STRETCHED-EXPONENTIAL MIXING FOR $\mathscr{C}^{1+\alpha}$ SKEW PRODUCTS WITH DISCONTINUITIES

PEYMAN ESLAMI

ABSTRACT. Consider the skew product $F : \mathbb{T}^2 \to \mathbb{T}^2$, $F(x, y) = (f(x), y + \tau(x))$, where $f : \mathbb{T}^1 \to \mathbb{T}^1$ is a piecewise $\mathscr{C}^{1+\alpha}$ expanding map on a countable partition and $\tau : \mathbb{T}^1 \to \mathbb{R}$ is piecewise $\mathscr{C}^1$. It is shown that if τ is not Lipschitz-cohomologous to a piecewise constant function on the joint partition of f and τ, then F is mixing at a stretched-exponential rate.

1. INTRODUCTION

An important problem in the statistical study of chaotic dynamical systems is obtaining a quantitative estimate on the rate of decay of correlations of the system. Such an estimate describes how fast the system looses memory of its past and opens the door to further statistical description of the system. Ideally, one would like to prove an exponential rate of mixing for systems with "enough" hyperbolicity.

The first results on exponential decay of correlations were obtained for uniformly expanding maps and hyperbolic maps (see [16] and references therein). Slower rates of mixing were also obtained for non-uniformly hyperbolic maps [26, 27, 23].

For flows most of the existing results on exponential decay of correlations pertain to smooth systems or those with a Markov structure (see [12, 17, 3, 1, 25] and references therein). For systems with singularities, Chernov [7] obtained a stretched-exponential rate of decay for certain Billiard flows, Baladi and Liverani [2] for piecewise cone-hyperbolic contact flows, while Obayashi [18] obtained exponential decay of correlations for suspension semiflows over piecewise expanding $\mathscr{C}^2$ maps of the interval using a tower construction and applying the main result of [3].

The goal of this article is to introduce a method by which rates of decay of correlations can be obtained for systems with a neutral direction that have discontinuities and are of low regularity (without assuming the existence of a Markov structure). Such systems appear in practice and are of physical relevance. Indeed, the flow of the Lorenz system of ordinary differential equations (see [5]) and Billiard flows are examples of such systems. Our motivation is to put forward a method to eventually prove exponential mixing rates for these systems; however, in this article we consider the simplest case – that of a 2D skew product with a neutral direction. The skew product $F : \mathbb{T}^2 \to \mathbb{T}^2$ is defined by $F(x, y) = (f(x), y + \tau(x))$, where the *base map* f is an expanding piecewise $\mathscr{C}^{1+\alpha}(\mathbb{T}^1)$ map and the *roof function* τ is piecewise $\mathscr{C}^1(\mathbb{T}^1, \mathbb{R})$. Under very mild conditions on f and τ we prove a stretched-exponential mixing rate for F. At the moment, it is not clear whether stretched-exponential is the optimal decay rate under our assumptions (listed in

---

2000 *Mathematics Subject Classification.* Primary: 37A25; Secondary: 37D50.

Key words and phrases. Decay of Correlations, Skew-Product, Partial Hyperbolicity, Standard Pairs, Oscillatory Cancellation.

It is my pleasure to thank Carlangelo Liverani and Oliver Butterley for their guidance and support in the preparation of this article. Without them this work would not have been possible. I also thank the anonymous referee whose suggestions greatly improved the presentation of this article. Research supported by INdAM-COFUND Marie Curie Fellowship.

Section 2) or whether an exponential bound can be obtained. However, we hope that our method can be refined to obtain exponential mixing rates.

Since the study of the decay rate for the Lorenz flow can be reduced to a suspension semi-flow over an expanding map, our method may be suitable for proving a stretched exponential bound for the Lorenz flow. However, some work is needed to weaken our assumptions that do not immediately apply to the Lorenz case. The main condition that is not satisfied for the Lorenz system is our assumption (2.9). The Lorenz semi-flow corresponds to the 2D skew product F where τ'/f' is not uniformly bounded; it tends to infinity as x tends to 0 hence the expression on the left hand side of (2.9) becomes unbounded at certain points.

The method proposed here is a combination of the point of view of standard pairs due to Dolgopyat [13] and further developed by Chernov [10, 8, 9], and the oscillatory cancellation mechanism due to Dolgopyat [12]. Essentially the proof can be reduced to obtaining one crucial estimate which is a result of oscillatory cancellations. The main estimate of this article is Proposition 6, an exponential version of which was first proven in the smooth setting in [12]. The reduction to this estimate is standard and it is detailed in Section 3, but the proof of this estimate in our setting is new and involves some modifications and reinterpretations of the ideas of Chernov and Dolgopyat.

The complications of proving the main estimate arise from the low regularity of the dynamics and the presence of discontinuities. To overcome these difficulties we introduce the notion of complex standard pairs (Definition 8) which are (densities of) measures supported on small intervals and enjoy some regularity. Indeed, written in polar form, their argument is $\mathscr{C}^1$, while the logarithm of their modulus is only Hölder. Such notion of regularity is forced on us by the $\mathscr{C}^\alpha$ jacobian of the dynamics and the $\mathscr{C}^1$ regularity of τ (also the need to exhibit oscillatory cancellations). The collection of standard pairs form standard families (Definition 9) whose evolution is suitably defined and shown to be invariant under the dynamics (Proposition 14). In this way the evolution of densities under the twisted transfer operator $\mathscr{L}_b^n$ (defined by (2.11)) is modelled by the evolution of standard families, with which it is easier to work due to their regularity. Standard families are especially suitable for studying systems with discontinuities because one can control such discontinuities using *growth lemma*'s (see (4.18), or (4.29)). The connection between the crucial estimate of Proposition 6 and standard families is that the total weight of a standard family under evolution bounds the $\mathbf{L}^1$-norm of the iterates of the twisted transfer operator $\mathscr{L}_b^n$. In this way, we have translated the main estimate into the decay of the total weights of standard families under the dynamics. Such a decay can then be established by suitably modifying standard families, in a dynamical way, to reduce their total weight, without changing the complex measure that they represent (see (6.1) for the notion of equivalent standard families).

The key ingredient in weight reduction of standard families, which makes oscillatory cancellations possible, is the notion of transversality of standard pairs, which is a consequence of uniform non-integrability (due to Chernov [11]) and follows from the assumption that the roof function τ is not Lipschitz-cohomologous to a piecewise constant function on the partition of f. We remark that the weight reduction of standard families must be done with care as to not lose the essential regularity properties of standard families. In this regard we cannot perform certain operations on standard pairs. For example we cannot just add two overlapping standard pairs to form a new standard pair. This is due to the fact that standard pairs are less regular in modulus than in their argument so adding them will result in a less than $\mathscr{C}^1$ regular argument, which is not good for obtaining oscillatory cancellations (for a remedy, see the splitting that appears in (6.7)). In general we must take great care to modify standard families.

One advantage of the *standard pair approach* in this article is that the crucial estimate on the twisted transfer operator $\mathscr{L}_b^n$, for large $|b|$, is obtained without any resort to functional analytic estimates such as the Lasota-Yorke inequality. Such an inequality is used in this paper, but only to obtain estimates for $\mathscr{L}_b^n$ for finitely many values of b. It would be interesting to be able to treat this case also with a more straight forward approach. Another advantage of this method is that it may be useful to treat higher dimensional skew products with discontinuities.

In a separate article [6] with O. Butterley, we show an *exponential* mixing rate for skew products with discontinuities satisfying stronger regularity conditions. More precisely, it is assumed that f is a piecewise $\mathscr{C}^2$ expanding map on a finite partition, and that τ is piecewise $\mathscr{C}^2$ on a finite partition and not Lipschitz cohomologous to a piecewise constant function. The method of the proof is different from the current article (however, it shares some similarities). It relies heavily on the exponential decay of a certain "measure of non-transversality" first introduced by Tsujii [24] in the smooth setting (using ideas of Peres and Solomyak [19]). In [1, 25] it was also shown that such a condition is equivalent to τ not being cohomologous to a constant. In [6] we extended these results to the piecewise $\mathscr{C}^2$ setting showing that if τ is not Lipschitz cohomologous to a piecewise constant function on the joint partition of f and τ then the measure of non-transversality decays exponentially.[1] Finally, using a van der Corput-type oscillatory cancellation estimate we showed exponential mixing. Several comments are in order. Firstly, a van der Corput type estimate requires at least a $\mathscr{C}^{1+\alpha}$ regularity (piecewise) of the phase (the argument of the exponential function appearing in the oscillatory integral) hence that of τ, while in the current setting τ is only assumed to be $\mathscr{C}^1$ (piecewise). Secondly, our Tsujii type estimates of transversality in [6] used extensively the strictly positive lower bound on the invariant density and the upper bound on the derivative of the base map f. In order to obtain exponential mixing in the $\mathscr{C}^{1+\alpha}$ setting on a countable partition, one must overcome the fact that the invariant density may be zero at some points and that $|f'|$ is not uniformly bounded. If these obstacles are overcome, then one can obtain an exponential rate of mixing in the $\mathscr{C}^{1+\alpha}$ setting (for both f and τ) by replacing the cancellation mechanism of this article with a van der Corput estimate to treat transversal standard pairs and treat the non-transversal pairs with the exponentially decaying measure of non-transversality.

The outline of the article is as follows. The class of skew products under study is introduced in Section 2. In Section 3 the main theorem is introduced and proven assuming the crucial estimate of Proposition 6. The rest of the article is devoted to the set up of the machinery of standard pairs and the proof of this estimate. Section 4 introduces the terminology of standard pairs and standard families. Section 5 introduces the notion of transversality. Section 6 shows how one can use the oscillatory cancellation mechanism of Dolgopyat to modify standard families. Finally in Section 7 the main estimate is proven.

2. The setting

Consider the skew-product $F : \mathbb{T}^2 \to \mathbb{T}^2$ defined by

$$F : (x, y) \mapsto (f(x), y + \tau(x)) . \tag{2.1}$$

Assume that

$$f : \mathbb{T}^1 \to \mathbb{T}^1 \text{ is piecewise } \mathscr{C}^{1+\alpha}. \tag{2.2}$$

[1]The current article uses a more primitive version of transversality that was introduced by Chernov [11] and subsequently used by Dolgopyat [12], and others [21, 3, 1].

That is, $\mathbb{T}^1$ can be partitioned into countably many open intervals (modulo the countable set of endpoints of the intervals) such that each open interval is an interval of monotonicity of f and f is $\mathscr{C}^{1+\alpha}$ on the open interval, extendable to the closed interval. The collection of such intervals is referred to as the *partition of* f. Note that, for every $n \geq 1$, f^n is also piecewise $\mathscr{C}^{1+\alpha}$. Having the graph of f^n in mind, we denote by $\mathcal{H}^n$ the set of inverse branches of f^n. We choose to index the elements of the partition of f^n by elements of $\mathcal{H}^n$. So, we denote the partition of f^n by $\{O_h\}_{h\in\mathcal{H}^n}$. Note that the domain and range of h are $f^n(O_h)$ and O_h, respectively.

Assume that f is expanding. That is, there exist λ such that

$$\ln 2 < \lambda \text{ and } e^{\lambda n} \leq |(f^n)'| . \tag{2.3}$$

The assumption that $\lambda > \ln 2$ is not crucial; it only makes the calculations more convenient. It is enough to assume $\lambda > 0$. Note that if $|f'| > 1$, then there exists $n_0 \in \mathbb{N}$ such that $|(f^{n_0})'| > 2$.

Assume that f satisfies the following distortion bound.[2] There exists a constant $D \geq 0$ such that for every $n \in \mathbb{N}$

$$\frac{|h'(x)|}{|h'(y)|} \leq e^{D|x-y|^\alpha}, \text{ for } h \in \mathcal{H}^n, \text{ and } x, y \in f^n(O_h). \tag{2.4}$$

Remark 1. *Condition* (2.4) *is implied by a similar condition for the first iterate of f with slightly worse constant D.*

Proof. Suppose $h \in \mathcal{H}^n$ and $x, y \in f^n(O_h)$, then

$$\log |h'(x)| = \sum_{j=1}^{n} \log \left|h_j'\right| \circ h_{j-1} \circ \cdots \circ h_1(x), \text{ for } h_1, \ldots, h_n \in \mathcal{H}^1.$$

Suppose that (2.4) holds for $h \in \mathcal{H}^1$ with constants D and α. Let $\tilde{D} = D/(1-e^{-\lambda\alpha})$. Since $h_{j-1} \circ \cdots \circ h_1(x) \in O_{h_{j-1}}$ with $h_{j-1} \in \mathcal{H}^1$, and using the contraction of the inverse branches, we have

$$\begin{aligned} |\log |h'(x)| - \log |h'(y)|| &\leq \sum_{j=1}^{n} D \left|h_{j-1} \circ \cdots \circ h_1(x) - h_{j-1} \circ \cdots \circ h_1(y)\right|^\alpha \\ &\leq D \sum_{j=1}^{n} e^{-\lambda\alpha(j-1)} |x-y|^\alpha \\ &\leq \tilde{D} |x-y|^\alpha . \end{aligned}$$

□

Remark 2. *Condition* (2.4) *for $n = 1$ implies that there exists a uniform constant, which we denote by $|1/f'|_\alpha$, such that for every $h \in \mathcal{H}^1$ and every $x, y \in O_h$ holds $|1/|f'|(x) - 1/|f'|(y)| \leq |1/f'|_\alpha |x-y|^\alpha$.*

[2] For a particular system if this condition is not satisfied for the maximum partition of monotonicity of f, then it may be possible to consider a countable refinement of this partition for which the condition is satisfied. For the Lorenz map, this can be done by refining the partition of f near the singularity by intersection with the polynomial parition $\{(\pm(k+1)^{-p}, \pm k^{-p})\}_{k\in\mathbb{N}}$, where $p \in \mathbb{N}$ and $p \geq 2$.

Proof. Suppose $h \in \mathcal{H}^1$ and $x, y \in O_h$, then

$$\begin{aligned} \left| \frac{1}{|f'(x)|} - \frac{1}{|f'(y)|} \right| &\leq \frac{1}{|f'(x)|} \left| 1 - \frac{f'(x)}{f'(y)} \right| \\ &\stackrel{(2.4)}{\leq} \frac{1}{|f'(x)|} \left| e^{D|f(x)-f(y)|^\alpha} - 1 \right| \\ &\leq De^D \frac{1}{|f'(x)|} |f(x) - f(y)|^\alpha . \end{aligned}$$

Note that (2.4) implies $1/|f'(x)| \leq e^D |x-y| / |f(x) - f(y)|$, for every $x, y \in O_h$ (the simple proof follows from Lemma 10; see (4.9)). Since $1/|f'(x)| \leq (1/|f'(x)|)^\alpha$, it follows that $1/|f'(x)| |f(x) - f(y)|^\alpha \leq e^{D\alpha} |x-y|^\alpha$. Therefore,

$$\left| \frac{1}{|f'(x)|} - \frac{1}{|f'(y)|} \right| \leq De^{D(1+\alpha)} |x-y|^\alpha , \text{ for every } x, y \in O_h.$$

□

Also, assume that, for every $n \in \mathbb{N}$,

$$\sum_{h \in \mathcal{H}^n} \sup_{f^n(O_h)} |h'| < \infty. \tag{2.5}$$

This condition is trivially satisfied when f has finitely many branches. This condition is also implied by the similar condition for the first iterate. Indeed, if $\sum_{h \in \mathcal{H}} \sup_{f(O_h)} |h'| \leq M < \infty$, then $\sum_{h \in \mathcal{H}^n} \sup_{f^n(O_h)} |h'| \leq M^n < \infty$.

Assume also that f satisfies the following covering condition.

$$\text{For every open interval } J \subset \mathbb{T}^1 \text{ there exists } N \in \mathbb{N}, \text{ such that } f^N J = \mathbb{T}^1. \tag{2.6}$$

This condition is sometimes called *locally eventually onto* and it is weaker than the covering condition appearing in [16]. Note that by compactness we can assume N depends only on the length of the interval J and not on its position.[3]

Assume that

$$\tau : \mathbb{T}^1 \to \mathbb{R} \text{ is piecewise } \mathscr{C}^1 \text{ with respect to the partition of } f. \tag{2.7}$$

Assume that τ is not Lipschitz-cohomologous to a piecewise constant function. That is,

$$\nexists \phi \in Lip(\mathbb{T}^1, \mathbb{R}) \text{ such that } \tau(x) = \phi \circ f(x) - \phi(x) + \psi(x), \tag{2.8}$$

where $\psi(x)$ is a piecewise constant function on the partition of f. Such a condition can sometimes be checked by considering the values of τ on periodic orbits of f; see Example 1.

For every $n \in \mathbb{N}$, define $\tau_n := \sum_{j=0}^{n-1} \tau \circ f^j$. Assume that there exists C_τ such that for every $n \in \mathbb{N}$

$$|(\tau_n \circ h)'| \leq C_\tau, \text{ for every } h \in \mathcal{H}^n. \tag{2.9}$$

This condition is easily satisfied if $|\tau'|$ is bounded. Note that in more general cases mentioned earlier, e.g. the case of the Lorenz flow, this condition is not satisfied (see for example [5]). In this article, condition (2.9) is only used in the proof of Proposition 14 to show invariance of the standard family.[4]

---

[3] Following the proofs closely, it follows that for a specific f there will be some $\delta_0 > 0$ such that condition (2.6) need only be checked for open intervals J of length $|J| > \delta_0$.

[4] One may generalize condition (2.9) to allow for some growth of $|(\tau_n \circ h)'|$, by requiring that there exist constants $0 \leq s < 1$ and C_τ such that $|(\tau_n \circ h)'| |f^n(O_h)|^s \leq C_\tau, \ \forall x \in f^n(O_h)$. Accordingly, one would have to modify also the definition of standard pairs and check for invariance. This can be done by modifying (4.5) to $|\arg(\rho)'| \leq a|b| |I|^{-s}, \ 0 \leq s < 1$.

Let $\mathscr{L} : \mathbf{L}^1(\mathbb{T}^2) \to \mathbf{L}^1(\mathbb{T}^2)$ be the transfer operator associated to the skew product F. That is,

$$\mathscr{L}g(x,y) = \sum_{z\in F^{-1}(x,y)} g(z)|\det DF^{-1}(z)| = \sum_{w\in f^{-1}(x)} \frac{1}{|f'(w)|}g(w, y-\tau(w)). \quad (2.10)$$

The twisted (or weighted) transfer operator is defined by

$$\mathscr{L}_b^n g = \sum_{h\in\mathcal{H}^n} e^{ib\tau_n\circ h} \cdot g\circ h \cdot |h'| \cdot \mathbb{1}_{O_h} \circ h. \quad (2.11)$$

Banach space assumptions. Let $\|\cdot\|_{\mathscr{C}^\alpha} = |\cdot|_\alpha + \|\cdot\|_{\mathscr{C}^0}$, where $|\cdot|_\alpha$ is the usual Hölder semi-norm. Suppose there exists a Banach space $\mathbf{B} \subset \mathbf{L}^1$ such that the following hold.

(1) For every $g \in \mathbf{B}$, $\|g\|_{\mathbf{L}^1} \le \|g\|_{\mathbf{B}}$. For every g in $\mathscr{C}^\alpha$, $\|g\|_{\mathbf{B}} \le \|g\|_{\mathscr{C}^\alpha}$.
(2) For every $b \in \mathbb{Z}$, the weighted transfer operator $\mathscr{L}_b : \mathbf{B} \to \mathbf{B}$ associated to f (see (2.11)) with weight $\xi(x) = e^{ib\tau(x)}$ is bounded, has a spectral radius ≤ 1 and has essential spectral radius strictly < 1.
(3) For every $\beta > 0$ there exists a Banach space $\mathbf{B}_\beta$ containing $\mathbf{B}$ such that if $g \in \mathbf{B}$ and $u \in \mathscr{C}^\beta$, then $ug \in \mathbf{B}_\beta$ and it is possible to approximate $ug \in \mathbf{B}_\beta$ with $g_\varepsilon \in \mathscr{C}^3$ such that $\|ug - g_\varepsilon\|_{\mathbf{L}^1} \le \|ug\|_{\mathbf{B}_\beta} \varepsilon^\beta$ and $\|g_\varepsilon\|_{\mathscr{C}^3} \le C\|ug\|_{\mathbf{B}_\beta} \varepsilon^{-3}$.

Remark 3. *Under the assumptions* (2.2)–(2.5) *the Banach space of functions of generalized bounded variation satisfies the assumptions (1)–(3) above.*

Proof. As in [5], the space of functions of generalized bounded variation is defined as follows.

$$\mathbf{B}_\alpha := \{g \in \mathbf{L}^1(\mathbb{T}^1) : |g|_{\mathbf{B}_\alpha} < \infty\},$$

where

$$|g|_{\mathbf{B}_\alpha} := \sup_{\epsilon\in(0,\epsilon_0)} \epsilon^{-\alpha} \int_{\mathbb{T}^1} \mathrm{osc}\,[g, B_\epsilon(x)]dx, \text{ for every } \alpha \in (0,1), \text{ some fixed } \epsilon_0; \text{ and,}$$

$$\mathrm{osc}[g, J] := \operatorname{ess\,sup}\{|g(x) - g(y)| : x, y \in J\}, \text{ for every interval } J \subset \mathbb{T}^1.$$

Note that for a α-Hölder function g, $|g|_{\mathbf{B}_\alpha} \le |g|_\alpha$. It is known [15, 5] that $\mathbf{B}_\alpha$ with the norm $\|g\|_{\mathbf{B}_\alpha} := |g|_{\mathbf{B}_\alpha} + \|g\|_{\mathbf{L}^1(\mathbb{T}^1)}$ is a Banach space. Moreover, $\mathbf{B}_\alpha \subset \mathbf{L}^\infty(\mathbb{T}^1) \subset \mathbf{L}^1(\mathbb{T}^1)$ and the embedding of $\mathbf{B}_\alpha$ into $\mathbf{L}^1$ is compact.

That assumption (1) is satisfied is a consequence of the definitions.

Assumption (2) follows from [5, Section 4.2] if we show that $|(\xi/f')(x) - (\xi/f')(y)| \le C_\xi |x-y|^\alpha$, for x, y in the same partition element of f. Indeed, this follows from our assumptions. Suppose $h \in \mathcal{H}^1$ and $x, y \in O_h$. We have,

$$\left|\frac{e^{ib\tau(x)}}{f'(x)} - \frac{e^{ib\tau(y)}}{f'(y)}\right| = \left|e^{ib\tau(x)}\left(\frac{1}{f'(x)} - \frac{1}{f'(y)}\right) + \frac{1}{f'(y)}\left(e^{ib\tau(x)} - e^{ib\tau(y)}\right)\right|. \quad (2.12)$$

Using the triangle inequality, the right hand side satisfies

$$RHS \le \left|\frac{1}{f'(x)} - \frac{1}{f'(y)}\right| + \frac{1}{|f'(y)|}\left|e^{ib\tau(x)} - e^{ib\tau(y)}\right|.$$

The first term is bounded by $|1/f'|_\alpha |x-y|^\alpha$, while the second term is estimated as follows using (2.9) and (2.4):

$$\begin{aligned} \frac{1}{|f'(y)|}\left|e^{ib\tau(x)} - e^{ib\tau(y)}\right| &\leq \frac{1}{|f'(y)|}\left|\int_x^y \left(e^{ib\tau(s)}\right)' ds\right| \\ &\leq |b| \int_x^y \frac{|\tau'(s)|}{|f'(s)|}\frac{|f'(s)|}{|f'(y)|} ds \\ &\overset{(2.9)}{\leq} |b|\, C_\tau\, |x-y| \sup_{s\in O_h} \frac{|f'(s)|}{|f'(y)|} \\ &\overset{(2.4)}{\leq} |b|\, C_\tau\, |x-y|\, e^D. \end{aligned}$$

Therefore, (2.12) is bounded by

$$|1/f'|_\alpha |x-y|^\alpha + |b|\, C_\tau e^D |x-y| \leq C_{D,\alpha} |b|\, |x-y|^\alpha .$$

We have proved our claim with $C_\xi = C_{D,\alpha}|b|$.

Assumption (3) is checked as follows. For $\beta > 0$ let $\mathbf{B}_\beta$ be the Banach space defined above with parameter $\beta \leq \alpha$. Suppose $u \in \mathscr{C}^\beta$, then since $g \in \mathbf{B}_\alpha \subset \mathbf{B}_\beta$, a simple calculation shows that $ug \in \mathbf{B}_\beta$. Let $g_\epsilon := \epsilon^{-1}\int \varphi(\epsilon^{-1}(x-y))g(y)dy$, where φ is a positive mollifier. We have

$$\begin{aligned} g_\epsilon(x) - g(x) &= \epsilon^{-1}\int \varphi(\epsilon^{-1}(x-y))(g(y)-g(x))dy \\ &\leq \epsilon^{-1}\int \varphi(\epsilon^{-1}(x-y))\,\mathrm{osc}\,[g, B_\epsilon(x)]dy \\ &\leq \mathrm{osc}[g, B_\epsilon(x)]. \end{aligned}$$

Therefore, applying the definition of the norms, $\|g-g_\epsilon\|_{\mathbf{L}^1} \leq \epsilon^\beta |g|_{\mathbf{B}_\beta} \leq \epsilon^\beta \|g\|_{\mathbf{B}_\beta}$. For estimating $\|g_\epsilon\|_{\mathscr{C}^3}$, differentiate g_ϵ three times to obtain that

$$\|g\|_{\mathscr{C}^3} \leq C\, \|ug\|_{\mathbf{L}^\infty} \epsilon^{-3} \leq \|ug\|_{\mathbf{B}_\beta} \epsilon^{-3}.$$

□

In a standard manner (see e.g. [14] or [4, Section 7]), our Banach space assumptions imply the quasi-compactness of $\mathscr{L}_0$ on $\mathbf{B}$. Moreover, the covering condition (2.6) implies that f preserves a unique absolutely continuous measure μ with a density $\wp \in \mathbf{B}$ and (f,μ) is mixing. It can be easily shown that F preserves the absolutely continuous measure $\nu = \mu \times m$, where m is the Lebesgue measure. We will also denote the density of ν by $\wp$ since it is constant in the vertical direction. The main result of this note, Theorem 4, is to prove a stretched-exponential decay of correlations for the skew product F. Such an estimate implies (F,ν) is mixing.

Let us give one example of a skew product that fits into our setting.

Example 1. *Let $\beta \in (0,1)$. Define $f : \mathbb{T}^1 \to \mathbb{T}^1$ by $f(x) = (5x)^\beta$ for $x \in (0, 1/5)$ and by $5x \mod 1$ otherwise. Suppose $\bar\tau : \mathbb{T}^1 \to \mathbb{R}$ is a constant function. Then there exists a perturbation of $\bar\tau$ which satisfies condition* (2.8)*. We take τ to be this perturbation of $\bar\tau$. The skew product F given by the pair (f,τ) satisfies* (2.1)*–*(2.9)*.*

Proof. Denote by $\mathcal{P} = \{(j/5,(j+1)/5)\}_{j=0}^4$ the maximal partition of monotonicity of f. Note that due to the unboundedness of f' at 0, we cannot hope for f to satisfy our assumptions with respect to the partition $\mathcal{P}$. However we can consider f on a refined partition $\mathcal{P}' = \{((k+1)^{-2}, k^{-2})\}_{k=5}^\infty \cup \{(1/25, 1/5)\} \cup \{(j/5,(j+1)/5)\}_{k=1}^4$. Note that every element of $\mathcal{P}$ contains a unique fixed point of f. Denote by p_1 the fixed point of f which lies in $(1/5, 2/5)$. Moreover, there exists a periodic point q_1 of minimal period 2 such that $q_1 \in (1/5, 2/5)$ and $f(q_1) \in (4/5, 1)$. Without loss of generality let us assume $\bar\tau$ is constant and equal to 1. Define τ_ε to be equal to 1 except on the interval $(q_1-\varepsilon, q_1+\varepsilon)$ on which it takes the value $|x-q_1|+1-\varepsilon$. We claim that there exists $\varepsilon_0>0$ such that τ_{ε_0} is not Lipschitz cohomologous to a piecewise constant function on the partition $\mathcal{P}'$. Indeed, suppose $\tau=\tau_\varepsilon$ were cohomologous to a piecewise constant (on $\mathcal{P}'$) function $\psi=\psi_\varepsilon$, then $\tau(q_1)+\tau\circ f(q_1)=\psi(q_1)+\psi\circ f(q_1)$. The function ψ being piecewise constant implies that $\psi(q_1)+\psi\circ f(q_1)=\psi_1(p_1)+\psi_1(p_4)$, where p_4 is the unique fixed point of f inside the partition element $(4/5,1)$ which also contains $f(q_1)$. Since τ is cohomologous to ψ, and τ is equal to 1 outside $(q_1-\varepsilon,q_1+\varepsilon)$, if ε is small enough that $p_1,p_4\notin(q_1-\varepsilon,q_1+\varepsilon)$, then $\psi_1(p_1)=\tau_1(p_1)=1$ and $\psi_1(p_4)=\tau_1(p_4)=1$. We have shown that $\tau(q_1)+\tau\circ f(q_1)=2$, but this is a contradiction because $\tau(f(q_1))=1$, but $\tau(q_1)=1-\varepsilon$. Therefore τ_1 is not cohomologous to a piecewise constant.

Let us check condition (2.4). Note that there is a bijective relation between $\mathcal{H}^1$ and $\mathcal{P}'$ and checking condition (2.4) is only nontrivial for the inverse branches $h\in\mathcal{H}^1$ that correspond to the partition elements $\{((k+1)^{-2},k^{-2})\}_{k=5}^{\infty}$. Suppose $h\in\mathcal{H}^1$ is an inverse branch with range $((k+1)^{-2},k^{-2})$. Note that $h(u)=(1/5)u^{1/\beta}$, for $u\in\mathbb{T}^1$. Suppose $u,v\in\mathbb{T}^1$ with $x=h(u)$, $y=h(v)$, then $|h'(u)/h'(v)|=|x/y|^{(1/\beta)-1}$. We have $|x/y|\le|x-y|/|y|+1\le e^{|x-y|/|y|}=e^{|x-y|^{1/3}\cdot|x-y|^{2/3}/|y|}$. Therefore, $|x/y|^{(1/\beta)-1}\le e^{((1/\beta)-1)|x-y|^{1/3}\cdot|x-y|^{2/3}/|y|}$. Note that $|y|\ge(k+1)^{-2}$, and $|x-y|\le 2k^{-3}$. Therefore, $|x-y|^{2/3}/|y|\le 2^{2/3}k^{-2}/(k+1)^{-2}\le 2^{2+2/3}$. Therefore, $|h'(u)/h'(v)|=|x/y|^{(1/\beta)-1}\le e^{D|x-y|^\alpha}\le e^{D|u-v|^\alpha}$, where $D=2^{2+2/3}((1/\beta)-1)$ and $\alpha=1/3$.

Condition (2.5) simply follows from the fact that $\sum_k k^{-2}<\infty$. The other conditions are easy to check. □

3. Decay of correlations

For two observables ϕ and ψ the correlation coefficients with respect to the invariant measure $\nu=\wp dm$ are defined by

$$\mathrm{cor}_{\phi,\psi}(n)=\int_{\mathbb{T}^2}\phi\cdot\psi\circ F^n\,d\nu-\int_{\mathbb{T}^2}\phi\,d\nu\int_{\mathbb{T}^2}\psi\,d\nu.$$

Theorem 4. *Suppose the skew product F satisfies assumptions (2.1)–(2.9). Then, for every $\beta>0$ there exists $\gamma_3>0$ such that for every $\phi\in\mathscr{C}^\beta(\mathbb{T}^2)$ with $\int\phi d\nu=0$ and $\psi\in\mathbf{L}^\infty(\mathbb{T}^2)$,*

$$|\mathrm{cor}_{\phi,\psi}(n)|\le Ce^{-\gamma_3\sqrt{n}}\left\|\phi\wp\right\|_{\mathbf{B}_\beta}\left\|\psi\right\|_{\mathbf{L}^\infty}. \tag{3.1}$$

Remark 5. *It suffices to estimate $\left|\int_{\mathbb{T}^2}\phi\cdot\psi\circ F^n\,dm\right|$, where m is the 2D Lebesgue measure, $\phi\in\mathscr{C}^3$, $\int\phi d\nu=0$ and $\psi\in\mathbf{L}^\infty$.*

Proof. Suppose for the moment that there exists $\gamma>0$ and $C>0$ such that for every $\phi\in\mathscr{C}^3(\mathbb{T}^2)$, $\int\phi d\nu=0$ and $\psi\in\mathbf{L}^\infty(\mathbb{T}^2)$, holds $\left|\int_{\mathbb{T}^2}\phi\cdot\psi\circ F^n\,dm\right|\le Ce^{-\gamma\sqrt{n}}\left\|\phi\right\|_{\mathscr{C}^3}\left\|\psi\right\|_{\mathbf{L}^\infty}$. Suppose $\nu=\wp dm$ with $\wp\in\mathbf{B}=\mathbf{B}_\alpha$ and $\phi\in\mathscr{C}^\beta$, $0<\beta<\alpha$. By condition (3) of the Banach space assumptions, $\phi\wp\in\mathbf{B}_\beta$ and there exists $\wp_\varepsilon\in\mathscr{C}^3$ such that $\left\|\phi\wp-\wp_\varepsilon\right\|_{\mathbf{L}^1}\le\left\|\phi\wp\right\|_{\mathbf{B}_\beta}\varepsilon^\beta$ and $\left\|\wp_\varepsilon\right\|_{\mathscr{C}^\beta}<C\left\|\phi\wp\right\|_{\mathbf{B}_\beta}\varepsilon^{-3}$. Therefore,

$$\begin{aligned}\left|\int_{\mathbb{T}^2}\phi\cdot\psi\circ F^n\,d\nu\right| &\le \left|\int_{\mathbb{T}^2}\psi\circ F^n\wp_\varepsilon\,dm\right|+\left|\int_{\mathbb{T}^2}\psi\circ F^n(\phi\wp-\wp_\varepsilon)\,dm\right|\\ &\le Ce^{-\gamma\sqrt{n}}\left\|\wp_\varepsilon\right\|_{\mathscr{C}^\beta}\left\|\psi\right\|_{\mathbf{L}^\infty}+\left\|\wp-\wp_\varepsilon\right\|_{\mathbf{L}^1}\left\|\psi\right\|_{\mathbf{L}^\infty}\\ &\le C\left\|\phi\wp\right\|_{\mathbf{B}_\beta}\left\|\psi\right\|_{\mathbf{L}^\infty}\left(e^{-\gamma\sqrt{n}}\varepsilon^{-3}+\varepsilon^\beta\right)\\ &\le C\left\|\phi\wp\right\|_{\mathbf{B}_\beta}\left\|\psi\right\|_{\mathbf{L}^\infty}e^{\frac{-\gamma\beta\sqrt{n}}{3+\beta}},\end{aligned}$$

where we have chosen $\varepsilon = e^{-\gamma(3+\beta)^{-1}\sqrt{n}}$. □

proof of Theorem 4. Using Parseval's Theorem and Fourier transformation in the y-direction we may write

$$\left|\int \phi \cdot \psi \circ F^n \; dm\right| = \left|\sum_{b\in\mathbb{Z}} \int \mathscr{L}_b^n(\hat{\phi}_b) \cdot \hat{\psi}_{-b} \; dm\right|,$$

where $\{\hat{\phi}_b = \int_{\mathbb{T}^1} e^{-iby}\phi(x,y)dy\}_{b\in\mathbb{Z}}$ are the Fourier coefficients of ϕ in the y-direction, and $\mathscr{L}_b^n$ is the twisted transfer operator as defined in (2.11). Noting that the $\mathbf{B}$-norm is stronger than the $\mathbf{L}^1$-norm (by Banach space assumption (1)),

$$\begin{aligned}\left|\sum_{b\in\mathbb{Z}} \int \mathscr{L}_b^n(\hat{\phi}_b) \cdot \hat{\psi}_{-b} \; dm\right| &\leq \sum_{b\in\mathbb{Z}} \left\|\mathscr{L}_b^n(\hat{\phi}_b)\right\|_{\mathbf{L}^1} \left\|\hat{\psi}_{-b}\right\|_{\mathbf{L}^\infty} \\ &\leq \sum_{|b|<b_0} \|\mathscr{L}_b^n\|_{\mathbf{B}} \left\|\hat{\phi}_b\right\|_{\mathbf{B}} \left\|\hat{\psi}_{-b}\right\|_{\mathbf{L}^\infty} \\ &+ \sum_{|b|\geq b_0} \|\mathscr{L}_b^n\|_{\mathscr{C}^1\to\mathbf{L}^1} \left\|\hat{\phi}_b\right\|_{\mathscr{C}^1} \left\|\hat{\psi}_{-b}\right\|_{\mathbf{L}^\infty}.\end{aligned} \tag{3.2}$$

We estimate the second sum above using the following.

Proposition 6. *Suppose assumptions* (2.1)–(2.9) *are satisfied. There exists* $\gamma_2 > 0$, $C > 0$ *and* $b_0 \in \mathbb{N}$ *given by* (6.6), *such that for every* $|b| \geq b_0$ *and for every* $n \in \mathbb{N}$,

$$\|\mathscr{L}_b^n\|_{\mathscr{C}^1\to\mathbf{L}^1} \leq Ce^{-\frac{\gamma_2}{\ln|b|}n}. \tag{3.3}$$

This is the main estimate of the article and the rest of the article is devoted to its proof. Assume that this statement holds and let us finish the proof.

Using the regularity of ϕ, ψ, there exist a constant C such that for every $b \in \mathbb{Z}$, $b \neq 0$,

$$\left\|\hat{\phi}_b\right\|_{\mathscr{C}^1} \leq C\|\phi\|_{\mathscr{C}^3}|b|^{-2}, \left\|\hat{\psi}_{-b}\right\|_{\mathbf{L}^\infty} \leq \|\psi\|_{\mathbf{L}^\infty}. \tag{3.4}$$

Using the estimate of Proposition 6,

$$\sum_{|b|\geq b_0} \|\mathscr{L}_b^n\|_{\mathscr{C}^1\to\mathbf{L}^1} \left\|\hat{\phi}_b\right\|_{\mathscr{C}^1} \left\|\hat{\psi}_{-b}\right\|_{\mathbf{L}^\infty} \leq \sum_{|b|\geq b_0} C\|\phi\|_{\mathscr{C}^3}\|\psi\|_{\mathbf{L}^\infty} e^{-\frac{\gamma_2 n}{\ln|b|}}|b|^{-2}. \tag{3.5}$$

For estimating the sum over $|b| < b_0$ in (3.2), first note that the term corresponding to $b = 0$ decays exponentially due to the exponential decay of correlations for the base map f and due to $\int \phi d\nu$ being equal to zero. For $1 \leq |b| < b_0$, we use the following result, which is proven in Section 7. Note that the constants C and r below depend on b and that is why for large $|b|$ we need a different argument.

Proposition 7. *Suppose assumptions* (2.1)–(2.9) *are satisfied. For all* $b \neq 0$, *there exists* C *and* $r > 0$, *both depending on* b, *such that for every* $n \in \mathbb{N}$,

$$\|\mathscr{L}_b^n\|_{\mathbf{B}} \leq Ce^{-rn}. \tag{3.6}$$

Using the estimate of Proposition 7 for $1 \leq |b| < b_0$, the exponential mixing of f for $b = 0$, and the estimate of Proposition 6 for $|b| > b_0$, it follows that

$$|\mathrm{cor}_{\phi,\psi}(n)| \leq \sum_{|b|\geq b_0} C_{b_0}C\|\phi\|_{\mathscr{C}^3}\|\psi\|_{\mathbf{L}^\infty} e^{-\frac{\gamma_2 n}{\ln|b|}}|b|^{-2}, \text{ for all } n \in \mathbb{N}.$$

Estimating the above sum yields a stretched-exponential decay. Indeed, one way to estimate $\sum_{|b|\geq b_0} e^{-\frac{\gamma_2 n}{\ln|b|}}|b|^{-2}$ is to split the sum into two parts $|b| \leq L$ and $|b| \geq L+1$ to get

$$\sum_{|b|\geq b_0} e^{-\frac{\gamma_2 n}{\ln|b|}}|b|^{-2} \leq Le^{-\frac{\gamma_2 n}{\ln|L|}} + L^{-1}.$$

Now choose L so that the two parts of the sum are equal. The solution is $L = e^{\sqrt{\frac{\gamma_2 n}{2}}}$, and gives

$$Le^{-\frac{\gamma_2 n}{\ln |L|}} + L^{-1} = 2L^{-1} \leq 2e^{-\sqrt{\frac{\gamma_2 n}{2}}}.$$

Therefore,

$$|\mathrm{cor}_{\phi,\psi}(n)| \leq 2C_{b_0} C \|\phi\|_{\mathscr{C}^3} \|\psi\|_{\mathbf{L}^\infty} e^{-\sqrt{\frac{\gamma_2}{2}}\sqrt{n}}.$$

□

4. Iteration of Standard Families

In this section we introduce standard families and their dynamics. We are essentially modelling the evolution of densities under $\mathscr{L}_b^n$ with the iteration of standard families. Note that on one hand we need the invariance of standard families under the $\mathscr{C}^{1+\alpha}$ dynamics, i.e. a $\mathscr{C}^\alpha$ jacobian; on the other hand, we need at least $\mathscr{C}^1$ regularity in order to take advantage of the oscillatory cancellation mechanism due to Dolgopyat. Our remedy for this issue is to introduce complex standard pairs with a notion of regularity where the modulus of the standard pair, written in polar form, is less regular than its argument. We also use (2.4) in an essential way to ensure the invariance of (4.6), via Equation (4.8).

For $\alpha \in (0,1)$, and a function $\rho : I \to \mathbb{C}$ define

$$H(\rho) = \sup_{x,y \in I} \frac{|\ln|\rho(x)| - \ln|\rho(y)||}{|x-y|^\alpha}. \tag{4.1}$$

Let $\arg(\rho) \in [0, 2\pi)$ be the argument of ρ written in polar form. All integrals where the measure is not indicated are with respect to the Lebesgue measure. For any measurable set A, $|A|$ denotes the Lebesgue measure of A.

Definition 8 (Standard pair)**.** *A* standard pair *with associated parameters $a > 0, b \in \mathbb{Z}, \varepsilon_0 > 0$ is a pair (I, ρ) consisting of an open interval I and a function $\rho \in \mathbf{L}^1(I, \mathbb{C})$ such that*

$$|I| < \varepsilon_0 \leq 1; \tag{4.2}$$

$$\int_I |\rho| = 1; \tag{4.3}$$

$$H(\rho) \leq a; \tag{4.4}$$

$$|\arg(\rho)'| \leq a|b|. \tag{4.5}$$

Definition 9 (Standard family)**.** *A* standard family *$\mathcal{G}$ is a set of standard pairs $\{(I_j, \rho_j)\}_{j\in\mathcal{J}}$ and an associated measure $w_\mathcal{G}$ on a countable set $\mathcal{J}$. We require that there exists a constant $B > 0$ such that,*

$$|\partial_\varepsilon \mathcal{G}| := \sum_{j\in\mathcal{J}} w_\mathcal{G}(j) \int_{\partial_\varepsilon I_j} |\rho_j| \leq B\varepsilon, \text{ for all } \varepsilon < \varepsilon_0, \tag{4.6}$$

where $\partial_\varepsilon I_j$ denotes the set of all points in I_j that have distance less than ε to the endpoints of I_j. If $w_\mathcal{G}$ is a probability measure, then $\mathcal{G}$ is called a standard probability family*. Each standard family induces an absolutely continuous (complex) measure on $\mathbb{T}^1$ with the density*[5]*:*

$$\rho_\mathcal{G} = \sum_{j\in\mathcal{J}} w_\mathcal{G}(j)\rho_j. \tag{4.7}$$

The total weight of a standard family is denoted $|\mathcal{G}| := \sum_{j\in\mathcal{J}} w_\mathcal{G}(j)$. The set of standard families with associated parameters a, b, B, ε_0 is denoted $\mathcal{M}_{a,b,B,\varepsilon_0}$.

[5]By the sum (4.7) we really mean the sum of the trivially extended standard pairs to densities defined on all of $\mathbb{T}^1$, that is we set them equal to zero outside their domain.

The following lemma (Lemma 10) and its consequences will be useful to preserve (4.6). Before we state the lemma we introduce some notation. Suppose $a > 0$. For positive quantities A, B, we shall write $A \asymp_a B$ and say that A and B are a-comparable if $e^{-a}A \le B \le e^a A$. Note the following simple facts.

(1) If $A \asymp_a B$ and $a' > a$, then $A \asymp_{a'} B$.
(2) If $A \asymp_a B$, then $B \asymp_a A$.
(3) If $A \asymp_a B$ and $A \le C \le B$, then $A \asymp_a C \asymp_a B$. That is, A, C, and B are pairwise a-comparable. It follows that all values between A and B are pairwise a-comparable.
(4) If $A \asymp_a B$ and $B \asymp_{a'} C$, then $A \asymp_{a+a'} C$. Therefore, $A \asymp_{a+a'} B \asymp_{a+a'} C$.
(5) If $A_1 \asymp_a B_1$ and $A_2 \asymp_{a'} B_2$, then $A_1 + A_2 \asymp_{\max\{a,a'\}} B_1 + B_2$.
(6) If $A_1 \asymp_a B_1$ and $A_2 \asymp_{a'} B_2$, then $A_1 A_2 \asymp_{a+a'} B_1 B_2$.

Lemma 10. *If (I, ρ) satisfies (4.4), then for every $J, J' \subset I$ with $|J|\,|J'| \neq 0$,*

$$\inf_I |\rho| \asymp_a \mathcal{A}_J |\rho| \asymp_a \mathcal{A}_{J'} |\rho| \asymp_a \sup_I |\rho| , \tag{4.8}$$

where $\mathcal{A}_J |\rho| = |J|^{-1} \int_J |\rho|$ is the average of $|\rho|$ on J.

Proof. Note that (4.4) implies that for every $x, y \in I$, $e^{-a|x-y|^\alpha} |\rho|(y) \le |\rho|(x) \le e^{a|x-y|^\alpha} |\rho|(y)$. This implies $\inf_I |\rho| \asymp_a \sup_I |\rho|$. For the rest, observe that for every $J \subset I$, $\inf_I |\rho| \le \inf_J |\rho| \le \mathcal{A}_J |\rho| \le \sup_J |\rho| \le \sup_I |\rho|$; therefore, lying between a-comparable quantities, the averages are also a-comparable. □

Note that since (2.4) implies $H(h') \le D$, we may apply Lemma 10 to $(f^n(O_h), h')$. It follows that

$$\sup_{f^n(O_h)} |h'| \asymp_D \mathcal{A}_{f^n(O_h)} |h'| = \frac{|O_h|}{|f^n(O_h)|}, \text{ for every } n \in \mathbb{N}, h \in \mathcal{H}^n. \tag{4.9}$$

For the last equality we have used that f^n is one-to-one on O_h.

Given a standard family $\mathcal{G} \in \mathcal{M}_{a,b,B,\varepsilon_0}$, we define its n-th iterate as follows.

Definition 11 (Iteration). *Let $\mathcal{G}$ be a standard family with index set $\mathcal{J}$ and weight $w_{\mathcal{G}}$. For $(j, h) \in \mathcal{J} \times \mathcal{H}^n$ such that $|f^n(I_j \cap O_h)| \ge \varepsilon_0$, let $\mathcal{U}_{(j,h)}$ be the index set of a finite partition $\{U_\ell\}_{\ell \in \mathcal{U}_{(j,h)}}$ of the interval $f^n(I_j \cap O_h)$ into open intervals[6] of size*

$$\varepsilon_0/3 \le |U_\ell| < \varepsilon_0. \tag{4.10}$$

For $(j, h) \in \mathcal{J} \times \mathcal{H}^n$ such that $0 < |f^n(I_j \cap O_h)| < \varepsilon_0$ set $\mathcal{U}_{(j,h)} = \emptyset$. Define

$$\mathcal{J}_n := \{(j, h, \ell) | (j, h) \in \mathcal{J} \times \mathcal{H}^n, \ell \in \mathcal{U}_{(j,h)}, I_j \cap O_h \neq \emptyset\}.^7 \tag{4.11}$$

For every $j_n := (j, h, \ell) \in \mathcal{J}_n$, define

$$\begin{aligned} I_{j_n} &:= f^n(I_j \cap O_h) \cap U_\ell, \\ \rho_{j_n} &:= e^{ib\tau_n \circ h} \rho_j \circ h |h'| z_{j_n}^{-1}, \text{ where } z_{j_n} := \int_{I_{j_n}} |\rho_j| \circ h\, |h'| . \end{aligned}$$

Define $\mathcal{G}_n := \{(I_{j_n}, \rho_{j_n})\}_{j_n \in \mathcal{J}_n}$ and associate to it the measure given by

$$w_{\mathcal{G}_n}(j_n) = z_{j_n} w_{\mathcal{G}}(j). \tag{4.12}$$

Remark 12. *Comparing (2.11) with the definition of $\mathcal{G}_n$ and the measure associated to it (4.7), we have*

$$\mathscr{L}_b^n \rho_{\mathcal{G}} = \rho_{\mathcal{G}_n}. \tag{4.13}$$

This is the main connection between the evolution of densities under $\mathscr{L}_b^n$ and the evolution of standard families.

---

[6] Modulo a finite set of endpoints.

[7] When $\mathcal{U}_{(j,h)} = \emptyset$, by (j, h, ℓ) we mean (j, h).

4.1. **Invariance.** The first thing to show is the invariance of $\mathcal{M}_{a,b,B,\varepsilon_0}$ under iterations of $\mathscr{L}_b^n$ for large enough a, B, n and small enough ε_0.

Remark 13. *In this section, by a, σ, n large and ε_0 small we mean values that satisfy the following inequalities simultaneously.*

(1) $e^{-\lambda\alpha n} + D/a < 1$,
(2) $e^{-\lambda n} + C_\tau/a < 1$,
(3) $e^{3a}(2^n e^{-\lambda n} + e^{-\sigma}) < 1$,
(4) $\sigma > 0$ *is such that there exists* $\mathcal{H}_\sigma^n \subset \mathcal{H}^n$ *such that* $\mathcal{H}_0^n := \mathcal{H}^n \setminus \mathcal{H}_\sigma^n$ *is finite, and* $\sum_{h\in\mathcal{H}_\sigma^n} \sup |h'| < e^{-\sigma}$,
(5) ε_0 *is such that for every interval* I *with* $|I| < \varepsilon_0$, $\#\{h \in \mathcal{H}_0^n : I \cap O_h \neq \emptyset\} \leq 2^n$.

The constants D, C_τ were introduced in Section 2 by (2.4) *and* (2.9). *One may first choose a and n large enough that the first two inequalities hold. Then also choose σ (and n) large enough that $e^{3a}(2^n e^{-\lambda n} + e^{-\sigma}) < 1$. The value of ε_0 is then determined by n and σ.*

Proposition 14. *Suppose $\mathcal{G} \in \mathcal{M}_{a,b,B,\varepsilon_0}$ is a standard family. For every $n \in \mathbb{N}$, for every $(I_{j_n}, \rho_{j_n}) \in \mathcal{G}_n$ we have*

$$\int_{I_{j_n}} |\rho_{j_n}| = 1, \tag{4.14}$$

$$H(\rho_{j_n}) \leq a(e^{-\lambda\alpha n} + a^{-1}D), \tag{4.15}$$

$$|\arg(\rho_{j_n})'| \leq a|b|(e^{-\lambda n} + a^{-1}C_\tau). \tag{4.16}$$

For every $n \in \mathbb{N}$,

$$|\mathcal{G}_n| = |\mathcal{G}|. \tag{4.17}$$

For every a and n large, for every $\sigma > 0$, if $\varepsilon_0 > 0$ is small enough, then

$$|\partial_\varepsilon \mathcal{G}_n| \leq 2e^{3a}(2^n + e^{-\sigma}e^{\lambda n})\,|\partial_{e^{-\lambda n}\varepsilon}\mathcal{G}| + 6e^{a+D}\varepsilon_0^{-1}\varepsilon\,|\mathcal{G}|\,, \text{ for all } \varepsilon < \varepsilon_0. \tag{4.18}$$

Proof. Property (4.14) follows from the definition.

To show (4.15), note that

$$H(\rho_{j_n}) = H\left(h' \cdot (\rho_j \circ h)\right). \tag{4.19}$$

Using the definition of $H(\cdot)$ and noting its properties under multiplication and composition, it follows that

$$H(\rho_{j_n}) \leq H(h') + e^{-\lambda\alpha n} H(\rho_j).$$

By (2.4) we have $H(h') \leq D$, and by assumption $H(\rho_j) \leq a$, finishing the proof of (4.15).

To show (4.16), note that $\arg(\rho_{j_n}) = b\tau_n \circ h + \arg(\rho) \circ h$. Therefore using condition (2.9),

$$|\arg(\rho_{j_n})'| \leq |b|\,|(\tau_n \circ h)'| + |\arg(\rho)'|\,|h'| \leq |b|C_\tau + a|b|e^{-\lambda n}.$$

To show (4.17), write

$$\begin{aligned}\sum_{j_n\in\mathcal{J}_n} w_{\mathcal{G}_n}(j_n) &= \sum_{j_n\in\mathcal{J}_n} w_{\mathcal{G}}(j)\int_{I_{j_n}} |\rho_j|\circ h\,|h'| \\ &= \sum_{(j,h)\in\mathcal{J}\times\mathcal{H}^n}\ \sum_{\ell\in\mathcal{U}_{(j,h)}} w_{\mathcal{G}}(j)\int_{f^n(I_j\cap O_h)\cap U_\ell} |\rho_j|\circ h\,|h'| \\ &= \sum_{(j,h)\in\mathcal{J}\times\mathcal{H}^n} w_{\mathcal{G}}(j)\int_{f^n(I_j\cap O_h)} |\rho_j|\circ h\,|h'| \qquad (4.20)\\ &= \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j)\sum_{h\in\mathcal{H}^n}\int_{f^n(I_j\cap O_h)} |\rho_j|\circ h\,|h'| \\ &= \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j)\int_{I_j} |\rho_j| = \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j).\end{aligned}$$

Note that in the second line we wrote $\mathcal{J}\times\mathcal{H}^n$ instead of $\{(j,h)\in\mathcal{J}\times\mathcal{H}^n | I_j\cap O_h \neq \emptyset\}$. We can do this because if $I_j\cap O_h=\emptyset$, then the corresponding terms are zero. The third equality follows by summing over all ℓ since the intervals $f^n(I_j\cap O_h)\cap U_\ell$ form a partition of the interval $f^n(I_j\cap O_h)$. The last line is a consequence of change of variables and $\int|\rho_j|$ being equal to 1.

To prove (4.18), suppose $\sigma>0$. Then, (2.5) implies that there exists $\mathcal{H}^n_\sigma\subset\mathcal{H}^n$ such that

$$\mathcal{H}^n_0 := \mathcal{H}^n\setminus\mathcal{H}^n_\sigma \text{ is finite,} \qquad (4.21)$$

and

$$\sum_{h\in\mathcal{H}^n_\sigma}\ \sup_{f^n(O_h)} |h'| < e^{-\sigma}. \qquad (4.22)$$

Since $\mathcal{H}^n_0$ is finite, choose $\varepsilon_0=\varepsilon_0(\mathcal{H}^n_0)$ such that[8] for every interval I with $|I|<\varepsilon_0$,

$$\#\{h\in\mathcal{H}^n_0 : I\cap O_h\neq\emptyset\}\leq 2^n. \qquad (4.23)$$

Suppose $\varepsilon<\varepsilon_0$. We have, by definition,

$$|\partial_\varepsilon\mathcal{G}_n| := \sum_{j_n\in\mathcal{J}_n} w_{\mathcal{G}_n}(j_n)\int_{\partial_\varepsilon I_{j_n}} |\rho_{j_n}|.$$

We split the sum into two parts according to whether $\mathcal{U}_{(j,h)}\neq\emptyset$ or $\mathcal{U}_{(j,h)}=\emptyset$. The two parts are respectively,

$$\sum_{\{j_n\in\mathcal{J}_n|\mathcal{U}_{(j,h)}\neq\emptyset\}} w_{\mathcal{G}_n}(j_n)\int_{\partial_\varepsilon I_{j_n}}|\rho_{j_n}| = \sum_{j\in\mathcal{J}}\ \sum_{\{h\in\mathcal{H}^n|f^n(I_j\cap O_h)\geq\varepsilon_0\}}\ \sum_{\ell\in\mathcal{U}_{(j,h)}} w_{\mathcal{G}_n}(j_n)\int_{\partial_\varepsilon I_{j_n}}|\rho_{j_n}|,$$

where $j_n=(j,h,\ell)$ (as defined in (4.11)) and

$$\sum_{\{j_n\in\mathcal{J}_n|\mathcal{U}_{(j,h)}=\emptyset\}} w_{\mathcal{G}_n}(j_n)\int_{\partial_\varepsilon I_{j_n}}|\rho_{j_n}| = \sum_{j\in\mathcal{J}}\ \sum_{\{h\in\mathcal{H}^n|f^n(I_j\cap O_h)<\varepsilon_0\}} w_{\mathcal{G}_n}(j_n)\int_{\partial_\varepsilon I_{j_n}}|\rho_{j_n}|.$$

---

[8]There is some freedom here to choose ε_0. The value of ε_0 depends on the partition $\{O_h\}_{h\in\mathcal{H}^n}$ and the value of σ. Note that since we only use (4.18) with a fixed n, the value of ε_0 causes no problems even if it is very small. The optimal value depends on the underlying system.

The case $\mathcal{U}_{(j,h)} \neq \emptyset$: By Definition 11 of the iteration of standard families, and since $\sup_{I_j} |\rho_j| \asymp_a |I_j|^{-1}$ (recall (4.8) and (4.3)),

$$\begin{aligned} w_{\mathcal{G}_n}(j_n) \int_{\partial_\varepsilon I_{j_n}} |\rho_{j_n}| &= w_{\mathcal{G}}(j) \int_{\partial_\varepsilon I_{j_n}} |\rho_j| \circ h \, |h'| \\ &\leq_a w_{\mathcal{G}}(j) \, |\partial_\varepsilon I_{j_n}| \, |I_j|^{-1} \sup_{f^n(I_j \cap O_h)} |h'| \, . \\ &\leq_a w_{\mathcal{G}}(j) 2\varepsilon \, |I_j|^{-1} \sup_{f^n(I_j \cap O_h)} |h'| \, . \end{aligned} \tag{4.24}$$

Note that by the notation $\leq_a C$ we mean $\leq e^a C$. To finish the estimate, we need to sum (4.24) over ℓ, then sum over h such that $f^n(I_j \cap O_h) \geq \varepsilon_0$ and then over $j \in \mathcal{J}$. Note that since each interval $f^n(I_j \cap O_h)$ is chopped into intervals of size $\geq \varepsilon_0/3$, the number of elements in $\mathcal{U}_{(j,h)}$ is bounded by $3\varepsilon_0^{-1} |f^n(I_j \cap O_h)|$. Therefore, summing (4.24) over all $\ell \in \mathcal{U}_{(j,h)}$, and then using the distortion estimate (4.9) yields,

$$2\varepsilon 3\varepsilon_0^{-1} w_{\mathcal{G}}(j) \, |I_j|^{-1} \sup_{f^n(I_j \cap O_h)} |h'| \, |f^n(I_j \cap O_h)| \leq_{a+D} 6\varepsilon\varepsilon_0^{-1} w_{\mathcal{G}}(j) \, |I_j|^{-1} \, |I_j \cap O_h| \, .$$

Summing over h such that $f^n(I_j \cap O_h) \geq \varepsilon_0$ and noting that this is no greater than summing over all $h \in \mathcal{H}^n$ yields

$$\leq_{a+D} 6\varepsilon\varepsilon_0^{-1} w_{\mathcal{G}}(j) \, |I_j|^{-1} \, |I_j| = 6\varepsilon\varepsilon_0^{-1} w_{\mathcal{G}}(j).$$

Finally, summing over $j \in \mathcal{J}$ yields,

$$\sum_{\{j_n \in \mathcal{J}_n | \mathcal{U}_{(j,h)} \neq \emptyset\}} w_{\mathcal{G}_n}(j_n) \int_{\partial_\varepsilon I_{j_n}} |\rho_{j_n}| \leq_{a+D} 6\varepsilon\varepsilon_0^{-1} |\mathcal{G}| \, . \tag{4.25}$$

The case $\mathcal{U}_{(j,h)} = \emptyset$: We need to estimate:

$$\sum_{\{j_n \in \mathcal{J}_n | \mathcal{U}_{(j,h)} = \emptyset\}} w_{\mathcal{G}_n}(j_n) \int_{\partial_\varepsilon I_{j_n}} |\rho_{j_n}| = \sum_{j \in \mathcal{J}} \sum_{\{h \in \mathcal{H}^n | f^n(I_j \cap O_h) < \varepsilon_0\}} w_{\mathcal{G}_n}(j_n) \int_{\partial_\varepsilon I_{j_n}} |\rho_{j_n}| \, .$$

We further split the second sum into two parts, one over $\mathcal{H}_0^n$ and the other over $\mathcal{H}_\sigma^n$.

For the sum over $\mathcal{H}_0^n$, we have the bound

$$\begin{aligned} &\sum_{j \in \mathcal{J}} w_{\mathcal{G}}(j) \sum_{\{h \in \mathcal{H}_0^n | f^n(I_j \cap O_h) < \varepsilon_0\}} \int_{\partial_\varepsilon f^n(I_j \cap O_h)} |\rho_j| \circ h \, |h'| \\ &\leq \sum_{j \in \mathcal{J}} w_{\mathcal{G}}(j) \sum_{\{h \in \mathcal{H}_0^n | f^n(I_j \cap O_h) < \varepsilon_0\}} \left[\int_{f^n\left(\partial_{e^{-\lambda n}\varepsilon} I_j \cap O_h\right)} + \int_{f^n\left(\partial_{e^{-\lambda n}\varepsilon} O_h \cap I_j\right)} \right] |\rho_j| \circ h \, |h'| \\ &\leq |\partial_{e^{-\lambda n}\varepsilon} \mathcal{G}| + \sum_{j \in \mathcal{J}} w_{\mathcal{G}}(j) \sum_{\{h \in \mathcal{H}_0^n | f^n(I_j \cap O_h) < \varepsilon_0\}} \int_{\partial_{e^{-\lambda n}\varepsilon} O_h \cap I_j} |\rho_j| \, . \end{aligned}$$

The first inequality holds because if a point is at a distance less than ε from the boundary of $f^n(I_j \cap O_h)$, then its preimage must be at a distance $e^{-\lambda n}\varepsilon$ from the boundary of I_j or from the boundary of O_h. The second inequality is a consequence of change of variables and f^n being one-to-one on $I_j \cap O_h$.

Since $|I_j| < \varepsilon_0$, by the choice of ε_0, it follows that I_j intersects at most 2^n intervals O_h. Also, (4.8) implies that $\mathcal{A}_{\partial_{e^{-\lambda n}\varepsilon} O_h \cap I_j} |\rho_j| \asymp_a \mathcal{A}_{\partial_{e^{-\lambda n}\varepsilon} I_j} |\rho_j|$. This in turn implies $\int_{\partial_{e^{-\lambda n}\varepsilon} O_h \cap I_j} |\rho_j| \leq_a \int_{\partial_{e^{-\lambda n}\varepsilon} I_j} |\rho_j|$ because $|\partial_{e^{-\lambda n}\varepsilon} O_h \cap I_j| \leq |\partial_{e^{-\lambda n}\varepsilon} I_j|$.

Therefore,

$$\begin{aligned}&\sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j) \sum_{\{h\in\mathcal{H}_0^n | f^n(I_j\cap O_h)<\varepsilon_0\}} \int_{\partial_\varepsilon f^n(I_j\cap O_h)} |\rho_j|\circ h\,|h'| \\ &\leq_a |\partial_{e^{-\lambda n}\varepsilon}\mathcal{G}| + 2^n \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j) \int_{\partial_{e^{-\lambda n}\varepsilon} I_j} |\rho_j| \\ &\leq_a 2^n 2\,|\partial_{e^{-\lambda n}\varepsilon}\mathcal{G}|\,.\end{aligned} \tag{4.26}$$

For the sum over $\mathcal{H}_\sigma^n$, similarly to (4.24), we have the bound

$$\sum_{\{j_n|\mathcal{U}_{(j,h)}=\emptyset\}} w_{\mathcal{G}_n}(j_n)\int_{\partial_\varepsilon I_{j_n}} |\rho_{j_n}| \leq_a \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j) \sum_{\{h\in\mathcal{H}_\sigma^n | f^n(I_j\cap O_h)<\varepsilon_0\}} |I_j|^{-1}\,|\partial_\varepsilon I_{j_n}| \sup_{f^n(I_j\cap O_h)} |h'|\,.$$

Multiplying and dividing the right hand side by $|\partial_\varepsilon I_j|$ and using $|\partial_\varepsilon I_{j_n}|\,|\partial_\varepsilon I_j|^{-1} \leq 1$, the right hand side is

$$\leq_a \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j)\,|\partial_\varepsilon I_{j_n}|\,|I_j|^{-1} \sum_{\{h\in\mathcal{H}_\sigma^n | f^n(I_j\cap O_h)<\varepsilon_0\}} \sup_{f^n(I_j\cap O_h)} |h'|\,.$$

Notice that $|\partial_\varepsilon I_{j_n}|\,|I_j|^{-1} \asymp_a \int_{\partial_\varepsilon I_j} |\rho_j|$. Therefore, the above quantity is

$$\leq_{2a} \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j)\int_{\partial_\varepsilon I_j} |\rho_j| \sum_{\{h\in\mathcal{H}_\sigma^n | f^n(I_j\cap O_h)<\varepsilon_0\}} \sup_{f^n(I_j\cap O_h)} |h'|\,.$$

Using (4.22), the estimate for the sum over $\mathcal{H}_\sigma^n$ is

$$\leq_{2a} e^{-\sigma} \sum_{j\in\mathcal{J}} w_{\mathcal{G}}(j)\int_{\partial_\varepsilon I_j} |\rho_j| = e^{-\sigma}\,|\partial_\varepsilon\mathcal{G}| \leq_a e^{-\sigma}e^{\lambda n}\,|\partial_{e^{-\lambda n}\varepsilon}\mathcal{G}|\,. \tag{4.27}$$

Finally, adding (4.25), (4.26) and (4.27) together, we arrive at

$$|\partial_\varepsilon\mathcal{G}_n| \leq 2e^{3a}(2^n + e^{-\sigma}e^{\lambda n})\,|\partial_{e^{-\lambda n}\varepsilon}\mathcal{G}| + 6e^{a+D}\varepsilon_0^{-1}\varepsilon\,|\mathcal{G}|\,. \tag{4.28}$$

□

Lemma 15. *Suppose $\mathcal{G} \in \mathcal{M}_{a,b,B,\varepsilon_0}$ with a sufficiently large and ε_0 sufficiently small. Then, there exist C, $\bar{C}$, and $0 < \beta < \lambda$ such that,*

$$|\partial_\varepsilon\mathcal{G}_m| \leq Ce^{\beta m}\,|\partial_{e^{-\lambda m}\varepsilon}\mathcal{G}| + \bar{C}\varepsilon,\ \textit{for all } m\in\mathbb{N},\ \textit{and } \varepsilon<\varepsilon_0. \tag{4.29}$$

Proof. The result follows by choosing a fixed n as in Remark 13 and iterating (4.18) with this fixed n. Choose n, σ large such that $2e^{3a}(2^n e^{-\lambda n} + e^{-\sigma}) < 1$ as in Remark 13. Let β be such that $e^\beta = (2e^{3a})^{1/n}(2^n + e^{-\sigma}e^{\lambda n})^{1/n}$. Then $e^{\beta n}e^{-\lambda n} < 1$. That is, $\beta < \lambda$. For every $m \in \mathbb{N}$, write $m = kn + r$, $0 \leq r < n$. Applying (4.18) for r, we have

$$|\partial_\varepsilon\mathcal{G}_m| \leq 2e^{3a}(2^r + e^{-\sigma}e^{\lambda r})\,|\partial_{e^{-\lambda r}\varepsilon}\mathcal{G}_{kn}| + 6e^{a+D}\varepsilon_0^{-1}\varepsilon\,|\mathcal{G}_{kn}|\,. \tag{4.30}$$

Let $\tilde{C} = 2e^{3a}(2^n + e^{-\sigma}e^{\lambda n})$. Applying (4.18) for n (k times), we get

$$\begin{aligned}|\partial_\varepsilon\mathcal{G}_m| &\leq \tilde{C}e^{\beta kn}\,|\partial_{e^{-\lambda r}e^{-\lambda kn}\varepsilon}\mathcal{G}| + 6e^{a+D}\varepsilon_0^{-1}\varepsilon(e^{-\lambda r}/(1-e^{\beta-\lambda})+1)\,|\mathcal{G}| \\ &\leq Ce^{\beta m}\,|\partial_{e^{-\lambda m}\varepsilon}\mathcal{G}| + \bar{C}\varepsilon\,|\mathcal{G}|\,,\end{aligned} \tag{4.31}$$

where $\bar{C} = 6e^{a+D}\varepsilon_0^{-1}(e^{-\lambda r}/(1-e^{\beta-\lambda})+1)$. □

Remark 16. *The invariance of $\mathcal{M}_{a,b,B,\varepsilon_0}$ under iterations by $\mathscr{L}_b^m$ follows by taking a, n, B large, and ε_0 small as in Remark 13. Note that $m = m(n)$ and $B = B(n)$ must be large enough that $Ce^{(\beta-\lambda)m} + \bar{C}/B < 1$ to guarantee $|\partial_\varepsilon\mathcal{G}_m| \leq B\varepsilon$ for all $\varepsilon < \varepsilon_0$.*

5. Transversality

Since the dynamics of F in the y-direction is a rigid rotation, it is possible that F does not mix or mixes at an arbitrary slow rate. For, example if τ is constant or cohomologous to a constant function, it is easy to see that F will not be mixing. More generally, there are examples[9] of f piecewise expanding and τ cohomologous to a piecewise constant where F mixes at arbitrary slow rates [22, 20]. To avoid such scenarios we assume that τ is not Lipschitz-cohomologous to a piecewise constant function on the partition of f. In this section, we show that this condition implies a uniform non-integrability condition,[10] which in turn implies the transversality notion that we use later to obtain a stretched-exponential decay of correlations. The use of the term transversality and the use of invariant cones is influenced by [25]. In [6] a stronger notion of transversality (inspired by Tsujii [25]) is used to obtain an exponential rate of mixing in the setting that f and τ are piecewise $\mathscr{C}^2$ on a finite partition. However, as explained in the introduction, the same procedure is not immediately applicable to the setting of this paper.

Note that

$$DF_{(x,y)} = \begin{pmatrix} f'(x) & 0 \\ \tau'(x) & 1 \end{pmatrix}, \tag{5.1}$$

which is independent of the second coordinate. Also, note that DF preserves the cone $\mathscr{K}_\eta = \{(u,v) : |v| \leq \eta\,|u|\}$, where $\eta = \|\tau'/f'\|_\infty / (1 - \|1/f'\|_\infty)$.

Lemma 17. *Suppose that for every $n \in \mathbb{N}$, for every $x \in \mathbb{T}^1$, and inverse branches $h_1, h_2 \in \mathcal{H}^n$ both containing x in their domain,*

$$DF^n_{h_1(x)}\mathscr{K}_\eta \cap DF^n_{h_2(x)}\mathscr{K}_\eta \neq \{0\}.$$

Then, τ is Lipschitz-cohomologous to a piecewise constant function on the partition of f.

Proof. The hypothesis implies that for every $x \in \mathbb{T}^1$, $\cap_{h\in\mathcal{H}^n} DF^n_{h(x)}\mathscr{K}_\eta \neq \{0\}$. That is, the intersection contains a common direction $\theta(x,n)$. Moreover, since the cones $DF^n_{h(x)}\mathscr{K}_\eta$ contract uniformly under iteration, the intersection of cones $\cap_{n\in\mathbb{N}}\cap_{h\in\mathcal{H}^n} DF^n_{h(x)}\mathscr{K}_\eta$ contains a unique direction $(1,\theta(x))$. Since $\theta(x)$ is invariant under DF, we have

$$f' \cdot \theta \circ f = \tau' + \theta. \tag{5.2}$$

Define $\phi(x) = \int_0^x \theta(t)dt$. Note that since $\theta(x)$ is bounded, ϕ is Lipschitz. Let p be the left endpoint of a partition element of the partition of f and x be a point in the same partition element. We may write $\tau(x) = \tau(p) + \int_p^x \tau'$, where $\tau(p)$ is interpreted as the value obtained by taking a one-sided limit. Substituting $\tau' = f' \cdot \theta \circ f - \theta$ from (5.2) into this equation, and doing a change of variables yields,

$$\tau(x) = \phi \circ f(x) - \phi(x) + (\tau(p) + \phi(p) - \phi \circ f(p)). \tag{5.3}$$

It follows that τ is Lipschitz-cohomologous to a piecewise constant function on the partition of f. □

Lemma 18. *Suppose for $x \in \mathbb{T}^1$, $n \in \mathbb{N}$, and inverse branches $h_1, h_2 \in \mathcal{H}^n$ holds*

$$DF^n_{h_1(x)}\mathscr{K}_\eta \cap DF^n_{h_2(x)}\mathscr{K}_\eta = \{0\}.$$

Then, there exists $C_0 := C_0(n,x)$ such that

$$|(\tau_n \circ h_1)'(x) - (\tau_n \circ h_2)'(x)| > C_0.$$

[9] I do not know whether *all* skew products of the current setting necessarily mix subexponentially (at best) when condition (2.8) fails.

[10] The connection between not being cohomologous to a piecewise constant function and uniform non-integrability (in the smooth setting) was first established in [1].

Proof. The hypothesis implies that the two cones $DF^n_{h_1(x)}\mathscr{K}_\eta$ and $DF^n_{h_2(x)}\mathscr{K}_\eta$ are a distance $C_0 := C_0(n,x)$ apart. Suppose $v_1, v_2 \in \mathbb{R}$ satisfy $|v_1|, |v_2| \le \eta$. Then $(1,v_1),(1,v_2) \in \mathscr{K}_\eta$. Observe that for $j \in \{1,2\}$, $DF^n_{h_j(x)}(1,v_j) = ((f^n)'\circ h_j(x), \tau_n' \circ h_j(x) + v_j)$. Therefore $(1, (\tau_n \circ h_j)'(x) + v_j h_j') \in DF^n_{h_j(x)}\mathscr{K}_\eta$. Therefore, the two vectors are also C_0 apart; that is,

$$|(\tau_n \circ h_1)'(x) + v_1 h_1'(x) - (\tau_n \circ h_2)'(x) - v_2 h_2'(x)| > C_0.$$

Using the triangle inequality,

$$|(\tau_n \circ h_1)'(x) - (\tau_n \circ h_2)'(x))| > C_0 - |v_1 h_1'(x) - v_2 h_2'(x)|.$$

Taking $v_1 = v_2 = 0$ implies the result. □

Lemma 19. *Suppose τ is not Lipschitz-cohomologous to a piecewise constant function on the partition of f. Then, there exist $x_0 \in \mathbb{T}^1$, there exists $n_1 \in \mathbb{N}$, inverse branches $h_1, h_2 \in \mathcal{H}^{n_1}$, a neighbourhood V_{n_1} of x_0 contained in the open set $f^{n_1}(O_{h_1}) \cap f^{n_1}(O_{h_2})$, and a constant $C_1 := C_1(n_1, x_0)$ such that*

$$\left|(\tau_{n_1} \circ h_1 - \tau_{n_1} \circ h_2)'(x)\right| > C_1 \text{ for every } x \in V_{n_1}. \tag{5.4}$$

Proof. Suppose τ is not Lipschitz-cohomologous to a piecewise constant function on the partition of f. Lemma 17 implies that there exists $n_1 \in \mathbb{N}$, $x_0 \in \mathbb{T}^1$, and inverse branches $h_1, h_2 \in \mathcal{H}^{n_1}$ such that

$$DF^{n_1}_{h_1(x_0)}\mathscr{K}_\eta \cap DF^{n_1}_{h_2(x_0)}\mathscr{K}_\eta = \{0\}. \tag{5.5}$$

Lemma 18 implies that there exists $C_0 = C_0(n_1, x_0)$ such that

$$\left|(\tau_{n_1} \circ h_1 - \tau_{n_1} \circ h_2)'(x_0)\right| \ge C_0. \tag{5.6}$$

By continuity of $(f^{n_1})'$ and τ_{n_1}' at $h_1(x_0)$ and $h_2(x_0)$ the cones $DF^{n_1}_{h_1(x_0)}\mathscr{K}_\eta$ and $DF^{n_1}_{h_2(x_0)}\mathscr{K}_\eta$ vary continuously in a neighbourhood of x_0 and so does the distance between them (i.e. $C(n,\cdot)$ varies continuously in a neighbourhood of x_0). It follows that there exists a neighbourhood V_{n_1} of x_0 and a constant, which we again denote by $C_1 := C_1(n_1, x_0)$ such that

$$\left|(\tau_{n_1} \circ h_1 - \tau_{n_1} \circ h_2)'(x)\right| \ge C_1 \text{ for every } x \in V_{n_1}. \tag{5.7}$$

□

Corollary 20. *Suppose τ is not Lipschitz-cohomologous to a piecewise constant function on the partition of f. Then, there exists $x_0 \in \mathbb{T}^1$, there exists $n_1 \in \mathbb{N}$, inverse branches $h_1, h_2 \in \mathcal{H}^{n_1}$, a constant $C_1 := C_1(n_1, x_0)$, and for every $n \ge n_1$ and every $l_1, l_2 \in \mathcal{H}^{n-n_1}$ there exists a neighbourhood V_n of x_0 contained in $f^n(O_{l_1\circ h_1}) \cap f^n(O_{l_2\circ h_2})$ such that*

$$\left|(\tau_n \circ l_1 \circ h_1 - \tau_n \circ l_2 \circ h_2)'(x)\right| > C_1 \text{ for every } x \in V_n. \tag{5.8}$$

Proof. Let x_0, n_1, V_{n_1} and C_1 be as in Lemma 19. For every $n \ge n_1$ and $l_1, l_2 \in \mathcal{H}^{n-n_1}$, by invariance of the cone, we have $DF^{n-n_1}_{l_j\circ h_j(x)}\mathscr{K}_\eta \subset \mathscr{K}_\eta$ for every $x \in f^n(O_{l_1\circ h_1}) \cap f^n(O_{l_2\circ h_2})$. If also $x \in V_{n_1}$ (the neighbourhood of x_0 from Lemma 19), then $DF^{n_1}_{h_1(x)}DF^{n-n_1}_{l_1\circ h_1(x)}\mathscr{K}_\eta \cap DF^{n_1}_{h_2(x)}DF^{n-n_1}_{l_2\circ h_2(x)}\mathscr{K}_\eta = \{0\}$. That is, the cones $DF^n_{l_1\circ h_1(x)}\mathscr{K}_\eta$ and $DF^n_{l_2\circ h_2(x)}\mathscr{K}_\eta$ are at least distant C_1 apart. As in Lemma 18, this transversality of the cones implies

$$\left|(\tau_{n_1} \circ l_1 \circ h_1 - \tau_{n_1} \circ l_2 \circ h_2)'(x)\right| > C_1 \text{ for every } x \in V_n, \tag{5.9}$$

where $V_n := V_{n_1} \cap f^n(O_{l_1\circ h_1}) \cap f^n(O_{l_2\circ h_2})$. □

The following shows that any interval of positive length maps forward, while getting cut and expanded, in a way that at least two of its pieces overlap and simultaneously satisfy a condition similar to (5.4).

Proposition 21. *Suppose τ is not Lipschitz cohomologous to a piecewise constant function on the partition of f; and, in addition, f is covering. There exists a constant C_1 such that for every interval I with $0 < \delta < |I| \le \varepsilon_0$, there exists n_δ such that for every $n \ge n_\delta$, there exist $h_1, h_2 \in \mathcal{H}^n$, such that $O_{h_1}, O_{h_2} \subset I$ and $f^n(O_{h_1}) \cap f^n(O_{h_2})$ contains an interval I_* of size $0 < \Delta \le |I_*|$ on which holds*[11]

$$\left|(\tau_n \circ h_1 - \tau_n \circ h_2)'\right| > C_1.$$

Proof. Corollary 20 implies that there exists x_0, n_1, inverse branches $\tilde{h}_1, \tilde{h}_2 \in \mathcal{H}^{n_1}$; there exists a constant C_1; and, for every $n \ge n_1$ and every $l_1, l_2 \in \mathcal{H}^{n-n_1}$, there exists a neighbourhood V_n of x_0 contained in $f^n(O_{l_1 \circ \tilde{h}_1}) \cap f^n(O_{l_2 \circ \tilde{h}_2})$, such that

$$\left|\left(\tau_n \circ l_1 \circ \tilde{h}_1 - \tau_n \circ l_2 \circ \tilde{h}_2\right)'(x)\right| > C_1 \text{ for every } x \in V_n. \tag{5.10}$$

Since f is covering, there exists $N(\delta)$ (recall that δ is the lower bound on the length of I) such that for every $n \ge N(\delta) + n_1 =: n_\delta$ and every $l_1, l_2 \in \mathcal{H}^{n-n_1}$

$$l_j \circ \tilde{h}_j(V_n) \subset O_{l_j \circ h_j} \subset I, \text{ for } j \in \{1, 2\}.$$

Note that the first inclusion is a consequence of the property that V_n, $n \ge n_1$, is contained in $f^n(O_{l_1 \circ \tilde{h}_1}) \cap f^n(O_{l_2 \circ \tilde{h}_2})$. Set $h_1 := l_1 \circ \tilde{h}_1$, $h_2 := l_2 \circ \tilde{h}_2$ and $I_* = V_n$. Then, we have

$$\left|(\tau_n \circ h_1 - \tau_n \circ h_2)'\right| > C_1 \text{ for every } x \in I_*.$$

Denote the length of I_*, the overlap interval, by Δ. Note that Δ depends on δ, n and the choice of the initial interval I. □

5.1. **Transversality of standard pairs.** In this subsection we will state the transversality condition of Proposition 21 in terms of standard pairs. We will also get rid of the dependence of Δ on the choice of the interval I using a compactness argument. The following proposition provides us with two standard pairs that overlap and are transversal on an interval of length no smaller than $\Delta = \Delta(k_\delta, n_\delta)$.

Proposition 22 (Transversality of standard pairs)**.** *Suppose τ is not Lipschitz cohomologous to a piecewise constant function on the partition of f and that parameters $a, \sigma, B, n, \varepsilon_0$ satisfy conditions of Remark 13 and Remark 16. Consider a standard family $\mathcal{G}$. For every $\delta > 0$, there exists $n_\delta \in \mathbb{N}$, a finite number $k := k_\delta$ of pairs of inverse branches*

$$\{(h_{1,1}, h_{1,2}) \dots (h_{k,1}, h_{k,2})\} \subset \mathcal{H}^{n_\delta} \times \mathcal{H}^{n_\delta}$$

such that for any standard pair $(I, \rho) \in \mathcal{G}$ with $|I| > 3\delta$, there exist $l \in \{1, \dots, k_\delta\}$, $\Delta := \Delta(k_\delta, n_\delta) > 0$, standard pairs $(I_j, \rho_j) \in \mathcal{G}_{n_\delta}$, $j \in \{1, 2\}$ such that there exists U with $\varepsilon_0/3 \le |U| \le \varepsilon_0$,[12] *such that*

$$I_j := f^{n_\delta}(O_{h_{l,j}}) \cap U, \rho_j := z_{h_{l,j}}^{-1} e^{ib\tau_{n_\delta} \circ h_{l,j}} \rho \circ h_{l,j} |(h_{l,j})'|, \text{ and } O_{h_{l,j}} \subset I.$$

Furthermore, $I_1 \cap I_2$ contains an interval I_* of size Δ on which holds

$$\left|(\tau_{n_\delta} \circ h_{l,1} - \tau_{n_\delta} \circ h_{l,2})'\right| > C_1,$$

where C_1 was defined in Lemma 19. Denote

$$M(n_\delta) := \min_{\substack{j \in \{1,2\} \\ l \in \{1, \cdots k_\delta\}}} \{\left|O_{h_{l,j}} \cap h_{l,j}(U_j)\right|\}. \tag{5.11}$$

[11] The quantity Δ depends on δ, n_δ and the choice of I. Later in Proposition 22 we will get rid of the dependence on I by a compactness argument.

[12] U is a choice of cutting and can be taken to be equal to $\mathbb{T}^1$ if no cutting is necessary; that is, when $\left|f^{n_\delta}(O_{h_{l,j}})\right| < \varepsilon_0$.

Proof. Divide $\mathbb{T}^1$ into intervals of length δ. Denote the finite collection of intervals by $\{J_l\}_{l=1}^{k_\delta}$. For each interval apply Proposition 21. It follows that there exist $n := n_\delta$ and finitely many inverse branches

$$\{(h_{1,1}, h_{1,2}) \dots (h_{k,1}, h_{k,2})\} \subset \mathcal{H}^{n_\delta} \times \mathcal{H}^{n_\delta}$$

such that $O_{h_{l,1}}, O_{h_{l,2}} \subset J_l$ and $f^n(O_{h_{l,1}}) \cap f^n(O_{h_{l,2}})$ contains an interval of size $\Delta_l > 0$[13] on which holds

$$\left|(\tau_n \circ h_{l,1} - \tau_n \circ h_{l,2})'\right| > C_1.$$

Let $\Delta = \min_{l \in \{1,\cdots,k_\delta\}} \Delta_l$. For any standard pair (I, ρ) with $|I| > 3\delta$, I contains at least one of the intervals J_l of length δ. As mentioned above, J_l contains a pair of partition intervals $O_{h_{l,1}}, O_{h_{l,2}}$ whose images overlap over an interval of length $\varepsilon_0/3$ and are transversal. If these images are of length $< \varepsilon_0$, by definition, they are the support of standard pairs:

$$I_j := f^n(O_{h_{l,j}}), \rho_j := z_{h_{l,j}}^{-1} e^{ib\tau_n \circ h_{l,j}} \rho \circ h_{l,j} |(h_{l,j})'|.$$

However, if one of the images is of size greater than ε_0, it must be shortened. In this case we may choose a cutting interval U, with $\varepsilon_0/3 \leq |U| \leq \varepsilon_0$ that does not cut the overlap if $\Delta < \varepsilon_0/3$. We also require that the cutting does not create other pieces of length $< \varepsilon_0/3$. This can be done if $\Delta < \varepsilon_0/3$ and if $\Delta \geq \varepsilon_0/3$, we can consider a smaller overlap interval of size $< \varepsilon_0/3$. With these considerations, we have obtained two standard pairs such that

$$I_j := f^n(O_{h_{l,j}}) \cap U, \rho_j := z_{h_{l,j}}^{-1} e^{ib\tau_n \circ h_{l,j}} \rho \circ h_{l,j} |(h_{l,j})'|,$$

and such that $O_{h_{l,1}}, O_{h_{l,2}} \subset I$ and $I_1 \cap I_2$ contains an interval of length Δ on which holds

$$\left|(\tau_n \circ h_{l,1} - \tau_n \circ h_{l,2})'\right| > C_1.$$

□

Remark 23. *The actual value of 3δ for which Proposition 22 above is used, is determined in Proposition 27 and will be fixed. Note that Δ depends on n_δ and b_0 will depend on Δ via* (6.6).

6. Weight reduction of standard families

In this section our goal is to replace a standard family, after certain number of iterations, with an equivalent standard family of lower total weight. This is the oscillatory cancellation mechanism that is used to show the decay of $\|\mathscr{L}_b^n\|_{\mathbf{L}^1}$. Existence of full phase oscillations is shown in Claim 1 where it is used that the argument of a standard pair is $\mathscr{C}^1$ (see also the paragraph after Claim 1 and also Claim 3).

Definition 24 (Equivalence)**.** *Two standard families $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are said to be* equivalent *if $\rho_{\mathcal{G}} = \rho_{\tilde{\mathcal{G}}}$, i.e. if*

$$\sum_{j \in \mathcal{J}} w_{\mathcal{G}}(j)\rho_j = \sum_{k \in \tilde{\mathcal{J}}} w_{\tilde{\mathcal{G}}}(k)\tilde{\rho}_k. \tag{6.1}$$

Remark 25. *In this section we assume the conditions on parameters of Remark 13 and Remark* 16*, but in addition we need to slightly increase the value of the parameter a. More precisely, we need $\left(e^{-\lambda n} + C_\tau/a + C_\kappa/a\right)\alpha_1^{-1} < 1$. This does not cause any problems since we could have chosen a larger to begin with in Remark 13. In regards to n, we need $n > n_\delta$ (δ is fixed; see Remark 23) and we choose it large enough that the above inequality holds and also $ae^{-\lambda n} < C_1/4$. Finally, we assume*

[13] Δ_l depends on δ and n_δ in addition to l.

that $|b| \geq b_0 = 4\pi/(C_1\Delta)$, where b_0 appears in (6.6), C_1 was defined Lemma 19, and $\Delta = \Delta(k_\delta, n_\delta)$ in Proposition 22.

Lemma 26. *Suppose $\mathcal{G} = \{(I,\rho)\} \in \mathcal{M}_{a,b,B,\varepsilon_0}$ is a singleton standard family with $\delta \leq |I|$ for which Proposition 22 holds. Assume the parameter conditions from Remark 13, Remark 16 and Remark 25. Then there exists constants $\gamma > 0$, $C = C_\delta > 0$, $b_0 > 0$ (given by (6.6)) such that for $|b| > b_0$, letting $n_b = C_\delta \ln|b|$, there exists a standard family $\mathcal{G}^*_{n_b}$ equivalent to $\mathcal{G}_{n_b}$ such that*

$$\sum_{j \in \mathcal{J}^*_{n_b}} w_{\mathcal{G}^*_{n_b}}(j) \leq e^{-\gamma} w_{\mathcal{G}}.$$

Proof. Let $(I_1,\rho_1),(I_2,\rho_2) \in \mathcal{G}_n$ (where $n > n_\delta$) be the transversal standard pairs provided by Proposition 22 applied to $\mathcal{G} = \{(I,\rho)\}$. Let $w_1 = w_{\mathcal{G}_n}(h_1)$ and $w_2 = w_{\mathcal{G}_n}(h_2)$ be the weights of these standard pairs.[14] Let I_* be the interval of length Δ on which Proposition 22 holds. Let $\theta_1 = \arg(\rho \circ h_1)$ and $\theta_2 = \arg(\rho \circ h_2)$. Then, on the interval I_*, we may write $|w_1\rho_1 + w_2\rho_2| = \left|e^{i\Theta_b} w_1 |\rho_1| + w_2 |\rho_2|\right|$, where

$$\Theta_b = b(\tau_n \circ h_1 - \tau_n \circ h_2) - (\theta_1 - \theta_2). \tag{6.2}$$

Our goal is to take out ρ_1 and ρ_2 from the family $\mathcal{G}_n$ and replace them with other standard pairs (formed by combining parts of ρ_1 and ρ_2) and obtain a standard family $\mathcal{G}^*_n$ which is still equivalent to $\mathcal{G}_n$, but has a total weight strictly less than that of $\mathcal{G}_n$.

We will first show that the phase difference Θ_b grows at a certain rate.

Claim 1 (Full phase oscillation)**.** *Let C_1 be as in Lemma 19. For large n and for every $b \neq 0$,*

$$\frac{|b|C_1}{2} \leq |\Theta_b'| \leq 2|b|(C_\tau + C_1) \text{ on the interval } I_*. \tag{6.3}$$

Proof. Note that by (4.5), on I_*,

$$\begin{aligned} |\theta_1'| &= \left|\arg(\rho \circ h_1)'\right| \leq a|b||h_1'| \leq a|b|e^{-\lambda n}, \\ |\theta_2'| &\leq a|b|e^{-\lambda n}. \end{aligned} \tag{6.4}$$

Choose n large enough that

$$ae^{-\lambda n} < C_1/4. \tag{6.5}$$

Then, $|\theta_1 - \theta_2| < |b|C_1/2$. Also, Proposition 22 implies $|b(\tau_n \circ h_1 - \tau_n \circ h_2)'| > |b|C_1$ hence $|\Theta_b'| > |b|C_1/2$. Finally, a simple estimate shows that $|\Theta_b'| \leq 2|b|(C_\tau + C_1)$. □

Note that since Θ_b is $\mathscr{C}^1$ and satisfies the bounds (6.3), Θ_b' does not change sign in I_*. Divide the range of Θ_b into intervals of length between 2π and 3π, then the bounds on Θ_b' imply that I_* will be divided into corresponding intervals I_m of length $K_1|b|^{-1} \leq |I_m| \leq K_2|b|^{-1}$, where $K_1 := \pi/(C_\tau + C_1)$ and $K_2 := 6\pi/C_1$. To clarify, these are intervals on which Θ_b makes one full oscillation, but less than one and a half full oscillations. Of course we must make sure I_* is large enough to fit at least one such interval I_m. That is we need $K_2 |b|^{-1} \leq |I_*|$. This can be accomplished by choosing b large enough:[15]

$$|b| \geq K_2/|I_*| = 4\pi/(C_1\Delta) := b_0. \tag{6.6}$$

We would like to combine some part of ρ_1 and ρ_2 to take advantage of their cancellation. However, since the modulus of these standard pairs are not smooth, if we combine them blindly, we might lose the $\mathscr{C}^1$-smoothness required for the

[14] We are using $\mathcal{H}^n$ for the index set of $\mathcal{G}_n$ to keep the notation simpler. To be precise, write $w_{\mathcal{G}_n}(j_n)$ with $j_n \in \mathcal{J}_n$ as defined above.

[15] This is the only restriction on b.

argument of a standard pair. For this reason we do the following *splitting trick*. We split the standard pairs into good parts, with a constant modulus, which we can combine; and bad parts, which we do not combine in this round.

For a function $\rho \in \mathbf{L}^1(I,\mathbb{C})$ with $\int_I |\rho| \neq 0$, denote $N(\rho) = \rho / \int_I |\rho|$. For $j \in \{1,2\}$, we split (I_j, ρ_j) into two standard pairs $(I_j, N(\tilde\rho_j))$, $(I_j, N(\bar\rho_j))$, such that:

$$\bar\rho_j = ce^{i\Theta_j}, \tilde\rho_j = (|\rho_j| - c)e^{i\Theta_j}, \text{ where } c = \frac{e^{-a}}{2}, \Theta_j = b(\tau_n \circ h_j) + \theta_j. \tag{6.7}$$

Associate to them the weights $\bar w_j = w_j \int_{I_j} |\bar\rho_j|$, $\tilde w_j = w_j \int_{I_j} |\tilde\rho_j|$.

Claim 2 (After splitting)**.** *For $j \in \{1,2\}$,*

$$\begin{aligned}
\rho_j &= \bar\rho_j + \tilde\rho_j && (6.8)\\
w_j &= \bar w_j + \tilde w_j && (6.9)\\
w_j\rho_j &= \bar w_j N(\bar\rho_j) + \tilde w_j N(\tilde\rho_j) && (6.10)\\
H(\bar\rho_j) &\leq a && (6.11)\\
H(\tilde\rho_j) &\leq 4a && (6.12)\\
\left|\Theta_j'\right| &\leq a|b|\left(e^{-\lambda n} + \frac{C_\tau}{a}\right) && (6.13)
\end{aligned}$$

Proof. The first four statements are easy to prove. To prove (6.12), note that (by Lemma 10) $c = (1/2)e^{-a} \leq (1/2)\inf|\rho_j|$. Therefore,

$$\frac{||\rho_j(x)| - c|}{||\rho_j(y)| - c|} \leq \frac{||\rho_j(x)| - |\rho_j(y)|| + |\rho_j(y)| - c}{|\rho_j(y)| - c} = 1 + \frac{||\rho_j(x)| - |\rho_j(y)||}{|\rho_j(y)| - c} \tag{6.14}$$

But, $|\rho_j(y)| - (1/2)|\rho_j(y)| = (1/2)|\rho_j(y)| \geq \inf|\rho_j| \geq c$. Hence, $|\rho_j(y)| - c \geq \frac{1}{2}|\rho_j(y)|$, and we have:

$$\begin{aligned}
\frac{||\rho_j(x)| - c|}{||\rho_j(y)| - c|} &\leq 1 + 2\frac{||\rho_j(x)| - |\rho_j(y)||}{|\rho_j(y)|} \leq 1 + 2\left|\frac{|\rho_j(x)|}{|\rho_j(y)|} - 1\right| \\
&\leq 1 + 2\left|e^{a|x-y|^\alpha} - 1\right| = 2e^{a|x-y|^\alpha} - 1 \\
&\leq e^{2a|x-y|^\alpha}.
\end{aligned} \tag{6.15}$$

The inequality (6.13) follows from (6.4) and (2.9). □

We now describe how to combine $\bar\rho_1$ and $\bar\rho_2$. Note that the modulus of these functions is constant and equal to c. We need the following result.

Claim 3 (J_m. Preparing for a controlled cancellation of $\bar\rho_1$ and $\bar\rho_2$)**.** *Suppose $w_2 \leq w_1$.*[16] *For every $\alpha_1 \in (0,1/2]$ and $\alpha_2 \in [(\sqrt{7}-1)/2, 1)$, for every m, there exists a subinterval $J_m \subset I_m$ with $K_3|b|^{-1} \leq |J_m| \leq K_4|b|^{-1}$ such that for every $\kappa_0 \geq w_2/(2w_1)$*

$$\alpha_1 c(\kappa_0 w_1 + w_2) \leq |\kappa_0 w_1 \bar\rho_1 + w_2\bar\rho_2| \leq c(\kappa_0 w_1 + \alpha_2 w_2). \tag{6.16}$$

Proof. Choose K_3, K_4 such that $1/4 \leq \cos(\Theta_b) \leq 1/2$ on J_m. This can be done because the phase difference Θ_b makes a full oscillation in I_m. The left side of (6.16) is easy to prove and does not require a restriction on κ_0. Let us prove the right side.

Note that, on one hand, using $\cos(\Theta_b) \leq 1/2$,

$$\begin{aligned}
|\kappa_0 w_1\bar\rho_1 + w_2\bar\rho_2|^2 &= \kappa_0^2 w_1^2 |\bar\rho_1|^2 + w_2^2|\bar\rho_2|^2 + 2\kappa_0 w_1 w_2 |\bar\rho_1||\bar\rho_2|\cos(\Theta_b) \\
&= c^2\left(\kappa_0^2 w_1^2 + w_2^2 + 2\kappa_0 w_1 w_2\cos(\Theta_b)\right) \\
&\leq c^2\left(\kappa_0^2 w_1^2 + w_2^2 + \kappa_0 w_1 w_2\right).
\end{aligned}$$

[16] Otherwise, interchange indices and do the same proof.

On the other hand,

$$(c(\kappa_0 w_1 + \alpha_2 w_2))^2 = c^2\left(\kappa_0^2 w_1^2 + \alpha_2^2 w_2^2 + 2\alpha_2\kappa_0 w_1 w_2\right).$$

Hence it suffices to show

$$0 \leq w_2(\alpha_2^2 - 1) + \kappa_0 w_1(2\alpha_2 - 1).$$

Solving for α_2, it suffices to show

$$\alpha_2 \geq \sqrt{1 + \frac{\kappa_0 w_1}{w_2} + \left(\frac{\kappa_0 w_1}{w_2}\right)^2} - \frac{\kappa_0 w_1}{w_2}. \tag{6.17}$$

If $w_2/(2w_1) \leq \kappa_0$, then $(\kappa_0 w_1)/w_2 \geq 1/2$. From the graph of $x \mapsto \sqrt{1+x+x^2} - x$, for $x \in [0,\infty)$, it is clear that the right hand side of (6.17) is at most $(\sqrt{7}-1)/2$. Therefore (6.17) is satisfied for any $\alpha_2 \geq (\sqrt{7}-1)/2$. □

Fix $\kappa_0 = w_2/(2w_1)$. Choose a smooth function $\kappa \in \mathscr{C}^1(I_*, [1-\kappa_0, 1])$ such that $\kappa = 1 - \kappa_0$ on the middle third of J_m, denoted J'_m, and $\kappa = 1$ outside J_m. Note that, taking the length of a connected component of $J_m \setminus J'_m$ into account, κ can be chosen such that $|\kappa'| < C_\kappa \kappa_0 |b|$ on I_m. Define (recall that we are assuming $w_2 \leq w_1$):

$$\bar\rho_{1*} := \kappa\bar\rho_1, \text{ and } \bar\rho_{2*} := \bar\rho_2 + (1-\kappa)(w_1/w_2)\bar\rho_1. \tag{6.18}$$

The domain of the definition above is the overlap interval I_*; however, for $j \in \{1,2\}$, we may extend the domain of $\bar\rho_{j*}$ to the interval I_j so that $\bar\rho_{j*} = \bar\rho_j$. This should be clear from the definition of κ and (6.18). We intend to replace $\bar\rho_1, \bar\rho_2$ with $\bar\rho_{1*}, \bar\rho_{2*}$. We will not touch $\tilde\rho_1, \tilde\rho_2$ except to normalize them.

Define a new family

$$\begin{aligned} \mathcal{G}_n^* := {} & (\mathcal{G}_n \setminus \{(I_1,\rho_1),(I_2,\rho_2)\}) \cup \\ & \{(I_1, N(\bar\rho_{1*})), (I_2, N(\bar\rho_{2*})), (I_1, N(\tilde\rho_1)), (I_2, N(\tilde\rho_2)\}. \end{aligned} \tag{6.19}$$

with associated weight measure $w_{\mathcal{G}_n^*}$ that is the same as $w_{\mathcal{G}_n}$ except for the modified standard pairs. For the modified standard pairs, define the new weights by $\bar w_{1*} := w_1 \int_{I_1} |\bar\rho_{1*}|$, $\bar w_{2*} := w_2 \int_{I_2} |\bar\rho_{2*}|$, $\tilde w_1 := w_1 \int_{I_1} |\tilde\rho_1|$, $\tilde w_2 := w_2 \int_{I_2} |\tilde\rho_2|$.

Now we check that the new collection $\mathcal{G}_n^*$ is a standard family equivalent to $\mathcal{G}_n$.

Claim 4 (After cancellation)**.** *We have:*

$$\begin{aligned} w_1\rho_1 + w_2\rho_2 &= w_1\bar\rho_{1*} + w_1\tilde\rho_1 + w_2\bar\rho_{2*} + w_2\tilde\rho_2. && (6.20)\\ H(\bar\rho_{j*}) &\leq a|b|, \forall j \in \{1,2\} && (6.21)\\ \left|\Theta'_{j*}\right| &\leq a|b|, \forall j \in \{1,2\} && (6.22)\\ |\partial_\varepsilon \mathcal{G}_n^*| &\leq C_* \left|\partial_\varepsilon \mathcal{G}_n\right|. && (6.23) \end{aligned}$$

Proof. The equality (6.20) follows from definition. Indeed, (6.18) implies

$$w_1\bar\rho_{1*} + w_2\bar\rho_{2*} = w_1\bar\rho_1 + w_2\bar\rho_2,$$

which in turn implies (6.20). Hence $\mathcal{G}_n^*$ and $\mathcal{G}_n$ are equivalent.

To prove (6.21), note that by construction $\bar\rho_{1*}, \bar\rho_{2*}$ are $\mathscr{C}^1$. Hence it suffices to show $\left|\bar\rho'_{j*}\right| \leq a|b|\,|\bar\rho_{j*}|$.

For $\bar\rho_{1*}$ this condition is easier to check than for $\bar\rho_{2*}$. Let us check, for $\bar\rho_{2*}$ the stronger condition: $|\bar\rho'_{2*}| \leq a|b|\,|\bar\rho_{2*}|$ for a and n large enough. Outside J_m, $\bar\rho_{2*}$ satisfies this condition because $\bar\rho_2$ does. On J_m, differentiating $\bar\rho_{2*}$ and using (6.7) and (6.13) yield,

$$\begin{aligned} |\bar\rho'_{2*}| &\leq c\left(|\Theta'_2| + |1-\kappa|\frac{w_1}{w_2}|\Theta'_1|\right) + \frac{w_1}{w_2}|\kappa'|\,|\rho_1| \\ &\leq a|b|\left(e^{-\lambda n} + \frac{C_\tau}{a}\right)c\left(1 + |1-\kappa|\frac{w_1}{w_2}\right) + c\frac{w_1}{w_2}|\kappa'|. \end{aligned}$$

Now observe that since $K_3|b|^{-1} \le |J_m| \le K_4|b|^{-1}$ and $1 - w_2/(2w_1) \le \kappa \le 1$, we have $|\kappa'| < C_\kappa w_2/(2w_1)|b|$. Also, $|1-\kappa| \le w_2/(2w_1)$. Therefore,

$$\frac{w_1}{w_2}|\kappa'| \le C_\kappa \left(1 + |1-\kappa| \frac{w_1}{w_2}\right)|b|$$

Therefore,

$$\begin{aligned} |\bar\rho'_{2*}| &\le a|b| \left(e^{-\lambda n} + \frac{C_\tau}{a} + \frac{C_\kappa}{a}\right) c \left(1 + |1-\kappa| \frac{w_1}{w_2}\right) \\ &\le a|b| \left(e^{-\lambda n} + \frac{C_\tau}{a} + \frac{C_\kappa}{a}\right) \frac{|\bar\rho_{2*}|}{\alpha_1}, \end{aligned}$$

where the last inequality follows from the left hand side of (6.16). Recall that the left hand side of (6.16) requires no restriction on κ_0. It simply follows from the phase difference satisfying $\cos(\Theta_b) = \cos(\Theta_1 - \Theta_2) \ge 1/4$ and $\alpha_1 < 1/2$. Take a, n large to conclude.

The inequality (6.22) is a consequence of $\left|\bar\rho'_{j*}\right| \le a|b|\,|\bar\rho_{j*}|$ since $\left|\Theta'_{j*}\right| \le \left|\bar\rho'_{j*}\right| / |\bar\rho_{j*}|$.

To prove (6.23), note that

$$|\partial_\varepsilon \mathcal{G}^*_n| \le |\partial_\varepsilon \mathcal{G}_n| + \int_{\partial_\varepsilon I_1} w_1 |\tilde\rho_1| + \int_{\partial_\varepsilon I_2} w_2 |\tilde\rho_2| + \int_{\partial_\varepsilon I_1} w_1 |\bar\rho_{1*}| + \int_{\partial_\varepsilon I_2} w_2 |\bar\rho_{2*}| .$$

The first three terms are simply bounded by $3|\partial_\varepsilon \mathcal{G}_n|$. The last two terms are bounded by $|\partial_\varepsilon \mathcal{G}_n| + \int_{\partial_\varepsilon I_*}(w_1|\bar\rho_{1*}| + w_2|\bar\rho_{2*}|)$. Note that using (6.18), $w_1|\bar\rho_{1*}| + w_2|\bar\rho_{2*}| \le w_1|\bar\rho_1| + w_2|\bar\rho_1|$. Putting all this together and noting that $\bar\rho_1 = \bar\rho_2 = c$, we have

$$\begin{aligned} |\partial_\varepsilon \mathcal{G}^*_n| &\le 3|\partial_\varepsilon \mathcal{G}_n| + |\partial_\varepsilon \mathcal{G}_n| + \int_{\partial_\varepsilon I_*} (w_2|\bar\rho_2| + w_1|\bar\rho_1|) \\ &\le 4|\partial_\varepsilon \mathcal{G}_n| + c\varepsilon(w_1 + w_2) \\ &\le 5c|\partial_\varepsilon \mathcal{G}_n| . \end{aligned}$$

□

Let us check that the total weight of the new standard family is less than the old one.

Claim 5. $\sum_{j \in \mathcal{J}'_{n_b}} w_{\mathcal{G}'_{n_b}}(j) \le e^{-\gamma} w_{\mathcal{G}}$.

Proof. First, note that on J'_m, $\kappa = 1 - \kappa_0$. Therefore on J'_m (and using the right hand side of (6.16)) we have

$$\begin{aligned} w_1|\bar\rho_{1*}| + w_2|\bar\rho_{2*}| &\le w_1(1-\kappa_0)|\bar\rho_1| + w_2\left|\bar\rho_2 + \kappa_0\frac{w_1}{w_2}\bar\rho_1\right| \\ &\le w_1(1-\kappa_0)|\bar\rho_1| + \alpha_2 w_2|\bar\rho_2| + w_1\kappa_0|\bar\rho_1| \\ &= w_1|\bar\rho_1| + \alpha_2 w_2 |\bar\rho_2| . \end{aligned}$$

Outside J'_m, by definition, $w_1|\bar\rho_{1*}| + w_2|\bar\rho_{2*}| \le w_1|\bar\rho_1| + w_2|\bar\rho_2|$.

Then,

$$\begin{aligned} \int_{I_*}(w_1|\bar\rho_{1*}| + w_2|\bar\rho_{2*}|) &\le \sum_m \int_{I_m \setminus J'_m}(w_1|\bar\rho_1| + w_2|\bar\rho_2|) + \int_{J'_m}(w_1|\bar\rho_1| + \alpha_2 w_2|\bar\rho_2|) \\ &= \int_{I_*} w_1|\bar\rho_1| + \sum_m \int_{I_m\setminus J'_m} w_2|\bar\rho_2| + \int_{J'_m}\alpha_2 w_2|\bar\rho_2| \\ &= \int_{I_*} w_1|\bar\rho_1| + \sum_m \int_{I_m} w_2|\bar\rho_2| - (1-\alpha_2)\int_{J'_m} w_2|\bar\rho_2| . \end{aligned}$$

Since $K_1|b|^{-1} \le |I_m| \le K_2|b|^{-1}$ and $(1/3)K_3|b|^{-1} \le |J'_m| \le (1/3)K_4|b|^{-1}$, and $|\bar\rho_2|$ is a constant, Lemma 10 implies that $\int_{J'_m} w_2 |\bar\rho_2| \ge (|J'_m| / |I_m|) \int_{I_m} w_2 |\bar\rho_2| \ge K_3/(3K_2) \int_{I_m} w_2 |\bar\rho_2|$. Therefore,

$$\int_{I_*} (w_1 |\bar\rho_{1*}| + w_2 |\bar\rho_{2*}|) \le \int_{I_*} w_1 |\bar\rho_1| + \sum_m \int_{I_m} w_2 |\bar\rho_2| (1 - (1-\alpha_2)K_3/(3K_2))$$

Set $\alpha'_2 := 1 - (1-\alpha_2)K_3/(3K_2)$. Note that $0 < \alpha'_2 < 1$ and it does not depend on m. Pulling it out of the integral, we have

$$\int_{I_*} (w_1 |\bar\rho_{1*}| + w_2 |\bar\rho_{2*}|) \le \int_{I_*} w_1 |\bar\rho_1| + \alpha'_2 \int_{I_*} w_2 |\bar\rho_2| .$$

Add $\int_{I_1 \setminus I_*} w_1 |\bar\rho_{1*}| + \int_{I_2 \setminus I_*} w_2 |\bar\rho_{2*}| = \int_{I_1 \setminus I_*} w_1 |\bar\rho_1| + \int_{I_2 \setminus I_*} w_2 |\bar\rho_2|$ to the above inequality and apply the same argument (using Lemma 10) to conclude that there exists $0 < \alpha''_2 < 1$ such that

$$\bar w_{1*} + \bar w_{2*} = w_1 \int_{I_1} |\bar\rho_{1*}| + w_2 \int_{I_2} |\bar\rho_{2*}| \le w_1 \int_{I_1} |\bar\rho_1| + \alpha''_2 w_2 \int_{I_2} |\bar\rho_2| = \bar w_1 + \alpha''_2 \bar w_2.$$

Observe that since $\alpha''_2 < 1$ and the rest of the standard pairs in the family were not modified, the total weight of the new standard family is less than the original one. Estimating the total weight more precisely, for large n,

$$\begin{aligned} |\mathcal{G}^*_n| &= |\mathcal{G}_n| - (w_1 + w_2) + \bar w_{1*} + \tilde w_1 + \bar w_{2*} + \tilde w_2 \\ &\le |\mathcal{G}_n| - (w_1 + w_2) + \bar w_1 + \tilde w_1 + \alpha''_2 \bar w_2 + \tilde w_2 \\ &\le |\mathcal{G}_n| - (1 - \alpha''_2)\bar w_2 \\ &\le w_{\mathcal{G}} - (1 - \alpha''_2) C M(n) w_{\mathcal{G}} \\ &\le e^{-\gamma} w_{\mathcal{G}}. \end{aligned}$$

Recall that $w_{\mathcal{G}}$ is the weight of the original standard family consisting of a single standard pair. In the next to last inequality we used the lower bound on w_2. Indeed, by definition of $\bar w_2$, Definition 11, change of variables, we have

$$\begin{aligned} \bar w_2 = w_2 \int_{I_2} |\bar\rho_2| = c|I_2| w_2 \quad &\ge \quad c|I_2| w \int_{I_2} |\rho| \circ h_2 \, |h'_2| \\ &\ge \quad c\Delta w \inf_I |\rho| \, |(I \cap O_{h_2}) \cap h_2(U_\ell)| . \end{aligned}$$

Recalling from Proposition 22 that $M(n) := \min_{\substack{j \in \{1,2\} \\ l \in \{1, \cdots k_\delta\}}} \{ |O_{h_{l,j}} \cap h_{l,j}(U_j)| \}$, we bound the right hand side and arrive at

$$\bar w_2 \ge c\Delta w \inf_I |\rho| \, M(n) := C M(n) w_{\mathcal{G}}$$

□

There is one last issue to resolve. The members of $\mathcal{G}^*_n$ may not satisfy $H(\rho) \le a$. In order to achieve this, we simply iterate the family for a time $C_\delta \ln |b|$ (recall that $n > n_\delta$). Indeed, suppose $(I, \rho) \in \mathcal{G}^*_n$. Note that $H(\rho) \le a|b|$. Following the proof of property (4.15) in Proposition 14, note that after $\tilde n$ more iterations, every image pair $\tilde\rho \in \mathcal{G}^*_{n+\tilde n}$ satisfies:

$$H(\tilde\rho) \le a \left(|b| e^{-\lambda \tilde n} + \frac{D}{a} \right).$$

Now, it is clear that, for $b \ge b_0$ if $\tilde n > \lambda^{-1} \ln \left(|b|(1 - Da^{-1})^{-1} \right) =: \tilde n_b$, then $H(\tilde\rho) \le a$. Finally, note that if C is chosen large enough, then $n_b := C \ln |b|$ dominates $n + \tilde n_b$. We have shown the existence of the standard family $\mathcal{G}^*_{n_b}$ as claimed in Lemma 26. □

Proposition 27. *There exist* $0 < \gamma_1 < \gamma$ *and* $\hat{C} > C_\delta$ *such that for* $\hat{n}_b = \hat{C} \ln |b|$ *with* $|b| \geq b_0$ *(defined by* (6.6)*), for any standard probability family* $\mathcal{G} \in \mathcal{M}_{a,b,B,\varepsilon_0}$*, there exists a standard family* $\mathcal{G}^*_{n_b} \in \mathcal{M}_{a,b,B,\varepsilon_0}$ *equivalent to* $\mathcal{G}_{n_b}$ *such that* $\left|\mathcal{G}^*_{n_b}\right| \leq e^{-\gamma_1}$.

Proof. Recall the assumptions on parameters of Remark 13, Remark 16 and Remark 25. Choose $\delta > 0$ small enough that $3B\delta < 1/4$. Then, by Lemma 15, for large m, $|\partial_{3\delta}\mathcal{G}_m| \leq 3B\delta < 1/4$. It follows that the total weight of the standard pairs of $\mathcal{G}_m$ that have width[17] larger than 3δ is $L \geq 3/4$. Let S denote the total weight of standard pairs that have width $\leq 3\delta$. Note that $|\mathcal{G}| = |\mathcal{G}_m| = S + L = 1$. Fix $\tilde{n}_b > n_{3\delta}$ as in Lemma 26. Using Lemma 26,

$$\left|\mathcal{G}^*_{m+\tilde{n}_b}\right| \leq S + e^{-\gamma} L \leq (S + L) - (1 - e^{-\gamma})L \leq 1 - (1 - e^{-\gamma})3/4 =: e^{-\gamma_1}.$$

Take $n_b = m + \tilde{n}_b$. □

7. Proofs of the main propositions

In this section we give the proofs of Proposition 6 and Proposition 7. We remark that the proof of Proposition 7 is based on the proof of Lemma 7.21 in [1]. First we need the following lemma.

Lemma 28. *There exists* $C > 0$, $\gamma_2 > 0$*, such that if* $\{(I, \rho)\} \in \mathcal{M}_{a,b,B,\varepsilon_0}$ *with the parameters satisfying conditions of Remark 13, Remark 16 and Remark 25, then for every* $n \in \mathbb{N}$,

$$\|\mathscr{L}_b^n \rho\|_{\mathbf{L}^1} \leq C e^{-\frac{\gamma_2}{\ln |b|} n}. \tag{7.1}$$

Proof. The result follows by repeatedly applying Proposition 27 and renormalizing the total weight of the standard family at every step. Indeed, let $\hat{n}_b = \hat{C} \ln |b|$ as in Proposition 27. After k repetitions we have

$$\begin{aligned}\left\|\mathscr{L}_b^{(k)\hat{n}_b} \rho\right\|_{\mathbf{L}^1} &\leq \sum_{j_k \in \mathcal{J}_k} w_{\mathcal{G}^*_{(k)\hat{n}_b}}(j_k) \\ &\leq \sum_{j_{k-1} \in \mathcal{J}_{k-1}} e^{-\gamma_1} w_{\mathcal{G}^*_{(k-1)\hat{n}_b}}(j_{k-1}) \\ &\leq e^{-k\gamma_1}.\end{aligned}$$

This means, for every $m \in \mathbb{N}$ ($m = k\hat{n}_b + r_b$, $0 \leq r_b < \hat{n}_b$),

$$\begin{aligned}\|\mathscr{L}_b^m \rho\|_{\mathbf{L}^1} &\leq \left\|\mathscr{L}_b^{r_b} \mathscr{L}_b^{k\hat{n}_b} \rho\right\|_{\mathbf{L}^1} \leq \left\|\mathscr{L}_0^{r_b} \left|\mathscr{L}_b^{k\hat{n}_b} \rho\right|\right\|_{\mathbf{L}^1} = \left\|\mathscr{L}_b^{k\hat{n}_b} \rho\right\|_{\mathbf{L}^1} \\ &\leq e^{-k\gamma_1} \leq e^{-(\frac{m}{\hat{n}_b} - 1)\gamma_1} \leq e^{\gamma_1} e^{-\frac{\gamma_1}{\hat{n}_b} m} \leq C_{\gamma_1} e^{-\frac{\gamma_1}{\hat{n}_b} m}.\end{aligned}$$

Note that here by $\mathscr{L}_b^0 \rho$, we mean ρ. Set $\gamma_2 = \gamma_1/\hat{C}$, where $\hat{n}_b = \hat{C} \ln |b|$, to conclude. □

Proof of Proposition 6. We will show that there exists a constant $C > 0$, such that for every $b \geq b_0$ (as defined by (6.6)), for every $n \in \mathbb{N}$,

$$\|\mathscr{L}_b^n\|_{\mathscr{C}^1 \to \mathbf{L}^1} \leq C e^{-\frac{\gamma_2}{\ln |b|} n}. \tag{7.2}$$

Suppose $g \in \mathscr{C}^\alpha$. Write $g = g - c + c$, where $c = 1 + |g|_\alpha / a + \sup |g|$ ($|g|_\alpha$ being the α-Hölder constant of g). Note that both $|c|$ and $|g - c| = 1 + |g|_\alpha/a + \sup|g| - g$ are bounded below by 1. Of course, $H(c) = 0$. Calculating as in (6.14), we get

$$\frac{|g(x) - c|}{|g(y) - c|} \leq \frac{|g(x) - g(y)|}{|g(y) - c|} + 1 \leq \frac{|g|_\alpha |x - y|^\alpha}{|g(y) - c|} + 1 \leq e^{\frac{|g|_\alpha |x-y|^\alpha}{|g(y)-c|}}.$$

By the choice of c, it follows that $|g|_\alpha / (|g(y) - c|) \leq a$. Therefore, $H(g - c) \leq a$.

[17] By the width of a standard pair (I, ρ) we mean $|I|$.

Let $g_c = g - c$. Partition the domains of g_c and c into intervals of length $\varepsilon_0 > 0$ (small enough as in Remark 13) and renormalize the restricted functions. Then $g_c = \sum_{j=1}^{L} w_j \rho_j$ and $c = \sum_{k=1}^{L} \tilde{w}_k \tilde{\rho}_k$, where $L = \lceil \varepsilon_0^{-1} \rceil$, $\{\rho_j\}, \{\tilde{\rho}_k\}$ are standard families with parameters a, b, B, ε_0 as defined by Remark 13, Remark 16 and Remark 25; and $\{w_j\}, \{\tilde{w}_k\}$ are their associated weights. Then $g = \sum w_j \rho_j + \sum \tilde{w}_k \tilde{\rho}_k$, where $\sum w_j + \sum \tilde{w}_k = \|g_c\|_{\mathbf{L}^1} + \|c\|_{\mathbf{L}^1}$. Apply Lemma 28 and obtain for $b \geq b_0$, for every $n \in \mathbb{N}$,

$$\begin{aligned}
\|\mathscr{L}_b^n g\|_{\mathbf{L}^1} &\leq \sum_{j=1}^{L} w_j \|\mathscr{L}_b^n \rho_j\|_{\mathbf{L}^1} + \sum_{j=1}^{L} \tilde{w}_k \|\mathscr{L}_b^n \tilde{\rho}_k\|_{\mathbf{L}^1} \\
&\leq C_{\gamma_1} e^{-\frac{\gamma_2}{\ln|b|} n} \left(\sum_{j=1}^{L} w_j + \sum_{j=1}^{L} \tilde{w}_k \right) \\
&\leq C_{\gamma_1} e^{-\frac{\gamma_2}{\ln|b|} n} \left(\|g_c\|_{\mathbf{L}^1} + \|c\|_{\mathbf{L}^1} \right) \\
&\leq C C_{\gamma_1} e^{-\frac{\gamma_2}{\ln|b|} n} \|g\|_{\mathscr{C}^\alpha} \leq C e^{-\frac{\gamma_2}{\ln|b|} n} \|g\|_{\mathscr{C}^1}
\end{aligned}$$

□

Proof of Proposition 7. It suffices to show that $\mathscr{L}_b$ has spectral radius $e^{-r} < 1$ when $b \neq 0$. By the assumptions on $\mathbf{B}$, we know that $\mathscr{L}_b$ has essential spectral radius $e^{-r'} < 1$ and spectral radius at most 1. This means that there are only finitely many eigenvalues outside the disk of radius $e^{-r'}$. Therefore, if we show that there are no eigenvalues on the unit circle, it follows that the spectral radius is strictly less than 1. In turn, this implies the existence of C such that $\|\mathscr{L}_b^n\|_{\mathbf{B}} < Ce^{-rn}$.

Let us show that there are no eigenvalues of $\mathscr{L}_b$ on the unit circle for $b \neq 0$. Fix $b \geq b_0$ and suppose that there exists g and $\lambda \in \mathbb{C}$ satisfying $|\lambda| = 1$ such that $\mathscr{L}_b g = \lambda g$. Since $|\mathscr{L}_b g| \leq \mathscr{L}_0 |g|$, it follows that $|g| \leq \mathscr{L}_0 |g|$. Since, $\int |g| = \int \mathscr{L}_0 |g|$, this implies that $|g| = \mathscr{L}_0 |g|$. Since $|\lambda| = 1$, this means that $|\mathscr{L}_b g| = \mathscr{L}_0 |g|$.

Using the definition of $\mathscr{L}_b$ observe that, for every y, the arguments of the complex numbers $g(x)e^{ib\tau(x)}$ must be equal for all x such that $f(x) = y$. Choose some $k \in \mathbb{N}$ such that $bk > b_0$, where b_0 is as in (6.6). The arguments of the complex numbers $g^k(x)e^{ibk\tau(x)}$ are equal for all $f(x) = y$. This means that $\left|\mathscr{L}_{kb} g^k\right| = \mathscr{L}_0 \left|g^k\right|$. It also means that $\left|\mathscr{L}_{kb}^n g^k\right| = \mathscr{L}_0^n \left|g^k\right|$ for any $n \in \mathbb{N}$ (we could have considered the n-th power from the start of the argument). We have that

$$\int \left|\mathscr{L}_{kb}^n g^k\right| dm = \int \mathscr{L}_0^n \left|g^k\right| dm = \int \left|g^k\right| dm \qquad \text{for all } n \in \mathbb{N}.$$

Using Proposition 6, if $g \in \mathbf{B}$, then the left hand side vanishes as $n \to \infty$ whereas the right hand side is fixed and non-zero. □

References

[1] A. Avila, S. Gouëzel, and J.-C. Yoccoz. Exponential mixing for the Teichmüller flow. *Publ. Math. Inst. Hautes Études Sci.*, (104):143–211, 2006.

[2] V. Baladi and C. Liverani. Exponential decay of correlations for piecewise cone hyperbolic contact flows. *Comm. Math. Phys.*, 314(3):689–773, 2012.

[3] V. Baladi and B. Vallée. Exponential decay of correlations for surface semi-flows without finite Markov partitions. *Proc. Amer. Math. Soc.*, 133(3):865–874 (electronic), 2005.

[4] A. Boyarsky and P. Góra. *Laws of chaos.* Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1997. Invariant measures and dynamical systems in one dimension.

[5] O. Butterley. Area expanding $\mathcal{C}^{1+\alpha}$ suspension semiflows. *Comm. Math. Phys.*, 325(2):803–820, 2014.

[6] O. Butterley and P. Eslami. Exponential mixing for skew products with discontinuities. *Preprint.*

[7] N. Chernov. A stretched exponential bound on time correlations for billiard flows. *J. Stat. Phys.*, 127(1):21–50, 2007.
[8] N. Chernov and D. Dolgopyat. Particle drift in self-similar billiards. *Ergodic Theory Dynam. Systems*, 28(2):389–403, 2008.
[9] N. Chernov and D. Dolgopyat. Brownian Brownian motion. I. *Mem. Amer. Math. Soc.*, 198(927):viii+193, 2009.
[10] N. Chernov and R. Markarian. *Chaotic billiards*, volume 127 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.
[11] N. I. Chernov. Markov approximations and decay of correlations for Anosov flows. *Ann. of Math. (2)*, 147(2):269–324, 1998.
[12] D. Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998.
[13] D. Dolgopyat. On differentiability of SRB states for partially hyperbolic systems. *Invent. Math.*, 155(2):389–449, 2004.
[14] C. T. Ionescu Tulcea and G. Marinescu. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math. (2)*, 52:140–147, 1950.
[15] G. Keller. Generalized bounded variation and applications to piecewise monotonic transformations. *Z. Wahrsch. Verw. Gebiete*, 69(3):461–478, 1985.
[16] C. Liverani. Decay of correlations for piecewise expanding maps. *J. Statist. Phys.*, 78(3-4):1111–1129, 1995.
[17] C. Liverani. On contact Anosov flows. *Ann. of Math. (2)*, 159(3):1275–1312, 2004.
[18] I. Obayashi. Exponential decay of correlations for surface semiflows with an expanding direction. *J. Math. Kyoto Univ.*, 49(2):427–440, 2009.
[19] Y. Peres and B. Solomyak. Absolute continuity of Bernoulli convolutions, a simple proof. *Math. Res. Lett.*, 3(2):231–239, 1996.
[20] M. Pollicott. On the rate of mixing of Axiom A flows. *Invent. Math.*, 81(3):413–426, 1985.
[21] M. Pollicott. On the mixing of Axiom A attracting flows and a conjecture of Ruelle. *Ergodic Theory Dynam. Systems*, 19(2):535–548, 1999.
[22] D. Ruelle. Flots qui ne mélangent pas exponentiellement. *C. R. Acad. Sci. Paris Sér. I Math.*, 296(4):191–193, 1983.
[23] O. Sarig. Subexponential decay of correlations. *Invent. Math.*, 150(3):629–653, 2002.
[24] M. Tsujii. Fat solenoidal attractors. *Nonlinearity*, 14(5):1011–1027, 2001.
[25] M. Tsujii. Decay of correlations in suspension semi-flows of angle-multiplying maps. *Ergodic Theory Dynam. Systems*, 28(1):291–317, 2008.
[26] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)*, 147(3):585–650, 1998.
[27] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.

Peyman Eslami, Dipartimento di Matematica, II Università di Roma (Tor Vergata), Via della Ricerca Scientifica, 00133 Roma, Italy.
E-mail address: eslami@mat.uniroma2.it