Conditional mixing in deterministic chaos

Caroline L. Wormell*

June 5, 2023

Abstract

While on the one hand, chaotic dynamical systems can be predicted for all time given exact knowledge of an initial state, they are also in many cases rapidly mixing, meaning that smooth probabilistic information (quantified by measures) on the system's state has negligible value for predicting the long-term future. However, an understanding of the long-term predictive value of intermediate kinds of probabilistic information is necessary in various physical problems, and largely remains lacking.

Of particular interest in data assimilation and linear response theory are the conditional measures of the SRB measure on zero sets of general smooth functions of the phase space. In this paper we give rigorous and numerical evidence that such measures generically converge back under the dynamics to the full SRB measures, exponentially quickly. We call this property conditional mixing. While conditional mixing typically cannot be proven from standard transfer operator theory, we will prove that conditional mixing holds in a class of generalised baker's maps, and demonstrate it numerically in some non-Markovian piecewise hyperbolic maps. Conditional mixing provides a natural limit on the effectiveness of long-term forecasting of chaotic systems via partial observations, and appears key to proving the existence of linear response outside the setting of smooth uniform hyperbolicity.

By definition, chaotic dynamical systems' future states are generally impossible to predict over the long term. For this reason they must be studied probabilistically, that is, in terms of measures that evolve under the dynamics. Many probabilistic questions about topologically mixing chaotic systems $f : M \circlearrowleft$ are quantitatively related to fast mixing (i.e. decay of correlations) for smooth observables with respect to some invariant measure μ [2]. Mathematically, mixing means the following decay of sufficiently regular (e.g. C^∞) observables $A, B : M \to \mathbb{R}$:

$$\int_M A \circ f^n \, B \, \mathrm{d}\mu - \int_M A \, \mathrm{d}\mu \int_M B \, \mathrm{d}\mu \xrightarrow{n \to \infty} 0.$$

Alternatively, this can be reformulated as weak convergence of smoothly re-weighted versions of μ back to μ under the actions the dynamics:

$$\int_M A \, \mathrm{d}f_*^n(B\mu) \xrightarrow{n \to \infty} \int_M A \, \mathrm{d}\mu$$

for any A, B with $\int_M B \, \mathrm{d}\mu = 1$, where f_* pushes measures forward under f.

A particularly important invariant measure is the SRB measure ρ. This is typically the most physically relevant invariant measure: among other things, assuming no periodicity in the system, measures with smooth enough Lebesgue densities commonly converge to the physical measure over time (typically exponentially quickly), that is,

$$\int_M A \, \mathrm{d}f_*^n(B \, \mathrm{Leb}) \xrightarrow{n \to \infty} \int_M A \, \mathrm{d}\rho$$

---

*Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, CNRS
email: caroline.wormell@anu.edu.au, ca.wormell@gmail.com
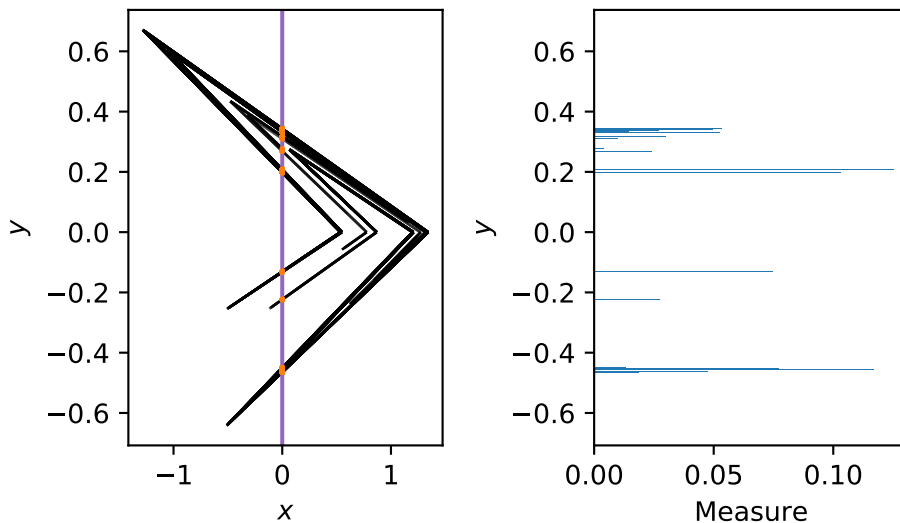ORCID: 0000-0003-2953-6493

Figure 1: Left: picture of the Lozi attractor at $a = 1.7, b = 0.5$ (black), the singular line $\{x = 0\}$ (purple line), their intersection (orange Cantor set). Right: histogram of $\rho(y \mid 0)$, obtained from 200,000 iterates of the unstable manifold dynamics $\vec{f}$, binned at width 0.0025.

for all $A, B \in C^\infty$ with $\int_M B \, d\,\mathrm{Leb} = 1$. It is a standard meta-result in smooth ergodic theory that for sufficiently hyperbolic maps with d_u positive Lyapunov exponents, the same result also holds for any measure μ with a regular conditional density along d_u-dimensional submanifolds that are tangent to expanding direction in phase space:

$$\int_M A \, df_*^n(B \, \mathrm{Leb}) \xrightarrow{n \to \infty} \int_M A \, d\rho \tag{1}$$

However, establishing fast mixing (or the related decay of iterates of a transfer operator) appears to be insufficient to answer various questions that depend on the long-term behaviour of a small problematic subset of the system's attractor, such as the response problem in non-uniformly hyperbolic systems [21]. We might therefore ask if other measures μ also satisfy (1).

Perhaps the simplest to describe example of a small subset of an attractor would be its restriction to a submanifold of phase space—which will generically be transverse to both stable and unstable manifolds. If this submanifold comes from some foliation (e.g. of level sets of an observable), we can disintegrate the physical measure and study its conditional measure μ on this submanifold. An example of such a measure is shown in Figure 1: note that unlike Gibbs invariant measures, the measure is supported on a non-invariant Cantor set without a product structure.

If (1) obtains for such a conditional measure μ, we call it conditional mixing. For dissipative systems, conditional mixing appears to lie outside the scope of traditional study by transfer operators, and this basic problem has hitherto seen very little study, notwithstanding work for other classes of μ in the specific case of linear one-dimensional maps [15, 7, 22]

Nevertheless, through a mixture of rigorous and theoretical study we will show that conditional mixing holds for a range of maps, and in particular, has some connection with the fractal-geometric theory of Fourier dimension [22, 19]. From our results it seems that conditional mixing is likely to hold for a large set of maps and submanifolds (perhaps even almost all those for which there is no "obvious" reason why it could not hold).

We also describe some consequences of conditional mixing. One physically meaningful consequence of conditional mixing is that the capability of Bayesian filters to make predictions about chaotic systems in the long term is limited if they only make partial observations (see Section 2); it also implies set-filling results for chaotic attractors (see Section 1). Another physical appli-

cation, to proving the widely-believed existence of linear response outside of smooth hyperbolic systems, will be considered in [24].

We stress that, except for conservative systems, conditional mixing *does not follow* from traditional Banach space approaches [2, 3]: these typically rely on some sort of regularity of the initial measure (e.g. smoothness of a density along a foliation). The conditional measures we will study in this paper have no regularity of this sort, and indeed are Cantor measures supported on totally disconnected sets. Banach space approaches furthermore typically proceed by showing that this kind of regularity is preserved or improved, and results usually hold while the map's structure remains the same: on the other hand, we present an example in Section 3.1 where one choice of map fails to have conditional mixing even though all the structurally identical maps around it do. We note that the Fourier dimension theory for nonlinear maps [22, 7] uses the Dolgopyat method (in dynamics more commonly used to prove rapid mixing for flows) [11], where one is interested in joint non-integrability.

The paper is structured as follows: in Section 1 we give a mathematical definition of conditional mixing and give a simple, illustrative consequence of it involving set filling; in Section 2 we present the application of conditional mixing to the area of forecasting. In Sections 3 and 4 we give evidence for conditional mixing in various systems, respectively presenting a theorem for a class of (potentially nonlinear) baker's maps and numerical evidence for some piecewise hyperbolic maps. We discuss our results in Section 5. The novel algorithms we use to obtain for our numerical results are given in Appendix A.

1 Definition and an illustrative consequence

Let $T : M \to M$ be a dynamical system with SRB measure ρ, and let $H : M \to \mathbb{R}^d$ be a C^2 function with no critical points on the level set

$$\ell_H := \{x \in M : H(x) = 0\}.$$

Suppose that for some $c : \mathbb{R}^+ \to \mathbb{R}^+$ the following limit exists in the C^0-weak topology:

$$\mu = \lim_{\delta \to 0} c(\delta)^{-1} \mathbb{1}(|H(\cdot)| \le \delta)\rho$$

and that μ is a finite measure. Any two sequences c, c' will yield μ, μ' identical up to scaling, as would changing the kernel $k(\delta^{-1}|H(x)|) = \mathbb{1}(\delta^{-1}|H(x)| < 1)$ to another bounded, compactly-supported decreasing function. If μ is a probability measure then we denote it as $\mathrm{d}\mu(x) = \mathrm{d}\rho(x \mid H(x) = 0)$.

Definition 1.1. (T, ρ) has *conditional mixing* with respect to H if for all $A, B \in C^\infty$,

$$\left| \int_M A \circ T^n B \, \mathrm{d}\mu - \int_M A \, \mathrm{d}\rho \int_M B \mathrm{d}\mu \right| \xrightarrow{n \to \infty} 0. \tag{2}$$

This is to say that repeatedly pushing forward any weighted conditional measure $B\mu$ by T, we converge back to the SRB measure ρ (up to a multiplicative constant) in the weak topology with respect to C^1.

It is natural to wish to put some quantitative bounds on mixing. By analogy with the usual sort of mixing, a natural decay rate is exponential:

Definition 1.2. (T, ρ) has *exponential conditional mixing* with respect to H if there exist $\xi < 1$, C and $r < \infty$ such that for all $A, B \in C^r$,

$$\left| \int_M A \circ T^n B \, \mathrm{d}\mu - \int_M A \, \mathrm{d}\rho \int_M B \, \mathrm{d}\mu \right| \le C\xi^n \|A\|_{C^r} \|B\|_{C^r}.$$

Note that (exponential) conditional mixing can already be expected to hold from the classical theory if T is conservative. For example, we have the following (proven in Appendix D):
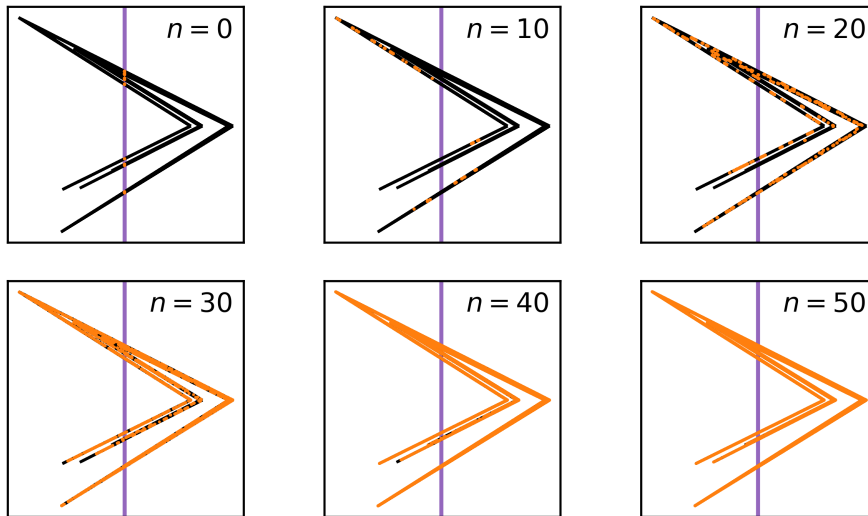
Figure 2: Exponentially fast filling of the Lozi attractor $\Lambda \subset \mathbb{R}^2$ (black) by the set intersection $\Lambda \cap \{x = 0\}$ (orange), pushed forward n times by the same Lozi map. The level set $\{x = 0\}$ is given in purple. Lozi parameters $a = 1.8$, $b = 0.35$ are used.

Proposition 1.1. *Suppose $T : M \to M$ is a C^3 conservative topologically mixing Anosov diffeomorphism with M a compact C^∞ manifold. Suppose $H \in C^2(M, \mathbb{R}^d)$ has no singular points on its zero set ℓ_H, the zero set is everywhere transverse to stable manifolds, and d is less than or equal to the number of stable directions. Then (T, Leb) has exponential conditional mixing with respect to H.*

The measure-based conditional mixing implies an interesting set-convergence property of the intersection of the level set ℓ_H with the support of ρ, which we notate as Λ and which is often an attractor of T. We find that iterates of the intersection of the line and Λ, i.e. iterates of a slice of Λ, converge back in Hausdorff distance to the full support of ρ. An example of this phenomenon is plotted in Figure 2.

Proposition 1.2. *Conditional mixing with respect to H implies that*

$$\lim_{n \to \infty} d_{\mathrm{Haus}}(T^n(\ell_H \cap \Lambda), \Lambda) = 0.$$

Under a reasonably general assumption on the regularity of the SRB measure, exponential conditional mixing gives us a quantitative version of this result as well:

Proposition 1.3. *Suppose ρ is lower-Ahlfors regular: i.e. there exist C, d such that $\rho(B(x, \delta)) > C\delta^d$ for all x. Exponential conditional mixing with respect to H implies that for some $\xi_1 < 1$ and C_1,*

$$d_{\mathrm{Haus}}(T^n(\ell_H \cap \Lambda), \Lambda) \leq C_1 \xi_1^n.$$

The proofs of these two propositions are given in Appendix D.

2 Forecasting with perfect partial observations

We now consider a fundamental practical problem to which the notion of conditional mixing is directly applicable: that of forecasting chaotic dynamics. Many forecasting methods have been developed that assimilate information obtained from observations. In general, these methods achieve this assimilation by approximating the Bayesian filter, also known as the optimal filter
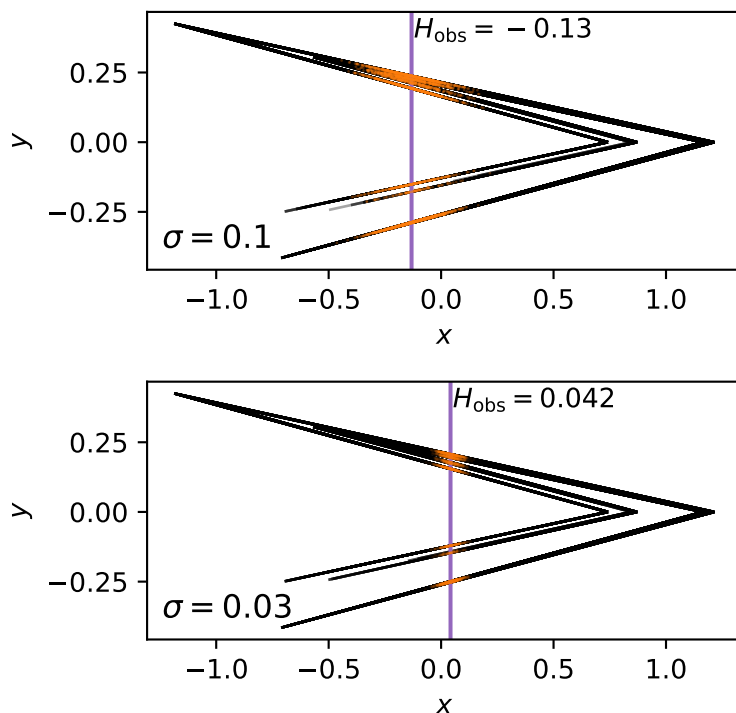
Figure 3: Posterior distribution (orange) after one observation step of $H(u,v) = u + \zeta$ where $\zeta \sim \mathcal{N}(0, \sigma^2)$, for varying values of σ. Line corresponding to $H(x,y) = H_{\mathrm{obs}}$ in purple, attractor in black. Convergence to the conditional measure as in Figure 1 can be seen as the noise $\sigma \to 0$.

[12]. The stability of the Bayesian filter with small, finite observation error has recently been studied in Anosov systems [20].

To understand this in the simplest instance, let us suppose that our system $T : \mathcal{M} \circlearrowleft$ for $\mathcal{M} \subset \mathbb{R}^d$ has exponential mixing, and at time $n = 0$, we have some prior probabilistic knowledge of the state of our system, given by some (presumably "nice") measure $\mathrm{d}\mu^-(x)$. If we start with an unobserved system at statistical equilibrium, the natural choice of prior is $\mu^- = \rho$, the SRB measure.

We can now make a noisy and perhaps partial observation of our system, given as a value $y = H(x) + \zeta \in \mathbb{R}^e$, where ζ is random with probabilities given by a smooth kernel $p(\zeta \mid x)\,\mathrm{d}\zeta$.

Assimilating the observation y, the posterior probability measure of x is given by Bayes' theorem as

$$\mathrm{d}\mu(x) = Z(y)^{-1} p(y - H(x) \mid x)\,\mathrm{d}\mu^-(x), \tag{3}$$

with normalising constant $Z(y) = \int_M p(y - H(w) \mid w)\,\mathrm{d}\mu^-(w)$. An example is given in Figure 3. Our best guess of the state of the system at future time n (i.e. of $T^n(x)$) is then given by $T^n_* \mu$; the expected value of some (nice) observable A at time n is therefore

$$\int_M A(T^n(w))\,\mathrm{d}\mu(w) = \int_M A(T^n(w)) \frac{p(y - H(w) \mid w)}{Z(y)}\,\mathrm{d}\mu^-(w). \tag{4}$$

Depending on how effective a measurement $y = H(x)$ is of x, μ is likely to be concentrated on a smaller set than μ^-, and this may improve forward estimates of the system's state over the short to medium term. However, under our assumptions for T, A, μ^-, p, exponential mixing results give that (4) will eventually converge at a fixed exponential rate to the SRB measure expectation $\int_M A\,\mathrm{d}\rho$.

But as we make our observations more and more precise, i.e. reduce the noise in H to zero, we might ask what posterior we end up with and what quality of forecast we can make with

it. This is trivial if $H(x)$ specifies x: we will know our value of x exactly, and therefore $T^n(x)$ exactly for all time. However, it is typical in high-dimensional systems for H to be only a partial observation.

In a zero-noise limit the kernel p is no longer smooth, with $p(\zeta \mid w) = \delta(\zeta)$. Our posterior measure $\mathrm{d}\mu(x)$ is then the simply the conditional probability measure of $\mathrm{d}\mu^-(x)$ given that $H(x) = y$:

$$\int_M A(f^n(w))\,\mathrm{d}\rho(w \mid H(w) = y).$$

If conditional mixing (2) holds, this must converge to the expectation of A with respect to the SRB measure ρ, i.e. in the long-term we end up back with the default no-information guess. Conditional mixing thus codifies the intuition that typical incomplete information on the system should wash out over time.

On the other hand, if conditional mixing did not hold generically, and thus partial observations could with some positive likelihood be predictively useful for all time, there might be significant practical consequences for chaotic systems. However, what our exploratory mathematical results in Sections 3 and 4 will suggest is that conditional mixing should in fact hold, excluding this possibility.

In practice, of course, physical observations will always have some random error to them. This is also true of the evolution of physical systems, which are nonetheless considered worth studying in their zero noise limit. In any case, if the observation uncertainty is small enough it will take some time to manifest in the prediction, and the behaviour until then is described by zero-noise limit we have discussed here (see Figure 4).

Nevertheless, prediction over the long-term with small, non-zero observation errors does admit an interesting question for which general results can be proven using standard dynamics techniques. If T is a C^∞ Anosov diffeomorphism and $A \in C^\infty(M)$, then over the long-term, the convergence of the forward predictions (4) are governed by how much μ picks up the transfer operator eigenfunctions of T. Mathematically, we have for all $K \in \mathbb{N}$ that as $n \to \infty$ [14]:

$$\int A \circ T^n\,\mathrm{d}\mu - \int A\,\mathrm{d}\rho \sim \sum_{k=1}^{K} \alpha_k(A)\beta_k(\tfrac{\mathrm{d}\mu}{\mathrm{d}\rho})\,n^{r_k}\lambda_k^n + o(n^{r_K}\lambda_K^n), \qquad (5)$$

where $r_k \in \mathbb{N}$ and $1 > |\lambda_1| \geq |\lambda_2| \geq \cdots \to 0$. Both the α_k and β_k are independent of the Bayesian posterior μ: the α_k, β_k are functionals, and in particularly β_k is an integration against a hyperdistribution that is C^∞ in the unstable direction. While the problem is technical and beyond the scope of this paper, this latter fact suggests that as the observation error decreases to zero (and thus the density $\frac{\mathrm{d}\mu}{\mathrm{d}\rho}$ converges to a conditional distribution supported on ℓ_H), the integral $\beta_k(\frac{\mathrm{d}\mu}{\mathrm{d}\rho})$ could potentially converge to a finite limit when ℓ_H is transverse to stable manifolds, at least for Anosov maps with a C^∞ unstable foliation. If this were the case, it would suggest that reducing the observation error does not strengthen Bayesian filter's *asymptotic* long-term predictive power—i.e. after any transients, which will at least initially approximate the zero-noise limit, and may take arbitrarily long to dissipate.

However, C^∞ uniformly hyperbolic systems are probably unique in that their transfer operators have no non-removable essential spectrum, and K can therefore be taken to infinity in (5). We can thus contrast the C^∞ Anosov case to our numerical evidence for the Lozi map in Figure 4, where the predictive power appeared to be improved by smaller error at times when the conditional mixing regime had washed out. Note that the Lozi map has a rough unstable foliation [4] and can usually be expected to lack a finite Markov partition [17], leading to essential spectrum for its transfer operator [8].

This argument would not yield a proof of conditional mixing *a priori*, as bounds for the error in the asymptotic expansion (5) will blow up as the observation noise goes to zero.
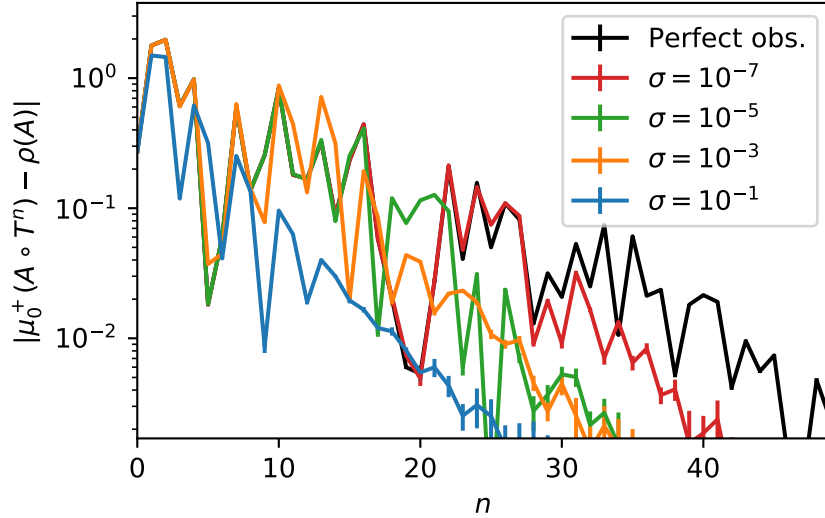
6

Figure 4: Decay of mean forecasts of $A(x, y) = 2x$ to SRB measure averages for Bayesian filters with various levels of observation noise. The observation function used was $H(x, y) = x$.

3   Rigorous results for a toy model: baker's map

As a simple model to study conditional mixing, let us consider baker's maps $b : D := [0, 1]^2 \circlearrowleft$ of the following form:

$$b(x, y) = (kx \mod 1, v_{\lceil kx \rceil}(y)) \tag{6}$$

where $k \geq 2$ is an integer, the $v_i, i = 1, \ldots, k$ are (possibly nonlinear) contractions with all $|v_i'| \leq \mu < 1$. We will also assume that the open images $v_i((0, 1))$ are disjoint, and the contractions have bounded distortion (i.e. the $\log |v_i'|$ are C^1). An example of such a map is plotted in Figure 5. For smooth enough transverse foliations, it is possible to define in a natural way conditional measures for all leaves. For simplicity, we assume we have a foliation $\Psi(t; y)$ of a subset of D into graphs $x = \psi(y) - t$, with the parameter t translating the graph in the x direction.

Proposition 3.1. *Suppose that $\Psi(t; y) = (\psi(y) - t, y)$ is a foliation of some subset $D_\Psi \subseteq D$ for $|t| \leq t_* \in [0, 1]$ and $y \in [0, 1]$. Let ρ be the (unique) SRB measure of b.*

Then for every $t \in [-t_*, t_*]$ there exists a unique probability measure ρ_t supported on $\Psi(t, [0, 1])$ such that

a. For all $t \in (-t_*, t_*)$ and all continuous functions $A : D \to \mathbb{R}$,

$$\int_D A \, d\rho_t = \lim_{\delta \to 0} \frac{1}{2\delta} \int_{\{\Psi(s, y) : |s - t| < \delta, y \in [0, 1]\}} A \, d\rho, \tag{7}$$

b. For all Borel sets $E \subseteq D_\Psi$,

$$\rho(E) = \int_{-t_*}^{t_*} \rho_t(E) \, dt.$$

Let us be as general as we can about the functions against which the conditional measures weakly converge back to the full measure. For $\alpha, \beta \in (0, 1]$ we define the following norm on continuous functions $\phi : D \to \mathcal{C}$:

$$\|\phi\|_{\alpha;\beta} = |\phi|_{\alpha,x} + |\phi|_{\beta,y} + \|\phi\|_{L^\infty}$$
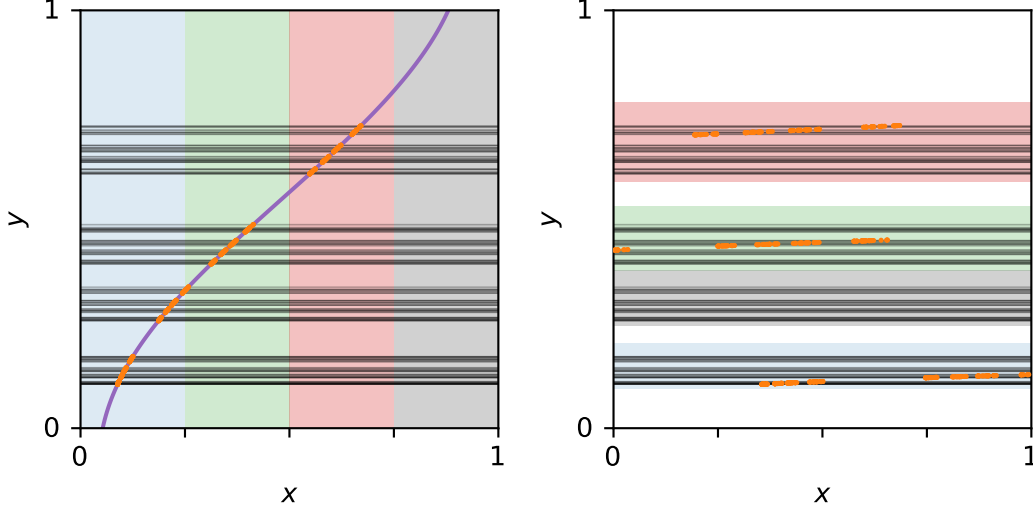
7

Figure 5: Picture of a baker's map of the form (6) showing the attractor (black), an intersecting line $x = \psi(y) - t_0$, a leaf of a transverse foliation of the form in Proposition 3.1, (purple), and the action of the map on pieces of the domain (in pale colours), and on the conditional measure of the SRB measure on the intersecting line (orange).

where the directional Hölder semi-norms are given by

$$|\phi|_{\alpha,x} = \sup_{x,x',y \in [0,1]} \frac{|\phi(x,y) - \phi(x',y)|}{|x - x'|^\alpha}$$

$$|\phi|_{\beta,y} = \sup_{x,y,y' \in [0,1]} \frac{|\phi(x,y) - \phi(x,y')|}{|y - y'|^\beta}.$$

The Banach space $C^{\alpha;\beta}$ will then consist of all continuous functions $\phi : D \to \mathcal{C}$ with $\|\phi\|_{\alpha;\beta} < \infty$: that is, functions that are α-Hölder in the x direction and β-Hölder in the y direction. In particular, the Banach space of C^1 functions is continuously embedded in $C^{\alpha;\beta}$.

The following theorem, proved in Appendix B, gives that for certain baker's maps, exponential conditional mixing holds for all conditional SRB measures on a smooth foliation transversal to unstable lines:

Theorem 3.2. *Suppose that* $\Psi(t; y) = (\psi(y) - t, y)$ *is a foliation of some subset of D for* $|t| \le t_* \in [0,1]$ *and* $y \in [0,1]$, *and* ψ *is* C^2 *with* $\psi' \ne 0$. *Suppose one of the following conditions hold:*

   I. *The contractions* v_i *are totally nonlinear* and* C^2 *and* $\cup_i v_i([0,1]) = [0,1]$.

   II. *The contractions* v_i *are totally nonlinear and analytic, as is* ψ.

   III. *The contractions are linear with* $v_i(x) := \mu x + o_i$ *for* $o_i \in [0, 1 - \mu]$, *and* $\psi'' \ne 0$.

   Let ρ *be the (unique) SRB measure of a modified baker's map b and let* $\{\rho_t\}_{t \in [-t_*, t_*]}$ *be the conditional measures of ρ on the foliation. Then* (b, ρ) *has exponential conditional mixing with respect to the level set of* $H(x, y) = \psi(y) - t - x$ *for all* $t \in (-t_*, t_*)$.

---

   *That is, no C^2 function exists conjugating all the v_i simultaneously to linear functions.

More specifically, there exists $d^* > 0$ such that for all $\gamma \in (1 - d^*, 1], \beta \in (0, 1], \alpha \in (\gamma - d_*, 1],$ there exist $C > 0, \xi \in (0, 1)$ such that for all $t \in (-t_*, t_*), A \in C^{\alpha;\beta}, B \in C^\gamma, n \in \mathbb{N},$

$$|\rho_t(A \circ b^n\, B) - \rho_t(B)\rho(A)| \leq C\|A\|_{C^{\alpha;\beta}}\|B\|_{C^\gamma}\xi^n.$$

Intuitively it seems that exponential conditional mixing holds wherever there is no obvious reason for it not to hold: that is to say, wherever the conditional measure μ is not compatible with the algebraic or symbolic structure of the unstable dynamics. Indeed, the Fourier decay results we use to prove Theorem 3.2 rely on joint non-integrability properties [22]). This incompatibility can be obtained either if ρ_t is generated by nonlinear contractions (where the expanding dynamics are linear), as in I–II, or if the leaf $x = \psi(y) - t$ transforms the self-affine stable measure in a nonlinear way (as in III). We give an example where the conditional measure picks up the algebraic structure of both the stable and unstable dynamics in Section 3.1.

From this and Proposition 1.3, exponential convergence in Hausdorff distance of the slice sets follow.

Corollary 3.3. *Under the conditions of Theorem 3.2, there exist C_1 and ξ_1 such that for all $t \in (-t_*, t_*),$*

$$d_{\mathrm{Haus}}(b^n(\Psi(t; \mathbb{R}) \cap \Lambda_b), \Lambda_b) \leq C_1\xi_1^n,$$

where Λ_b is the attractor of b.

Note that the classic piecewise affine baker's maps fall under condition III of the theorem, although for conservative maps the application of Fourier dimension theory is unnecessary due to the Lebesgue-absolute continuity of ρ.

It can also be seen from the proof that the constant d_* is half the Fourier dimension of ρ_t projected onto the x coordinate. (A choice of test functions with more specific Fourier decay properties might yield d_* to be exactly Fourier dimension). This Fourier dimension is bounded from above by the Hausdorff dimension of ρ_t, which, given the product structure of the measure, is easily seen also to be the stable dimension of the systems [23]. Thus, a larger stable dimension suggests conditional mixing holds with respect to increasingly less regular observables, with consequences for linear response theory [21, 24].

In fact, for case III, d^* as well as the asymptotic rate of conditional mixing ξ are independent of the conditioning foliation [19, Theorem 3.1].

It is worth mentioning that it is also possible to decompose sufficiently regular (e.g. real analytic) curves that are tangent to stable or unstable manifolds away from the support of the attractor into a finite set of curves that uniformly avoid tangencies. In fact, the stipulations that $\psi' \neq 0$ or indeed that the foliation is a graph in y are mere artefacts of the proof and can almost certainly be relaxed.

To prove Theorem 3.2, we use essentially two facts. The first fact, used in Lemma B.6, is that the x component of b is a tupling map, whose action on Fourier coefficients of measures is well known. The second fact (Proposition B.3) is that the SRB measure is a product of uniform measure in the x direction and a Gibbs measure (in fact, a measure of maximal entropy) of an iterated function system in the y direction. This allows us to bring some recent results on Fourier dimension of Gibbs measures [19, 22, 18].

The Fourier dimension of Gibbs measures is an area in progress whose results have not yet been consolidated, hence the somewhat particular set of alternatives. We remark that if $B(x, y)$ depends only on x then in case III the $\psi' \neq 0$ restriction can be dropped: that is, quadratic tangencies with stable manifolds (lines of constant y) are allowed here.

3.1 A case where conditional mixing fails

However, it is clear that some stipulations on the contractions and foliation of a similar flavour to I–III must remain. In particular, either the contractions or the foliation should avoid preserving too good a linear structure, or conditional mixing will fail. An example where this occurs is
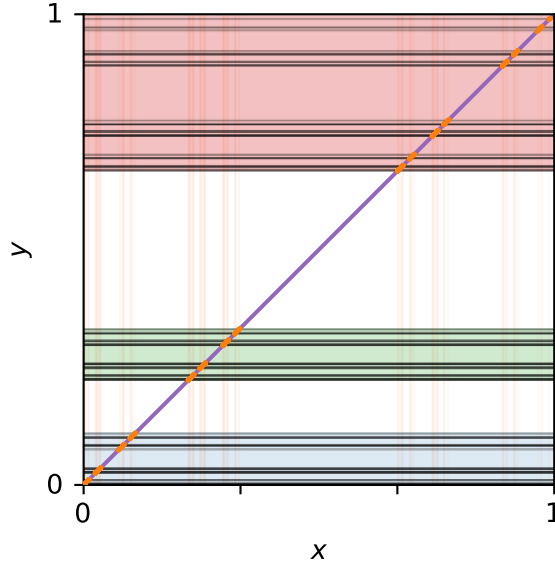
Figure 6: Picture showing the attractor $[0,1] \times C_{1/3}$ of the piecewise linear baker's map b_{bad} (black) with the images of its pieces (pale colours), and the linear level set $x - y = 0$ (purple). The conditional measure of the SRB measure on the intersecting line is shown in orange dots, and, in pale vertical orange lines, the b_{bad}-invariant subset $C_{1/3} \times [0,1]$ in which the support of the conditional measure is contained.

given for the map b_{bad}, given by $k = 3$ and

$$v_1(y) = y/9$$
$$v_2(y) = (y+2)/9$$
$$v_3(y) = (2y+1)/3.$$

The attractor and dynamical picture for this map are plotted in Figure 6.

Here, we have the following:

Proposition 3.4. $(b_{\text{bad}}, \rho_{\text{bad}})$ *has no conditional mixing with respect to* $H(x,y) = y - x$.

This is because the x and y directions, which are both linear, both preserve the middle-thirds Cantor set, which is copied across the directions using the linear foliation:

Proof. The contractions v_1, v_2, v_3 together generate the middle-thirds Cantor set $C_{1/3}$: if $f_1(x) = x/3$, $f_2(x) = (x+2)/3$ are the classical generators, then $v_1 = f_1^2$, $v_2 = f_1 \circ f_2$ and $v_3 = f_2$. The attractor $\Lambda_{b_{\text{bad}}}$ of b_{bad} is therefore $[0,1] \times C_{1/3}$ (e.g. using Proposition B.3). Our level set is $L = \{x - y = 0\}$ so the support of our conditional measure is $L \cap \Lambda_{b_{\text{bad}}} = \{(x,x) : x \in C_{1/3}\}$.

The expanding dynamics of b_{bad} is just the tripling map $\kappa(x) = 3x \mod 1$, which has $C_{1/3}$ as an invariant set. This means $C_{1/3} \times [0,1]$ is a proper invariant closed subset for the full baker's map. Our conditional measure's support $L \cap \Lambda_{b_{\text{bad}}}$ is contained in this set and so for all time, $(b_{\text{bad}}^n(L \cap \Lambda_{b_{\text{bad}}}) \subseteq C_{1/3} \times [0,1]$. We note that $p = (\frac{1}{3}, \frac{1}{2})$ lies in $\Lambda_{b_{\text{bad}}}$, but $d(p, (b_{\text{bad}}^n(L \cap \Lambda_{b_{\text{bad}}})) \geq d(p, C_{1/3} \times [0,1]) = 1/6$, so $(b_{\text{bad}}^n(L \cap \Lambda_{b_{\text{bad}}})$ cannot converge in Hausdorff distance to $\Lambda_{b_{\text{bad}}}$. By the contrapositive of Proposition 1.2, no conditional mixing holds. $\quad\square$

Nevertheless, if any nonlinear perturbation of either the level set or any of the v_j's is made, Theorem 3.2 will apply and conditional mixing will manifest. Conditional mixing relies on the lack of joint algebraic structure, rather than simply on regularising properties.

4 Numerical example: Lozi map

While baker's maps' special structure (e.g. as skew products) allow us to find rigorous results for them by borrowing existing theory, it is not immediately clear how to mathematically generalise our results. To study maps with less structure we now therefore turn to rigorously justified numerics, and consider the commonly-studied and numerically amenable class of Lozi maps.

These are piecewise hyperbolic affine maps $f : \mathbb{R}^2 \to \mathbb{R}^2$ with

$$f(x, y) = (1 + y - a|x|, bx),\ b \neq 0. \tag{8}$$

For $a \in (1, 2)$ and $b \in (0, \min\{a - 1, 4 - 2a\})$ the Lozi map f has chaotic dynamics on a compact region in phase space: when additionally $b \in (0, \sqrt{2}(a - \sqrt{2}))$ this has a single mixing SRB measure [16, Theorem 5], all unstable manifolds have positive length when $b \in (0, a - \sqrt{2})$ [25, Theorem]. A Lozi attractor is shown in Figure 1.

Lozi maps are continuous, with a jump in the Jacobian across the singularity set $\mathcal{S} = \{x = 0\}$.

In a similar fashion to those we defined for the baker's map (Proposition 3.1), conditional measures of the SRB measure $\mathrm{d}\rho(x, y)$ on sets $\ell_{x_0} := \{x = x_0\}$ are well-defined for all $x_0 \in \mathbb{R}$ intersecting the support of the Lozi attractor [24, Theorem 2.1]. Let us notate these conditional measures as $\mathrm{d}\rho(y \mid x_0)$. These conditional measures can be expected to have Hausdorff dimension strictly between 0 and 1: in particular, they lack any manifold structure. A histogram of $\mathrm{d}\rho(y \mid 0) = \mathrm{d}\rho(y \mid (x, y) \in \mathcal{S})$ is plotted in Figure 1: linear response for the Lozi map is determined from the mixing properties of the conditional measure on $\mathcal{S}$ [24], so we will be most interested in this particular conditional measure.

We specifically conjecture that measures $\rho(\cdot \mid x_0)$, when pushed forward under the Lozi map f, converge back to the full SRB measure ρ, and that this convergence happens at an exponential rate:

Conjecture 4.1. *For generic Lozi parameters (a, b) and Lebesgue-almost all $x_0 \in \mathbb{R}$, the Lozi map has conditional mixing with respect to level curves $x = x_0$ (i.e. the measures $\rho(\cdot \mid x_0)$).*

We have strong and direct numerical evidence in favour of this conjecture: Figure 7 shows exponential decay of the correlation

$$\int A \circ T^n \, \mathrm{d}\rho(\cdot \mid 0) - \int A \, \mathrm{d}\rho \int \mathrm{d}\rho(\cdot \mid 0) \tag{9}$$

by four orders of magnitude, with reliable error quantification. In fact, it seems that A need only be piecewise C^1. The consequence of Conjecture 4.1 holding on the singular line $x = 0$ (up to a technical generalisation to second-order mixing), we will show in [24, Theorem 2.3], is that the Lozi map a formal linear response to bounded dynamical perturbations [24].

It should be noted that obtaining valid samples of the quantities in (9) is tricky. To begin with, we are sampling from the conditional measure $\rho(\cdot \mid x_0)$, which is a codimension-one object on the attractor. We then must iterate forward under the Lozi map, which is chaotic and thus unstable. To validly perform this sampling rigorously, and quantify the associated numerical error, we have developed novel algorithms, presented in Appendix A. These algorithms could, we imagine, with care be extended to general hyperbolic dynamics.

5 Conclusion

While in a numerical example, one can only consider a particular case, our results on baker's maps suggest that conditional mixing is a very robust property. Indeed, Theorem 3.2 shows that for an open dense set of such baker's maps (perhaps all baker's maps), conditional mixing holds on an open dense set of analytic curves. While these skew product maps are rather special, it seems from the Fourier dimension results that conditional mixing occurs when structure is *broken*, and so the situation might yet be better for conditional mixing in more general maps. To this end, we make the following conjecture:
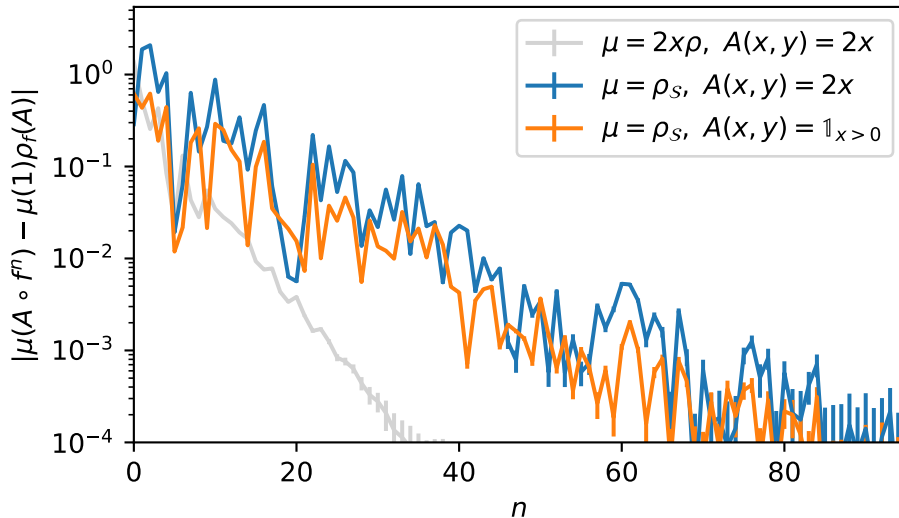
Figure 7: In orange and blue, for the parameters of the Lozi map $a = 1.8$ and $b = 0.35$, $|\rho_{\mathcal{S}}(A \circ f^n) - \rho_{\mathcal{S}}(1)\rho(A)|$ plotted for varying n and two different observables A, where $\rho_{\mathcal{S}} = \rho(\cdot \mid 0)$. The measure was sampled using 100 time series of 10^7 iterates of the segment dynamics $\vec{f}$, implemented in high-precision (196 bit) validated interval arithmetic. Details of the algorithm are given in Appendix A. Two forms of error are plotted: with error bars, 99% confidence intervals for the sampling error, and with block error (not visible), the error arising from the interval arithmetic. In grey, decay of autocorrelations for a smooth observable $A(x) = 2x$ against the SRB measure, using 100 time series of 10^8 iterates of the full dynamics f.

Conjecture 5.1. *For all analytic Anosov diffeomorphisms on compact surfaces, conditional mixing holds on an open dense set of functions H with no critical points on their zero level set.*

On the other hand, it is interesting that the exponential rates of decay for the conditional measures are substantially slower than for smooth observables against the full SRB measure: for example, in Figure 7 the rate of exponential convergence for a smooth observable $A(x, y) = x$ is much slower when the initial measure μ is a conditional measure rather than for the full SRB measure $\mu \sim \rho$. Furthermore, it appears that at least for the baker's map in some cases that this decay rate is independent of the conditioning submanifold (see discussion in Section B). In a related fashion, mixing rates against the SRB measure, which depend on the essential spectrum of the transfer operator in relevant function spaces, tend to be slower for smaller Hölder regularities of observables [1]: thinking of the conditional measure $\rho(x \mid H(x) = 0)$ as equivalent to the SRB measure multiplied by a distribution $\delta(H(x))$ provides some connection between these two phenomena. Therefore, while the main connection we have seen appears to be to Fourier dimension, perhaps an appropriate functional analytic approach could yield fruit in studying this basic property of a chaotic system.

One might also ask what happens when the codimension of the conditioning submanifold is increased (in our study it has always been one). This is natural for the Bayesian filter problem since one typically makes repeated observations when observing a system [12, 20]: we expect to end up with, say, a vector of one-dimensional observations $y = (H(x), H(f(x)), H(f^2(x)), \ldots, H(f^m(x)))$. If this dimension m is no greater than the unstable dimension (i.e. the number of positive Lyapunov exponents), then the construction of the conditional measure as in Proposition 3.1 and [24, Theorem 2.1] will go through, and we might feel empowered to say we expect conditional mixing to hold generically. On the other hand, if m is more than twice the box-counting dimension of the attractor, then the probabilistic Takens embedding theorem would tells us that y specifies x exactly ρ-almost surely [5]. In the intermediate case where M lies between the unstable dimension and the attractor dimension, it could be that some kind of generic intersection

property à la blenders [6] holds to give a conditional set of positive fractal dimension, which could also allow for some kind of conditional mixing. Numerical study in higher-dimensional systems may shed light on the situation.

A Simulation of Lozi map dynamics

Rather than attempt to compute deterministic estimates of these systems we will proceed by Birkhoff–Monte Carlo sampling of the quantities we are interested in. Because we need to sample from measures $\rho(\cdot \mid x_0)$ conditioned on a codimension-1 manifold, it will be necessary to simulate not point dynamics but dynamics on sets of higher dimensions, the natural choice being local unstable manifolds. Helpfully, because the Lozi map is piecewise affine, local unstable manifolds are straight line segments. A "segment dynamics" is proposed in Section A.1 and a numerical implementation given in Section A.2.

However, we would like to be sure we are not merely percieving artifacts of sampling error or numerical imprecision: to achieve this, we will also need to quantify the statistical and deterministic errors associated with our numerical simulations. Section A.3 gives, surprisingly, a stable algorithm to simulate the chaotic Lozi dynamics that is compatible with validated interval arithmetic, and Section A.4 explains the quantification of random sampling errors.

A.1 Segment dynamics

Let Λ be the attractor of the Lozi map f, and for points p whose orbits do not intersect the singular line $\mathcal{S} := \{x = 0\}$, define the local unstable manifold of a point $p \in \Lambda$ to be

$$\mathcal{W}^u_{\mathrm{loc}}(p) := \left\{ q \in \Lambda : \lim_{n \to \infty} |f^{-n}(q) - f^{-n}(p)| = 0, \forall n \in \mathbb{N}^+ \ \sigma_{f^{-n}(p)} = \sigma_{f^{-n}(q)} \right\}.$$

where $\sigma_{(x,y)} := \mathrm{sign}\, x$. These are segments of the full unstable manifolds which have always remained on the same side of the singular line $\mathcal{S}$.

Let $\vec{\mathcal{G}}$ be the set of directed open line segments in $\mathbb{R}^2$, i.e., open intervals where start and end points are distinguished. Then we can define a set of directed local unstable manifolds

$$\vec{\mathcal{L}} = \left\{ \vec{I} \in \vec{\mathcal{G}} : \exists p \in \Lambda \ I = \mathcal{W}^u_{\mathrm{loc}}(p) \right\},$$

which captures ρ-almost all local unstable manifolds, since Lozi maps are piecewise affine, and almost all unstable manifolds are of positive length.

Let us also define the following product space

$$\vec{\Lambda} = \vec{\mathcal{L}} \times (0, 1),$$

which we are going to use to parametrise each $\vec{I} \in \vec{\mathcal{L}}$. $\vec{\Lambda}$ is almost everywhere a two-to-one cover of Λ by the map $\pi(\vec{I}_{p,q}, t) = (1-t)p + tq$, where we denote the directed segment from point p to point q by $\vec{I}_{p,q}$.

As a result, up to a set of ρ-measure zero, we can lift the f-dynamics to $\vec{\Lambda}$, by a map of the form

$$\vec{f}(\vec{I}, t) = (f(\vec{I} \cap \mathcal{M}_{\pi(\vec{I},t)}), \lambda_{\vec{I}}(t)), \tag{10}$$

where $\lambda_{\vec{I}} : [0, 1] \circlearrowleft$ is a full-branch expanding interval map. Let us define this a little more explicitly.

When $\vec{I}_{p,q} \cap \mathcal{S}$ is non-empty, then we know it has exactly one element which we denote by $s \in \mathcal{S}$ with $s = \pi(\vec{I}_{p,q}, t_*)$ for some $t_* \in (0, 1)$. The segment dynamics $\vec{f}$ can then be written explicitly as

$$\vec{f}(\vec{I}_{p,q}, t) = \begin{cases} (\vec{I}_{f(p),f(q)}, t), & \vec{I}_{p,q} \cap \mathcal{S} = \emptyset \\ (\vec{I}_{f(p),f(s)}, t/t_*), & \vec{I}_{p,q} \cap \mathcal{S} \neq \emptyset \text{ and } t < t_* \\ (\vec{I}_{f(s),f(q)}, (t - t_*)/(1 - t_*)), & \vec{I}_{p,q} \cap \mathcal{S} \neq \emptyset \text{ and } t > t_*. \end{cases} \tag{11}$$

It turns out that the SRB measure ρ also can be lifted to an invariant measure of $\vec{f}$ by

$$\int_{\vec{\Lambda}} \Psi \, \mathrm{d}\vec{\rho} = \int_{\Lambda} \frac{\sum_{(\vec{I},t) \in \pi^{-1}(x)} \Psi(\vec{I}, t)}{2} \mathrm{d}\rho(x),$$

because π^{-1} is two-to-one ρ-almost everywhere we have that $(\vec{\Lambda}, \vec{f}, \vec{\rho})$ has most two ergodic components (which are identical up to reversing the direction of the segments), and we will be able to sample $\vec{\rho}$ by iterating $\vec{f}$.

In defining a stable numerical method it will be useful to us that almost every point's local unstable manifold has endpoints originating from the singular line:

Proposition A.1. *The set*

$$\left\{ (\vec{I}_{p,q}, t) \in \vec{\Lambda} : p, q \in \cup_{n=1}^{\infty} f^n(\mathcal{S}) \right\}$$

has full $\vec{\rho}$ measure.

This proposition is proved in Appendix E.

A.2 Simulation of segment dynamics

Given our interval dynamics, we can sample the conditional measures using ρ via the function $\kappa_{x_0}(\vec{I}) := \vec{I} \cap \{x = x_0\}$ and $\ell(\vec{I}) = |\vec{I}|$ the length of the segment. Because the SRB measure is uniformly distributed along unstable manifolds, we have that

$$\int_{\vec{\Lambda}} A(x_0, y) \, \mathrm{d}\rho(y \mid x_0) = \frac{\int_{\vec{\Lambda}} A \circ \kappa_{x_0} / \ell \, \mathrm{d}\vec{\rho}}{\int_{\vec{\Lambda}} 1/\ell \, \mathrm{d}\vec{\rho}} \tag{12}$$

assuming that no contribution to the sum is made when $\kappa_{x_0}(\vec{I}) = \emptyset$ (i.e, $\vec{I}$ does not intersect with the line we want to sample a conditional measure from). Then, for $\vec{\rho}$-almost-all starting values $(\vec{I}_0, t_0)$ we can estimate these expectations via a Birkhoff sum

$$\int_{\vec{\Lambda}} \Psi \, \mathrm{d}\vec{\rho} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \Psi(\vec{f}^n(\vec{I}_0, t_0))), \tag{13}$$

with the segment dynamics that can be simulated using (11). The convergence (13) for the Ψ we are interested in uses the Birkhoff ergodic theorem and the fact that κ_{x_0}, ℓ do not depend on the direction of the interval (so the ergodic component of $\vec{\Lambda}$ we sample from is immaterial).

Because from (11) the t dynamics are generated by full-branch interval maps that preserve Lebesgue measure, the branch dynamics are Markovian, with transition probabilities that are explicitly given. Because of this, the random dynamics

$$\vec{f}(\vec{I}_{p_{n+1}, q_{n+1}}, T_{n+1}) = \begin{cases} (\vec{I}_{f(p_n), f(q_n)}, T_n), & \vec{I}_{p_n, q_n} \cap \mathcal{S} = \emptyset \\ (\vec{I}_{f(p_n), f(s_n)}, T_n) \text{ with probability } t_*, & \vec{I}_{p_n, q_n} \cap \mathcal{S} \neq \emptyset \\ (\vec{I}_{f(s_n), f(q_n)}, T_n) \text{ with probability } 1 - t_*, & \vec{I}_{p_n, q_n} \cap \mathcal{S} \neq \emptyset. \end{cases} \tag{14}$$

generate the $\vec{f}$ dynamics at equilibrium, where the T_n are uniformly distributed, dependent hidden variables which are given by

$$T_{n+1} = \lambda_{\vec{I}_{p_n, q_n}}(T_n).$$

To sample that $\vec{I}$ we do not actually need to know the T_n, but they can be reconstructed from a time series of the segments by sampling the final value $T_{n_{\text{final}}} \sim \text{Uniform}(0, 1)$ and iterating backwards according to (11): inverting the $\lambda_{\vec{I}}$ yields a contraction. This means we do not have to directly simulate an expanding map (which would have been problematic for rigorously validated simulation of the dynamics). In practice this is an very effective way to simulate the segment dynamics.

14

A.3 Validated numerical implementation of segment dynamics

However, computers can only encode real numbers to finite precision. Thus, at every computational step the results must be rounded to a given tolerance, introducing small errors, which may invalidate fine numerical results such as we wish to obtain. Validated interval arithmetic provides a vehicle to quantify the errors, but to use it we must first produce a deterministically stable algorithm to simulate a generic long chaotic time series. In this subappendix will we present such an algorithm, introducing first the notion of interval arithmetic.

Let the set of closed intervals in $\mathbb{R}$ be $\mathcal{I}$. The idea of validated interval arithmetic is to represent real numbers α by an interval $\mathfrak{a} \in \mathcal{I}$ such that we know $\alpha \in \mathfrak{a} \subset \mathbb{R}$. Such an interval $\mathfrak{a}$ is given by its upper and lower bounds, and we can restrict the set of allowed intervals so these upper and lower bounds are representable in the finite precision computer encoding. A function $g : \mathbb{R}^d \to \mathbb{R}^e$ can be implemented in validated arithmetic through a function $\mathfrak{g} : \mathcal{I}^d \to \mathcal{I}^e$ such that $\mathfrak{g}(\mathfrak{a})$ will always contain $g(\mathfrak{a})$. In such a way one can be absolutely certain that α is contained in a set $\mathfrak{g}(\mathfrak{a})$ and so on.

The unstable manifolds $\vec{I}_{p,q}$ are defined by their endpoints, which update under the chaotic dynamics f. By definition, f is exponentially stretching, making it very difficult in general to obtain a rigorously validated time series. However, the fact that we are constantly resetting the segment endpoints to the critical line (11) makes efficiently obtaining such a time series quite possible.

We will store our segments $\vec{I}_{p,q}$ as $Q\vec{I}_{p',q'}$, where Q is an orthogonal transformation of $\mathbb{R}^2$ (i.e. a rotation matrix), and $p', q' \in \mathbb{R}^2$ have identical second coordinate. (Of course, an segment can be stored as a 2×2 matrix of its endpoints' coordinates.) Thus, Q rotates the phase space so that the unstable direction on the segment is along the first coordinate. If our segment at the next step is $\vec{I}_{p_1,q_1} = Q_1\vec{I}_{p'_1,q'_1}$, we do not compute the quantities on the right-hand side from f explicitly, but rather, since we have on $f^{-1}(\vec{I}_{p_1,q_1})$ that the Lozi map is affine, having for some J that

$$f(x) = Jx + e_1,$$

we make the QR decomposition

$$Q_1 R_1 = JQ, \tag{15}$$

where Q_1 is a rotation matrix and R_1 is upper triangular, and set

$$\vec{I}_{p'_1,q'_1} = R_1\vec{I}_{p',q'} + Q_1^\top e_1.$$

This means the dynamics in the endpoints' shared second coordinate is contracting (and thus numerically stable), as in fact are the dynamics of Q, and that except for the (discrete) choice of branch $\mathcal{M}_\pm$, these are both independent of the points' first coordinates. This means the second coordinate as well as the rotation matrix Q can be stably approximated in interval arithmetic*.

The first coordinates of p' and q', on the other hand, have expanding dynamics. However, when the segment $\vec{I}_{p,q}$ is cut by the singular line at a point r, we can define $r' = Q'r$ without reference to these first coordinates. In particular, the singular line in the transformed coordinates is $Q^\top \mathcal{S}$ which solves some equation $x' = \beta y'$, with β bounded because unstable manifolds are uniformly transversal to the singular line [25]. Then, if the shared second coordinate of p', q' is y', then we can write

$$s' = (\beta y', y').$$

Notably, this point s' is generated using only quantities whose numerical error remains stable. It thus replaces either one of p, q which contain dynamics where the error grows. By Proposition A.1, almost every p, q will eventually be replaced by such an s, resetting the size of its error and therefore ensuring it does not grow too big.

Implementation of the algorithm described above in validated interval arithmetic is straightforward, because it is stable, but here we must be careful: when our local unstable manifold

---

*It is however necessary to explicitly code the QR decomposition (15) in a way optimised for interval arithmetic, as many standard `qr` routines give sub-par interval bounds that will lead to numerical blow-up of the algorithm.

is split and the choice of child manifold is to be made, the t_* used to determine the choice is interval-valued (i.e. in this set $\mathcal{I}$ and likely of positive width). The natural way to choose the branch to continue with is made by sampling a uniform random variable U and comparing it with t_*: the choice of segment is then clear except where there is an overlap between U and t_*. A simple way to deal with this problem is to choose the floating-point precision small enough to make an overlap unlikely enough to invite references to the age of the universe (twice the bits of the standard double-precision is enough). More comprehensive handling of this eschatological edge-case may be done in various ways, including using importance sampling on multiple time series[*]

We will therefore be able to find an interval hypercube containing an exact time series $\{\vec{f}^n(\vec{I}, t_0)\}_{n=0,\dots,N}$ from (14), where t_0 is implicitly defined by the random selection when the segment is cut.

A.4 Quantification of statistical error

As the $\vec{f}$ dynamics we sample is just a two-to-one lift from the f dynamics, which have a spectral gap [4], we expect that for large N the error between Birkhoff means and true expectations (13) obey a central limit theorem [9]. If we have several long time series

$$\Psi^{(r)} = \frac{1}{N} \sum_{n=0}^{N-1} \Psi(\vec{f}^n(\vec{I}^{(r)})))$$

for R independent samples from $I^{(r)} \sim \vec{\rho}$, then for sufficiently large N, the sample mean $\bar{\Psi}$ of the $\Psi^{(r)}$ will have expectation $\vec{\rho}(\Psi)$ (from the initialisation of the time series at equilibrium), and will differ from this by a factor of $\mathcal{O}(1/\sqrt{RN})$. Helpfully, we can quantify this deviation *a posteriori*: using the Gaussian behaviour of the $\Psi^{(r)}$ we have that if s_Ψ^2 is the sample variance, then for large N we have

$$\frac{\bar{\Psi} - \vec{\rho}(\Psi)}{\sqrt{R} s_\Psi} \sim t_{R-1},$$

where t_{R-1} is a Student t distribution with $R-1$ degrees of freedom. This allows us to put confidence intervals on our estimates, as in Figure 7. Such a principle has been used to test for linear response in previous work [13].

To obtain an accurate sample from $\vec{\rho}$ in initialising our time series we begin by initialising $(\vec{I}_{p_0, f(p_0)}, t^{(r)}) \in \vec{\Lambda}$, where $p_0 = (\frac{2}{2+a-a^2+4b}, 0)$ and $t^{(r)} \sim \text{Uniform}(0,1)$ (in fact, implicitly using the above random choice methods). This initial measure lies in the Banach space that converges exponentially quickly to the physical measure [10] and so by making the spin-up time n_{init} sufficiently long, we can ensure that our sampling initialisations $\vec{I}^{(r)} = \vec{f}^{n_{init}}(\vec{I}_{p_0, p_1}, t^{(r)})$ come from a distribution exponentially close to $\vec{\rho}$.

B Proof of baker's map result

In this appendix we will prove Theorem 3.2 on exponential conditional mixing for baker's maps of the form (6).

For concision when quantitatively referring to Fourier dimension, let us say that a measure-function pair (ν, ψ) has (η, C) Fourier decay if for all $j \in \mathbb{Z} \backslash \{0\}$,

$$\left| \int_0^1 e^{2\pi i j \psi(x)} \, d\nu(x) \right| \leq C|j|^{-\eta}, \tag{16}$$

and $\int_0^1 |d\nu| \leq C$. This implies that the Fourier dimension of $\psi_* \nu$ is at least equal to 2η.

Fourier decay is invariant under translations of ψ:

---

[*]One compares U with some real number $t_{**} \in t_*$ and reweights the time series by t_*/t_{**} or $(1-t_*)/(1-t_{**})$ as appropriate: note that the weights are also interval-valued.

Proposition B.1. *Suppose that (ν, ψ) has (C, η) Fourier decay. Then for all $t \in \mathbb{R}$, so does $(\nu, \psi + t)$.*

Proof. We have

$$\left| \int_0^1 e^{2\pi i j (\psi(x) + t)} \, \mathrm{d}\nu(x) \right| = \left| \int_0^1 e^{2\pi i j \psi(x)} \, \mathrm{d}\nu(x) \right| \leq C |j|^{-\eta},$$

and the integral of $|\nu|$ remains no greater than C, as required. $\qquad\square$

To prove Theorem 3.2 we will employ a separate Fourier dimension theorem [19, 22] for each of the conditions in the theorem's statement. The common component is the following lemma (into which any new Fourier dimension results may also be substituted):

Lemma B.2. *Suppose one has a modified baker's map b with contracting maps v_i. Let ν_0 be the probability measure such that $v_i^* \nu_0 = k^{-1} \nu_0$ for all $i = 1, \ldots, k$, and let $\psi \in C^2$ be such that $\psi' \neq 0$ and (ν_0, ψ) has (C, η) Fourier decay.*
Let $\gamma \in (1 - \eta, 1)$, $\beta \in (0, 1]$ and $\alpha \in (2 - \eta - \gamma, 1)$.
Then there exists $\xi \in (0, 1)$ and C' depending only on $C, \eta, \alpha, \beta\gamma, \psi'$ such that for all $A \in C^{\alpha;\beta}$ and $B \in C^\gamma$

$$|\rho_0(A \circ b^n \, B) - \rho_0(B)\rho(A)| \leq C' \|A\|_{C^{\alpha;\beta}} \|B\|_{C^\gamma} \xi^n$$

where ρ_0 is defined as in Theorem 3.2.

Proof of Theorem 3.2. The measures ν_0 from Lemma B.2 are the measures of maximal entropy of these expanding iterated function schemes: in particular, they are Gibbs (with constant weights) and atomless. If we have that (ν_0, ψ) has Fourier decay then so do (ν_0, ψ) uniformly from Proposition B.1, and Lemma B.2 then secures us the theorem. Since ν_0 is a probability measure, it is only necessary to check that (16) holds, which we do procedurally from existing results.

That (16) holds for option III is a simple application of [19, Theorem 3.1] (and in fact here, η is independent of ψ).

To see it for options I and II requires a little more cunning. We have that ψ is a diffeomorphism onto its image: let $\omega(x) = \omega_0 + \omega_1 x$ map $\psi([0,1])$ linearly onto $[0,1]$, so that $\tilde{\psi} = \omega \circ \psi : [0,1] \circlearrowleft$ is a diffeomorphism. If $\{v_{\mathbf{i}}\}_{\mathbf{i} \in \{1,\ldots,k\}^n}$ are n-fold compositions of the contractions v_i then for some large enough n, the n-fold compositions $\{\tilde{\psi} \circ v_{\mathbf{i}} \circ \tilde{\psi}^{-1}\}_{\mathbf{i} \in \{1,\ldots,k\}^n}$ are uniformly contracting. They are also totally nonlinear and C^2 with bounded distortion. If I holds then their ranges fill $[0,1]$ and if II holds then they are analytic. Furthermore, under either option they remain totally nonlinear. By [22, Theorem 1.1] their measure of maximal entropy $\tilde{\nu}_0$ (which is Gibbs and atomless) therefore has polynomial decay of its Fourier transform, that is, for all $l \in \mathbb{R} \backslash \{0\}$,

$$\left| \int_0^1 e^{-2\pi i l y} \mathrm{d}\tilde{\nu}_0(y) \right| \leq C |l|^{-\eta}$$

for some $\eta > 0$ and $C < \infty$.

Now, this measure $\tilde{\nu}_0$ is also the measure of maximal entropy of the conjugated single iterates $\{\tilde{\psi} \circ v_i \circ \tilde{\psi}^{-1}\}_{i=1,\ldots,k}$; from the conjugacy we therefore know that $\tilde{\nu}_0 = \tilde{\psi}^* \nu_0$. Hence,

$$\int_0^1 e^{-2\pi i l y} \mathrm{d}\tilde{\nu}_0(y) = \int_0^1 e^{-2\pi i l \omega(\psi(y))} \mathrm{d}\nu_0(y)$$

$$= e^{-2\pi i l \omega_0} \int_0^1 e^{-2\pi i l \omega_1 \psi(y)} \mathrm{d}\nu_0(y),$$

so setting $l = j / \omega_1$ we obtain that

$$\left| \int_0^1 e^{-2\pi i j \omega_1 \psi(y)} \mathrm{d}\nu_0(y) \right| \leq C |\omega_1|^\eta |j|^{-\eta},$$

as required. $\qquad\square$

The relevance of ν_0 is because it is the cross-section of the SRB measure along lines of constant x (i.e. local stable manifolds):

Proposition B.3. *Let ν_0 be as in Lemma B.2. Then $\rho = \mathrm{Leb} \times \nu_0$ is the SRB measure of b.*

Henceforth we will find it useful to notate the unstable dynamics $\kappa(x) = kx \mod 1$.

Proof of Proposition B.3. ρ is conditionally absolutely continuous along unstable manifolds (which are lines of fixed y), and solves

$$
\begin{aligned}
\int_D A \circ b \, \mathrm{d}\rho &= \int_0^1 \int_0^1 A(\kappa(x), v_{\lceil kx \rceil}(y)) \, \mathrm{d}\nu_0(y) \, \mathrm{d}x \\
&= \sum_{i=1}^k \int_{(i-1)/k}^{i/k} \int_0^1 A(\kappa(x), v_i(y)) \, \mathrm{d}\nu_0(y) \, \mathrm{d}x \\
&= \sum_{i=1}^k \int_0^1 k^{-1} \int_0^1 A(x, v_i(y)) \, \mathrm{d}\nu_0(y) \, \mathrm{d}x \\
&= \sum_{i=1}^k \int_0^1 k^{-1} \int_0^1 A(x, y) \, \mathrm{d}(v_i^* \nu_0)(y) \, \mathrm{d}x \\
&= \int_0^1 \int_0^1 A(x, y) k^{-1} \sum_{i=1}^k \mathrm{d}(v_i^* \nu_0)(y) \, \mathrm{d}x \\
&= \int_0^1 \int_0^1 A(x, y) \, \mathrm{d}\nu_0(y) \, \mathrm{d}x = \int_D A \, \mathrm{d}\rho.
\end{aligned}
$$

Hence, it is an SRB measure. $\qquad\square$

This allows us to prove the existence of our conditional measures ρ_t:

Proof of Proposition 3.1. To prove (7), noticing that ρ is just a product measure of the uniform measure in x (therefore t) and ν_0 in y, we have that

$$
\frac{1}{2\delta} \int_{\{\Psi(s,y) \mid |s-t| < \delta, y \in [0,1]\}} A(x,y) \, \mathrm{d}\rho = \int_{[0,1]} \frac{1}{2\delta} \int_{[s-\delta, s+\delta]} A(\psi(y) - t, y) \, \mathrm{d}t \, \mathrm{d}\nu_0(y),
$$

where ν_0 is defined in Lemma B.2. This integral is absolutely bounded by $\sup |A|$, and so by the dominated convergence theorem

$$
\begin{aligned}
\lim_{\delta \to 0} \frac{1}{2\delta} \int_{\{\Psi(s,y) \mid |s-t| < \delta, y \in [0,1]\}} A(x,y) \, \mathrm{d}\rho &= \int_{[0,1]} \lim_{\delta \to 0} \frac{1}{2\delta} \int_{[t-\delta, t+\delta]} A(\psi(y) - s, y) \, \mathrm{d}s \, \mathrm{d}\nu_0(y) \\
&= \int_{[0,1]} A(\psi(y) - t, y) \, \mathrm{d}\nu_0(y) \\
&= \int_D A \, \mathrm{d}(\Psi(t; \cdot)_* \nu_0).
\end{aligned}
$$

This means we must define

$$
\rho_t := \Psi(t; \cdot)_* \nu_0, \tag{17}
$$

and so get (7).

To prove the second part, we have that for any $E \subseteq D_\Psi$,

$$
\rho_t(E) = \int_{[0,1]} \mathbb{1}_E(\Psi(t; y)) \, \mathrm{d}\nu_0(y),
$$

so

$$
\int_{-t_*}^{t_*} \rho_t(E) \, \mathrm{d}t = \int_{-t_*}^{t_*} \int_{[0,1]} \mathbb{1}_E(\Psi(t; y)) \, \mathrm{d}\nu_0(y) \, \mathrm{d}t.
$$

By a change of coordinates $(x, y) = \Psi(t, y) = (\psi(y) - t, y)$, and using Proposition B.3, we have

$$\int_{-t_*}^{t_*} \rho_t(E) \, \mathrm{d}t = \int_{D_\Psi} \mathbb{1}_E(x, y) \, \mathrm{d}\rho(x, y) = \rho(E).$$

$\square$

The following technical lemmas will be of use in following proofs. We will prove them in Appendix C.

Lemma B.4. *Suppose that for some $\alpha \in (0, 1]$, $\phi : [0, 1] \to \mathbb{R}$ is a piecewise α-Hölder function with a finite number of jumps, and ν is an integrable atomless measure. Let $\hat{\phi}_j, \hat{\nu}_j$ be the respective Fourier coefficients of ϕ and ν. Then*

$$\int \phi \, \mathrm{d}\nu = \sum_{j \in \mathbb{Z}} \hat{\phi}_{-j} \hat{\nu}_j, \tag{18}$$

provided the sum is absolutely convergent.

For $\alpha \in (0, 1]$ let the Hölder semi-norm on a set $E \subseteq [0, 1]$ be defined as follows:

$$|\phi|_{C^\alpha(E)} := \sup_{[x,y] \subseteq E} \frac{|\phi(x) - \phi(y)|}{|x - y|^\alpha}. \tag{19}$$

Lemma B.5. *Suppose $\phi : [0, 1] \to \mathbb{R}$ is piecewise α-Hölder with jumps on a set $S \subset (0, 1)$ (including at 1 if it is not periodic). Then for $j \in \mathbb{Z} \backslash \{0\}$,*

$$\left| \int_0^1 \phi(x) e^{-2\pi i j x} \, \mathrm{d}x \right| \leq |\phi|_{C^\alpha(S^c)} |j|^{-\alpha} + |S| \|\phi\|_{L^\infty} |j|^{-1}.$$

With these lemmas in hand, we will try and prove exponential conditional mixing in the projection of the baker's map onto the x coordinate. This next lemma is the heart of the proof

Lemma B.6. *Suppose $\phi : \mathbb{R}/\mathbb{Z}$ is as in Lemma B.5, and ν is an integrable atomless measure with Fourier coefficients $\hat{\nu}_j$, (id, ν) has (C_ν, η) Fourier decay, and $\alpha > 1 - \eta$. Then*

$$\left| \int_0^1 \phi \circ \kappa^n \, \mathrm{d}\nu - \int_0^1 \phi \, \mathrm{d}x \int_0^1 \mathrm{d}\nu \right| \leq \frac{4 C_\nu}{\alpha + \eta - 1} k^{-n\eta} (|\phi|_{C^\alpha(S^c)} + |S| \|\phi\|_{L^\infty}).$$

Proof. Recall that $\kappa^n(x) = k^n x \mod 1$. The Fourier coefficients of $\phi \circ \kappa^n$ are zero except for those whose indices are multiples of k^n:

$$\int_0^1 \phi(\kappa_n(x)) e^{-2\pi i j k^n x} \, \mathrm{d}x = \hat{\phi}_{k^n j}.$$

These decay as $\mathcal{O}(|j|^{-\alpha})$, whereas the Fourier coefficients of ν are $\mathcal{O}(|j|^{-\eta})$, so we know their convolution is summable. By Lemma B.4 we therefore have

$$\int_0^1 \phi \circ \kappa^n \, \mathrm{d}\nu = \sum_{j \in \mathbb{Z}} \hat{\phi}_{-j} \hat{v}_{k^n j}.$$

This means

$$\left| \int_0^1 \phi \circ \kappa^n \, \mathrm{d}\nu - \hat{\phi}_0 \hat{\nu}_0 \right| \leq \sum_{j=1}^\infty |\hat{\phi}_{-j}||\hat{v}_{k^n j}| + |\hat{\phi}_j||\hat{v}_{-k^n j}|$$

$$\leq \sum_{j=1}^\infty 2 \left(|\phi|_{C^\alpha(S^c)} |j|^{-\alpha} + |S| \|\phi\|_{L^\infty} |j|^{-1} \right) C_\nu |k^n j|^{-\eta}$$

$$\leq 2 C_\nu k^{-n\eta} \left(\frac{\alpha + \eta}{\alpha + \eta - 1} |\phi|_{C^\alpha(S^c)} + \frac{1 + \eta}{\eta} |S| \|\phi\|_{L^\infty} \right).$$

Elementary inequalities on the fractions, and the zeroth Fourier coefficient's definition as the total integral give the required result. $\square$

We now attempt to connect this one-dimensional picture in κ to the two-dimensional picture of the baker's map. In this proposition we define a one-dimensional observable $A_{m,y_0}(x)$ that in the following proposition we find will closely approximate $A(b^m(x,y))$ for any y, when m is large enough.

Lemma B.7. *Suppose that ν, α are as in Lemma B.6. Suppose that $A : D \to \mathbb{R}$ has $|A|_{\alpha,x} < \infty$ and let*

$$A_{m,y_0}(x) := A(b^m(x,y_0)).$$

Then

$$\left| \int_0^1 A_{m,y_0} \circ \kappa^n \, d\nu - \int_0^1 A_{m,y_0} \, dx \int_0^1 \, d\nu \right| \leq \frac{4C_\nu}{\alpha + \eta - 1} k^{m - n\eta} (|A|_{\alpha,x} + \|A\|_{L^\infty}).$$

Proof. It is clear that A_{m,y_0} is piecewise α-Hölder with jumps at $S_m := \{i/k^m : i = 1, \ldots, k^m\}$.

We can also bound its Hölder constant. Suppose $[x, z] \subset (0,1] \backslash S_m$. This means that $b^l(x, y_0)$ and $b^l(x, z_0)$ lie on the same piece of b for all $0 \leq l < m$, and therefore that $b^m(x, y_0)$ and $b^m(z, y_0)$ have the same z component. As a result,

$$\begin{aligned}
|A_{m,y_0}(x) - A_{m,y_0}(z)| &\leq |A(b^m(x, y_0)) - A(\kappa^m(z), y_0)| \\
&\leq |A|_{\alpha,x} |\kappa^m(x) - \kappa^m(z)|^\alpha \\
&\leq |A|_{\alpha,x} k^{m\alpha} |x - z|^\alpha.
\end{aligned}$$

From (19), this means that $|A_{m,y_0}|_{C^\alpha(S_m^c)} \leq |A|_{\alpha,x} k^{m\alpha}$.

Applying Lemma B.6, we get that

$$\begin{aligned}
\left| \int_0^1 A_{m,y_0} \circ \kappa^n \, d\nu - \int_0^1 A_{m,y_0} \, dx \int_0^1 \, d\nu \right| &\leq \frac{4C_\nu}{\alpha + \eta - 1} |k|^{-n\eta} (|A_{m,y_0}|_{C^\alpha(S_m^c)} + |S_m| \|A_{m,y_0}\|_{L^\infty}) \\
&\leq \frac{4C_\nu}{\alpha + \eta - 1} k^{-n\eta} (|A|_{\alpha,x} k^{m\alpha} + k^m \|A\|_{L^\infty}) \\
&\leq \frac{4C_\nu}{\alpha + \eta - 1} k^{m - n\eta} (|A|_{\alpha,x} + \|A\|_{L^\infty}),
\end{aligned}$$

as required. $\qquad\square$

Proposition B.8. *For all $\beta > 0$, $m \in \mathbb{N}$, A with finite $|\cdot|_{\beta,y}$ norm, and $x, y, y_0 \in [0,1]$,*

$$|A_{m,y_0}(x) - (A \circ b^m)(x, y)| \leq \mu^{\beta m} |A|_{\beta,y} |y - y_0|^\beta.$$

Proof. We prove this by induction on m. We have that

$$A_{0,y_0}(x) - A(x, y) = A(x, y_0) - A(x, y),$$

which is bounded for all $y, y_0 \in [0,1]$ by $|A|_{\beta,y} |y - y_0|^\beta$. Suppose then that our proposition holds for some m. Then for any $y \in [0,1]$,

$$b^{m+1}(x, y) = b^m(\kappa(x), \nu_i(y)),$$

where the map branch $i = \lceil kx \rceil$. As a consequence, $A_{m+1,y_0}(x) = A_{m,\nu_i(y_0)}(\kappa(x))$, and so

$$\begin{aligned}
|A_{m+1,y_0}(x) - (A \circ b^{m+1})(x, y)| &= A_{m,\nu_i(y_0)}(\kappa(x)) - (A \circ b^m)(\kappa(x), \nu_i(y)) \\
&\leq \mu^{\beta m} |A|_{\beta,y} |\nu_i(y) - \nu_i(y_0)|^\beta \\
&\leq \mu^{\beta(m+1)} |A|_{\beta,y} |y - y_0|^\beta
\end{aligned}$$

as required for the inductive step, where we used that the ν_i contract points by a factor of μ. $\quad\square$

We can then put Lemma B.7 and Proposition B.8 together to prove a primitive version of Lemma B.2:

Proposition B.9. *Suppose $\alpha + \eta > 1$. Then there exists $\xi < 1$ depending only on η, k, μ and β and there also exists C such that if ψ is C^1 with $\psi' = 0$ on a finite set, and ν is an atomless measure such that (ν, ψ) has $(\eta, C_{\nu,\psi})$ Fourier decay, then*

$$\left| \int_D (A \circ b^n)(\psi(y), y) \, \mathrm{d}\nu(y) - \int A \, \mathrm{d}\rho \int \mathrm{d}\nu \right| \leq C C_{\nu,\psi} \xi^n \|A\|_{\alpha;\beta}. \tag{20}$$

Proof. We will divide $n = m + l$ and the difference in (20) up into several pieces that we will bound largely using previous results.

To begin with, we have as an application of Proposition B.8 that for any y_0,

$$\left| \int_D \left((A \circ b^{m+l})(\psi(y), y) - (A_{m,y_0} \circ b^l)(\kappa^n(\psi(y))) \right) \, \mathrm{d}\nu(y) \right| \leq \mu^{\beta m} |A|_{\beta,y} \int |\mathrm{d}\nu|. \tag{21}$$

Now, $\psi_* \nu$ is atomless with $(\mathrm{id}, \psi_* \nu)$ having $(\eta, C_{\nu,\psi})$ Fourier decay, so we can apply Lemma B.7 to obtain

$$\left| \int_0^1 (A_{m,y_0} \circ b^l)(\kappa^n(\psi(y))) \, \mathrm{d}\nu(y) - \int_0^1 A_{m,y_0} \, \mathrm{d}x \int \mathrm{d}\psi_* \nu \right| \leq \frac{4 C_{\nu,\psi}}{\alpha + \eta - 1} k^{m - l\eta} (|A|_{\alpha,x} + \|A\|_{L^\infty}). \tag{22}$$

Next,

$$\int_0^1 A_{m,y_0} \, \mathrm{d}x = \int_0^1 A(b^m(x, y_0)) \, \mathrm{d}x$$
$$= \int_D A(b^m(x, y_0)) \, \mathrm{d}\rho(x, y)$$

because from Proposition B.3, the SRB measure ρ projects to Lebesgue measure in the x coordinate. With this, we have that

$$\left| \int_0^1 A_{m,y_0} \, \mathrm{d}x - \int_D A \circ b^m \, \mathrm{d}\rho \right| = \left| \int_D (A_{m,y_0} - A \circ b^m) \, \mathrm{d}\rho \right|$$
$$\leq \mu^{\beta m} |A|_{\beta,y}$$

using Proposition B.8. Recalling also that ρ is b-invariant, and pushing ν forward preserves its total integral, we can say that

$$\left| \int_0^1 A_{m,y_0} \, \mathrm{d}x \int \mathrm{d}\psi_* \nu - \int_D A \, \mathrm{d}\rho \int_D \mathrm{d}\nu \right| \leq \mu^{\beta m} |A|_{\beta,y} \int |\mathrm{d}\nu|. \tag{23}$$

Putting (21), (22) and (23) together, we have that

$$\left| \int_D (A \circ b^{m+l})(\psi(y), y) \, \mathrm{d}\nu - \int_D A \, \mathrm{d}\rho \int_D \mathrm{d}\nu \right| \leq 2 \int |\mathrm{d}\nu| \, \mu^{\beta m} |A|_{\beta,y} + \frac{4 C_{\nu,\psi}}{\alpha + \eta - 1} k^{m - l\eta} (|A|_{\alpha,x} + \|A\|_{L^\infty}).$$

By setting $l = \lceil (1 - \frac{\eta \log k}{\log \mu^{-1} \beta + (1+\eta) \log k}) n \rceil$ we obtain that there exists a constant C depending on $\alpha, \beta, \eta, \mu, k$ such that (20) holds with

$$\xi = k^{-\eta / (\beta + (1+\eta) \log k / \log \mu^{-1})}.$$

$\square$

At this point, if we set $B \equiv 1$, we could prove Lemma B.2 already. However, to incorporate it we need to show that we can multiply ν_0 by sufficiently smooth Hölder functions and still retain adequate Fourier decay.

Lemma B.10. *Suppose that $\psi : [0,1] \circlearrowleft$ is a C^1 diffeomorphism onto its image, and ν_0 is an atomless measure such that (ψ, ν) has $(\eta, C_{\nu_0,\psi})$ Fourier decay. Then for all $\gamma \in (1 - \eta, 1]$ there exists C such that for all $B \in C^\gamma(D)$,*

$$\left| \int_0^1 e^{2\pi i j \psi(y)} B(y, \psi(y)) \, d\nu(y) \right| \leq C_{\nu_0, \psi} C \|B\|_{C^\gamma} |j|^{-(\eta + \gamma - 1)},$$

that is, if the measure $d\sigma(y) := B(y, \psi^{-1}(y)) d\nu(y)$, then (σ, ψ) has $(\eta + \gamma - 1, C_{\nu, \psi} C)$ decay; furthermore σ is atomless.

Proof. We have that σ is atomless because it is defined as an atomless measure multiplied by a bounded function.

Define function $B_\psi \in C^\gamma([0,1])$ such that $B_\psi(x) = B(\psi^{-1}(x), x)$ on $\psi([0,1])$. It is possible to do this so that

$$|B_\psi|_{C^\gamma([0,1])} = |B_\psi|_{C^\gamma(\psi([0,1]))} \leq C'_{\psi'} |B|_{C^\gamma},$$

where $C'_{\psi'} = 1 + \|1/\psi'\|_{L^\infty} \geq 1$, and $\|B_\psi\|_{L^\infty} \leq \|B\|_{L^\infty}$.

By Lemma B.5 we have that $\hat{b}_l$, the Fourier coefficients of B_ψ, have a certain bound,

$$|\hat{b}_l| \leq \begin{cases} \|B\|_{L^\infty}, & l = 0 \\ |l|^{-\gamma} C'_{\psi'} |B|_{C^\gamma}, & \text{else} \end{cases} \leq \min\{1, |l|\}^{-\gamma} C'_{\psi'} \|B\|_{C^\gamma}$$

and so in particular for any j, the Fourier coefficients of $e^{2\pi i j \cdot} B_\psi$, which are just shifts of those of B_ψ, decay as $\mathcal{O}(l^{-\gamma})$.

Therefore, we can apply Lemma B.4 to get that

$$\int_0^1 e^{2\pi i j \psi(y)} B(y, \psi(y)) \, d\nu(y) = \sum_{l \in \mathbb{Z}} \hat{b}_{-j-l} \hat{v}_l$$

and so

$$\left| \int_0^1 e^{2\pi i j \psi(y)} B(y, \psi(y)) \, d\nu(y) \right| \leq C_{\nu, \psi} C'_{\psi'} \|B\|_{C^\gamma} \sum_{l=-\infty}^\infty \min\{|j+l|^{-\gamma}, 1\} \min\{|l|^{-\eta}, 1\}$$

$$\leq C_{\nu, \psi} C'_{\psi'} \|B\|_{C^\gamma} C'' \min\{|j|^{-(\eta+\gamma-1)}, 1\}$$

for some C'' depending on γ, η, giving what is required. $\qquad \square$

This is all we need to prove Lemma B.2.

Proof of Lemma B.2. By the definition of ρ_t in (17),

$$\int_D A(x, y) \, d\rho_t(x, y) = \int_0^1 A(\psi(y) - t, y) \, d\nu_0(y),$$

and so

$$\rho_0(A \circ b^n B) - \rho(B)\rho(A) = \int_0^1 (A \circ b^n)(\psi(y), y) \, d\sigma(y) - \int_D A \, d\rho \int_0^1 d\sigma(y)$$

where $\sigma := B(\cdot, \psi^{-1}(\cdot))\nu_0$.

Since by assumption, (ν_0, ψ) has $(\eta, C_{\nu_0, \psi})$ Fourier decay, σ has $(\eta + \gamma - 1, C_{\nu_0, \psi} C)$ decay for some C depending on η, γ, ψ as a result of Lemma B.10. From this, σ is also atomless. An application of Proposition B.9 gives us our result. $\qquad \square$

Finally, the construction of ρ also allows us to prove the corollary:

Proof of Corollary 3.3. Using Proposition 1.3 it is enough to prove that the SRB measure ρ is lower-Ahlfors regular. Since ρ is a Cartesian product of Lebesgue measure with the Gibbs measure ν_0, ρ is lower-Ahlfors regular if ν_0 is.

Fix $\delta > 0$, and let $M = \lceil \frac{\log \delta}{\log \mu} \rceil$. For any y in the support of ν_0 we have that there exists a composition of M contractions in $\{v_j\}$, which we notate as $v_{\mathbf{j}}$, such that $y = v_{\mathbf{j}}(y_M)$ for some $y_M \in [0,1]$. Set inclusion tells us that we must therefore have $y \in v_{\mathbf{j}}([0,1])$.

The uniform contraction of the v_j by a factor of ν means that the diameter of this set $v_{\mathbf{j}}([0,1])$ must be bounded by μ^M, which by construction is smaller than δ. Hence, $v_{\mathbf{j}}([0,1])$ must be contained in the ball $B(y, \delta)$.

On the other hand, the ν_0-measure of this set $v_{\mathbf{j}}([0,1])$ can be given using the constitutive relation of ν_0 as $k^{-M}\nu_0([0,1]) = k^{-M}$.

We can therefore say that

$$
\begin{aligned}
\nu_0(B(y,\delta)) &\geq \nu_0(v_{\mathbf{j}}([0,1])) \\
&= k^{-M} \geq k^{-1}\delta^{-\log k/\log \mu},
\end{aligned}
$$

as required for lower Ahlfors regularity.

Since the constants here are independent of t, we obtain the uniform-in-t convergence asked for. $\qquad\square$

C   Proofs of some integration lemmas

Here we collect some proofs of lemmas used in Appendix C involving integration.

Proof of Lemma B.4. For $l \in \mathbb{N}^+$ the 1-periodic Fejér kernel is given by

$$
F_l(x) = \frac{1 - \cos 2\pi l x}{l(1 - \cos 2\pi x)} = \sum_{j=-l}^{l} \left(1 - \frac{|j|}{l} \right) e^{2\pi i j x}.
$$

It is non-negative with total integral equal to 1, and for all $u \in [0, 1/2]$,

$$
\lim_{l\to\infty} \int_{-u}^{u} F_l \, \mathrm{d}x = 1.
$$

As a result, $\phi * F_l$ converges pointwise to ϕ as $l \to \infty$ at all points of continuity of ϕ: which is to say, ν-almost everywhere, since ϕ has a finite number of jumps and ν has no atoms. Furthermore, the functions $\phi * F_l$ are uniformly bounded by the constant function $\|\phi\|_\infty$ (which is ν-integrable). As a result, we can apply the dominated convergence theorem to say that

$$
\int \phi \, \mathrm{d}\nu = \lim_{l\to\infty} \int_0^1 \phi * F_l \, \mathrm{d}\nu \tag{24}
$$

Now,

$$
\begin{aligned}
\int_0^1 \phi * F_l \, \mathrm{d}\nu &= \int_0^1 \sum_{j=-l}^{l} \left(1 - \frac{|j|}{l} \right) \hat{\phi}_{-j} e^{-2\pi i j x} \, \mathrm{d}\nu(x) \\
&= \sum_{j=-l}^{l} \left(1 - \frac{|j|}{l} \right) \hat{\phi}_{-j} \hat{\nu}_j, \tag{25}
\end{aligned}
$$

where in the last line we could interchange integration and (finite) summation. Now, (25) is none other than the lth Césaro sum of $\hat{\psi}_{-j}\hat{\nu}_j$, whose limit is therefore the full sum, the full sum being absolutely convergent. Substituting this limit into (24) we obtain (18) as required. $\quad\square$

Proof of Lemma B.5. We can divide up the interval of integration into j even pieces as such:

$$\int_0^1 \phi(x)e^{-2\pi ijx}\,\mathrm{d}x = \sum_{l=0}^{|j|-1}\int_{l/|j|}^{(l+1)/|j|}\phi(x)e^{-2\pi ijx}\,\mathrm{d}x.$$

We always have on any of these segments that

$$\left|\int_{l/j}^{(l+1)/j}\phi(x)e^{-2\pi ijx}\,\mathrm{d}x\right| \leq \frac{1}{|j|}\|\phi\|_\infty. \tag{26}$$

Let the index set of segments with jumps be

$$J_j = \left\{l : \left(\tfrac{l}{|j|}, \tfrac{l+1}{|j|}\right)\cap X \neq \emptyset\right\}.$$

Clearly we have that the cardinality of J_j is smaller than that of X for all j. For $l \notin J_j$, we can do the usual Hölder continuity bound on the integral, using to begin with that $e^{2\pi ijx}$ is mean zero:

$$\int_{l/|j|}^{(l+1)/|j|}\phi(x)e^{-2\pi ijx}\,\mathrm{d}x = \int_{l/|j|}^{(l+1)/|j|}\left(\phi(x) - \phi(l/|j|)\right)e^{-2\pi ijx}\,\mathrm{d}x$$

and then that

$$|\phi(x) - \phi(l/|j|)| \leq |\phi|_{C^\alpha(X^c)}(x - l/|j|)^\alpha \leq |\phi|_{\alpha,X^c}|j|^{-\alpha}$$

to get that

$$\left|\int_{l/|j|}^{(l+1)/|j|}\phi(x)e^{-2\pi ijx}\,\mathrm{d}x\right| \leq |\phi|_{C^\alpha(X^c)}|j|^{-\alpha-1}. \tag{27}$$

Combining (26) for the segments with jumps and (27) otherwise, we get that

$$\left|\int_0^1\phi(x)e^{-2\pi ijx}\,\mathrm{d}x\right| \leq |J_j|\frac{1}{|j|}\|\phi\|_\infty + (|j| - |J_j|)|\phi|_{C^\alpha(X^c)}|j|^{-\alpha-1}$$

$$\leq |X|\|\phi\|_\infty|j|^{-1} + |\phi|_{C^\alpha(X^c)}|j|^{-\alpha}$$

as required. $\qquad\square$

D Proof of results in Section 1

Proof of Proposition 1.1. Since T is conservative and topologically mixing it must have one physical measure ρ which is Lebesgue measure.

Because H has no critical points on ℓ_H, the conditional measure $\rho(x \mid H(x) = 0)$ can be well-defined as a C^0-weak limit of

$$\mu_\delta(x) = \frac{1}{c(\delta)}\psi(|H(x)|/\delta), \tag{28}$$

for some bump function $\psi \in C_c^\infty$, with $c(\delta) = \mathcal{O}(\delta^d)$. This is constant along manifolds $\{x : H(x) = c\}, |c| \leq \delta$, which are C^1 manifolds for small enough δ.

We know that the stable vector bundle of T is continuous and ℓ_H is compact. Hence, the manifolds $\{x : H(x) = c\}, |c| \leq \delta$ must decompose locally into submanifolds of the same dimensionality as the unstable dimension, which are *uniformly* transverse to stable vector fields, and therefore some set of admissible stable leaves in the sense of [14]. Furthermore, $c(\delta) \sim C\delta^d$ for some $C > 0$.

It can then be shown by a computation that the family (28) is convergent in the $\mathcal{B}^{1,1}$ norm of [14], including when multiplied by C^1 functions B, and so the conditional measure lies in $\mathcal{B}^{1,1}$, on which the Perron-Frobenius operator of T has a spectral gap [14, Theorem 2.3]. Hence, exponential conditional mixing obtains with $r = 1$. $\qquad\square$

Proof of Proposition 1.2. We will consider convergence of $d_{\text{Haus}}(\Lambda, \overline{T^n(\ell_H \cap \Lambda)})$, as it is the same thing as without the set closure.

From its definition, the conditional measure's support $\operatorname{supp}\mu$ must be contained in $\ell_H \cap \Lambda$, and so $\operatorname{supp} T_*^n\mu$ is contained in $\overline{T^n(\ell_H \cap \Lambda)}$.

Now, $\overline{T^n(\ell_H \cap \Lambda)} \subset \overline{T^n(\Lambda)} \subseteq \overline{\Lambda} = \Lambda$, so it is enough to show that

$$\lim_{n\to\infty} \inf_{x\in\Lambda} d(x, \operatorname{supp} T_*^n\mu) = 0. \tag{29}$$

Fix $\epsilon > 0$, and let $\{B(\xi, \epsilon)\}_{\xi\in\Xi}$ be a finite open cover of Λ. Then, if for each $\xi \in \Xi$ we can show that $B(\xi, \epsilon)$ has positive $T_*^n\mu$-measure for every n large enough, we have that $d_{\text{Haus}}(T^n(\ell_H \cap \Lambda), \Lambda) < 2\epsilon$ for these n, and so we are done.

Conditional mixing implies that if $\psi_{\xi,\epsilon}$ is any C^∞ non-negative bump function bounded by 1 whose support is $B(\xi, \epsilon)$, then

$$\lim_{n\to\infty} \int \psi_{\xi,\epsilon} \circ T^n \, \mathrm{d}\mu = \int \psi_{\xi,\epsilon} \mathrm{d}\rho \int \mathrm{d}\mu > 0,$$

because $\psi_{\xi,\epsilon} > 0$ on an open set overlapping with $\Lambda = \operatorname{supp}\rho$. This means that for n large enough,

$$0 = \int \psi_{\xi,\epsilon} \circ T^n \, \mathrm{d}\mu = \int \psi_{\xi,\epsilon} \, \mathrm{d}T_*^n\mu \leq T_*^n\mu(B(\xi,\epsilon)),$$

as required. $\qquad\square$

Proof of Proposition 1.3. By translation and dilation we can construct C^∞ bump functions $\psi_{\xi,\epsilon}$ such that $\psi_{\xi,\epsilon} = 1$ on $B(\xi, \epsilon/2)$, and their C^r norms are bounded $\|\psi_{\xi,\epsilon}\|_{C^1} \leq K\epsilon^{-r}$ for constant K.

Lower-Ahlfors regularity of ρ gives us that

$$\int \psi_{\xi,\epsilon} \, \mathrm{d}\rho \geq \rho(B(\xi, \epsilon/2)) \geq C(\epsilon/2)^d,$$

and the exponential conditional mixing assumption then gives us that for all ξ, ϵ,

$$\int \psi_{\xi,\epsilon} \circ T^n \, \mathrm{d}\mu > C(\epsilon/2)^d - C\xi^n K\epsilon^{-r}.$$

This will all be positive when $\epsilon \geq K_1(\xi^{1/(d+r)})^n$ for some constant K_1, giving us uniform exponential decay in n of the bound 2ϵ on the Hausdorff distance. $\qquad\square$

E   Proof of Proposition A.1

To prove Proposition A.1 we will first require the following:

Proposition E.1. *There exist $C > 0$, $\lambda > 1$ such that for all $I \in \hat{\mathcal{L}}$ and $n \geq 0$ such that $f^m I \cap \mathcal{S} = \emptyset$ for $m < n$,*

$$|f^n I| > C\lambda^n |I|.$$

Proof. The segments I are unstable manifolds: because f is piecewise uniformly hyperbolic [25], these segments are eventually expanded by the action of f. $\qquad\square$

Proof of Proposition A.1. Define the observable on $\vec{\Lambda}$

$$P(\vec{I}_{p,q}, t) = \mathbb{1}(\mathcal{M}_p \neq \mathcal{M}_{\pi((\vec{I}_{p,q},t))}) \tag{30}$$

This measures whether $\vec{I}$ is cut by $\mathcal{S}$ on the left-hand side. Clearly, if $P(\vec{f}^{-n}(\vec{I}, t)) = 1$, then $p_{\vec{I}} = f^n(s)$ for some $s \in \mathcal{S}$. This holds *a fortiori* if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} P(\vec{f}^{-n}(\vec{I}, t)) > 0 \tag{31}$$

In the following we will show that this limit is almost always positive.

Fix $\epsilon > 0$, and let $\vec{\Lambda}_{\mathrm{erg}}$ be an ergodic component of $\vec{\Lambda}$. Because almost all unstable manifolds have positive measure, the definition of $\vec{\Lambda}_{\mathrm{erg}}$, there exists a set $E \subset \vec{\Lambda}$ of positive measure such that $|\vec{I}| > \epsilon$ for all $(\vec{I}, t) \in E$. It is clear that we can choose $E = E_I \times (0,1)$, E_I being a collection of directed segments.

We know that for any segment $\vec{J}$, if $\vec{J}$ does not cross $\mathcal{S}$ then $\vec{f}(\vec{J}, t) = (f(\vec{J}), t)$. By Proposition E.1, this means that for $\vec{I} \in A_I$, $f^n(\vec{I}, t) = (f^n(\vec{I})t)$ with $|f^n(\vec{I})| \geq C\epsilon\lambda^n$, unless $\vec{I}$ is cut by $\mathcal{S}$ for some $m < n$. However, there is some m_* sufficiently large such that $C\epsilon\lambda^{m_*}$ is greater than the diameter of the attractor Λ, which brings about a contradiction, as $f^n(I) = \mathcal{W}_{\mathrm{loc}}^u(f^n(\pi(\vec{I}, 0.5)))$ is an segment subset of Λ. Thus, $f^m(\vec{I})$ is cut by $\mathcal{S}$ for some $m > n$.

Now, recalling that our segments are open segments, $\mathcal{S}$ will cut $f^m\vec{I}$ at some point $t' \in (0,1)$. As a result, $P(\vec{f}^{m+1}(I, t)) = 1$ for $t \in (0, t')$. This holds for any $\vec{I} \in E_I$.

The consequence is that, for at least one m' between 0 and m_*, there exists $B_P \subset f^{m'+1}(A) \subset \vec{\Lambda}_{\mathrm{erg}}$ of positive measure such that $P = 1$ on B_P. By applying the Birkhoff ergodic theorem to $(\vec{f}^{-1}, \vec{\Lambda}_{\mathrm{erg}}, \vec{\rho}|_{\vec{\Lambda}_{\mathrm{erg}}})$, the limit (31) must hold for $\vec{\rho}$-almost all points on on $\vec{\Lambda}_{\mathrm{erg}}$.

By taking a union over all ergodic components of $\vec{\Lambda}$ we therefore have that for almost all $\vec{I}_{p,q} \in \vec{\Lambda}$, $p_{\vec{I}} = f^n(s)$ for some $s \in \mathcal{S}$. This is to say that $p_{\vec{I}}$ lies in the forward orbit of the singular line. Using that reversing the direction of segments is a measure isometry, this result equally holds true for q. $\qquad\square$

Acknowledgements

Data availability statement

The datasets generated and/or analysed during the current study are available from the author on reasonable request.

Conflict of interest statement

The author has no conflicts of interest to declare.

References

[1] Viviane Baladi. Optimality of Ruelle's bound for the domain of meromorphy of generalized zeta functions. *Portugaliae mathematica*, 49(1):69–83, 1992.

[2] Viviane Baladi. *Positive transfer operators and decay of correlations*, volume 16. World Scientific, 2000.

[3] Viviane Baladi. *Dynamical zeta functions and dynamical determinants for hyperbolic maps*. Springer, 2018.

[4] Viviane Baladi and Sébastien Gouëzel. Good Banach spaces for piecewise hyperbolic maps via interpolation. In *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, volume 26, pages 1453–1481. Elsevier, 2009.

[5] Krzysztof Barański, Yonatan Gutman, and Adam Śpiewak. A probabilistic Takens theorem. *Nonlinearity*, 33(9):4940, 2020.

[6] Christian Bonatti, Sylvain Crovisier, Lorenzo J Díaz, and Amie Wilkinson. What is... a blender? *Notices of the American Mathematical Society*, 63(10):1175–1178, 2016.

[7] Jean Bourgain and Semyon Dyatlov. Fourier dimension and spectral gaps for hyperbolic surfaces. *Geometric and Functional Analysis*, 27(4):744–771, 2017.

[8] Oliver Butterley, Giovanni Canestrari, and Sakshi Jain. Discontinuities cause essential spectrum. *Communications in Mathematical Physics*, 398(2):627–653, 2023.

[9] Nikolai I Chernov. Limit theorems and Markov approximations for chaotic dynamical systems. *Probability Theory and Related Fields*, 101(3):321–362, 1995.

[10] Mark Demers and Carlangelo Liverani. Stability of statistical properties in two-dimensional piecewise hyperbolic maps. *Transactions of the American Mathematical Society*, 360(9):4777–4814, 2008.

[11] Dmitry Dolgopyat. On decay of correlations in anosov flows. *Annals of mathematics*, 147(2):357–390, 1998.

[12] Arnaud Doucet, Nando De Freitas, Neil James Gordon, et al. *Sequential Monte Carlo methods in practice*, volume 1. Springer, 2001.

[13] Georg A Gottwald, Caroline L Wormell, and Jeroen Wouters. On spurious detection of linear response and misuse of the fluctuation–dissipation theorem in finite time series. *Physica D: Nonlinear Phenomena*, 331:89–101, 2016.

[14] Sébastien Gouëzel and Carlangelo Liverani. Banach spaces adapted to anosov systems. *Ergodic Theory and Dynamical Systems*, 26(1):189–217, 2006.

[15] Michael Hochman and Pablo Shmerkin. Equidistribution from fractal measures. *Inventiones mathematicae*, 202(1):427–479, 2015.

[16] Michal Misiurewicz. Strange attractors for the Lozi mappings. *Annals of the New York Academy of Sciences*, 357(1):348–358, 1980.

[17] Michał Misiurewicz and Sonja Štimac. Symbolic dynamics for lozi maps. *Nonlinearity*, 29(10):3031, 2016.

[18] Carolina A Mosquera and Andrea Olivo. Fourier decay of self-similar measures on the complex plane. *arXiv preprint arXiv:2207.11570*, 2022.

[19] Carolina A Mosquera and Pablo Shmerkin. Self-similar measures: asymptotic bounds for the dimension and Fourier decay of smooth images. *Annales Academiæ Scientiarum Fennicæ Mathematica*, 43:823–834, 2018.

[20] Lea Oljača, Tobias Kuna, and Jochen Bröcker. Exponential stability and asymptotic properties of the optimal filter for signals with deterministic hyperbolic dynamics. *arXiv preprint arXiv:2103.01190*, 2021.

[21] David Ruelle. Linear response theory for diffeomorphisms with tangencies of stable and unstable manifolds—a contribution to the Gallavotti-Cohen chaotic hypothesis. *Nonlinearity*, 31(12):5683, 2018.

[22] Tuomas Sahlsten and Connor Stevens. Fourier transform and expanding maps on Cantor sets. *arXiv preprint arXiv:2009.01703*, 2020.

[23] Jörg Schmeling and Serge Troubetzkoy. Dimension and invertibility of hyperbolic endomorphisms with singularities. *Ergodic Theory and Dynamical Systems*, 18(5):1257–1282, 1998.

[24] Caroline L Wormell. On convergence of linear response formulae in some piecewise hyperbolic maps. *arXiv preprint arXiv:2206.09292*, 2022.

[25] Lai-Sang Young. Bowen-Ruelle measures for certain piecewise hyperbolic maps. In *The Theory of Chaotic Attractors*, pages 265–272. Springer, 1985.