$$S_H = \bigcup_{n=n_0}^{\infty} \bigcup_{j=1}^{4^n} \{(I \cap \tilde{\partial}_{\varepsilon\Lambda_n} H_{n,j}) \setminus \partial_{\varepsilon\Lambda} I\}, \qquad S_V = \bigcup_{n=n_0}^{\infty} \bigcup_{j=1}^{4^n} \{(I \cap \tilde{\partial}_{\varepsilon\Lambda_n} V_{n,j}) \setminus \partial_{\varepsilon\Lambda} I\}.$$

We estimate S_H and S_V separately. The curve $V_n = \bigcup_{j=1}^{4^n} V_{j,n}$ is smooth with Lipschitz constant bounded by L_0, and

$$S_V = \bigcup_{n=n_0}^{\infty} \mathrm{Leb}\{(I \cap \tilde{\partial}_{\varepsilon\Lambda_n} V_n) \setminus \partial_{\varepsilon\Lambda} I\}.$$

By Proposition A.5 (taking $\bar{\xi} = \Lambda_n/\Lambda$ and replacing ε by $\varepsilon\Lambda$),

$$\mathrm{Leb}((I \cap \tilde{\partial}_{\varepsilon\Lambda_n} V_n) \setminus \partial_{\varepsilon\Lambda} I) \leq \frac{2\Lambda_n}{\Lambda}(1 + L_0)\,\mathrm{Leb}\,\partial_{\varepsilon\Lambda} I.$$

Hence, by the choice of n_0,

$$\frac{\mathrm{Leb}\,S_V}{\mathrm{Leb}\,\partial_{\varepsilon\Lambda} I} \leq \sum_{n=n_0}^{\infty} \frac{2\Lambda_n}{\Lambda}(1 + L_0) \leq \frac{1}{4}.$$

Shrinking ε_0 further if necessary, it follows from the skew-product structure of F (where vertical distances are contracted by $4^{-\varphi}$) that Λ_n can be improved to 4^{-n} in the formula for S_H leading to the estimate $\frac{\mathrm{Leb}\,S_H}{\mathrm{Leb}\,\partial_{\varepsilon\Lambda} I} \leq \frac{1}{4}$.

Hence again we obtain the desired complexity bound

$$\sum_a \frac{\mathrm{Leb}(F_a^{-1}(\partial_\varepsilon F(I \cap a)) \setminus \partial_{\varepsilon\Lambda} I)}{\mathrm{Leb}\,\partial_{\varepsilon\Lambda} I} \leq \frac{1}{2} < \Lambda^{-1} - 1.$$

A.4 Set Z

The construction of Z and Z' proceeds as follows: Recall that the partition elements in α^Y accumulate on the left vertical side $\{\frac{3}{4}\} \times \mathbb{T}$ of Y. Let $c = 1/100$ and let S_0, S_1 denote open squares with side lengths $c\delta$ and $2c\delta$, respectively, and centred at $l_0 = (\frac{3}{4}, 0)$. Let $Z = S_0 \cap Y$ and $Z' = S_1 \cap Y$. It is immediate that $Z' \supset Z$, $\mathrm{Leb}\, Z < \mathrm{Leb}\, Z'$, and $\mathrm{diam}\, Z' \leq \delta$.

Now Z is a rectangle with vertex l_0, and elements of α^Y accumulate at l_0 and shrink in diameter. Hence there exists $n_0 \geq 2$, $i_0 \geq 1$ such that $Y_{n,i} \subset Z$ for all $n \geq n_0$, $i = 1, \dots, 4\,(\mathrm{mod}\, i_n)$, where $i_n = 4^{n-n_0}(i_0 - 1)$. For $n \geq n_0$,

$$F(Z \cap \mathrm{Int}\{\varphi = n\}) \supset F\Big(\bigcup_{i=1}^4 Y_{n,i_n+i}\Big) = Y$$

proving (A.1). Setting $a_1 = Y_{n_0,i_{n_0}+1}$ and $a_2 = Y_{n_0+1,(i_{n_0+1})+1}$, we have $a_i \subset Z$ and $Fa_i \supset [\frac{3}{4}, \frac{15}{16}] \times \mathbb{T} \supset Z'$ for $i = 1, 2$. Moreover, $\varphi|_{a_1} = n_0$ and $\varphi|_{a_2} = n_0 + 1$, verifying (A.2).

B Boundary terms

In this appendix we recall some standard estimates for computing integrals around the boundary of "rectangular" domains. Consider a domain of the form

$$a = \{(y, \theta) \in \mathbb{R} \times [C, D] : \psi_1(\theta) \leq y \leq \psi_2(\theta)\},$$

where $\psi_1, \psi_2 : [C, D] \to \mathbb{R}$ are C^1 with $\psi_1 < \psi_2$. Define $M : [C, D] \to \mathbb{R}$, $M(\theta) = \max\{|\psi_1'(\theta)|, |\psi_2'(\theta)|\}$.

Theorem B.1 *Let $v : \mathbb{R}^2 \to \mathbb{R}$ be a C^1 function. Then*

$$\begin{aligned}\int_{\partial a} |v| \leq & 2\Big\{(D - C)^{-1} + \big|((1 + M^2)^{1/2} + 2M)/(\psi_2 - \psi_1)\big|_\infty\Big\}|1_a v|_1 \\ & + \sqrt{2}\big((1 + |M|_\infty^2)^{1/2} + |M|_\infty\big)|1_a \nabla v|_1.\end{aligned}$$

First, we consider the special case where a is a rectangle.

Proposition B.2 *Suppose that $a = [A, B] \times [C, D]$ is a rectangle and that v is C^1. Write $\partial a = H_a \cup V_a$ where H_a is the union of the two horizontal edges and V_a is the union of the two 'vertical' edges. Then*

$$\int_{H_a} |v| \leq 2(D - C)^{-1}|1_a v|_1 + |1_a \partial v/\partial\theta|_1, \quad \int_{V_a} |v| \leq 2(B - A)^{-1}|1_a v|_1 + |1_a \partial v/\partial y|_1.$$

Consequently, $\int_{\partial a} |v| \leq K|1_a v|_1 + \sqrt{2}|1_a \nabla v|_1$ where $K = 2(B - A)^{-1} + 2(D - C)^{-1}$.

Proof We give the details for the horizontal edges. The vertical edges are dealt with in the identical manner. The final statement follows from the fact that $|x| + |y| \le \sqrt{2}(x^2 + y^2)^{\frac{1}{2}}$.

Note that $\int_{H_a} |v| = \int_A^B |v(y, C)|\, dy + \int_A^B |v(y, D)|\, dy$. (Throughout, we work in the coordinate system (y, θ).) On the bottom edge,

$$\begin{aligned}\int_A^B |v(y, C)|\, dy &= [(D - C)/2)]^{-1} \int_A^B \Big\{ \int_C^{(C+D)/2} |v(y, C)|\, d\theta \Big\}\, dy \\ &\le [(D - C)/2)]^{-1} \Big\{ |1_{a_1} v|_1 + \int_A^B \Big\{ \int_C^{(C+D)/2} |v(y, \theta) - v(y, C)|\, d\theta \Big\}\, dy \Big\}\end{aligned}$$

where a_1 is the rectangle $[A, B] \times [C, (C + D)/2]$. But

$$|v(y, \theta) - v(y, C)| = \Big| \int_C^\theta \frac{\partial v}{\partial \theta}(y, \psi)\, d\psi \Big| \le \int_C^{(C+D)/2} \Big| \frac{\partial v}{\partial \theta}(y, \psi) \Big|\, d\psi,$$

and it follows that

$$\int_A^B |v(y, C)|\, dy \le 2(D - C)^{-1} |1_{a_1} v|_1 + |1_{a_1} \partial v / \partial \theta|_1.$$

The same estimate holds for the top edge $\int_A^B |v(y, D)|\, dy$, but with a_1 replaced by $a_2 = [A, B] \times [(C + D)/2, D]$. Since $a_1 \cup a_2 = a$ we obtain the required estimate for $\int_{H_a} |v|$. ∎

To prove Theorem B.1, we introduce the diffeomorphism $g : a \to [-1, 1] \times [C, D]$ given by

$$g(y, \theta) = \Big(\frac{2y - (\psi_2(\theta) + \psi_1(\theta))}{\psi_2(\theta) - \psi_1(\theta)}, \theta \Big), \quad g^{-1}(y, \theta) = (h(y, \theta), \theta),$$

where $h(y, \theta) = \frac{1}{2}(\psi_2(\theta) - \psi_1(\theta))y + \frac{1}{2}(\psi_2(\theta) + \psi_1(\theta))$. Note that $Jg = 1/\partial_y h = 2/(\psi_2 - \psi_1)$.

Proposition B.3 $|\partial_\theta h(y, \theta)| \le M(\theta) = \max\{|\psi_1'(\theta)|, |\psi_2'(\theta)|\}$ *for all* $y \in [-1, 1]$, $\theta \in [C, D]$.

Proof Write $\partial_\theta h = \frac{1}{2}(\psi_2' - \psi_1')y + \frac{1}{2}(\psi_2' + \psi_1')$. If $\psi_2' > \psi_1'$, then the maximum value m_2 and minimum value m_1 are obtained at $y = 1$ and $y = -1$ respectively yielding $m_2 = \psi_2'$ and $m_1 = \psi_1'$ respectively. The values are reversed if $\psi_2' < \psi_1'$. ∎

Vertical edges Let γ_1 be the left edge and γ_2 the right edge and write $V_a = \gamma_1 \cup \gamma_2$.

Lemma B.4 *Let* $v : \mathbb{R}^2 \to \mathbb{R}$ *be a* C^1 *function. Then*

$$\int_{V_a} |v| \le 2\big| ((1 + M^2)^{1/2} + M)/(\psi_2 - \psi_1) \big|_\infty |1_a v|_1 + |(1 + M^2)^{1/2}|_\infty |1_a \partial_y v|_1.$$

Proof

$$\begin{aligned}\int_{\gamma_2} |v| &= \int_C^D |v(\psi_2(\theta),\theta)|(1+(\psi_2'(\theta))^2)^{1/2}\,d\theta \\ &= \int_C^D |v(g^{-1}(1,\theta))|(1+[(\partial_\theta h)(1,\theta)]^2)^{1/2}\,d\theta = \int_C^D |w(1,\theta)|\,d\theta,\end{aligned}$$

where $w:[-1,1]\times[C,D]\to\mathbb{R}$ is given by $w = v\circ g^{-1}\,(1+(\partial_\theta h)^2)^{1/2}$. Similarly, $\int_{\gamma_1}|v| = \int_C^D |w(-1,\theta)|\,d\theta$.

By Proposition B.2,

$$\int_{V_a} |v| \le |1_{[-1,1]\times[C,D]}w|_1 + |1_{[-1,1]\times[C,D]}\partial_y w|_1.$$

For the first term,

$$\begin{aligned}|1_{[-1,1]\times[C,D]}w|_1 &= \int_{[-1,1]\times[C,D]} |v\circ g^{-1}|\,(1+(\partial_\theta h)^2)^{1/2} \\ &\le \int_{[-1,1]\times[C,D]} |v\circ g^{-1}|\,(1+M^2)^{1/2} = \int_a |v|\,(1+M^2)^{1/2} Jg \\ &\le 2\int_a |v|\,(1+M^2)^{1/2}/(\psi_2-\psi_1) \le 2|(1+M^2)^{1/2}/(\psi_2-\psi_1)|_\infty |1_a v|_1.\end{aligned}$$

Next, we have

$$\partial_y w = (\partial_y v)\circ g^{-1}\,\partial_y h(1+(\partial_\theta h)^2)^{1/2} + v\circ g^{-1}(1+(\partial_\theta h)^2)^{-1/2}\partial_\theta h\,\partial_\theta\partial_y h.$$

Hence

$$|1_{[-1,1]\times[C,D]}\partial_y w|_1 \le I_1 + I_2$$

where

$$\begin{aligned}I_1 &= \int_{[-1,1]\times[C,D]} |(\partial_y v)\circ g^{-1}|\,|\partial_y h|\,(1+(\partial_\theta h)^2)^{1/2} \\ &= \int_a |\partial_y v|\,(1+(\partial_\theta h)^2\circ g)^{1/2} \le |(1+M^2)^{1/2}|_\infty |1_a \partial_y v|_1,\end{aligned}$$

and

$$\begin{aligned}I_2 &= \int_{[-1,1]\times[C,D]} |v\circ g^{-1}|\,(1+(\partial_\theta h)^2)^{-1/2}|\partial_\theta h|\,|\partial_\theta\partial_y h| \le \int_{[-1,1]\times[C,D]} |v\circ g^{-1}|\,|\partial_\theta\partial_y h| \\ &= \int_a |v|\,|\partial_\theta\partial_y h|\,Jg = \int_a |v|\,|\psi_2'-\psi_1'|/(\psi_2-\psi_1) \\ &\le |(\psi_2'-\psi_1')/(\psi_2-\psi_1)|_\infty |1_a v|_1 \le 2|M/(\psi_2-\psi_1)|_\infty |1_a v|_1.\end{aligned}$$

∎

Horizontal edges Next we let H_a denote the union of the horizontal edges.

Lemma B.5 *Let $v:\mathbb{R}^2\to\mathbb{R}$ be a C^1 function. Then*

$$\int_{H_a}|v|\le 2\Big\{(D-C)^{-1}+\big|M/(\psi_2-\psi_1)\big|_\infty\Big\}|1_a v|_1+|1_a\partial_\theta v|_1+|M|_\infty|1_a\partial_y v|_1.$$

Proof For the bottom edge, we write

$$\int_{\psi_1(C)}^{\psi_2(C)}|v(t,C)|\,dt=\int_{-1}^{1}|v(h(y,C),C)||(\partial_y h)(y,C)|\,dy=\int_{-1}^{1}|w(y,C)|\,dy,$$

where $w=v\circ g^{-1}\partial_y h$. Similarly, $\int_{\psi_1(D)}^{\psi_2(D)}|v(t,D)|\,dt=\int_{-1}^{1}|w(y,D)|\,dy$.

By Proposition B.2,

$$\int_{H_a}|v(t,C)|\,dt\le 2(D-C)^{-1}|1_{[-1,1]\times[C,D]}w|_1+|1_{[-1,1]\times[C,D]}\partial_\theta w|_1.$$

Now

$$|1_{[-1,1]\times[C,D]}w|_1=\int_{[-1,1]\times[C,D]}|v\circ g^{-1}|\,|\partial_y h|=\int_a|v|\,|\partial_y h|Jg=|1_a v|_1.$$

Also,

$$\partial_\theta w=\partial_y v\circ g^{-1}\,\partial_\theta h\,\partial_y h+\partial_\theta v\circ g^{-1}\,\partial_y h+v\circ g^{-1}\,\partial_\theta\partial_y h,$$

and so

$$|1_{[-1,1]\times[C,D]}\partial_\theta w|_1\le I_1+I_2+I_3,$$

where

$$\begin{aligned}I_1&=\int_{[-1,1]\times[C,D]}|\partial_y v\circ g^{-1}||\partial_\theta h||\partial_y h|=\int_a|\partial_y v||\partial_\theta h|\circ g|\partial_y h|Jg\\&=\int_a|\partial_y v||\partial_\theta h|\circ g\le|M|_\infty|1_a\partial_y v|_1,\end{aligned}$$

$$I_2=\int_{[-1,1]\times[C,D]}|\partial_\theta v\circ g^{-1}||\partial_y h|=\int_a|\partial_\theta v||\partial_y h|Jg=\int_a|\partial_\theta v|=|1_a\partial_\theta v|_1,$$

and

$$\begin{aligned}I_3&=\int_{[-1,1]\times[C,D]}|v\circ g^{-1}||\partial_\theta\partial_y h|=\int_a|v||\partial_\theta\partial_y h|Jg\le|(\psi_2'-\psi_1')/(\psi_2-\psi_1)|_\infty|1_a v|_1\\&\le 2|M/(\psi_2-\psi_1)|_\infty|1_a v|_1.\end{aligned}$$

∎

Proof of Theorem B.1 We combine the contributions from Lemmas B.4 and B.5. The coefficient of the $|1_a v|_1$ term is immediate. The remaining terms yield

$$|1_a\partial_\theta v|_1+\big((1+|M|_\infty^2)^{1/2}+|M|_\infty\big)|1_a\partial_y v|_1.$$

The result follows since $|\partial_\theta v|+|\partial_y v|\le\sqrt{2}|\nabla v|$. ∎

C Convergence to a stable law

In this appendix, we describe a general functional-analytic framework for establishing convergence to a stable law. Our presentation follows [2, Theorem 6.1] with a simplification due to [32].

Let $F : Y \to Y$ be an ergodic measure-preserving transformation on a probability space (Y, μ_Y) with transfer operator $R : L^1(Y) \to L^1(Y)$. Let $\mathcal{B}(Y) \subset L^1(Y)$ be a Banach space containing constant functions. In particular, 1 is a simple eigenvalue for $R : \mathcal{B}(Y) \to \mathcal{B}(Y)$. We assume that there is a spectral gap for $R : \mathcal{B}(Y) \to \mathcal{B}(Y)$, so $\operatorname{spec} R \subset \{1\} \cup B_\kappa(0)$ for some $\kappa < 1$.

Let $\psi \in L^1(Y)$ with $\int_Y \psi \, d\mu_Y = 0$, and suppose that there are constants $\sigma_1, \sigma_2 \geq 0$ with $\sigma_1 + \sigma_2 > 0$, and $\alpha \in (1, 2)$, such that

$$\mu_Y(\psi > x) = (\sigma_1 + o(1))x^{-\alpha} \quad \text{and} \quad \mu_Y(\psi < -x) = (\sigma_2 + o(1))x^{-\alpha} \quad \text{as } x \to \infty. \tag{C.1}$$

Define

$$\sigma = (\sigma_1 + \sigma_2)\Gamma(1 - \alpha) \cos \tfrac{\alpha\pi}{2}, \qquad \beta = (\sigma_1 - \sigma_2)/(\sigma_1 + \sigma_2).$$

It follows from these assumptions on ψ (see [40, Theorem 2.6.5]) that

$$\int_Y e^{it\psi} \, d\mu_Y = 1 - \sigma|t|^\alpha(1 - i\beta \operatorname{sgn} t \tan \tfrac{\alpha\pi}{2}) + o(|t|^\alpha) \quad \text{as } t \to 0.$$

Define the twisted transfer operators $R_t : L^1(Y) \to L^1(Y)$, $t \in \mathbb{R}$, by $R_t v = R(e^{it\psi}v)$. Our final assumption is that there exists $t_0 > 0$, $\alpha' \in (\frac{1}{2}\alpha, 1]$ and $C > 0$ such that R_t restricts to an operator $R_t : \mathcal{B}(Y) \to \mathcal{B}(Y)$ and $\|R_t - R\|_{\mathcal{B}} \leq C|t|^{\alpha'}$ for all $|t| < t_0$. Let $\psi_n = \sum_{j=0}^{n-1} \psi \circ F^j$.

Theorem C.1 *Under the above assumptions, $n^{-1/\alpha}\psi_n \to_d \sigma^{1/\alpha} G_{\alpha,\beta}$ where $G_{\alpha,\beta}$ is the α-stable law with characteristic function $\mathbb{E}(e^{itG_{\alpha,\beta}}) = \exp\{-|t|^\alpha(1 - i\beta \operatorname{sgn} t \tan \frac{\alpha\pi}{2})\}$.*

Proof The argument is by now standard. Since we could not find the result stated in the literature, we give the details.

Since $t \mapsto R_t : \mathcal{B}(Y) \to \mathcal{B}(Y)$ is continuous at $t = 0$, there exists $t_1 \in (0, t_0]$, $\kappa_0 \in (\kappa, 1)$ and $\lambda_t \in B_1(0)$, such that λ_t is a simple isolated eigenvalue for R_t and $\operatorname{spec} R_t \subset \{\lambda_t\} \cup B_{\kappa_0}(0)$ for all $|t| < t_1$. Moreover, $|\lambda_t - 1| \ll |t|^{\alpha'}$.

Let $w_t \in \mathcal{B}(Y)$ denote the family of eigenfunctions corresponding to λ_t with $w_0 = 1$. Shrinking t_1 if necessary, we can ensure that $w_t > 0$. In particular, we can normalize so that $\int_Y w_t \, d\mu_Y = 1$ for all $|t| < t_1$.

Let P_t be the corresponding family of spectral projections with $P_0 v = \int_Y v \, d\mu_Y$. Again $\|P_t - P_0\|_{\mathcal{B}} \ll |t|^{\alpha'}$. We have

$$R_t^n = \lambda_t^n P_t + R_t^n(I - P_t).$$

Let $\kappa_1 \in (\kappa_0, 1)$. Then there exists a constant $C > 0$ and functions $a_1(t)$, $a_2(t,n)$ such that

$$\int_Y R_t^n 1\, d\mu_Y = \lambda_t^n(1 + a_1(t)) \;+\; a_2(t,n)$$

and

$$|a_1(t)| \le C|t|^{\alpha'}, \qquad |a_2(t,n)| \le C\kappa_1^n,$$

for all $|t| < t_1$, $n \ge 1$.

Next,

$$\begin{aligned}\lambda_t &= \int_Y \lambda_t w_t\, d\mu_Y = \int_Y R_t w_t\, d\mu_Y = \int_Y R_t 1\, d\mu_Y + \int_Y (R_t - R)(w_t - w_0)\, d\mu_Y \\ &= \int_Y e^{it\psi}\, d\mu_Y + O(t^{2\alpha'}) = \int_Y e^{it\psi}\, d\mu_Y = 1 - \sigma|t|^\alpha(1 - i\beta \operatorname{sgn} t \tan \tfrac{\alpha\pi}{2}) + o(t^\alpha).\end{aligned}$$

Now fix $t \in \mathbb{R}$. Then

$$\begin{aligned}\int_Y e^{itn^{-1/\alpha}\psi_n}\, d\mu_Y &= \int_Y R^n(e^{itn^{-1/\alpha}\psi_n})\, d\mu_Y = \int_Y R^n_{tn^{-1/\alpha}} 1\, d\mu_Y \\ &= \lambda^n_{tn^{-1/\alpha}}(1 + a_1(tn^{-1/\alpha})) \;+\; a_2(tn^{-1/\alpha}, n) = \lambda^n_{tn^{-1/\alpha}}(1 + O(n^{-\alpha'/\alpha})) \;+\; O(\kappa_1^n) \\ &= \big(1 - \sigma|t|^\alpha n^{-1}(1 - i\beta \operatorname{sgn} t \tan \tfrac{\alpha\pi}{2}\big)^n(1 + O(n^{-\alpha'/\alpha})) \;+\; O(\kappa_1^n) \\ &\to \exp\{-\sigma|t|^\alpha(1 - i\beta \operatorname{sgn} t \tan \tfrac{\alpha\pi}{2})\},\end{aligned}$$

as $n \to \infty$. Replacing t by $t\sigma^{-1/\alpha}$, it follows from the Lévy continuity theorem that $\sigma^{-1/\alpha} n^{-1/\alpha}\psi_n \to_d G_{\alpha,\beta}$. ∎

Remark C.2 Similarly, following [3], if $\mu_Y(|\psi| > x) \sim (\sigma^2 + o(1))x^{-2}$ as $x \to \infty$, then $\lambda_t = 1 + \sigma^2 t^2 \log|t| + o(t^2 \log|t|)$. (Here, we require that $\|R_t - R\|_{\mathcal{B}} \le C|t|$.) The above argument then shows that $(n \log n)^{-1/2}\psi_n \to_d N(0, \sigma^2)$.

Remark C.3 We have restricted to tails of the form $\mu_Y(|\psi| > x) = \ell(x)x^{-\alpha}$ where $\lim_{x\to\infty} \ell(x) = c$ for some $c > 0$, since this suffices for our examples. The general case with ℓ slowly varying goes through as in [2, 3].

Acknowledgements The research of PE and IM was supported in part by a European Advanced Grant *StochExtHomog* (ERC AdG 320977).

References

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory*. Math. Surveys and Monographs **50**, Amer. Math. Soc., 1997. [cited on p. 16]

[2] 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. [cited on p. 46, 47]

[3] J. Aaronson and M. Denker. A local limit theorem for stationary processes in the domain of attraction of a normal distribution. *Asymptotic methods in probability and statistics with applications (St. Petersburg, 1998)*, Stat. Ind. Technol., Birkhäuser Boston, Boston, MA, 2001, pp. 215–223. [cited on p. 47]

[4] M. Antoniou and I. Melbourne. Rate of convergence in the weak invariance principle for deterministic systems. *Comm. Math. Phys.* **369** (2019) 1147-1165. [cited on p. 22]

[5] W. Bahsoun, C. Bose and Y. Duan. Decay of correlation for random intermittent maps. *Nonlinearity* **27** (2014) 1543–1554. [cited on p. 6]

[6] W. Bahsoun and C. Bose. Mixing rates and limit theorems for random intermittent maps. *Nonlinearity* **29** (2016) 1417–1433. [cited on p. 6]

[7] W. Bahsoun, C. Bose and M. Ruziboev. Quenched decay of correlations for slowly mixing systems. *Trans. Amer. Math. Soc.* **372** (2019) 6547–6587. [cited on p. 6]

[8] P. Bálint and I. P. Tóth. Exponential decay of correlations in multi-dimensional dispersing billiards. *Ann. Henri Poincaré* **9** (2008) 1309–1369. [cited on p. 38, 39]

[9] P. Bálint, O. Butterley and I. Melbourne. Polynomial decay of correlations for flows, including Lorentz gas examples. *Comm. Math. Phys.* **368** (2019) 55–111. [cited on p. 20]

[10] H. Bruin, I. Melbourne and D. Terhesiu. Sharp polynomial bounds on decay of correlations for multidimensional nonuniformly hyperbolic systems and billiards. *Ann. Henri Lebesgue* **4** (2021) 407–451. [cited on p. 6, 7, 18]

[11] J. Buzzi. Absolutely continuous invariant probability measures for arbitrary expanding piecewise $\mathbf{R}$-analytic mappings of the plane. *Ergodic Theory Dynam. Systems* **20** (2000) 697–708. [cited on p. 6]

[12] N. Chernov. Statistical properties of piecewise smooth hyperbolic systems in high dimensions. *Discrete Contin. Dynam. Systems* **5** (1999) 425–448. [cited on p. 7]

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

[14] N. I. Chernov and H.-K. Zhang. Improved estimates for correlations in billiards. *Comm. Math. Phys.* **77** (2008) 305–321. [cited on p. 19]

[15] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang. Multiscale systems, homogenization, and rough paths. In: Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016, eds. P. Friz et al. *Springer Proc. in Maths. & Stat.* **283** (2019) 17–42. [cited on p. 23]

[16] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2. Preprint, 2020. [cited on p. 23]

[17] C. Cuny, J. Dedecker, A. Korepanov and F. Merlevède. Rates in almost sure invariance principle for slowly mixing dynamical systems. *Ergodic Theory Dynam. Systems* **40** (2020) 2317–2348. [cited on p. 22]

[18] C. Cuny and F. Merlevède. Strong invariance principles with rate for "reverse" martingales and applications. *J. Theor. Probab.* (2015) 137–183. [cited on p. 22]

[19] W. J. Cowieson. Stochastic stability for piecewise expanding maps in $\mathbf{R}^d$. *Nonlinearity* **13** (2000) 1745–1760. [cited on p. 6, 27]

[20] W. J. Cowieson. Absolutely continuous invariant measures for most piecewise smooth expanding maps. *Ergodic Theory Dynam. Systems* **22** (2002) 1061–1078. [cited on p. 6, 27]

[21] J. Dedecker and F. Merlevède. Moment bounds for dependent sequences in smooth Banach spaces. *Stochastic Process. Appl.* **125** (2015) 3401–3429. [cited on p. 2, 22]

[22] D. Dolgopyat. On dynamics of mostly contracting diffeomorphisms. *Comm. Math. Phys.* **213** (2000) 181–201. [cited on p. 7]

[23] P. Eslami. Inducing schemes for multi-dimensional piecewise expanding maps. Preprint, 2019. To appear in *Discrete Contin. Dyn. Syst.* [cited on p. 7, 36, 37]

[24] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992. [cited on p. 32]

[25] E. Giusti. *Minimal surfaces and functions of bounded variation.* Volume 80 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel, 1984. [cited on p. 26, 27, 31]

[26] P. Góra and A. Boyarsky. Absolutely continuous invariant measures for piecewise expanding C^2 transformation in $\mathbf{R}^N$. *Israel J. Math.* **67** (1989) 272–286. [cited on p. 6, 27]

[27] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. London A* (2013) 20130201. [cited on p. 23]

[28] S. Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* **139** (2004) 29–65. [cited on p. 2, 34]

[29] S. Gouëzel. Central limit theorem and stable laws for intermittent maps. *Probab. Theory Relat. Fields* **128** (2004) 82–122. [cited on p. 2]

[30] S. Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005) 997–1024. [cited on p. 2, 22]

[31] S. Gouëzel. Statistical properties of a skew product with a curve of neutral points. *Ergodic Theory Dynam. Systems* **27** (2007) 123–151. [cited on p. 6]

[32] S. Gouëzel. Characterization of weak convergence of Birkhoff sums for Gibbs-Markov maps. *Israel J. Math.* **180** (2010) 1–41. [cited on p. 46]

[33] S. Gouëzel. Correlation asymptotics from large deviations in dynamical systems with infinite measure. *Colloq. Math.* **125** (2011) 193–212. [cited on p. 2, 34]

[34] S. Gouëzel and I. Melbourne. Moment bounds and concentration inequalities for slowly mixing dynamical systems. *Electron. J. Probab.* **19** (2014) 1–30. [cited on p. 2, 22, 23]

[35] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.* **118** (1993) 627–634. [cited on p. 31]

[36] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.* **180** (1982) 119–140. [cited on p. 36]

[37] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergodic Theory Dynam. Systems* **24** (2004) 495–524. [cited on p. 2]

[38] H. Hu and S. Vaienti. Absolutely continuous invariant measures for non-uniformly expanding maps. *Ergodic Theory Dynam. Systems* **29** (2009) 1185–1215. [cited on p. 6]

[39] H. Hu and S. Vaienti. Lower bounds for the decay of correlations in non-uniformly expanding maps. *Ergodic Theory Dynam. Systems* **39** (2019) 1936–1970. [cited on p. 6]

[40] I. A. Ibragimov and Y. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen, 1971. [cited on p. 46]

[41] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2016) 479–520. [cited on p. 23]

[42] D. Kelly and I. Melbourne. Homogenization for deterministic fast-slow systems with multidimensional multiplicative noise. *J. Funct. Anal.* **272** (2017) 4063–4102. [cited on p. 23]

[43] A. Korepanov. Equidistribution for nonuniformly expanding dynamical systems, and application to the almost sure invariance principle. *Comm. Math. Phys.* **359** (2018) 1123–1138. [cited on p. 22]

[44] A. Korepanov, Z. Kosloff and I. Melbourne. Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **35** (2018) 859–885. [cited on p. 22]

[45] A. Korepanov, Z. Kosloff and I. Melbourne. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1. Preprint, 2020. To appear in *Ann Inst. H. Poincaré (B) Probab. Statist.* [cited on p. 23]

[46] C. Liverani. Multidimensional expanding maps with singularities: a pedestrian approach. *Ergodic Theory Dynam. Systems* **33** (2013) 168–182. [cited on p. 6, 32]

[47] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** (1999) 671–685. [cited on p. 2, 3]

[48] R. Markarian. Billiards with polynomial decay of correlations. *Ergodic Theory Dynam. Systems* **24** (2004) 177–197. [cited on p. 18]

[49] I. Melbourne. Large and moderate deviations for slowly mixing dynamical systems. *Proc. Amer. Math. Soc.* **137** (2009) 1735–1741. [cited on p. 2, 22]

[50] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005) 131–146. [cited on p. 2, 22]

[51] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.* **360** (2008) 6661–6676. [cited on p. 2, 22]

[52] I. Melbourne and D. Terhesiu. Operator renewal theory and mixing rates for dynamical systems with infinite measure. *Invent. Math.* **189** (2012) 61–110. [cited on p. 2, 34]

[53] I. Melbourne and A. Török. Convergence of moments for Axiom A and nonuniformly hyperbolic flows. *Ergodic Theory Dynam. Systems* **32** (2012) 1091–1100. [cited on p. 2, 22]

[54] 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. [cited on p. 35, 36]

[55] I. Melbourne and R. Zweimüller. Weak convergence to stable Lévy processes for nonuniformly hyperbolic dynamical systems. *Ann Inst. H. Poincaré (B) Probab. Statist.* **51** (2015) 545–556. [cited on p. 2]

[56] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980) 189–197. [cited on p. 1]

[57] M . R. Rychlik. *Invariant measures and the variational principle for Lozi mappings*. ProQuest LLC, Ann Arbor, MI, 1983, Ph. D. Thesis, University of California, Berkeley. [cited on p. 17, 18]

[58] O. M. Sarig. Subexponential decay of correlations. *Invent. Math.* **150** (2002) 629–653. [cited on p. 2, 34]

[59] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* **116** (2000) 223–248. [cited on p. 6]

[60] A. V. Skorohod. Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen.* **1** (1956) 289–319. [cited on p. 5]

[61] D. Szász, T. Varjú. Limit laws and recurrence for the planar Lorentz process with infinite horizon. *J. Stat. Phys.* **129** (2007) 59–80. [cited on p. 19]

[62] M. Thaler. Estimates of the invariant densities of endomorphisms with indifferent fixed points. *Israel J. Math.* **37** (1980) 303–314. [cited on p. 1]

[63] M. Tsujii. Absolutely continuous invariant measures for piecewise real-analytic expanding maps on the plane. *Comm. Math. Phys.* **208** (2000) 605–622. [cited on p. 6]

[64] 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. [cited on p. 5]

[65] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188. [cited on p. 2, 17, 18, 22]

[66] R. Zweimüller. Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points. *Nonlinearity* **11** (1998) 1263–1276. [cited on p. 1]

[67] R. Zweimüller. Mixing limit theorems for ergodic transformations. *J. Theoret. Probab.* **20** (2007) 1059–1071. [cited on p. 4]