Ergod. Th. & Dynam. Sys. (2021), **41**, 981–1024
doi:10.1017/etds.2019.115

Statistical properties for compositions of standard maps with increasing coefficient

ALEX BLUMENTHAL

University of Maryland, College Park, USA
(e-mail: alexb123@math.umd.edu)

(*Received* 10 *May* 2018 *and accepted in revised form* 22 *November* 2019)

Abstract. The Chirikov standard map is a prototypical example of a one-parameter family of volume-preserving maps for which one anticipates chaotic behavior on a non-negligible (positive-volume) subset of phase space for a large set of parameters. Rigorous analysis is notoriously difficult and it remains an open question whether this chaotic region, the stochastic sea, has positive Lebesgue measure for any parameter value. Here we study a problem of intermediate difficulty: compositions of standard maps with increasing coefficient. When the coefficients increase to infinity at a sufficiently fast polynomial rate, we obtain a strong law, a central limit theorem, and quantitative mixing estimates for Holder observables. The methods used are not specific to the standard map and apply to a class of compositions of 'prototypical' two-dimensional maps with hyperbolicity on 'most' of phase space.

Key words: non-autonomous dynamics, non-uniform hyperbolicity, statistical properties of deterministic dynamics
2010 Mathematics Subject Classification: 37C60, 37A25, 37D25 (Primary); 60F05 (Secondary)

1. *Introduction and statement of results*

Let $f : M \to M$ be a smooth dynamical system. In many systems of interest, the dynamics of f does not tend to a stable or periodic equilibrium, as evidenced, e.g., when observables $\phi : M \to \mathbb{R}$ of such systems fluctuate indefinitely, i.e., $\phi \circ f^n(x)$ fluctuates as $n \to \infty$ for a 'large' set of $x \in M$. In such cases, the asymptotic dynamics of the system is best described not by equilibria, but by a 'physical' measure μ for f: an f-invariant probability measure μ on M is called *physical* if for a positive Lebesgue measure set of $x \in M$ (the 'basin' of μ) and any observable $\phi : M \to \mathbb{R}$, we have that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi \circ f^i(x) = \int \phi \, d\mu. \tag{1}$$

Treating the sequence of observations $\{\phi \circ f^i\}_{i\geq 0}$ as a sequence of random variables, (1) above is a strong law of large numbers. Pursuing this interpretation, it is natural to ask whether finer statistical properties hold, e.g.:

- central limit theorems (CLTs) pertaining to the convergence in distribution of $(1/\sqrt{n})\sum_{i=0}^{n-1}(\phi \circ f^i(X) - m)$, where X is distributed in M with some given law ν and $m \in \mathbb{R}$ is a centering constant; and
- decay of correlations, i.e., estimates on the decay of $|\int \phi \circ f^n \cdot \psi \, d\mu - \int \phi \, d\mu \int \psi \, d\mu|$ as $n \to \infty$ for some class of observables ϕ, ψ on M.

These properties are by now classical for maps f with uniform hyperbolicity, e.g., expanding, Anosov, or Axiom A maps (see, e.g., [**23**]). Outside the 'uniform' setting, an extremely important tool in the exploration of statistical properties of deterministic dynamical systems is non-uniformly hyperbolic theory, also known as Pesin theory [**6**, **33**]. Assuming some control on the (typically non-uniform) rate of hyperbolicity, techniques have been developed for use in conjunction with non-uniform hyperbolicity to probe finer statistical properties of deterministic dynamical systems (e.g., the technique of countable Markov extensions, also known as Young towers [**34**]).

1.1. *Difficulties and challenges.* Use of these tools requires establishing non-uniform hyperbolicity, which is notoriously difficult to verify even for maps which 'appear' to be hyperbolic on most (but not all) of phase space. In the volume-preserving category, the difficulties involved are exemplified by the Chirikov standard map family $\{F_L\}_{L>0}$ of volume-preserving maps on the torus $\mathbb{T}^2$ [**13**]. For large L, the map F_L exhibits strong hyperbolicity (i.e., F_L admits a continuous, invariant family of cones with strong expansion) on a large but non-invariant subset of phase space. A key difficulty is that typical orbits will enter a set where cone invariance is violated (e.g., the vicinity of an elliptic fixed point for F_L), and the previously expanding invariant cone is potentially 'twisted' towards the strongly contracting direction, after which all the growth accumulated may be destroyed. Thus, the estimation of Lyapunov exponents (i.e., verifying non-uniform hyperbolicity) amounts to a delicate cancellation problem between 'growth phases', when tangent directions are roughly parallel to expanding directions, and 'contraction phases', during which tangent directions experience contraction.

These challenges are real, as the following results illustrate. Duarte [**20**] showed that for a residual set of large coefficients L, the set of elliptic periodic points for the standard map F_L is approximately $L^{-1/3}$-dense in $\mathbb{T}^2$. On the other hand, Gorodetski [**21**] showed that for a residual set of sufficiently large L, the set of hyperbolic points has full Hausdorff dimension and is $L^{-1/3}$-dense. In particular, taking L large does not alleviate the problem of elliptic dynamics: for all such L, even quite large, chaotic behavior intermingles in a convoluted way with elliptic-type behavior. It stands to reason, then, that distinguishing one regime from the other should be extremely challenging. Indeed, it remains an open question (the standard map conjecture) whether F_L is non-uniformly hyperbolic on a positive-volume subset for *any* value of $L > 0$. We remark, however, that it has recently been proved [**8**] that the standard map F_L is C^r close to non-uniformly hyperbolic maps for a large subset of L.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

1.2. *Results in this paper.* In the interest of studying a problem of intermediate difficulty between the classical uniformly hyperbolic settings and the presently intractable two-dimensional non-uniformly hyperbolic setting exemplified by the standard map, we propose to study compositions of standard maps with *increasing* coefficient. Cone twisting does occur on a positive-volume subset of phase space at each time step and so we contend with many of the same problems described above for systems away from the 'uniform setting'. Indeed, our hypotheses do not preclude the existence of elliptic fixed points for our compositions. Important for our analysis, however, is the fact that increasing the coefficient at each time step both increases the strength of expansion and decreases the size of phase space committing 'cone twisting'—a crucial feature of this model is that a generic trajectory reaches these 'bad' regions at most finitely many times when the increasing coefficients $\{L_n\}$ are inverse summable (see §2.1).

Our main results pertain to the situation when the sequence of coefficients increases sufficiently rapidly: we are able to establish a strong law of large numbers, a central limit theorem, and decay of correlations (Theorems A, B, and C, respectively). Our methods are quite flexible and only rely on the bulk geometry of hyperbolicity on successively larger-volume subsets of phase space. As such, our results apply to a class of volume-preserving maps which are qualitatively similar to the standard map family. For this reason, the techniques of this paper are able to handle effectively 'non-autonomous' dynamics, i.e., dynamics whose behavior is allowed to change with time.

Along the way towards proving the main results, certain 'finite-time' decay of correlations estimates are obtained for standard maps with *fixed* coefficient L, i.e., correlation estimates providing sharp bounds at all times $n \leq N_L$ (in our results, N_L grows as a fractional power of L). This result (formulated as Theorem D) is of independent interest: although it fails to be a true *asymptotic* result, these estimates demonstrate that for large L, the standard map F_L is strongly mixing on a relatively long time scale.

1.3. *Related prior work.* The study of non-autonomous dynamical systems is still in its infancy and many open questions remain. That being said, the statistical properties explored in this paper are closest to those on memory loss for non-autonomous compositions of hyperbolic maps [**4**, **5**, **30**] (see also [**3**]), Sinai billiard systems with slowly moving scatterers [**11**, **31**, **32**], and polynomial loss of memory for intermittent-type maps of the interval with a neutral fixed point at the origin [**1**, **26**]. We have benefited especially from the techniques in [**14**], which studies statistical properties of sequential piecewise-expanding compositions in one dimension.

Pertaining to the Chirikov standard map, there is a large literature on this and related systems (e.g., Schroedinger cocycles), which we do not include here. See, e.g., the citations in [**10**] for a small sampling of such results and some additional discussion.

Random dynamical systems can be thought of as a version of non-autonomous dynamics with some stationarity properties; see, e.g., [**2**, **22**]. Lyapunov exponents of random perturbations of the standard map with large coupling coefficient were studied in [**10**]. We also note [**19**], which established quenched (sample-wise) statistical properties for a large class of stochastic differential equations in both the volume-preserving and dissipative regimens.

The analysis in this paper bears some qualitative similarities with that used in [**9**], which studies Lyapunov exponents and statistical properties of random perturbations of dissipative two-dimensional maps with qualitatively similar features to the Henon map; these results apply as well to the standard map. As it turns out, statistical properties of the corresponding Markov chain can be deduced from *finite-time mixing estimates* for the dynamics, very much in keeping with the spirit of the analysis in the present paper (especially Theorem D).

Lastly, we mention that the techniques in this paper may be useful in future studies of 'bouncing ball' models of Fermi acceleration [**15**, **16**, **18**]. As it turns out, the static wall approximation of bouncing ball models in a potential field gives rise to a Poincaré return map bearing strong qualitative similarities to the standard map (see [**15**] for a detailed derivation) and so it is conceivable that the analysis in this paper may shed insight on open problems related to 'escaping trajectories' for such models.

1.4. *Statement of results.*

1.4.1. *Definition of model.* Let $M_0 \in \mathbb{N}$, $K_0, K_1 > 0$ be fixed constants. Let $L_0 > 0$, which should be thought of as sufficiently large, and let $\{L_n\}$ be a non-decreasing sequence, i.e.,

$$L_0 \leq L_1 \leq L_2 \leq \cdots \leq L_n \leq \cdots .$$

In our results, we will assume that $L_n \to \infty$ at a sufficiently fast polynomial rate in n.

For each $n \geq 1$, let $f_n : \mathbb{T}^1 \to \mathbb{R}$ be a C^3 function for which:

(H1) $\|f_n'\|_{C^1} = \|f_n'\|_{C^0} + \|f_n''\|_{C^0} \leq K_0 L_n$;
(H2) $\mathcal{C}_n := \{\hat{x} \in \mathbb{T}^1 : f_n'(\hat{x}) = 0\}$ is finite, with cardinality $\leq M_0$; and
(H3) for any $n \geq 1$, $x \in \mathbb{T}^1$, we have $|f_n'(x)| \geq L_n K_1^{-1} d(x, \mathcal{C}_n)$.

We will consider the non-autonomous composition of the maps $F_n : \mathbb{T}^2 \to \mathbb{T}^2$ defined by setting

$$F_n(x, y) = \begin{pmatrix} f_n(x) - y \pmod 1 \\ x \end{pmatrix}.$$

Above, (mod 1) refers to the projection $\mathbb{R} \to \mathbb{T}^1$ defined by $x \mapsto x - \lfloor x \rfloor$†, having abused notation somewhat and parametrized $\mathbb{T}^1$ by $[0, 1)$. We will continue to use this convention throughout the paper.

We note that conditions (H1)–(H3) are satisfied by the family $f_n(x) := L_n \sin(2\pi x) + 2x$, in which case F_n is (up to conjugation by a linear toral automorphism) the standard map with coefficient L_n. These conditions are also satisfied for the family $f_n(x) := L_n \psi(x) + a_n$, where $\{a_n\} \subset [0, 1)$ is any subsequence and $\psi : \mathbb{T}^1 \to \mathbb{R}$ is a map satisfying some C^3-generic conditions—details are left to the reader. The hypotheses (H1)–(H3) are similar to those for [**10**, Theorem 1].

For $n \geq m \geq 1$, we write $F_m^n = F_n \circ F_{n-1} \circ \cdots \circ F_m$, and write $F^n = F_1^n$. We adopt the conventions $F_n^{n-1} = \mathrm{Id}$, $F^0 = \mathrm{Id}$.

† Here for $x \in \mathbb{R}$ we define the floor function $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

1.4.2. *Results.* Our first result is a strong law of large numbers, which can be thought of as an ergodicity-type property for the non-autonomous compositions $\{F^n\}$.

THEOREM A. *Let $\alpha \in (0, 1]$. Assume that L_0 is sufficiently large, depending on the system parameters K_0, K_1, M_0 and on α. Let $\phi : \mathbb{T}^2 \to \mathbb{R}$ be α-Holder continuous with $\int \phi \, d\,\mathrm{Leb}_{\mathbb{T}^2} = 0$.*

(a) *If $N^2 L_N^{-\alpha/(3\alpha+4)} \to 0$, then $(1/N)\sum_{i=0}^N \phi \circ F^i \to 0$ in L^2.*

(b) *If $N^{4+\epsilon} L_N^{-\alpha/(3\alpha+4)} \to 0$ for some arbitrary $\epsilon > 0$, then $(1/N)\sum_{i=0}^N \phi \circ F^i \to 0$ Lebesgue almost everywhere.*

Example 1.1. Fix $\alpha \in (0, 1]$, $p > 0$. Define $L_n = \max\{L_0, n^p\}$ for some $p > 1$ and L_0 sufficiently large. Then Theorem A(a) holds when $p > \alpha^{-1}(6\alpha + 8)$ and Theorem A(b) holds when $p \geq \alpha^{-1}(12\alpha + 16)$. The exponent of n is minimized when $\alpha = 1$ (i.e., ϕ is Lipschitz); here $p \geq 14$ suffices for (a) and $p \geq 32$ for (b).

Next is a central limit theorem for Holder observables.

THEOREM B. *Let $\alpha \in (0, 1]$. Let L_0 be as in Theorem A and, additionally, assume that*

$$\lim_{N\to\infty} N^8 L_N^{-\alpha/(3\alpha+4)} = 0.$$

Let ϕ be an α-Holder continuous function on $\mathbb{T}^2$ for which $\int \phi \, d\,\mathrm{Leb}_{\mathbb{T}^2} = 0$. Let X be a uniformly distributed $\mathbb{T}^2$-valued random variable. Then $(1/\sigma\sqrt{N})\sum_{i=0}^N \phi \circ F^i(X)$ converges in distribution to a standard Gaussian as $N \to \infty$ with

$$\sigma^2 = \int \phi(x, y)^2 \, dx \, dy + 2 \int \phi(x, z)\phi(z, y) \, dx \, dy \, dz,$$

provided that $\sigma > 0$. Moreover, we have $\sigma = 0$ if and only if $\phi(x, y) = \psi(x) - \psi(y)$ for some continuous $\psi : \mathbb{T}^1 \to \mathbb{R}$.

These conditions are satisfied for L_n as in Example 1.2 when $p > 8\alpha^{-1}(3\alpha + 4)$. The asymptotic variance σ appearing in Theorem B comes from an appropriate interpretation of the 'singular' limit of the maps F_n as $n \to \infty$. The condition $\phi(x, y) = \psi(x) - \psi(y)$ has the connotation of a coboundary condition for this singular limit. See the discussion in §3.1 (in particular Remark 3.2 and Lemma 3.3) for more details. Theorem B is proved in §5. In the setting of Example 1.2, Theorem B holds when $p > 8(3\alpha + 4)/\alpha$; the result is optimal when $\alpha = 1$, in which case $p > 56$ suffices.

Finally, we present a decay of correlations estimate for the compositions $\{F^n\}$.

THEOREM C. *Fix $\eta \in (1/2, 1)$. Let L_0 be sufficiently large in terms of the system parameters and η. Assume that $\sum_n L_n^{-(1/2)(1-\eta)} < \infty$. Then there is a constant $C = C(K_0, K_1, M_0, \eta)$ for which the following holds.*

Let $\alpha \in (0, 1]$ and let φ, ψ be α-Holder continuous functions on $\mathbb{T}^2$. Then

$$\left| \int \psi \circ F^n \cdot \varphi - \int \varphi \int \psi \right| \leq C \|\psi\|_\alpha \|\varphi\|_\alpha \max\left\{ L_{\lfloor n/2 \rfloor}^{1-2\eta}, \left(\sum_{i=\lfloor n/8 \rfloor}^{\infty} L_i^{-(1/2)(1-\eta)} \right)^{\alpha/(\alpha+2)} \right\}$$

for all $n \geq 0$.

Above, all integrals $\int$ are with respect to $\mathrm{Leb}_{\mathbb{T}^2}$ and we have written

$$[\varphi]_\alpha := \sup_{p,q\in\mathbb{T}^2} \frac{|\varphi(p)-\varphi(q)|}{d_{\mathbb{T}^2}(p,q)^\alpha} \quad \text{and} \quad \|\varphi\|_\alpha = \|\varphi\|_{C^0} + [\varphi]_\alpha$$

for $\alpha \in (0, 1]$ the Holder moduli and norms, respectively, and $d_{\mathbb{T}^2}$ is the geodesic distance on $\mathbb{T}^2$ endowed with the flat geometry of $\mathbb{R}^2/\mathbb{Z}^2$.

Example 1.2. Fix $\eta \in (1/2, 1)$. Define $L_n = \max\{L_0, n^p\}$ for some $p > 4$ and L_0 sufficiently large. In particular, $\sum_n L_n^{-(1/2)(1-\eta)} < \infty$ if and only if $p > 2/(1-\eta)$. One obtains that the $\max\{\cdots\}$ term on the right-hand side is

$$\leq \text{Const. } \|\psi\|_\alpha \|\phi\|_\alpha n^{\max\{p(1-2\eta),(\alpha/(\alpha+2))(1-(1/2)p(1-\eta))\}}.$$

The exponent of n is optimized at $\eta = (3\alpha p + 4p - 2\alpha)/(5\alpha p + 8p)$ at the value $(4-p)\alpha/(5\alpha+8)$ (valid since here $p > 2/(1-\eta)$ reduces to $p > 4$, which has been assumed), leading to the estimate

$$\leq \text{Const. } \|\psi\|_\alpha \|\phi\|_\alpha n^{-(\alpha(p-4)/(5\alpha+8))}.$$

The exponent of n is minimized when $\alpha = 1$, in which case decay of correlations is summable if $p > 17$.

1.4.3. *Finite-time decay of correlations estimates for fixed-coefficient standard maps.* Our estimates in this paper can also be used to obtain the following *finite-time* decay of correlations estimate for Holder observables.

THEOREM D. *Let* $\alpha \in (0, 1]$ *and let* L *be sufficiently large in terms of* α. *Let* ϕ, ψ *be* α*-Holder-continuous functions on* $\mathbb{T}^2$. *Then*

$$\left| \int \phi \circ F_L^n \cdot \psi - \int \phi \int \psi \right| \leq C \|\phi\|_\alpha \|\psi\|_\alpha \cdot n L^{-\alpha/(3\alpha+4)}$$

for all $n \geq 2$, *where* $C = C(\alpha) > 0$ *is a constant independent of* L, ψ, ϕ.

For each fixed $L > 0$, Theorem D provides a non-trivial upper bound on correlations for times $n \ll L^{\alpha/(3\alpha+4)}$, and thus gives information on the mixing properties of the standard map in the so-called anti-integrable limit. Like before, the result is strongest at $\alpha = 1$.

1.4.4. *Plan for the paper.* We collect preliminaries and basic hyperbolicity results in §2, with an emphasis on the geometry of iterates of curves roughly parallel to the strongly expanding direction (called *horizontal curves*) for the dynamics.

In §3 we develop finite-time mixing estimates for the composition $\{F^n\}$; this verifies Theorem D and also lets us provide a statistical description of the 'singular' limit of the maps F_n as $n \to \infty$. In §4 we deduce the strong law (Theorem A) and in §5 we prove the central limit theorem (Theorem B).

The proof of Theorem C, carried out in §§6 and 7, is logically independent of §§3–5; indeed, it should not be surprising that the 'finite-time mixing' estimates in these sections do not yield the long-time asymptotic correlation estimate in Theorem C. The proof of the latter requires a more careful study of the 'shape' of iterates of small, sufficiently nice sets $S \subset \mathbb{T}^2$. This is carried out in §6 and the proof of Theorem C is completed in §7.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

1.4.5. *Notation and conventions.* We parametrize $\mathbb{T}^1$ as $[0, 1)$ throughout the paper. The torus $\mathbb{T}^2$ carries the flat geometry of $\mathbb{R}^2/\mathbb{Z}^2$ and we identify all tangent spaces with the same copy of $\mathbb{R}^2$. We write $d_{\mathbb{T}^1}$, $d_{\mathbb{T}^2}$ for the geodesic metrics on $\mathbb{T}^1$, $\mathbb{T}^2$ respectively.

We repeatedly use big-O notation: a quantity $\beta \in \mathbb{R}$ is said to be $O(\kappa)$ for some $\kappa > 0$, written $\beta = O(\kappa)$, if there is a constant $C > 0$, depending only on the *system parameters* K_0, K_1, M_0, for which $|\alpha| \leq C\kappa$. Similarly, the letter C is reserved for any positive constant depending only on the parameters K_0, K_1, M_0.

We write Leb or $\text{Leb}_{\mathbb{T}^2}$ for the Lebesgue measure on $\mathbb{T}^2$, although, unless otherwise stated, any integral $\int$ over $\mathbb{T}^2$ should be assumed to be with respect to Lebesgue. When $\gamma \subset \mathbb{T}^2$ is a C^2 curve, we write Leb_γ for the (unnormalized) induced Lebesgue measure on γ.

If $p \in \mathbb{T}^2$, we typically write $p_0 := p$ and $p_n := F^{n-1} p_0$; in general, objects with a subscript n should be thought of as belonging to the 'domain' of F_n. Given $\gamma \subset \mathbb{T}^2$, we likewise write $\gamma_0 = \gamma$ and $\gamma_n = F^{n-1}\gamma$.

Lastly, the parameter $L_0 > 1$ is assumed fixed and will be taken sufficiently large in a finite number of places in the proofs to come. Whenever L_0 is enlarged, it is done so in a way that depends only on the system parameters K_0, K_1, M_0, the Holder exponent α, and the auxiliary parameter η introduced below in §2.1.1.

From this point forward, we will assume that $\{L_n\}$, $\{f_n\}$ *are as in (H1)–(H3) and that* $\{L_n\}$ *is a non-decreasing sequence.*

2. *Predominant hyperbolicity*

For all large n, the maps F_n are *predominantly hyperbolic*, which is to say that the derivative maps dF_n exhibit strong expansion along roughly horizontal directions on an increasingly large (but non-invariant) proportion of phase space. Our purpose in this section is to make this idea precise and collect some preliminary results.

In §2.1 we essentially deal with hyperbolicity on the linear level: when $L_n \to \infty$ sufficiently fast, we show that the compositions $\{F^n\}$ possess non-zero (in fact, infinite) Lyapunov exponents at Lebesgue-almost every point. On the other hand, the *rate* at which this hyperbolicity is expressed is non-uniform across phase space and so, in analogy with standard non-uniformly hyperbolic theory in the stationary setting, we develop in §2.1 the notion of *uniformity set* to control this non-uniformity.

In §§2.2 and 2.3, we consider the nonlinear picture: the time evolution of curves roughly parallel to the unstable (horizontal) direction. The basic idea is that sufficiently long 'horizontal curves' proliferate rapidly through phase space: this is precisely the mixing mechanism one anticipates when working with this model and is used repeatedly throughout the paper. Standard hyperbolic theory preliminaries are given in §2.2, while in §2.3 this mixing mechanism is more precisely laid out in the form of a mixing estimate for Lebesgue measure supported on a sufficiently long horizontal curve.

2.1. *Predominant hyperbolicity of the maps F_n.*

Let us begin by identifying subsets of phase space where the maps F_n exhibit uniformly strong hyperbolicity. For $L > 0$ and

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

$n \geq 1$, define the *critical strips*

$$S_{n,L} = \{(x, y) \in \mathbb{T}^2 : d(x, \mathcal{C}_n) \leq K_1 L_n^{-1} L\}$$

and note that by (H3), for $(x, y) \notin S_{n,L}$, we have $|f_n'(x)| \geq L$. For each n, outside $S_{n,L}$, we have that F_n is strongly expanding in the horizontal direction: to wit, for any L sufficiently large ($L \geq 10$ will do for our purposes) and any $n \geq 1$, $p \notin S_{n,L}$, the cone

$$C_h = \{v = (v_x, v_y) \in \mathbb{R}^2 : |v_y| \leq \tfrac{1}{10}|v_x|\}$$

is preserved by $(dF_n)_p$ and all vectors in the cone are expanded by a factor $\geq L/4$.

In particular, observe that $\text{Leb}(S_{n,L}) \approx L/L_n$. Thus, for fixed L, the proportion of phase space $\mathbb{T}^2 \setminus S_{n,L}$ on which F_n preserves and expands C_h increases as n increases. When the sequence L_n increases sufficiently rapidly, this implies an infinite Lyapunov exponent almost everywhere, as follows.

LEMMA 2.1. *Assume that* $\sum_{n=1}^{\infty} L_n^{-1} < \infty$. *Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \|dF_p^n\| = \infty \tag{2}$$

for Leb-*almost every* $p \in \mathbb{T}^2$.

Proof. For each $L > 0$, we have

$$\sum_{n=1}^{\infty} \text{Leb}(F^{n-1})^{-1} S_{n,L} = \sum_{n=1}^{\infty} \text{Leb}\, S_{n,L} \leq 2K_1 M_0 L \sum_{n=1}^{\infty} L_n^{-1} < \infty.$$

The Borel–Cantelli lemma thus applies to the sequence of sets $\{(F^{n-1})^{-1} S_{n,L}\}_{n\geq 1}$ and so the set $S_L = \{p \in \mathbb{T}^2 : F^{n-1} p \in S_{n,L}$ *infinitely often*$\}$ has zero Lebesgue measure. Taking $S = \bigcup_{N=1}^{\infty} S_N$, it is now simple to check that (2) holds for all $p \in \mathbb{T}^2 \setminus S$. □

Let us emphasize, however, that the limit (2) is highly non-uniform in x, due to the fact that the critical strips $S_{n,L}$ have positive mass for all $n \geq 1$. We encode this non-uniformity in a way analogous to that of *uniformity sets* (alternatively called *Pesin sets*) for non-uniformly hyperbolic dynamics.

2.1.1. *Construction of uniformity sets for the composition* $\{F^n\}$. For our purposes in this paper, it is expedient to 'fatten' the critical strips $S_{n,L}$ as follows. Let $\eta \in (0, 1)$ and, for $n \geq 1$, define

$$B_n(\eta) = \{(x, y) \in \mathbb{T}^2 : d(x, \mathcal{C}_n) \leq 2K_1 L_n^{-1+\eta}\}.$$

For $p = (x, y) \notin B_n(\eta)$, we have $|f_n'(x)| \geq 2L_n^{\eta}$. In particular, for such p, we have that $(dF_n)_p$ preserves the cone C_h and expands tangent vectors in C_h by a factor $\geq L_n^{\eta}$. The parameter η dictates the proportion of expansion we recover in $(B_n(\eta))^c$ and hence the tradeoff: the larger η, the more expansion we demand away from the bad sets $B_n(\eta)$, but the larger the bad sets $B_n(\eta)$ become. We note that η appears throughout the paper and is often fixed in advance; as such, for simplicity we often write $B_n = B_n(\eta)$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

For $p \in \mathbb{T}^2$, define

$$\tau(p) := 1 + \max\{m \geq 1 : F^{m-1} p \in B_m\}$$
$$= \min\{k \geq 1 : F^{n-1} p \notin B_n \text{ for all } n \geq k\};$$

in particular, for a given orbit $\{p_n = F^{n-1} p\}$, the derivative mapping $(dF_n)_{p_n}$ is uniformly expanding along the horizontal cone C_h for all $n \geq \tau(p)$. In this way, the sets

$$\Gamma_N = \{\tau(p) \leq N\}$$

can be thought of as *uniformity sets* for the composition $\{F^n\}_{n \geq 1}$. Repeating the proof of Lemma 2.1 yields the following.

LEMMA 2.2. *Fix* $\eta \in (0, 1)$ *and assume that* $\sum_{n=1}^{\infty} L_n^{-1+\eta} < \infty$. *Then* $\tau < \infty$ *almost surely and* $\bigcup_N \Gamma_N$ *has full Lebesgue measure.*

Indeed, we have the estimate

$$\text{Leb}\{\tau > N\} \leq \sum_{n=N}^{\infty} \text{Leb}(F^{n-1})^{-1} B_n = \sum_{n=N}^{\infty} \text{Leb } B_n = O\bigg(\sum_{n=N}^{\infty} L_n^{-1+\eta}\bigg).$$

2.2. *Horizontal curves.* Curves roughly parallel to unstable directions, sometimes called u-curves in the literature, are an effective and well-used tool for describing the mixing mechanism of hyperbolic dynamical systems: the elongation of such curves under successive applications of hyperbolic dynamics leads to their proliferation through phase space, resulting in mixing. These ideas are standard for (autonomous) smooth dynamical systems exhibiting hyperbolicity; see, e.g., [**17**, **27**, **29**].

In the setting of this paper, *horizontal curves* play the role of u-curves. Although much of the material in this section is standard for iterates of a single map, we note that the maps F_n in our compositions become more singular as n increases. So, it is important to ensure that the necessary estimates (e.g., distortion control) do not worsen with n. For this reason, we re-prove below in §2.2 what are otherwise standard results in hyperbolic dynamics.

The point of departure is an identification of a class of curves 'roughly parallel to unstable (horizontal) directions'.

Definition 2.3. A *horizontal curve* is a connected C^2 curve $\gamma \subset \mathbb{T}^2$ with the property that $\gamma = \{(x, h_\gamma(x)) : y \in I_\gamma\}$ for some (open, proper) subarc $I_\gamma \subset \mathbb{T}^1$ and some Lipschitz continuous function $h_\gamma : I_\gamma \to \mathbb{T}^1$ with $\|h'_\gamma\|_{C^0} \leq 1/10$. Note that $I_\gamma = (0, 1)$ is allowed.

The plan is as follows. In Lemma 2.4 below we describe the evolution of horizontal curves under successive iterates of our non-autonomous compositions $\{F_m^n, m \leq n\}$ when these curves are assumed to avoid the critical strips B_n for each n. Lemma 2.5 is a distortion estimate between trajectories evolving on the same horizontal curve. Finally, Lemma 2.7 considers the time evolution of *sufficiently long* horizontal curves which are allowed to meet bad sets.

The following is a description of the geometry of successive images of horizontal curves which do not meet the bad sets $\{B_n\}$.

LEMMA 2.4. (Forward graph transform) *Fix $\eta \in (0, 1)$; then the following holds whenever L_0 is sufficiently large (depending on η). Let $N \geq 1$ and let $\gamma \subset \mathbb{T}^2$ be a C^2 horizontal curve of the form $\gamma = \gamma_N = \text{graph}\, g_N = \{(x, g_N(x)) : x \in I_N\}$, where $I_N \subset \mathbb{R}$ and $g_N : I_N \to \mathbb{R}$ is a C^2 function for which $\|g_N'\|_{C^0} \leq 1/10$ and $\|g_N''\|_{C^0} \leq 1$.*

Let $n > N$ and assume that for all $N \leq k \leq n-1$, we have that

$$F_N^{k-1}(\gamma) \cap B_k = \emptyset.$$

Then, for each $N \leq k \leq n$, we have that $\gamma_k = F_N^{k-1}(\gamma)$ is a horizontal curve of the form $\text{graph}\, g_k = \{(x, g_k(x)) : x \in I_k\}$ for an interval $I_k \subset \mathbb{T}^1$ and a C^2 function $g_k : I_k \to \mathbb{T}^1$ for which:

(a) *we have the bounds*

$$\|g_k'\|_{C^0} \leq L_{k-1}^{-\eta} \quad \text{and} \quad \|g_k''\|_{C^0} \leq 2K_0 L_{k-1}^{-3\eta+1}; \quad \text{and}$$

(b) *for any $p_N^i \in \gamma$, $i = 1, 2$, writing $F_N^{k-1} p_N^i = p_k^i$, we have that*

$$\|p_k^1 - p_k^2\| \leq L_k^{\eta} \|p_{k+1}^1 - p_{k+1}^2\|.$$

Proof. The proof is a standard graph transform argument, which we recall here. It suffices to describe the induction step, that is, the procedure for obtaining g_{k+1} from g_k for $N \leq k \leq n-1$.

To start, define the 'lifted' map $\tilde{F}_k : \mathbb{T}^2 \to \mathbb{R} \times \mathbb{T}^1$ by setting $\tilde{F}_k(x, y) = (f_k(x) - y, x)$ (that is, without the '(mod 1)' in the first coordinate). Projecting $\tilde{F}_k(x, g_k(x))$ to the first coordinate results in a function $\tilde{f}_k : I_k \to \mathbb{R}$ of the form $\tilde{f}_k(x) = f_k(x) - g_k(x)$.

Since $\gamma_k \cap B_k = \emptyset$, we have $|f_k'| \geq 2L_k^{\eta}$ on I_k and so $|\tilde{f}_k'| \geq 2L_k^{\eta} - 1/10 > 0$ (on taking $L_0 > 1$). In particular, $\tilde{f}_k : I_k \to \mathbb{R}$ is invertible on its image $\tilde{I}_{k+1}$. Defining $I_{k+1} \subset \mathbb{T}^1$ to be the projection of $\tilde{I}_{k+1}$ to $\mathbb{T}^1$, we define $g_{k+1} : I_{k+1} \to \mathbb{T}^1$ to be the (uniquely determined) mapping for which $g_{k+1}(\tilde{f}_k(x) \pmod 1) = x$ for all $x \in I_k$. This completes the description of the induction step.

The estimates in item (a) are now derived from the implicit derivatives

$$g_k'(x) = \frac{1}{(f_{k-1}' - g_{k-1}')(g_k(x))} \quad \text{and} \quad g_k''(x) = -\frac{(f_{k-1}'' - g_{k-1}'')}{(f_{k-1}' - g_{k-1}')^3}(g_k(x)).$$

The estimate in (b) follows from the bound $|(\tilde{f}_k)'| \geq 2L_k^{\eta} - 1/10 \geq L_k^{\eta}$. All estimates require taking L_0 sufficiently large depending on η. □

Next we obtain distortion estimates along forward iterates of horizontal leaves in the setting of Lemma 2.4.

LEMMA 2.5. *Assume the setting of Lemma 2.4. Let $p_N^i \in \gamma$, $i = 1, 2$, and write $p_n^i = F_N^{n-1} p_N^i$. Then*

$$\left| \log \frac{\|(dF_N^{n-1})_{p_N^1}|_{T\gamma}\|}{\|(dF_N^{n-1})_{p_N^2}|_{T\gamma}\|} \right| = O(L_N^{1-2\eta} \|p_n^1 - p_n^2\|).$$

Remark 2.6. The above bound is quite poor unless $\eta \in (1/2, 1)$, which is why in Theorem C, and indeed throughout the paper, we will work exclusively in the setting where

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

$\eta \in (1/2, 1)$. Of course, the lower the value of η, the stronger the decay of correlations estimate in Theorem C. It is likely that lowering η is possible: one way to accommodate the distortion estimate in Lemma 2.5 is to further subdivide images of the curve γ into pieces of size $\ll L_n^{1-2\eta}$.

Proof of Lemma 2.5. Write $p_k^i = F_N^{k-1}(p_N^i) = (x_k^i, y_k^i)$. Let $\gamma_k = F_N^{k-1}(\gamma)$ and $g_k : I_k \to \mathbb{T}^1$, $I_k \subset \mathbb{T}^1$ be as in Lemma 2.4. Then

$$\|(dF_k)_{p_k^i}|_{T\gamma_k}\| = \sqrt{\frac{1 + (g_{k+1}'(x_{k+1}^i))^2}{1 + (g_k'(x_k^i))^2}} |f_k'(x_k^i) - g_k'(x_k^i)|$$

and so

$$\log \frac{\|(dF_N^{n-1})_{p_N^1}|_{T\gamma}\|}{\|(dF_N^{n-1})_{p_N^2}|_{T\gamma}\|} = \frac{1}{2}\bigg(\log \frac{1 + (g_N'(x_N^2))^2}{1 + (g_N'(x_N^1))^2} + \log \frac{1 + (g_n'(x_n^1))^2}{1 + (g_n'(x_n^2))^2}\bigg) + \sum_{k=N}^{n-1} \log \frac{f_k'(x_k^1) - g_k'(x_k^1)}{f_k'(x_k^2) - g_k'(x_k^2)}. \tag{3}$$

For the first two terms, observe that for $\beta_1, \beta_2 \in [0, \infty)$, we have the elementary bound $|\log(1 + \beta_1) - \log(1 + \beta_2)| \le |\beta_1 - \beta_2|$ and so for $k = N, n$ we have

$$\left|\log \frac{1 + (g_k'(x_k^1))^2}{1 + (g_k'(x_k^2))^2}\right| \le |(g_k'(x_k^1))^2 - (g_k'(x_k^2))^2| \le 2|g_k'(x_k^1) - g_k'(x_k^2)| \le 2 \operatorname{Lip}(g_k'') \cdot |x_k^1 - x_k^2|.$$

Applying the expansion estimate along images of horizontal curves as in Lemma 2.4(a),

$$|x_k^1 - x_k^2| \le L_k^{-\eta} |x_{k+1}^1 - x_{k+1}^2| \le \cdots \le L_k^{-\eta} \cdots L_{n-1}^{-\eta} |x_n^1 - x_n^2| \tag{4}$$

and the estimate $\operatorname{Lip}(g_k'') \le 2K_0 L_k^{1-3\eta}$ coming from Lemma 2.4, we obtain the following upper bound for the first two terms in (3):

$$\operatorname{Lip}(g_N'') \cdot |x_N^1 - x_N^2| + \operatorname{Lip}(g_n'') \cdot |x_n^1 - x_n^2| \le 2K_0 L_N^{1-3\eta}(1 + L_N^{-(n-N)\eta})|x_n^1 - x_n^2|.$$

Thus, these terms are $O(L_N^{1-3\eta})$.

We now estimate the summation term in (3). With $\tilde{f}_k = f_k - g_k : I_k \to \mathbb{R}$ as in the proof of Lemma 2.4, we have that

$$|\log \tilde{f}_k'(x_k^1) - \log \tilde{f}_k'(x_k^2)| \le \frac{\sup_{\zeta \in I_k} |\tilde{f}_k''(\zeta)|}{\inf_{\zeta \in I_k} |\tilde{f}_n'(\zeta)|} \cdot |x_k^1 - x_k^2| \le 2K_0 L_k^{1-\eta} |x_k^1 - x_k^2|.$$

Applying (4) and collecting,

$$\begin{aligned}\left|\log \frac{(\tilde{f}_N^{n-1})'(x_N^1)}{(\tilde{f}_N^{n-1})'(x_N^2)}\right| &\le 2K_0 \bigg(\sum_{k=N}^{n-1} \frac{L_k^{1-\eta}}{L_k^\eta L_{k+1}^\eta \cdots L_{n-1}^\eta} \bigg) |x_n^1 - x_n^2| \\ &\le 2K_0 L_N^{1-2\eta} \bigg(\sum_{k=N}^{n-1} L_N^{-(n-1-k)\eta} \bigg) |x_n^1 - x_n^2| \\ &\le 3K_0 L_N^{1-2\eta} \|p_n^1 - p_n^2\|\end{aligned}$$

when L_0 is taken suitably large. This completes the estimate. □

The above results describe the dynamics of a horizontal curve γ which 'avoids' the bad sets $\{B_n\}$ for some amount of time. On the other hand, if a given horizontal curve is allowed to meet the bad sets along its trajectory, then we lose control over the geometry where these iterates meet bad sets. Below, we describe an algorithm for excising those parts of a curve which fall into the bad set and describe the geometry of the parts of γ with a 'good' trajectory.

We say that a horizontal curve γ is *fully crossing* if $I_\gamma = (0, 1)$ (all notation here and below is as in Definition 2.3).

LEMMA 2.7. *Fix $\eta \in (1/2, 1)$. Let γ be a horizontal curve. Then, for any $m \geq 1$, $k \geq m$, there are a set $\mathcal{B}^k_m(\gamma) \subseteq \gamma$ and a partition (possibly empty) $\bar{\Gamma}^k_m(\gamma)$ of $F^k_m(\gamma \setminus \mathcal{B}^k_m(\gamma))$ into fully crossing curves with the following properties.*

(a) *For any $\bar{\gamma} \in \bar{\Gamma}^k_m(\gamma)$, we have $\|h'_{\bar{\gamma}}\|_{C^0} \leq L_k^{-\eta}$.*

(b) *We have the estimate†*

$$\mathrm{Leb}_\gamma(\mathcal{B}^k_m(\gamma)) = O\bigg(\sum_{i=m}^{k} L_i^{-1+\eta}\bigg).$$

(c) *For any $\bar{\gamma} \in \bar{\Gamma}^k_m(\gamma)$ and any $p, p' \in (F^k_m)^{-1}\bar{\gamma}$, we have*

$$\frac{\|(dF^k_m)_p|_{T\gamma}\|}{\|(dF^k_m)_{p'}|_{T\gamma}\|} = 1 + O(L_m^{1-2\eta}).$$

When $k = m$, we write $\bar{\Gamma}_m(\gamma) = \bar{\Gamma}^m_m(\gamma)$, $\mathcal{B}_m(\gamma) = \mathcal{B}^m_m(\gamma)$ for short.

Observe that Lemma 2.7 is inherently limited in two ways: (i) it is a *finite-time result*: for a given curve γ and fixed $m \geq 1$, we have $\mathcal{B}^k_m(\gamma) = \gamma$ for all k sufficiently large; and (ii) if γ is too short, then we may even have $\gamma = \mathcal{B}_m(\gamma)$.

Proof of Lemma 2.7. Below, $\tilde{F}_m : \mathbb{T}^2 \to \mathbb{R} \times \mathbb{T}^1$ is as defined in the proof of Lemma 2.4. To start, we define $\bar{\Gamma}_m(\gamma)$, $\mathcal{B}_m(\gamma)$ as follows.

For each connected component γ_i, $1 \leq i \leq K$, of $\gamma \setminus B_m$, the image $\tilde{\gamma}_i = \tilde{F}_m(\gamma_i)$ is of the form graph $\tilde{h}_i$, where $\tilde{h}_i : \tilde{I}_i \to \mathbb{T}^1$ for an interval $\tilde{I}_i \subset \mathbb{R}$ of the form $(a_i - r_i, b_i + s_i)$, where $a_i, b_i \in \mathbb{Z}$, $r_i, s_i \in [0, 1)$.

If $a_i = b_i$, i.e., $\tilde{F}_m(\gamma_i) \subset [a, a + 1) \times \mathbb{T}^1$ for some $a \in \mathbb{Z}$, then we set $\bar{\Gamma}_m(\gamma) = \emptyset$ and $\mathcal{B}_m(\gamma) = \gamma$, checking that if this is indeed the case, then $\mathrm{Leb}_\gamma(\gamma) = O(L_m^{-1+\eta})$ follows.

When $a_i < b_i$, we define $\bar{\Gamma}_m(\gamma)$ to be the collection of curves of the form graph $\tilde{h}_i(\cdot + l)$ (projected to $\mathbb{T}^2$) for $l = a_i, \ldots, b_i - 1$. We set

$$\mathcal{B}_m(\gamma) = (\gamma \cap B_m) \cup \bigcup_{i=1}^{K} (\tilde{F}_m)^{-1} \,\mathrm{graph}(\tilde{h}_i|_{(a_i - r_i, a_i) \cup (b_i, b_i + s_i)}).$$

For each curve of the form $\hat{\gamma} = (\tilde{F}_m)^{-1}(\mathrm{graph}\, \tilde{h}_i|_{(a_i - r_i, a_i)})$, we have

$$\mathrm{Leb}_\gamma(\hat{\gamma}) = O(L_m^{-\eta})$$

† Recall that Leb_γ is the unnormalized arc length along γ.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

since $\gamma_i \cap B_m = \emptyset$ and similarly for curves of the form $\hat{\gamma} = (\tilde{F}_m)^{-1}(\text{graph } \tilde{h}_i|_{(b_i, b_i+s_i)})$. Combining this with the bound $\text{Leb}_\gamma(\gamma \cap B_m) = O(L_m^{-1+\eta})$, we conclude that

$$\text{Leb}_\gamma(\mathcal{B}_m(\gamma)) = O(L_m^{-1+\eta}).$$

Lastly, item (c) holds for $k = m$ by Lemma 2.5.

Let us now describe the induction procedure for obtaining $\bar{\Gamma}_m^{l+1}(\gamma)$, $\mathcal{B}_m^{l+1}(\gamma)$ $l < k$, assuming that $\bar{\Gamma}_m^l(\gamma)$ and $\mathcal{B}_m^l(\gamma)$ have been defined and that item (c) holds for $k = l$. We define

$$\bar{\Gamma}_m^{l+1}(\gamma) := \bigcup_{\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)} \bar{\Gamma}_{l+1}(\bar{\gamma}),$$

$$\mathcal{B}_m^{l+1}(\gamma) = \mathcal{B}_m^l(\gamma) \cup (F_m^l)^{-1} \bigcup_{\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)} \mathcal{B}_{l+1}(\bar{\gamma}).$$

Repeating the above steps until step $l = k$, we have that $\bar{\Gamma}_m^k(\gamma)$ is composed of fully crossing horizontal curves $\bar{\gamma}$ for which $\|h'_{\bar{\gamma}}\|_{C^0} \le L_k^{-\eta}$. Item (c) similarly follows by the distortion estimate in Lemma 2.5.

It remains to estimate the size of $\mathcal{B}_m^k(\gamma)$. We have for each $m \le l < k$ that

$$\text{Leb}_\gamma(\mathcal{B}_m^{l+1}(\gamma)) = \text{Leb}_\gamma(\mathcal{B}_m^l(\gamma)) + \text{Leb}_\gamma(F_m^l)^{-1} \bigcup_{\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)} \mathcal{B}_{l+1}(\bar{\gamma}).$$

For each $\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)$, we have $\text{Leb}_{\bar{\gamma}}\ \mathcal{B}_{l+1}(\bar{\gamma}) = O(L_{l+1}^{-1+\eta})$ and so

$$\begin{aligned}
\text{Leb}_\gamma(F_m^l)^{-1} \bigcup_{\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)} \mathcal{B}_{l+1}(\bar{\gamma}) &= \sum_{\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)} \text{Leb}_{\bar{\gamma}}(F_m^l)^{-1}(\mathcal{B}_{l+1}(\bar{\gamma})) \\
&= (1 + O(L_m^{1-2\eta})) \sum_{\bar{\gamma} \in \bar{\Gamma}_m^l(\gamma)} \text{Leb}_\gamma((F_m^l)^{-1}\bar{\gamma}) \\
&\quad \cdot \frac{\text{Leb}_{\bar{\gamma}}(\mathcal{B}_{l+1}(\bar{\gamma}))}{\text{Leb}_{\bar{\gamma}}(\bar{\gamma})} \\
&= O(L_{l+1}^{-1+\eta}),
\end{aligned}$$

having applied the distortion estimate in item (c) with $k = l$. This completes the estimate. □

2.3. *Decay of correlations for curves.* The proliferation of horizontal curves throughout phase space is a mixing mechanism for our system. The estimates below justify this in the following sense: the Lebesgue mass along a given fully crossing horizontal curve spreads around throughout phase space in such a way as to approximate Lebesgue measure very closely for Holder-continuous observables.

PROPOSITION 2.8. *Let* $\eta \in (1/2, 1)$. *Assume that* $L_1 \ge \bar{L}_0$, *where* $\bar{L}_0 = \bar{L}_0(M_0, K_0, K_1, \eta)$. *Let* γ *be a fully crossing horizontal curve and let* $\psi : \mathbb{T}^2 \to \mathbb{R}$ *be* α*-Holder continuous. For* $1 \le m \le n$, *we have*

$$\left| \int_\gamma \psi \circ F_m^n d\, \text{Leb}_\gamma - \text{Len}(\gamma) \cdot \int \psi \right| \le C \|\psi\|_\alpha \left(L_n^{-\alpha(1-\eta)/(\alpha+2)} + L_m^{1-2\eta} + \sum_{k=m}^{n-1} L_k^{-1+\eta} \right).$$

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Note that Proposition 2.8 does not require any conditions on the summability of the tail of $\{L_n\}$.

Remark 2.9. As one can see in the following proof, Proposition 2.8 boils down to the case $m = n$, i.e., that for fully crossing horizontal curves γ, the image curve $F_m(\gamma)$ is relatively dense in phase space. Informally, observe that Lemma 2.7 implies that $F_m(\gamma)$ consists (mostly, up to $O(L_m^{-1+\eta})$ error) of fully crossing branches $\check{\gamma}_i$ stacked on top of each other, ordered so that $F_m^{-1}(\check{\gamma}_i)$ is situated to the left of $F_m^{-1}(\check{\gamma}_{i+1})$. Since the x-coordinates of $F_m^{-1}(\check{\gamma}_i)$ are smaller than those of $F_m^{-1}(\check{\gamma}_{i+1})$, it follows from the form of our maps F_m that $\check{\gamma}_i$ is situated *below* $\check{\gamma}_{i+1}$. The vertical distance between $\check{\gamma}_i$, $\check{\gamma}_{i+1}$ is $O(L_m^{-\min\{\eta,1-\eta\}}) = O(L_m^{-(1-\eta)})$ since (i) the approximate length of the subarcs $F_m^{-1}(\check{\gamma}_i)$, $F_m^{-1}(\check{\gamma}_{i+1}) \subset \gamma$ is $O(L_m^{-\eta})$ and (ii) the 'gaps' due to the critical set B_m are of width at most $O(L_m^{-(1-\eta)})$. Since the original curve γ is fully crossing, all x-coordinates are realized on γ and hence all y-coordinates are realized by $F(\gamma)$. It follows from this argument that the image $F(\gamma)$ is approximately $O(L_m^{-(1-\eta)})$-dense in all of $\mathbb{T}^2$.

Proof of Proposition 2.8. With ψ fixed and γ a fully crossing horizontal curve, let $K \in \mathbb{N}$, to be specified shortly, and $\ell = K^{-1}$.

Let $I_1, \ldots, I_K$ denote the partition of $[0, 1)$ into K intervals of length ℓ each. For $1 \le i, j \le K$, let $R_{i,j} = I_i \times I_j$. Note that with $\psi_{i,j} = \inf\{\psi(p) : p \in R_{i,j}\}$, we have

$$\Big\| \psi - \sum_{1 \le i,j \le K} \psi_{i,j} \chi_{R_{i,j}} \Big\|_{L^\infty} = O(\ell^\alpha \|\psi\|_\alpha).$$

Thus,

$$(*) := \int_\gamma \psi \circ F_m^n d\,\mathrm{Leb}_\gamma = O(\ell^\alpha \|\psi\|_\alpha) + \sum_{1 \le i,j \le K} \psi_{i,j} \int_\gamma \chi_{R_{i,j}} \circ F_m^n \, d\,\mathrm{Leb}_\gamma .$$

Form $\bar{\Gamma}_m^{n-1}(\gamma)$, $\mathcal{B}_m^{n-1}(\gamma)$ as in Lemma 2.7, so that for each (i, j)-summand, we have

$$\int_\gamma \chi_{R_{i,j}} \circ F_m^n \, d\,\mathrm{Leb}_\gamma = O\bigg(\sum_{k=m}^{n-1} L_k^{-1+\eta} \bigg) + \sum_{\bar{\gamma} \in \bar{\Gamma}_m^{n-1}(\gamma)} \int_{\bar{\gamma}} \chi_{R_{i,j}} \circ F_n \frac{d\,\mathrm{Leb}_{\bar{\gamma}}}{\|dF_m^{n-1}\| \circ (F_m^{n-1})^{-1}},$$

so that

$$(*) = \|\psi\|_\alpha \cdot O\bigg(\ell^\alpha + \sum_{k=m}^{n-1} L_k^{-1+\eta} \bigg) + \sum_{1 \le i,j \le K} \psi_{i,j} \sum_{\bar{\gamma} \in \bar{\Gamma}_m^{n-1}(\gamma)} \int_{\bar{\gamma}} \frac{d\,\mathrm{Leb}_{\bar{\gamma}}}{\|dF_m^{n-1}\| \circ (F_m^{n-1})^{-1}} \chi_{R_{i,j}} \circ F_n .$$

By the distortion estimate in Lemma 2.7(c), the $(i, j, \bar{\gamma})$-summand equals

$$(1 + O(L_m^{1-2\eta})) \cdot \mathrm{Leb}_\gamma((F_m^{n-1})^{-1}\bar{\gamma}) \underbrace{\int_{\bar{\gamma}} \chi_{R_{i,j}} \circ F_n \, d\,\mathrm{Leb}_{\bar{\gamma}}}_{(**)} .$$

To estimate $(**)$, observe that $\bar\gamma \cap F_n^{-1} R_{i,j} = \bar\gamma|_j$, where for a set $S \subset \mathbb{T}^2$ we write $S|_i = S \cap (I_i \times [0,1))$. Form now the collection $\bar\Gamma_n(\bar\gamma|_j)$ and the set $\mathcal{B}_n(\bar\gamma|_j)$. We obtain

$$
\begin{aligned}
(**) &= \int_{\bar\gamma} \chi_{R_{i,j}} \circ F_n d\,\mathrm{Leb}_{\bar\gamma} = O(\mathcal{B}_n(\bar\gamma|_j)) + \sum_{\bar\gamma' \in \bar\Gamma_n(\bar\gamma|_j)} \int_{\bar\gamma'} \frac{d\,\mathrm{Leb}_{\bar\gamma'}}{\|dF_n|_{T\bar\gamma}\| \circ F_n^{-1}} \chi_{R_{i,j}} \\
&= O(L_n^{-1+\eta}) + (1 + O(L_n^{1-2\eta})) \cdot \sum_{\bar\gamma' \in \bar\Gamma_n(\bar\gamma|_j)} \mathrm{Leb}_{\bar\gamma'}(\bar\gamma'|_i) \cdot \mathrm{Leb}_{\bar\gamma}(F_n^{-1}\bar\gamma')
\end{aligned}
$$

by the distortion estimate in Lemma 2.7(c). Since $\|h'_{\bar\gamma'}\|_{C^0} \le L_n^{-\eta}$ for each $\bar\gamma' \in \bar\Gamma_n(\bar\gamma|_j)$, we easily estimate $\mathrm{Leb}_{\bar\gamma'}(\bar\gamma'|_i) = (1 + O(L_n^{-\eta}))\ell$, so that

$$
\begin{aligned}
(**) &= O(L_n^{-1+\eta}) + (1 + O(L_n^{1-2\eta}))(1 + O(L_n^{-\eta})) \cdot \ell \cdot \mathrm{Leb}_{\bar\gamma}(\bar\gamma|_j \setminus \mathcal{B}_n(\bar\gamma|_j)) \\
&= O(L_n^{-1+\eta}) + (1 + O(L_n^{1-2\eta})) \cdot \ell \cdot \mathrm{Leb}_{\bar\gamma}(\bar\gamma|_j \setminus \mathcal{B}_n(\bar\gamma|_j)).
\end{aligned}
$$

Now, $\mathrm{Leb}_{\bar\gamma}(\mathcal{B}_n(\bar\gamma|_j)) = O(L_n^{-1+\eta})$, so

$$
\begin{aligned}
\mathrm{Leb}_{\bar\gamma}(\bar\gamma|_j \setminus \mathcal{B}_n(\bar\gamma|_j)) &= \mathrm{Leb}_{\bar\gamma}(\bar\gamma|_j) + O(L_n^{-1+\eta}) = (1 + O(L_{n-1}^{-\eta}))\ell + O(L_n^{-1+\eta}) \\
&= (1 + O(L_{n-1}^{-\eta} + \ell^{-1} L_n^{-1+\eta}))\ell,
\end{aligned}
$$

having used the estimate $\|h'_{\bar\gamma}\|_{C^0} \le L_{n-1}^{-\eta}$. Consolidating our estimates,

$$
\begin{aligned}
(**) &= O(L_n^{-1+\eta}) + (1 + O(L_n^{1-2\eta})) \cdot (1 + O(L_{n-1}^{-\eta} + \ell^{-1} L_n^{-1+\eta})) \cdot \ell^2 \\
&= (1 + O(L_n^{1-2\eta} + L_{n-1}^{-\eta} + \ell^{-2} L_n^{-1+\eta}))\ell^2.
\end{aligned}
$$

This establishes the constraint $\ell^{-2} L_n^{-1+\eta} \ll 1$. Inserting the above estimate back into the expression for $(*)$ and using this constraint gives

$$
\begin{aligned}
(*) &= \|\psi\|_\alpha \cdot O\left(\ell^\alpha + \sum_{k=m}^{n-1} L_k^{-1+\eta}\right) \\
&\quad + (1 + O(L_m^{1-2\eta} + L_{n-1}^{-\eta} + \ell^{-2} L_n^{-1+\eta}))\,\mathrm{Leb}_\gamma(\gamma \setminus \mathcal{B}_m^{n-1}(\gamma)) \cdot \sum_{1 \le i,j \le K} \psi_{i,j} \ell^2 \\
&= \|\psi\|_\alpha \cdot O\left(\ell^\alpha + \sum_{k=m}^{n-1} L_k^{-1+\eta}\right) \\
&\quad + (1 + O(L_m^{1-2\eta} + \ell^{-2} L_n^{-1+\eta}))\left(\mathrm{Len}(\gamma) + O\left(\sum_{k=m}^{n-1} L_k^{-1+\eta}\right)\right) \cdot \int \psi \\
&= \|\psi\|_\alpha \cdot O\left(\ell^\alpha + \sum_{k=m}^{n-1} L_k^{-1+\eta}\right) + (1 + O(L_m^{1-2\eta} + \ell^{-2} L_n^{-1+\eta}))\,\mathrm{Len}(\gamma) \cdot \int \psi \\
&= \mathrm{Len}(\gamma) \cdot \int \psi + \|\psi\|_\alpha \cdot O\left(\ell^\alpha + \sum_{k=m}^{n-1} L_k^{-1+\eta} + L_m^{1-2\eta} + \ell^{-2} L_n^{-1+\eta}\right).
\end{aligned}
$$

On setting $K = \ell^{-1} = \lceil L_n^{(1-\eta)/(\alpha+2)} \rceil$, the proof is complete. □

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

3. *Singular limit of* $\{F_n\}$*; finite-time mixing estimates*

Although the compositions $\{F^n\}$ consist of maps becoming increasingly singular by design, we argue in this section that the individual maps F_n do converge, in a sense to be made precise, to some *stationary* process. This we formulate in a precise way in §3.1. As we argue below, these considerations naturally follow from finite-time mixing properties of the partial compositions F_m^n for m, n very large, $m \le n$; we state and prove these mixing estimates in §3.2, verifying the convergence mode described in §3.1.

As they are of independent interest, these finite-time mixing estimates are re-formulated for the standard maps F_L, $L > 0$, as Theorem D.

3.1. *Singular limit of* $\{F_n\}$. As n increases, the maps $F_n(x, y) = (f_n(x) - y \pmod 1, x)$ become more and more singular due to the fact that $L_n \to \infty$; in particular, $\lim_{n\to\infty} F_n$ does not exist in any meaningful topology on diffeomorphisms of $\mathbb{T}^2$. To motivate a meaningful convergence notion, let us consider the action in the x-coordinate given by the map $f_n : \mathbb{T}^1 \to \mathbb{T}^1$.

Observe that for n extremely large, $f_n : \mathbb{T}^1 \to \mathbb{T}^1$ is predominantly an expanding map and so in one time iterate the value of $f_n(x)$, $x \in \mathbb{T}^1$ is increasingly sensitive to $x \in \mathbb{T}^1$. Cast in a different light, f_n is increasingly 'randomizing' on $\mathbb{T}^1$, to the point where x and $f_n(x)$ are increasingly decorrelated as $n \to \infty$. One might expect, then, that in the limit, $f_n(x)$ can be modeled by a random variable *independent of* x. A step towards a precise formulation might be as follows: for some class of continuous observables $\phi, \psi : \mathbb{T}^1 \to \mathbb{R}$, we should expect that

$$\lim_{n\to\infty} \int_{\mathbb{T}^1} \phi \circ f_n(x) \cdot \psi(x) = \int \phi \int \psi.$$

Morally speaking, we expect that when X is a random variable distributed in a 'nice' way on $\mathbb{T}^1$, we have that the joint law of the pair $(X, f_n(X))$ converges, in a weak sense, to the joint law of a pair (X, Z) for which Z is independent of X.

Let us now return to the implications for the full maps $F_n : \mathbb{T}^2 \to \mathbb{T}^2$ and make things more precise. The above discussion motivates modeling F_n for n large by a *Markov chain* $\{Z_n = (X_n, Y_n)\}$ defined as follows. Let $\beta_1, \beta_2, \ldots$ be independent and identically distributed random variables uniformly distributed on $\mathbb{T}^1$. Given an initial condition $Z_0 = (X_0, Y_0) \in \mathbb{T}^2$, we iteratively define

$$Z_{n+1} = (X_{n+1}, Y_{n+1}) = (\beta_{n+1}, X_n)$$

for $n \ge 0$. The form of this Markov chain agrees with the idea, argued above, that $X, f_n(X)$ are 'roughly independent' for large n. Note that $Z_{n+2} = (\beta_{n+2}, \beta_{n+1})$ is independent of $Z_n = (\beta_n, \beta_{n-1})$ for all $n \ge 1$.

Let P denote the transition operator associated with Z_n, so that

$$P((x, y), A \times B) = \mathrm{Leb}(A) \cdot \delta_x(B)$$

for Borel $A, B \subset \mathbb{T}^2$, where δ_x denotes the Dirac mass at x. Write P^k for the kth iterate of P. For $\phi : \mathbb{T}^2 \to \mathbb{R}$, $k \ge 1$, we define $P^k\phi : \mathbb{T}^2 \to \mathbb{R}$ by $P^k\phi(x, y) = \int P^k((x, y), d\bar{x}\, d\bar{y})\phi(\bar{x}, \bar{y})$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

PROPOSITION 3.1. *Fix $k \geq 1$ and let $\phi, \psi : \mathbb{T}^2 \to \mathbb{R}$ be continuous. Assume that $L_m \to \infty$ as $m \to \infty$. Then*

$$\lim_{m\to\infty} \int \psi \circ F_m^{m+k-1} \cdot \phi = \int P^k \psi \cdot \phi.$$

That is, the maps F_n converge to the Markov chain $(Z_n)_n$ in the sense that the associated Koopman operators converge to the transition operator P for Holder observables in a way reminiscent of the weak operator topology. Proposition 3.1 is proved in §3.2 below.

Remark 3.2. The convergence described in Proposition 3.1 suggests that the asymptotic variance of sums $(1/\sqrt{N}) \sum_{i=0}^{N-1} \phi \circ F^i$ as in the central limit theorem (Theorem B) should coincide with the asymptotic variance $\hat{\sigma}^2(\phi)$ of $(1/\sqrt{N}) \sum_{i=0}^{N-1} \phi(Z_i)$, $Z_0 \sim \text{Leb}_{\mathbb{T}^2}$. Developing the Green–Kubo formula for $\hat{\sigma}^2(\phi)$, we obtain

$$\begin{aligned}\hat{\sigma}^2(\phi) &= \mathbb{E}(\phi(Z_0)^2) + 2\sum_{l=1}^{\infty} \mathbb{E}(\phi(Z_0)\phi(Z_l)) \\ &= \mathbb{E}(\phi(Z_0)^2) + 2\mathbb{E}(\phi(Z_0)\phi(Z_1)) \\ &= \int \phi^2 + 2\int \phi(x, y)\phi(y, z)\, dx\, dy\, dz,\end{aligned}$$

where we have used the fact that Z_k, Z_0 are independent when $k \geq 2$. This is precisely the form of σ^2 given in Theorem B. Here $\mathbb{E}$ refers to the expectation where $Z_0 \sim \text{Leb}_{\mathbb{T}^2}$.

This perspective also explains the 'coboundary condition' $\phi(x, y) = \psi(x) - \psi(y)$ for some bounded $\psi : \mathbb{T}^1 \to \mathbb{R}$. If ϕ has this form, then the sums in the CLT for this Markov chain telescope: $\phi(Z_0) + \phi(Z_1) + \cdots + \phi(Z_{n-1}) = -\psi(Y_0) + \psi(X_n)$ and so the asymptotic variance is zero. Let us now check that this is also a necessary condition for the asymptotic variance $\hat{\sigma}^2(\phi)$ to be zero.

LEMMA 3.3. *Let $\phi : \mathbb{T}^2 \to \mathbb{R}$ be a Holder continuous function with $\int \phi\, dxdy = 0$. Then $\hat{\sigma}^2(\phi) = 0$ if and only if $\phi(x, y) = \psi(x) - \psi(y)$, where $\psi : \mathbb{T}^1 \to \mathbb{R}$ is some Holder continuous function.*

Proof. We have the identity

$$\hat{\sigma}^2(\phi) = \int \bigg(\phi(x, y) + \int \phi(z, x)\, dz - \int \phi(w, y)\, dw\bigg)^2 dx\, dy,$$

the verification of which is a straightforward (albeit tedious) computation omitted for brevity. Now, $\hat{\sigma}^2(\phi) = 0$ implies that $\phi(x, y) = \psi(x) - \psi(y)$ pointwise (since ϕ is continuous), where $\psi(x) := -\int \phi(z, x)\, dx$. □

3.2. *Finite-time mixing estimates.* The limiting notion described in Proposition 3.1 has at its core the statement that finite compositions F_m^n, $m \leq n$ are 'mixing' in the limit $m, n \to \infty$. We will, in fact, prove something much stronger: a concrete estimate on the correlation of (x, y) to $F_m^n(x, y)$ for m, n large with respect to C^α observables.

PROPOSITION 3.4. *Fix $\eta \in (1/2, 1)$ and $\alpha \in (0, 1]$. Let L_0 be sufficiently large, depending on α, η. Let $m \geq 1$ and let $\phi_1, \phi_2 : \mathbb{T}^2 \to \mathbb{R}$ be α-Holder continuous functions. Then there exists a constant $C > 0$, depending only on K_0, K_1, M_0, such that the following hold.*

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

(a) *We have*

$$\left| \int \phi_1 \circ F_m \cdot \phi_2 - \int \phi_1(x, z)\phi_2(z, y)\, dx\, dy\, dz \right| \\ \leq C\|\phi_1\|_\alpha \|\phi_2\|_\alpha L_m^{-\min\{2\eta-1, \alpha(1-\eta)/(2+\alpha)\}}.$$

(b) *Let $n > m$. Then*

$$\left| \int \phi_1 \circ F_m^n \cdot \phi_2 - \int \phi_1 \int \phi_2 \right| \\ \leq C\|\phi_1\|_\alpha \|\phi_2\|_\alpha \left(L_m^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \right).$$

Observe that Proposition 3.1 follows easily from Proposition 3.4. Moreover, as we leave to the reader to check, the proof of Proposition 3.4 requires only that the sequence $\{L_n\}$ be *non-decreasing* and so applies equally well in the case when $L_m = L_{m+1} = \cdots = L_n = L$ for some fixed $L > 0$. Thus, Theorem D follows.

Items (a) and (b) are proved separately in §§3.2.1 and 3.2.2 below, respectively.

3.2.1. *Proof of Proposition 3.4(a).* Throughout §§3.2.1 and 3.2.2, we let I_i, $R_{i,j}$ be as in the proof of Proposition 2.8, where $\ell = K^{-1}$ and $K \in \mathbb{N}$ will be specified at the end (twice, once for part (a) and again for part (b)).

With $\alpha \in (0, 1]$ and ϕ_1, ϕ_2 fixed, for $l = 1, 2$, we define $\phi_{i,j}^l = \inf_{R_{i,j}} \phi_l$, so that

$$\left\| \phi_l - \sum_{i,j} \phi_{i,j}^l \chi_{R_{i,j}} \right\|_{L^\infty} = O(\|\phi\|_\alpha \ell^\alpha).$$

To begin, we estimate

$$\int \phi_1 \circ F_m \cdot \phi_2 = O(\|\phi_1\|_\alpha \|\phi_2\|_\alpha \ell^\alpha) + \sum_{1 \leq i,j,i',j' \leq K} \phi_{i,j}^1 \phi_{i',j'}^2 \int \chi_{R_{i,j}} \circ F_m \cdot \chi_{R_{i',j'}} \\ = O(\|\phi_1\|_\alpha \|\phi_2\|_\alpha \ell^\alpha) + \sum_{1 \leq i_0, i_1, i_2 \leq K} \phi_{i_2 i_1}^1 \phi_{i_1 i_0}^2 \int \chi_{R_{i_1 i_0}} \chi_{R_{i_2 i_1}} \circ F_m,$$

where in passing from the first line to the second we have used that $F_m(R_{i,j}) \subset [0, 1) \times I_i$.

Fixing i_0, i_1, i_2, let $y_0 \in I_{i_0}$ and set $H = I_{i_1} \times \{y_0\}$. Applying Lemma 2.7,

$$(*) = \int_H \chi_{R_{i_2 i_1}} \circ F_m d\,\mathrm{Leb}_H = O(\mathrm{Leb}_H(\mathcal{B}_m(H))) \\ + \sum_{\bar\gamma \in \bar\Gamma_m(H)} \int_{\bar\gamma} \frac{d\,\mathrm{Leb}_{\bar\gamma}}{\|dF_m|_{TH}\| \circ F_m^{-1}} \chi_{R_{i_2 i_1}} d\,\mathrm{Leb}_{\bar\gamma} \\ = O(L_m^{-1+\eta}) + \sum_{\bar\gamma \in \bar\Gamma_m(H)} (1 + O(L_m^{1-2\eta}))\, \mathrm{Leb}_H(F_m^{-1}\bar\gamma) \cdot \int_{\bar\gamma} \chi_{I_{i_2} \times [0,1)}\, d\,\mathrm{Leb}_{\bar\gamma},$$

having used again that $F_m(I_i \times [0, 1)) \subset [0, 1) \times I_i$ to develop the integrand on the far right. Estimating $\mathrm{Leb}_{\bar\gamma}(\bar\gamma \cap I_{i_2} \times [0, 1)) = (1 + O(L_m^{-\eta}))\ell$ (having used that $\|h'_{\bar\gamma}\|_{C^0} =$

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

$O(L_m^{-\eta})$), we obtain

$$
\begin{aligned}
(*) &= O(L_m^{-1+\eta}) + (1 + O(L_m^{1-2\eta}))(1 + O(L_m^{-\eta}))\, \mathrm{Leb}_H(H \backslash \mathcal{B}_m(H)) \cdot \ell \\
&= O(L_m^{-1+\eta}) + (1 + O(L_m^{1-2\eta}))(1 + O(L_m^{-\eta}))(\ell + O(L_m^{-1+\eta})) \cdot \ell \\
&= \ell^2(1 + O(L_m^{1-2\eta} + \ell^{-2} L_m^{-1+\eta}).
\end{aligned}
$$

Integrating over $y_0 \in I_{i_0}$, we conclude that

$$
\mathrm{Leb}(R_{i_0 i_1} \cap F_m^{-1} R_{i_2 i_1}) = \ell^3(1 + O(\ell^{-2} L_m^{-1+\eta} + L_m^{1-2\eta})).
$$

Summing now over $1 \le i_0, i_1, i_2 \le K$ gives

$$
\begin{aligned}
\int \phi_1 \circ F_m \cdot \phi_2 &= O(\|\phi_1\|_\alpha \|\phi_2\|_\alpha (\ell^\alpha + \ell^{-2} L_m^{-1+\eta} + L_m^{1-2\eta})) + \sum_{1 \le i_0, i_1, i_2 \le K} \phi^1_{i_2 i_1} \phi^2_{i_1 i_0} \ell^3 \\
&= O(\|\phi_1\|_\alpha \|\phi_2\|_\alpha (\ell^\alpha + \ell^{-2} L_m^{-1+\eta} + L_m^{1-2\eta})) \\
&\quad + \int \phi_1(x, z) \phi_2(z, y)\, dx\, dy\, dz.
\end{aligned}
$$

The proof is complete on setting $K = \ell^{-1} = \lceil L_m^{(1-\eta)/(2+\alpha)} \rceil$.

3.2.2. *Proof of Proposition 3.4(b).* All notation is as in the beginning of §3.2.1. We estimate

$$
(**) = \int \phi_1 \circ F_m^n \cdot \phi_2 = O(\|\phi_1\|_\alpha \|\phi_2\|_\alpha \ell^\alpha) + \sum_{1 \le i, j \le K} \phi^2_{i,j} \int_{R_{i,j}} \phi_1 \circ F_m^n.
$$

Fix $1 \le i, j \le K$. For $y_0 \in I_j$, write $H = H(y_0) = I_i \times \{y_0\}$. Then

$$
\int_{R_{i,j}} \phi_1 \circ F_m^n = \int_{y \in I_j} \int_{H(y_0)} \phi_1 \circ F_m^n d\, \mathrm{Leb}_{H(y_0)}\, dy_0.
$$

Developing the inner integral and applying Lemma 2.7,

$$
\begin{aligned}
\int_{H(y_0)} \phi_2 \circ F_m^n d\, \mathrm{Leb}_{H(y_0)} &= O(\|\phi_1\|_0\, \mathrm{Leb}_{H(y_0)}\, \mathcal{B}_m(H(y_0))) \\
&\quad + \sum_{\bar\gamma \in \bar\Gamma_m(\gamma(y_0))} \int_{\bar\gamma} \frac{d\, \mathrm{Leb}_{\bar\gamma}}{\|dF_m|_{TH(y_0)}\| \circ F_m^{-1}} \phi_1 \circ F_{m+1}^n \\
&= O(\|\phi_1\|_{C^0} L_m^{-1+\eta}) + (1 + O(L_m^{1-2\eta})) \\
&\quad \times \sum_{\bar\gamma \in \bar\Gamma_m(H(y_0))} \mathrm{Leb}_{H(y_0)}(F_m^{-1} \bar\gamma) \int_{\bar\gamma} \phi_1 \circ F_{m+1}^n d\, \mathrm{Leb}_{\bar\gamma}\, .
\end{aligned}
$$

The curves $\bar\gamma$ cross the full horizontal extent of $\mathbb{T}^2$ and so fall under the purview of Proposition 2.8. Applying the estimate there, we obtain

$$
\begin{aligned}
&\int_{\bar\gamma} \phi_1 \circ F_{m+1}^n d\, \mathrm{Leb}_{\bar\gamma} \\
&= \mathrm{Len}(\bar\gamma) \int \phi_1 + O\bigg(\|\phi_1\|_\alpha \bigg(L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_{m+1}^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta}\bigg)\bigg) \\
&= \int \phi_1 + O\bigg(\|\phi_1\|_\alpha \bigg(L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_{m+1}^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta}\bigg)\bigg).
\end{aligned}
$$

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Summing over $\bar{\gamma}$, we obtain that $\int_{H(y_0)} \phi_1 \circ F_m^n d\,\mathrm{Leb}_{H(y_0)}$ equals

$$
\begin{aligned}
&O(\|\phi_1\|_{C^0} L_m^{-1+\eta}) + (1 + O(L_m^{1-2\eta})) \sum_{\bar{\gamma} \in \bar{\Gamma}_m(H(y_0))} \mathrm{Leb}_{H(y_0)}(F_m^{-1}\bar{\gamma}) \\
&\quad \cdot \bigg(\int \phi_1 + O\bigg(\|\phi_1\|_\alpha \bigg(L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_{m+1}^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \bigg)\bigg)\bigg) \\
&= O\bigg(\|\phi_1\|_\alpha \bigg(L_m^{-1+\eta} + \ell \bigg(L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_{m+1}^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \bigg)\bigg)\bigg) \\
&\quad + (1 + O(L_m^{1-2\eta}))\, \mathrm{Leb}_{H(y_0)}(H(y_0) \backslash \mathcal{B}_m(H(y_0))) \int \phi_1 \\
&= O\bigg(\|\phi_1\|_\alpha \bigg(L_m^{-1+\eta} + \ell \bigg(L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_{m+1}^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \bigg)\bigg)\bigg) \\
&\quad + (1 + O(L_m^{1-2\eta}))(1 + O(\ell^{-1} L_m^{-1+\eta})) \cdot \ell \int \phi_1 \\
&= \ell \cdot \bigg\{ O\bigg(\|\phi_1\|_\alpha \bigg(\ell^{-1} L_m^{-1+\eta} + L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_m^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \bigg)\bigg) \\
&\quad + \int \phi_1 \bigg\}.
\end{aligned}
$$

Integrating over $y_0 \in I_j$ yields the same estimate for $\int \chi_{R_{i,j}} \phi_1 \circ F_m^n$ with an additional factor of ℓ. Summing over $1 \le i, j \le K$, we have that $\int \phi_1 \circ F_m^n \cdot \phi_2$ equals

$$
\begin{aligned}
&O\bigg(\|\phi_1\|_\alpha \|\phi_2\|_\alpha \bigg(\ell^\alpha + \ell^{-1} L_m^{-1+\eta} + L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_m^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \bigg)\bigg) \\
&\quad + \sum_{i,j=1}^{K} \ell^2 \phi_{i,j}^2 \int \phi_1 \\
&= O\bigg(\|\phi_1\|_\alpha \|\phi_2\|_\alpha \bigg(\ell^\alpha + \ell^{-1} L_m^{-1+\eta} + L_n^{-\alpha(1-\eta)/(2+\alpha)} + L_m^{1-2\eta} + \sum_{k=m+1}^{n-1} L_k^{-1+\eta} \bigg)\bigg) \\
&\quad + \int \phi_1 \int \phi_2.
\end{aligned}
$$

The proof is complete on setting $K = \lceil L_m^{(1-\eta)/(1+\alpha)} \rceil$.

4. *Law of large numbers*

We continue our study of the statistical properties of the composition $\{F^n\}$ by proving Theorem A, a pair of formulations of a 'law of large numbers' for time averages of observables.

In this section $\alpha \in (0, 1]$ *is fixed, as is a sequence of* α*-Holder continuous observables* $\phi_i : \mathbb{T}^2 \to \mathbb{R}$, $i \ge 0$ *with* $\int \phi_i = 0$ *for all* i *and* $\sup_{i \ge 0} \|\phi_i\|_\alpha \le C_0$ *for a constant* $C_0 > 0$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

For $0 \le M \le N$, we define

$$\hat{S}_{M,N} = \phi_M \circ F^M + \cdots + \phi_N \circ F^N$$

and set $\hat{S}_N = \hat{S}_{0,N}$. Noting that the simple estimate

$$|\hat{S}_N - \hat{S}_{M,N}| = \left| \sum_{i=0}^{M-1} \phi_i \circ F^i \right| \le C_0 M$$

holds pointwise on $\mathbb{T}^2$, it follows that to prove a strong law for $\hat{S}_N$, it suffices to prove a strong law for $\hat{S}_{M,N}$, where $M = M(N) = \lfloor \sqrt{N} \rfloor$. Similarly, a weak law for $\hat{S}_N$ follows from a weak law for $\hat{S}_{M,N}$. More precisely, to prove Theorem A, it suffices to prove the following.

PROPOSITION 4.1. *For $N \ge 1$, let $M = M(N) = \lfloor \sqrt{N} \rfloor$.*

(a) *If $N^2 L_N^{-\alpha/(3\alpha+4)} \to 0$, then $(1/(N-M))\hat{S}_{M,N}$ converges in L^2 to 0.*

(b) *If $N^{4+\epsilon} L_{\lfloor \sqrt{N} \rfloor}^{-\alpha/(3\alpha+4)} \to 0$ as $N \to \infty$ for some $\epsilon > 0$, then $(1/(N-M))\hat{S}_{M,N}$ converges almost surely to 0.*

Proof of Proposition 4.1. To start, we expand

$$\int \hat{S}_{M,N}^2 = \sum_{n=M}^{N} \int \phi_n^2 \circ F_M^n + 2 \sum_{M \le m < n \le N} \int \phi_n \circ F_{m+1}^n \cdot \phi_m.$$

For the first term, each summand is precisely $\int \phi_n^2 \le C_0^2$. For the second term, the (m, n)-summand is bounded

$$\|\phi_n\|_\alpha \|\phi_m\|_\alpha \cdot O\left(L_{m+1}^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}} + \sum_{k=m+2}^{n-1} L_k^{-1+\eta} \right)$$

by Proposition 3.4(b) and so the entire summation is bounded

$$\begin{aligned} C_0^2 (N-M)^2 O\left(L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}} + \sum_{k=M}^{N} L_k^{-1+\eta} \right) \\ = C_0^2 (N-M)^3 O(L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}}). \end{aligned}$$

Optimizing in η, the function $\eta \mapsto \min\{\alpha(1-\eta)/(2+\alpha), 2\eta - 1\}$ is maximized at the value $\alpha/(3\alpha+4)$ at the point $\eta = (2\alpha+2)/(3\alpha+4)$. Hereafter this value of η is fixed.

Setting $M = M(N) = \lfloor \sqrt{N} \rfloor$, we obtain that $N^{-2} \int \hat{S}_{M,N}^2 \to 0$ as $N \to \infty$ so long as $N L_{\lfloor \sqrt{N} \rfloor}^{-\alpha/(3\alpha+4)} \to 0$, as we have in the hypotheses of item (a). For (b), our estimates imply that the sequence $\{N^{-2} \int \hat{S}_{M,N}^2\}_{N \ge 1}$ is summable whenever $N^{2+\epsilon} L_{\lfloor \sqrt{N} \rfloor}^{-\alpha/(3\alpha+4)} \to 0$ for some $\epsilon > 0$ (which we have from the condition in (b)). Summability implies fast convergence in probability, which implies almost sure convergence (using the Borel–Cantelli lemma). This completes the proof. □

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

5. *Central limit theorem*

Here we carry out the proof of the central limit theorem in Theorem B. A standard technique, attributed to Gordin, for proving the central limit theorem for a deterministic dynamical system is to look for *reverse martingale difference approximations* for sums of observables and then to use probability theory tools for proving the central limit theorem for sums of reverse martingale differences (see, e.g., [**24**] for an exposition).

For expediency, we pursue a slightly different method: we construct here an array of *forward* martingale difference approximations. The corresponding *forward* filtrations are composed (mostly) of fully crossing horizontal curves. The filtration is constructed in §5.1.1. Our martingale difference approximation is constructed in §5.1.2 and in §5.1.3 we show how the CLT for our approximation implies the CLT as in Theorem B. The CLT for our martingale difference approximation is proved in §5.2.

Throughout this section $\alpha \in (0, 1]$ *is fixed and* $\phi : \mathbb{T}^2 \to \mathbb{R}$ *is assumed to be an* α*-Holder continuous observable with* $\int \phi = 0$*. The value* $\eta \in (1/2, 1)$ *is assumed fixed; as we did in the previous section, in §5.1.3 we will specialize to a particular value of* η *depending on* α*.*

Notation. We write $\mathbb{E}$ below for the expectation with respect to Lebesgue measure on $\mathbb{T}^2$. When $\mathcal{G}$ is a sub-sigma-algebra of the Borel sigma-algebra, we write $\mathbb{E}(\cdot|\mathcal{G})$ for the conditional expectation with respect to $\mathcal{G}$.

5.1. *Preliminaries for CLT: construction of a martingale approximation.*

5.1.1. *Construction of the increasing filtrations* $\{\hat{\mathcal{G}}_{M,k}, k \geq M\}$.

We will produce an increasing filtration of (most of) $\mathbb{T}^2$ by horizontal curves with a small and controlled exceptional set. Below, $M \in \mathbb{N}$ should be thought of as large.

First, we will construct a sequence of *partitions* $\zeta_{M,M}, \zeta_{M,M+1}, \ldots, \zeta_{M,k}, \ldots$ of $\mathbb{T}^2$ with the following properties for each $M \leq k \leq N$:

(A) the partition $\zeta_{M,k}$ is 'mostly' composed of fully crossing horizontal curves; and

(B) $\zeta_{M,k} \leq F_k^{-1} \zeta_{M,k+1}$†.

Once the $\zeta_{M,k}$ are constructed, we define $\mathcal{G}_{M,k}$ to be the sigma-algebra of measurable unions of elements in $\zeta_{M,k}$ and, finally,

$$\hat{\mathcal{G}}_{M,k} = (F_M^k)^{-1} \mathcal{G}_{M,k+1},$$

so that $\{\hat{\mathcal{G}}_{M,k}\}_{k \geq M}$ is an increasing filtration on $\mathbb{T}^2$. This is the filtration we will use in the following to construct our forward martingale difference approximation.

Construction of $\{\zeta_{M,k}, k \geq M\}$ *satisfying (A) and (B).* Set $\zeta_{M,M}$ to be the partition of $\mathbb{T}^2 \setminus \{x = 0\}$ into horizontal line segments. Applying Lemma 2.7, for each $\zeta \in \zeta_{M,M}$, form $\mathcal{B}_M(\zeta)$ and $\bar{\Gamma}_M(\zeta)$, writing

$$G_{M,M+1} = \bigcup_{\substack{\zeta \in \zeta_{M,M} \\ \bar{\zeta} \in \bar{\Gamma}_M(\zeta)}} \bar{\zeta}, \quad B_{M,M+1} = \bigcup_{\zeta \in \zeta_{M,M}} F_M(\mathcal{B}_M(\zeta)).$$

† Here '$\leq$' refers to the partial order on partitions: two partitions ζ, ζ' satisfy $\zeta \leq \zeta'$ if any atom of ζ is a union of ζ' atoms.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Defining the partition $\mathcal{H}_{M,M+1} = \{G_{M,M+1}, B_{M,M+1}\}$, we now define the partition $\zeta_{M,M+1} \geq \mathcal{H}_{M,M+1}$ as follows:

$$\zeta_{M,M+1}|_{G_{M,M+1}} = \{\bar{\zeta} : \bar{\zeta} \in \bar{\Gamma}_M(\zeta), \zeta \in \zeta_{M,M}\},$$
$$\zeta_{M,M+1}|_{B_{M,M+1}} = \{F_M(\zeta) \cap B_{M,M+1} : \zeta \in \zeta_{M,M}\}.$$

Iterating, assume that $\zeta_{M,k}$ has been formed, where $k \geq M+2$, along with the partition $\mathcal{H}_{M,k} = \{G_{M,k}, B_{M,k}\}$ for which $\zeta_{M,k} \geq \mathcal{H}_{M,k}$. For each $\zeta \in \zeta_{M,k}|_{G_{M,k}}$, form $\bar{\Gamma}_k(\zeta)$ and define

$$G_{M,k+1} = \bigcup_{\substack{\zeta \in \zeta_{M,k} \\ \bar{\zeta} \in \bar{\Gamma}_k(\zeta)}} \bar{\zeta}, \quad B_{M,k+1} = \bigcup_{\zeta \in \zeta_{M,k}} F_k(\mathcal{B}_k(\zeta))$$

and define $\zeta_{M,k+1}$ by

$$\zeta_{M,k+1}|_{G_{M,k+1}} = \{\bar{\zeta} : \bar{\zeta} \in \bar{\Gamma}_k(\zeta), \zeta \in \zeta_{M,k}|_{G_{M,k}}\},$$
$$\zeta_{M,k+1}|_{B_{M,k+1}} = \{F_k(\zeta) \cap B_{M,k+1} : \zeta \in \zeta_{M,k}\}.$$

Below, we formulate and verify properties (A) and (B) above for the sequence $\zeta_{M,k}$, $k \geq M$ constructed above.

LEMMA 5.1. *The partitions* $\{\zeta_{M,k}\}_{k \geq M}$, $\mathcal{H}_{M,k} = \{G_{M,k}, B_{M,k}\}$ *are measurable and have the following properties for each* $k \geq M$.

(a) *Every atom* $\zeta \in \zeta_{M,k}|_{G_{M,k}}$ *is a fully crossing horizontal curve for which* $\|h'_\zeta\|_{C^0} \leq L_{k-1}^{-\eta}$.
(b) *We have* $\zeta_{M,k} \leq F_k^{-1}\zeta_{M,k+1}$.
(c) *We have the estimate*

$$\mathrm{Leb}(B_{M,k}) = O\bigg(\sum_{i=M}^{k-1} L_i^{-1+\eta}\bigg).$$

Proof. Measurability is not hard to check. Items (a) and (b) follow from the construction. For the estimate in item (c), observe that for each $k \geq M$, $\zeta \in \zeta_{M,k}|_{G_{M,k}}$, we have $\mathrm{Leb}_\zeta(\mathcal{B}_k(\zeta)) = O(L_k^{-1+\eta})$ and hence $(\mathrm{Leb}_{\mathbb{T}^2})_\zeta(\mathcal{B}_k(\zeta)) \leq (1 + O(L_{k-1}^{-\eta})) \cdot O(L_k^{-1+\eta}) = O(L_k^{-1+\eta})$, where here $(\mathrm{Leb}_{\mathbb{T}^2})_\zeta$ is the disintegration measure of $\mathrm{Leb}_{\mathbb{T}^2}|_{G_{M,k}}$ with respect to $\zeta \in \zeta_{M,k}|_{G_{M,k}}$. We conclude that

$$\mathrm{Leb}(G_{M,k+1}) = (1 + O(L_k^{-1+\eta}))\,\mathrm{Leb}(G_{M,k})$$

and hence

$$\mathrm{Leb}(G_{M,m}) = \prod_{k=M}^{m-1} (1 + O(L_k^{-1+\eta})) \geq 1 + O\bigg(\sum_{k=M}^{m-1} L_k^{-1+\eta}\bigg). \qquad \square$$

The choice of $\hat{\mathcal{G}}_{M,k}$ is made so that $F_M^{k-1}\hat{\mathcal{G}}_{M,k} = F_k^{-1}\mathcal{G}_{M,k+1}$ is a very 'fine' sigma-algebra. Before proceeding, we record the following estimate.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

LEMMA 5.2. *Let ϕ be α-Holder continuous, $k \geq M$. Then*

$$|\phi - \mathbb{E}(\phi | F_k^{-1} \mathcal{G}_{M,k+1})| = O(\|\phi\|_\alpha L_k^{-\eta\alpha})$$

on $F_k^{-1} G_{M,k+1}$.

Proof. Let $\zeta \in \mathcal{G}_{M,k+1}|_{G_{M,k+1}}$. Then $F_k^{-1}\zeta$ is, by our construction, a subsegment of a fully crossing curve $\zeta' \in \zeta_{M,k}|_{G_{M,k}}$ with diameter $O(L_k^{-\eta})$. So, for any points $p,\ p' \in F_k^{-1}\zeta$, we have $|\phi(p) - \phi(p')| = O(\|\phi\|_\alpha L_k^{-\eta\alpha})$. □

5.1.2. *Approximation by sum of martingale differences.* For a bounded observable $\phi : \mathbb{T}^2 \to \mathbb{R}$, convergence in distribution of $(1/\sqrt{N}) S_N(X)$, $X \sim \mathrm{Leb}_{\mathbb{T}^2}$, where

$$S_N = \sum_{n=1}^{N} \phi \circ F^{n-1},$$

is equivalent to convergence in distribution of $(1/\sqrt{N}) S_{M,N}(X)$, $X \sim \mathrm{Leb}_{\mathbb{T}^2}$, where

$$S_{M,N} = \sum_{n=M}^{N} \phi \circ F_M^{n-1}$$

and $M = M(N)$ is a sequence satisfying $M(N) \ll \sqrt{N}$. Here '$X \sim \mathrm{Leb}_{\mathbb{T}^2}$' means that X is a $\mathbb{T}^2$-valued random variable with law $\mathrm{Leb}_{\mathbb{T}^2}$.

Thus, for Theorem B, it suffices to prove convergence in distribution of $(1/\sqrt{N}) S_{M,N}(X)$; for this, we approximate $S_{M,N}$ by a sum of martingale differences with respect to the increasing filtrations $\hat{\mathcal{G}}_{M,k}$, $k \geq M$.

PROPOSITION 5.3. *Let $M \leq N$. Define*

$$\tilde{S}_{M,N} = \sum_{n=M}^{N} \mathbb{E}(\phi | (F_n)^{-1} \mathcal{G}_{M,n+1}) \circ F_M^{n-1} = \sum_{n=M}^{N} \mathbb{E}(\phi \circ F_M^{n-1} | \hat{\mathcal{G}}_{M,n}).$$

(a) *The sum $\tilde{S}_{M,N}$ admits the representation $\tilde{S}_{M,N} = \sum_{n=M}^{N} U_{M,N,n}$, where*

$$U_{M,N,n} = \sum_{m=n-1}^{N-1} (\mathbb{E}(\phi \circ F_M^m | \hat{\mathcal{G}}_{M,n}) - \mathbb{E}(\phi \circ F_M^m | \hat{\mathcal{G}}_{M,n-1})).$$

The sequence $\{U_{M,N,n},\ M \leq n \leq N\}$ is a forward martingale difference adapted to $(\hat{\mathcal{G}}_{M,n},\ M \leq n \leq N)$. Precisely, $\mathbb{E}(U_{M,N,n} | \hat{\mathcal{G}}_{M,n}) = U_{M,N,n}$ and $\mathbb{E}(U_{M,N,n} | \hat{\mathcal{G}}_{M,n-1}) = 0$.

(b) *We have*

$$|S_{M,N} - \tilde{S}_{M,N}| = O\bigg((N - M)\|\phi\|_\alpha \sum_{m=M}^{N} L_m^{-\eta\alpha}\bigg)$$

on $G_{M,N}$.

Above, we use the convention that $\hat{\mathcal{G}}_{M,M-1} = \{\emptyset, \mathbb{T}^2\}$ is the trivial sigma-algebra on $\mathbb{T}^2$. For notational simplicity, when M, N are fixed, we write $U_n = U_{M,N,n}$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Proof. Item (b) is a simple consequence of Lemma 5.2. For item (a), the relation $\tilde{S}_{M,N} = \sum_{M\le n\le N} U_{M,N,n}$ can be verified by a direct computation.

Alternatively, following the analogue of the derivation of a reverse martingale difference approximation given in [**14**] for forward martingale differences, one can look for a martingale difference $U_n = \mathbb{E}(\phi\circ F_M^n|\hat{\mathcal{G}}_{M,n}) + h_n - h_{n+1}$, where $(h_n)_{M\le n\le N+1}$ is some sequence of 'coboundary' functions to be determined. Making the ansatz $h_{N+1}=0$ and 'solving' the conditions $\mathbb{E}(U_n|\hat{\mathcal{G}}_{M,n}) = U_n$, $\mathbb{E}(U_n|\hat{\mathcal{G}}_{M,n-1}) = 0$ for each n, we deduce formally that

$$h_n = -\sum_{m=n-1}^{N-1} \mathbb{E}(\phi\circ F_M^m|\hat{\mathcal{G}}_{M,n-1}).$$

Inserting this formula into the relation $U_n = \mathbb{E}(\phi\circ F_M^n|\hat{\mathcal{G}}_{M,n}) + h_n - h_{n+1}$ yields the form of U_n given above. The choice $\hat{\mathcal{G}}_{M,M-1} = \{\emptyset, \mathbb{T}^2\}$ ensures that $h_M = 0$ and hence $\tilde{S}_{M,N} = \sum_{M\le n\le N} U_n + h_M - h_{N+1} = \sum_{M\le n\le N} U_n$ holds. □

5.1.3. *Deducing Theorem B from the martingale approximation.* We will deduce Theorem B from the following.

PROPOSITION 5.4. *Assume that $N^8 L_N^{-\alpha/(3\alpha+4)} \to 0$ as $N\to\infty$. For $N>0$, let $M = M(N) = \lfloor \sqrt[4]{N} \rfloor$. Then*

$$\frac{1}{\sqrt{\sum_{n=M}^N \mathbb{E}U_{M,N,n}^2}} \sum_{n=M}^N U_{M,N,n}(X), \qquad X\sim \mathrm{Leb}_{\mathbb{T}^2}$$

converges weakly to a standard Gaussian as $N\to\infty$.

Proposition 5.4 is proved in the next section. Let us first complete the proof of Theorem B.

Throughout, $M = \lfloor \sqrt[4]{N} \rfloor$. For the remainder of §5, we specialize to the value $\eta = (2\alpha+2)/(3\alpha+4)$, noting that this value maximizes the function $\eta\mapsto \min\{2\eta-1, \alpha(1-\eta)(\alpha+2)\}$. In particular, $N^8 L_N^{-\min\{2\eta-1,\alpha(1-\eta)/(\alpha+2)\}} \to 0$ as $N\to\infty$ under the conditions of Proposition 5.4.

As we noted at the beginning of §5.1.2, it suffices to prove the CLT for $(1/\sqrt{N})S_{M,N}$ since here $M\approx N^{1/4} \ll \sqrt{N}$. Thus, to prove Theorem B, it suffices to check that:

(I) $\|S_{M,N} - \tilde{S}_{M,N}\|_{L^2} \to 0$ as $N\to\infty$; and

(II) $(1/N)\sum_{n=M}^N \mathbb{E}U_{M,N,n}^2 \to \sigma^2$ as $N\to\infty$, where σ^2 is as in Theorem B.

For (I), we estimate $\|S_{M,N} - \tilde{S}_{M,N}\|_{L^2}$ as follows:

$$\begin{aligned}\|S_{M,N} - \tilde{S}_{M,N}\|_{L^2} &\le C(N-M)\|\phi\|_0 \,\mathrm{Leb}(B_{M,N}) + C(N-M)\|\phi\|_\alpha \sum_{m=M}^N L_m^{-\eta\alpha} \\ &\le C(N-M)\|\phi\|_\alpha \sum_{m=M}^N L_m^{-\min\{\alpha\eta,1-\eta\}},\end{aligned}$$

applying first Proposition 5.3(b) and then Lemma 5.1. The above converges to 0 as $N\to\infty$ by the hypotheses of Proposition 5.4.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

For (II), we observe that

$$\sum_{n=M}^{N} \mathbb{E} U_{M,N,n}^2 = \mathbb{E}\bigg(\sum_{n=M}^{N} U_{M,N,n}\bigg)^2 = \int_{\mathbb{T}^2} \tilde{S}_{M,N}^2 d\,\mathrm{Leb}_{\mathbb{T}^2}$$
$$= \int_{\mathbb{T}^2} S_{M,N}^2 d\,\mathrm{Leb}_{\mathbb{T}^2} + O(\|S_{M,N} - \tilde{S}_{M,N}\|_{L^2}^2).$$

From (I), it follows that $\lim_{N\to\infty}(\|\tilde{S}_{M,N}\|_{L^2}^2 - \|S_{M,N}\|_{L^2}^2) = 0$. It remains to compute $\|S_{M,N}\|_{L^2}^2$, which we do below.

LEMMA 5.5. *Assume the setting of Proposition 5.4. With $M = M(N) = \lfloor \sqrt[4]{N} \rfloor$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \int S_{M,N}^2 d\,\mathrm{Leb} = \sigma^2 = \int \phi^2 + 2\int \phi(x,z)\phi(z,y)\,dx\,dy\,dz.$$

Proof. We have

$$\int S_{M,N}^2$$
$$= (N - M + 1)\int \phi^2 + 2\sum_{M\le m<n\le N} \int \phi\circ F_M^m \cdot \phi\circ F_M^n$$
$$= (N - M + 1)\int \phi^2 + 2\sum_{n=M+1}^{N} \int \phi\cdot\phi\circ F_n + 2\sum_{\substack{M\le m<n\le N\\ m<n-1}} \int \phi\cdot\phi\circ F_{m+1}^n d\,\mathrm{Leb}.$$

Applying Proposition 3.4(a) to the middle summation, we obtain the estimate

$$2(N-M)\int \phi(x,z)\phi(z,y)\,dx\,dy\,dz + O(\|\phi\|_\alpha^2 (N-M) L_M^{-\min\{2\eta-1,\alpha(1-\eta)/(2+\alpha)\}}).$$

Applying Proposition 3.4(b) to the (m, n)-summand in the third term,

$$\int \phi\cdot\phi\circ F_{m+1}^n d\,\mathrm{Leb} = O\bigg(\|\phi\|_\alpha^2\bigg(L_{m+1}^{-\min\{\alpha(1-\eta)/(2+\alpha),2\eta-1\}} + \sum_{k=m+2}^{n-1} L_k^{-1+\eta}\bigg)\bigg)$$
$$= O(\|\phi\|_\alpha^2 (N-M) L_M^{-\min\{\alpha(1-\eta)/(2+\alpha),2\eta-1\}}),$$

and applying the summation, the third term is bounded

$$O(\|\phi\|_\alpha^2 (N-M)^2 L_M^{-\min\{\alpha(1-\eta)/(2+\alpha),2\eta-1\}}).$$

All error terms go to 0 under the hypothesis of Proposition 5.4. □

5.2. *Proof of Proposition 5.4.* We use the following criterion for the CLT for arrays of martingale differences.

THEOREM 5.6. (McLeish) *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\{k_n\}_{n\ge1}$ be an increasing sequence of whole numbers tending to infinity and, for each $n \ge 1$, let $\mathcal{F}_{1,n} \subset \mathcal{F}_{2,n} \subset \cdots \subset \mathcal{F}_{k_n,n} \subset \mathcal{F}$ be an increasing sequence of sub-σ-algebras. For each such n, i, let $X_{i,n}$ be a random variable, measurable with respect to $\mathcal{F}_{i,n}$, for which $\mathbb{E}(X_{i,n}|\mathcal{F}_{i-1,n}) = 0$ and write $Z_n = \sum_{1\le i\le k_n} X_{i,n}$. Assume that:*

(a) $\max_{i \le k_n} |X_{i,n}|$ *is uniformly bounded, in n, in the* L^2 *norm;*
(b) $\max_{i \le k_n} |X_{i,n}| \to 0$ *in probability as* $n \to \infty$*; and*
(c) $\sum_i X_{i,n}^2 \to 1$ *in probability as* $n \to \infty$.
Then Z_n *converges weakly to a standard Gaussian.*

We apply this to the array

$$\frac{1}{\sqrt{\sum_{m=M(N)}^{N} \mathbb{E}U_{M(N),N,m}^2}} U_{M(N),N,n}(X), \quad M(N) \le n \le N, \quad X \sim \mathrm{Leb}_{\mathbb{T}^2},$$

where as before $M(N) = \lfloor \sqrt[4]{N} \rfloor$.

A preliminary asymptotic estimate for U_n is given in §5.2.1. The verification of (a)–(c) as in Theorem 5.6 is given in §5.2.2.

5.2.1. *An asymptotic estimate for* U_n. The following approximation is extremely useful in the coming arguments.

LEMMA 5.7. *Set* $\hat{U}_n = U_n \circ (F_M^{n-1})^{-1}$. *Then*

$$\hat{U}_n = \phi - \psi + \psi \circ F_{n+1} + O(N^2 \|\phi\|_\alpha L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}})$$

with uniform constants on $F_n^{-1} G_{M,n+1}$, *independently of n, where* $\psi(y) = \int \phi(\bar{x}, y)\, d\bar{x}$.

Proof. We have

$$\begin{aligned}\hat{U}_n = {} & \mathbb{E}(\phi | F_n^{-1} \mathcal{G}_{M,n+1}) - \mathbb{E}(\phi | \mathcal{G}_{M,n}) + \mathbb{E}(\phi | \mathcal{G}_{M,n+1}) \circ F_n \\ & + \sum_{m=n+1}^{N-1} \mathbb{E}(\phi \circ F_{n+1}^m | \mathcal{G}_{M,n+1}) \circ F_n + \sum_{m=n}^{N-1} \mathbb{E}(\phi \circ F_n^m | \mathcal{G}_{M,n}). \end{aligned} \tag{5}$$

As we will show, the terms in the top line approximate to $\phi - \psi + \psi \circ F_{n+1}$, while the terms in the second line are small.

For the first term in (5), we have from Lemma 5.2 that $|\mathbb{E}(\phi | F_n^{-1} \mathcal{G}_{M,n+1}) - \phi| = O(\|\phi\|_\alpha L_n^{-\alpha\eta})$ on $F_n^{-1} G_{M,n+1}$.

For the second term in (5), we have that

$$\mathbb{E}(\phi | \mathcal{G}_{M,n}) = \frac{1}{\mathrm{Len}(\gamma)} \int_\gamma \phi \, d\, \mathrm{Leb}_\gamma$$

on $G_{M,n}$, where γ is a fully crossing horizontal curve with $\|h'_\gamma\|_{C^0} \le L_{n-1}^{-\eta}$. Let now $p \in \gamma$, $p = (x_0, y_0)$. Noting that $|\phi(x, h_\gamma(x)) - \phi(x, y_0)| \le \|\phi\|_\alpha |h_\gamma(x) - h_\gamma(x_0)|^\alpha \le C\|\phi\|_\alpha L_{n-1}^{-\alpha\eta}$, we have

$$\begin{aligned}\frac{1}{\mathrm{Len}(\gamma)} \int_\gamma \phi d\, \mathrm{Leb}_\gamma &= (1 + O(\|\phi\|_\alpha L_{n-1}^{-\eta})) \int_0^1 \phi(x, h_\gamma(x))\, dx \\ &= (1 + O(\|\phi\|_\alpha L_{n-1}^{-\alpha\eta})) \int_0^1 \phi(x, y_0)\, dx;\end{aligned}$$

we therefore conclude that

$$|\mathbb{E}(\phi | \mathcal{G}_{M,n}) - \psi| \le C\|\phi\|_\alpha L_{n-1}^{-\alpha\eta}$$

on $G_{M,n}$. Similarly, for the third term in (5), we obtain the bound

$$|\mathbb{E}(\phi|\mathcal{G}_{M,n+1}) \circ F_n - \psi \circ F_n| \le C\|\phi\|_\alpha L_n^{-\alpha\eta}$$

on $F_n^{-1} G_{M,n+1}$.

For the fourth term in (5), we estimate from Proposition 2.8 that on $G_{M,n}$,

$$\begin{aligned}\mathbb{E}(\phi \circ F_n^m|\mathcal{G}_{M,n}) &= \frac{1}{\text{Len}(\gamma)} \int_\gamma \phi \circ F_n^m d\,\text{Leb}_\gamma \\ &= O\bigg(\|\phi\|_\alpha \cdot \bigg(L_m^{-\alpha(1-\eta)/(2+\alpha)} + L_n^{1-2\eta} + \sum_{k=n}^{m-1} L_k^{-1+\eta}\bigg)\bigg) \\ &= O(\|\phi\|_\alpha (N-M) L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}})\end{aligned}$$

for some $\gamma \in \zeta_{M,n}$. Estimating similarly the fifth term in (5), we deduce that on $F_n^{-1} G_{M,n+1}$, the contribution of the fourth and fifth terms combined is

$$O(\|\phi\|_\alpha (N-M)^2 L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}}). \qquad \square$$

5.2.2. *Verifying properties (a)–(c) in Theorem 5.6.*

Properties (a) and (b). By Lemma 5.7, we have that on $(F_M^{N-1})^{-1} G_{M,N}$,

$$\begin{aligned}|U_n| &= O(\|\phi\|_{C^0} + \|\phi\|_\alpha N^2 L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}}) \\ &= O(\|\phi\|_\alpha N^2 L_M^{-\min\{\alpha(1-\eta)/(2+\alpha), 2\eta-1\}}),\end{aligned}$$

which is uniformly bounded in n, N. Property (b) is now immediate since $\text{Leb}(G_{M,N}) \to 1$ as $N \to \infty$.

For property (a), off $(F_M^{N-1})^{-1} G_{M,N}$, we have

$$|U_n| \le CN\|\phi\|_\alpha,$$

so

$$\bigg\| \max_{M \le n \le N} |U_{M,N,n}| \bigg\|_{L^2} \le C\|\phi\|_\alpha \cdot N\sqrt{\text{Leb}(G_{M,N}^c)} + C\|\phi\|_\alpha.$$

Property (a) follows from the estimate $\text{Leb}(G_{M,N}^c) = O(\sum_M^{N-1} L_k^{-1+\eta}) = O((N-M) L_M^{-1+\eta})$ in Lemma 5.1. $\square$

Below is a formulation of property (c).

PROPOSITION 5.8. (Strong law for $\{U_n^2\}$) *We have*

$$\lim_{N\to\infty} \frac{\sum_{n=M}^N U_{M,N,n}^2}{\mathbb{E} \sum_{n=M}^N U_{M,N,n}^2} = 1$$

in probability.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Proof. We prove the stronger property of convergence in L^2. To start, we evaluate

$$\int \bigg(\sum_M^N U_n^2 - \sum_M^N \mathbb{E}(U_n^2) \bigg)^2 d\,\mathrm{Leb}$$
$$= \sum_{M \le m,n \le N} \int (U_n^2 - \mathbb{E}(U_n^2))(U_m^2 - \mathbb{E}(U_m^2))\, d\,\mathrm{Leb}$$
$$= \sum_{n=M}^N (\mathbb{E}(U_n^4) - \mathbb{E}(U_n^2)^2) + 2 \sum_{M \le m < n \le N} \int (\hat{U}_m^2 - \mathbb{E}(U_m^2))(\hat{U}_n^2 - \mathbb{E}(U_n^2)) \circ F_m^{n-1} d\,\mathrm{Leb}.$$

We start with bounding $\mathbb{E}(U_n^2), \mathbb{E}(U_n^4)$. For N sufficiently large, we have on $(F_M^{N-1})^{-1} G_{M,N}$ that $|U_n| = O(\|\phi\|_\alpha)$ by Lemma 5.7, while on the complement we have $|U_n| = O(N\|\phi\|_\alpha)$ and so, applying the estimate on $\mathrm{Leb}(G^c_{M,N})$, we obtain

$$\mathbb{E}(U_n^2) = O(\|\phi\|_\alpha^2 (N^3 L_M^{-1+\eta} + 1)) \quad \text{and} \quad \mathbb{E}(U_n^4) = O(\|\phi\|_\alpha^4 (N^5 L_M^{-1+\eta} + 1)).$$

Thus, the first summation is bounded like

$$O(\|\phi\|_\alpha^4 N (N^5 L_M^{-1+\eta} + 1)).$$

For the second summation, let us write $\phi_*(x, y) := \phi(x, y) - \psi(y) + \psi(x)$ in the notation of Lemma 5.7. Since this quantity appears repeatedly, let us also use the shorthand $c = \alpha/(3\alpha + 4)$, noting that under the hypotheses of Theorem B we have that $N^2 L_M^{-c} \to 0$ as $N \to \infty$. We estimate

$$\hat{U}_n^2 - \phi_*^2 = (\hat{U}_n + \phi_*)(\hat{U}_n - \phi_*) = O(\|\phi\|_\alpha^2 N^2 L_M^{-c}(1 + N^2 L_M^{-c})) = O(\|\phi\|_\alpha^2 N^2 L_M^{-c})$$

on $(F_n^{N-1})^{-1} G_{M,N}$ and so

$$|\mathbb{E}(\hat{U}_n^2) - \mathbb{E}(\phi_*^2)| \le C N^3 \|\phi\|_\alpha^2 L_M^{-1+\eta} + C \|\phi\|_\alpha^2 N^2 L_M^{-c} = O(\|\phi\|_\alpha^2 (N^3 L_M^{-1+\eta} + N^2 L_M^{-c});$$

hence,

$$|\mathbb{E}(\hat{U}_n^2)\mathbb{E}(\hat{U}_m^2) - \mathbb{E}(\phi_*^2)^2| \le \mathbb{E}(\hat{U}_n^2)|\mathbb{E}(\hat{U}_m^2) - \mathbb{E}(\phi_*^2)| + \mathbb{E}(\phi_*^2)|\mathbb{E}(\hat{U}_n^2) - \mathbb{E}(\phi_*^2)|$$
$$= O(\|\phi\|_\alpha^4 (1 + N^3 L_M^{-1+\eta})(N^3 L_M^{-1+\eta} + N^2 L_M^{-c})).$$

On $(F_m^{N-1})^{-1} G_{M,N}$, we have

$$|\hat{U}_m^2 \cdot \hat{U}_n^2 \circ F_m^{n-1} - \phi_*^2 \cdot \phi_*^2 \circ F_m^{n-1}|$$
$$\le \hat{U}_m^2 |\hat{U}_n^2 \circ F_m^{n-1} - \phi_*^2 \circ F_m^{n-1}| + \phi_*^2 \circ F_m^{n-1} \cdot |\hat{U}_m^2 - \phi_*^2|$$
$$= O(\|\phi\|_\alpha^4 (1 + N^2 L_M^{-c}) N^2 L_M^{-c}) = O(\|\phi\|_\alpha^4 N^2 L_M^{-c}).$$

Collecting,

$$\int \hat{U}_m^2 \hat{U}_n^2 \circ F_m^{n-1} - \mathbb{E}(U_m^2)\mathbb{E}(U_n^2) - \bigg(\int \phi_*^2 \cdot \phi_*^2 \circ F_m^{n-1} - \mathbb{E}(\phi_*^2)^2 \bigg)$$
$$= O(\|\phi\|_\alpha^4 (1 + N^3 L_M^{-1+\eta})(N^3 L_M^{-1+\eta} + N^2 L_M^{-c})).$$

Applying now Proposition 3.4(b), we obtain the estimate

$$\bigg| \int \phi_*^2 \cdot \phi_*^2 \circ F_m^{n-1} - \bigg(\int \phi_*^2 \bigg)^2 \bigg| = O(\|\phi\|_\alpha^4 (N L_M^{-1+\eta} + L_M^{-c})),$$

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

so we conclude that

$$\int \hat{U}_m^2 \hat{U}_n^2 \circ F_m^{n-1} - \mathbb{E}(U_m^2)\mathbb{E}(U_n^2) = O(\|\phi\|_\alpha^4 (1 + N^3 L_M^{-1+\eta})(N^3 L_M^{-1+\eta} + N^2 L_M^{-c})).$$

Summing over the $\approx N^2$ terms and noting that $(\sum_M^N \mathbb{E}(U_n^2))^2 \approx \sigma^4 N^2$ for N large, we obtain

$$\frac{1}{\|\phi\|_\alpha^4} \left\| \frac{\sum_M^N U_n^2}{\sum_M^N \mathbb{E} U_n^2} - 1 \right\|_{L^2}^2$$
$$= O((1 + N^3 L_M^{-1+\eta})(N^3 L_M^{-1+\eta} + N^2 L_M^{-c}) + N^{-1} + N^4 L_M^{-1+\eta}).$$

The proof goes through if all terms on the right-hand side go to 0 as $N \to \infty$. For this, it suffices that $N^2 L_M^{-c} \to 0$ as $N \to \infty$: to see this, observe that $N^4 L_M^{-1+\eta} \leq N^4 L_M^{-2c}$ holds for any $\eta \in (1/2, 1)$, $\alpha \in (0, 1)$. The latter clearly goes to 0 when $N^2 L_M^{-c} \to 0$. □

6. *Hyperbolicity and the shape of successive iterates of a set*

We close this paper with the proof of Theorem C, given in §§6 and 7.

We argued in §2 (see Remark 2.9) that fully crossing horizontal curves proliferate throughout phase space in a roughly uniform way and that this proliferation is the mixing mechanism for the compositions $\{F^n\}$. In this section we flesh out this picture by showing the following: given a set $S \subset \mathbb{T}^2$ with a suitably nice boundary and n large enough, the nth image $F^n(S)$ is 'mostly' foliated by disjoint fully crossing horizontal curves.

The plan is as follows. In §6.1 we construct for each n a foliation of $S_n = F^{n-1}S$ by horizontal curves. It is shown in §6.2 that for n sufficiently large, a large proportion of the curves in the foliation of S_n are 'sufficiently long', in the sense that in one time step such curves become fully crossing. In §6.3 we show that on disintegrating Lebesgue measure restricted to S_n, the disintegration densities on the leaves of our horizontal foliation are controlled. These results are synthesized in Proposition 6.11 in §6.4, the main result of this section.

This last result is a primary ingredient in the proof of Theorem C, the proof of which will be completed in §7.

6.1. *Construction of foliations by horizontal curves.*

Let $S \subset \mathbb{T}^2$ be an open subset and write ν_S for normalized Lebesgue measure on S. Our aim is to build a foliation of the nth image $F^n(S)$ by horizontal curves with the property that for n sufficiently large, 'most' of the foliating curves are sufficiently long.

6.1.1. *Standing assumptions for §6.*

The parameter $\eta \in (1/2, 1)$ is fixed. The open set $S \subset \mathbb{T}^2$ is such that the topological boundary $\partial S = \bar{S} \backslash S$ is the finite union of smooth curves and, moreover, is assumed to have the following property: for any $l > 0$,

$$\nu_S\{p \in S : d(p, \partial S) \leq l\} \leq C_S l, \tag{6}$$

where $C_S > 0$ is a constant independent of l. Let us write $S_1 = S$ and $F^{n-1}S_1 = S_n$ for $n \geq 1$, noting that $\partial S_n = F^{n-1}\partial S_1$ since each F^n is a diffeomorphism.

For $n \geq 1$, we write $\mathcal{B}_n$ for the partition of $\mathbb{T}^2$ into the connected components of B_n and B_n^c, noting that each is a partition of $\mathbb{T}^2$ into vertical cylinders (sets of the form $I \times \mathbb{T}^1$ for

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

a proper connected subinterval $I \subset \mathbb{T}^1$). We also abuse notation somewhat and write $\partial \mathcal{B}_n$ for the union of the boundaries of each atom of $\mathcal{B}_n$; that is, $\partial \mathcal{B}_n$ is the union of circles of the form $\{\hat{x}_n \pm 2K_1 L_n^{-1+\eta}\} \times \mathbb{T}^1$ as $\hat{x}_n$ varies over $\mathcal{C}_n$.

Define the sequence of partitions $\{\mathcal{P}_n\}_{n\geq 1}$ of $\mathbb{T}^2$ as follows:

$$\mathcal{P}_1 = \mathcal{B}_1 \vee \{S_1, S_1^c\}$$

and, for $n \geq 2$,

$$\mathcal{P}_n = \mathcal{B}_n \vee F_{n-1}(\mathcal{P}_{n-1}).$$

Above, $\vee$ refers to the *join* of partitions. Hereafter for $q \in \mathbb{T}^2$, we write $\mathcal{P}_n(q)$ for the atom of $\mathcal{P}_n$ containing q. Again we abuse notation somewhat and write $\partial \mathcal{P}_n$ for the union over the collection of boundaries of each atom comprising $\mathcal{P}_n$.

Additional notation. For $q = (x, y) \in \mathbb{T}^2$, let us write $H_q = \mathbb{T}^1 \times \{y\}$ for the horizontal circle containing q. When $\mathcal{P}$ is a partition of $\mathbb{T}^2$ and $p \in \mathbb{T}^2$, we write $\mathcal{P}(p)$ for the atom of $\mathcal{P}$ containing p. We write '$\leq$' for the partial order on partitions: for partitions $\mathcal{P}$, $\mathcal{Q}$, we write $\mathcal{P} \leq \mathcal{Q}$ if each atom in $\mathcal{P}$ is a union of $\mathcal{Q}$ atoms.

6.1.2. *Algorithm for foliating S_n by horizontal curves.* We now define, for each $n \geq 1$, a foliation (partition) $\hat{\gamma}_n$ of S_n by horizontal curves.

For $n = 1$, we define $\hat{\gamma}_1$ to be the partition of S_1 consisting of atoms of the form

$$\hat{\gamma}_1(p) = H_p \cap \mathcal{P}_1(p)$$

for $p \in S_1$. Clearly $\hat{\gamma}_1$ is a measurable partition of S_1 and $\hat{\gamma}_1 \leq \mathcal{P}_1|_{S_1}$ (here $\leq$ indicates the partial order on partitions in terms of refinement and $\mathcal{P}_1|_{S_1}$ denotes the restriction of $\mathcal{P}_1$ to S_1). Inductively, assume that $\hat{\gamma}_1, \ldots, \hat{\gamma}_n$ have been constructed and that $\hat{\gamma}_n \geq \mathcal{P}_n|_{S_n}$. To define $\hat{\gamma}_{n+1}(p_{n+1})$ for $p_{n+1} \in S_{n+1}$, we distinguish two cases. Below, we write $p_n = F_n^{-1}(p_{n+1})$.

Case 1. $p_n \notin B_n$. By construction, $\hat{\gamma}_n(p_n) \cap B_n = \emptyset$ and so $F_n(\hat{\gamma}_n(p_n))$ is a horizontal curve (Lemma 2.4). In preparation for the next iterate, we cut this image curve by $\mathcal{P}_{n+1}$; that is,

$$\hat{\gamma}_{n+1}(p_{n+1}) = F_n(\hat{\gamma}_n(p_n)) \cap \mathcal{P}_{n+1}(p_{n+1}).$$

Equivalently, $\hat{\gamma}_{n+1}|_{F_n(B_n^c \cap S_n)} = F_n(\hat{\gamma}_n \cap B_n^c) \vee \mathcal{P}_{n+1}|_{F_n(B_n^c \cap S_n)}$.

Case 2. $p_n \in B_n$. In this case $\hat{\gamma}_n(p_n) \subset B_n$ and so we lose our control on the image curve $F_n(\hat{\gamma}_n(p_n))$. The procedure here is to repartition the entire image of B_n by horizontal line segments cut by $\mathcal{P}_{n+1}$ in preparation for the next iterate. Precisely, we define

$$\hat{\gamma}_{n+1}(p_{n+1}) = H_{p_{n+1}} \cap \mathcal{P}_{n+1}(p_{n+1}).$$

Equivalently, $\hat{\gamma}_{n+1}|_{F_n(B_n \cap S_n)}$ is the join of $\mathcal{P}_{n+1}|_{F_n(B_n \cap S_n)}$ with the partition of $F_n(B_n)$ into horizontal circles (sets of the form $\mathbb{T}^1 \times \{y\} \subset \mathbb{T}^2$ for $y \in \mathbb{T}^1$).

This induction procedure bootstraps because $\hat{\gamma}_{n+1}$ is a partition of S_{n+1} into horizontal curves for which $\hat{\gamma}_{n+1} \geq \mathcal{P}_{n+1}|_{S_{n+1}}$. All partitions mentioned are measurable [**28**] and so we have the following.

LEMMA 6.1. *For each $n \geq 1$, the partition $\hat{\gamma}_n$ of S_n as above is defined and is a measurable partition of S_n into connected, smooth horizontal curves for which $\hat{\gamma}_n \geq \mathcal{P}_n|_{S_n}$.*

6.2. *Estimating time to curve length growth.* As indicated in the procedure laid out above, the curves of $\hat{\gamma}_{n+1}$ coming from $\hat{\gamma}_n|_{S_n \cap B_n^c}$ have been elongated by the strong expansion of F_n along horizontal directions. However, this elongation of curves competes with the 'cutting' of curves near bad sets (case 1) and the occasional 'repartitioning' of the images of the bad sets $S_n \cap B_n$ by horizontal line segments (case 2). Our aim now is to show that for large n, the expansion wins out and 'most' of the curves comprising the foliation $\hat{\gamma}_n$ are of sufficiently long horizontal extent.

6.2.1. *Preparations.* For a connected C^1 curve $\gamma \subset \mathbb{T}^2$ and a point $q = (x, y) \in \gamma$, we define

$$\mathrm{Rad}_q(\gamma) = d_\gamma(q, \partial\gamma).$$

Here d_γ denotes the Euclidean distance on γ and $\partial\gamma$ denotes the end points of γ; that is, if $\gamma = \mathrm{graph}\, h_\gamma$ for $h_\gamma : I_\gamma \to \mathbb{T}^1$, then $\partial\gamma = \{(\hat{x}, h_\gamma(\hat{x})) : \hat{x} \in \partial I_\gamma\}$. Recall that $I_\gamma \subset \mathbb{T}^1$ is always a proper connected subarc, so ∂I_γ, and hence $\partial\gamma$, consists of exactly two points.

Additionally, let us define the following alternative of the time τ defined in §2.1: for $p \in \mathbb{T}^2$, we define

$$\begin{aligned}\bar{\tau}(p) &= 1 + \max\{m \geq 1 : d(F^{m-1}(x, y), B_m) < K_1 L_m^{-1+\eta'}\} \\ &= \min\{k \geq 1 : d(F^{n-1}(p), B_n) \geq K_1 L_n^{-1+\eta'} \text{ for all } n \geq k\}.\end{aligned}$$

Here we have set

$$\eta' = \frac{\eta + 1}{2}.$$

Clearly $\tau \leq \bar{\tau}$. A straightforward variation of the argument for Lemma 2.2 implies that $\bar{\tau}$ is almost surely finite and satisfies an analogous tail estimate to that of τ whenever $\sum_n L_n^{-1+\eta'} < \infty$. Precisely, we have

$$\mathrm{Leb}\{\bar{\tau} > N\} \leq \sum_{n=N}^{\infty} 6K_1 L_n^{-1+\eta'} = O\bigg(\sum_{n \geq N} L_n^{-1+\eta'}\bigg). \tag{7}$$

For the remainder of §6, we shall assume that the sequence $\{L_n\}$ is such that the right-hand side of (7) is finite.

6.2.2. *The curve growth time σ_S.*

Definition 6.2. Given $p \in S_1$, we define the *curve growth time* $\sigma_S(p)$ by

$$\sigma_S(p) = \min\{k \geq \bar{\tau}(p) : \mathrm{Rad}_{p_k}(\hat{\gamma}_k(p_k)) \geq K_1 L_k^{-1+\eta'}\},$$

where above we write $p_k = F^{k-1}(p)$.

In this section we write $\sigma = \sigma_S$ for short.

Our definition of σ is motivated by the following consideration. Let $p \in S_1$, $p_n = F^{n-1}(p)$ and assume that $\sigma(p) = n$. Then $\hat{\gamma}_n(p_n) \cap B_n = \emptyset$ and $|I_{\hat{\gamma}_n(p_n)}| \geq 2K_1 L_n^{-1+\eta'}$: this implies that $F_n(\hat{\gamma}_n(p_n))$ is a union of approximately $L_n^{2\eta-1} \gg 1$ fully crossing horizontal curves. Thus, σ has the connotation of a *mixing time*: the set $\{\sigma \leq n\} \subset S$ is a region of S which has proliferated throughout $\mathbb{T}^2$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

A possible obstruction to mixing is that once this mass has proliferated, it could become 'trapped' again by the bad sets B_n. This is not possible, however, due to the way that σ is defined. Precisely, we have the following.

LEMMA 6.3. *Let $p \in S$ and assume that $\sigma(p) = n$ for some $n \geq 1$. Then* $\mathrm{Rad}_{p_k}(\hat{\gamma}_k(p_k)) \geq K_1 L_k^{-1+\eta'}$ *for all $k \geq n$.*

Proof. It suffices to show that for any $k \geq \bar{\tau}(p)$, we have that $\mathrm{Rad}_{p_k}(\hat{\gamma}_k(p_k)) \geq K_1 L_k^{-1+\eta'}$ implies that $\mathrm{Rad}_{p_{k+1}}(\hat{\gamma}_k(p_k)) \geq K_1 L_k^{-1+\eta'}$. This is implied directly by Lemma 2.7. □

The main result of §6.2 is the following estimate on the tail of σ.

PROPOSITION 6.4. *There is a constant C, depending only on K_1, M_0, such that the following holds. Let L_0 be sufficiently large. Then, for any $n \geq 1$, we have that*

$$\nu\{\sigma(p) > 4n\} \leq \left(\frac{C}{\mathrm{Leb}(S)} + C_S\right) \sum_{i=n}^{\infty} L_i^{-1+\eta'}.$$

Proposition 6.4 bears a strong resemblance to the volume lemma in billiard dynamics, used to control the lengths of unstable manifolds; see, e.g., [**12**].

Remark 6.5. Let us draw a comparison between the present situation and that of a typical non-uniformly hyperbolic system for which correlation decay and statistical properties are known, e.g., systems admitting Young towers with controllable 'good' return times to their bases [**34**]. Roughly speaking, the typical situation is that a given 'lump' of mass can fail to proliferate: for example, nice hyperbolic geometry can be spoiled (as happens for Henon maps; see, e.g., [**7**]) or mass may become 'trapped' somewhere (as happens for intermittent maps; see, e.g., [**25**]). In a typical situation admitting a Young tower, a given 'lump' of mass experiences infinitely many 'proliferations' (returns to the base), followed by some possibly unbounded 'reset' time (sojourn up the tower) before the next proliferation takes place. Thus, correlation decay estimates depend critically on the delicate balance between these two behaviors.

In contrast, the situation for our composition $\{F^n\}$ is simpler: at any time, some positive proportion of ν_n is 'trapped' in a bad region, but, as time evolves, an increasingly larger proportion of the mass of ν_n has 'permanently proliferated' throughout $\mathbb{T}^2$.

6.2.3. *Proof of Proposition 6.4.* We require two estimates:

(A) for any $p_n \in S_n$, $n \geq 1$, a 'bad' a priori estimate on $\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n))$; and

(B) for Leb-almost every $p \in S_1$, a 'good' estimate for $\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n))$ for $n \gg \bar{\tau}(p)$ (where $p_n = F^{n-1}(p)$.

Afterwards, we will (C) synthesize these estimates to obtain the desired estimate on the tail of σ.

Let us briefly elaborate on this strategy. Before time $\bar{\tau}(p)$, we have no control whatsoever on the orbit of p and so our procedure may indeed produce very short curves $\hat{\gamma}_n(p_n)$, $p_n = F^{n-1}(p)$ for such n. As a result, we have access to only the 'worst possible' estimates for $\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n))$. We carry these estimates out in (A) below. Once $\bar{\tau}(p)$ has

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

elapsed, we will leverage our control on the orbit of p after time $\bar{\tau}(p)$ to grow the curves $\hat{\gamma}_n(p_n)$ to sufficient horizontal extent—this is carried out in part (B).

(A) 'Bad' a priori length estimate for $\hat{\gamma}_n(p_n)$ for all n. Here we prove the following estimate.

LEMMA 6.6. *Let $p_1 \in S_1$ and write $p_k = F^{k-1} p_1$ for $k > 1$. Then, for any $n \geq 1$,*

$$\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n)) \geq \min\left\{ \min_{1 \leq i \leq n} \left\{ \left(\prod_{j=i}^{n-1} 2K_0 L_j \right)^{-1} d(p_i, \partial\mathcal{B}_i) \right\}, \left(\prod_{j=1}^{n-1} 2K_0 L_j \right)^{-1} d(p_1, \partial S_1) \right\}.$$

Lemma 6.6 will be obtained from the corresponding identical estimate for $d(p_n, \partial\mathcal{P}_n)$.

LEMMA 6.7. *In the setting of Lemma 6.6, we have*

$$d(p_n, \partial\mathcal{P}_n) \geq \min\left\{ \min_{1 \leq i \leq n} \left\{ \left(\prod_{j=i}^{n-1} 2K_0 L_j \right)^{-1} d(p_i, \partial\mathcal{B}_i) \right\}, \left(\prod_{j=1}^{n-1} 2K_0 L_j \right)^{-1} d(p_1, \partial S_1) \right\}.$$

In both of Lemmas 6.6 and 6.7, the empty product $\prod_{j=n}^{n-1}$ is to interpreted as equal to 1.

Proof of Lemma 6.7. To prove this estimate, recall that for $k \geq 1$, we have $\partial\mathcal{P}_k = \partial\mathcal{B}_k \cup F_{k-1}(\partial\mathcal{P}_{k-1})$; thus,

$$d(p_k, \partial\mathcal{P}_k) = \min\{d(p_k, \partial\mathcal{B}_k), d(p_k, F_{k-1}(\partial\mathcal{P}_{k-1})\}.$$

Noting that $\mathrm{Lip}(F_{k-1}^{-1}) \leq 2K_0 L_{k-1}$, we obtain

$$d(p_k, F_{k-1}(\partial\mathcal{P}_{k-1})) \geq (2K_0 L_{k-1})^{-1} d(p_{k-1}, \partial\mathcal{P}_{k-1}).$$

Thus, for all $n \geq 2$, we obtain the following. Below, we write $a \wedge b = \min\{a, b\}$ for short.

$$\begin{aligned} d(p_n, \partial\mathcal{P}_n) &\geq \min\{d(p_n, \partial\mathcal{B}_n), (2K_0 L_{n-1})^{-1} d(p_{n-1}, \partial\mathcal{P}_{n-1})\} \\ &\geq \min\{d(p_n, \partial\mathcal{B}_n), (2K_0 L_{n-1})^{-1} d(p_{n-1}, \partial\mathcal{B}_{n-1}), \\ &\quad (2K_0 L_{n-1})^{-1}(2K_0 L_{n-2})^{-1} d(p_{n-2}, \partial\mathcal{P}_{n-2})\} \\ &\geq \cdots \geq d(p_n, \partial\mathcal{B}_n) \wedge \min_{2 \leq i \leq n-1} \left\{ \left(\prod_{j=i}^{n-1} 2K_0 L_j \right)^{-1} d(p_i, \partial\mathcal{B}_i) \right\} \\ &\wedge \left(\prod_{j=1}^{n-1} 2K_0 L_j \right)^{-1} d(p_1, \partial\mathcal{P}_1). \end{aligned}$$

The desired estimate now follows from the fact that $\partial\mathcal{P}_1 = \partial S_1 \cup \partial\mathcal{B}_1$. □

Proof of Lemma 6.6. With $n \in \mathbb{N}$ fixed, define

$$n_1 = \max\{1 \leq k \leq n-1 : p_k \in B_k\},$$

where we use the ad hoc convention $n_1 = 1$ if $p_k \notin B_k$ for all $1 \le k \le n-1$. Observe that $\hat{\gamma}_{n_1+1}(p_{n_1+1})$ is formed by using Case 2 in the algorithm and that $\hat{\gamma}_k(p_k)$ is formed by using Case 1 for every $k \ge n_1 + 2$. In particular,

$$\mathrm{Rad}_{p_{n_1+1}}(\hat{\gamma}_{n_1+1}(p_{n_1+1})) \ge d(p_{n_1+1}, \partial\mathcal{P}_{n_1+1})$$

and, for every $n_1 + 2 \le k \le n$, we have

$$\mathrm{Rad}_{p_k}(\hat{\gamma}_k(p_k)) \ge \min\{d(p_k, \partial\mathcal{P}_k), \mathrm{Rad}_{p_k}(F_{k-1}(\hat{\gamma}_{k-1}(p_{k-1}))\}.$$

To prove Lemma 6.6, it suffices to show that

$$\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n)) \ge \min_{n_1+1 \le k \le n}\{d(p_k, \partial\mathcal{P}_k)\}. \tag{8}$$

Once (8) is proved, Lemma 6.6 follows on inserting the estimates for $d(p_k, \partial\mathcal{P}_k)$ for $1 \le k \le n$.

Turning to (8): if $n_1 = n-1$, then there is nothing left to show. If $n_1 < n-1$, then we estimate

$$\begin{aligned}\mathrm{Rad}_{p_n}(\hat{\gamma}_n) &\ge d(p_n, \partial\mathcal{P}_n) \wedge L^{\eta}_{n-1}\,\mathrm{Rad}_{p_{n-1}}(\hat{\gamma}_{n-1}(p)),\\ &\ge d(p_n, \partial\mathcal{P}_n) \wedge L^{\eta}_{n-1} d(p_{n-1}, \partial\mathcal{P}_{n-1}) \wedge L^{\eta}_{n-1}L^{\eta'}_{n-2}\,\mathrm{Rad}_{p_{n-2}}(\hat{\gamma}_{n-2}) \ge \cdots\\ &\ge d(p_n, \partial\mathcal{P}_n) \wedge \min_{n_1+2 \le i \le n-1}\left\{\left(\prod_{j=i}^{n-1} L^{\eta}_j\right) d(p_i, \partial\mathcal{P}_i)\right\}\\ &\wedge \mathrm{Rad}_{p_{n_1+1}}(\hat{\gamma}_{n_1+1}(p_{n_1+1})).\end{aligned}$$

Here we have used the simple estimate

$$\mathrm{Rad}_{p_{j+1}}\, F_j(\hat{\gamma}_j(p_j)) \ge L^{\eta}_j\,\mathrm{Rad}_{p_j}(\hat{\gamma}_j(p_j)), \tag{9}$$

which follows from the expansion estimate along horizontal curves in Lemma 2.4. Replacing all L^{η}_j terms with 1, we obtain (8). □

(B) Good length estimate for $\hat{\gamma}_n(p_n)$ for $n \gg \tau(p)$. Here we prove the following.

LEMMA 6.8. *Let $N \ge 1$ and let $p \in S_1$ be such that $\bar{\tau}(p) \le N < \infty$. Then, for any $n \ge N$,*

$$\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n)) \ge \min\left\{d(p_n, \partial\mathcal{B}_n), \left(\prod_{k=N}^{n-1} L^{\eta}_k\right)\mathrm{Rad}_{p_N}(\hat{\gamma}_N(p_N))\right\}. \tag{10}$$

Proof of Lemma 6.8. The proof leans on the following claim.

CLAIM 6.9. *Let $p \in S_1$ be such that $\bar{\tau}(p) \le N < \infty$. Then, for all $n \ge N$, we have*

$$\mathrm{Rad}_{p_{n+1}}(\hat{\gamma}_{n+1}(p_{n+1})) \ge \min\{d(p_{n+1}, \partial\mathcal{B}_{n+1}), \mathrm{Rad}_{p_n}(F_n(\hat{\gamma}_n(p_n)))\}.$$

Proof of Claim. Observe that since $n \ge \bar{\tau}(p) \ge \tau(p)$, we always use Case 1 in the construction of $\hat{\gamma}_{n+1}(p_{n+1})$, i.e., $\hat{\gamma}_{n+1}(p_{n+1}) = F_n(\hat{\gamma}_n(p_n)) \cap \partial\mathcal{P}_{n+1}(p_{n+1})$. Moreover, $\hat{\gamma}_n(p_n) \subset \mathcal{P}_n(p_n)$ by construction and hence $F_n(\hat{\gamma}_n(p_n)) \subset F_n(\mathcal{P}_n(p_n))$ and so we arrive at

$$\hat{\gamma}_{n+1}(p_{n+1}) = F_n(\hat{\gamma}_n(p_n)) \cap \mathcal{B}_{n+1}(p_{n+1}).$$

The desired estimate now follows. □

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Fixing $n \geq N$, we now estimate

$$\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n)) \geq \min\{d(p_n, \partial\mathcal{B}_n), \mathrm{Rad}_{p_n}(F_{n-1}(\hat{\gamma}_{n-1}(p_{n-1})))\}.$$

Observe that since $d(p_{n-1}, \partial\mathcal{B}_{n-1}) \geq K_1 L_{n-1}^{-1+\eta}$, it follows that

$$\mathrm{Rad}_{p_n}(F_{n-1}(\hat{\gamma}_{n-1}(p_{n-1}))) \geq L_{n-1}^{\eta} \, \mathrm{Rad}_{p_{n-1}}(\hat{\gamma}_{n-1}(p_{n-1})),$$

on applying (9). Iterating,

$$\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n)) \geq d(p_n, \mathcal{B}_n) \wedge \min_{N \leq k \leq n-1} \left\{ \left(\prod_{i=k}^{n-1} L_k^{\eta} \right) d(p_k, \partial\mathcal{B}_k) \right\}$$
$$\wedge \left(\prod_{i=N}^{n-1} L_i^{\eta} \right) \mathrm{Rad}_{p_N}(\hat{\gamma}_N(p_N)).$$

Note however that for $N \leq k \leq n-1$, we have that

$$d(p_k, \partial\mathcal{B}_k) \geq K_1 L_k^{-1+\eta'}$$

and hence $L_k^{\eta} \cdot d(p_k, \partial\mathcal{B}_k) \geq K_1 L^{2(\eta+\eta')-1} \gg 1$ (recall that $\eta > 1/2$, so $\eta + \eta' - 1 > 2\eta - 1 > 0$) when L_0 is sufficiently large in terms of K_1, η. This yields the desired estimate. □

(C) Final estimates on the tail of σ. We are now in position to prove our estimate on $\mathrm{Leb}\{p \in S_1 : \sigma(p) > 4n\}$. Assume that $p \in S_1$ and $\bar{\tau}(p) \leq n < \infty$; finally, assume that $\sigma(p) > 4n$. From Lemma 6.8, it follows that

$$\mathrm{Rad}_{p_n}(\hat{\gamma}_n(p_n)) < K_1 L_{4n}^{-1+\eta'} \cdot \left(\prod_{k=n}^{4n-1} L_k^{\eta} \right)^{-1} \leq \left(\prod_{k=n}^{4n-1} L_k^{\eta} \right)^{-1}$$

for L_0 sufficiently large since here we always have $d(p_{4n}, \partial\mathcal{B}_{4n}) \geq K_1 L_{4n}^{-1+\eta'}$ by definition of $\bar{\tau}, \sigma$. Inserting our estimate from Lemma 6.6, there are two cases to consider.

Case (a). For some $1 \leq k \leq n$, we have

$$d(p_k, \partial\mathcal{B}_k) < \frac{\prod_{i=k}^{n-1} 2K_0 L_i}{\prod_{i=n}^{4n-1} L_i^{\eta}}$$

(again the empty product $\prod_{i=n}^{n-1}$ is taken to equal 1).

Case (b). We have

$$d(p_1, \partial S_1) < \frac{\prod_{i=1}^{n-1} 2K_0 L_i}{\prod_{i=n}^{4n-1} L_i^{\eta}}.$$

By volume preservation, it follows that for each $1 \leq k \leq n-1$,

$$\mathrm{Leb} \left\{ \begin{matrix} p \in S_1 : \tau(p) \leq n, \sigma(p) > 4n, \\ \text{and Case (a) holds for value } k \end{matrix} \right\} \leq 2\#(\mathcal{C}_k) \cdot \frac{\prod_{i=k}^{n-1} 2K_0 L_i}{\prod_{i=n}^{4n-1} L_i^{\eta}}.$$

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

Additionally, using the estimate (6), we have

$$\text{Leb}\left\{\begin{matrix} p \in S_1 : \tau(p) \le n, \sigma(p) > 4n, \\ \text{and Case (b) holds}\end{matrix}\right\} \le \text{Leb}\left\{p \in S_1 : d(p, \partial S_1) \le \frac{\prod_{i=1}^{n-1} 2K_0 L_i}{\prod_{i=n}^{4n-1} L_i^\eta}\right\}$$
$$\le C_S \,\text{Leb}(S) \frac{\prod_{i=1}^{n-1} 2K_0 L_i}{\prod_{i=n}^{4n-1} L_i^\eta}.$$

Thus,

$$\text{Leb}\{p \in S_1 : \tau(p) \le n, \sigma(p) > 4n\} \le (2nM_0 + C_S \,\text{Leb}(S)) \frac{\prod_{i=1}^{n-1} 2K_0 L_i}{\prod_{i=n}^{4n-1} L_i^\eta}. \quad (11)$$

To develop the right-hand side, observe that

$$\frac{\prod_{i=1}^{n-1} 2K_0 L_i}{\prod_{i=n}^{3n-1} L_i^\eta} \le \prod_{i=1}^{n} 2K_0 L_i^{1-2\eta} \le 1,$$

using that $\{L_i\}$ is a non-decreasing sequence, on taking L_0 sufficiently large so that $2K_0 L_0^{1-2\eta} \le 1$. For the terms $i = 3n, \ldots, 4n-1$, we estimate

$$\prod_{i=3n}^{4n-1} 2L_i^{-\eta} = \left(\prod_{i=3n}^{4n-1} L_i^{-\eta n}\right)^{1/n} \le \frac{1}{n}\sum_{i=3n}^{4n-1} L_i^{-\eta n} \le \frac{1}{n}\sum_{i=3n}^{4n-1} L_i^{-1+\eta}$$

by arithmetic mean-geometric mean inequality, on noting that $L_n^{-\eta n} < L_n^{-1+\eta}$ for all $n \ge 1$. Thus,

$$(11) \le (2M_0 + C_S \,\text{Leb}(S)) \sum_{i=3n}^{4n-1} L_i^{-1+\eta}.$$

For the final estimate, observe that

$$\text{Leb}\{p \in S_1 : \sigma(p) > 4n\} \le \text{Leb}\{p \in S_1 : \bar\tau(p) \le n, \sigma(p) > 4n\} + \text{Leb}\{p \in S_1 : \bar\tau(p) > n\}$$
$$\le (2M_0 + C_S) \sum_{i=3n}^{4n-1} L_i^{-1+\eta} + 6K_1 M_0 \sum_{i=n}^{\infty} L_i^{-1+\eta'}$$
$$\le (2M_0 + C_S \,\text{Leb}(S) + 6K_1 M_0) \sum_{i=n}^{\infty} L_i^{-1+\eta'}$$

on using (7) and that $\{L_i\}$ is non-decreasing. This completes the proof of Proposition 6.4.

6.3. *Disintegration of Lebesgue measure along horizontal foliation* $\hat\gamma_n$. To complete our description of the foliation $\hat\gamma_n$ of S_n, we describe here how $\hat\gamma_n$ disintegrates Lebesgue measure $\nu_n = F_*^{n-1}\nu_S = (1/\,\text{Leb}_{\mathbb{T}^2}(S_n))\,\text{Leb}_{\mathbb{T}^2}|_{S_n}$ on S_n.

Below, for $n \ge 1$ and an atom $\gamma \in \hat\gamma_n$, we write $(\nu_n)_\gamma$ for the disintegration measure of ν_n on γ; the disintegration measures $(\nu_n)_\gamma$ are the (almost surely) unique family of probability measures, supported on the $\gamma \in \hat\gamma_n$, which satisfy

$$\nu_n(K) = \int_{\gamma \in S_n/\hat\gamma_n} (\nu_n)_\gamma(\gamma \cap K)\, d\nu_n^T(\gamma)$$

for Borel $K \subset \mathbb{T}^2$; here ν_n^T is the pushforward of ν_n onto the quotient space of equivalence classes $S_n/\hat\gamma_n$.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

LEMMA 6.10. *Let $n \geq 1$ and fix $\gamma \in \hat{\gamma}_n$. Let ρ^n_γ denote the density of $(\nu_n)_\gamma$ with respect to* Leb_γ. *Then, for any $p, q \in \gamma$, we have that*

$$\frac{\rho^n_\gamma(p)}{\rho^n_\gamma(q)} = \frac{\det(dF^{n-1}_{n_1+1}|_{T\gamma_{n_1+1}}) \circ (F^{n-1}_{n_1+1})^{-1}(q)}{\det(dF^{n-1}_{n_1+1}|_{T\gamma_{n_1+1}}) \circ (F^{n-1}_{n_1+1})^{-1}(p)}.$$

Here $n_1 = \max(\{0\} \cup \{1 \leq k \leq n-1 : p_k \in B_k\})$, $p_n \in \gamma$ is an (arbitrary) representative, $p_k \in S_k$ is such that $F^{n-1}_k p_k = p_n$ for each $k \leq n$, and γ_{n_1+1} is the atom in $\hat{\gamma}_{n_1+1}$ for which $F^{n-1}_{n_1+1}(\gamma_{n_1+1}) \supset \gamma$.

Proof. To start, let us describe the disintegration measures $(\nu_1)_{\hat{\gamma}_1(p_1)}$ for $p_1 \in S_1$. It is clear that

$$(\nu_1)_{\hat{\gamma}_1(p)} = \frac{1}{\text{Leb}_{H_{p_1}}(\hat{\gamma}_1(p_1))} \text{Leb}_{H_{p_1}}|_{\hat{\gamma}_1(p_1)}, \tag{12}$$

where H_{p_1} is as in §6.1 and $\text{Len}(\gamma)$ denotes the arc length of a smooth connected curve $\gamma \subset \mathbb{T}^2$. Thus, Lemma 6.10 holds trivially in this case with $n_1 = 1$.

Inductively, let us express the disintegration ν_{n+1} in terms of that for ν_n. Observe that

$$\nu_{n+1} = (F_n)_* \nu_n|_{S_n \cap B_n} + (F_n)_* \nu_n|_{S_n \setminus B_n};$$

since $S_n \cap B_n, S_n \setminus B_n \in \mathcal{P}_n$, it suffices to consider these separately in working out the disintegration measures $(\nu_{n+1})_\gamma$, $\gamma \in \hat{\gamma}_{n+1}$.

On $F_n(S_n \cap B_n)$, Case 2 is applied in constructing $\hat{\gamma}_{n+1}|_{F_n(S_n \cap B_n)}$ and so disintegration measures are obtained using the analogue of (12) with $n+1$ replacing 1.

On $F_n(S_n \setminus B_n)$, we apply Case 1 in the construction of $\hat{\gamma}_{n+1}$, i.e., $\hat{\gamma}_{n+1} = \mathcal{P}_{n+1}|_{F_n(S_n \cap B_n)} \vee F_n(\hat{\gamma}_n|_{S_n \cap B_n})$. In particular, the disintegration $(\nu_{n+1}|_{F_n(S_n \setminus B_n)})_\gamma$, $\gamma \in \hat{\gamma}_{n+1}$ can be obtained by disintegrating, for each $\check{\gamma} \in \hat{\gamma}_n$, the measures $(F_n)_*((\nu_n)_{\check{\gamma}})$ against the (finite) partition $\mathcal{P}_{n+1}|_{F_n(\check{\gamma})}$. To wit, if $\gamma \in \hat{\gamma}_{n+1}|_{F_n(S_n \setminus B_n)}$ has $\gamma \subset F_n(\check{\gamma})$ for $\check{\gamma} \in \hat{\gamma}_{n+1}$, then

$$(\nu_{n+1})_\gamma = \frac{1}{(\nu_n)_{\check{\gamma}}(F_n^{-1}\gamma)} (F_n)_*((\nu_n)_{\check{\gamma}})|_\gamma.$$

In particular, we have shown that for any $p, q \in \gamma$, we have that

$$\frac{\rho^{n+1}_\gamma(p)}{\rho^{n+1}_\gamma(q)} = \frac{\det(dF_n|_{T\check{\gamma}}) \circ F_n^{-1}(q)}{\det(dF_n|_{T\check{\gamma}}) \circ F_n^{-1}(p)} \cdot \frac{\rho^n_{\check{\gamma}} \circ F_n^{-1}(p)}{\rho^n_{\check{\gamma}} \circ F_n^{-1}(q)}.$$

Lemma 6.10 follows by iterating the above relations from $n_1 + 1$ to $n - 1$. □

6.4. *Description of $(F^n)_* \nu_S$.* Here we synthesize the results of §§6.1–6.3 into our main result, a precise description of the bulk of $(F^n)_* \nu_S$ as foliated by a collection of fully crossing horizontal curves with controlled disintegration densities.

PROPOSITION 6.11. *Let $n \geq 2$. Then there are a measurable set $G \subset F^n S$ and a measurable partition $\mathcal{G}$ of G with the following properties.*

(a) *Each atom $\gamma \in \mathcal{G}$ is of the form* graph h_γ, *where $h_\gamma : (0, 1) \to \mathbb{T}^1$ is a C^2, fully crossing horizontal curve with $\|h'_\gamma\|_{C^0} = O(L_n^{-\eta})$.*

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

(b) *We have the estimate*

$$\nu_{n+1}(G) \geq 1 - O(L_n^{-(1/2)(1-\eta)}) - \nu_S\{\sigma > n\}$$
$$\geq 1 - \left(O(1) + C_S + \frac{C}{\mathrm{Leb}(S)}\right) \sum_{i=\lfloor n/4 \rfloor}^{\infty} L_i^{-(1/2)(1-\eta)} \tag{13}$$

on inserting the estimate in Proposition 6.4.

(c) *Let ν_G denote the restriction $\nu_{n+1}|_G$ and let $\{(\nu_G)_\gamma\}_{\gamma\in\mathcal{G}}$ denote the canonical disintegration of ν_G with respect to $\mathcal{G}$ by probability measures supported on each $\gamma \in \mathcal{G}$. Let $\rho_\gamma : \gamma \to [0, \infty)$ denote the density of $(\nu_G)_\gamma$ with respect to* Leb_γ. *Then, for any $p_1, p_2 \in \gamma$, we have*

$$\frac{\rho(p_1)}{\rho(p_2)} \leq e^{CL_n^{1-2\eta}}.$$

Proof. To start, define

$$\hat{\mathcal{G}}_n = \{\hat{\gamma} \in \hat{\gamma}_n : (\nu_n)_{\hat{\gamma}} F^{n-1}\{\sigma \leq n\} > 0\} \quad \text{and} \quad \hat{G}_n = \bigcup_{\hat{\gamma}\in\hat{\mathcal{G}}_n} \hat{\gamma}.$$

By Lemma A.1 in the appendix, we have

$$\nu_n(\hat{G}_n) = \nu_n^T\{\hat{\gamma} \in \hat{\mathcal{G}}_n\} \geq \nu_1\{\sigma \leq n\}.$$

Recalling the notation in Lemma 2.7, we define $\mathcal{G}$ by

$$\mathcal{G} = \bigcup_{\hat{\gamma}\in\hat{\mathcal{G}}_n} \bar{\Gamma}_n(\hat{\gamma}) \quad \text{and} \quad G = \bigcup_{\gamma\in\mathcal{G}} \gamma = \bigcup_{\hat{\gamma}\in\hat{\mathcal{G}}_n} F_n(\hat{\gamma}\backslash\mathcal{B}_n(\hat{\gamma})),$$

noting that $\mathcal{G}$ partitions G into horizontal curves γ which satisfy item (a) by construction.

To check item (b), for each $\hat{\gamma} \in \hat{\gamma}_n$ and subset $K \subset \hat{\gamma}$, we have that

$$(\nu_n)_{\hat{\gamma}}(K) \leq \frac{C}{\mathrm{Len}(\hat{\gamma})} \mathrm{Leb}_{\hat{\gamma}}(K)$$

on applying the distortion estimate in Lemma 2.5 to the density $\rho^n_{\hat{\gamma}}$ derived in Lemma 6.10. Since $\mathrm{Len}(\hat{\gamma})^{-1} = O(L_n^{1-\eta'})$ from the fact that $\hat{\gamma} \cap F^{n-1}\{\sigma \leq n\} \neq \emptyset$, we obtain the estimate

$$(\nu_n)_{\hat{\gamma}}(\mathcal{B}_n(\hat{\gamma})) = O\left(\frac{L_n^{-1+\eta}}{L_n^{-1+\eta'}}\right) = O(L_n^{-(1/2)(1-\eta)})$$

on inserting $K = \mathcal{B}_n(\hat{\gamma})$. Thus, (13) follows on noting that $-1 + \eta' = -1 + (1 + \eta)/2 = (\eta - 1)/2$.

For item (c), let $p_1, p_2 \in \gamma$ for some $\gamma \in \mathcal{G}$ and assume that $\gamma \in \bar{\Gamma}_n(\hat{\gamma})$ for $\hat{\gamma} \in \hat{\gamma}_n$. Then

$$\frac{\rho_\gamma(p_1)}{\rho_\gamma(p_2)} = \frac{\det(dF_n|_{T\hat{\gamma}}) \circ F_n^{-1}(p_2)}{\det(dF_n|_{T\hat{\gamma}}) \circ F_n^{-1}(p_1)} \cdot \frac{\rho^n_{\hat{\gamma}} \circ F_n^{-1}(p_1)}{\rho^n_{\hat{\gamma}} \circ F_n^{-1}(p_2)}$$

in the notation of §6.3. The first factor is bounded $\leq e^{CL_n^{1-2\eta}\|p_1-p_2\|}$ by Lemma 2.5. For the second factor, note that $\|F_n^{-1}(p_1) - F_n^{-1}(p_2)\| \leq L_n^{-\eta}\|p_1 - p_2\|$ by Lemma 2.4 and so Lemma 6.10 yields the estimate $\leq e^{CL_{n_1}^{1-2\eta}\cdot L_n^{-\eta}\|p_1-p_2\|} \leq e^{CL_n^{-\eta}}$. The estimate in item (c) follows. □

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

7. *Decay of correlation estimates*

Leaning on the mixing mechanism explored in the previous section, we complete here the proof of Theorem C.

In §7.1 we will show how to reduce Theorem C to the case when φ is the characteristic function of a small square (Proposition 7.1). In §7.2 we apply the results of §6 when S is a small square and give the proof of Proposition 7.1.

We assume throughout §7 that $\eta \in (1/2, 1)$ has been fixed and that $\{L_n\}$ has the property that $\sum_n L_n^{-1+\eta'} < \infty$, where $\eta' = (\eta + 1)/2$ is as in §6.2.1. These assumptions are consistent with the hypotheses of Theorem C.

7.1. *Reduction.*

We will show here that to prove Theorem C, it suffices to prove the following.

PROPOSITION 7.1. *Let R be a square in $\mathbb{T}^2$ of side length ℓ and let ν denote the normalized Lebesgue measure restricted to R. Let $\psi : \mathbb{T}^2 \to \mathbb{R}$ be α-Holder continuous. Then*

$$\left| \int \psi \circ F^n \, d\nu - \int \psi \right| \leq C \|\psi\|_\alpha \max \left\{ L_{\lfloor n/2 \rfloor}^{-\min\{2\eta - 1, \alpha(1-\eta)/(\alpha+2)\}}, \ell^{-2} \sum_{i=\lfloor n/8 \rfloor}^{\infty} L_i^{-(1/2)(1-\eta)} \right\}.$$

Proof of Theorem C assuming Proposition 7.1. Below, $n \geq 2$ is fixed, as are α-Holder continuous $\varphi, \psi : \mathbb{T}^2 \to \mathbb{R}$. Let us write $\mathcal{S}_n$ for the first element in the $\max\{\cdots\}$ in Proposition 7.1 and write $\mathcal{T}_n$ for the summation in the second term, so that the bound on the right-hand side reads as $\leq C \|\psi\|_\alpha \max\{\mathcal{S}_n, \ell^{-2} \mathcal{T}_n\}$.

With $K \in \mathbb{N}$ to be specified later, subdivide $\mathbb{T}^2$ into rectangles $R_{i,j}$, $1 \leq i, j \leq K$ of side length $\ell = 1/K$ each. We set

$$\varphi_{i,j} = \inf_{p \in R_{i,j}} \varphi(p).$$

Define $\hat{\varphi} := \sum_{i,j=1}^K \varphi_{i,j} \chi_{R_{i,j}}$, so that

$$\int (\varphi - \hat{\varphi}) \, d\,\mathrm{Leb} = O(\|\varphi\|_\alpha \cdot \ell^\alpha).$$

Let $\nu^{i,j}$ denote normalized Lebesgue measure on $R_{i,j}$. Then

$$\int \psi \circ F^n \cdot \varphi = \int \psi \circ F^n \cdot (\varphi - \hat{\varphi}) + \sum_{i,j=1}^K \ell^2 \varphi_{i,j} \int \psi \, d F^n_* \nu^{i,j}.$$

For the first term,

$$\int \psi \circ F^n \cdot (\varphi - \hat{\varphi}) = O(\|\psi\|_\alpha \|\varphi\|_\alpha \ell^\alpha).$$

Similarly, we estimate

$$\begin{aligned} \int \psi \cdot \int \varphi &= \int (\varphi - \varphi_k) \cdot \int \psi + \sum_{i,j=1}^K \ell^2 \varphi_{i,j} \int \psi \\ &= O(\|\psi\|_\alpha \|\varphi\|_\alpha \ell^\alpha) + \sum_{i,j=1}^K \ell^2 \varphi_{i,j} \int \psi; \end{aligned}$$

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

hence,

$$\left|\int \psi \circ F^n \cdot \psi - \int \psi \int \varphi\right| \le \sum_{i,j=1}^{K} \ell^2 \varphi_{i,j} \left|\int \psi \, dF_*^n \nu^{i,j} - \int \psi\right| + O(\|\psi\|_{C^0}[\varphi]_\alpha \ell^\alpha)$$
$$= \|\psi\|_\alpha \|\varphi\|_\alpha \cdot O(\mathcal{S}_n + \ell^{-2}\mathcal{T}_n + [\varphi]_\alpha \ell^\alpha).$$

Setting

$$K = \left\lfloor \left(\frac{[\varphi]_\alpha}{\mathcal{T}_n}\right)^{1/(2+\alpha)} \right\rfloor,$$

we obtain the estimate

$$\left|\int \psi \circ F^n \cdot \varphi - \int \varphi \int \psi\right| \le C\|\psi\|_\alpha \|\varphi\|_\alpha^{(4+\alpha)/(2+\alpha)} (\mathcal{T}_n^{\alpha/(2+\alpha)} + \mathcal{S}_n).$$

The only difference between this and our desired estimate is the exponent of $\|\varphi\|_\alpha$ on the right-hand side. To fix this, define $\check{\varphi} = \varphi/\|\varphi\|_\alpha$ and note that $\|\check{\varphi}\|_\alpha = 1$; for this function, we have

$$\left|\int \psi \circ F^n \cdot \check{\varphi} - \int \check{\varphi} \int \psi\right| \le C\|\psi\|_\alpha (\mathcal{T}_n^{\alpha/(2+\alpha)} + \mathcal{S}_n)$$

and so the desired estimate follows on multiplying both sides by $\|\varphi\|_\alpha$. To complete the proof, observe that

$$\max\{\mathcal{S}_n, \mathcal{T}_n^{\alpha/(\alpha+2)}\} \le \max\left\{L_{\lfloor n/2\rfloor}^{1-2\eta}, \left(\sum_{i=\lfloor n/8\rfloor} L_i^{-(1/2)(1-\eta)}\right)^{\alpha/(\alpha+2)}\right\}$$

since $\mathcal{T}_n^{\alpha/(\alpha+2)}$ always dominates $L_{\lfloor n/2\rfloor}^{-\alpha(1-\eta)/(\alpha+2)}$. □

7.2. *Proof of Proposition 7.1.* To complete the proof of Theorem C, it remains to prove Proposition 7.1. We combine the description in Proposition 6.11 of the foliation by long horizontal curves with the mixing estimate in Proposition 2.8 along those horizontal curves.

To wit: let $\psi : \mathbb{T}^2 \to \mathbb{R}$ be α-Holder continuous and let R be a square of side length ℓ as in the statement of Proposition 7.1. With ν denoting the Lebesgue measure restricted to R and (for notational convenience) applying the substitution $n \mapsto 2n$, we will estimate

$$\int \psi \circ F^{2n} \, d\nu = \int \psi \circ F_{n+1}^{2n} \, d(F_*^n \nu). \tag{14}$$

For each $k \ge 1$, define $\nu_k = F_*^{k-1}\nu_1$, where $\nu_1 = \nu$. Applying Proposition 6.11 to $S = R$, we obtain the collection $\mathcal{G}$ of horizontal curves foliating the set $G \subset F^n R$. In the notation of Proposition 6.4, we have $C_R = O(\ell^{-1})$ and so

$$\nu_{n+1}(G^c) = O\left(\ell^{-2} \sum_{i=\lfloor n/4\rfloor}^{\infty} L_i^{-(1/2)(1+\eta)}\right).$$

Returning to the estimate of (14),

$$(14) = O(\|\psi\|_\alpha \, \nu_{n+1}(G^c)) + \int \psi \circ F^{2n}_{n+1} \, d\nu_G$$
$$= O(\|\psi\|_\alpha \, \nu_{n+1}(G^c)) + \int_{G/\mathcal{G}} \bigg(\int_\gamma \psi \circ F^{2n}_{n+1} \, d(\nu_G)_\gamma \bigg) d\nu^T_G,$$

where the transversal measure ν^T_G is the pushforward of ν_G onto $G/\mathcal{G}$.

Fixing $\gamma \in \mathcal{G}$, we have by the density estimate in Proposition 6.11 that

$$\int_\gamma \psi \circ F^{2n}_{n+1} \, d(\nu_G)_\gamma = (1 + O(L_n^{1-2\eta})) \int_\gamma \psi \circ F^{2n}_{n+1} \, d\,\mathrm{Leb}_\gamma$$

and so applying Proposition 2.8 with $m \mapsto n+1$, $n \mapsto 2n$, we have

$$\int_\gamma \psi \circ F^{2n}_{n+1} \, d(\nu_G)_\gamma$$
$$= (1 + O(L_n^{1-2\eta}))\,\mathrm{Len}(\gamma) \cdot \int \psi + (1 + O(L_n^{1-2\eta}))\|\psi\|_\alpha$$
$$\cdot\, O\bigg(L_{2n}^{-\alpha(1-\eta)/(2+\alpha)} + L_{n+1}^{1-2\eta} + \sum_{k=n+1}^{2n-1} L_k^{-1+\eta} \bigg)$$
$$= \int \psi + \|\psi\|_\alpha \cdot O\bigg(L_{2n}^{-\alpha(1-\eta)/(2+\alpha)} + L_n^{1-2\eta} + \sum_{k=n+1}^{2n-1} L_k^{-1+\eta} \bigg).$$

Collecting these estimates, we conclude that

$$\bigg| \int \psi \circ F^{2n} \, d\nu - \int \psi \bigg| \le C\|\psi\|_\alpha \bigg(L_n^{-\min\{2\eta-1,\alpha(1-\eta)/(2+\alpha)\}} + \ell^{-2} \sum_{i=\lfloor n/4 \rfloor}^{\infty} L_i^{-(1/2)(1-\eta)} \bigg).$$

This completes the proof.

Acknowledgements. The author thanks Dmitry Dolgopyat for suggesting this problem and for many helpful discussions. This research was supported by NSF grant DMS-1604805.

A. *Appendix.*

LEMMA A.1. (Partition saturation) *Let X be a compact metric space,* $\mathrm{Bor}(X)$ *the Borel σ-algebra on X, and μ a probability on $(X, \mathrm{Bor}(X))$. Let ξ be a measurable partition of X and denote by $(\mu_C)_{C\in\xi}$ the canonical disintegration of μ with respect to ξ. Let μ^T denote the transverse measure on X/η.*

Let $Y \in \mathrm{Bor}(X)$. Then $\mu^T\{C \in X/\eta : \mu_C(Y) > 0\} \ge \mu(Y)$.

Proof. We estimate

$$\mu(Y) = \int_{X/\eta} \mu_C(Y) \, d\mu^T(C) = \int_{C \in X/\eta : \mu_C(Y) > 0} \mu_C(Y) \, d\mu^T(C)$$
$$\le \mu^T\{C \in X/\eta : \mu_C(Y) > 0\}.$$

□

REFERENCES

[**1**] R. Aimino, H. Hu, M. Nicol, A. Török and S. Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst. A* **35**(3) (2015), 793–806.

[**2**] L. Arnold. *Random Dynamical Systems*. Springer Science & Business Media, Berlin, 2013.

[**3**] P. Arnoux and A. M. Fisher. Anosov families, renormalization and non-stationary subshifts. *Ergod. Th. & Dynam. Sys.* **25**(3) (2005), 661–709.

[**4**] V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings, I. *Russ. Acad. Sci. Izv. Math.* **44**(2) (1995), 247–279.

[**5**] V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings, II. *Russ. Acad. Sci. Izv. Math.* **44**(3) (1995), 617–627.

[**6**] L. Barreira and Y. Pesin. *Nonuniform Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents (Encyclopedia of Mathematics and its Applications, 115)*. Cambridge University Press, Cambridge, 2007.

[**7**] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Ann. of Math. (2)* **133**(1) (1991), 73–169.

[**8**] P. Berger and D. Turaev. On Herman's positive entropy conjecture. *Adv. Math.* **349** (2019), 1234–1288.

[**9**] A. Blumenthal, J. Xue and L.-S. Young. Lyapunov exponents and correlation decay for random perturbations of some prototypical 2D maps. *Comm. Math. Phys.* **359**(1) (2018), 347–373.

[**10**] A. Blumenthal, J. Xue and L.-S. Young. Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. *Ann. of Math. (2)* **185**(1) (2017), 285–310.

[**11**] N. Chernov and D. Dolgopyat. *Brownian Brownian Motion-I*. American Mathematical Society, Providence, RI, 2009.

[**12**] N. Chernov and R. Markarian. *Chaotic Billiards (Mathematical Surveys and Monographs, 127)*. American Mathematical Society, Providence, RI, 2006.

[**13**] B. V. Chirikov. A universal instability of many-dimensional oscillator systems. *Phys. Rep.* **52**(5) (1979), 263–379.

[**14**] J.-P. Conze and A. Raugi. Limit theorems for sequential expanding dynamical systems. *Ergodic Theory and Related Fields: 2004–2006 Chapel Hill Workshops on Probability and Ergodic Theory, University of North Carolina, Chapel Hill, North Carolina (Contemporary Mathematics, 430)*. American Mathematical Society, Providence, RI, 2007, pp. 89–121.

[**15**] J. De Simoi. Stability and instability results in a model of Fermi acceleration. *Discrete Contin. Dyn. Syst. A* **25**(3) (2009), 719–750.

[**16**] J. De Simoi. Fermi acceleration in anti-integrable limits of the standard map. *Comm. Math. Phys.* **321**(3) (2013), 703–745.

[**17**] D. I. Dolgopyat. Averaging and invariant measures. *Mosc. Math. J.* **5**(3) (2005), 537–576.

[**18**] D. Dolgopyat. Bouncing balls in non-linear potentials. *Discrete Contin. Dyn. Syst.* **22**(1–2) (2008), 165–182.

[**19**] D. Dolgopyat, V. Kaloshin and L. Koralov. Sample path properties of the stochastic flows. *Ann. Probab.* **32** (2004), 1–27.

[**20**] P. Duarte. Plenty of elliptic islands for the standard family of area preserving maps. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **11** (1994), 359–409.

[**21**] A. Gorodetski. On stochastic sea of the standard map. *Comm. Math. Phys.* **309**(1) (2012), 155–192.

[**22**] Y. Kifer. *Ergodic Theory of Random Transformations (Progress in Probability and Statistics, 10)*. Springer Science & Business Media, Heidelberg, 2012.

[**23**] C. Liverani. Decay of correlations. *Ann. of Math. (2)* **142**(2) (1995), 239–301.

[**24**] C. Liverani. Central limit theorem for deterministic systems. *Int. Conf. Dynamical Systems (Montevideo, 1995) (Pitman Research Notes in Mathematics Series, 362)*. Longman, London, 1996, pp. 56–75.

[**25**] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergod. Th. & Dynam. Sys.* **19**(3) (1999), 671–685.

[**26**] M. Nicol, A. Török and S. Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. *Ergod. Th. & Dynam. Sys.* **38**(3) (2018), 1127–1153.

[**27**] Ya. B. Pesin and Ya. G. Sinai. Gibbs measures for partially hyperbolic attractors. *Ergod. Th. & Dynam. Sys.* **2**(3–4) (1982), 417–438.

[**28**] V. A. Rokhlin. On the fundamental ideas of measure theory. *Mat. Sb.* **67**(1) (1949), 107–150.

[**29**] D. Ruelle. A measure associated with Axiom-A attractors. *Amer. J. Math.* **98** (1976), 619–654.

[**30**] M. Stenlund. Non-stationary compositions of Anosov diffeomorphisms. *Nonlinearity* **24**(10) (2011), 2991–3018.

https://doi.org/10.1017/etds.2019.115 Published online by Cambridge University Press

[**31**] M. Stenlund. A vector-valued almost sure invariance principle for Sinai billiards with random scatterers. *Comm. Math. Phys.* **325**(3) (2014), 879–916.

[**32**] M. Stenlund, L.-S. Young and H. Zhang. Dispersing billiards with moving scatterers. *Comm. Math. Phys.* **322**(3) (2013), 909–955.

[**33**] L.-S. Young. Ergodic theory of differentiable dynamical systems. *Real and Complex Dynamical Systems*. Springer, Heidelberg, 1995, pp. 293–336.

[**34**] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)* **147**(3) (1998), 585–650.