Abstract

Chaotic systems notoriously tend to wash out inexact information about their state as they evolve. Probability measures, which encode such information, therefore tend to equilibrate as they are pushed forward, typically converging to a physical or SRB measure. Many statistical properties of the system reduce to the presence and rate of this convergence for certain measures. Typically these measures have some level of smoothness in one or more directions, which enables a proof of convergence, usually at an exponential rate.

In this paper we give evidence that conditional measures of the SRB measure along submanifolds of the phase space (that are not stable or unstable manifolds) generically also have exponential convergence back to the full SRB measures, a property of a system that we call conditioned decay of correlations (CDoC). We will prove that CDoC holds for a class of generalised baker’s maps, and give rigorous numerical evidence in its favour for some piecewise hyperbolic maps. CDoC is not necessarily expected under existing theory, but naturally encodes the idea of long-term prediction of systems using perfect partial observations, and appears key to a rigorous understanding of the emergence of linear response in high-dimensional systems.

Chaotic dynamics’ definitional problem is understanding the long-time behaviour of typical orbits. This behaviour is quantitatively encoded by SRB measures, a special class of physically meaningful invariant measures of a system. Many important problems in chaotic dynamics can be formulated in terms of these measures, including how long the system at equilibrium retains memory (mixing rates), and the mean response of the system to dynamical perturbations (the response problem).

Mathematically, a SRB probability measure $\rho$ of a chaotic diffeomorphism $f : M \to M$ is a measure whose conditional measures along unstable manifolds of $f$ are absolutely continuous with respect to the conditional Lebesgue measure. If $f$ is topologically mixing and $\rho$ is hyperbolic, an SRB measure is a physical measure in the sense that it is the limiting distribution of the empirical measure of the forward orbit of Lebesgue-almost any point [19]. In fact, for many systems we also know something stronger, specifically that $\rho$ is a weak limit of Lebesgue measure when pushed forward by $f$, with the convergence being exponential for smooth observables.

One might therefore ask the question: for what initial probability measures $\mu$ does the pushforward $f^n_{\ast} \mu \to \rho$ in some again suitably weak sense? We could formulate this as the convergence of future expectations of sufficiently regular (e.g. $C^1$) observables $A$:

$$\int_M A \circ f^n d\mu \xrightarrow{n \to \infty} \int_M A d\rho.$$ (1)

In other words, when does an initial guess $\mu$ of the system’s state tell us precisely that it is physically generic, and no more? We might also specify some speed on the convergence of (1), that is the rate of information loss: in the best case, this might be exponential.

For most systems with additive noise we know that the convergence (1) happens at an exponential rate for all probability measures $\mu$. On the other hand, for many chaotic systems our understanding is much more incomplete.

We can describe some measures $\mu$ for which this happens: for maps with some uniform hyperbolicity it is known that it includes those that are structurally similar to the SRB measure:

Decay of correlations after conditioning on generic submanifolds in deterministic chaos

Caroline Wormell

June 22, 2022
loosely, these measures are conditionally absolutely continuous on a set of submanifolds that are of the correct dimension and whose tangent spaces are uniformly transverse to stable manifolds, and where these conditional densities have a certain additional level of regularity. Rigorously speaking, these measures lie in certain anisotropic Banach spaces \([6, 2]\). Included in this class are smooth absolutely continuous measures, and often SRB measures of small perturbations of the map.

On the other hand, there are obvious examples where \([1]\) doesn’t obtain: for example, when \(\mu\) is a delta function, \(\mu\) is supported on a single stable manifold, or \(\mu\) is an invariant measure of \(f\) other than \(\rho\) (for example, an equilibrium measure with a different potential). One might even engineer measures having \(\{\int A \circ f^n \, d\mu\}_{n \in \mathbb{N}}\) being any kind of sequence, but these measures are generally arrived at via nefarious mathematical processes and are physically unlikely.

The question then arises: what about other sorts of measures? Of course there are many interesting classes with some physically meaningful derivation. The only class of measures for which this question has actually been studied are Gibbs measures of a nonlinear map, when \(f(x)\) is a linear interval map: this in the context of Fourier dimension \([15, 10]\) and understanding normality of various Cantor sets \([8, 3]\). Exponential convergence of \([1]\) appears to hold generically for measures in this particular class.

This paper will demonstrate this property, and show that it is worth understanding, for another class: conditional measures of \(f\)’s SRB measure along generic manifolds, i.e., those that are neither stable nor unstable manifolds. As well as being philosophically important in forecasting chaotic systems (discussed in Section \([3]\)), this kind of \textit{conditioned decay of correlations} (CDoC) is key to resolving the open fundamental question of the validity of linear response theory in dissipative dynamics (see Section \([4]\) for further discussion). Our study of these measures also supports a broad conjecture that exponential convergence of \([1]\) holds true generically for non-“adversarial” classes of measures \(\mu\).

The conditional measures we study typically lack the regularity necessary to apply the standard transfer operator techniques: for example, they are supported on Cantor sets (see Figure \([1]\)). New theoretical approaches will therefore be needed to fully rigorously answer the question. In lieu, we will prove a result for a toy class of baker’s maps by harnessing Fourier dimension theory; to consider a less special case, we will also perform some careful rigorously justified numerics on the Lozi maps. To this end, we present an efficient algorithm to sample a long time series from the Lozi map’s true SRB measure up to some uniform rigorous error bound: the existence of such an algorithm might be considered surprising.

The paper is structured as follows: in Sections \([1, 2]\) we give evidence that CDoC can be expected to hold, respectively presenting a theorem for a class of (potentially non-linear) baker’s maps and numerical evidence for some piecewise hyperbolic maps; in Sections \([3, 4]\) we give two major applications of CDoC, in the areas of prediction and linear response respectively; in Section \([5]\) we describe the algorithms used to obtain our numerical results; finally Section \([6]\) is dedicated to the proof of the baker’s map results.

1 Rigorous results for a toy model: baker’s map

As a simple model to study CDoC, let us consider baker’s maps \(b : D := [0, 1]^2 \cap \) of the following form:

\[
b(x, y) = (kx \mod 1, \nu_{\lfloor kx \rfloor}(y))
\]

where \(k \geq 2\) is an integer, the \(v_i, i = 1, \ldots, k\) are contractions with all \(|v'_i| \leq \mu < 1\). We will also assume that \(v_i((0, 1))\) are disjoint, and they have bounded distortion (i.e. the \(\log|v'_i|\) are \(C^1\)). An example of such a map is plotted in Figure \([2]\). For smooth enough transverse foliations, it is possible to define in a natural way conditional measures for all leaves:

**Proposition 1.1.** Suppose that \(\Psi(t; y) = (\psi(y) - t, y)\) is a foliation of some subset \(D_\psi \subseteq D\) for \(|t| \leq t_\ast \in [0, 1]\) and \(y \in [0, 1]\). Let \(\rho\) be the (unique) SRB measure of \(b\).

Then for every \(t \in [-t_\ast, t_\ast]\) there exists a unique probability measure \(\rho_t\) supported on \(\Psi(t, [0, 1])\) such that
Figure 1: Left: picture of the Lozi attractor at $a = 1.7, b = 0.5$ (black), the singular line $\{x = 0\}$ (purple), their intersection (orange). Right: histogram of $\rho(y \mid 0)$, obtained from 200,000 iterates of the unstable manifold dynamics $\vec{f}$, binned at width 0.0025.

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

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

(3)

b. For all Borel sets $E \subseteq D_y$,

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

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 $\phi : D \to \mathbb{C}$:

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

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 $\phi : D \to \mathbb{C}$ with $\|\phi\|_{\alpha, \beta} < \infty$: that is, functions that are $\alpha$-Hölder in the $x$ direction and $\beta$-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 Section 6, gives that for certain baker’s maps, exponential CDoC holds for all conditional SRB measures on a smooth foliation transversal to unstable lines:

**Theorem 1.2.** Suppose that $\Psi(t; y) = (\psi(y) - t, y)$ is a foliation of some subset of $D$ for $|t| \leq t^* \in [0, 1]$ and $y \in [0, 1]$, and $\psi$ is $C^2$ with $\psi' \neq 0$. Suppose one of the following conditions hold:

1. The contractions $v_i$ are $C^2$ and $\cup_i v_i([0, 1]) = [0, 1]$. There is no $v_i$ at $1$. This is the original version.
II. The contractions $v_i$ are analytic, as is $\psi$.

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

Let $\rho$ 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 $\rho$ on the foliation. Then 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 $|t| < t_*$, $A \in C^{\alpha, \beta}$, $B \in C^{\gamma}$, $n \in \mathbb{N}$,

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

Note that the classic piecewise affine baker’s maps fall under condition III, although for conservative maps the application of Fourier dimension theory is unnecessary due to the absolute continuity of $\rho$.

Let us remark that, from the proof, the constant $d_*$ is the Fourier dimension of $\rho_t$ projected onto the $x$ coordinate. This Fourier dimension is bounded from above by (and, perhaps, often equal to) the Hausdorff dimension of $\rho_t$, which, since $b$ is Markov, is the stable dimension of the system. Thus, a larger stable dimension suggests a decay of correlations with respect to less regular observables, as hoped for in Section 4 and [14].

To prove Theorem 1.2 we use essentially two facts. The first fact, used in Lemma 6.6, 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 6.3) 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 [10, 15].

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 II the $\psi'' \neq 0$ restriction can be dropped: in this case therefore, quadratic tangencies with stable manifolds (lines of constant $y$) are in fact allowed.
2 Numerical example: Lozi map

While we are able to find rigorous results for them, baker’s maps are rather special, not least in their skew product structure, and it is not obvious to us how to generalise these theoretical results. Therefore we instead go on to present rigorously justified numerical evidence for CDoC in a less rarified class: those of Lozi maps.

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

$$f(x, y) = (1 + y - a|x|, bx), \ b \neq 0. \quad (4)$$

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 $b \in (0, (a - \sqrt{2}))$ this has a single mixing SRB measure \[9, \text{Theorem 5}\], all unstable manifolds have positive length when $b \in (0, a - \sqrt{2})$ \[18, \text{Theorem}\]. A Lozi attractor is shown in Figure 3.

Lozi maps are continuous, with a jump in the Jacobian on the singularity set $S = \{x = 0\}$. In a similar fashion to those we defined for the baker’s map (Proposition 1.1), conditional measures of the SRB measure $d\rho(x, y)$ on sets $\ell_{x_0} := \{x = x_0\}, x_0 \in \mathbb{R}$ are well-defined for all $x_0$ intersecting the support of the Lozi attractor \[17, \text{Theorem 2.1}\]. Let us note these conditional measures as $d\rho(y | x_0)$. 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 $d\rho(y | 0) = d\rho(y | (x, y) \in S)$ is plotted in Figure 1: linear response for the Lozi map is determined from the mixing properties of the conditional measure on $S$ \[17\], so we will be most interested in this particular conditional measure.

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

Conjecture 2.1. For generic Lozi parameters $(a, b)$ and $x_0 \in \mathbb{R}$, there exist $C > 0, c \in (0, 1)$ such that for all $A, B \in C^1$ and $n \geq 0$,

$$\left| \int_{\mathbb{R}^2} (A \circ f^n) B \, d\rho(y | x_0) - \int_{\mathbb{R}^2} A \, d\rho \int_{\mathbb{R}^2} B(x_0, y) d\rho(y | x_0) \right| \leq C\|A\|_{C^1} \|B\|_{C^1} c^n. \quad (5)$$

We have strong and direct numerical evidence that this holds: Figure 4 shows exponential decay of the left-hand side of (5) by four orders of magnitude, with reliable error quantification. In fact, it seems that $A$ need only be piecewise $C^1$. The consequence of this (up to some
specialisation of Conjecture 2.1 on the singular line $\ell_0$ is that the Lozi map formally has linear response [17].

It should be noted that obtaining valid samples of the quantities in (5) is tricky: not only is one sampling from the conditional measure $\rho(y \mid x_0)$, which is a codimension-one object on the attractor. We will present the novel methods required to obtain this evidence in Section 5. These algorithms could, we imagine, be generalised to general piecewise hyperbolic dynamics.

It is interesting that the rates of convergence are substantially slower than the decay of correlation rates for smooth observables against the full SRB measure: for example, in Figure 4 the rate of exponential convergence for a smooth observable $A(x, y) = x$ is much slower when the initial measure $\mu$ is a conditional measure rather than for the full SRB measure $\mu \sim \rho_p$. Intuitively, mixing rates, which depend on the essential spectrum of the transfer operator in relevant function spaces, tend to be slower for functions of low (e.g. Hölder) regularity [1]: thinking of the conditional measure $\rho(y \mid x_0)$ as equivalent to the SRB measure multiplied with a distribution $\delta(y - x - x_0)$ this slow “decay of correlations” can be at least intuitively understood. It is of course likely that this mixing rate also seriously depends on the conditioning manifold, and it would be interesting to know to what extent one has continuity in mixing rates, for example [15].

3 Application I: prediction with a perfect partial observa-

Let us now consider some practical problems to which the notion of CDoC is applicable. The first of these is the fundamental problem of prediction of chaotic dynamics. Many prediction methods have been developed that assimilate information obtained from observations. In general, these methods achieve this assimilation by mimicking or approximating Bayes’ rule [16].
To understand this in the simplest instance, let us suppose that our system $f$ has exponential decay of correlations, and at time $n = 0$, we have some prior probabilistic knowledge of the state of our system, given by some (presumably “nice”) measure $d\mu^-(x)$. If we start with an unobserved system at statistical equilibrium, the natural choice would be $\mu^- = \rho$, the SRB measure.

We can now make a (perhaps partial) observation of our system, given as a value $y = H(x) \in \mathbb{R}^e$. We might reasonably assume that our observation $H$ is noisy, with probabilities given by a smooth kernel $p(y \mid x) dy$.

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

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

with $Z(y) = \int_M p(y \mid \xi) d\mu^-(\xi)$. Our best guess of the state of the system at future time $n$ (i.e. of $f^n(x)$) is then given by $f^n\mu$; the expected value of some (nice) observable $A$ at time $n$ is

$$\int_M A(f^n(\xi)) d\mu(\xi) = \int_M A(f^n(\xi)) \frac{p(y \mid \xi)}{Z(y)} d\mu^-(\xi).$$

Depending on how effective a measurement $y = H(x)$ is of $x$, $\mu$ is likely to be concentrated on a smaller set than $\mu^-$, and this may improve forward estimates of the system’s state over the short to medium term. However, under our assumptions for $f, A, \mu^-, p$, exponential decay of correlations results give that (7) will eventually converge to a fixed exponential rate to the SRB measure expectation $\int_M A d\rho$.

But as we make our observations more and more precise, i.e. reduce the noise in $H$ to zero, what do we end up with? This is trivial if $H(x)$ specifies $x$: we know $f^n(x)$ exactly for all time. It is common, particularly in large systems, for $H$ to be only partial observation.

In a zero-noise limit the kernel $p$ is no longer smooth, with $p(y \mid \xi) = \delta(y - H(\xi))$. Our posterior measure $d\mu(x)$ is then the conditional measure of $d\mu^-(x)$ given that $H(x) = y$. Our future expectations of observable $A$ then become

$$\int_M A(f^n(\xi)) d\rho(\xi \mid H(\xi) = y).$$

Following the intuition that incomplete information on the system should wash out over time, we would expect that for most choices of $H$ this converges to the SRB measure expectation for all $A$ (i.e. CDoC holds). Indeed, this is what our results above suggest. On the other hand, we don’t yet have a grip on this theoretically, and indeed if partial observations could have some predictive use for all time this might fundamentally change how we approach prediction.

Of course, in practice physical observations will always have some random error to them. However, this is also true of the evolution of physical systems, which are considered worth studying in their zero noise limit. In any case, if an error is small enough it may take a long time to manifest, and the zero-noise limit is what captures the medium term behaviour.

4 Motivation II: linear response theory

Less obviously, CDoC arises in response theory: that is, existence and computation of derivatives of the map $f \mapsto \rho_f$. In this section, we will sketch how this comes to be, and extend a conjecture of [14].

Supposing we have a one-parameter family of maps $\varepsilon \mapsto f_\varepsilon$ for $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, where the $f_\varepsilon$ are chaotic with SRB measures $\rho_\varepsilon$, we might ask about the regularity of the map

$$A : \varepsilon \mapsto \int_M A \, d\rho_\varepsilon$$

for smooth observables $A \in C^2$. If this regularity can be established to some order, then we can often obtain some formula that gives us derivatives of $A(\varepsilon)$, ideally in terms of the system at $\varepsilon = 0$.

*If $H, \mu^-$ is absolutely continuous, then $Z(y)$ remains a function and we may be able to recover some version of (6), see for example Proposition 4.4.*
For simplicity, let us consider the first derivative (the so-called linear response). Let us assume that $\varepsilon \to f_\varepsilon$ is $C^1$ in a suitable sense, and let $X : M \to TM$ be $\frac{df_\varepsilon}{d\varepsilon}|_{\varepsilon=0}$. Then, at least formally, we can write the derivative

$$A'(0) = \sum_{n=0}^{\infty} \int_M (\nabla A) \circ f^n \cdot DF^n X d\rho$$

(8)

In all instances where we have proofs, absolute convergence of this sum implies linear response. Nevertheless some massaging is required to obtain this, because the differential $DF^n$ explodes when applied to a generic vector as a result of $f$’s chaoticity. If we suppose that $f$ is (non-uniformly) hyperbolic with stable subspaces $E^s_x$ and unstable subspaces $E^u_x$ defined for $\rho$-almost all $x \in M$, then we can decompose

$$X = X^u + X^s$$

with $X^u(x) \in E^u_x$, $X^s(x) \in E^s_x$ for almost all $x \in M$. Then, $D_x f^n X^s(x) \to 0$ as $n \to \infty$, so the main point of contention in (8) is the contribution of $X^u$. Here, we can perform an integration by parts:

$$\int_M (\nabla A) \circ f^n \cdot DF^n X^u d\rho = \int_M \nabla (A \circ f^n) \cdot X^u d\rho = - \int_M A \circ f^n \nabla \cdot (X^u d\rho).$$

(9)

For there to be a linear response, we require these terms to decay to zero summably in $n$. This can be obtained if we can show that the (hyper-)distribution $\nabla \cdot (X^u d\rho)$ has fast decay of correlations with respect to $f$.

4.1 Smooth, uniformly hyperbolic maps

In the best-case scenario of smooth, uniformly hyperbolic maps, this is assured by two facts. Firstly, because the stable and unstable foliations are Hölder, $X^u$ is a regular, bounded function. Secondly, because the conditional densities of $\rho$ on unstable manifolds are bounded away from zero and at least Hölder, $\rho$ can be differentiated in unstable directions (such as in the direction of $X^u$). This means that $\nabla \cdot (X^u d\rho)$ is absolutely continuous with respect to $\rho$, with a Hölder density:

$$\frac{\nabla \cdot (X^u d\rho)}{d\rho} = \nabla \cdot X^u + \frac{X^u \cdot \nabla (d\rho)}{d\rho}.$$  

(10)

This is often known as the “unstable divergence” [11], an example of which is given for a hyperbolic map in Figure 5. When the unstable divergence is Hölder, classical exponential decay of correlations results ensure the susceptibility function converges [11].

4.2 Piecewise hyperbolic maps

On the other hand, in physical systems we largely expect tangencies between stable and unstable manifolds. This situation is very difficult to study, but in a first instance we can model the “folding” behaviour present in non-uniformly hyperbolic systems (see Figure 6) by jump singularities in the derivative of an otherwise hyperbolic map such as the Lozi map along some submanifold $S$ (see Figure 3).

These singularities in the map introduce various issues in the preceding picture. In particular, the unstable and stable subspaces no longer vary continuously in space. However, we can isolate these (jump) discontinuities respectively to the backwards and forwards orbits of the singular set (see Figure 5 for an example in the forward direction), and see that the discontinuities decay exponentially as one iterates forward and backward from $S$ [17]. Furthermore, the regularity of the conditional measure of $\rho$ just depends on historic regularity of unstable subspaces. Consequently, the unstable divergence [10] becomes a (convergent) sum of delta functions over the forward and backward orbits, plus an absolutely continuous background of little trouble. We
Figure 5: Heatmap of the density $\sum (X^u \, dp)$ for a $C^\infty$ injective hyperbolic map of the torus given in Appendix C and perturbation $X(x, y) = (\sin 2\pi x, 0)$. This density can clearly be seen to be regular (in fact, $C^\infty$).

have therefore that

$$\int_M A \circ f^n \nabla \cdot (X^u \, dp) \sim \sum_{m \in \mathbb{Z}} \int_M A(f^n(x)) \phi_m(x) \delta(f^{-m}(x) \in S) \, dp$$

$$= \sum_{m \in \mathbb{Z}} \int_M A \circ f^{n+m} \phi_m \circ f^m \delta_S \, dp. \quad (11)$$

Now, up to a normalisation constant $\delta_S \, dp$ is really just the conditional measure of $\rho$ along the singular line $S$. Given the $|\phi_m| \to 0$ quickly as $|m| \to \infty$, we see that, provided we have conditional exponential mixing, (11) decays exponentially as $n \to \infty$ [17, Theorem 2.3]. In Section 2 we give numerical evidence that this is the case.

4.3 Non-uniformly hyperbolic maps

When one moves away from hyperbolic systems, the picture becomes more clouded. To begin with, we cannot make a nice separation $X = X^u + X^s$: on the one hand, stable and unstable subspaces no longer vary continuously in space, so we cannot assume much regularity in our $X^u$; on the other hand, on a certain subset of the attractor $C$, the stable and unstable subspaces may overlap and thus fail to span the entire tangent space, so even to the extent some regularity is achievable we would expect $X^u$ and $X^s$ to blow up near these points of tangency. Furthermore, unstable manifolds are contracted in the forward orbit of points in $C$, introducing kinks which lead to singularities in the conditional density of the SRB measure $\rho$. These processes lead to an unstable divergence with many non-integrable singularities (see Figure 6).

However, we can theorise some order. If $f$ is a smooth diffeomorphism, the set $C$ is forward and backward invariant. Let us suppose that for each orbit $\{x_n\}_{n \in \mathbb{Z}} \subset C$ there is a point $x_0$ where the offending tangent vectors in $E^s \cap E^u$ (generically one-dimensional) transition from being expanded (as unstable vectors) to being contracted (as stable vectors); we might imagine that most of these points can be located on some sort of submanifold $C_0$ where the attractor is “folded” by the map [14]: a picture of the situation for the Hénon map is plotted in Figure 6.

The relative contribution of each of the singularities on this submanifold can be expected to be equivalent to $\rho$ conditioned on $C_0$: provided that some distributional derivative of this
conditional measure mixes sufficiently quickly when pushed forward by $f$, we might obtain summability of $\mathcal{C}_0$.

A similar argument was originally put forward by Ruelle [13, 14], the difference being that Marstrand’s theorem is invoked to say that the projection of the singularities onto a single unstable manifolds should have Hölder conditional density, which would give mixing of the singularities. Quantitatively speaking, this should occur when the Hausdorff dimension of $\mathcal{C}_0$ is greater than $\frac{1}{2}$. However, for linear response, exponential mixing of the derivative of the singularities’ density is required: Ruelle says this occurs “if the correlations for $\rho$ decay exponentially in a suitable sense”. In general, exponential mixing of a Hölder function certainly does not guarantee exponential mixing of any sort of derivative: however, because this function is given by (a convolution of) the singularities, a CDoC argument may provide the answer.

Figure 6: Heatmap (in log scale) of the density $\nabla \cdot (X_u \frac{d\rho}{d\rho})$ for the Hénon map $f(x, y) = (1 - 1.4x^2 + y, 0.3x)$ and perturbation $X(x, y) = (0, x)$. In black, an approximation of the line of folding, obtained as a line of tangencies of finite time Lyapunov exponents, and in grey two of its forward iterates (dotted lines) and reverse iterates (dashed lines). Blow-up of the density can be observed along the intersection of these lines with the attractor.

5 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 $\rho(\cdot | x_0)$ 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 5.1 and a numerical implementation given in Section 5.2.

However, we would like to be sure we are not merely perceiving 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 5.3 gives, surprisingly, a
stable algorithm to simulate the chaotic Lozi dynamics that is compatible with validated interval
arithmetic, and Section 5.4 explains the quantification of random sampling errors.

5.1 Segment dynamics
Let \( \Lambda \) be the attractor of the Lozi map \( f \), and define the local unstable manifold of a point \( p \in \Lambda \) to be
\[
W^u_{loc}(p) := \left\{ q \in \Lambda : \lim_{n \to \infty} |f^{-n}(q) - f^{-n}(p)| = 0, \forall n \in \mathbb{N}^+ \right\}.
\]
where \( \sigma_{(x,y)} := \text{sign } x \). These are segments of the full unstable manifolds which have always
remained on the same side of the singular line \( S = \{ x = 0 \} \).

Let \( \mathcal{G} \) 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
\[
\mathcal{E} = \left\{ I \in \mathcal{G} : \exists p \in \Lambda I = W^u_{loc}(p) \right\},
\]
which captures \( \rho \)-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
\[
\bar{\Lambda} = \mathcal{E} \times (0,1),
\]
which we are going to use to parametrise each \( I \in \mathcal{E} \). \( \bar{\Lambda} \) is almost everywhere a two-to-one cover
of \( \Lambda \) by the map \( \pi(I_{p,q}, t) = (1-t)p + tq \), where we denote the directed segment from point \( p \) to
point \( q \) by \( I_{p,q} \).

As a result, up to a set of \( \rho \)-measure zero, we can lift the \( f \)-dynamics to \( \bar{\Lambda} \), by a map of the form
\[
\tilde{f}(I, t) = (f(I \cap M_{\pi(I,t)}), \nu_I(t)),
\]
where \( \nu_I : [0,1] \to \mathbb{R} \) is a full-branch expanding interval map. Let us define this a little more
explicitly.

When \( I_{p,q} \cap S \) is non-empty, then we know it has exactly one element which we denote by
\( s \in S \) with \( s = \pi(I_{p,q}, t) \), for some \( t \in (0,1) \). The segment dynamics \( \tilde{f} \) can then be written
explicitly as
\[
\tilde{f}(I_{p,q}, t) = \begin{cases} 
(I_{f(p),f(q)}, t), & I_{p,q} \cap S = \emptyset \\
(I_{f(p),f(s)}, t/t_s), & I_{p,q} \cap S \neq \emptyset \text{ and } t < t_s \\
(I_{f(s),f(q)}, (t-t_s)/(1-t_s)), & I_{p,q} \cap S \neq \emptyset \text{ and } t > t_s.
\end{cases}
\]

(13)

It turns out that the SRB measure \( \rho \) also can be lifted to an invariant measure of \( \tilde{f} \) by
\[
\int_{\Lambda} \Psi d\tilde{\rho} = \int_{\Lambda} \frac{\sum_{(I,t) \in \pi^{-1}(x)} \Psi(I, t)}{2} d\rho(x),
\]
because \( \pi^{-1} \) is two-to-one \( \rho \)-almost everywhere we have that \( (\bar{\Lambda}, \tilde{f}, \tilde{\rho}) \) has most two ergodic
components (which are identical up to reversing the direction of the segments), and we will be
able to sample \( \tilde{\rho} \) by iterating \( \tilde{f} \).

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 5.1.** The set
\[
\left\{(I_{p,q}, t) : p, q \in \bigcup_{n=1}^{\infty} f^n(S) \right\}
\]
has full \( \tilde{\rho} \) measure.

This proposition is proved in Appendix B.
5.2 Simulation of segment dynamics

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

$$\int_{\tilde{I}} A(x_0, y) \, d\rho(y \mid x_0) = \frac{\int_{\tilde{I}} A \circ \kappa_{x_0} / \ell \, d\tilde{\rho}}{\int_{\tilde{I}} 1 / \ell \, d\tilde{\rho}}$$

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

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

with the segment dynamics that can be simulated using (13). The convergence (15) for the $\Psi$ we are interested in uses the Birkhoff ergodic theorem and the fact that $\kappa_{x_0}, \ell$ do not depend on the direction of the interval (so the ergodic component of $\tilde{\Lambda}$ we sample from is inmaterial).

Because from (13), 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

$$\tilde{f}(\tilde{I}_{p_{n+1}, q_{n+1}}, T_{n+1}) = \begin{cases} (\tilde{I}_{f_{p_{n}}, f(q_{n})}, T_{n}), & \tilde{I}_{p_{n}, q_{n}} \cap S = \emptyset \\ (\tilde{I}_{f_{p_{n}}, f_{s_{n}}}, T_{n}) \text{ with probability } t_{s}, & \tilde{I}_{p_{n}, q_{n}} \cap S = \emptyset \\ (\tilde{I}_{f_{s_{n}}, f(q_{n})}, T_{n}) \text{ with probability } 1-t_{s}, & \tilde{I}_{p_{n}, q_{n}} \cap S \neq \emptyset. \end{cases}$$

(16)

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

$$T_{n+1} = \nu_{\tilde{I}_{p_{n}, q_{n}}}(T_{n}).$$

To sample that $\tilde{f}$ 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 $T_{\text{final}} \sim \text{Uniform}(0,1)$ and iterating backwards according to (13); inverting the $\nu_T$ 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.

5.3 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 section 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 $\alpha$ by an interval $\alpha \in \mathcal{I}$ such that we know $\alpha \in \alpha \subset \mathbb{R}$. Such an interval $\alpha$ 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 $g : \mathbb{R}^d \to \mathbb{R}^c$ can be implemented in validated arithmetic through a function $\tilde{g} : \mathcal{I}^d \to \mathcal{I}^c$ such that $g(\alpha)$ will always contain $\tilde{g}(\alpha)$. In such a way one can be absolutely certain that $\alpha$ is contained in a set $\tilde{g}(\alpha)$ and so on.

The unstable manifolds $\tilde{I}_{p_{n}, q_{n}}$ 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 \([13]\) 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 \(\times\) 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 \(\vec{I}_{p_1,q_1}\), 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

\[
Q_1R_1 = JQ,
\]  

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

\[
\vec{I}_{p_1,q_1} = R_1\vec{I}_{p',q'} + Q_1 e_1.
\]

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 \(\mathcal{M}_b\), 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\(^4\).

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 \(Q^\top S\) which solves some equation \(x' = \beta y'\), with \(\beta\) bounded because unstable manifolds are uniformly transversal to the singular line \([16]\). 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 5.1, 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 in this section in validated interval arithmetic is straightforward, because it is stable, but here we must be careful: when our local unstable manifold is split and the choice of child manifold is to be made, the \(t_\star\) used to determine the choice is interval-valued (i.e. in this set \(I\) 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 \(t_\star\); the choice of segment is then clear except where there is an overlap between \(U\) and \(t_\star\). 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\(^5\).

We will therefore be able to find an interval hypercube containing an exact time series \(\{\hat{f}^n(I_0,t_0)\}_{n=0,...,N}\) from \([16]\), where \(t_0\) is implicitly defined by the random selection when the segment is cut.

---

\(^4\)It is however necessary to explicitly code the QR decomposition \([17]\) 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.

\(^5\)One compares \(U\) with some real number \(t_\star \in t_\star\) and reweights the time series by \(t_\star/t_\star\) or \((1-t_\star)/(1-t_\star)\) as appropriate: note that the weights are also interval-valued.
5.4 Quantification of statistical error

As the \( \tilde{f} \) dynamics we sample is just a two-to-one lift from the \( f \) dynamics, which have a spectral gap \([2]\), we expect that for large \( N \) the error between Birkhoff means and true expectations \([15]\) obey a central limit theorem \([4]\). If we have several long time series

\[
\Psi^{(r)} = \frac{1}{N} \sum_{n=0}^{N-1} \Psi(\tilde{f}^n(\tilde{f}^r)))
\]

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

\[
\frac{\Psi - \tilde{p}(\Psi)}{\sqrt{RN}} \sim t_{R-1},
\]

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 4. Such a principle has been used to test for linear response in previous work \([7]\).

To obtain an accurate sample from \( \tilde{p} \) in initialising our time series we begin by initialising \((\tilde{I}_{p_0,f(p_0,t^{(r)})}) \in \Lambda \), where \( p_0 = \left( \frac{2}{\pi(a + b + 4s)}, 0 \right) \) 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 with exponential decay of correlations \([5]\) and so by making the spin-up time \( n_{\text{inst}} \) sufficiently long, we can ensure that our sampling initialisations \( \tilde{f}^{(r)} = \tilde{f}^{n_{\text{inst}}}(\tilde{I}_{p_0,p_1,t^{(r)}}) \) come from a distribution exponentially close to \( \tilde{p} \).

6 Proof of baker’s map result

In this section we will prove Theorem 1.2 on exponential conditional mixing for baker’s maps of the form \([2]\).

For concision when quantitatively referring to Fourier dimension, let us say that a measure-function pair \((\nu, \psi)\) has \((\delta, C)\) Fourier decay if for all \( j \in \mathbb{Z}\setminus\{0\} \)

\[
\left| \int_{0}^{1} e^{2\pi ij\psi(x)} \, d\nu(x) \right| \leq C|j|^{-\delta},
\]

and \( \int_{0}^{1} |d\nu| \leq C \). This implies that the Fourier dimension of \( \psi \ast \nu \) is at least equal to \( \delta \).

Fourier decay is invariant under translations of \( \psi \):

\textbf{Proposition 6.1.} Suppose that \((\nu, \psi)\) has \((C, \delta)\) Fourier decay. Then for all \( t \in \mathbb{R} \), so does \((\nu, \psi + t)\).

\textit{Proof.} We have

\[
\left| \int_{0}^{1} e^{2\pi ij(\psi(x)+t)} \, d\nu(x) \right| = \left| \int_{0}^{1} e^{2\pi ij\psi(x)} \, d\nu(x) \right| \leq C|j|^{-\delta},
\]

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

To prove Theorem 1.2 we will employ a separate Fourier dimension theorem \([10]\) \([15]\) 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):

\textbf{Lemma 6.2.} Suppose one has a modified baker’s map \( b \) with contracting maps \( v_i \). Let \( v_0 \) be the probability measure such that \( v_i^*v_0 = k^{-1}v_0 \) for all \( i = 1, \ldots, k \), and let \( \psi \in C^2 \) be such that \( \psi' \neq 0 \) and \((v_0, \psi)\) has \((C, \delta)\) Fourier decay.
Let $\gamma \in (1 - \delta, 1)$, $\beta \in (0, 1]$ and $\alpha \in (2 - \delta - \gamma, 1)$.

Then there exists $\xi \in (0, 1)$ and $C'$ depending only on $C, \xi, \alpha, \beta, \gamma, \psi'$ 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 $\rho_0$ is defined as in Theorem 1.2.

Proof of Theorem 1.2. The measures $\nu_0$ from Lemma 6.2 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 $(\nu_0, \psi)$ has Fourier decay then so do $(\nu_0, \psi)$ uniformly from Proposition 6.1 and Lemma 6.2 then secures us the theorem. Since $\nu_0$ is a probability measure, it is only necessary to check that (18) holds, which we do procedurally from existing results.

That (18) holds for option III is a simple application of [10, Theorem 3.1] (and in fact here, $\delta$ is independent of $\psi$).

To see it for options I and II requires a little more cunning. We have that $\psi$ is a diffeomorphism onto its image: let $\omega(x) = \omega_0 + \omega_1 x$ map $\psi([0, 1])$ linearly onto $[0, 1]$, so that $\tilde{\psi} = \omega \circ \psi : [0, 1] \to [0, 1]$ is a diffeomorphism. If $\{\psi_i\}_{i \in \{1, \ldots, k\}}$ are $n$-fold compositions of the contractions $\psi_i$, then for some large enough $n$, the $n$-fold compositions $\{\psi \circ \psi_i \circ \psi^{-1}\}_{i \in \{1, \ldots, k\}}$ are uniformly contracting. They are also totally nonlinear and $C^2$ with bounded distortion. If (11) holds then their ranges fill $[0, 1]$ and if (11) holds then they are analytic. By [15] 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 $l \in \mathbb{R}\setminus\{0\}$,

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

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

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

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

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

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

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

as required.

The relevance of $\nu_0$ is because it is the cross-section of the SRB measure along lines of constant $x$ (i.e. local stable manifolds):

Proposition 6.3. Let $\nu_0$ be as in Lemma 6.2. 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 6.3 $\rho$ is conditionally absolutely continuous along unstable manifolds (which
are lines of fixed $y$, and solves
\[
\int_D A \circ b \, d\rho = \int_0^1 \int_0^1 A(\kappa(x), v_{|kx|}(y)) \, d\nu_0(y) \, dx
\]
\[
= \sum_{i=1}^{k} \int_{(i-1)/k}^{i/k} \int_0^1 A(\kappa(x), v_i(y)) \, d\nu_0(y) \, dx
\]
\[
= \sum_{i=1}^{k} \int_0^1 k^{-1} \int_0^1 A(x, v_i(y)) \, d\nu_0(y) \, dx
\]
\[
= \sum_{i=1}^{k} \int_0^1 k^{-1} \int_0^1 A(x, y) \, d(v^*_i \nu_0)(y) \, dx
\]
\[
= \int_0^1 \int_0^1 A(x, y) \, d\nu_0(y) \, dx = \int_D A \, d\rho.
\]
Hence, it is an SRB measure. \[\square\]

This allows us to prove the existence of our conditional measures $\rho_t$:

**Proof of Proposition 1.1** To prove (3), noticing that $\rho$ is just a product measure of the uniform measure in $x$ (therefore $t$) and $\nu_0$ in $y$, we have that
\[
\frac{1}{2\delta} \int_{\{\Psi(s,y) || s-t|<\delta, y \in [0,1]\}} A(x, y) \, d\rho = \int_{[0,1]} \frac{1}{2\delta} \int_{[|t-s|, t+s]} A(\psi(y) - s, y) \, ds \, d\nu_0(y),
\]
where $\nu_0$ is defined in Lemma 6.2. This integral is absolutely bounded by $\text{sup} |A|$, and so by the dominated convergence theorem
\[
\lim_{\delta \to 0} \frac{1}{2\delta} \int_{\{\Psi(s,y) || s-t|<\delta, y \in [0,1]\}} A(x, y) \, d\rho = \int_{[0,1]} \lim_{\delta \to 0} \frac{1}{2\delta} \int_{[|t-s|, t+s]} A(\psi(y) - s, y) \, ds \, d\nu_0(y)
\]
\[
= \int_{[0,1]} A(\psi(y) - t, y) \, d\nu_0(y)
\]
\[
= \int_D A \, d(\Psi(t, \cdot), \nu_0).
\]
This means we must define
\[
\rho_t := \Psi(t, \cdot)_* \nu_0,
\]
and so get (3).

To prove the second part, we have that for any $E \subseteq D_\Psi$,
\[
\rho_t(E) = \int_{[0,1]} 1_E(\Psi(t, y)) \, d\nu_0(y),
\]
so
\[
\int_{-t^*}^{t^*} \rho_t(E) \, dt = \int_{-t^*}^{t^*} \int_{[0,1]} 1_E(\Psi(t, y)) \, d\nu_0(y) \, dt.
\]
By a change of coordinates $(x, y) = \Psi(t, y) = (\psi(y) - t, y)$, and using Proposition 6.3 we have
\[
\int_{-t^*}^{t^*} \rho_t(E) \, dt = \int_{D^*_\Psi} 1_E(x, y) \, d\rho(x, y) = \rho(E).
\]
\[\square\]
The following technical lemmas will be of use in following proofs. We will prove them in Appendix A.

**Lemma 6.4.** Suppose that for some $\alpha \in (0, 1]$, $\phi : [0, 1] \to \mathbb{R}$ is a piecewise $\alpha$-Hölder function with a finite number of jumps, and $\nu$ is an integrable atomless measure. Let $\hat{\phi}, \hat{\nu}$ be the respective Fourier coefficients of $\phi$ and $\nu$. Then

$$
\int \phi \, d\nu = \sum_{j \in \mathbb{Z}} \hat{\phi}_j \hat{\nu}_j,
$$

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:

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

**Lemma 6.5.** Suppose $\phi : [0, 1] \to \mathbb{R}$ is piecewise $\alpha$-Hölder with jumps on a set $S \subset (0, 1]$ (including at 1 if it is not periodic). Then for $j \in \mathbb{Z}\setminus\{0\}$,

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

With these lemmas in hand, we will try and prove decay of correlations for a one-dimensional projection of the baker’s map. This next lemma is the heart of the proof.

**Lemma 6.6.** Suppose $\phi : \mathbb{R}/\mathbb{Z}$ is as in Lemma 6.5 and $\nu$ is an integrable atomless measure with Fourier coefficients $\hat{\nu}$, $(1d, \nu)$ has $(C_\nu, \delta)$ Fourier decay, and $\alpha > 1 - \delta$. Then

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

**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 $k^n$:

$$
\int_0^1 \phi(k^n(x))e^{-2\pi ijk^n x} \, dx = \hat{\phi}_{k^n j}.
$$

These decay as $O(|j|^{-\alpha})$, whereas the Fourier coefficients of $\nu$ are $O(|j|^{-\delta})$, so we know their convolution is summable. By Lemma 6.4 we therefore have

$$
\int_0^1 \phi \circ \kappa^n \, d\nu = \sum_{j \in \mathbb{Z}} \hat{\phi}_j \hat{\kappa^n j}.
$$

This means

$$
\left| \int_0^1 \phi \circ \kappa^n \, d\nu - \hat{\phi}_0 \hat{\nu}_0 \right| \leq \sum_{j=1}^\infty |\hat{\phi}_j| |\hat{\kappa^n j}| + |\hat{\phi}_0| |\hat{\kappa^n 0}|
$$

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

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

Elementary inequalities on the fractions, and the zeroth Fourier coefficient’s definition as the total integral give the required result. \qed
We now attempt to connect this one-dimensional picture in $\kappa$ 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 6.7.** Suppose that $\nu, \alpha$ are as in Lemma 6.6. 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

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

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

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

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

From [21], this means that $|A_{m,y_0}|_{C^\infty(S_m^\infty)} \leq |A|_{\alpha,x} k^{m\alpha}$.

Applying Lemma 6.6, we get that

$$\left| \int_0^1 A_{m,y_0} \circ \kappa^n \, d\nu - \int_0^1 A_{m,y_0} \, dx \int_0^1 \, dv \right| \leq \frac{4C\nu}{\alpha + \delta - 1} k^{-n\delta} (|A_{m,y_0}|_{C^\infty(S_m^\infty)} + |S_m| \|A_{m,y_0}\|_{L^\infty})$$

$$\leq \frac{4C\nu}{\alpha + \delta - 1} k^{-n\delta} (|A|_{\alpha,x} k^{m\alpha} + m \|A\|_{L^\infty})$$

$$\leq \frac{4C\nu}{\alpha + \delta - 1} k^{m-n\delta} (|A|_{\alpha,x} + \|A\|_{L^\infty}),$$

as required. \qed

**Proposition 6.8.** For all $\beta > 0$, $m \in \mathbb{N}$, $A$ with finite $|\cdot|_{\beta,y}$ norm, and $x, y, y_0 \in [0,1]$,

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

**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 $i = \lfloor kx \rfloor$. As a consequence, $A_{m+1,y_0}(x) = A_{m,y_0}(\kappa(x))$, and so

$$|A_{m+1,y_0}(x) - (A \circ b^{m+1})(x,y)| = |A_{m,y_0}(\kappa(x)) - (A \circ b^m)(\kappa(x), \nu_i(y))|$$

$$\leq \mu^m |A|_{\beta,y} |\nu_i(y) - \nu_i(y_0)|^\beta$$

$$\leq \mu^m |A|_{\beta,y} |y - y_0|^\beta,$$

as required for the inductive step, where we used that the $\nu_i$ contract points by a factor of $\mu$. \qed

We can then put Lemma 6.7 and Proposition 6.8 together to prove a primitive version of Lemma 6.2.

18
Proposition 6.9. Suppose $\alpha + \delta > 1$. Then there exists $\xi < 1$ depending only on $\delta, k, \mu$ and $\beta$ and there also exists $C$ such that if $\psi$ is $C^1$ with $\psi' = 0$ on a finite set, and $\nu$ is an atomless measure such that $(\nu, \psi)$ has $(\delta, C_{(\nu, \psi)}$ Fourier decay, then

$$
\left| \int_D (A \circ b^n)(\psi(y), y) \, d\nu(y) - \int_D A \, d\rho \int_D d\nu \right| \leq CC_{(\nu, \psi)} \xi^n \| A \|_{\alpha, \beta}.
$$

(22)

Proof. We will divide $n = m + l$ and the difference in (22) up into several pieces that we will bound largely using previous results.

To begin with, we have as an application of Proposition 6.8 that for any $y_0$,

$$
\left| \int_D ((A \circ b^{m+1})(\psi(y), y) - (A_{m,y_0} \circ b^l)(\kappa^n(\psi(y)))) \, d\nu(y) \right| \leq \mu^{\beta m} |A|_{\beta,y} \int |d\nu|.
$$

(23)

Now, $\psi, \nu$ is atomless with $(\id, \psi, \nu)$ having $(\delta, C_{(\nu, \psi)}$ Fourier decay, so we can apply Lemma 6.7 to obtain

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

(24)

Next,

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

because from Proposition 6.3, the SRB measure $\rho$ projects to Lebesgue measure in the $x$ coordinate. With this, we have that

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

using Lemma 6.2. Recalling also that $\rho$ is $b$-invariant, and pushing $\nu$ forward preserves its total integral, we can say that

$$
\left| \int_0^1 A_{m,y_0} \, dx \int D d\psi, \nu - \int_D A \, d\rho \int D d\nu \right| \leq \mu^{\beta m} |A|_{\beta,y} \int |d\nu|.
$$

(25)

Putting (23), (24) and (25) together, we have that

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

By setting $l = \left[ (1 - \frac{4 \log k}{\log \mu^{-1} + (1 + \delta) \log \xi}) n \right]$ we obtain that there exists a constant $C$ depending on $\alpha, \beta, \delta, \mu, k$ such that (22) holds with

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

At this point, if we set $B \equiv 1$, we could prove Lemma 6.2 already. However, to incorporate it we need to show that we can multiply $\nu_0$ by sufficiently smooth Hölder functions and still retain adequate Fourier decay.
Lemma 6.10. Suppose that \( \psi : [0, 1] \to \mathbb{R} \) is a \( C^1 \) diffeomorphism onto its image, and \( \nu_0 \) is an atomless measure such that \( (\psi, \nu) \) has \((\delta, C_{\nu_0, \psi})\) Fourier decay. Then for all \( \gamma \in (1 - \delta, 1] \) there exists \( C \) such that for all \( B \in C^\gamma(D) \),

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

that is, if the measure \( d\sigma(y) := B(y, \psi^{-1}(y)) \, d\nu(y) \), then \( (\sigma, \psi) \) has \((\delta + \gamma - 1, C_{\nu_0, \psi} C)\) decay; furthermore \( \sigma \) is atomless.

**Proof.** We have that \( \sigma \) 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

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

where \( C_{\psi} = 1 + \|1/\psi'\|_{L^\infty} \geq 1 \), and \( \left\| B_\psi \right\|_{L^\infty} \leq \left\| B \right\|_{L^\infty} \).

By Lemma 6.5 we have that \( \hat{b}_l \), the Fourier coefficients of \( B_\psi \), have a certain bound,

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

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

Therefore, we can apply Lemma 6.4 to get that

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

and so

\[
\left| \int_0^1 e^{2\pi i \psi(y)} B(y, \psi(y)) \, d\nu(y) \right| \leq C_{\nu_0, \psi} C_{\psi} \left\| B \right\|_{C^\gamma} \sum_{l=-\infty}^{\infty} \min\{|l|^{-\gamma}, 1\} \min\{|l|^{-\delta}, 1\}
\]

\[
\leq C_{\nu_0, \psi} C_{\psi} \left\| B \right\|_{C^\gamma} C'' \min\{|j|^{-\left(\delta + \gamma - 1\right)}, 1\}
\]

for some \( C'' \) depending on \( \gamma, \delta \), giving what is required. \( \square \)

This is all we need to prove Lemma 6.2.

**Proof of Lemma 6.2** By the definition of \( \rho_t \) in (19),

\[
\int_D A(x, y) \, d\rho_t(x, y) = \int_0^1 A(\psi(y) - t, y) \, d\rho_0(y),
\]

and so

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

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

Since by assumption, \( (\nu_0, \psi) \) has \((\delta, C_{\nu_0, \psi})\) Fourier decay, \( \sigma \) has \((\delta + \gamma - 1, C_{\nu_0, \psi} C)\) decay for some \( C \) depending on \( \delta, \gamma, \psi \) as a result of Lemma 6.10. From this, \( \sigma \) is also atomless. An application of Proposition 6.3 gives us our result. \( \square \)

**Acknowledgements**

This research has been supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 787304), as well as by the ETH Zurich Institute for Theoretical Studies.

CW thanks Viviane Baladi for her comments on the manuscript.
A Proofs of some lemmas associated with the baker’s map

Proof of Lemma 6.4. For \( l \in \mathbb{N}^+ \) the 1-periodic Fejér kernel is given by

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

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

\[
\lim_{l \to \infty} \int_{-u}^{u} F_l \, dx = 1.
\]

As a result, \( \phi \ast F_l \) converges pointwise to \( \phi \) as \( l \to \infty \) at all points of continuity of \( \phi \): which is to say, \( \nu \)-almost everywhere, since \( \phi \) has a finite number of jumps and \( \nu \) has no atoms. Furthermore, the functions \( \phi \ast F_l \) are uniformly bounded by the constant function \( \|\phi\|_\infty \) (which is \( \nu \)-integrable).

As a result, we can apply the dominated convergence theorem to say that

\[
\int \phi \, d\nu = \lim_{l \to \infty} \int_{0}^{1} \phi \ast F_l \, d\nu \tag{26}
\]

Now,

\[
\int_{0}^{1} \phi \ast F_l \, d\nu = \int_{0}^{1} \sum_{j=-l}^{l} \left( 1 - \frac{|j|}{l} \right) \hat{\phi}_{-j} e^{-2\pi i j x} \, d\nu(x)
\]

\[
= \sum_{j=-l}^{l} \left( 1 - \frac{|j|}{l} \right) \hat{\phi}_{-j} \hat{\nu}_j, \tag{27}
\]

where in the last line we could interchange integration and (finite) summation. Now, \( (27) \) is none other than the \( l \)th Césaro sum of \( \hat{\nu}_j \), whose limit is therefore the full sum, the full sum being absolutely convergent. Substituting this limit into \( (26) \) we obtain \( (20) \) as required. \( \square \)

Proof of Lemma 6.5. We can divide up the interval of integration into \( j \) even pieces as such:

\[
\int_{0}^{1} \phi(x) e^{-2\pi i j x} \, dx = \sum_{l=0}^{\frac{|j|-1}{|j|}} \int_{l/|j|}^{(l+1)/|j|} \phi(x) e^{-2\pi i j x} \, dx.
\]

We always have on any of these segments that

\[
\left| \int_{l/|j|}^{(l+1)/|j|} \phi(x) e^{-2\pi i j x} \, dx \right| \leq \frac{1}{|j|} \|\phi\|_\infty \tag{28}
\]

Let the index set of segments with jumps be

\[
J_j = \left\{ l : \left( \frac{1}{[l/|j|]}, \frac{1}{[l/|j|]} \right) \cap X \neq \emptyset \right\}.
\]

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 \( e^{2\pi i j x} \) is mean zero:

\[
\int_{l/|j|}^{(l+1)/|j|} \phi(x) e^{-2\pi i j x} \, dx = \int_{l/|j|}^{(l+1)/|j|} (\phi(x) - \phi(l/|j|)) e^{-2\pi i j x} \, dx
\]

and then that

\[
|\phi(x) - \phi(l/|j|)| \leq |\phi|_{C^0(X \setminus \{l/|j|\})} \leq |\phi|_{C^0(X \setminus \{l/|j|\})}^{\alpha}
\]

to get that

\[
\left| \int_{l/|j|}^{(l+1)/|j|} \phi(x) e^{-2\pi i j x} \, dx \right| \leq |\phi|_{C^0(X \setminus \{l/|j|\})}^{\alpha - 1} \tag{29}
\]

\[
\int_{l/|j|}^{(l+1)/|j|} \phi(x) e^{-2\pi i j x} \, dx.
\]
Combining (28) for the segments with jumps and (29) otherwise, we get that
\[
\left| \int_0^1 \phi(x)e^{-2\pi i j x} \, dx \right| \leq |J| \left( \frac{1}{|j|} \|\phi\|_\infty + (|j| - |J|)\|\phi\|_{C^\infty(X^s)} |j|^{-\alpha - 1} \right)
\leq |X| \|\phi\|_\infty |j|^{-1} + |\phi|_{C^\infty(X^s)} |j|^{-\alpha}
\]
as required. \hfill \square

\section{Proof of Proposition 5.1}

To prove Proposition 5.1 we will first require the following:

\begin{proposition}
There exist $C > 0$, $\lambda > 1$ such that for all $I \in \mathcal{L}$ and $n \geq 0$ such that $f^m I \cap S = \emptyset$ for $m < n$, \[ |f^n I| > C\lambda^n |I|. \]
\end{proposition}

\begin{proof}
The segments $I$ are unstable manifolds: because $f$ is piecewise uniformly hyperbolic \cite{13}, these segments are eventually expanded by the action of $f$.
\end{proof}

\begin{proof}[Proof of Proposition 5.1]
Define the observable on $\tilde{A}$
\[ P(\tilde{I}_{p,q}, t) = 1(\mathcal{M}_p \neq \mathcal{M}_{x((\tilde{I}_{p,q}, t)))}) \]
This measures whether $\tilde{I}$ is cut by $S$ on the left-hand side. Clearly, if $P(\tilde{f}^{-n}(\tilde{I}, t)) = 1$, then $p_{\tilde{I}} = f^n(s)$ for some $s \in S$. This holds \textit{a fortiori} if
\[ \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N P(\tilde{f}^{-n}(\tilde{I}, t)) > 0 \]
In the following we will show that this limit is almost always positive.

Fix $\epsilon > 0$, and let $\tilde{A}_{\text{erg}}$ be an ergodic component of $\tilde{A}$. Because almost all unstable manifolds have positive measure, the definition of $\tilde{A}_{\text{erg}}$, there exists a set $E \subset \tilde{A}$ of positive measure such that $|\tilde{I}| > \epsilon$ for all $(\tilde{I}, t) \in E$. 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 $\tilde{J}$, if $\tilde{J}$ does not cross $S$ then $\tilde{f}(\tilde{J}, t) = (f(\tilde{J}), t)$. By Proposition 5.1 this means that for $\tilde{I} \in A_I$, $f^m(\tilde{I}, t) = (f^m(\tilde{I}), t)$ with $|f^m(\tilde{I})| \geq C\lambda^m$, unless $\tilde{I}$ is cut by $S$ for some $m < n$. However, there is some $m_\ast$ sufficiently large such that $C\epsilon \lambda^{m_\ast}$ is greater than the diameter of the attractor $\Lambda$, which brings about a contradiction, as $f^m(\tilde{I}) = W_{\text{loc}}^s(f^m(\pi(\tilde{I}, 0.5)))$ is an segment subset of $\Lambda$. Thus, $f^m(\tilde{I})$ is cut by $S$ for some $m > n$.

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

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

By taking a union over all ergodic components of $\tilde{A}$ we therefore have that for almost all $I_{p,q} \in \tilde{A}$, $p_{\tilde{I}} = f^n(s)$ for some $s \in S$. This is to say that $p_{\tilde{I}}$ 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 $q$. \hfill \square

\section{$C^\infty$ hyperbolic map of the torus}

Defining on the torus $(\mathbb{R}/\mathbb{Z})^2$ the cat map $f_0(x, y) = (2x + y, x + y) \mod 1$, a distance function constant on the cat map’s unstable manifolds $c(x, y) = (\frac{2x}{\sqrt{1+\sqrt{5}} - y} \mod 1$, and a perturbation
\[ \tilde{f}(x, y) = (x, y + e^{4-1/x(1-x)}(c(x, y) - 1/2)) \mod 1 \]
we have that $f = \tilde{f} \circ f_0$ is an injective but not surjective $C^\infty$ hyperbolic map of an open subset of the torus. We chose this rather than an Anosov diffeomorphism for better comparison with the Hénon map.

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 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.

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

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

[5] 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.

[6] Mark F Demers. A gentle introduction to anisotropic banach spaces. *Chaos, Solitons & Fractals*, 116:29–42, 2018.

[7] 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.

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

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

[10] 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.

[11] David Ruelle. Differentiation of SRB states. *Communications in Mathematical Physics*, 187(1):227–241, 1997.

[12] David Ruelle. A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, 22(4):855–870, 2009.

[13] David Ruelle. Singularities of the susceptibility of a Sinai–Ruelle–Bowen measure in the presence of stable–unstable tangencies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1935):482–493, 2011.

[14] 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.

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

[16] Simo Särkkä. *Bayesian filtering and smoothing*. Number 3. Cambridge university press, 2013.
[17] Caroline L Wormell. On convergence of linear response formulae in some piecewise hyperbolic maps. *arXiv preprint*, 2022.

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

[19] Lai-Sang Young. What are SRB measures, and which dynamical systems have them? *Journal of statistical physics*, 108(5):733–754, 2002.