



Doubly Intermittent Full Branch Maps with Critical Points and Singularities

Douglas Coates³ , Stefano Luzzatto¹, Mubarak Muhammad^{1,2}

¹ Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy. E-mail: luzzatto@ictp.it

² Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy. E-mail: mmuhamma@ictp.it, mubarak@aims.edu.gh

³ University of Exeter, Exeter, UK. E-mail: dc485@exeter.ac.uk

Received: 19 October 2022 / Accepted: 14 May 2023
Published online: 28 June 2023 – © The Author(s) 2023

Abstract: We study a class of one-dimensional full branch maps admitting two indifferent fixed points as well as critical points and/or unbounded derivative. Under some mild assumptions we prove the existence of a unique invariant mixing absolutely continuous probability measure, study its rate of decay of correlation and prove a number of limit theorems.

1. Introduction

The purpose of this paper is to study the ergodic properties of a large class of full branch interval maps with two branches, including maps with *two* indifferent fixed points (which, as we shall see below, affects both the results and the construction of the induced map which we require). We also allow the derivative to go to *zero* as well as to *infinity* at the boundary between the two branches, and we do not assume any symmetry, even the domains of the branches can be of arbitrary length. Such maps are known to exhibit a wide range of behaviour from an ergodic point of view and many of them have been extensively studied, we give a detailed literature review below.

In Sect. 1.1 we give the precise definition of the class of maps we consider, which includes many cases already studied in the literature as well as many cases which have not yet been studied; in section 1.2 we give the precise statements of our results; in section 1.3 we give a literature review of related results and include specific examples of maps in our family; in Sect. 2 we give a detailed outline of our proof, emphasising several novel aspects of our construction and arguments. Then in Sect. 3 we give the construction and estimates related to our “double-induced” map and in Sect. 4 apply these estimates to complete the proofs of our results

1.1. Full branch maps. We start by defining the class of maps which we consider in this paper. Let I , I_- , I_+ be compact intervals, let $\overset{\circ}{I}$, $\overset{\circ}{I_-}$, $\overset{\circ}{I_+}$ denote their interiors, and suppose that $I = I_- \cup I_+$ and $\overset{\circ}{I_-} \cap \overset{\circ}{I_+} = \emptyset$.

(A0) $g : I \rightarrow I$ is *full branch*: the restrictions $g_- : \overset{\circ}{I}_- \rightarrow \overset{\circ}{I}$ and $g_+ : \overset{\circ}{I}_+ \rightarrow \overset{\circ}{I}$ are orientation preserving C^2 diffeomorphisms and the only fixed points are the endpoints of I .

To simplify the notation we will assume that

$$I = [-1, 1], \quad I_- = [-1, 0], \quad I_+ = [0, 1]$$

but our results and proofs will be easily seen to hold in the general setting.

(A1) There exists constants $\ell_1, \ell_2 \geq 0, \iota, k_1, k_2, a_1, a_2, b_1, b_2 > 0$ such that:

(i) if $\ell_1, \ell_2 \neq 0$ and $k_1, k_2 \neq 1$, then

$$g(x) = \begin{cases} x + b_1(1+x)^{1+\ell_1} & \text{in } U_{-1}, \\ 1 - a_1|x|^{k_1} & \text{in } U_{0-}, \\ -1 + a_2x^{k_2} & \text{in } U_{0+}, \\ x - b_2(1-x)^{1+\ell_2} & \text{in } U_{+1}, \end{cases} \tag{1}$$

where

$$U_{0-} := (-\iota, 0], \quad U_{0+} := [0, \iota), \quad U_{-1} := g(U_{0+}), \quad U_{+1} := g(U_{0-}). \tag{2}$$

(ii) If $\ell_1 = 0$ and/or $\ell_2 = 0$ we replace the corresponding lines in (1) with

$$g|_{U_{\pm 1}}(x) := \pm 1 + (1 + b_1)(x + 1) \mp \xi(x), \tag{3}$$

where ξ is C^2 , $\xi(\pm 1) = 0, \xi'(\pm 1) = 0$, and $\xi''(x) > 0$ on U_{-1} and $\xi''(x) < 0$ on U_{+1} .

If $k_1 = 1$ and/or $k_2 = 1$, then we replace the corresponding lines in (1) with the assumption that $g'(0_-) = a_1 > 1$ and/or $g'(0_+) = a_2 > 1$ respectively, and that g is monotone in the corresponding neighbourhood.

Remark 1.1. It is easy to see that the definition in (1) yields maps with dramatically different derivative behaviour depending on the values of ℓ_1, ℓ_2, k_1, k_2 , including having neutral or expanding fixed points and points with zero or infinite derivative, see Remark 1.3 for a detailed discussion. For the moment we just remark that the assumptions described in part ii) of condition (A1) are consistent with (1) but significantly relax the definition given there as in these cases (1) would imply that the map is affine in the corresponding neighbourhood, whereas we only need expansivity. In particular this allows us to include uniformly expanding maps in our class of maps. In the calculations below we will explicitly consider the cases $\ell_1 = 0$ and/or $\ell_2 = 0$, which correspond to assuming that one or both the fixed points are expanding instead of neutral, since they yield different estimates (several quantities decay exponentially rather than polynomially in these cases) and different results, and still include some maps which, as far as we know, have not been studied in the literature. For simplicity, on the other hand, we will not consider explicitly the cases $k_1 = 1$ and/or $k_2 = 1$, which just correspond to assuming the derivative at one or both sides of the discontinuity is finite instead of being zero or infinite. These correspond to much simpler special cases and the required estimates follow by arguments which are very similar to arguments and calculations we give here, and which are essentially already considered in the literature, but treating them explicitly would require a significant amount of additional notation and calculations.

Our final assumption can be intuitively thought of as saying that g is uniformly expanding outside the neighbourhoods $U_{0\pm}$ and $U_{\pm 1}$. This is however much stronger than what is needed, and therefore we formulate a weaker and more general assumption for which we need to describe some aspects of the topological structure of maps satisfying condition (A0). First of all we define

$$\Delta_0^- := g^{-1}(0, 1) \cap I_- \quad \text{and} \quad \Delta_0^+ := g^{-1}(-1, 0) \cap I_+. \tag{4}$$

Then we define iteratively, for every $n \geq 1$, the sets

$$\Delta_n^- := g^{-1}(\Delta_{n-1}^-) \cap I_- \quad \text{and} \quad \Delta_n^+ := g^{-1}(\Delta_{n-1}^+) \cap I_+ \tag{5}$$

as the n 'th preimages of Δ_0^- , Δ_0^+ inside the intervals I_- , I_+ . It follows from (A0) that $\{\Delta_n^-\}_{n \geq 0}$ and $\{\Delta_n^+\}_{n \geq 0}$ are mod 0 partitions of I_- and I_+ respectively, and that the partition elements depend *monotonically* on the index in the sense that $n > m$ implies that Δ_n^\pm is closer to ± 1 than Δ_m^\pm , in particular the only accumulation points of these partitions are -1 and 1 respectively. Then, for every $n \geq 1$, we let

$$\delta_n^- := g^{-1}(\Delta_{n-1}^+) \cap \Delta_0^- \quad \text{and} \quad \delta_n^+ := g^{-1}(\Delta_{n-1}^-) \cap \Delta_0^+. \tag{6}$$

Notice that $\{\delta_n^-\}_{n \geq 1}$ and $\{\delta_n^+\}_{n \geq 1}$ are mod 0 partitions of Δ_0^- and Δ_0^+ respectively and also in these cases the partition elements depend monotonically on the index in the sense that $n > m$ implies that δ_n^\pm is closer to 0 than δ_m^\pm , (and in particular the only accumulation point of these partitions is 0). Notice moreover, that

$$g^n(\delta_n^-) = \Delta_0^+ \quad \text{and} \quad g^n(\delta_n^+) = \Delta_0^-.$$

We now define two non-negative integers n_\pm which depend on the positions of the partition elements δ_n^\pm and on the sizes of the neighbourhoods $U_{0\pm}$ on which the map g is explicitly defined. If $\Delta_0^- \subseteq U_{0-}$ and/or $\Delta_0^+ \subseteq U_{0+}$, we define $n_- = 0$ and/or $n_+ = 0$ respectively, otherwise we let

$$n_+ := \min\{n : \delta_n^+ \subset U_{0+}\} \quad \text{and} \quad n_- := \min\{n : \delta_n^- \subset U_{0-}\}. \tag{7}$$

We can now formulate our final assumption as follows.

(A2) There exists a $\lambda > 1$ such that for all $1 \leq n \leq n_\pm$ and for all $x \in \delta_n^\pm$ we have $(g^n)'(x) > \lambda$.

Notice that (A2) is an expansivity condition for points outside the neighbourhoods $U_{0\pm}$ and $U_{\pm 1}$ but is much weaker than assuming that the derivative of g is greater than 1 outside these neighbourhoods, which would be unnatural and unnecessarily restrictive in the presence of critical points. This completes the set of conditions which we require, and for convenience we let

$$\widehat{\mathfrak{F}} := \{g : I \rightarrow I \text{ which satisfy (A0)-(A2)}\}$$

The class $\widehat{\mathfrak{F}}$ contains many maps which have been studied in the literature, including uniformly expanding maps and various well known intermittency maps with a single neutral fixed point. We will give a more in-depth literature review in Sect. 1.3. Here we make a few technical remarks concerning these assumptions before proceeding to state our results in the next subsection.

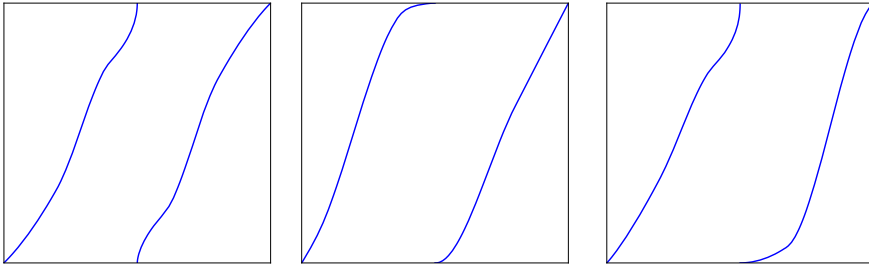


Fig. 1. Graph of g for various possible values of parameters

Remark 1.2 (Remark on notation). To simplify many statements which will be made through the paper, it will be useful to recall some relatively standard notation as follows. Given sequences (s_n) and (t_n) of non-negative terms, we write $s_n = O(t_n)$, or $s_n \lesssim t_n$, if s_n/t_n is uniformly bounded above; $s_n \approx t_n$ if s_n/t_n is uniformly bounded away from 0 and ∞ ; $s_n = o(t_n)$ if $s_n/t_n \rightarrow 0$ as $n \rightarrow \infty$; and $s_n \sim t_n$ if $s_n/t_n = 1 + o(1)$, i.e. if s_n/t_n converges to 1 as $n \rightarrow \infty$.

Remark 1.3. Changing the parameter values ℓ_1, ℓ_2, k_1, k_2 gives rise to maps with quite different characteristics. For example, if $\ell_1 > 0$, we have

$$g'|_{U_{-1}}(x) = 1 + b_1(1 + \ell_1)(1 + x)^{\ell_1} \quad \text{and} \quad g''|_{U_{-1}}(x) = b_1(1 + \ell_1)\ell_1(1 + x)^{\ell_1-1}. \tag{8}$$

Then $g'(-1) = 1$ and the fixed point -1 is a *neutral* fixed point. Similarly, when $\ell_2 > 0$ the fixed point 1 is a neutral fixed point. On the other hand, when $\ell_1 = 0$, from (3) we have

$$g'|_{U_{-1}}(x) = 1 + b_1 + \xi'(x) \quad \text{and} \quad g''|_{U_{-1}}(x) = \xi''(x) \tag{9}$$

and thus the fixed point -1 is *hyperbolic repelling* with $g'(-1) = 1 + b$. When $k_1 \neq 1$ we have

$$g'|_{U_{0-}}(x) = a_1 k_1 |x|^{k_1-1} \quad \text{and} \quad g''|_{U_{0-}}(x) = a_1 k_1 (k_1 - 1) |x|^{k_1-2}. \tag{10}$$

Then $k_1 \in (0, 1)$ implies that $|g'|_{U_{0-}}(x)| \rightarrow \infty$ as $x \rightarrow 0$, in which case we say that $g|_{U_{0-}}$ has a (one-sided) *singularity* at 0, whereas $k_1 > 1$ implies that $|g'|_{U_{0-}}(x)| \rightarrow 0$ as $x \rightarrow 0$, and therefore we say that $g|_{U_{0-}}$ has a (one-sided) *critical point* at 0. Analogous observations hold for the various values of ℓ_2 and k_2 and Fig. 1 shows the graph of g for various combinations of these exponents.

For future reference we mention also some additional properties which follow from (A1). First of all notice that if $\ell_1 \in (0, 1)$ we have $g''(x) \rightarrow \infty$ but if $\ell_1 > 1$ we have $g''(x) \rightarrow 0$, as $x \rightarrow -1$ and, as we shall see, this qualitative difference in the higher order derivative plays a crucial role in the ergodic properties of g . Analogous observations apply to $g|_{U_1}$ when $\ell_2 > 0$. Secondly, notice also that for every $x \in U_{-1}$ we have

$$g''(x)/g'(x) \lesssim (1 + x)^{\ell_1-1} \tag{11}$$

and an analogous bound holds for $x \in U_1$. Similarly, in U_0 we have

$$|g''(x)|/|g'(x)| \lesssim x^{-1}, \tag{12}$$

and notice that in this case the bound does not actually depend on the value of k_1 or k_2 and in particular does not depend on whether we have a critical point or a singularity. Finally, we note that when $\ell_1 = 0$, it follows from (9) and from the assumption that $\xi''(x) > 0$ that

$$\xi'(x) > \xi(x)/(1+x) \tag{13}$$

for every $x \in U_{-1}$. Indeed, notice that $1+x$ is just the distance between x and -1 and thus $\xi(x)/(1+x)$ is the slope of the straight line joining the point $(-1, 0)$ to $(x, \xi(x))$ in the graph of ξ , which is exactly the *average* derivative of ξ in the interval $[-1, x]$. Since $\xi'' > 0$, the derivative is monotone increasing and thus the derivative ξ' is maximal at the endpoint x , which implies (13). The same statement of course holds for $\ell_2 = 0$ and for all $x \in U_{+1}$.

1.2. Statement of results. Our first result is completely general and applies to all maps in $\widehat{\mathfrak{F}}$.

Theorem A. *Every $g \in \widehat{\mathfrak{F}}$ admits a unique (up to scaling by a constant) invariant measure which is absolutely continuous with respect to Lebesgue; this measure is σ -finite and equivalent to Lebesgue.*

This is perhaps not completely unexpected but also certainly not obvious in the full generality of the maps in $\widehat{\mathfrak{F}}$, especially for maps which admit critical points (which can, moreover, be of arbitrarily high order). Our construction gives some additional information about the measure given in Theorem A, in particular the fact that its density with respect to Lebesgue is locally Lipschitz and unbounded only at the endpoints ± 1 . We will show that, depending on the exponents k_1, k_2, ℓ_1, ℓ_2 , the density may or may not be integrable and so the measure may or may not be finite. More specifically, let

$$\beta_1 := k_2 \ell_1, \quad \beta_2 := k_1 \ell_2, \quad \text{and} \quad \beta := \max\{\beta_1, \beta_2\}.$$

We will show that the density is Lebesgue integrable at -1 or 1 respectively if and only if β_1 and β_2 respectively are < 1 . In particular, letting

$$\mathfrak{F} := \{g \in \widehat{\mathfrak{F}} \text{ with } \beta < 1\}$$

we have the following result.

Theorem B. *A map $g \in \widehat{\mathfrak{F}}$ admits a unique ergodic invariant probability measure μ_g absolutely continuous with respect to (indeed equivalent to) Lebesgue if and only if $g \in \mathfrak{F}$.*

Notice that the condition $\beta < 1$ is a restriction only on the *relative* values of k_1 with respect to ℓ_2 and of k_2 with respect to ℓ_1 . It still allows k_1 and/or k_2 to be *arbitrarily large*, thus allowing arbitrarily “degenerate” critical points, as long as the corresponding exponents ℓ_2 and/or ℓ_1 are sufficiently small, i.e. as long as the corresponding neutral fixed points are not too degenerate.

We now give several non-trivial results about the statistical properties maps $g \in \mathfrak{F}$ with respect to the probability measure μ_g . To state our first result recall that the *measure-theoretic entropy* of g with respect to the measure μ is defined as

$$h_\mu(g) := \sup_{\mathcal{P}} \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\omega_n \in \mathcal{P}_n} -\mu(\omega_n) \ln \mu(\omega_n) \right\}$$

where the supremum is taken over all finite measurable partitions \mathcal{P} of the underlying measure space and $\mathcal{P}_n := \mathcal{P} \vee f^{-1}\mathcal{P} \vee \dots \vee f^{-n}\mathcal{P}$ is the dynamical refinement of \mathcal{P} by f .

Theorem C. *Let $g \in \mathfrak{F}$. Then μ_g satisfies the Pesin entropy formula: $h_{\mu_g}(g) = \int \log |g'| d\mu_g$.*

For Hölder continuous functions $\varphi, \psi : [-1, 1] \rightarrow \mathbb{R}$ and $n \geq 1$, we define the *correlation function*

$$C_n(\varphi, \psi) := \left| \int \varphi \psi \circ g^n d\mu - \int \varphi d\mu \int \psi d\mu \right|.$$

It is well known that μ_g is *mixing* if and only if $C_n(\varphi, \psi) \rightarrow 0$ as $n \rightarrow \infty$. We say that μ_g is *exponentially mixing*, or satisfies *exponential decay of correlations* if there exists a $\lambda > 0$ such that for all Hölder continuous functions φ, ψ there exists a constant $C_{\varphi, \psi}$ such that $C_n(\varphi, \psi) \leq C_{\varphi, \psi} e^{-\lambda n}$. We say that μ_g is *polynomially mixing*, or satisfies *polynomial decay of correlations*, with rate $\alpha > 0$ if for all Hölder continuous functions φ, ψ there exists a constant $C_{\varphi, \psi}$ such that $C_n(\varphi, \psi) \leq C_{\varphi, \psi} n^{-\alpha}$.

Theorem D. *Let $g \in \mathfrak{F}$. If $\beta = 0$ then μ_g is exponentially mixing, if $\beta \in (0, 1)$ then μ_g is polynomially mixing with rate $(1 - \beta)/\beta$.*

Notice that the polynomial rate of decay of correlations $(1 - \beta)/\beta$ itself decays to 0 as β approaches 1, which is the *transition parameter* at which the invariant measure ceases to be finite. Intuitively, as $\beta \rightarrow 1$, the measure, while still equivalent to Lebesgue, is increasingly concentrated in neighbourhoods of the neutral fixed points, which *slow down* the decay of correlations.

Our final result concerns a number of *limit theorems* for maps $g \in \mathfrak{F}$, which depend on the parameters of the map and, in some cases, also on some additional regularity conditions. These are arguably some of the most interesting results of the paper, and those in which the existence of two indifferent fixed points, instead of just one, really comes into play, giving rise to quite a complex scenario of possibilities. We start by recalling the relevant definitions. For integrable functions φ with $\int \varphi d\mu = 0$ we define the following limit theorems.

CLT φ satisfies a *central limit theorem* with respect to μ if there exists a $\sigma^2 \geq 0$ and a $\mathcal{N}(0, \sigma^2)$ random variable V such that

$$\lim_{n \rightarrow \infty} \mu \left(\frac{\sum_{k=0}^{n-1} \varphi \circ g^k}{\sqrt{n}} \leq x \right) = \mu(V \leq x),$$

for every $x \in \mathbb{R}$ for which the function $x \mapsto \mu(V_{\sigma^2} \leq x)$ is continuous.

CLT_ns φ satisfies a *non-standard central limit theorem* with respect to μ if there exists a $\sigma^2 \geq 0$ and a $\mathcal{N}(0, \sigma^2)$ random variable V such that

$$\lim_{n \rightarrow \infty} \mu \left(\frac{\sum_{k=0}^{n-1} \varphi \circ g^k}{\sqrt{n \log n}} \leq x \right) = \mu(V \leq x),$$

for every $x \in \mathbb{R}$ for which the function $x \mapsto \mu(V_{\sigma^2} \leq x)$ is continuous.

SL_α φ satisfies a *stable law* of index $\alpha \in (1, 2)$, with respect to a measure μ , if there exists a stable random variable W_α such that

$$\lim_{n \rightarrow \infty} \mu \left(\frac{\sum_{k=0}^{n-1} \varphi \circ g^k}{n^{1/\alpha}} \leq x \right) = \mu(W_\alpha \leq x),$$

for every $x \in \mathbb{R}$ for which the function $x \mapsto \mu(W_\alpha \leq x)$ is continuous.

Finally, we say that an observable $\varphi : [-1, 1] \rightarrow \mathbb{R}$ is a *co-boundary* if there exists a measurable function $\chi : [-1, 1] \rightarrow \mathbb{R}$ such that $\varphi = \chi \circ g - \chi$. We are now ready to state our result on the various limit theorems which hold under some conditions on the parameters and on the observable φ . In order to state these conditions it is convenient to introduce the following variable:

$$\beta_\varphi := \begin{cases} 0 & \text{if } \varphi(-1) = 0 \text{ and } \varphi(1) = 0 \\ \beta_1 & \text{if } \varphi(-1) \neq 0 \text{ and } \varphi(1) = 0 \\ \beta_2 & \text{if } \varphi(-1) = 0 \text{ and } \varphi(1) \neq 0 \\ \beta & \text{if } \varphi(-1) \neq 0 \text{ and } \varphi(1) \neq 0, \end{cases} \tag{14}$$

We can then state our results in all cases in a clear and compact way as follows.

Theorem E. *Let $g \in \mathfrak{F}$ and $\varphi : [-1, 1] \rightarrow \mathbb{R}$ be Hölder continuous with $\int \varphi d\mu = 0$ and satisfying*

$$(\mathcal{H}) \quad \nu_1 > (\beta_1 - 1/2)/k_2 \quad \text{and} \quad \nu_2 > (\beta_2 - 1/2)/k_1,$$

where ν_1, ν_2 are the Hölder exponents of $\varphi|_{[-1,0]}$ and $\varphi|_{(0,1]}$ respectively. Then

1. if $\beta_\varphi \in [0, 1/2)$ then φ satisfies CLT,
2. if $\beta_\varphi = 1/2$ then φ satisfies CLT_{ns} ,
3. if $\beta_\varphi \in (1/2, 1)$ then φ satisfies SL_{1/β_φ} .

In case 3 we can replace the Hölder continuity condition (\mathcal{H}) by the weaker (in this case) condition

$$(\mathcal{H}') \quad \nu_1 > (\beta_1 - \beta_\varphi)/k_2 \quad \text{and} \quad \nu_2 > (\beta_2 - \beta_\varphi)/k_1.$$

Moreover, in all cases where (CLT) holds we have that $\sigma^2 = 0$ if and only if φ is a coboundary.

Remark 1.4. Our results highlight the fundamental significance of the value of the observable φ at the two fixed points, and how *the fixed point at which φ is non-zero*, in some sense *dominates*, and determines the kind of limit law which the observable satisfies. If φ is non-zero at both fixed points, then it is the larger exponent which dominates.

Remark 1.5. Note that (\mathcal{H}) and (\mathcal{H}') are automatically satisfied for various ranges of β_1, β_2 , for example if $\beta \leq 1/2$ then (\mathcal{H}) always holds and if $\beta = \beta_\varphi$ then (\mathcal{H}') always holds. These Hölder continuity conditions arise as technical conditions in the proof and it is not clear to us if they are really necessary and what could be proved without them. It may be the case, for example, that some limit theorems still hold under weaker regularity conditions on φ .

Remark 1.6. We remark also that the compact statement of Theorem E somewhat “conceals” quite a large number of cases which express an intricate relationship between the map parameters and the values and regularity of the observable. For example, the case $\beta_\varphi = 0$ allows all possible values $\beta_1, \beta_2 \in [0, 1)$ and the case $\beta_\varphi = \beta_1$ allows all possible values of $\beta_2 \in [0, 1)$. We therefore have a huge number of possible combinations which do not occur in the case of maps with just a single intermittent fixed point.

1.3. Examples and literature review. There is an extensive literature on the dynamics and statistical properties of full branch maps, which have been studied systematically since the 1950s. Their importance stems partly from the fact that they occur very naturally, for example any smooth non-invertible local diffeomorphism of \mathbb{S}^1 is a full branch map, but also, and perhaps most importantly, because many arguments in Smooth Ergodic Theory apply in this setting in a particularly clear and conceptually straightforward way. Indeed, arguably, most existing techniques used to study hyperbolic (including non-uniformly hyperbolic) dynamical systems are essentially (albeit often highly non-trivial) extensions and generalisations of methods first introduced and developed in the setting of one-dimensional full branch maps.

Our class of maps $\widehat{\mathfrak{F}}$ is quite general and includes many one-dimensional full branch maps which have been studied in the literature as well as many maps which have *not* been previously studied. We give below a brief survey of some of these examples and indicate for which choices of parameters these correspond to maps in our family.¹

Arguably one of the very first and simplest general class of maps for which the existence of an invariant ergodic and absolutely continuous probability measure was proved are *uniformly expanding full branch maps* with derivatives uniformly bounded away from 0 and infinity, a result often referred to as the *Folklore Theorem* and generally attributed to Renyi. Some particularly simple examples of uniformly expanding maps are piecewise affine maps such as those given by

$$g(x) = \begin{cases} ax & \text{for } x \in [0, 1/a] \\ \frac{a}{a-1} \left(x - \frac{1}{a}\right) & \text{for } x \in (1/a, 1] \end{cases} \tag{15}$$

for parameters $a > 1$, see Fig. 2a. These are easily seen to be contained in the class $\widehat{\mathfrak{F}}$ with parameters $(\ell_1, \ell_2, k_1, k_2, a_1, a_2, b_1, b_2) = (0, 0, 1, 1, a, a/(a - 1), a - 1, a/(a - 1) - 1)$.

In the late '70s, physicists Maneville and Pomeau [PM80] introduced a simple but extremely interesting generalisation consisting of a class of full branch one-dimensional maps $g : [0, 1] \rightarrow [0, 1]$, which they called *intermittency maps*, defined by

$$g(x) = x(1 + x^\alpha) \pmod 1 \tag{16}$$

for $\alpha > 0$, see Fig. 2b (notice that for $\alpha = 0$ this just gives the map $g(x) = 2x \pmod 1$, which is just (15) with $a = 2$). These maps can be seen to be contained in our class $\widehat{\mathfrak{F}}$ by taking the parameters $(\ell_1, \ell_2, k_1, k_2, a_1, a_2, b_1, b_2) = (\alpha, 0, 1, 1, a, a, 1, 1)$, where $a = g'(x_0)$, and $x_0 \in (0, 1)$ is the boundary of the intervals on which the two branches of the map are defined. The Maneville-Pomeau maps are interesting because the uniform expansivity condition fails at a single fixed point on the boundary of the interval, where we have $g'(0) = 1$. Their motivation was to model fluid flow where long period of stable flow is followed with an intermittent phase of turbulence, and they showed that this simple model indeed seemed to exhibit such dynamical behaviour. It was then shown in [Pia80] that for $\alpha > 2$, the intermittency maps *failed* to have an invariant ergodic and absolutely continuous probability measure and satisfies the extremely remarkable property that *the time averages of Lebesgue almost every point converge to the Dirac-delta measure δ_0 at the neutral fixed point*, even though these orbits are dense in $[0, 1]$ and the fixed point is topologically repelling.

¹ Recall that we have fixed the domains of the branches of our maps as $[-1, 0)$ and $(0, 1]$ for convenience. In the examples below, when listing parameters, we slightly abuse notation and assume an affine change of coordinates which transforms the given domains into the ones used in our class.

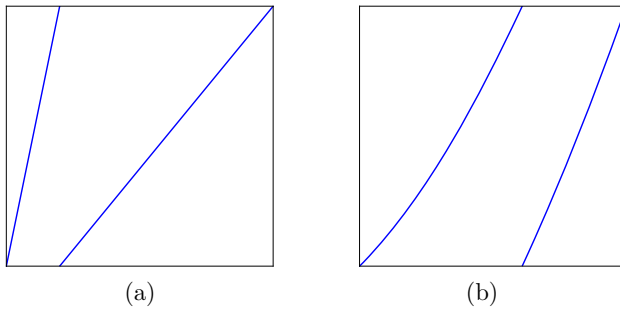


Fig. 2. **a** Graph of (15) with $a = 5$. **b** Graph of (16) with $\alpha = 9/10$. Graphs of a piecewise affine, Manneville–Pomeau, and Liverani–Saussol–Vaienti maps

Various variations of intermittency maps have been studied extensively from various points of views and with different techniques yielding quite deep results, see e.g. [LSV99, You99, Sar01, Mel09, PS09, FL01, Gou04a, Gou04b, NTV18, CHT21, Fre+16, Kor16, BS16, Ter16, SS13, Zwe03]. One well known version is the so-called Liverani–Saussol–Vaienti (LSV) map $g : [0, 1] \rightarrow [0, 1]$ introduced in [LSV99] and defined by

$$g(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{for } x \in [0, 1/2] \\ 2x - 1 & \text{for } x \in (1/2, 1) \end{cases} \tag{17}$$

with parameter $\alpha > 0$, see Fig. 3a. This maintains the essential features of the Maneville–Pomeau maps (16), i.e. it is uniformly expanding except at the neutral fixed point at the origin, but in slightly simplified form where the two branches are always defined on the fixed domains $[0, 1/2]$ and $1/2, 1)$ and the second branch is affine, both of which make the map family easier to study, including the effect of varying the parameter. The family of LSV maps (17) can be seen to be contained in our class $\widehat{\mathfrak{F}}$ by taking the parameters $(\ell_1, \ell_2, k_1, k_2, a_1, a_2, b_1, b_2) = (\alpha, 0, 1, 1, 2, 2, 2^\alpha, 1)$.

In an earlier paper [Pik91], Pikovsky had introduced the maps $g : \mathbb{S}^1 \rightarrow \mathbb{S}^1$, defined (in a somewhat unwieldy way) by the implicit equation

$$x = \begin{cases} \frac{1}{2\alpha}(1 + g(x))^\alpha & \text{for } x \in [0, 1/2\alpha] \\ g(x) + \frac{1}{2\alpha}(1 - g(x))^\alpha & \text{for } x \in (1/2\alpha, 1) \end{cases} \tag{18}$$

for $x \in [0, 1)$, and then by the symmetry $g(x) = g(-x)$ for $x \in (-1, 0]$, see Fig. 3b. These maps have a neutral fixed point at the left end point, like in (16) and (17) but with the added complication of having unbounded derivative at the boundary between the domains of the two branches. On the other hand the definition is specifically designed in such a way that the order of intermittency is the inverse of the order of the singularity and, together with the symmetry of the two branches, this implies that Lebesgue measure is invariant for all values of the parameter $\alpha > 0$. Ergodic and statistical properties of these maps were studied in [AA04, Cri+10, BM14] and they can be seen to be contained within our class $\widehat{\mathfrak{F}}$ by taking the parameters $(\ell_1, \ell_2, k_1, k_2, a_1, a_2, b_1, b_2) = (\alpha - 1, \alpha - 1, 1/\alpha, 1/\alpha, (2\alpha)^{1/\alpha}, (2\alpha)^{1/\alpha}, 1/2\alpha, 1/2\alpha)$.

Finally, [Ino92, Cui21] consider a class of maps, see Fig. 3c for an example, with a single intermittent fixed point and multiple critical points with each critical point mapping to the fixed point. These include some maps which are more general than those

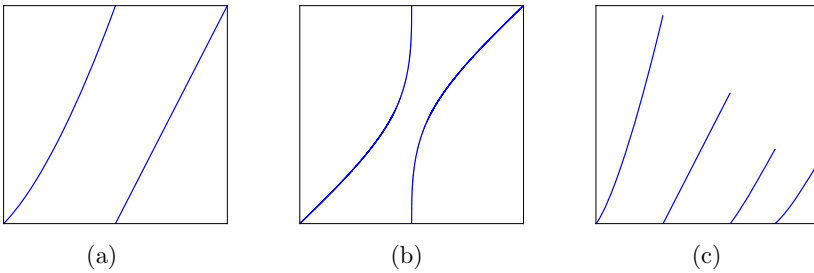


Fig. 3. **a** Graph of (17) with $\alpha = 9/10$. **b** Graph of (18) with $\alpha = 3$. **c** An example from [Cui21]. Graphs of previously studied generalisations of the Manneville-Pomeau map

we consider here as they are defined near the fixed and critical points through some bounds rather than explicitly as we do here, but are also more restrictive as they only allow for a single neutral fixed point. Under a condition on the product of the orders of the neutral and (the most degenerate) critical point which is exactly analogous to our condition $\beta < 1$, the existence of an invariant ergodic probability measure is proved which exhibits decay of correlations but no bounds are given for the rate of decay and no limit theorems are obtained.

2. Overview of the Proof

We discuss here our overall strategy and prove our Theorems modulo some key technical Propositions which we then prove in the rest of the paper. Our argument can be naturally divided into three main steps which we describe in some detail in the following three subsections.

2.1. The induced map. The first step of our arguments is the construction of an *induced* full branch Gibbs-Markov map, also known as a *Young Tower*. This is relatively standard for many systems, including intermittent maps, however, the inducing domain which we are obliged to use here due to the presence of two indifferent fixed points is *different from the usual inducing domains* and requires a more sophisticated *double inducing* procedure, which we outline here and describe and carry out in detail in Sect. 3. Recall the definition of Δ_0^- in (4) and, for $x \in \Delta_0^-$, let

$$\tau(x) := \min\{n > 0 : g^n(x) \in \Delta_0^-\}$$

be the first return time to Δ_0^- . Then we define the *first-return induced map*

$$G : \Delta_0^- \rightarrow \Delta_0^- \quad \text{by} \quad G(x) := g^{\tau(x)}(x). \tag{19}$$

We say that a first return map (or, more generally, any induced map), *saturates* the interval I if

$$\bigcup_{n \geq 0} \bigcup_{i=0}^{n-1} g^i(\{\tau = n\}) = \bigcup_{n \geq 0} g^n(\{\tau > n\}) = I \pmod{0}. \tag{20}$$

Intuitively, saturation means that the return map “reaches” every part of the original domain of the map g , and thus the properties and characteristics of the return map reflect, to some extent, all the relevant characteristics of g .

Remark 2.1. If G is a first return induced map, as in our case, then all sets of the form $g^i(\{\tau = n\})$ are pairwise disjoint and therefore form a partition of $I \bmod 0$.

The first main result of the paper is the following.

Proposition 2.2. *Let $g \in \widehat{\mathcal{F}}$. Then $G : \Delta_0^- \rightarrow \Delta_0^-$ is a first return induced Gibbs-Markov map which saturates I .*

We give the precise definition of Gibbs-Markov map, and prove Proposition 2.2, in Sect. 3. In Sect. 3.1 we describe the topological structure of G and show that it a full branch map with countably many branches which saturates I (we will define G as a composition of two full branch maps, see (37) and (40), which is why we call the construction a double inducing procedure); in Sect. 3.2 we obtain key estimates concerning the sizes of the partition elements of the corresponding partition; in Sect. 3.3 we show that G is uniformly expanding; in Sect. 3.4 we show that G has bounded distortion. From these results we get Proposition 2.2 from which we can then obtain our first main Theorem.

Proof of Theorem A. By standard results G admits a unique ergodic invariant probability measure $\hat{\mu}_-$, supported on Δ_0^- , which is equivalent to Lebesgue measure m and which has Lipschitz continuous density $\hat{h}_- = d\hat{\mu}_-/dm$ bounded above and below. We then “spread” the measure over the original interval I by defining the measure

$$\tilde{\mu} := \sum_{n=0}^{\infty} g_*^n(\hat{\mu}_-|\{\tau \geq n\}) \tag{21}$$

where $g_*^n(\hat{\mu}_-|\{\tau \geq n\})(E) := \hat{\mu}_-(g^{-n}(E) \cap \{\tau \geq n\})$. Again by standard arguments, we have that $\tilde{\mu}$ is a sigma-finite measure which is ergodic and invariant for g and, using the non-singularity of g , it is absolutely continuous with respect to Lebesgue. The fact that G saturates I implies moreover that $\tilde{\mu}$ is equivalent to Lebesgue, which completes the proof. \square

Remark 2.3. We emphasize that we are not assuming any symmetry in the two branches of the map g . It is not important that the branches are defined on intervals of the same length and, depending on the choice of constants, we might even have a critical point in one branch and a singularity with unbounded derivative on the other. Interestingly, however, there is some symmetry in the construction in the sense that for $x \in \Delta_0^+$, we can define the first return map $G_+ : \Delta_0^+ \rightarrow \Delta_0^+$ in a completely analogous way to the definition of G above (see discussion in Sect. 3.1). Moreover, the conclusions of Proposition 2.2 hold for G_+ and thus G_+ admits a unique ergodic invariant probability measure $\hat{\mu}_+$ which is equivalent to Lebesgue measure m and such that the density $\hat{h}_+ := d\hat{\mu}_+/dm$ is Lipschitz continuous and bounded above and below. The two maps G and G_+ are clearly distinct, as are the measures $\hat{\mu}_-$ and $\hat{\mu}_+$, but exhibit a subtle kind of symmetry in the sense that the corresponding measure $\tilde{\mu}$ obtained by substituting $\hat{\mu}_-$ by $\hat{\mu}_+$ in (21) is, up to a constant scaling factor, exactly the same measure.

Corollary 2.4. *The density \tilde{h} of $\tilde{\mu}|_{\Delta_0^- \cup \Delta_0^+}$ is Lipschitz continuous and bounded and $\tilde{\mu}|_{\Delta_0^-} = \hat{\mu}$.*

Proof. Since G is a first return induced map it follows that the measure $\tilde{\mu}$ defined in (21) satisfies $\tilde{\mu}|_{\Delta_0^-} = \hat{\mu}$ and so the density \tilde{h} of $\tilde{\mu}$ is Lipschitz continuous and bounded

away from both 0 and infinity on Δ_0^- . Moreover, as mentioned in Remark 2.3, $\tilde{\mu}|_{\Delta_0^+}$ is equal, up to a constant, to the measure $\hat{\mu}_+$ and so the density of $\tilde{\mu}|_{\Delta_0^+}$ is also Lipschitz continuous and bounded away from 0 and infinity. \square

Remark 2.5. We have used above the notation G rather than G_- for simplicity as this is the map which plays a more central role in our construction, see Remark 3.3 below. Similarly, we will from now on simply use the notation $\hat{\mu}$ to denote the measure $\hat{\mu}_-$.

2.2. Orbit distribution estimates. The second step of the argument is aimed at establishing conditions under which the measure $\tilde{\mu}$ is finite, and can therefore be renormalized to a probability measure $\mu := \tilde{\mu}/\tilde{\mu}(I)$, and aimed at studying the ergodic and statistical properties of μ . Our approach here differs even more significantly from existing approaches in the literature, although it does have some similarities with the argument of [Cri+10]: rather than starting with estimates of the tail of the inducing time (which would themselves anyway be significantly more involved than in the usual examples of intermittency maps with a single critical point due to our double inducing procedure), we carry out *more general estimates* on the *distribution* of iterates of points in I_- and I_+ before they return to Δ_0^- . More precisely, we define the functions $\tau^\pm(x) : \Delta_0^- \rightarrow \mathbb{N}$ by

$$\tau^+(x) := \#\{1 \leq i \leq \tau : g^i(x) \in I_+\}, \quad \text{and} \quad \tau^-(x) := \#\{1 \leq j \leq \tau : g^j(x) \in I_-\}. \tag{22}$$

These functions *count the number of iterates of x in I_- and I_+ respectively before returning to Δ_0^-* . Then for any $a, b \in \mathbb{R}$ we define *weighted combination* $\tau_{a,b} : \Delta_0^- \rightarrow \mathbb{R}$ by

$$\tau_{a,b}(x) = a\tau^+(x) + b\tau^-(x) \tag{23}$$

As we shall see as part of our construction of the induced map, both of these functions are *unbounded* and their level sets have a non-trivial structure in Δ_0^- and, moreover, the *inducing time* function $\tau : \Delta_0^- \rightarrow \mathbb{N}$ of the induced map G_- corresponds exactly to $\tau_{1,1}$ so that

$$\tau(x) = \tau_{1,1}(x) = \tau^+(x) + \tau^-(x). \tag{24}$$

The key results of this part of the proof consists of explicit and sharp asymptotic bounds for the distribution of $\tau_{a,b}$ for different values of a, b , from which we can then obtain as an immediate corollary the rates of decay of the inducing time function τ , and which will also provide the core estimates for the various distributional limit theorems. To state our results, let

$$B_1 := a_1^{-1/k_1}(\ell_2 b_2)^{-1/\beta_2} \quad \text{and} \quad B_2 := a_2^{-1/k_2}(\ell_1 b_1)^{-1/\beta_1}, \tag{25}$$

(the expressions defining the constants B_1, B_2 will appear in the proof of Proposition 3.5 below).

Recall from Corollary 2.4 that the density \tilde{h} of $\tilde{\mu}$ is bounded on $\Delta_0^- \cup \Delta_0^+$ and let $\tilde{h}(0^-)$ and $\tilde{h}(0^+)$ denote the values of this density on either side of 0. Then, for any $a, b \geq 0$, we let

$$C_a := \tilde{h}(0^-)B_1 a^{1/\beta_2}, \quad \text{and} \quad C_b := \tilde{h}(0^+)B_2 b^{1/\beta_1}. \tag{26}$$

Then we have the following distributional estimates.

Proposition 2.6. *Let $g \in \widehat{\mathfrak{F}}$. Then for every $a, b \geq 0$ we have the following distribution estimates.*

For every $\gamma \in [0, 1)$

$$\tilde{\mu}(a\tau^+ + b\tau^- > t) = \begin{cases} C_b t^{-1/\beta_1} + C_a t^{-1/\beta_2} + o(t^{-\gamma-1/\beta}) & \text{if } \ell_1, \ell_2 > 0 \\ C_b t^{-1/\beta_1} + o(t^{-\gamma-1/\beta_1}) & \text{if } \ell_1 > 0, \ell_2 = 0 \\ C_a t^{-1/\beta_2} + o(t^{-\gamma-1/\beta_2}) & \text{if } \ell_1 = 0, \ell_2 > 0 \\ O((1+b_1)^{-t/k_2} + (1+b_2)^{-t/k_1}) & \text{if } \ell_1 = 0, \ell_2 = 0 \end{cases} \tag{27}$$

$$\tilde{\mu}(a\tau^+ - b\tau^- > t) = \begin{cases} C_a t^{-1/\beta_2} + o(t^{-\gamma-1/\beta}) & \text{if } \ell_2 > 0 \\ O((1+b_2)^{-t/k_1}), & \text{if } \ell_2 = 0 \end{cases} \tag{28}$$

$$\tilde{\mu}(a\tau^+ - b\tau^- < -t) = \begin{cases} C_b t^{-1/\beta_1} + o(t^{-\gamma-1/\beta}) & \text{if } \ell_1 > 0 \\ O((1+b_1)^{-t/k_2}), & \text{if } \ell_1 = 0. \end{cases} \tag{29}$$

Remark 2.7. We have assumed in Proposition 2.6 that $a, b \geq 0$ to avoid stating explicitly too many cases, but one can easily read off the tails for $\tau_{a,b}$ for arbitrary $a, b \in \mathbb{R}$. For example, if $a < 0$ and $b > 0$ we can write $\tilde{\mu}(-a\tau^+ + b\tau^- > t) = \tilde{\mu}(a\tau^+ - b\tau^- < -t)$ and get the corresponding estimate from (29). Notice moreover, that the estimates for $\ell_1 = 0$ and/or $\ell_2 = 0$ are exponential.

Recall from Corollary 2.4 that $\hat{\mu} = \tilde{\mu}$ on the inducing domain Δ_{0^-} and therefore all the above estimates hold for $\hat{\mu}$ with exactly the same constants. In particular by Proposition 2.6 and (24), we immediately get the corresponding estimates for the tail $\hat{\mu}(\tau > t) = \tilde{\mu}(\tau > t)$.

Corollary 2.8. *If $\beta = 0$ then $\hat{\mu}(\tau > t)$ decay exponentially as $t \rightarrow +\infty$. If $\beta > 0$ then there exists a positive constant C_τ (which can be computed explicitly) such that*

$$\hat{\mu}(\tau > t) \sim C_\tau t^{-1/\beta}.$$

Proposition 2.6 will be proved in Sect.4.1, here we show how it implies Theorems B, C, D.

Proof of Theorems B, C, and D. From the definition of $\tilde{\mu}$ in (21) and since $g^{-n}(I) = I$ we have

$$\tilde{\mu}(I) := \sum_{n=0}^{\infty} \hat{\mu}_-(g^{-n}(I) \cap \{\tau > n\}) = \sum_{n=0}^{\infty} \hat{\mu}_-(I \cap \{\tau > n\}) = \sum_{n=0}^{\infty} \tilde{\mu}_-(\tau > n).$$

By Corollary 2.8, if $\beta = 0$, the quantities $\hat{\mu}_-(\tau > n)$ decay exponentially and, if $\beta > 0$ we have

$$\tilde{\mu}(I) = C \sum_{n=1}^{\infty} n^{-\frac{1}{\beta}} (1 + o(1)),$$

for some $C > 0$. This implies that $\tilde{\mu}(I) < \infty$ if and only if $\beta \in [0, 1)$, i.e. if and only if $g \in \mathfrak{F}$. Thus, for $g \in \mathfrak{F}$ we can define the measure $\mu_g := \tilde{\mu}/\tilde{\mu}(I)$, which is an invariant ergodic probability measure for g , and is unique because it is equivalent to Lebesgue, thus proving Theorem B. Theorem C follows from Theorem A in [AM21] by

noticing that $\mathcal{P} = \{(-1, 0), (0, 1)\}$ is a Lebesgue mod 0 generating partition such that $H_{\mu_g}(\mathcal{P}) < \infty$ and $h_{\mu_g}(g, \mathcal{P}) < \infty$, and therefore $h_{\mu_g}(g) < \infty$. Finally, Theorem D follows by well known results [You99] which show that the decay rate of the tail of the inducing times provides upper bounds for the rates of decay of correlations as stated. \square

2.3. Distribution of induced observables. The last part of our argument is focused on obtaining the limit theorems stated in Theorem E. When $\beta = 0$ the decay of correlations is exponential and the result follows from [YouRecurrenceTimesRates1999]. Similarly, after having established Proposition 2.2 and Corollary 2.8, the case that only one of ℓ_1, ℓ_2 is positive implies that there is only one intermittent fixed point, and thus essentially reduces to the argument given in [Gou04a, Theorem 1.3] for the LSV map. We only therefore need to consider the case that both $\ell_1, \ell_2 > 0$, which implies in particular that $\beta \in (0, 1)$.

Given an observable $\varphi : [0, 1] \rightarrow \mathbb{R}$, we define the *induced observable* $\Phi : \Delta_0^- \rightarrow \mathbb{R}$ by

$$\Phi(x) := \sum_{k=0}^{\tau(x)-1} \varphi \circ g^k.$$

Definition 2.9. We write $\Phi \in \mathcal{D}_\alpha$ if $\exists c_1, c_2 \geq 0$, with *at least one* of c_1, c_2 non-zero, such that

$$\hat{\mu}(\Phi > t) = c_1 t^{-\alpha} + o(t^{-\alpha}) \quad \text{and} \quad \hat{\mu}(\Phi < -t) = c_2 t^{-\alpha} + o(t^{-\alpha}), \quad (30)$$

In certain settings, limit theorems can be deduced from properties of the induced observable Φ . In particular, it is proved in Theorems 1.1 and 1.2 of [Gou04a] that, precisely in our setting²:

$$\text{if } \Phi \in L^2(\hat{\mu}) \text{ then } \varphi \text{ satisfies (CLT),} \quad (31)$$

$$\text{if } \Phi \in \mathcal{D}_2 \text{ then } \varphi \text{ satisfies (CLT}_{\text{ns}}), \quad (32)$$

$$\text{if } \Phi \in \mathcal{D}_\alpha \text{ with } \alpha \in (1, 2) \text{ then } \varphi \text{ satisfies } SL_\alpha. \quad (33)$$

We will argue that in each case of Theorem E, the induced observable Φ satisfies one of the above. To prove this, we first decompose a general observable $\varphi : [-1, 1] \rightarrow \mathbb{R}$ by letting $a := \varphi(1)$ and $b := \varphi(-1)$ and writing

$$\varphi = \varphi_{a,b} + \tilde{\varphi} \quad \text{where} \quad \varphi_{a,b} := b\chi_{[-1,0)} + a\chi_{[0,1]} \text{ and } \tilde{\varphi} := \varphi - \varphi_{a,b}, \quad (34)$$

where $\chi_{[-1,0)}$, $\chi_{[0,1]}$ are the characteristic functions of the intervals $[-1, 0)$ and $(0, 1]$ respectively. The induced observable of φ is the sum of the induced observables of $\varphi_{a,b}$ and $\tilde{\varphi}$ giving

$$\Phi(x) = \sum_{k=0}^{\tau(x)-1} \varphi_{a,b} \circ g^k(x) + \sum_{k=0}^{\tau(x)-1} \tilde{\varphi} \circ g^k(x) = \tau_{a,b} + \tilde{\Phi}, \quad (35)$$

² The assumptions of [Gou04a, Theorems 1.1 and 1.2] are that φ is Hölder continuous and G is an induced Gibbs-Markov map with invariant absolutely continuous probability measure $\hat{\mu}$ and return time satisfying $\hat{\mu}(\tau > n) = O(n^{-\gamma})$ for some $\gamma > 1$, which holds in our case by Corollary 2.8 and the fact that $\beta \in (0, 1)$.

where $\tilde{\Phi}$ denote the induced observable of $\tilde{\varphi}$, and $\tau_{a,b}$ is defined in (23), indeed, $\varphi_{a,b} \circ g^k(x)$ takes only two possible values, a or b , depending on whether $g^k(x) \in (0, 1]$ or $g^k(x) \in [-1, 0)$, and therefore the corresponding induced observable is precisely $\tau_{a,b}$.

To prove Theorem E we obtain regularity and distribution results for the induced observables $\tau_{a,b}$ and $\tilde{\Phi}$ and substitute them into (35) to get the various cases (31)-(33). The motivation for the decomposition (34) is given by the observation that $\tilde{\varphi}(-1) = \tilde{\varphi}(1) = 0$, which allows us to prove the following estimate for the corresponding induced observable $\tilde{\Phi}$.

Proposition 2.10. *Let $g \in \mathfrak{F}$ with $\beta \in (0, 1)$ and let $\tilde{\varphi} : [-1, 1] \rightarrow \mathbb{R}$ be a Hölder continuous observable such that $\tilde{\varphi}(-1) = \tilde{\varphi}(1) = 0$. Then*

$$(\mathcal{H}) \implies \tilde{\Phi} \in L^2 \quad \text{and} \quad (\mathcal{H}') \implies \hat{\mu}(\pm\tilde{\Phi} > t) = o(t^{-1/\beta_\varphi}). \tag{36}$$

Proposition 2.6 gives results for $\tau_{a,b}$.

Corollary 2.11 (Corollary to Proposition 2.6). *If at least one of $a:=\varphi(1)$, $b:=\varphi(-1)$ is non-zero then:*

$$\beta_\varphi \in [0, 1/2) \implies \tau_{a,b} \in L^2(\hat{\mu}), \quad \text{and} \quad \beta_\varphi \in [1/2, 1) \implies \tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}.$$

We prove Corollary 2.11 and Proposition 2.10 in Sect. 4.2. For now we show how they imply Theorem E.

Proof of Theorem E. If $\varphi(-1) = \varphi(1) = 0$ then $\tau_{a,b} \equiv 0$ and so $\Phi = \tilde{\Phi}$, Proposition 2.10 implies that $\Phi \in L^2(\hat{\mu})$ and so (31) holds. If at least one of $\varphi(-1)$, $\varphi(1)$ is non-zero, we have two cases. If $\beta_\varphi \in (0, 1/2)$, Proposition 2.10 and Corollary 2.11 give that both $\tau_{a,b}$, $\tilde{\Phi} \in L^2(\hat{\mu})$, which implies that $\Phi \in L^2(\hat{\mu})$ and therefore (31) holds. If $\beta_\varphi \in [1/2, 1)$ then $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$ by Corollary 2.11 and $\hat{\mu}(\pm\tilde{\Phi} > t) = o(t^{-1/\beta_\varphi})$ by Proposition 2.10, and therefore $\Phi = \tau_{a,b} + \tilde{\Phi} \in \mathcal{D}_{1/\beta_\varphi}$ since the tail of $\tilde{\Phi}$ is negligible compared to that of $\tau_{a,b}$. Whence, (32) holds when $\beta_\varphi = 1/2$ and (33) holds otherwise. \square

Remark 2.12. The relation between Φ and $\tau_{a,b}$ is given formally in (35) but it can be useful to have a heuristic idea of this relationships. Given a point $x \in \delta_{i,j}$ with i, j both large we know that most of the first i iterates $x, g(x), \dots, g^{i-1}(x)$ will lie near the fixed point 1. Similarly, most of the next j iterates $g^i(x), \dots, g^{i+j-1}$ will lie near the fixed point -1 . Thus, if we assume that φ is “sufficiently well behaved” near 1 and -1 (in a sense that is made precise by conditions (\mathcal{H}) and (\mathcal{H}')), it is reasonable to hope that the induced observable Φ at the point x will behave like $\Phi(x) = \sum_{k=0}^{n-1} \varphi \circ g^k \approx ai + bj = \tau_{a,b}(x)$ when $a = \varphi(1)$, $b = \varphi(-1)$ are not both zero.

3. The Induced Map

In this section we prove Proposition 2.2. We begin by recalling one of several essentially equivalent definitions of Gibbs-Markov map.

Definition 3.1. An interval map $F : I \rightarrow I$ is called a (full branch) Gibbs-Markov map if there exists a partition \mathcal{P} of $I \pmod 0$ into open subintervals such that:

1. F is full branch: for all $\omega \in \mathcal{P}$ the restriction $F|_\omega : \omega \rightarrow \text{int}(I)$ is a C^1 diffeomorphism;

2. F is *uniformly expanding*: there exists $\lambda > 1$ such that $|F'(x)| \geq \lambda$ for all $x \in \omega$ for all $\omega \in \mathcal{P}$;
3. F has *bounded distortion*: there exists $C > 0, \theta \in (0, 1)$ s.t. for all $\omega \in \mathcal{P}$ and all $x, y \in \omega$,

$$\log \left| \frac{F'(x)}{F'(y)} \right| \leq C\theta^{s(x,y)},$$

where $s(x, y) := \inf\{n \geq 0 : F^n x \text{ and } F^n y \text{ lie in different elements of the partition } \mathcal{P}\}$.

We will show that the first return map G defined in (19) satisfies all the conditions above as well as the saturation condition (20). In Sect. 3.1 we describe the topological structure of G and show that it is a full branch map with countably many branches which saturates I ; this will require only the very basic topological structure of g provided by condition (A0). In Sect. 3.2 we obtain estimates concerning the sizes of the partition elements of the corresponding partition; this will require the explicit form of the map g as given in (A1). In Sect. 3.3 we show that G is uniformly expanding; this will require the final condition (A2). Finally, in Sect. 3.4 we use the estimates and results obtained to show that G has bounded distortion.

3.1. Topological construction. In this section we give an explicit and purely topological construction of the first return maps $G^- : \Delta_0^- \rightarrow \Delta_0^-$ and $G^+ : \Delta_0^+ \rightarrow \Delta_0^+$ which essentially depends only on condition (A0), i.e. the fact that g is a full branch map with two orientation preserving branches. Recall first of all the definitions of the sets Δ_n^\pm and δ_n^\pm in (5) and (6). It follows immediately from the definitions and from the fact that each branch of g is a C^2 diffeomorphism, that for every $n \geq 1$, the maps $g : \delta_n^- \rightarrow \Delta_{n-1}^+$ and $g : \delta_n^+ \rightarrow \Delta_{n-1}^-$ are C^2 diffeomorphisms, and, for $n \geq 2$, the same is true for the maps $g^{n-1} : \Delta_{n-1}^- \rightarrow \Delta_0^-$, and $g^{n-1} : \Delta_{n-1}^+ \rightarrow \Delta_0^+$, which implies that for every $n \geq 1$, the maps

$$g^n : \delta_n^- \rightarrow \Delta_0^+ \quad \text{and} \quad g^n : \delta_n^+ \rightarrow \Delta_0^-$$

are C^2 diffeomorphisms. We can therefore define two maps

$$\tilde{G}^- : \Delta_0^- \rightarrow \Delta_0^+ \quad \text{and} \quad \tilde{G}^+ : \Delta_0^+ \rightarrow \Delta_0^- \quad \text{by} \quad \tilde{G}^\pm|_{\delta_n^\pm} := g^n. \tag{37}$$

Notice that these are *full branch* maps although they have *different domains and ranges*, indeed the domain of one is the range of the other and viceversa. The fact that they are full branch allows us to *pullback* the partition elements δ_n^\pm *into each other*: for every $m, n \geq 1$ we let

$$\delta_{m,n}^- := g^{-m}(\delta_n^+) \cap \delta_m^- \quad \text{and} \quad \delta_{m,n}^+ := g^{-m}(\delta_n^-) \cap \delta_m^+.$$

Then, for $m \geq 1$, the sets $\{\delta_{m,n}^-\}_{n \geq 1}$ and $\{\delta_{m,n}^+\}_{n \geq 1}$ are partitions of δ_m^- and δ_m^+ respectively and so

$$\mathcal{P}^- := \{\delta_{m,n}^-\}_{m,n \geq 1} \quad \text{and} \quad \mathcal{P}^+ := \{\delta_{m,n}^+\}_{m,n \geq 1} \tag{38}$$

are partitions of Δ_0^-, Δ_0^+ respectively, with the property that for every $m, n \geq 1$, the maps

$$g^{m+n} : \delta_{m,n}^- \rightarrow \Delta_0^- \quad \text{and} \quad g^{m+n} : \delta_{m,n}^+ \rightarrow \Delta_0^+ \tag{39}$$

are C^2 diffeomorphisms. Notice that $m + n$ is the *first return time* of points in $\delta_{m,n}^-$ and $\delta_{m,n}^+$ to Δ_0^- and Δ_0^+ respectively and we have thus constructed two *full branch first return induced maps*

$$G^- := \tilde{G}^+ \circ \tilde{G}^- : \Delta_0^- \rightarrow \Delta_0^- \quad \text{and} \quad G^+ := \tilde{G}^- \circ \tilde{G}^+ : \Delta_0^+ \rightarrow \Delta_0^+. \tag{40}$$

for which we have $G^-|_{\delta_{m,n}^-} = g^{m+n}$ and $G^+|_{\delta_{m,n}^+} = g^{m+n}$.

Lemma 3.2. *The maps G^- and G^+ are full branch maps which saturate I*

Proof. The full branch property follows immediately from (39). It then also follows from the construction that the families

$$\{g^j(\delta_{m,n}^-)\}_{\substack{m,n \geq 1 \\ 0 \leq j < m+n}} \quad \text{and} \quad \{g^j(\delta_{m,n}^+)\}_{\substack{m,n \geq 1 \\ 0 \leq j < m+n}}$$

of the images of the partition elements (38) are each formed by a collection of *pairwise disjoint* intervals which satisfy

$$\bigcup_{\delta_{m,n}^- \in \mathcal{P}^-} \bigcup_{j=0}^{m+n-1} g^j(\delta_{m,n}^-) = \bigcup_{\delta_{m,n}^+ \in \mathcal{P}^+} \bigcup_{j=0}^{m+n-1} g^j(\delta_{m,n}^+) = I \pmod 0$$

and therefore clearly satisfy (20), giving the saturation. □

Remark 3.3. Notice that the map G^- is exactly the first return map G defined in (19) and therefore Lemma 3.2 implies the first part of Proposition 2.2.

3.2. Partition estimates. The construction of the full branch induced maps $G^\pm : \Delta_0^\pm \rightarrow \Delta_0^\pm$ in the previous section is purely topological and works for any map g satisfying condition (A0). In this section we proceed to estimate the sizes and positions of the various intervals defined above, and this will require more information about the map, especially the forms of the map as given in (A1). Before stating the estimates we introduce some notation. First of all, we let $(x_n^-)_{n \geq 0}$ and $(x_n^+)_{n \geq 0}$ be the boundary points of the intervals Δ_n^-, Δ_n^+ so that $\Delta_0^- = (x_0^-, 0)$, $\Delta_0^+ = (0, x_0^+)$ and, for every $n \geq 1$ we have

$$\Delta_0^- = (x_0^-, 0), \quad \Delta_0^+ = (0, x_0^+), \quad \Delta_n^- = (x_n^-, x_{n-1}^-), \quad \Delta_n^+ = (x_{n-1}^+, x_n^+). \tag{41}$$

The following proposition gives the speed at which the sequences (x_n^+) , (x_n^-) converge to the fixed points $1, -1$ respectively and gives estimates for the size of the partition elements Δ_n^\pm for large n in terms of the values of ℓ_1 and ℓ_2 . To state the result we let

$$C_1 = (\ell_1 b_1)^{-1/\ell_1}, \quad C_2 = (\ell_2 b_2)^{-1/\ell_2}, \quad C_3 = \ell_1^{-(1+1/\ell_1)} b_1^{-1/\ell_1} \quad C_4 = \ell_2^{-(1+1/\ell_2)} b_2^{-1/\ell_2}.$$

Proposition 3.4. *If $\ell_1 = 0$, then*

$$(1 + b_1 + \epsilon)^{-n} \lesssim 1 + x_n^- \lesssim (1 + b_1)^{-n} \quad \text{and} \quad |\Delta_n^-| \lesssim (1 + b_1)^{-n}. \quad (42)$$

If $\ell_1 > 0$, then

$$1 + x_n^- \sim C_1 n^{-1/\ell_1} \quad \text{and} \quad |\Delta_n^-| \sim C_3 n^{-(1+1/\ell_1)}. \quad (43)$$

If $\ell_2 = 0$, then

$$(1 + b_2 + \epsilon)^{-n} \lesssim 1 - x_n^+ \lesssim (1 + b_2)^{-n} \quad \text{and} \quad |\Delta_n^+| \lesssim (1 + b_2)^{-n}. \quad (44)$$

If $\ell_2 > 0$, then

$$1 - x_n^+ \sim C_2 n^{-1/\ell_2} \quad \text{and} \quad |\Delta_n^+| \sim C_4 n^{-(1+1/\ell_2)}. \quad (45)$$

Proof. We will prove (44) and (45) and then (44) and (45) follow by exactly the same arguments. Notice first of all that from (7) we have $\delta_n^+ \subset U_{0+}$ for all $n > n_+$ and, since from (2) we have that $U_{-1} := g(U_{0+})$, this implies that $\Delta_{n-1}^- := g(\delta_n^+) \subset U_{-1}$ for all $n > n_+$, and thus $\Delta_n^- \subset U_{-1}$ for all $n \geq n_+$ which, by the definition of x_n in (41), implies that $x_n^- \in U_{-1}$ for $n \geq n_+$.

Now suppose that $\ell_1 > 0$. For $n \geq n_+$, by definition of the x_n^- , we have that $g(x_{n+1}^-) = x_n^-$, and so $1 + x_n^- = 1 + x_{n+1}^- + b_1(1 + x_{n+1}^-)^{1+\ell_1}$. Setting $z_n = 1 + x_n^-$ we can write this as $z_n = z_{n+1}(1 + b_1 z_{n+1}^{\ell_1})$ and, taking the power $-\ell_1$ and expanding we get

$$\frac{1}{z_n^{\ell_1}} = \frac{1}{z_{n+1}^{\ell_1}} (1 + b_1 z_{n+1}^{\ell_1})^{-\ell_1} = \frac{1}{z_{n+1}^{\ell_1}} \left(1 - \ell_1 b_1 z_{n+1}^{\ell_1} + O(z_{n+1}^{2\ell_1}) \right) = \frac{1}{z_{n+1}^{\ell_1}} - \ell_1 b_1 + o(1).$$

From the above we know that $z_n^{-\ell_1} = z_{n-1}^{-\ell_1} + b_1 \ell_1 + o(1)$ and applying this relation recursively we obtain that $z_n^{-\ell_1} = \ell_1 b_1 n + o(n)$ which yields $x_n^- + 1 = (\ell_1 b_1 n)^{-1/\ell_1} (1 + o(1))$, thus giving the first statement in (45). Now, by definition $\Delta_n^- = [x_n^-, x_{n-1}^-] = [x_n^-, g(x_n^-)]$, so, for all n large enough, $|\Delta_n^-| = g(x_n^-) - x_n^- = b_1(1 + x_n^-)^{1+\ell_1}$. Inserting $x_n^- + 1 \sim C_1 n^{-1/\ell_1}$ into this expression for $|\Delta_n^-|$ then yields $|\Delta_n^-| \sim \ell_1^{-(1+1/\ell_1)} b_1 n^{-1-1/\ell_1}$, completing the proof of (45).

Now, for $\ell_1 = 0$, since $g(x_n^-) = x_{n-1}^-$, the mean value theorem implies $(1 + b_1) \leq (x_{n-1}^- + 1)/(x_n^- + 1) \leq (1 + b_1 + o(1))$ which can be written as $(1 + b_1 + o(1))^{-1} (x_{n-1}^- + 1) \leq x_n^- + 1 \leq (1 + b_1)^{-1} (x_{n-1}^- + 1)$. Iterating this relation we obtain the claimed bounds for $x_n^- + 1$. As in the previous case we may calculate using (3) that $|\Delta_n^-| = g(x_n^-) - x_n^- = -1 + (1 + b_1)(1 + x_n^-) + \xi(x_n^-) - x_{n-1}^- = b_1(1 + x_n^-) + o(1) \lesssim (1 + b_1)^{-n}$, which concludes the proof. \square

To get analogous estimates for the intervals δ_n^-, δ_n^+ , we let $(y_n^-)_{n \geq 0}$ and $(y_n^+)_{n \geq 0}$ be the boundary points of the intervals δ_n^-, δ_n^+ respectively, so that for every $n \geq 1$ we have

$$\delta_n^- = (y_{n-1}^-, y_n^-) \quad \text{and} \quad \delta_n^+ = (y_n^+, y_{n-1}^+).$$

In particular, $y_0^- = x_0^-, y_0^+ = x_0^+$, and $g(y_n^-) = x_{n-1}^-, g(y_n^+) = x_{n-1}^-$ for $n \geq 1$. Then we let

$$B_1 = a_1^{-1/k_1} (\ell_2 b_2)^{-1/\beta_2}, \quad B_2 = a_2^{-1/k_2} (\ell_1 b_1)^{-1/\beta_1}, \quad B_3 = B_1/\beta_2, \quad B_4 = B_2/\beta_1.$$

Recall that B_1, B_2 have already been defined in (25).

Proposition 3.5. *If $\ell_1 = 0$, then for every $\varepsilon > 0$*

$$\left(\frac{1}{1+b_2+\varepsilon}\right)^{\frac{n}{k_1}} \lesssim -y_n^- \lesssim \left(\frac{1}{1+b_2}\right)^{\frac{n}{k_1}} \quad \text{and} \quad |\delta_n^-| \lesssim \left(\frac{1}{1+b_2}\right)^n. \tag{46}$$

If $\ell_1 > 0$, then

$$y_n^- \sim -B_1 n^{-\frac{1}{\beta_2}}, \quad \text{and} \quad |\delta_n^-| \sim B_3 n^{-\left(1+\frac{1}{\beta_2}\right)}. \tag{47}$$

If $\ell_2 = 0$, then for every $\varepsilon > 0$

$$\left(\frac{1}{1+b_1+\varepsilon}\right)^{\frac{n}{k_2}} \lesssim y_n^+ \lesssim \left(\frac{1}{1+b_1}\right)^{\frac{n}{k_2}}, \quad \text{and} \quad |\delta_n^+| \lesssim \left(\frac{1}{1+b_1}\right)^n. \tag{48}$$

If $\ell_2 > 0$, then

$$y_n^+ \sim B_2 n^{-\frac{1}{\beta_1}}, \quad \text{and} \quad |\delta_n^+| \sim B_4 n^{-\left(1+\frac{1}{\beta_1}\right)}, \tag{49}$$

Proof. We will prove (48) and (47), as (48) and (49) follow by analogous arguments. Suppose first that $\ell_1 > 0$. As $x_n^+ \rightarrow 1$, and as $g_-(y_n^-) = x_{n-1}^+$ we know that for all n sufficiently large we have $g_-(y_n^-) = 1 - a_1(-y_n^-)^{k_1} = x_{n-1}^+$. Solving for y_n this gives

$$y_n^- = -\left((1 - x_{n-1}^+)/a_1\right)^{1/k_1} = -a_1^{-1/k_1} (\ell_2 b_2 n)^{-1/\ell_2 k_1} (1 + o(1))$$

which is the first statement in (47). Now we turn our attention to the size of the intervals δ_n^- . First let us note that for any $\gamma > 0$ we have that $n^{-\gamma} - (n+1)^{-\gamma} = n^{-\gamma} [1 - (1+1/n)^{-\gamma}] = n^{-\gamma} [1 - (1 - \gamma/n + O(n^{-2}))] = \gamma n^{-(1+\gamma)} (1 + O(1/n))$ and therefore

$$|\delta_n| = y_n^- - y_{n+1}^- = B_1 (n^{-1/\ell_2 k_1} - (n+1)^{-1/\ell_2 k_1}) (1 + o(1)) = \frac{B_1}{\ell_2 k_1} n^{-(1+1/\ell_2 k_1)} (1 + o(1))$$

which completes the proof of (47). Now for $\ell_2 = 0$ we proceed as before, and by (44) we get

$$(1 + b_2 + \varepsilon)^{-n/k_1} \lesssim -y_n^- = \left((1 - x_{n-1}^+)/a_1\right)^{1/k_1} \lesssim (1 + b_2)^{-n/k_1}.$$

For the size of the interval δ_n^- , we may use the mean value theorem to conclude that

$$g'(u_n) = \frac{x_{n-2}^+ - x_{n-1}^+}{y_{n-1}^- - y_n^-} = \frac{|\Delta_{n-1}^+|}{|\delta_n^-|},$$

for some $u_n \in \delta_n^-$. As g' is monotone on U_0^- we know, from the above and (44), that

$$|\delta_n^-| \lesssim |\Delta_{n-1}^+|/g'(y_n^-) \lesssim (1 + b_2 + \varepsilon)^{-n/k_1} (1 + b_2)^{-n} \lesssim (1 + b_2)^{-n}.$$

which concludes the proof. □

$$\begin{array}{ccc}
 \delta_{n+1}^+ & \xrightarrow{\phi} & \delta_n^+ \\
 \downarrow g & & \downarrow g \\
 \Delta_n^- & \xrightarrow{g} & \Delta_{n-1}^-
 \end{array}$$

Fig. 4. Definition of the map $\phi : \Delta_0^+ \setminus \delta_1^+ \rightarrow \Delta_0^+$

3.3. Expansion estimates.

Proposition 3.6. *For every $g \in \mathfrak{F}$ the first return map $G : \Delta_0^- \rightarrow \Delta_0^-$ is uniformly expanding.*

It is enough to prove uniform expansivity for the two maps \tilde{G}^-, \tilde{G}^+ , recall (37), since this implies the same property for their composition $\tilde{G} = \tilde{G}^-$, recall (40). To simplify the notation we will only prove the statement for \tilde{G}^+ , i.e. we will prove that $x \in \delta_n^+ \Rightarrow (g^n)'(x) > \lambda$. The fact that $x \in \delta_n^- \Rightarrow (g^n)'(x) > \lambda$ follows by an identical argument.

Notice first of all that if $k_2 \in (0, 1)$, the derivative in Δ_0^+ is greater than 1 and therefore the uniform expansion of \tilde{G}^+ is immediate. So throughout this section we will assume that $k_2 \geq 1$ which means that g has a critical point and the derivative of g in Δ_0^+ can be arbitrarily small. For points outside the neighbourhood U_{0+} on which the map g has a precise form, more precisely for $1 \leq n \leq n_+$ and for $x \in \delta_n^+$, the expansivity is automatically guaranteed by condition (A2), but for points close to 0 where the derivative can be arbitrarily small the statement is non-trivial. It ultimately depends on writing $\tilde{G}^+(x) := g^n(x)$ for $x \in \delta_n^+$, so that $(\tilde{G}^+)'(x) = (g^n)'(x) = (g^{n-1})'(g(x))g'(x)$, and then showing that the small derivative $g'(x)$ near the critical point is compensated by sufficiently large number of iterates where the derivative is > 1 . This clearly relies very much on the partition estimates in Sect. 3.2 which provide a relation between the position of points, and therefore their derivatives, and the corresponding values of n . A relatively straightforward computation using those estimates shows that we get expansion for sufficiently large $n \geq 1$, which is quite remarkable but not enough for our purposes as it does not give a complete proof of expansivity for \tilde{G}^+ at every point in Δ_0^+ . We therefore need to use a somewhat more sophisticated approach that shows that the derivative of \tilde{G}^+ has a kind of “monotonicity” property in the following sense. Define the function $\phi : \Delta_0^+ \setminus \delta_1^+ \rightarrow \Delta_0^+$ given implicitly by $g^2 = g \circ \phi$ and explicitly by

$$\phi := (g|_{U_{0+}})^{-1} \circ g|_{U_{-1}} \circ g|_{U_{0+}} \tag{50}$$

Notice that ϕ is the bijection which makes the diagram in Fig. 4 commute.

The key step in the proof of Proposition 3.6 is the following lemma.

Lemma 3.7. *For all $n \geq n^+$ and $x \in \delta_{n+1}^+$ we have*

$$(g^2)'(x) > g'(\phi(x)).$$

Remark 3.8. Lemma 3.7 is equivalent to $(g^2)'(x)/g'(\phi(x)) > 1$ which is equivalent to

$$\frac{g'(x)}{g'(\phi(x))}g'(g(x)) > 1. \tag{51}$$

Notice that the ratio $g'(x)/g'(\phi(x))$ is < 1 and measures how much derivative is “lost” when choosing the initial condition x instead of the initial condition $\phi(x)$ (since $\phi(x) > x$ and the derivative is monotone increasing), whereas $g'(g(x)) > 1$ measures how much derivative is “gained” from performing an extra iteration of g . The Lemma says that the gain is more than the loss.

Proof. To simplify the notation let us set $a = a_2, b = b_1, k = k_2$, and $\ell = \ell_1$. Notice first of all that by the form of g in U_{0+} given in (A1) we have

$$\frac{g'(x)}{g'(\phi(x))} = \left(\frac{x}{\phi(x)}\right)^{k-1} = \left(\frac{\phi(x)}{x}\right)^{1-k} \tag{52}$$

Recall that $k > 1$ and $x < \phi(x)$ and so the ratio above is < 1 . To estimate $g'(g(x))$ we consider two cases depending on ℓ . If $\ell > 0$, using the form of g given in (A1) and plugging into 50 we get

$$\phi(x) = \left[x^k + ba^\ell x^{k(\ell+1)}\right]^{1/k} \quad \text{and therefore} \quad \left(\frac{\phi(x)}{x}\right)^k = 1 + ba^\ell x^{k\ell}$$

and, therefore, using the form of G in U_{-1} , this gives

$$g'(g(x)) = g'(-1 + ax^k) = 1 + ba^\ell x^{k\ell} + b\ell a^\ell x^{k\ell} = \left(\frac{\phi(x)}{x}\right)^k + b\ell a^\ell x^{k\ell}. \tag{53}$$

From (52) and (53) and the fact that $x < \phi(x)$ we immediately get

$$\frac{g'(x)}{g'(\phi(x))}g'(g(x)) = \left(\frac{\phi(x)}{x}\right)^{1-k} \left[\left(\frac{\phi(x)}{x}\right)^k + b\ell a^\ell x^{k\ell}\right] > \frac{\phi(x)}{x} > 1$$

which establishes (51) and completes the case that $\ell > 0$. For $\ell = 0$, proceeding as above we obtain

$$\begin{aligned} \phi(x) &= \left[(1+b)x^k + \xi(g(x))/a\right]^{1/k} \quad \text{and therefore} \\ \left(\frac{\phi(x)}{x}\right)^k &= (1+b) + \frac{\xi(g(x))}{ax^k}. \end{aligned} \tag{54}$$

Since $g(x) = -1 + ax^k$, from (13) we have $\xi'(g(x)) \geq \xi(g(x))/(1+g(x)) = \xi(g(x))/ax^k$, and so

$$g'(g(x)) = (1+b) + \xi'(g(x)) \geq (1+b) + \frac{\xi(g(x))}{ax^k} = \left(\frac{\phi(x)}{x}\right)^k.$$

Together with (52), as above, we get the statement in this case also. □

As an almost immediate consequence of Lemma 3.7 we get the following.

Corollary 3.9. For all $n \geq n^+$ and $x \in \delta_{n+1}^+$ we have

$$(\tilde{G}^+)'(x) > (\tilde{G}^+)'(\phi(x)).$$

Proof. By Lemma 3.7 and (51), for any $1 \leq m \leq n$ we have

$$(g^{m+1})'(x) = g'(x)g'(g(x)) \cdots g'(g^m(x)) = \frac{g'(x)g'(g(x))}{g'(\phi(x))} (g^m)'(\phi(x)) > (g^m)'(\phi(x)). \tag{55}$$

□

Proof of Proposition 3.6. Condition (A2) implies that $(\tilde{G}^+)'(x) \geq \lambda$ for all $x \in \delta_n^+$ for $1 \leq n \leq n^+$. Then, for $x \in \delta_{n+1}^+$ we have $\phi(x) \in \delta_n^+$ and therefore

$$(\tilde{G}^+)'(x) > (\tilde{G}^+)'(\phi(x)) \geq \lambda$$

Proceeding inductively we obtain the result. □

3.4. Distortion estimates.

Proposition 3.10. For all $g \in \widehat{\mathfrak{S}}$ there exists a constant $\mathfrak{D} > 0$ such that for all $0 \leq m < n$ and all $x, y \in \delta_n^\pm$,

$$\log \frac{(g^{n-m})'(g^m(x))}{(g^{n-m})'(g^m(y))} \leq \mathfrak{D}|g^n(x) - g^n(y)|.$$

As a consequence we get that G is a Gibbs-Markov map with constants $C = \mathfrak{D}\lambda$ and $\theta = \lambda^{-1}$.

Corollary 3.11. For all $x, y \in \delta_{i,j} \in \mathcal{P}$ with $x \neq y$ we have

$$\log \left| \frac{G'(x)}{G'(y)} \right| \leq \mathfrak{D}\lambda^{-s(x,y)+1}.$$

Proof. Let $n:=s(x, y)$. Since G is uniformly expanding, we have $1 \geq |G^n(x) - G^n(y)| = |(G^{n-1})'(u)||G(x) - G(y)| \geq \lambda^{n-1}|G(x) - G(y)|$ and therefore $|G(x) - G(y)| \leq \lambda^{-n+1}$. By Proposition 3.10 this gives $\log |G'(x)/G'(y)| \leq \mathfrak{D}|G(x) - G(y)| \leq \mathfrak{D}\lambda^{-n+1} = \mathfrak{D}\lambda^{-s(x,y)+1}$. □

Proof of Proposition 3.10. We begin with a couple of simple formal steps. First of all, by the chain rule, we can write

$$\log \frac{(g^{n-m})'(g^m(x))}{(g^{n-m})'(g^m(y))} = \log \prod_{i=m}^{n-1} \frac{g'(g^i(x))}{g'(g^i(y))} = \sum_{i=m}^{n-1} \log \frac{g'(g^i(x))}{g'(g^i(y))}.$$

Then, since $g^i(x), g^i(y)$ are both in the same smoothness component of g , by the Mean Value Theorem, there exists $u_i \in (g^i(x), g^i(y))$ such that

$$\log \frac{g'(g^i(x))}{g'(g^i(y))} = \log g'(g^i(x)) - \log g'(g^i(y)) = \frac{g''(u_i)}{g'(u_i)} |g^i(x) - g^i(y)|.$$

Substituting this into the expression above, and writing $\mathfrak{D}_i := g''(u_i)/g'(u_i)$ for simplicity, we get

$$\log \frac{(g^{n-m})'(g^m(x))}{(g^{n-m})'(g^m(y))} = \sum_{i=m}^{n-1} \mathfrak{D}_i |g^i(x) - g^i(y)| \leq \sum_{i=0}^{n-1} \mathfrak{D}_i |g^i(x) - g^i(y)|. \tag{56}$$

We will bound the sum above in two steps. First of all we will show that it admits a uniform bound $\widehat{\mathfrak{D}}$ independent of m, n . We will then use this bound to improve our estimates and show that by paying a small price (increasing the uniform bound to a larger bound $\mathfrak{D} := \widehat{\mathfrak{D}}^2/|\Delta_0^-|$) we can include the term $|g^n(x) - g^n(y)|$ as required. Ultimately this gives a stronger result since it takes into account the closeness of the points x, y .

Let us suppose first for simplicity that $x, y \in \delta_n^+$, the estimates for δ_n^- are identical. Then for $1 \leq i < n$ we have that $g^i(x), g^i(y), u_i \in \Delta_{n-i}^-$ and therefore we can bound (56) by

$$\sum_{i=0}^{n-1} \mathfrak{D}_i |g^i(x) - g^i(y)| \leq \mathfrak{D}_0 |x - y| + \sum_{i=1}^{n-1} \mathfrak{D}_i |g^i(x) - g^i(y)| \leq \mathfrak{D}_0 \delta_n^+ + \sum_{i=1}^{n-1} \mathfrak{D}_i |\Delta_{n-i}^-|. \tag{57}$$

From (12) and using the relationship between the y_n^+ and the x_n^- we may bound the first term by

$$\mathfrak{D}_0 |\delta_n^+| \lesssim u_0^{-1} |\delta_n^+| \lesssim \frac{y_n^+ - y_{n+1}^+}{y_{n+1}^+} \lesssim \left(\frac{1 + x_n^-}{1 + x_{n+1}^-} \right)^{1/k} - 1 \rightarrow c < \infty \tag{58}$$

where we have used the fact that for some sequence $\xi_n \rightarrow -1$ we have $(1 + x_n^-)/(1 + x_{n+1}^-) = g'(\xi_n)$ which converges to 1 if $\ell > 0$ (and therefore $c = 0$) or $1 + b_1$ otherwise (and therefore $c = b_1$). If $\ell_1 = 0$ then \mathfrak{D}_i is uniformly bounded for $i > 0$, if $\ell_1 > 0$, then from (11) and (45) we know that

$$\mathfrak{D}_i |\Delta_{n-i}^-| \lesssim (1 + u_i)^{\ell_1 - 1} |\Delta_{n-i}^-| \lesssim (n - i)^{-\frac{\ell_1 - 1}{\ell_1}} (n - i)^{-(1 + \frac{1}{\ell_1})} = (n - i)^{-2}. \tag{59}$$

Then by (58) and (59) we find that

$$\widehat{\mathfrak{D}} := \exp \left\{ \mathfrak{D}_0 \delta_n^+ + \sum_{i=1}^{\infty} \mathfrak{D}_i |\Delta_{n-i}^-| \right\} \leq \exp \left\{ \mathfrak{D}_0 + \sum_{i=1}^{\infty} \mathfrak{D}_i |\Delta_{n-i}^-| \right\} < \infty. \tag{60}$$

Substituting this back into (57) and then into (56) we get

$$\log \left| \frac{(g^{n-m})'(g^m(x))}{(g^{n-m})'(g^m(y))} \right| \leq \log \widehat{\mathfrak{D}} \tag{61}$$

which completes the first step in the proof, as discussed above. We now take advantage of this bound to improve our estimates as follows. By a standard and straightforward application of the Mean Value Theorem, (61) implies that the diffeomorphisms $g^n : \delta_n^+ \rightarrow \Delta_0^-$ and $g^{n-m} : \Delta_{n-m}^- \rightarrow \Delta_0^-$ all have uniformly bounded distortion in the sense that for every $x, y \in \delta_n^+$ and $1 \leq m < n$ we have

$$\frac{|x - y|}{|\delta_n^+|} \leq \widehat{\mathfrak{D}} \frac{|g^n(x) - g^n(y)|}{|\Delta_0^-|} \tag{62}$$

and

$$\frac{|g^m(x) - g^m(y)|}{|\Delta_{n-m}^+|} \leq \widehat{\mathfrak{D}} \frac{|g^{n-m}(g^m(x)) - g^{n-m}(g^m(y))|}{|\Delta_0^-|} = \widehat{\mathfrak{D}} \frac{|g^n(x) - g^n(y)|}{|\Delta_0^-|}. \quad (63)$$

Therefore

$$|x - y| \leq \frac{\widehat{\mathfrak{D}}}{|\Delta_0^-|} |g^n(x) - g^n(y)| |\delta_n^+| \quad \text{and} \quad |g^m(x) - g^m(y)| \leq \frac{\widehat{\mathfrak{D}}}{|\Delta_0^-|} |g^n(x) - g^n(y)| |\Delta_{n-m}^-|.$$

Substituting these bounds back into (56) (with $i = m$), and letting $\mathfrak{D} := \widehat{\mathfrak{D}}^2/|\Delta_0^-|$, we get

$$\begin{aligned} \log \frac{(g^{n-m})'(g^m(x))}{(g^{n-m})'(g^m(y))} &\leq \sum_{i=0}^{n-1} \mathfrak{D}_i |g^i(x) - g^i(y)| = \mathfrak{D}_0 |x - y| + \sum_{i=1}^{n-1} \mathfrak{D}_i |g^i(x) - g^i(y)| \\ &\leq \mathfrak{D}_0 \frac{\widehat{\mathfrak{D}}}{|\Delta_0^-|} |g^n(x) - g^n(y)| |\delta_n^+| + \sum_{i=1}^{n-1} \mathfrak{D}_i \frac{\widehat{\mathfrak{D}}}{|\Delta_0^-|} |g^n(x) - g^n(y)| |\Delta_{n-i}^-| \\ &= \frac{\widehat{\mathfrak{D}}}{|\Delta_0^-|} \left[\mathfrak{D}_0 |\delta_n^+| + \sum_{i=1}^{n-1} \mathfrak{D}_i |\Delta_{n-i}^-| \right] |g^n(x) - g^n(y)| \\ &\leq \frac{\widehat{\mathfrak{D}}^2}{|\Delta_0^-|} |g^n(x) - g^n(y)| = \mathfrak{D} |g^n(x) - g^n(y)|. \end{aligned}$$

Notice that the last inequality follows from (60). This completes the proof. □

We state here also a simple corollary of Propositions 3.5 and 3.10 which we will use in Sect. 4.

Lemma 3.12. *For all $i, j \geq 1$ we have*

$$\tilde{\mu}(\delta_{i,j}) \lesssim \begin{cases} i^{-(1+1/\beta_2)} j^{-(1+1/\beta_1)}, & \text{if } \ell_1, \ell_2 > 0 \\ (1 + b_2)^{-i}, & \text{if } \ell_1 = 0, \ell_2 > 0 \\ (1 + b_1)^{-j}, & \text{if } \ell_2 = 0, \ell_1 > 0 \\ \min\{(1 + b_1), (1 + b_2)\}^{-i-j}, & \text{if } \ell_1 = 0, \ell_2 = 0 \end{cases} \quad (64)$$

Proof. Proposition 3.10 implies that $|\delta_{ij}| \approx |\delta_i^-| |\delta_j^+|$ uniformly for all $i, j \geq 1$. Indeed, more precisely, it implies $\mathfrak{D}^{-1} |g^i(\delta_{ij})|/|g^i(\delta_i)| \leq |\delta_{ij}|/|\delta_i| \leq \mathfrak{D} |g^i(\delta_{ij})|/|g^i(\delta_i)|$ which implies $|\delta_i^-| |\delta_j^+|/\mathfrak{D} |\Delta_0^+| \leq |\delta_{ij}| \leq \mathfrak{D} |\delta_i^-| |\delta_j^+|/|\Delta_0^+|$. As $\tilde{\mu}$ is equivalent to Lebesgue on $\Delta_0^- \cup \Delta_0^+$ we obtain the Lemma immediately from Proposition 3.5. □

4. Statistical Properties

In Sect. 4.1 we prove Proposition 2.6 and in Sect. 4.2 we prove Proposition 2.10 and Corollary 2.11. As discussed in Sect. 2.3 this completes the proof of Theorem E.

4.1. *Distribution and tail estimates.* In this section we prove Proposition 2.6. We will only explicitly prove (27) and (28) as the proof of (29) is identical to that of (28). For $a, b \geq 0$ consider the following decompositions

$$\tilde{\mu}(a\tau^+ + b\tau^- > t) = \tilde{\mu}(a\tau^+ > t) + \tilde{\mu}(b\tau^- > t) \tag{65}$$

$$- \tilde{\mu}(a\tau^+ > t, b\tau^- > t) + \tilde{\mu}(a\tau^+ + b\tau^- > t, \max\{a\tau^+, b\tau^-\} \leq t) \tag{66}$$

and

$$\tilde{\mu}(a\tau^+ - b\tau^- > t) = \tilde{\mu}(a\tau^+ > t) \tag{67}$$

$$- \tilde{\mu}(a\tau^+ > t, a\tau^+ - b\tau^- \leq t). \tag{68}$$

We can then reduce the proof to two further Propositions. First of all we give precise asymptotic estimates of the terms $\mu(a\tau^+ > t)$, $\mu(b\tau^- > t)$ which make up (65) and (67).

Proposition 4.1. *For every $a, b \geq 0$ and for every $\gamma \in (0, 1)$*

$$\tilde{\mu}(a\tau^+ > t) = \begin{cases} C_a t^{-1/\beta_2} + o(t^{-\gamma-1/\beta_2}) & \text{if } \ell_1 > 0 \\ O((1 + b_2)^{-t/ak_1}) & \text{if } \ell_1 = 0 \end{cases} \tag{69}$$

and

$$\tilde{\mu}(b\tau^- > t) = \begin{cases} C_b t^{-1/\beta_1} + o(t^{-\gamma-1/\beta_1}) & \text{if } \ell_2 > 0 \\ O((1 + b_1)^{-t/bk_2}) & \text{if } \ell_2 = 0 \end{cases} \tag{70}$$

Then, we show that the remaining terms (66) and (68) in the decompositions above have negligible contribution to the leading order asymptotics of the tail.

Proposition 4.2. *If at least one of ℓ_1, ℓ_2 are not zero, then for every $a, b \geq 0, \gamma \in (0, 1)$ we have*

$$\tilde{\mu}(a\tau^+ > t, b\tau^- > t) + \tilde{\mu}(a\tau^+ + b\tau^- > t, \max\{a\tau^+, b\tau^-\} \leq t) = o(t^{-\gamma-1/\beta}) \tag{71}$$

and

$$\tilde{\mu}(a\tau^+ > t, a\tau^+ - b\tau^- \leq t) = o(t^{-\gamma-1/\beta}) \tag{72}$$

As we shall see, (72) actually holds for all $\gamma \in (0, 1/\beta)$ (where $1/\beta > 1$ since $\beta \in (0, 1)$ by assumption) but we will not need this stronger statement. We prove Proposition 4.1 in Sect. 4.1.1 and Proposition 4.2 in Sect. 4.1.2, but first we show how they imply Proposition 2.6.

Proof of Proposition 2.6. To prove (27), first suppose that least one of ℓ_1, ℓ_2 is non-zero. Substituting the corresponding lines of (70) and (70) into (66) and substituting (71) into (68) we obtain (27) in this case. If $\ell_1 = \ell_2 = 0$ we only need to establish an upper bound for $\tilde{\mu}(a\tau^+ + b\tau^- > t)$ rather than an asymptotic equality and therefore, instead of the decomposition in (65) and (66), we can use the fact that

$$\tilde{\mu}(a\tau^+ + b\tau^- > t) \leq \tilde{\mu}(a\tau^+ > t) + \tilde{\mu}(b\tau^- > t) \tag{73}$$

The result then follows by inserting the corresponding lines of (70) and (70) into (73).

To prove (28), if $\ell_2 > 0$ the result follows by substituting the corresponding line of (70) into (67) and substituting (72) into (68). Again, if $\ell_2 = 0$ we only need to establish an upper bound for $\tilde{\mu}(a\tau^+ - b\tau^- > t)$ rather than an asymptotic equality and therefore, instead of the decomposition in (67) and (68), we can use the fact that

$$\tilde{\mu}(a\tau^+ - b\tau^- > t) \leq \tilde{\mu}(a\tau^+ > t) \tag{74}$$

The result then follows by inserting the corresponding line of (70) into (74). \square

4.1.1. Leading order asymptotics We prove Proposition 4.1 via two lemmas which show in particular how the values $\tilde{h}(0^-)$, $\tilde{h}(0^+)$ of the density of the measure $\tilde{\mu}$ turn up in the constants C_a, C_b defined in (26). Our first lemma shows that the tails of the distributions $\tilde{\mu}(\tau^+ > t)$ and $\tilde{\mu}(\tau^- > t)$ have a very geometric interpretation.

Lemma 4.3. *For every $t > 0$ we have*

$$\tilde{\mu}(\tau^+ > t) = \tilde{\mu}(y_{[t]}^-, 0) \quad \text{and} \quad \tilde{\mu}(\tau^- > t) = \tilde{\mu}(0, y_{[t]}^+). \tag{75}$$

Remark 4.4. While the first statement in (75) is relatively straightforward, the second statement is not at all obvious since τ^- is defined on Δ_0^- and there is no immediate connection with the interval $(0, y_{[t]}^+)$ in Δ_0^+ . As we shall see, the proof of Lemma 4.3 requires a subtle and interesting argument.

Remark 4.5. Since $\tilde{\mu}$ is equivalent to Lebesgue measure on Δ_0^- and Δ_0^+ , we immediately have that $\tilde{\mu}(y_{[t]}^-, 0) \approx |y_{[t]}^-|$ and $\tilde{\mu}(0, y_{[t]}^+) \approx y_{[t]}^+$, and we can then use (47) and (49), and Lemma 4.3, to get upper bounds for the distributions $\tilde{\mu}(\tau^+ > t)$ and $\tilde{\mu}(\tau^- > t)$. This is however not enough for our purposes as we require sharper estimates for the distributions, and we therefore need a more sophisticated argument which yields the statement in the following lemma.

Lemma 4.6. *For every $t > 0$ we have*

$$\tilde{\mu}(y_{[t]}^-, 0) = y_{[t]}^- \tilde{h}(0^-) + O((y_{[t]}^-)^2) \quad \text{and} \quad \tilde{\mu}(0, y_{[t]}^+) = y_{[t]}^+ \tilde{h}(0^+) + O((y_{[t]}^+)^2).$$

Before proving these two lemmas we show how they imply Proposition 4.1.

Proof of Proposition 4.1. Let us first show (70). Recall from the definition of C_a in (26) that $a = 0 \Rightarrow C_a = 0$, so if $a = 0$ there is nothing to prove. Let us suppose then that $a > 0$. By Lemmas 4.3 and 4.6 we have

$$\tilde{\mu}(a\tau^+ > t) = \tilde{\mu}(\tau^+ > t/a) = y_{[t/a]}^- \tilde{h}(0^-) + O((y_{[t/a]}^-)^2).$$

Then, using the asymptotic estimates (48) and (47) for y_n^- in Proposition 3.5; and since $O(t^{-2/\beta_1}) = o(t^{-\gamma-1/\beta_1})$ for every $\gamma \in (0, 1)$; and by the definition of C_a in (26), we obtain

$$\mu(a\tau^+ > t) = \begin{cases} B_1 \tilde{h}(0^-)(t/a)^{-1/\beta_2} + O(t^{-2/\beta_2}) = C_a t^{-1/\beta_2} + o(t^{-\gamma-1/\beta_2}) & \text{if } \ell_1 > 0 \\ O((1+b_2)^{-t/ak_1}) & \text{if } \ell_1 = 0 \end{cases}$$

yielding (70). To show (70) we can proceed similarly to the above. As before, if $b = 0$ there is nothing to prove so we assume $b > 0$ in which case Lemmas 4.3 and 4.6 we have $\tilde{\mu}(b\tau^- > t) = \tilde{\mu}(\tau^- > t/b) = y_{[t/b]}^+ \tilde{h}(0^+) + O((y_{[t/b]}^+)^2)$. Now using (48) and (49), and arguing as above we find that (29) holds for every $\gamma \in (0, 1)$. \square

We complete this section with the proofs of Lemmas 4.3 and 4.6.

Proof of Lemma 4.3. By definition, recall (22), $\tau^+(x) = i$, $\tau^-(x) = j$ for all $x \in \delta_{i,j}$, and therefore

$$\tilde{\mu}(\tau^+ > t) = \sum_{i>t} \sum_{j=1}^{\infty} \tilde{\mu}(\delta_{i,j}) \quad \text{and} \quad \tilde{\mu}(\tau^- > t) = \sum_{j>t} \sum_{i=1}^{\infty} \tilde{\mu}(\delta_{i,j}). \tag{76}$$

We claim that for every $i, j \geq 1$ we have

$$\sum_{j=1}^{\infty} \tilde{\mu}(\delta_{i,j}) = \tilde{\mu}(\delta_i^-) \quad \text{and} \quad \sum_{i=1}^{\infty} \tilde{\mu}(\delta_{i,j}) = \tilde{\mu}(\delta_j^+). \tag{77}$$

Then, substituting (77) into (76) we get

$$\tilde{\mu}(\tau^+ > t) = \sum_{i>t} \sum_{j=1}^{\infty} \tilde{\mu}(\delta_{i,j}) = \sum_{i>t} \tilde{\mu}(\delta_i^-) = \tilde{\mu}(y_{\lceil t \rceil}^-, 0), \tag{78}$$

and

$$\tilde{\mu}(\tau^- > t) = \sum_{j>t} \sum_{i=1}^{\infty} \tilde{\mu}(\delta_{i,j}) = \sum_{j>t} \tilde{\mu}(\delta_j^+) = \tilde{\mu}(0, y_{\lceil t \rceil}^+)$$

which is exactly the statement (75) in the Lemma.

Thus it only remains to prove (77). As already mentioned in Remark 4.4, despite the apparent symmetry between the two statements, the situation in the two expressions is actually quite different. Indeed, from the topological construction of the induced map, for each $i \geq 1$ we have

$$\delta_i^- = \bigcup_{j=1}^{\infty} \delta_{i,j} \tag{79}$$

which, since the intervals $\delta_{i,j}$ are pairwise disjoint, clearly implies the first equality in (77). The second equality is not immediate since, for each fixed $j \geq 1$, the intervals $\delta_{i,j}$ are spread out in Δ_0^- , with each $\delta_{i,j}$ lying inside the corresponding interval δ_i^- , and indeed the $\delta_{i,j}$ do not even belong to δ_j^+ and therefore we cannot just substitute i and j to get a corresponding version of (79). We use instead a simple but clever argument inspired by a similar argument in [Cri+10, Lemma 8] which takes advantage of the invariance of the measure $\tilde{\mu}$. Recall first of all from the construction of the induced map, that $g^{-1}(\delta_j^+)$ consists of exactly two connected components, one is exactly the interval $\delta_{1,j}$ and the other one is a subinterval of Δ_1^+ . So for any $j \geq 1$ we have

$$g^{-1}(\delta_j^+) = \delta_{1,j} \cup \{x : \Delta_1^+ : g(x) \in \delta_j^+\}.$$

By the invariance of the measure $\tilde{\mu}$, and since these two components are disjoint, this implies

$$\tilde{\mu}(\delta_j^+) = \tilde{\mu}(g^{-1}(\delta_j^+)) = \tilde{\mu}(\delta_{1,j}) + \tilde{\mu}(\{x : \Delta_1^+ : g(x) \in \delta_j^+\}) \tag{80}$$

The preimage of the set $\{x : \Delta_1^+ : g(x) \in \delta_j^+\}$ itself also has two disjoint connected components

$$g^{-1}\{x \in \Delta_1^+ : g(x) \in \delta_j^+\} = \delta_{2,j} \cup \{x \in \Delta_2^+ : g^2(x) \in \delta_j^+\}$$

and therefore, again by the invariance of $\tilde{\mu}$, we get

$$\tilde{\mu}(g^{-1}\{x \in \Delta_1^+ : g(x) \in \delta_j^+\}) = \tilde{\mu}(\delta_{2,j}) + \tilde{\mu}(\{x \in \Delta_2^+ : g^2(x) \in \delta_j^+\})$$

and, substituting this into (80), we get

$$\tilde{\mu}(\delta_j^+) = \tilde{\mu}(g^{-1}(\delta_j^+)) = \tilde{\mu}(\delta_{1,j}) + \tilde{\mu}(\delta_{2,j}) + \tilde{\mu}(\{x \in \Delta_2^+ : g^2(x) \in \delta_j^+\}).$$

Repeating this procedure n times gives

$$\tilde{\mu}(\delta_j^+) = \sum_{i=1}^n \tilde{\mu}(\delta_{i,j}) + \tilde{\mu}(\{x \in \Delta_n^+ : g^n(x) \in \delta_j^+\})$$

and therefore inductively, we obtain (77), thus completing the proof. \square

Proof of Lemma 4.6. From Lemma 4.3 we can give precise estimates for $\tilde{\mu}(\tau^\pm > t)$ in terms of the $y_{[\tau]}$ by making use of the fact that \tilde{h} is Lipschitz on Δ_0^\pm (see Corollary 2.4). Indeed,

$$\tilde{\mu}(\tau^- > t) = \tilde{\mu}(0, y_{[\tau]}^+) = \int_0^{y_{[\tau]}^+} \tilde{h}(x) dx = y_{[\tau]}^+ \tilde{h}(0^+) + \int_0^{y_{[\tau]}^+} \tilde{h}(x) - \tilde{h}(0^+) dx.$$

Using the fact that the density is Lipschitz we have

$$\left| \int_0^{y_{[\tau]}^+} \tilde{h}(x) - \tilde{h}(0^+) dx \right| \lesssim \int_0^{y_{[\tau]}^+} x dx \lesssim (y_{[\tau]}^+)^2$$

and so $\tilde{\mu}(\tau^- > t) = y_{[\tau]}^+ \tilde{h}(0^+) + O((y_{[\tau]}^+)^2)$. The statement for $\mu(\tau^+ > t)$ follows in the same way. \square

4.1.2. Higher order asymptotics In this subsection we prove Proposition 4.2. For clarity we prove (71) and (72) in two separate lemmas. We will make repeated use of some upper bounds for the measure $\tilde{\mu}(\delta_{i,j})$ of the partition elements which are given in Lemma 3.12

Lemma 4.7. *If at least one of ℓ_1, ℓ_2 are not zero, then for every $a, b \geq 0$*

$$\tilde{\mu}(a\tau^+ > t, b\tau^- > t) + \tilde{\mu}(a\tau^+ + b\tau^- > t, \max\{a\tau^+, b\tau^-\} \leq t) = o(t^{-\gamma-1/\beta}) \quad (81)$$

for any $\gamma \in (0, 1)$.

Proof. First note that if one of a, b is 0 then (81) is automatically satisfied.

Now suppose that $a, b > 0$. For the first term in (81), from Lemma 3.12 we get

$$\tilde{\mu}(a\tau^+ > t, b\tau^- > t) = \sum_{i=t/a}^{\infty} \sum_{j=t/b}^{\infty} \tilde{\mu}(\delta_{i,j}) \lesssim \sum_{i=t/a}^{\infty} \sum_{j=t/b}^{\infty} (ij)^{-(1+1/\beta)} \lesssim t^{-2/\beta} \quad (82)$$

which is $o(t^{-\gamma-1/\beta})$ for every $\gamma \in (0, 1/\beta)$ and therefore in particular for every $\gamma \in (0, 1)$.

For the second term in (81) we obtain from Lemma 3.12 that

$$\begin{aligned} \tilde{\mu}(a\tau^+ + b\tau^- > t, \max\{a\tau^+, b\tau^-\} \leq t) &= \sum_{i=1}^{t/a} \sum_{j=\frac{t-ai+1}{b}}^{t/b} \tilde{\mu}(\delta_{i,j}) \lesssim \sum_{i=1}^{t/a} \sum_{j=\frac{t-ai+1}{b}}^{t/b} i^{-1-1/\beta} j^{-1-1/\beta} \\ &\lesssim \sum_{i=1}^{t/a} i^{-1-1/\beta} \left[\left(\frac{t-ai+1}{b} \right)^{-1/\beta} - \left(\frac{t}{b} \right)^{-1/\beta} \right] \\ &\lesssim t^{-1/\beta} \sum_{i=1}^{t/a} i^{-1-1/\beta} \left[\left(1 - \frac{ai-1}{t} \right)^{1/\beta} - 1 \right] \end{aligned} \tag{83}$$

Making the change of variables $k = \lceil ai - 1 \rceil$ and using that the first term in the sum is 0 we obtain

$$\tilde{\mu}(a\tau^+ + b\tau^- > t, \max\{a\tau^+, b\tau^-\} \leq t) \lesssim t^{-1/\beta} \sum_{k=1}^{t-1} k^{-1-1/\beta} \left[\left(1 + \frac{k}{t} \right)^{1/\beta} - 1 \right].$$

Let us set $a_k(t) := k^{-1-1/\beta} \left[\left(1 + \frac{k}{t} \right)^{1/\beta} - 1 \right]$ and use the binomial theorem to get

$$a_k(t) = \frac{1}{k^{1+1/\beta}} \sum_{m=1}^{\infty} \binom{-1/\beta}{m} \left(-\frac{k}{t} \right)^m = \frac{1}{tk^{1/\beta}} \sum_{m=1}^{\infty} \binom{-1/\beta}{m-1} \left(\frac{1}{m} \left(\frac{1}{\beta} - 1 \right) + 1 \right) \left(-\frac{k}{t} \right)^{m-1}.$$

As $\left(\frac{1}{\beta} - 1 \right) / m$ is uniformly bounded above by some constant depending only on β we obtain

$$a_k(t) \lesssim k^{-1/\beta} t^{-1} \sum_{k=0}^{\infty} \binom{-1/\beta}{m-1} \left(-\frac{k}{t} \right)^{m-1} = k^{-1/\beta} t^{-1} \left(1 - \frac{k}{t} \right)^{-1/\beta}.$$

Using the fact that $n/(n-1) < 2$ and that $1/\beta > 1$ we may conclude

$$\begin{aligned} \tilde{\mu}(a\tau^+ + b\tau^- > t, \max\{a\tau^+, b\tau^-\} \leq t) &\lesssim t^{-1-1/\beta} \sum_{k=1}^{t-1} \left(\frac{t}{k(t-k)} \right)^{1/\beta} \lesssim t^{-1-1/\beta} \sum_{k=1}^{t-1} \frac{t}{k(t-k)} \\ &\lesssim t^{-1-1/\beta} \int_1^{t-1} \frac{t}{x(t-x)} dx \lesssim t^{-1-1/\beta} \log(t) = o(t^{-\gamma-1/\beta}) \end{aligned}$$

for any $\gamma \in (0, 1)$. □

Lemma 4.8. *If ℓ_1, ℓ_2 are not both zero, then for every $a, b \geq 0, \gamma \in (0, 1/\beta)$ we have*

$$\tilde{\mu}(a\tau^+ > t, a\tau^+ - b\tau^- \leq t) = o(t^{-\gamma-1/\beta}).$$

Proof. By Lemma 3.12 we get

$$\begin{aligned} \tilde{\mu}(a\tau^+ > t, a\tau^+ - b\tau^- \leq t) &= \sum_{i>t/a} \tilde{\mu}(\tau^+ = i, b\tau^- \geq ai - t) = \sum_{i>t/a} \sum_{j \geq (ai-t)/b} \tilde{\mu}(\delta_{i,j}) \\ &\lesssim \sum_{i>t/a} i^{-(1+1/\beta)} (ai - t)^{-1/\beta} \lesssim \sum_{i=1}^{\infty} (i+t)^{-(1+1/\beta)} i^{-1/\beta}. \end{aligned}$$

We claim that

$$\sum_{i=1}^{\infty} (t+i)^{-1-1/\beta} i^{-1/\beta} \lesssim t^{-\gamma-1/\beta}$$

for every $0 < \gamma < 1/\beta$, which is equivalent to showing that

$$\sum_{i=1}^{\infty} \frac{t^{\gamma+1/\beta}}{(t+i)^{1+1/\beta} i^{1/\beta}} \leq C$$

for some $C > 0$ independent of t . Indeed, for every i ,

$$\frac{t^{\gamma+1/\beta}}{(t+i)^{1+1/\beta} i^{1/\beta}} \leq \frac{(t+i)^{\gamma+1/\beta}}{(t+i)^{1+1/\beta} i^{1/\beta}} = \frac{1}{(t+i)^{1-\gamma} i^{1/\beta}} = \frac{1}{(t/i+1)^{1-\gamma} i^{1-\gamma+1/\beta}} \leq \frac{1}{i^{1-\gamma+1/\beta}}$$

which is summable for every $0 < \gamma < 1/\beta$. This implies the claim and thus the lemma. \square

4.2. Estimates for the induced observables. In this section we prove Corollary 2.11 and Proposition 2.10. We recall (see paragraph at the beginning of Sect. 2.3) that we will only explicitly treat the case that $\ell_1, \ell_2 > 0$ (and thus in particular $\beta > 0$). Throughout this section we will assume that φ is a Hölder observable and define $a = \varphi(1), b = \varphi(-1)$.

4.2.1. Proof of Corollary 2.11 We first consider the case where $\beta_\varphi = \beta$. Recall from (14) that $\beta_\varphi = \beta$ occurs when φ is non-zero at a fixed point corresponding to the maximum of β_1, β_2 ; and that $\beta_\varphi \neq \beta$ occurs when $\beta_1 \neq \beta_2$ and φ is zero at the fixed point corresponding to the minimum of β_1, β_2 .

Lemma 4.9. *If $\beta_\varphi = \beta$, then $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$. In particular, if $\beta_\varphi \in (0, 1/2)$ then $\tau_{a,b} \in L^2(\hat{\mu})$.*

Proof. Notice first of all that if $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$ then, in particular, $\tilde{\mu}(\pm\tau_{a,b} > t) \lesssim t^{-1/\beta_\varphi}$ and so, if $\beta_\varphi \in (0, 1/2)$ we obtain that $\tau_{a,b} \in L^2(\hat{\mu})$. Thus we just need to prove that $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$.

Suppose first that a, b do not have opposite signs, i.e. either $a, b \geq 0$ or $a, b \leq 0$, in which case the distribution of $\tau_{a,b}$ is determined by the first case of (27). Therefore, taking $\gamma = 0$ we get

$$\tilde{\mu}(\tau_{a,b} > t) = C_a t^{-1/\beta_2} + C_b t^{-1/\beta_1} + o(t^{-1/\beta}) = c t^{-1/\beta} + o(t^{-1/\beta})$$

for some constant $c > 0$, which may be equal to C_a, C_b , or $C_a + C_b$, depending on the relative values of β_1, β_2 . If $a, b \geq 0$, and exactly the same tail for $\tilde{\mu}(\tau_{a,b} < -t)$ if $a, b \leq 0$. By (30) and the fact that $\beta_\varphi = \beta$ we get that $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$, thus proving the result in this case. If $a \geq 0, b \leq 0$ the distribution of $\tau_{a,b}$ is given by (28) and (29) and so, taking $\gamma = 0$ gives

$$\tilde{\mu}(\tau_{a,b} > t) = C_a t^{-1/\beta_2} + o(t^{-1/\beta}) = c_1 t^{-1/\beta} + o(t^{-1/\beta})$$

and

$$\tilde{\mu}(\tau_{a,b} < -t) = C_{|b|} t^{-1/\beta_1} + o(t^{-1/\beta}) = c_2 t^{-1/\beta} + o(t^{-1/\beta})$$

where $c_1 = C_a$ and $c_2 = C_{|b|}$ if $\beta_2 = \beta$ and $\beta_1 = \beta$ respectively, and equal to 0 otherwise. At least one of the c_1, c_2 has to be non-zero as $\beta_\varphi = \beta$ implies that φ is non-zero at a fixed point corresponding to the largest of β_1, β_2 and so if $\beta_1 = \max\{\beta_1, \beta_2\}$ we know from (26) that $c_2 = C_{|b|} > 0$ and if $\beta_2 = \max\{\beta_1, \beta_2\}$ we know from (26) that $c_1 = C_a > 0$. Thus, since $\beta = \beta_\varphi$, we get $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$. If $a \leq 0, b \geq 0$ the same argument holds exchanging the roles of the positive and negative tails. \square

Proof of Corollary 2.11. We have already proved the result for $\beta_\varphi = \beta$ in Lemma 4.9 so we can assume that $\beta_\varphi \neq \beta$. This implies that $\beta_1 \neq \beta_2$ and that φ is only non-zero at the fixed point corresponding to the smallest of the β_1, β_2 . This situation can arise in two ways: either (i) $a \neq 0, b = 0$ and $\beta_\varphi = \beta_2 < \beta_1$; or (ii) $a = 0, b \neq 0$ and $\beta_\varphi = \beta_1 < \beta_2$. We will assume (i) and give an explicit proof of the Lemma. The proof of the Lemma in situation (ii) then follows in the same way.

Under our assumptions we know from Proposition 2.6 that the tail of $\tau_{a,b}$ is determined by (27), and we recall from (26) that $C_b = 0$. If $\beta_\varphi = \beta_2 = 0$ then we know from the second line of (27) that

$$\hat{\mu}(\pm\tau_{a,b} > t) \lesssim t^{-\gamma-1/\beta_1}.$$

Since $\beta = \beta_1 < 1$ by assumption, we may choose $\gamma \in [0, 1)$ such that $\gamma + 1/\beta_1 > 2$ yielding $\tau_{a,b} \in L^2(\hat{\mu})$. If $\beta_\varphi \in (0, 1/2)$ then the first line of (27) gives that

$$\hat{\mu}(\pm\tau_{a,b} > t) \lesssim \max\{t^{-1/\beta_\varphi}, t^{-\gamma-1/\beta}\}$$

for any $\gamma \in [0, 1)$. Choosing γ as before so that $\gamma + 1/\beta_1 > 2$ we again obtain that $\tau_{a,b} \in L^2(\hat{\mu})$. If $\beta_\varphi = \beta_2 \in [1/2, 1)$ then, choosing $\gamma \in [0, 1)$ so that $\gamma + 1/\beta_1 > 1/\beta_2$ we know from (27) that the non-zero tail of $\tau_{a,b}$ is given by

$$\hat{\mu}(\pm\tau_{a,b} > t) \lesssim C_a t^{-1/\beta_2} + o(t^{-\gamma-1/\beta_1}) = C_a t^{-1/\beta_2} + o(t^{-1/\beta_2})$$

yielding $\tau_{a,b} \in \mathcal{D}_{1/\beta_\varphi}$. \square

4.2.2. Proof of Proposition 2.10

Proof of Proposition 2.10. For a point $x \in \delta_{i,j}^-$ we know that $\tau(x) = i + j$ and that

$$g^k(x) \in \Delta_{i-k}^+ \quad \forall 1 \leq k \leq i; \quad \text{and} \quad g^{i+k}(x) \in \Delta_{j-k}^- \quad \forall 1 \leq k \leq j - 1. \quad (84)$$

Recall that by Proposition 3.4 we have $1 - x_n^+ \lesssim n^{-1/\ell_2}$, and $|\Delta_n^+| \lesssim n^{-(1+1/\ell_2)} \ll 1 - x_n^+$, which means that we can use the fact that $\tilde{\varphi}(1) = 0$ and the Hölder continuity of $\tilde{\varphi}_{(0,1]}$ to obtain

$$|\tilde{\varphi} \circ g^k(x)| \lesssim (1 - x_{i-k}^+)^{\nu_2} \leq (i - k)^{-\nu_2/\ell_2}, \quad (85)$$

for all $1 \leq k \leq i - 1$. Similarly, using the fact that $\tilde{\varphi}(-1) = 0$ and the Hölder continuity of $\tilde{\varphi}_{[-1,0)}$,

$$|\tilde{\varphi} \circ g^{i+k}(x)| \lesssim (1 + x_{j-k}^+)^{\nu_1} \leq (j - k)^{-\nu_1/\ell_1}, \quad (86)$$

for all for $1 \leq k \leq j - 1$. For $x \in \delta_{i,j}$ we know from (85) and (86) that

$$|\tilde{\Phi}(x)| \lesssim \sum_{k=1}^{i-1} (i - k)^{-\frac{\nu_2}{\ell_2}} + \sum_{k=1}^{j-1} (j - k)^{-\frac{\nu_1}{\ell_1}}. \quad (87)$$

We now consider two cases. Suppose first that $\ell_1 < \nu_1$ and $\ell_2 < \nu_2$. Then $|\tilde{\Phi}(x)|$ is uniformly bounded in x , as both (85) and (86) are summable in k and therefore the sums in (87) both converge. Therefore $\tilde{\Phi} \in L^q(\hat{\mu})$ for every $q > 0$, in particular $\tilde{\Phi} \in L^2(\hat{\mu})$ giving the first implication in (36), and by Chebyshev’s inequality, $\hat{\mu}(\pm\tilde{\Phi} > t) = O(t^{-q})$ for every $q > 0$, giving the second implication in (36). Notice that we have not required in this case the conditions (\mathcal{H}) and (\mathcal{H}') .

Now suppose that $\ell_1 \geq \nu_1$ and/or $\ell_2 \geq \nu_2$ and suppose also that

$$\nu_1 > \frac{\beta_1 - 1/q}{k_2} \quad \text{and} \quad \nu_2 > \frac{\beta_2 - 1/q}{k_1}. \tag{88}$$

Notice that for $q = 2$ this gives exactly condition (\mathcal{H}) and for $q = 1/\beta_\varphi$ this gives exactly (\mathcal{H}') . We can also suppose without loss of generality that in fact $\ell_1 > \nu_1$ and/or $\ell_2 > \nu_2$ since we can decrease slightly the Hölder exponent while still satisfying (88). In this case the sums in (87) diverge but admit the following bounds:

$$|\tilde{\Phi}(x)| \lesssim \sum_{k=1}^{i-1} (i-k)^{-\frac{\nu_2}{\ell_2}} + \sum_{k=1}^{j-1} (j-k)^{-\frac{\nu_1}{\ell_1}} \lesssim i^{1-\frac{\nu_2}{\ell_2}} + j^{1-\frac{\nu_1}{\ell_1}}.$$

We can then bound the integral by

$$\int_{\Delta_0^-} |\tilde{\Phi}(x)|^q dm \lesssim \left[\sum_{i=1}^\infty \sum_{j=1}^\infty |\delta_{i,j}| \left(i^{1-\frac{\nu_2}{\ell_2}} + j^{1-\frac{\nu_1}{\ell_1}} \right)^q \right]$$

Then, since $|\delta_{i,j}| = O(i^{-(1+1/\beta_2)} j^{-(1+1/\beta_1)})$ we get

$$\begin{aligned} \int_{\Delta_0^-} |\tilde{\Phi}(x)|^q dm &\lesssim \left[\sum_{i,j=1}^\infty \frac{i^{q-\frac{q\nu_2}{\ell_2}}}{i^{1+\frac{1}{\beta_2}} j^{1+\frac{1}{\beta_1}}} + \sum_{i,j=1}^\infty \frac{j^{q-\frac{q\nu_1}{\ell_1}}}{i^{1+\frac{1}{\beta_2}} j^{1+\frac{1}{\beta_1}}} \right] \lesssim \sum_{i=1}^\infty i^{q-\frac{q\nu_2}{\ell_2}-1-\frac{1}{\beta_2}} \\ &\quad + \sum_{j=1}^\infty j^{q-\frac{q\nu_1}{\ell_1}-1-\frac{1}{\beta_1}} \end{aligned}$$

The latter sums are bounded exactly when (88) holds. As mentioned above, for $q = 2$ this is exactly condition (\mathcal{H}) and therefore we get that $\tilde{\Phi} \in L^2(\hat{\mu})$. For $q = 1/\beta_\varphi$ this is exactly condition (\mathcal{H}') and therefore we get that $\tilde{\Phi} \in L^q(\hat{\mu})$. In fact if (88) holds for $q = 1/\beta_\varphi$ then there exists some $\varepsilon > 0$ such that (88) holds for all $q \in [1/\beta_\varphi, 1/\beta_\varphi + \varepsilon)$ and therefore $\tilde{\Phi} \in L^q(\hat{\mu})$ for every $q \in [1/\beta_\varphi, 1/\beta_\varphi + \varepsilon)$. From this and Chebyshev’s inequality we get $\hat{\mu}(\pm\tilde{\Phi} > t) = o(t^{-1/\beta_\varphi})$. □

Acknowledgements. We would like to thank Emanuel Carneiro, Sylvain Crovisier, and Santiago Martinchich for their comments and suggestions regarding early versions of this work. D.C. was partially supported by the ERC project 692925 *NUHGD* and by the Abdus Salam ICTP visitors program.

Declarations

Conflict of interest The authors declare that there are no other competing interests relevant to the research communicated in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- [AM21] Alves, J.F., Mesquita, D.: Entropy Formula for Systems with Inducing Schemes. (2021). <https://doi.org/10.48550/arXiv.2104.12629>. arXiv:2104.12629 [math]
- [AA04] Alves, J.F., Araújo, V.: Hyperbolic times: frequency versus integrability. *Ergodic Theory Dyn. Syst.* **24**(2), 329–346 (2004). <https://doi.org/10.1017/S0143385703000555>
- [BS16] Bahsoun, W., Saussol, B.: Linear response in the intermittent family: differentiation in a weighted C^0 -norm. *Discrete Contin. Dyn. Syst. Ser. A* **36**(12), 6657–6668 (2016). <https://doi.org/10.3934/dcds.2016089>
- [BM14] Bose, C., Murray, R.: First hyperbolic times for intermittent maps with unbounded derivative. *Dyn. Syst. Int. J.* **29**(3), 352–368 (2014). <https://doi.org/10.1080/14689367.2014.902038>
- [CHT21] Coates, D., Holland, M., Terhesiu, D.: Limit theorems for Wobbly interval intermittent maps. *Stud. Math.* **261**(3), 269–305 (2021). <https://doi.org/10.4064/sm200427-21-11>
- [Cri+10] Cristadoro, G., Haydn, N., Marie, P., Vaienti, S.: Statistical properties of intermittent maps with unbounded derivative. *Nonlinearity* **23**(5), 1071–1095 (2010). <https://doi.org/10.1088/0951-7715/23/5/003>
- [FL01] Fisher, A.M., Lopes, A.: Exact bounds for the polynomial decay of correlation, $1/F$ noise and the CLT for the equilibrium state of a non-Hölder potential. *Nonlinearity* **14**(5), 1071–1104 (2001). <https://doi.org/10.1088/0951-7715/14/5/310>
- [Fre+16] Freitas, A.C., Moreira, F., Jorge, M., Todd, M., Vaienti, S.: Rare events for the Manneville–Pomeau map. *Stoch. Process. Appl.* **126**(11), 3463–3479 (2016). <https://doi.org/10.1016/j.spa.2016.05.001>
- [Gou04a] Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. *Probab. Theory Relat. Fields* **128**(1), 82–122 (2004). <https://doi.org/10.1007/s00440-003-0300-4>
- [Gou04b] Gouëzel, S.: Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* **139**(1), 29–65 (2004). <https://doi.org/10.1007/BF02787541>
- [Ino92] Inoue, T.: Weakly attracting Repellers for Piecewise convex maps. *Jpn. J. Ind. Appl. Math.* **9**(3), 413–430 (1992). <https://doi.org/10.1007/BF03167275>
- [Kor16] Korepanov, A.: Linear response for intermittent maps with summable and nonsummable decay of correlations. *Nonlinearity* **29**(6), 1735–1754 (2016). <https://doi.org/10.1088/0951-7715/29/6/1735>
- [LSV99] Liverani, C., Saussol, B., Vaienti, S.: A probabilistic approach to intermittency. *Ergodic Theory Dyn. Syst.* **19**(3), 671–685 (1999). <https://doi.org/10.1017/S0143385799133856>
- [Mel09] Melbourne, I.: Large and moderate deviations for slowly mixing dynamical systems. *Proc. Am. Math. Soc.* **137**(5), 1735–1741 (2009). <https://doi.org/10.1090/S0002-9939-08-09751-7>
- [NTV18] Nicol, M., Török, A., Vaienti, S.: Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory Dyn. Syst.* **38**(3), 1127–1153 (2018). <https://doi.org/10.1017/etds.2016.69>
- [Pia80] Pianigiani, G.: First return map and invariant measures. *Israel J. Math.* **35**(1–2), 32–48 (1980). <https://doi.org/10.1007/BF02760937>
- [Pik91] Pikovsky, A.S.: Statistical properties of dynamically generated anomalous diffusion. *Phys. Rev. A* **43**(6), 3146–3148 (1991). <https://doi.org/10.1103/PhysRevA.43.3146>
- [PS09] Pollicott, M., Sharp, R.: Large deviations for intermittent maps. *Nonlinearity* **22**(9), 2079–2092 (2009). <https://doi.org/10.1088/0951-7715/22/9/001>

- [PM80] Pomeau, Y., Manneville, P.: Intermittent transition to turbulence in dissipative dynamical systems. *Commun. Math. Phys.* **74**(2), 189–197 (1980)
- [Sar01] Sarig, O.M.: Phase transitions for countable markov shifts. *Commun. Math. Phys.* **217**(3), 555–577 (2001). <https://doi.org/10.1007/s002200100367>
- [SS13] Shen, W., van Strien, S.: On stochastic stability of expanding circle maps with neutral fixed points. *Dyn. Syst. Int. J.* **28**(3), 423–452 (2013). <https://doi.org/10.1080/14689367.2013.806733>
- [Ter16] Terhesiu, D.: Mixing rates for intermittent maps of high exponent. *Probab. Theory Relat. Fields* **166**(3–4), 1025–1060 (2016). <https://doi.org/10.1007/s00440-015-0690-0>
- [You99] Young, L.-S.: Recurrence times and rates of mixing. *Israel J. Math.* **110**, 153–188 (1999). <https://doi.org/10.1007/BF02808180>
- [Zwe03] Zweimüller, R.: Stable limits for probability preserving maps with indifferent fixed points. *Stoch. Dyn.* **3**(1), 83–99 (2003). <https://doi.org/10.1142/S0219493703000620>

Communicated by C. Liverani