# EXPONENTIAL MIXING IMPLIES BERNOULLI

D. DOLGOPYAT, A. KANIGOWSKI, F. RODRIGUEZ-HERTZ

ABSTRACT. Let $f$ be a $C^{1+\alpha}$ diffeomorphism of a compact manifold $M$ preserving a smooth measure $\mu$. We show that if $f : (M, \mu) \to (M, \mu)$ is exponentially mixing then it is Bernoulli.

## 1. INTRODUCTION

1.1. **Main result.** Let $f : (M, \mu) \to (M, \mu)$ be a $C^{1+\alpha}$ diffeomorphism of a compact manifold $M$ that preserves a smooth measure $\mu$. We say that $f$ is *exponentially mixing* (for smooth functions) if there exists [1] $\mathbf{r} \in \mathbb{N}$, $C > 0$, and $\eta > 0$ such that for any $\phi, \psi \in C^{\mathbf{r}}(M)$

$$(1.1) \qquad \left| \int_M \phi(x)\psi(f^n x)d\mu - \int_M \phi \, d\mu \int_M \psi \, d\mu \right| \leqslant Ce^{-\eta n}\|\phi\|_{\mathbf{r}}\|\psi\|_{\mathbf{r}},$$

where $\| \cdot \|_{\mathbf{r}}$ is the norm on $C^{\mathbf{r}}(M)$.

Recall that $(f, M, \mu)$ is a *Bernoulli system* (or Bernoulli) if for some $m$ it is measure theoretically isomorphic to the shift on $\{1, \ldots, \ell\}^{\mathbb{Z}}$ with the measure $\mathbf{p}^{\mathbb{Z}}$ where $\mathbf{p} = (p_1, \ldots, p_\ell)$ is a probability vector. Our main result is the following:

**Theorem 1.1.** *If $f$ is exponentially mixing then it is Bernoulli.*

1.2. **Broader context.** One of the central discoveries made in the last century in the theory of dynamical systems is that smooth systems can exhibit chaotic behavior. The strongest ergodic property that describes chaoticity is the Bernoulli property, i.e. being (measure-theoretically) isomorphic to a Bernoulli shift. Some weaker ergodic properties describing chaoticity are (see e.g. a survey article by Ya. Sinai, [51]): the $K$-property, positive entropy, mixing of all orders, mixing, weak mixing and ergodicity. It is easy to see that Bernoulli implies $K$ and that mixing of all orders implies mixing which implies weak mixing which implies ergodicity. It follows by [47] that $K$-property implies positive entropy. Moreover $K$-property also implies mixing of all orders, see e.g [16] (this inclusion is probably least trivial from all the inclusions mentioned above). It is not known if mixing implies mixing of all orders; this is known as the Rokhlin problem, [45]. Except for the Rokhlin problem it is known that all the above inclusions are strict (also in the smooth setting see e.g. the discussion in [20]). All the above mentioned properties do not require a smooth structure and can be defined for an arbitrary measure preserving system.

---

[1] A standard interpolation argument (see e.g. Lemma B.1 in Appendix B) shows that if (1.1) holds for some $\mathbf{r}$ then it holds for all $\mathbf{r} > 0$ (but taking $\mathbf{r}$ small would require making $\eta$ small).

Classical statistical properties that require a smooth structure (see e.g. [51]) are: central limit theorem, large deviations and exponential mixing. These properties provide quantitative information on the system. All the three properties imply ergodicity, but central limit theorem and large deviations do not imply weak mixing and hence also do not imply stronger ergodic properties, see e.g. [20]. In this paper we focus on consequences of exponential mixing. Notice that trivially exponential mixing implies mixing. However it was not known if it implies any stronger ergodic properties. Our main result (see Theorem 1.1) shows that exponential mixing implies the strongest ergodic property: Bernoullicity. In particular, it has the following non-trivial corollary:

**Corollary 1.2.** *Let $f \in C^{1+\alpha}(M)$. If $f$ is exponentially mixing with respect to a smooth measure $\mu$ then it is mixing of all orders and also has positive entropy.*

We note that [27] shows that mixing implies mixing of all orders for systems whose spectral measure is singular. Corollary 1.2 treats the opposite case where the spectral measure has analytic density for smooth observables.

We in fact show in Section 3 that if $f \in C^{1+\alpha}(M)$ is exponentially mixing for a $f$-invariant measure $\mu$ which is not supported on a fixed point of $f$, then $f$ has a non-zero Lyapunov exponent with respect to $\mu$, i.e. we have:

**Proposition 1.3.** *If $f : M \to M$ is a $C^{1+\alpha}$ diffeomorphism which is exponentially mixing with respect to a non atomic measure $\mu$ then $f$ has at least one positive Lyapunov exponent (for the measure $\mu$).*

The following questions are natural:

**Question 1.4.** *Let $f : X \to X$ be a $C^{1+\alpha}$ map preserving a non atomic measure $\mu$. Assume that $(f, \mu)$ is exponentially mixing for Hölder observables. Does $f$ have positive topological entropy? Does $(f, \mu)$ have positive metric entropy? Is $(f, \mu)$ a $K$ system? Is it Bernoulli?*

Our main result provides positive answers to all those questions if the measure $\mu$ is smooth, but it is interesting to weaken assumptions on the invariant measure.

We also remark that while our results show that exponential mixing implies mixing of all orders we do not get any quantitative bounds on the rate of multiple mixing. In particular, the following question is natural. We say that $(f, \mu)$ is *exponentially mixing of order $k$* if there exist constants $\mathbf{r}_k, C_k, \eta_k$ such that if $\phi_1, \phi_2, \ldots, \phi_k \in C^{\mathbf{r}}(M)$ then

$$(1.2) \qquad \left| \int_M \left( \prod_{j=1}^k \phi_j(f^{n_j}x) \right) d\mu - \prod_{j=1}^k \int_M \phi_j \, d\mu \right| \leqslant C_k e^{-\eta_k L} \prod_{j=1}^k \|\phi_j\|_{\mathbf{r}_k},$$

where $L = \min_{1 \leqslant j \leqslant k-1} (n_{j+1} - n_j)$.

**Question 1.5.** *Does exponential mixing imply exponential mixing of all orders?*

We note that exponential mixing of all orders implies several statistical properties such as the Central Limit Theorem [7] and Poisson Limit Theorem for close returns [21].

The reason why our method does not provide quantitative bounds on multiple mixing is because we rely on the Pesin theory, which in particular uses a Multiplicative Ergodic Theorem which is a non constructive result. It seems of interest to obtain quantitative bounds assuming some estimates on the measure of points where the convergence in the Multiplicative Ergodic Theorem is slow. Such results were previously obtained in [1, 2] where instead of exponential mixing the authors assume non-uniform hyperbolicity and dominated splitting.

The Bernoulli property was shown to hold for many classes of natural dynamical systems: ergodic toral automorphisms [32], Axiom $A$ diffeomorphisms [8], quadratic maps[2] with absolutely continuous invariant measure [34], geodesic flows on surfaces of constant negative curvature [36], geodesic flows on higher dimensional manifolds (without focal points) [40], Anosov flows [44], non-uniformly hyperbolic maps and flows (with singularities) [15]. Recently in [28] it was shown that partially hyperbolic homogeneous systems are Bernoulli. The above list is not complete but it contains the main examples of smooth Bernoulli systems.

We note that in the last 25 years there has been a significant progress in proving K property for partially hyperbolic systems. This study was initiated in [50, 10, 25]. Currently the strongest result is due to [13] and says that a partially hyperbolic center bunched volume preserving diffeomorphism with essential accessibility property is K. Recall that $f$ is *partially hyperbolic* if there is a $Df$ invariant splitting $TM = E^u \oplus E^c \oplus E^s$ and positive functions $\nu(x), \hat{\nu}(x), \gamma(x),, \hat{\gamma}(x)$ such that

$$\nu, \hat{\nu} < 1, \quad \nu < \gamma < \hat{\gamma}^{-1} < \hat{\nu}^{-1},$$

and

$$\|Df(v)\| < \nu\|v\| \text{ if } v \in E^s, \quad \|Df(v)\| > \hat{\nu}^{-1}\|v\| \text{ if } v \in E^u,$$
$$\gamma\|v\| < \|Df(v)\| < \hat{\gamma}^{-1}\|v\| \text{ if } v \in E^c.$$

A partially hyperbolic system is called *center bunched* if the above functions could be chosen so that $\nu < \gamma\hat{\gamma}$ and $\hat{\nu} < \gamma\hat{\gamma}$. A key inspiration for our approach comes from the remark that any system with non zero Lyapunov exponents could be regarded as a non-uniformly partially hyperbolic system (enjoying the center bunching). This allows one to extend several tools from the theory of partially hyperbolic systems to the non-uniform setting and plays an important role in our proof.

Recall that for a partially hyperbolic system one can define an accessibility class of a point $x$ as the set of points which can be joined to $x$ by a piecewise smooth curve such that each piece belongs to either one stable leaf or one unstable leaf. Essential accessibility means that every measurable set which consists of whole accessibility classes has measure zero or one. This is weaker than accessibility which means that there is only one accessibility class. We note that the essential accessibility is insufficient for the Bernoulli property, see [30, 48, 29]. Thus a natural next step is to understand which additional features of the system are responsible for Bernoullicity. The recent works [28, 22, 20] seem to indicate that an important role is played by the competition between the rate of mixing and the complexity of the system restricted to the subspace

---

[2]For non invertible systems the Bernoulli property means that the natural extension of $f$ is isomorphic to a Bernoulli shift.

with zero exponents. In particular, in the present paper we show that the exponential mixing implies the Bernoulli property as the growth in the zero exponents directions is always sub-exponential. However, there are still many open questions related to the Bernoulli property for smooth systems. Below we mention a few which seem to play a central role in the theory.

**Question 1.6.** *Is exponential mixing assumption in our main theorem optimal? In particular, does there exists a diffeomorphism which enjoys a stretched exponential mixing (i.e. at rate $e^{-n^\alpha}$ for $\alpha \in (0,1)$) but is not Bernoulli?*

We note that [20] constructs non Bernoulli systems with arbitrary fast polynomial mixing rate. However, in order to get the mixing rate of $n^{-\alpha}$ [20] considers manifolds of dimension growing quadratically with $\alpha$. In fact, lowering the dimension of the phase space makes it more difficult to construct non Bernoulli systems. In [29] the authors construct $K$ non-Bernoulli examples in dimension 4. On the other hand it follows from the Pesin theory that $K$ implies Bernoulli in dimension 2. Hence the following classical question is of a central importance:

**Question 1.7.** *Does there exist a $K$ non Bernoulli diffeomorphism preserving a smooth measure in dimension three?*

The next question is important for the theory of partially hyperbolic systems.

**Question 1.8.** *Is every volume preserving partially hyperbolic system with accessibility property Bernoulli?*

1.3. **Outline of the proof.** Since our approach requires rather technical results from the Pesin theory of $C^{1+\alpha}$ diffeomorphisms we will outline the main steps in the proof for convenience of the reader.

A standard approach for proving the Bernoulli property (developed in [36]) is to verify the *very weak Bernoulli (vwB) property* which means that for each $S \geqslant 1$ the itinerary of the orbit during the time interval $[0, S]$ is asymptotically independent of the remote past. More precisely, if $\mathcal{P}$ is a finite partition then for large enough $N_2 > N_1$, the distribution of the itineraries on time $[0, S]$ are almost the same for most atoms of $\mathcal{P}_{N_1,N_2} := \bigvee_{i=N_1}^{N_2} f^i \mathcal{P}$ (here the closeness of the distributions on itineraries is induced by the topology in which two itineraries are close if their Hamming distance is small). In this paper we also verify vwB property, but to do so we need to develop geometric structure of our system which follows from exponential mixing. We divide our argument into several steps.

The first main step in the proof (conducted in Section 3) is to show that exponential mixing implies that some exponents are non-zero. For the proof of this we only require that $f \in C^{1+\alpha}(M)$ is exponentially mixing for a measure $\mu$ which is not supported on a single point. We first show that if $f$ is exponentially mixing, then there is $\eta_1 > 0$ and a set $B$ of positive measure such that for all sufficiently large $n$ the balls $O_n$ centered in $B$ with radius $e^{-\eta_1 n}$ satisfy that the diameter of $f^n(O_n)$ becomes macroscopic. Next we show that if exponents were zero then, for each $\varepsilon > 0$ the images of balls of size

$\sim e^{-\varepsilon n}$ centered at Oseledets typical points remain exponentially small after iterating $n$ times. Taking $\varepsilon < \eta_1$ leads to a contradiction with the fact that such an image must become of order 1 (macroscopic) in diameter.

The existence of non zero exponents makes the Pesin theory applicable to our problem. In particular, almost every point has an unstable manifold of positive size and the unstable lamination is absolutely continuous. We note that all the necessary facts from Pesin theory needed in our paper can be obtained by standard techniques. In particular our presentation relies heavily on Barreira-Pesin book [6]. On the other hand the statements in our paper are less restrictive than in most other references as we only assume that the system has some non-zero exponents. In particular we extend the theory of fake center stable foliations developed in the ergodic theory of partially hyperbolic systems to the non-uniform setting. One technical novelty in our argument is that we define center foliations that work for finitely many iterates, and they have good absolute continuity properties (see the descriptions of Sections 5 and 6 below). We would like to emphasize that the results of Sections 4–6 (as well as Section 8 and Appendix A) are valid for any diffeomorphism preserving a smooth measure with some non-zero exponents and as such are of independent interest.

Next, a standard backwards contraction argument going back to [3, 36, 50] shows that the remote past partition $\mathcal{P}_{N_1,N_2}$ is almost $u$-saturated, meaning that if $\xi$ is sufficiently small then for most points the unstable manifold of size $\xi$ around $x$ belongs to the same atom as $x$. We present this argument in Section 4.

Given almost u-saturation, the natural idea for verifying the vwB property is to show that for any large $n$ and any two "typical" unstable pieces $\mathcal{W}_1, \mathcal{W}_2$ of size $\xi$ there exists an almost measure-preserving map $\theta : (\mathcal{W}_1, m^u_{\mathcal{W}_1}) \to (\mathcal{W}_2, m^u_{\mathcal{W}_2})$ such that the points $x$ and $\theta x$ remain close for $1 - \varepsilon$ proportion of the first $n$ iterates. More precisely, the existence of such maps allows to control $S$–itineraries for sufficiently large $S$, while for small $S$ we can use the $K$-property, see Corollary 2.22 for details. For many of the Bernoulli examples mentioned in §1.2 this map can be constructed taking $n' \ll n$, subdividing $f^{n'}\mathcal{W}_s = \bigcup_j \mathcal{W}_{s,j}$ so that $\mathcal{W}_{1,j}$ is close to $\mathcal{W}_{2,j}$ and defining $\theta : f^{-n'}\mathcal{W}_{1,j} \to f^{-n'}\mathcal{W}_{2,j}$ using the center-stable holonomy. In this approach almost measure preservation comes from the absolute continuity of the center stable foliation, the closeness of $f^j x$ and $f^j(\theta x)$ comes from the fact that the center stable direction is non-expanding and the possibility of subdividing $f^{n'}\mathcal{W}_1$ and $f^{n'}\mathcal{W}_2$ into the pieces which are close to each other comes from the minimality of the unstable lamination. In our case none of the above properties is available. That is, we do not know if the center stable distribution is uniquely integrable and if so, it is not clear if the resulting lamination is absolutely continuous. Moreover the vectors in the subspace corresponding to the zero exponent could grow albeit at a subexponential rate. Finally we do not know if the unstable lamination is minimal. To overcome these difficulties we establish weakened analogues of the above properties which are nonetheless sufficient for our purposes.

First of all, in Section 5 we introduce the crucial notion of *fake center-stable foliation* at time $n$. This (locally defined) foliation mimics the behavior of the center-stable foliation for $n$ iterates. Fake center-stable foliation was previously used in the study of ergodic properties of partially hyperbolic systems in [13] where these foliations were constructed near an arbitrary orbit. In our setting this foliation is defined only near Lyapunov regular orbits. The main result of Section 5 is that we can define this lamination outside of a set of an arbitrary small measure so that there is a unique leaf passing through each point. This statement is non trivial even for partially hyperbolic systems in case the center stable distribution is not uniquely integrable.

The properties of the fake foliation are studied in Section 6. For each fixed $n$ the center stable foliation is obtained by pulling back a smooth foliation, hence it is absolutely continuous. If the center stable distribution is uniquely integrable then the fake foliations approach the real center stable foliation as $n$ tend to infinity. Therefore if the center stable foliation is not absolutely continuous, then we could expect the fake center stable jacobians to deteriorate on the unit scale. However, we show in Proposition 6.4 that if we take two submanifolds which are exponentially close then the jacobian of the fake center stable holonomy is close to 1. Accordingly the fake center stable holonomy between nearby typical unstable leaves is almost measure preserving (Proposition 6.8). Another consequence of the local absolute continuity is local product structure of the measure on the small scale established in Corollary 6.6.

The next key step in the analysis is the exponential almost equidistribution of unstable leaves established in Section 7. The main result of that section is the Main Proposition (Proposition 7.3) which says that given a typical unstable leaf $\mathcal{W}$ and a partition of the phase space into cubes of size $r \geqslant e^{-\varepsilon n}$ we can discard a small proportion of cubes so that the proportion of $f^n \mathcal{W}$ inside the remaining cubes is approximately equal to the measure of the cube. We also point out that the Main Proposition also implies that for **any** cube of size of order 1, $f^n \mathcal{W}$ is equdistributed in the cube.

A standard approach to proving equidistribution of the unstable leaves is the following. Take exponentially narrow tube $\mathcal{T}$ around $\mathcal{W}$. By exponential mixing the image of $\mathcal{T}$ is equidistributed in the phase space. Next every point $z$ in $\mathcal{T}$ belongs to the center stable leaf of a point $z'$ in $\mathcal{W}$, and since the Lyapunov exponents in the center stable direction are non-positive, we expect $f^n z$ and $f^n z'$ to be close. Unfortunately, this closeness only holds if $z'$ is sufficiently regular point and while the contribution of non regular points is small it does not have to be exponentially small. These necessitates discarding the cubes in our partition which attract an anomalously large proportion of non regular points. The Main Proposition is the crucial result in establishing both the $K$ and the Bernoulli properties. Namely, the $K$-property follows from the equidistribution of the image of unstable leaves on the unit scale, as we explain in Section 8. To verify the very weak Bernoulli we follow the strategy described above by constructing the coupling between the nearby pieces of $f^{n'}\mathcal{W}_1$ and $f^{n'}\mathcal{W}_2$ with $n' = \varepsilon n$. Here the possibility of the subdivision

$$ f^{n'} W_s = \Big( \bigcup_j \mathcal{W}_{s,j} \Big) \cup \{\text{small unmatched part}\} $$

so that $\mathcal{W}_{1,j}$ is close to $\mathcal{W}_{2,j}$ comes from the exponential almost equidistribution, the coupling map is almost measure preserving due to the local absolute continuity of the fake center stable foliation and the closeness of $x$ and $\theta x$ comes from the fact that $f^{n'}x$ and $f^{n'}\theta(x)$ are exponentially close at time $n'$ and the divergence in the center stable direction is subexponential. This argument is presented in Section 9.

1.4. **The choice of parameters.** We finish our outline by specifying the dependence of parameters that appear in our proofs. Our proofs rely heavily on the Pesin theory and the very weak Bernoulli property. They involve several scales of smallness (or largeness) of parameters. To make it easier for the reader we summarize the dependences that appear in the paper.

(1) We start with $f \in C^{1+\alpha}(M)$ which is exponentially mixing on $C^{\mathbf{r}}(M)$ with exponent $\eta > 0$. In the text we introduce the number $\hat{\eta}$ (depending only on $\eta, \mathbf{r}$ and $\dim M$). Namely in Lemma 3.1 we obtain as a consequence of exponential mixing, that if $\hat{\eta}$ is small enough then the images of balls (centered at typical points) of size $e^{-\hat{\eta} m}$ become macroscopic after $n$ iterates. Also in Lemma B.2 we show that exponential mixing for smooth functions implies mixing on parallelograms of size $e^{-\eta_2 n}$ provided that $\eta_2$ is small enough. [3] We use the notation $\alpha_i$ to denote Hölder exponents of certain functions that we define. All the $\alpha_i$ depend only on $\|f\|_{C^1}$, the Lyapunov exponents of $f$, $\alpha$, and $\dim M$. In Proposition 6.4 we will also use the number $\beta > 0$ which can be taken to be $1/2 \min(\alpha, \min_i \alpha_i)$.

(2) Our proof proceeds by verifying the Ornstein-Weiss criterion (see Corollary 2.22). Thus given $\varepsilon > 0$ we need to construct a map $\theta$ verifying (2.22). This $\varepsilon$ defines the next level of smallness. We note that if the conditions of Corollary 2.22 are satisfied for some $\varepsilon > 0$ then they also hold for all $\varepsilon' > \varepsilon$. Therefore in the proof we assume that $\varepsilon$ is sufficiently small and we will use estimates like $\varepsilon^2 < \varepsilon/100$ without additional explanations.

(3) Having fixed $\varepsilon > 0$ we take $\delta = \varepsilon^{100}$. We then apply Lemma 2.4 for this choice of $\delta$ to get the corresponding function $\mathfrak{r}(\cdot)$ describing the size of Pesin chart for $(\lambda, \delta)$-regular points. Next we pick the parameter $\tau > 0$ for the Pesin sets in (2.4) and (2.14) respectively. For points in these sets the function $\mathfrak{r}(\cdot)$ is larger than $\tau$. We again want $\tau = \tau(\varepsilon)$ to be small enough in terms of $\varepsilon$ so that $\mu(P_\tau) \geqslant 1 - \varepsilon^{10^{10}}$ (see (2.15)).

(4) Having fixed all the above parameters we now choose a sufficiently large $n_0$ (largeness depending on all the previously fixed parameters) and we conduct the proof for $n \geqslant n_0$.

Throughout the paper we use the following abbreviations:
$$\xi_n = e^{\varepsilon^2 n - \eta_2 \varepsilon n}, \quad \tilde{r}_n = e^{-\eta_2 \varepsilon n - \varepsilon^2 n}.$$

The above parameters are the sizes of the parallelograms $B(\xi_n, \tilde{r}_n)$ that we will consider: the parameter $\xi_n$ is the size in $u$ direction and $\tilde{r}_n$ is the size in the (fake) $cs$ direction. It is important that both are exponentially small, but the $u$ direction is

---

[3]According to (B.5) and (B.4) we can take $\hat{\eta} = \min\left(\dfrac{1}{10D}, \dfrac{1}{8\mathbf{r}}\right)$, and $\eta_2 = \dfrac{\eta}{1 + 2\mathbf{r}}\left(\dfrac{D}{10} - \dfrac{1}{15}\right)$, where $D = \dim(M)$, but the precise values of these constants are not important for our argument.

longer than the $cs$ (this simplifies some arguments using Hölder continuity of the $u$ foliation).

1.5. **Notation.** The following notation is used throughout the paper:

$\mathbf{b} = 10^{10}$.

$B(\xi, r)$–the parallelograms in the phase space (see (5.1)).

$\bar{B}_s$ – disjoint parallelograms whose union covers most of the space (see family $\mathbb{B}_1$ after Lemma 5.1)

$\mathbb{B}_2$ – the family of parallelograms defined after equation (5.7).

$B_i(r)$ – parallelograms defined after equation (5.11).

$\mathcal{C}_{\mathfrak{b}}^u, \mathcal{C}_{\mathfrak{b}}^{cs}$ –unstable and center-stable cones of aperture $\mathfrak{b}$ (see (2.16), (2.17)).

$\mathcal{C}_x^u(y), \mathcal{C}_x^{cs}(y)$–the cones of aperture $1/2$ of $x$ shifted to $y$ (see (2.18)).

$D = \dim(M)$, $d^u = \dim(E^u)$, $d^{cs} = d(E^{cs})$.

$E^u(x)$, $E^{cs}(x)$–the Oseledets subspaces at $x$ corresponding to positive and non-positive exponents respectively.

$\tilde{E}_x^{cs,n,\delta}(\bar{y}) = T_{\bar{y}} \tilde{W}_x^{cs,n,\delta}(\bar{y})$.

$\tilde{f}$–representation of $f$ in Lyapunov charts (see Lemma 2.4).

$\mathcal{F}_{i,j}$–the leaves of fake center stable foliations (obtained by reindexing of $\mathcal{F}_{i,j,s}$ from (5.13), see end of Section 5).

$h_{x,\delta}$–Lyapunov coordinate map (see Lemma 2.4).

$L_{x,\delta}$–the linear map rectifying the angles between $E^u(x)$ and $E^{cs}(x)$ (see Lemma 2.4).

LyapReg–the set of Lyapunov regular points (see Definition 2.1).

$\mathcal{L}_{n,\tau} = P_\tau \cap f^{-n} P_\tau \cap f^{-\varepsilon n} P_\tau$ (see (5.3)).

$m_x^u$ – Lebesgue measure on $W_x^u$.

$m_{\mathcal{W}}^u$ – the conditional of $\mu$ on the unstable leaf $\mathcal{W}$ (see (2.19)).

$P_\tau$ – the Pesin set of points with $\mathfrak{r} \geqslant \tau$ (see (2.15)).

$Q_x$ – Lyapunov neighborhood of the point $x$ (see (2.6)).

$Q_x^n$ – Bowen ball with respect to the Lyapunov neighborhood $x$ (see (2.6)).

$Q_x(r_u, r_{cs})$ – images of the parallelograms in the tangent space (see (2.7)).

$r_u(x)$ – size of unstable manifold of $x$ defined in Lemma 4.4.

$\mathfrak{r}_\delta(x)$ – sizes of Lyapunov charts at $x$ (see (2.3)).

$\mathfrak{R}_\delta(x)$ – a function measuring the nonuniformity in the Multiplicative Ergodic Theorem for $x$ (see Definition 2.1 and Theorem 2.2).

$\tilde{R}_{i,j}(r)$ – the domains of $\mathcal{F}_{i,j}$ (see (5.16)).

$\tilde{r}_n = e^{-\eta_2 \varepsilon n - \varepsilon^2 n}$ (see (7.1)).

$\mathcal{R}_n$ - union of $\mathcal{F}$–saturations of $\tilde{R}_{i,j}(r)$ defined by (7.14).

$\widetilde{W}_x^{cs,n,\delta}(\bar{y})$ – the image of fake center stable foliation in the Lyapunov charts (see Definition 2.5).

$\widetilde{W}^u(x)$ – the image of (fake) unstable manifold in the exponential coordinates (see Lemma 2.8).

$W_{x,R}^u$ – the local stable manifold of $x$ of size $R$ (see (2.11)).

$W_x^u$ – the local unstable manifold of $x$ (see (2.12)).

$\widehat{\mathcal{W}}_i$ – reference unstable manifolds (see (5.10)).

$\alpha$ – Hölder regularity of $Df$.

$\alpha_1$ – the regularity of the Oseledets spaces on Pesin set (see Lemma 2.3).

$\alpha_2 = \min(\alpha/2, \alpha_1)$ – regularity of the rectifying maps $L_x$ (see (A.5)).

$\alpha_3$ – the (controlled) regularity of center stable manifolds (see Lemma 2.6).

$\alpha_4$ – the regularity of unstable manifolds (see Lemma 2.8).

$\alpha_5$ – cone contraction stability exponent (see Lemma 2.11).

$\alpha_6$ – the (controlled) regularity of the tangent spaces of fake center stable manifolds (see Lemma A.5).

$\alpha_7$ – the (controlled) regularity of the tangent spaces of admissible manifolds (see Lemma A.10).

$\beta$ – regularity exponent of admissible manifolds in Definition 6.2.

$\eta_x^u : \mathbb{R}^u \to \mathbb{R}^{cs}$ – a function whose graph defines $\widetilde{W}^u(x)$ (see Lemma 2.8).

$\tilde{\eta}_{x,\bar{y}}^{cs,n,\delta} : \mathbb{R}^{cs} \to \mathbb{R}^u$ – the function defining the graph of $\tilde{W}_x^{cs,n,\delta}(\bar{y})$ (see Lemma 2.6).

$\eta_2$ – the smallness exponent sufficient to guarantee the exponential mixing on parallelograms (see Lemma B.2).

$\lambda$ – the smallest positive exponent (see Lemma 2.4).

$\xi_n = e^{\varepsilon^2 n - \eta_2 \varepsilon n}$ (see (5.6)).

$\boldsymbol{\rho}(x)$ – density of $\mu$ in the product coordinates (see (5.2)).

$\tau$ – the parameter of the Pesin set so that $\mu(P_\tau) \geqslant 1 - \varepsilon^{\mathbf{b}}$.

1.6. **Layout of the paper.** The proof of Theorem 1.1 is carried out in Sections 3–9 following the outline given above. Section 2 contains the necessary background. Technical estimates from the Pesin theory needed in our proof are collected in the appendix. Section 10 describes several skew products with Anosov base and homogenous dynamics in the fiber where the Bernoulli property follows from our main result. These examples seem unaccessible by other methods. The results are sharp as the failure of exponential mixing also leads to the failure of the Bernoulli property.

## 2. PRELIMINARIES

2.1. **The Pesin theory.** Let $f$ be a $C^{1+\alpha}$ diffeomorphism of $(M, \mu)$ with $\dim(M) = D$. We denote the derivative of $f$ by $Df_x : T_x M \to T_{fx} M$ and let $Df_x^n := Df_{f^{n-1}x} \circ \ldots \circ Df_x$. We assume throughout that $\mu$ is smooth and that $f$ is ergodic with respect to $\mu$.

**Definition 2.1.** [12, Thm 2.4 & Prop 2.6][6, Thm 3.4.10, Thm 3.5.5, Prop 3.5.8.] *Let* $\lambda > 0$. *A point* $x$ *is said to be* $(\lambda, \delta)$-*Lyapunov regular if there is a splitting*

$$T_{f^k x} M = E^{cs}(f^k x) \oplus E^u(f^k x)$$

*for* $k \in \mathbb{Z}$ *and there are numbers* $\mathfrak{R}_\delta(f^k x) > 0$ *such that*

*(1)* $\mathfrak{R}_\delta(f^{k+n}x) \leqslant e^{\delta|k|}\mathfrak{R}_\delta(f^n x)$ *for every* $k \in \mathbb{Z}$, $n \in \mathbb{Z}$;

*(2)* $Df^k E^\sigma(x) = E^\sigma(f^k x)$ *for every* $k \in \mathbb{Z}$, $\sigma = cs, u$;

*(3) if* $v \in E^{cs}(f^k x)$, $n \geqslant 0$, *then*

$$\|Df^n v\| \leqslant \mathfrak{R}_\delta(f^{n+k}x)e^{\delta n}\|v\|;$$

*(4) if $v \in E^u(f^k x)$, $n \leqslant 0$, then*

$$\|Df^n v\| \leqslant \mathfrak{R}_\delta(f^{n+k}x)e^{n\lambda+|n|\delta}\|v\|;$$

*(5) $\angle(E^{cs}(f^k x), E^u(f^k x)) \geqslant \mathfrak{R}_\delta^{-1}(f^k(x))$.*

Note that $(\lambda, \delta)$-Lyapunov regular point is $(\lambda, \delta')$-Lyapunov regular for any $\delta' > \delta$.

**Theorem 2.2.** [6, Thm 3.4.10, Thm 3.5.5, Prop 3.5.8.] *Let $\lambda$ be the smallest positive Lyapunov exponent of $f$. Then for each $\delta > 0$ the set of $(\lambda, \delta)$-Lyapunov regular points has full measure. Moreover, the function $x \to \mathfrak{R}_\delta(x)$ can be chosen Borel measurable.*

Fix $\lambda > 0$ from the above theorem and let $\delta$ be a small number. Let

$$(2.1) \qquad LyapReg(\delta) := \{x \in M : x \text{ is } (\lambda, \delta/4)\text{-Lyapunov regular}\}.$$

Notice that since $f$ is $C^{1+\alpha}$, there is a constant $C_1 > 0$ such that denoting by $\exp_p$ the exponential map and letting $\hat{f}_p = \exp_{f_p^{-1}} \circ f \circ \exp_p$ we have that $\hat{f}_p$ is defined on the ball of radius $1/C_1$ in $T_p M$ and

$$(2.2) \qquad \|(D_{\hat{z}_1}\hat{f}_p)^{-1} - (D_{\hat{z}_2}\hat{f}_p)^{-1}\|_{T_{f_p}M \to T_p M} \leqslant C_1 |\hat{z}_1 - \hat{z}_2|_p^\alpha$$

if $|\hat{z}_1|_p, |\hat{z}_2|_p \leqslant \frac{1}{C_1}$. For $p \in LyapReg(\delta)$ let

$$(2.3) \qquad \mathfrak{r}_\delta(p) := \left(\delta^{-1}\frac{(\mathfrak{R}_{\delta/4}(p))^2}{\sqrt{1-e^{-\delta}}}(\sqrt{2})^{1+\alpha}C_1\right)^{-2/\alpha}.$$

The precise formula for $\mathfrak{r}_\delta(p)$ will not be important in our arguments. We will just use that it is uniformly bounded from below on Pesin sets defined in (2.4) below. We may assume that $\mathfrak{r}_\delta(p) \leqslant \frac{1}{C_1}$ (see e.g. the comment below Lemma A.2). Moreover by property (1) in Definition 2.1, we have that

$$e^{-\frac{\delta}{\alpha}}\mathfrak{r}_\delta(p) \leqslant \mathfrak{r}_\delta(fp) \leqslant e^{\frac{\delta}{\alpha}}\mathfrak{r}_\delta(p).$$

In what follows (see eg. Lemma 2.4) we will work with a rescaling of $\mathfrak{r}(\cdot)$, $\mathfrak{r}'_\delta(\cdot) = \mathfrak{r}_{\delta\alpha}(\cdot)$. To simplify notation and since we will work with the rescaled function from now on, we will denote the rescaling also by $\mathfrak{r}_\delta(\cdot)$.

Let us write $\mathbb{R}^D = \mathbb{R}^u \times \mathbb{R}^{cs}$. Sometimes we will write $\mathbb{R}^D_x$ to emphasize that it correspond to the tangent space of $x$.

We now define the Lyapunov norm. Let $x \in LyapReg(\delta/4)$. Define for $u \in E^u(x)$

$$|u|'^2_{x,\delta} = \sum_{m \leqslant 0} |D_x f^m u|^2_{f^m x} e^{-2\lambda m - 2\delta|m|}$$

and for $v \in E^{cs}(x)$

$$|v|'^2_{x,\delta} = \sum_{m \geqslant 0} |D_x f^m v|^2_{f^m x} e^{-2\delta|m|}.$$

These norms define inner products naturally and we extend it to an inner product on $T_x M = E^u(x) \oplus E^{cs}(x)$ by declaring these spaces orthogonal.

Next we define the Pesin sets. Let

$$(2.4) \qquad \hat{P}_\tau = \hat{P}_\tau^\delta := \{x \in M : \mathfrak{r}_\delta(x) \geqslant \tau\}.$$

Since $\mathfrak{r}_\delta(\cdot)$ is measurable it follows that $\mu(\hat{P}_\tau) \to 1$ when $\tau \to 0$. (Note also that $\hat{P}_{\tau'} \subset \hat{P}_\tau$ for $\tau < \tau'$).

**Lemma 2.3.** *There exists $\alpha_1 > 0$ such that for every $\tau, \delta > 0$, $\delta < \delta_0$ there is $K = K(\tau, \delta) > 0$ such that the maps $x \mapsto E^u(x)$ and $x \mapsto E^{cs}(x)$ are $(K, \alpha_1)$-Hölder continuous on $\hat{P}_\tau^\delta$ .*

*Proof.* The proof of this is the same as the proof of Theorem 5.3.2 in [6] using the observation that the proof of this theorem does not require that $E^{cs}$ is contracting. $\square$

**Lemma 2.4** (Lyapunov charts). *[see [12, Prop 5.1],[6, Thm 5.6.1] and [11, Sec 11.2] for one sided charts] There is $\alpha_2, \delta_0 > 0$ such that for every $\delta < \delta_0$ there are linear maps $L_{x,\delta} : \mathbb{R}^D \to T_x M$, such that*

   L1. *$L_{x,\delta}$ is an isometry between the standard metric in $\mathbb{R}^D$ and the $|\cdot|'_{x,\delta}$ metric;*
   L2. *for every sufficiently small $\tau > 0$, the maps $x \mapsto L_{x,\delta}$ are $\alpha_2$- Hölder continuous on $\hat{P}_\tau$.*

   *Moreover defining*

$$h_{x,\delta} = \exp_x \circ L_{x,\delta}, \quad \text{and} \quad \tilde{f}_{x,\delta} = h_{fx,\delta}^{-1} \circ f \circ h_{x,\delta},$$

*we have*

   *(1) $h_{x,\delta}(0) = x$;*
   *(2) $L_{x,\delta}(\mathbb{R}^u) = E^u(x)$ and $L_{x,\delta}(\mathbb{R}^{cs}) = E^{cs}(x)$;*
   *(3) $\max(\|L_{x,\delta}\|, \|L_{x,\delta}^{-1}\|) \leqslant \mathfrak{r}_\delta^{-1}(x)$;*
   *(4) $domain(\tilde{f}_{x,\delta}) \supset B_{\mathfrak{r}_\delta(x)}(0)$ and $domain(\tilde{f}_{x,\delta}^{-1}) \supset B_{\mathfrak{r}_\delta(x)}(0)$;*
   *(5) $e^{\lambda-\delta}|v| \leqslant |D_0 \tilde{f}_{x,\delta}(v)|$ for $v \in \mathbb{R}^u$, $\quad |D_0 \tilde{f}_{x,\delta}(v)| \leqslant e^\delta |v|$ for $v \in \mathbb{R}^{cs}$;*
   *(6) $H\ddot{o}l_{\alpha_2}(D\tilde{f}_{x,\delta}) \leqslant \delta$, $Lip(\tilde{f}_{x,\delta} - D_0\tilde{f}_{x,\delta}) \leqslant \delta$ and $Lip(\tilde{f}_{x,\delta}^{-1} - D_0\tilde{f}_{x,\delta}^{-1}) \leqslant \delta$;*
   *(7) $\max\left(Lip(h_{x,\delta}), Lip(h_{x,\delta}^{-1})\right) \leqslant \mathfrak{r}_\delta^{-1}(x)$.*

While this lemma is standard (see the references above) we recall the proof (the statement of Lemma 2.4 is obtained by combining Lemmas A.3, A.4,A.1 from Appendix A) in Appendix A since the intermediate steps in its proof are also important in the derivation of other estimates which will be used in the paper and are described below.

We extend $\tilde{f}_{x,\delta}$ to all $\mathbb{R}^D$ by making it linear outside of the ball of radius $2\mathfrak{r}_\delta(x)$ and with same bounds as in Lemma 2.4 (note that this is possible by taking a smaller $\alpha_2$ if necessary, see the definition of $\mathfrak{r}_\delta(\cdot)$).

## 2.2. **Hadamard-Perron, center-stable foliation.**

**Definition 2.5.** *Given $0 < \delta < \delta_0$, $x \in LyapReg(\delta)$ and $n \geqslant 0$, we define the foliation $\widetilde{W}_x^{cs,n,\delta}$ on $\mathbb{R}_x^D$ by by pulling back the foliation by planes parallel to $\mathbb{R}_x^{cs}$ via $\tilde{f}_{x,\delta}^{(n)}$, i.e.*

$$\widetilde{W}_x^{cs,n,\delta}(\bar{y}) = (\tilde{f}_{x,\delta}^{(n)})^{-1}\left(\mathbb{R}^{cs} + \tilde{f}_{x,\delta}^{(n)}(\bar{y})\right).$$

*Let $\tilde{E}_x^{cs,n,\delta}(\bar{y}) = T_{\bar{y}} \tilde{W}_x^{cs,n,\delta}(\bar{y})$.*

Notice that for $k \in [0, n]$,

$$\tilde{f}_{x,\delta}^{(k)}(\tilde{W}_x^{cs,n,\delta}(\bar{y})) = \tilde{W}_{f^k x}^{cs,n-k,\delta}(\tilde{f}_{x,\delta}^{(k)}(\bar{y})).$$

The following lemma is crucial:

**Lemma 2.6.** *There exists $\alpha_3, \delta_f > 0$ such that the following holds: for every $\delta \in (0, \delta_f)$, any $x \in LyapReg(\delta)$, every $\bar{y} \in \mathbb{R}^D$, there is $\tilde{\eta}_{x,\bar{y}}^{cs,n,\delta} : \mathbb{R}^{cs} \to \mathbb{R}^u$ such that*

$$\tilde{W}_x^{cs,n,\delta}(\bar{y}) = \text{graph}\left(\tilde{\eta}_{x,\bar{y}}^{cs,n,\delta}\right),$$

*and*

$$\|D\tilde{\eta}_{x,\bar{y}}^{cs,n,\delta}\|_{C^0} \leqslant \frac{3\delta}{1 - e^{-\lambda+\sqrt{\delta}}}$$

*and*

(2.5) $$[D\tilde{\eta}_{x,\bar{y}}^{cs,n,\delta}]_{C^{\alpha_3}} \leqslant \frac{12\delta}{1 - e^{-\lambda+\sqrt{\delta}}}$$

We will prove the above lemma in §A.2 of the appendix.

We finish this subsection with the following straightforward lemma:

**Lemma 2.7.** *Let $\phi : GL(\mathbb{R}^D) \times Grass^u(\mathbb{R}^D) \to \mathbb{R}$ be $\phi(A, E) = \log|\det(A|E)|$. Given $C_0$ there is a constant $N(C_0)$ such that if $\|A\|, \|B\|, \|A^{-1}\|, \|B^{-1}\| \leqslant C_0$ and $E$ and $F$ are graphs of the maps $L_E, L_F : \mathbb{R}^a \to \mathbb{R}^b$, respectively, where $\mathbb{R}^D = \mathbb{R}^a \times \mathbb{R}^b$, then*

$$|\phi(A, E) - \phi(B, F)| \leqslant N(C_0)(\|A - B\| + \|L_E - L_F\|).$$

In the proof of our main result (see the beginning of Section 5) we work with fixed $\varepsilon > 0$. We always assume that $\delta > 0$ (from the above results) is small enough in terms of $\varepsilon$ but fixed (e.g $\delta = \varepsilon^{10}$ would do). Therefore, we will omit it in the notation below.

Denote by $B^{cs}(z, R)$ the ball of radius $R > 0$ centered at $z \in \mathbb{R}^{cs}$, with an analogous notation for $B^u(z, R)$. If $z = 0$ we simply denote $B^{cs}(R) = B^{cs}(0, R)$ with an analogous notation for $B^u(R)$.

Let

(2.6) $$Q_x := h_x(B_{\mathfrak{r}_\delta(x)}(0)) \quad \text{and} \quad Q_x^{(n)} := \bigcap_{k=0}^{n-1} f^{-k} Q_{f^k x};$$

notice that $Q_x^{(n)} = h_x(\bigcap_{k=0}^{n-1} (\tilde{f}_x^{(k)})^{-1}(B_{\mathfrak{r}_\delta(f^k x)}))$. We also define

(2.7) $$Q_x(r_u, r_{cs}) := h_x(B^u(r_u) \times B^{cs}(r_{cs})),$$

(2.8) $$Q_x(r) = Q_x(r, r), \quad Q_x(A, r) = h_x(A \times B^{cs}(r)).$$

Let $d^u = \dim E^u$ and $d^{cs} = \dim E^{cs}$.

In the paper it will be more convenient to work with cubes rather than balls (i.e. balls in the *maximum* norm $\|x\|_\infty = \max|x_i|$). Let $C^u(R) \subset \mathbb{R}^u$ be a cube centered at 0 of side length $R$, i.e. a ball centered at 0 of radius $R/2$ in the metric $\|\cdot\|_\infty$.

**Lemma 2.8.** *(Lemma A.8 and Corollary A.9) There are constants $C_0, \alpha_4 > 0$ such that for every $x \in LyapReg$,*

$$\widetilde{W}_x^u := \{y \in \mathbb{R}^D : \limsup \frac{1}{n} \log |\tilde{f}_x^{(n)}(y)| < 0\}$$

*is the graph of a $C^{1+\alpha_4}$ function*

(2.9)
$$\eta = \eta_x^u : \mathbb{R}^u \to \mathbb{R}^{cs}$$

*with*

(2.10)
$$\|\eta\|_{C^{1+\alpha_4}} \leqslant C_0, \quad \eta(0) = 0, \quad D_0\eta = 0.$$

*Moreover if $\bar{z}_1, \bar{z}_2 \in \widetilde{W}_x^u$ then*

$$|(\tilde{f}_{f^{-1}x})^{-1}(\bar{z}_1) - (\tilde{f}_{f^{-1}x})^{-1}(\bar{z}_2)| \leqslant e^{(-\lambda+2\delta)}|\bar{z}_1 - \bar{z}_2|;$$

*and if we define*

(2.11)
$$\widetilde{W}_{x,R}^u := \mathrm{graph}(\eta, C^u(R)), \quad W_{x,R}^u = h_x(\widetilde{W}_{x,R}^u)$$

*then for every $x \in Lyapreg$ and $R > 0$,*

$$(\tilde{f}_x)^{-1}\widetilde{W}_{fx,R}^u \subset \widetilde{W}_{x,e^{(-\lambda+3\delta)}R}^u.$$

Set

(2.12)
$$W_x^u = W_{x,\mathfrak{r}_\delta(x)}^u.$$

Notice that $f^{-1}W_x^u \subset W_{f^{-1}x}^u$ since by Lemma 2.8 $f^{-1}(W_{x,\mathfrak{r}_\delta(x)}^u) \subset W_{f^{-1}x,\mathfrak{r}_\delta(f^{-1}x)}^u(f^{-n}x)$. In some papers $W_x^u$ is called *local* unstable manifold of $x$ and is denoted $W_{x,loc}^u$. We do not use the subscript *loc* since we will not need to consider global unstable manifolds of $x$.

By [6, Theorem 7.1.1] (see also Corollary A.9) there exists a measurable function $\mathfrak{K}(x)$ such that the size of the unstable manifold of $x$ is greater than $1/\mathfrak{K}(x)$ and moreover for each $y \in W_x^u$

(2.13)
$$d(f^{-n}x, f^{-n}y) \leqslant \mathfrak{K}(x)e^{(\delta-\lambda)n}d(x,y)$$

i.e. $W_x^u$ is exponentially contracted. Let

(2.14)
$$\tilde{P}_\tau = \{x : \mathfrak{K}(x) \leqslant \tau\}.$$

In what follows for a given $\varepsilon$ we will pick $\tau$ small enough so that if we define (see (2.4))

(2.15)
$$P_\tau := \hat{P}_\tau \cap \tilde{P}_{\tau^{-1}},$$

then $\mu(P_\tau) \geqslant 1 - \varepsilon^{\mathbf{b}}$, where $\mathbf{b} = 10^{10}$.

We start with the following observation:

**Lemma 2.9.** *There exists a constant $K'$ such that for every $x \in LyapReg$,*

$$\|D_0\tilde{f}_x\|, \|D_0\tilde{f}_x^{-1}\| < K'.$$

*Proof.* This follows from (A.1) in Lemma A.3. $\qquad\square$

2.3. **Cones.** Let

$$\mathcal{C}_{\mathfrak{b}}^u = \{v = v^u + v^{cs} \in \mathbb{R}^D : |v^{cs}| \leqslant \mathfrak{b}|v^u|\}, \tag{2.16}$$

and

$$\mathcal{C}_{\mathfrak{b}}^{cs} = \{v = v^u + v^{cs} \in \mathbb{R}^D : |v^u| \leqslant \mathfrak{b}|v^{cs}|\} \tag{2.17}$$

be the $\mathfrak{b}-$cones around $\mathbb{R}^u$ and $\mathbb{R}^{cs}$. For $x \in LyapReg$ and $y \in Q_x$, define the cones

$$\mathcal{C}_x^u(y) = (D_y h_x^{-1}) h_x \mathcal{C}_{1/2}^u, \quad \mathcal{C}_x^{cs}(y) = (D_y h_x^{-1}) h_x \mathcal{C}_{1/2}^{cs}. \tag{2.18}$$

We have the following:

**Lemma 2.10.** *For every $\delta, \tau > 0$ there exists $n_{\delta,\tau}$ such that for every $n \geqslant n_{\delta,\tau}$, every $x \in \hat{P}_\tau^\delta \cap f^{-n}\hat{P}_\tau^\delta$, and every $y \in Q_x^{(n)}$ we have*
  *(1) if $v \in \mathcal{C}_x^u(y)$ $\|D_y f^n v\| \geqslant \tau^2 e^{n(\lambda-3\delta)}\|v\|$;*
  *(2) if $v \in \mathcal{C}_x^{cs}(y)$ then for each $0 \leqslant k \leqslant n$ it holds $\|D_y f^k v\| \leqslant \tau^{-2} e^{2\delta k}\|v\|$;*
  *(3) $Df_x^n(\mathcal{C}_x^u(y)) \subset \mathcal{C}_{f^n(x)}^u(f^n(y))$;*
  *(4) $Df_x^n(\mathcal{C}_x^{cs}(y)) \supset \mathcal{C}_{f^n(x)}^{cs}(f^n(y))$;*

The proof is an immediate consequence of the definition of $h_x$ and the following Lemma:

**Lemma 2.11.** *There is $\alpha_5 > 0$ such that for every $\mathfrak{b} < 1$ there is $C_0 > 0$ such that for every $\delta \in (0, \delta_0)$ there exists $r_\delta > 0$ such that for every $r \in (0, r_\delta)$ satisfying $C_0 r^{\alpha_5} < \mathfrak{b} < 1$ every $x \in LyapReg$ and $|w| < r$,*

$$|D_w \tilde{f}_x(v)| \geqslant e^{\lambda-2\delta} \text{ for } v \in \mathcal{C}_{\mathfrak{b}}^u, \quad |D_w \tilde{f}_x(v)| \leqslant e^{2\delta} \text{ for } v \in \mathcal{C}_{\mathfrak{b}}^{cs},$$

$$D_y \tilde{f}_x \mathcal{C}_{\mathfrak{b}}^u \subset \mathcal{C}_{e^{-\lambda+4\delta}\mathfrak{b}}^u \quad \text{and} \quad \mathcal{C}_{\mathfrak{b}}^{cs} \subset D_y \tilde{f}_x \mathcal{C}_{e^{\lambda+4\delta}\mathfrak{b}}^{cs}.$$

*Proof.* This is a straightforward consequence of $f$ being $C^{1+\alpha}$ and Lemma 2.4. $\square$

2.4. **Conditional measure along unstables.** For $x \in LyapReg$, we let $m_x^u$ to be Lebesgue measure on $W_x^u$.

**Lemma 2.12.** [6, Thm 8.6.8 and its proof and Thm 8.6.13 and Thm 9.3.4] *For $x \in LyapReg$ and $z \in W_x^u$, we define $\rho(z, \cdot) : W_z^u \to \mathbb{R}$,*

$$\rho(z, y) = \prod_{j \geqslant 1} \frac{\det \left(Df_{f^{-n}(z)}|E^u(f^{-n}(z))\right)}{\det \left(Df_{f^{-n}(y)}|E^u(f^{-n}(y))\right)},$$

*it is Hölder continuous. Moreover, defining $\rho_x := \rho(x, \cdot)$ we have that for every $\tau$, $\hat{P}_\tau \ni x \to \rho_x$ is Hölder continuous and $\rho(x, y)\rho(y, z) = \rho(x, z)$.*

For $\mathcal{W} \subset W_x^u$, we define

$$m_{\mathcal{W}}^u(A) = \frac{\int_A \rho(x, y) dm_x^u(y)}{\int_{\mathcal{W}} \rho(x, y) dm_x^u(y)}. \tag{2.19}$$

Note that if $z \in W_x^u$ and we use $\rho_z$ to define $\mu_{\mathcal{W}}$ we will get the same formula.

For $x \in \hat{P}_\tau$, we take a transversal $T_x = h_x(\text{graph}(\psi))$, where $\psi : \mathbb{R}_x^{cs} \to \mathbb{R}_x^u$ is $C^1$ function with $\|\eta\|_{C^1} < C_1$, and we define $\tilde{T}_x^\tau = \bigcup_{y \in \hat{P}_\tau \cap Q_x} T_x \cap W_{y,\tau}^u$ and $\tilde{Q}_x^\tau = \bigcup_{y \in \hat{P}_\tau \cap Q_x} W_{y,\tau}^u$. We have the following lemma:

**Lemma 2.13.** *([39, §3.3]) Given $\tau > 0$ and $x \in \hat{P}_\tau$, if $A \subset Q_x$ then*

$$\mu(A \cap \tilde{Q}_x^\tau) = \int_{\tilde{T}_x^\tau} d\nu_x^\tau(z) \int_{W_{z,\tau}^u \cap A} \rho(z,y) dm_z^u(y)$$

*where $\nu_x^\tau$ is absolutely continuous w.r.t. Lebesgue on $\tilde{T}_x^\tau \subset T_x$. Also we have that*

$$\frac{m_{T_x}^{cs}(T_x \setminus \tilde{T}_x^\tau)}{m_{T_x}^{cs}(T_x)} \to 0$$

*as $\tau \to 0$ where $m_{T_x}^{cs}$ is Lebesgue measure on the transversal $T_x$.*

2.5. **Measure theory.**

**Definition 2.14.** *A map $\theta : (X_1, \nu_1) \to (X_2, \nu_2)$ between two measure spaces is called $\varepsilon$-measure preserving if there exists a set $E_1 \subset X_1$, $\nu_1(E_1) < \varepsilon$ and such that for every $A \in X_1 \setminus E_1$, we have*

$$\left| \frac{\nu_2(\theta(A))}{\nu_1(A)} - 1 \right| < \varepsilon.$$

The following fact will be useful in constructing $\varepsilon$-measure preserving maps. By *Lebesgue space* we mean a probability measure defined on a Borel $\sigma$-algebra in a Polish space.

**Lemma 2.15.** *([46]) Any two atomless Lebesgue spaces are isomorphic.*

2.6. *$K$ **and Bernoulli properties.*** Let $(X, \mathcal{B}, \mu)$ be a standard probability Borel space and let $\mathcal{P} = (P_1, \ldots P_k)$ and $\mathcal{Q} = (Q_1, \ldots, Q_k)$ be two finite measurable partitions of $X$. Let $\mathcal{P} \vee \mathcal{Q}$ be the partition into sets of the form $P_i \cap Q_j$, $i, j \in \{1, \ldots, k\}$. Let $T$ be an automorphism of $(X, \mathcal{B}, \mu)$. We say that $\mathcal{P}$ is *generating* if $\bigvee_{-\infty}^{+\infty} T^i \mathcal{P} = \mathcal{B}$.

We say that a property holds for $\varepsilon$ a.e. atom of a partition $\mathcal{Q}$ if the union of all atoms for which the property does not hold has measure $\leqslant \varepsilon$.

We recall the definition of $K$-property:

**Definition 2.16.** *Let $T$ be an automorphism of $(X, \mathcal{B}, \mu)$ and let $\mathcal{P}$ be a finite partition of $X$. We say that $\mathcal{P}$ is a $K$-partition if for every $D \in \bigvee_{-\infty}^{+\infty} T^i \mathcal{P}$ and every $\varepsilon > 0$ there exists $N_0 = N_0(\varepsilon, D)$ such that for every $N' \geqslant N \geqslant N_0$, $\varepsilon$ a.e. atom $A \in \bigvee_N^{N'} T^i \mathcal{P}$ satisfies*

$$\left| \frac{\mu(A \cap D)}{\mu(A)} - \mu(D) \right| < \varepsilon.$$

*We say that $T$ has the* Kolmogorov *property ($K$ property) if there exists a generating $K$-partition. It then follows that* every *partition is a $K$-partition.*

We will need the following simple modification of the original definition of the $K$-property:

**Lemma 2.17.** *$T$ has the $K$ property if there exists a* generating *partition $\mathcal{P}$ such that for every $D \in \mathcal{B}$ and every $\varepsilon > 0$ there exists $n \in \mathbb{N}$ and $N_0 = N_0(\varepsilon, D)$ such that for every $N' \geqslant N \geqslant N_0$, $\varepsilon$ a.e. atom $A \in \bigvee_N^{N'} T^i \mathcal{P}$ satisfies*

(2.20) $$\left| \frac{\mu(A \cap f^{-n}D)}{\mu(A)} - \mu(D) \right| < \varepsilon.$$

*Proof.* The proof is an immediate consequence of the following observation: if $A \in \bigvee_N^{N'} T^i \mathcal{P}$, then $A = f^n(A')$ where $A' \in \bigvee_{N-n'}^{N'-n} T^i \mathcal{P}$. Since $\mu(A \cap D) = \mu(A' \cap f^{-n}D)$ it easily follows that the statement of the lemma is equivalent to the original definition of the $K$ property. $\qquad\square$

By [47] the $K$ property is equivalent to *completely positive entropy*: every *factor* of $T \in Aut(X, \mathcal{B}, \mu)$ has positive entropy.

**Definition 2.18.** *$T \in Aut(X, \mathcal{B}, \mu)$ is Bernoulli if it is measure theoretically isomorphic to the* Bernoulli shift, *i.e. the shift map on the space* $(\{1, \ldots, \ell\}^{\mathbb{Z}}, \mathbf{p}^{\mathbb{Z}})$, *where* $\mathbf{p} = (p_1, \ldots, p_\ell)$ *is a probability vector.*

For $A \subset X$, $\mathcal{P}_{|A}$ denotes the induced partition of the space $(A, \mu_{|A})$, i.e.

$$\mathcal{P}_{|A} := (P_1 \cap A, \ldots, P_k \cap A) \text{ and } \mu_{|A}(B) = \frac{\mu(A \cap B)}{\mu(A)}.$$

We introduce the following distance on the space of partitions of $(X, \mu)$:

$$\bar{d}(\mathcal{P}, \mathcal{Q}) := \sum_{i=1}^{k} \mu(P_i \triangle Q_i).$$

Now let $\mathcal{P}^s = (\mathcal{P}_1^s, \ldots, \mathcal{P}_k^s)$, $s = 1, \ldots, S$ be a sequence of finite partitions of $(X, \mu)$ and $\mathcal{Q}^s = (\mathcal{Q}_1^s, \ldots, \mathcal{Q}_k^s)$, $s = 1, \ldots, S$ be a sequence of finite partitions of $(Y, \nu)$. If additionally $(X, \mu) = (Y, \nu)$, then

$$\bar{d}\left((\mathcal{P}^s)_{s=1}^{S}, (\mathcal{Q}^s)_{s=1}^{S}\right) := \frac{1}{S} \sum_{s=1}^{S} \bar{d}(\mathcal{P}^s, \mathcal{Q}^s).$$

More generally, if $(\mathcal{P}^s)_{s=1}^{S}$ and $(\mathcal{Q}^s)_{s=1}^{S}$ are partitions of different spaces, we say that $\mathcal{P}^s \sim \mathcal{Q}^s$ for $s = 1, \ldots, S$ if $\mu(P_i^s) = \nu(Q_i^s)$ for $i = 1, \ldots, k$ and $s = 1, \ldots, S$. We can then compare the distance between $(\mathcal{P}^s)_{s=1}^{S}$ and $(\mathcal{Q}^s)_{s=1}^{S}$ by setting

$$\bar{d}\left((\mathcal{P}^s)_{s=1}^{S}, (\mathcal{Q}^s)_{s=1}^{S}\right) = \inf_{\bar{\mathcal{Q}}^s \sim \mathcal{Q}^s, \ s=1,\ldots,S} \bar{d}\left((\mathcal{P}^s)_{s=1}^{S}, (\bar{\mathcal{Q}}^s)_{s=1}^{S}\right),$$

where the infimum is taken over sequences of partitions $\bar{\mathcal{Q}}^s$ of $(X, \mu)$. We denote by $T^n \mathcal{P}$ the partition given by $(T^n P_1, \ldots, T^n P_k)$.

**Definition 2.19** (Very weak Bernoulli, vwB)**.** *Let $T \in Aut(X, \mathcal{B}, \mu)$ and let $\mathcal{P}$ be a finite partition of $X$. Then $\mathcal{P}$ is a very weak Bernoulli partition (vwB partition) if for every $\varepsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that for every $N' \geqslant N \geqslant N_0$ every $S \geqslant 0$ and $\varepsilon$ a.e. atom $A$ of $\bigvee_N^{N'} T^i(\mathcal{P})$, we have*

$$\bar{d}\left(\{T^{-i}\mathcal{P}\}_{s=0}^{S}, \{T^{-i}\mathcal{P}_{|A}\}_{s=0}^{S}\right) < \varepsilon.$$

The following classical theorem is a crucial tool in establishing Bernoullicity of a system (see e.g. [36]). Recal that a sequence of partitions $(\mathcal{P}_k)_{k=1}^{+\infty}$ of $(X, \mathcal{B}, \mu)$ *converges to partition into points* if the smallest $\sigma$-algebra with respect to which all $\mathcal{P}_k$ are measurable, is $\mathcal{B}$.

**Theorem 2.20.** *If $(\mathcal{P}_k)_{k=1}^{+\infty}$ is a sequence of partitions of $(X, \mathcal{B}, \mu)$ converging to partition into points and, for every $k \geqslant 1$, $\mathcal{P}_k$ is VWB partition for $T \in Aut(X, \mathcal{B}, \mu)$, then $T$ is a Bernoulli system.*

We will now recall the main method of establishing VWB property, [36]. For a partition $\mathcal{P} = (P_1, \dots P_k)$ of $(X, \mu)$, an integer $S \geqslant 1$ and $x \in X$ the $S, \mathcal{P}$-name of $x$ is a sequence $(x_i^{\mathcal{P}})_{i=0}^S \in \{1, \dots, k\}^{S+1}$ given by the condition $T^i(x) \in P_{x_i^{\mathcal{P}}}$. Let $e : \mathbb{Z} \to \mathbb{Z}$ be given by $e(0) = 0$ and $e(n) = 1$ for $n \neq 0$.

We have the following lemma

**Lemma 2.21** (Lemma 1.3. in [36]). *Let $T \in Aut(X, \mathcal{B}, \mu)$ and $\mathcal{P}$ be a finite partition of $X$. If for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for every $N' \geqslant N$, $\varepsilon$ a.e. atom $A \in \bigvee_N^{N'} T^i \mathcal{P}$ and every $S \geqslant 1$ there exists an $\varepsilon$-measure preserving map $\theta = \theta(N, S, A) : (A, \mu_{|A}) \to (X, \mu)$ such that*

$$(2.21) \qquad \bar{d}_S(x, \theta(x)) := \frac{1}{S} \sum_{i=0}^{S-1} e\left(x_i^{\mathcal{P}} - (\theta(x))_i^{\mathcal{P}}\right) < \varepsilon.$$

*then $\mathcal{P}$ is a VWB partition.*

A finite partition $\mathcal{P}$ of $(X, \mathcal{B}, \mu, d)$ is called *regular* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\mu(V_\delta(\partial \mathcal{P})) < \varepsilon,$$

where for $A \subset X$, $V_\delta(A)$ denotes the $\delta$ neighborhood of $A$ (in the metric $d$). For existence results of regular partitions we refer the reader to e.g. [37, Lemma 4.1]. In what follows we will always consider only regular partitions.

We shall use the following form of Lemma 2.21.

**Corollary 2.22.** *Let $f : (X, \mathcal{B}, \mu, d) \to (X, \mathcal{B}, \mu, d)$ satisfy the $K$-property and $\mathcal{P}$ be a regular partition of $X$. If for every $\varepsilon > 0$ there exists $N, \tilde{N} \in \mathbb{N}$ such that for every $N' \geqslant N$, $\varepsilon$ a.e. atom $A \in \bigvee_N^{N'} T^i \mathcal{P}$ and every $S \geqslant \tilde{N}$ there exists an $\varepsilon$-measure preserving map $\theta = \theta(N, S, A) : (A, \mu_{|A}) \to (X, \mu)$ such that for everu $x \in A$,*

$$(2.22) \qquad \frac{1}{S} Card\left(\{i \in \{0, \dots, S-1\} : d(f^i x, f^i(\theta x)) < \varepsilon\}\right) > 1 - \varepsilon,$$

*then $\mathcal{P}$ is a VWB partition.*

*Proof.* Fix $\varepsilon > 0$. We shall assume that (2.22) holds with $\varepsilon' = \varepsilon^{10}$ instead of $\varepsilon$ and verify that the assumptions of Lemma 2.21 hold. Since $\mathcal{P}$ is regular, there exists $\delta > 0$ such that $\mu(V_\delta(\mathcal{P})) \leqslant \varepsilon^8$. By ergodicity of $f$ (recall that $f$ is $K$), there exists $m \in \mathbb{N}$ and a set $E \subset X$, $\mu(E) \geqslant 1 - \varepsilon^4$ such that for every $x \in E$ and every $n \geqslant m$,

$$(2.23) \qquad \left|\{0 \leqslant i \leqslant n-1 : f^i(x) \notin V_\delta(\mathcal{P})\}\right| \geqslant (1 - \varepsilon^2)n.$$

Let $N$ and $\tilde{N}$ come from the assumptions of the statement of the corollary for $\varepsilon' = \min(\delta, \varepsilon^{10})$. We assume WLOG that $\tilde{N} \geqslant m$. We will first show that for $S \leqslant \tilde{N}$, the

assumptions of Lemma 2.21 are satisfied by the $K$-property. Let $\{D_i\}_{j=1}^{\tilde{J}}$ be all the atoms of

$$\bigvee_{-\tilde{N}-1}^{0} f^i \mathcal{P}.$$

Notice that by definition, if $x, y \in D_i$, then $f^j x$ and $f^j y$ are in the same atom of $\mathcal{P}$ for ever $0 \leqslant j \leqslant \tilde{N}$. By the $K$-property (see Definition 2.16) for $D = D_i \in \bigvee_{-\infty}^{+\infty} f^i \mathcal{P}$, $i = 1, \ldots, \tilde{J}$ (simultaneously), it follows that there exists $N_0$, such that for every $N' \geqslant N_0$, $\varepsilon^2$ a.e. atom $A \in \bigvee_{N_0}^{N'} f^i \mathcal{P}$ satisfies: for every $i \leqslant \tilde{J}$, we have

$$(2.24) \qquad\qquad \left| \frac{\mu(A \cap D_i)}{\mu(A)} - \mu(D_i) \right| \leqslant \varepsilon^2.$$

For $S \leqslant \tilde{N}$ and $i \leqslant \tilde{J}$, let $\theta_i = \theta_i(N, S, A) : (A \cap D_i, \mu_{|A}) \to (D_i, \mu_{|D_i})$ be any $\varepsilon$-measure preserving map (such map exists by (2.24) and Lemma 2.15). We then naturally define for $x \in A$, $\theta(x) = \theta_i(x)$, where $x \in A \cap D_i$. By definition, $\theta$ is $\varepsilon$ measure preserving and moreover for $x \in A \cap D_i$, $\theta(x) \in D_i$. This however implies that $\bar{d}_S(x, \theta(x)) = 0$, by the definition of $D_i$ and since $S \leqslant \tilde{N}$. This implies that Lemma 2.21 holds for $S \leqslant \tilde{N}$.

Consider now the case $S > \tilde{N}$. Let $N \geqslant N'$. We say that an atom $A \in \bigvee_N^{N'} f^i \mathcal{P}$ is *good* if it satisfies (2.22) and moreover

$$\mu(A \cap E) \geqslant (1 - \varepsilon^2) \mu(A).$$

Since $\mu(E) \geqslant 1 - \varepsilon^4$ it follows that $\varepsilon$ a.e. atom $A \in \bigvee_N^{N'} f^i \mathcal{P}$ is good. For a good $A$, let $\bar{\theta} = \theta(N, S, A) : (A, \mu_{|A}) \to (X, \mu)$ be the $\varepsilon'$- measure preserving map from the assertion of Corollary 2.22. We define $\theta(x) := \bar{\theta}(x)$ if $x \in A \cap E$ and $\theta(x) = x$ otherwise. Since $\mu(A \cap E) \geqslant (1 - \varepsilon^2) \mu(A)$ it follows that $\theta : (A, \mu_{|A}) \to (X, \mu)$ is $\varepsilon$ measure preserving. Now, if $x \notin A \cap E$, then $\theta(x) = x$ and so (2.21) trivially holds. If $x \in A \cap E$ then (2.22) holds for $x$ and $\theta(x)$ and $\varepsilon = \varepsilon'$. But since $x \in E$, we get that (2.23) holds for $x$ and $S \geqslant \tilde{N} \geqslant m$. Therefore, if $i \leqslant S - 1$ is a time satisfying (2.22) and (2.23), then $f^i x$ and $f^i(\theta(x))$ are in the same atom of $\mathcal{P}$ (since $\varepsilon' \leqslant \delta$). It remains to notice that the total cardinality of such $i \leqslant S$ is at least $(1 - \varepsilon)S$. This shows (2.21) and hence finishes the proof of Corollary 2.22. $\qquad\square$

## 3. Exponential mixing implies non-zero exponents

In this section we will prove Proposition 1.3. Let $f : (M, \mu) \to (M, \mu)$ be exponentially mixing. We will show that $f$ has at least one non-zero Lyapunov exponent. In this section we don't need to assume that the measure $\mu$ is smooth; it is enough that $\mathrm{supp}(\mu)$ is not a single point.

For $z \in M$ and for $r > 0$, let $O_r(z)$ be the ball of radius $r$ centered at $z$. We have the following lemma:

**Lemma 3.1.** *If $f$ is exponentially mixing for some non-atomic measure then there exist $c, \hat{\eta} > 0$ such that for every $B \subset M$ with $\mu(B) \geqslant 1 - c$,*

$$(3.1) \qquad\qquad \inf_{n \in \mathbb{N}} \max_{z \in B} \mathrm{diam}\left( f^n \left( O_{e^{-\hat{\eta}n}}(z) \right) \right) \geqslant c.$$

Note that (3.1) is equivalent to saying that for each $n$

$$\mu\left(x : \operatorname{diam}\left(f^n\left(O_{e^{-\hat{\eta}n}}(z)\right)\right) \geqslant c\right) \geqslant c.$$

In other words, the lemma says that the image of balls of radius $e^{-\hat{\eta}n}$ become macroscopic for a set of centers of a sizable measure. We will prove Lemma 3.1 in Appendix B.

*Proof of Proposition 1.3:* Assume by contradiction that all exponents of $f$ are non positive and let $\zeta = \hat{\eta}/1000$. Let $\zeta' > 0$ be such that $[\sup \|Df_x\|]^{\zeta'} < e^{\zeta}$ and $\zeta' < c/2$ where $c$ is from Lemma 3.1.

By Oseledets theorem there exists $C > 0$, a set $A \subset M$, $\mu(A) \geqslant 1 - \zeta'/2$ and $n_1 \in \mathbb{N}$ such that for every $x \in A$ and every $n \geqslant n_1$,

$$\|Df_x^n\| < Ce^{\zeta n}$$

By enlarging $C$ if necessary, we may assume that the above also holds for $n \leqslant n_1$. Using ergodic theorem for the function $\chi_A$ it follows that there exists $n_2$ and a set $B$, $\mu(B) \geqslant 1 - c$ such that for every $x \in B$ and every $n \geqslant n_2$,

$$\left|\{0 \leqslant m \leqslant n \ : \ f^m x \in A\}\right| \geqslant (1 - \zeta')n.$$

Then it follows that for every $x \in B$, $n \geqslant n_2$, and every $\ell \leqslant n$,

$$(3.2) \qquad \qquad \|Df_{f^\ell x}^{n-\ell}\| \leqslant Ce^{2\zeta n}.$$

Indeed, for $\ell \leqslant n$, let $m \geqslant \ell$ be the smallest number such that $f^m x \in A$. Since $x \in B$ it follows that $|m - \ell| \leqslant \zeta' n$. Using that $f^m x \in A$ and the definition of $\zeta'$, we get

$$\|Df_{f^\ell x}^{n-\ell}\| \leqslant \|Df_{f^\ell x}^{m-\ell}\| \cdot \|Df_{f^m x}^{n-m}\| \leqslant [\sup \|Df_x\|]^{\zeta' n} \cdot Ce^{\zeta n - m} \leqslant Ce^{2\zeta n}.$$

Moreover, by enlarging $C$, we can assume that (3.2) holds also for $n \leqslant n_2$.

We will show that for every $n$ sufficiently large, (3.1) does not hold. For $x \in B$ let $k = k(x, n)$ be the smallest number such that

$$\operatorname{diam}\left(f^k\left(O_{e^{-\hat{\eta}n}}(x)\right)\right) \geqslant e^{-\frac{1}{2}\hat{\eta}n}.$$

By Lemma 3.1 there is $x \in B$ such that $k(x, n) \leqslant n$. Then by the definition of $k$ it follows that for every $y \in O_{e^{-\hat{\eta}n}}(x)$ and every $0 \leqslant i < k \leqslant n$,

$$(3.3) \qquad \qquad d(f^i x, f^i y) \leqslant e^{-\frac{1}{2}\hat{\eta}n}.$$

We will show that there is a constant $C'$ such that for every $z \in O_{e^{-\hat{\eta}n}}(x)$ and every $0 \leqslant i \leqslant k \leqslant n$,

$$(3.4) \qquad \qquad \|Df_z^i\| \leqslant C'e^{3\zeta k}.$$

Before we prove (3.4), let us show how it implies the proposition.

Notice that if $k \leqslant n$ then using (3.4) and the mean value theorem, for every $i \leqslant k$ and every $z \in O_{e^{-\hat{\eta}n}}(x)$

$$d(f^k x, f^k z) \leqslant C^* e^{3\zeta k} d(x, z) \leqslant C^* e^{3\zeta k - \hat{\eta}n} \leqslant e^{-\frac{2}{3}\hat{\eta}n} < e^{-\frac{1}{2}\hat{\eta}n},$$

which contradicts the definition of $k$. This contradiction shows that $f$ in fact has non-zero exponent proving the proposition.

It remains to establish (3.4). We will proceed by induction on $i$. The case $i = 1$ just follows by taking $C' \geqslant \sup_{x \in M} \|Df_x\|$. For $j \in \mathbb{N}$, let $A_j := Df_{f^j x}$ and $B_j = B_j(z) = Df_{f^j z}$ and assume that (3.4) holds for all $\ell \leqslant i - 1$. This implies that for every $\ell \leqslant i - 1$,

$$\left\| \prod_{j=0}^{\ell-1} B_j \right\| \leqslant C' e^{3\zeta k}.$$

Then

$$\|Df_x^i - Df_z^i\| = \left\| \prod_{j=0}^{i-1} A_j - \prod_{j=0}^{i-1} B_j \right\| = \left\| \sum_{\ell=0}^{i-1} \left( \prod_{j=\ell+1}^{i-1} A_j \right) (A_\ell - B_\ell) \prod_{j=0}^{\ell-1} B_j \right\|$$

$$\leqslant \sum_{\ell=0}^{i-1} \left\| \prod_{j=\ell+1}^{i-1} A_j \right\| \|A_\ell - B_\ell\| \left\| \prod_{j=0}^{\ell-1} B_j \right\|.$$

Note that since $\ell \leqslant i - 1 \leqslant k - 1$, it follows by (3.3) that

$$\|A_\ell - B_\ell\| \leqslant C'' e^{-\frac{1}{2}\hat{\eta} n}.$$

Moreover, by (3.2) (using that $i \leqslant k$),

$$\left\| \prod_{j=\ell+1}^{i-1} A_j \right\| \leqslant C \cdot e^{2\zeta k}$$

Using the inductive assumption and the above bound (and $i \leqslant k$) we get that

$$\|Df_x^i - Df_z^i\| \leqslant CC'C''i \cdot e^{2\zeta k - \frac{1}{2}\hat{\eta} n + 3\zeta k} \leqslant C' e^{3\zeta k},$$

since $i \leqslant k \leqslant n$ and $\zeta = \hat{\eta}/1000$. This proves (3.4) and finishes the proof of the proposition. □

Combining Proposition 1.3, with Pesin entropy formula we see that if $f$ is exponentially mixing for a smooth measure $\mu$ then it has positive entropy. Our main result, Theorem 1.1 gives a much stronger conclusion, namely $(f, \mu)$ is Bernoulli. In particular, the exponentially mixing diffeos with the same entropy are isomorphic.

## 4. Almost u-saturation

In Sections 4–6 we assume that $f$ is $C^{1+\alpha}$ diffeomorphism of a compact manifold $M$ preserving a smooth measure $\mu$, such that at least one Lyapunov exponent of $f$ is non-zero. In particular, the assumption that $f$ is exponentially mixing will not be used until Section 7 .

**Definition 4.1.** *A set $A \subset M$ is called $(\varepsilon, \xi)$ u-saturated if there exists a set $E \subset A$ with $\mu|_A(E) \geqslant 1 - \varepsilon$ such that if $x \in E$ and $\mathcal{W}$ is a unstable box of size $\xi$ containing $x$ then $\mathcal{W} \subset A$.*

*A partition $\Pi$ is called $(\varepsilon, \xi)$ u-saturated if there exists a set $E \subset M$ with $\mu(E) \geqslant 1 - \varepsilon$ and such that if $x \in E$ and $\mathcal{W}$ is a unstable box of size $\xi$ containing $x$ then $\mathcal{W} \subset \Pi(x)$.*

We note that if $\Pi$ is $(\xi, \varepsilon^2)$ u-saturated then $\varepsilon$ almost every atom of $\Pi$ is $(\xi, \varepsilon)$ u-saturated. We will need the following lemma which uses exponential contraction of the unstable foliation. Let $\mathcal{P}$ be any partition of $M$ with smooth boundaries.

**Lemma 4.2.** *For every $\varepsilon > 0$ there exists $\xi = \xi(\varepsilon) > 0$ there exists $\bar{N} = N_{\xi,\varepsilon} \in \mathbb{N}$ such that for every $N_1 \geqslant N_2 \geqslant \bar{N}$, the partition $\mathfrak{P} = \bigvee\limits_{i=N_1}^{N_2} f^i(\mathcal{P})$ is $(\xi, \varepsilon)$ u-saturated.*

*Proof.* The proof is similar to the proof of Lemma 2.1. in [36].

Let $\tilde{P}_{\tau^{-1}}$ be the set from (2.14) where $\tau$ such that $\mu(\tilde{P}_{\tau^{-1}}) \geqslant 1 - \varepsilon^4$ and let $\xi = \tau$.

Notice that if for some $k \geqslant N_1$ and some atom $A \in f^k(\mathcal{P})$ we have $x \in \tilde{P}_\tau \cap A$ but $W_{x,\xi}^u \notin A$, then by exponential contraction, $d(f^{-k}x, \partial \mathcal{P}) \leqslant c^{-k}$ where $c = e^{\delta - \lambda} < 1$. Since $\mathcal{P}$ is piecewise smooth, the measure of points satisfying the last condition for some $k \geqslant N_1$ is at most $C \sum\limits_{k=N_1}^{+\infty} c^{-k}$ for some $C > 0$. The last expression can be made smaller than $\varepsilon^4$ by taking $N_1$ large enough. In summary, the set of points $x$ such that $\mathfrak{P}(x) \not\supset W_{x,\xi}^u$ has measure which is smaller than $2\varepsilon^4$. By Markov inequality, $\mathfrak{P}$ is $(\xi, \varepsilon)$ u-saturated. $\qquad\square$

**Lemma 4.3.** *If $A$ is $(\xi, \varepsilon)$ u-saturated and*

$$(4.1) \qquad\qquad \mu_{|A}(P_\tau) \geqslant 1 - \varepsilon^2,$$

*then we can decompose $\mu|_A = \mu_g + \mu_r$ so that $\mu_r(M) \leqslant \varepsilon$ and there is a family $\{\mathcal{W}_t\}_{t \in \mathfrak{T}}$ of unstable boxes of size $\xi$ and a measure $\nu$ of $\mathfrak{T}$ such that*

$$(4.2) \qquad\qquad \mu_g(B) = \int_{\mathfrak{T}} m_{\mathcal{W}_t}^u(B) d\nu(t).$$

*Proof.* Let $\bar{E}_1 = A \setminus P_\tau$ and $\bar{E}_2 = E^c$, and $\bar{E} = \bar{E}_1 \cup \bar{E}_2$ where $E$ is from Definition 4.1 for $A$ and $E^c$ is the complement of $E$. Since $A$ is $(\xi, \varepsilon)$ saturated and (4.1) holds for $A$, we have that $\mu_{|A}(\bar{E}) \leqslant \varepsilon$. It then follows that $A \setminus \bar{E} = \bigcup_{x_i} \bigcup_{z \in T_{x_i} \cap P_\tau} W_{z,\xi}^u \cap P_\tau$.

Now absolute continuity of unstable foliation on $P_\tau \subset \hat{P}_\tau$ (see Lemma 2.13) gives

$$\mu(B) = \mu|_A \left(B \cap \bar{E}\right) + \sum_i \int_{T_{x_i} \cap P_\tau} d\nu_i(z) m^u(W_{z,\xi}^u \cap A \cap P_\tau) \mu_{\mathcal{W}}(B).$$

By the definition of $\bar{E}$, the first term is smaller than $\varepsilon$. This gives (4.2) with $\mathfrak{T} = \bigcup_i \bigcup_{z \in T_{x_i} \cap P_\tau} z$ and $d\nu = \sum_i d\nu_i$. $\qquad\square$

**Lemma 4.4.** *Let $\mathfrak{B}$ be as set with $\mu(\mathfrak{B}) \leqslant \hat{\varepsilon}^4$. Then there is $\xi_0$ such that for $\xi \leqslant \xi_0$ the set*

$$\mathcal{K} = \{x \in M \ : \ r_u(x) \geqslant \xi \text{ and } m_{W_{x,\xi}^u}^u(\mathfrak{B}) \leqslant \hat{\varepsilon}\}$$

*has measure greater than $1 - 4\hat{\varepsilon}$ where $r_u(x)$ is the size of the unstable manifold of $x$.*

*Proof.* Take a small $\bar{\xi}$. Then Lemma 4.3 gives a decomposition $\mu = \mu_g + \mu_r$ where $\mu_r(M) \leqslant \hat{\varepsilon}$ and

$$\mu_g(A) = \int_{\mathfrak{T}} m^u_{\mathcal{W}_t}(A) d\nu(t)$$

where $\mathcal{W}_t$ are unstable boxes of size $\bar{\xi}$. By Markov inequality we may assume, possibly increasing $\mu_r(M)$ by $\hat{\varepsilon}$, that for each $t \in \mathfrak{T}$ we have $m^u_{\mathcal{W}_t}(\mathfrak{B}^c) > 1 - \hat{\varepsilon}^3$. Next take $\xi_0$ so small that $m^u_{\mathcal{W}_t}(\partial_{\xi_0} \mathcal{W}_t) \leqslant \hat{\varepsilon}$. Now for each $t \in \mathfrak{T}$ we apply Lemma C.1 with

$$\nu = m^u_{\mathcal{W}_t}, \quad D = \{x \in \mathcal{W}_t : \ d(x, \partial \mathcal{W}_t) \geqslant \xi\}, \quad B = \mathfrak{B}$$

and conclude that for each $t$ we have

$$\mu^u_{\mathcal{W}_t}(\mathcal{K}^c) \leqslant \mu^u_{\mathcal{W}_t}(\mathcal{K}^c \cap D) + \mu^u_{\mathcal{W}_t}(D^c) \leqslant 2\hat{\varepsilon}.$$

Integrating over $t \in \mathfrak{T}$ and remembering that $\mu_r(M) < 2\hat{\varepsilon}$ we obtain the result. $\qquad\square$

## 5. Construction of fake center-stable foliation.

The results of this section are obtained under the assumption that $f$ is $C^{1+\alpha}$ diffeomorphism of a compact manifold $M$ preserving a smooth measure $\mu$ with some non-zero exponents. The reader may suppose in the arguments below that $\eta_2$ is a small constant and $\varepsilon \ll \eta_2$. Later we apply the results of this section in the case where $\eta_2$ satisfies the conditions of Lemma B.2. Before we go to the details of the construction we provide an outline.

**Outline of the construction of the fake cs-foliations:** The constructed foliation depends on the time parameter $n$. The idea is to choose a regular reference point $x$ and to bring back the foliation $\exp_x(\mathbb{R}^{cs})$ from time $n$ to time $0$. The problem with this approach is that the leaves constructed using two different reference points $x'$ and $x''$ may intersect. To overcome this problem we will create small buffers between the regions where we take different foliations. This will resolve the problem of intersections at the price that our foliation will be defined not on all of $M$ but on a set whose measure could be made arbitrary close to 1.

More precisely, we fix a family $\{\bar{B}_s\} = \{B_s(\xi_n, r)\}$ (see (5.1)) of disjoint parallelograms centered at points in $\mathcal{L}_{n,\tau}$ (see (5.3)) and with sizes $\bar{\xi}_n = e^{-\varepsilon^3 n}$ in the $\mathbb{R}^u$ direction and $r_n = \bar{\xi}_n^{1+\omega}$ (where $\omega$ is sufficiently small) in the $\mathbb{R}^{cs}$ direction. The exact choice of the size of $\bar{\xi}_n$ is not so important as long as it is exponentially small with $n$ (with the exponent much smaller than the exponential mixing exponent). We want the $r_n$ to be smaller than $\xi_n$ to be able to use absolute continuity of the unstable foliation. We also want that $\{\bar{B}_s\}$ cover most of the space (i.e.[4] $1 - 30\varepsilon^{\mathbf{b}/4}$). Existence of such a family is established in Lemma 5.1. We now take $\xi_n := e^{\varepsilon^2 n - \eta_2 \varepsilon n}$ and $\varepsilon' = \varepsilon^{100\mathbf{b}}$. We say that an unstable manifold $\mathcal{W}$ of size $\xi_n$ is *good* if

$$m^u_{\mathcal{W}}\left(\bigcup_s f^{-n}\left((1 - \varepsilon') \cdot \bar{B}_s\right) \cap \mathcal{L}_{n,\tau}\right) \geqslant 1 - 100\varepsilon^{\mathbf{b}/16}.$$

---

[4] Recall that $\mathbf{b} = 10^{10}$.

Using again Lemma 5.1 we show that there exists a family $\{\widehat{\mathcal{W}}_i = W_i(\xi_n)\}$ of good unstable manifolds such that the corresponding parallelograms $\{B_i = B_i(\widehat{\mathcal{W}}_i, r)\}$ with $r \leqslant e^{-\eta_2 \varepsilon n}$ cover most of the space. In fact in the applications in Sections 6 and 7 we will only work with two sizes of $r = e^{-\eta_2 \varepsilon n}$ and $\tilde{r} = e^{-\eta_2 \varepsilon n - \varepsilon^2 n}$. The extra $-\varepsilon^2 n$ term in $\tilde{r}$ is responsible for the fact that if we start with two points on the same fake center-stable leaf which are at distance $\tilde{r}$ then after $\varepsilon \cdot n$ iterates their distance will be less than $r$. This will be used in Proposition 7.1. Next, for each $i$ and $s$ we look at the maximal connected components $\{R_{i,s,j}\}$ of the set $\widehat{\mathcal{W}}_i \cap f^{-n}((1-\varepsilon')\bar{B}_s)$. For each such connected component $R_{i,s,j}$ we construct a fake cs-foliation by pulling the $\mathbb{R}^{cs}$ foliation for the set $f^n(\widehat{\mathcal{W}}_i) \cap \bar{B}_s$ (see (5.12), (5.13)). Here the fact that we hit $\bar{B}_s$ not too close to the boundary will create aforementioned buffers, while the fact that $\varepsilon'$ is small, will ensure that our foliations are defined on a set of a measure close to 1. Also we want $\bar{B}_s$ to be much larger than $B_i$ to ensure that the unstable leaves we construct will fully cross $B_i$. The next step is to glue the foliations for different $s, j$ to get a foliation of (most of) the sets $B_i(\widehat{\mathcal{W}}_i, r)$. The crucial properties of the foliation $\{\mathcal{F}_{i,s,j}\}$ are established in Lemma 5.2. Namely, we show that indeed one can glue the foliations over different $s, j$ as they don't intersect (property **F5.**). We also show that the growth on the leaves of the foliation $\{\mathcal{F}_{i,s,j}\}$ is sub-exponential (property **F4.**). In property **F3** we show that most of the set $\widehat{\mathcal{W}}_i$ intersects a leaf of $\mathcal{F}_{i,s,j}$. The crucial property of the constructed foliation is that it is locally absolutely continuous (on exponentially small scale), see Proposition 6.4. The bounded distortion property (property **F2** of Lemma 5.2) plays a key role in proving Proposition 6.4.

We now move to the details of the construction.

Recall that for $x \in \mathbb{R}^D$, $\|x\|_\infty = \max_i |x_i|$. For $\xi < \tau$, $x \in P_\tau$ let $\mathcal{W}(\xi) \subset W^u_{z,\tau}$ be an unstable cube of $z$ of size $\xi$. Let

$$(5.1) \quad B(\xi, r) = B(\mathcal{W}(\xi), r) := h_z \left( \{(a, b) \in \mathbb{R}^u \times \mathbb{R}^{cs} : \|a\|_\infty \leqslant \xi, \|b - \eta_z(a)\|_\infty \leqslant r \} \right)$$

where the unstable manifold of $z$ is given by the graph of $\eta_z$ in the exponential coordinates (see (2.9)). By Lemma 2.4 there exists a measurable function $\boldsymbol{\rho} : LyapReg \to \mathbb{R}$ and $\alpha_{\boldsymbol{\rho}} > 0$ such that for every $\tau > 0$ the restriction of $\boldsymbol{\rho}$ to $P_\tau$ is Hölder continuous with exponent $\alpha_{\boldsymbol{\rho}}$ (independent of $\tau$) and, by Fubini theorem for each $\hat{\varepsilon}$ and $\tau > 0$ and any $\xi$ and $r$ small enough (in terms of $\hat{\varepsilon}$ and $\tau$) so that

$$\mu(B(\xi, r)) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})\boldsymbol{\rho}(x) \cdot \int_{\|a\|_\infty \leqslant \xi} \mathrm{mes}(B^{cs}(\eta_z(a), r)) da$$

$$(5.2) \qquad = (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})\boldsymbol{\rho}(x) \cdot m^u(\mathcal{W}(\xi))\mathrm{mes}(B^{cs}(0, r)),$$

for arbitrary $x \in B(\xi, r) \cap P_\tau$. In fact it is enough to take $\boldsymbol{\rho}(x) := \det(L_x)$ (see Lemma 2.4 and Lemma 2.3).

Fix $\varepsilon$ and let $\tau$ be as in (2.15). For $n \in \mathbb{N}$ we define

$$(5.3) \qquad\qquad \mathcal{L}_{n,\tau} = P_\tau \cap f^{-n}P_\tau \cap f^{-\varepsilon n}P_\tau.$$

Note that by (2.15), $\mu(\mathcal{L}_{n,\tau}) \geqslant 1 - 3\varepsilon^{\mathbf{b}}$.

**Lemma 5.1.** *There exists $\omega > 0$ and $n_\varepsilon \in \mathbb{N}$ such that for every $n \geqslant n_\varepsilon$ and for any $\xi \in (0, e^{-\varepsilon^3 n})$, $r \in [\xi^{1+\omega}, \xi]$ and for any set $H \subset M$ there exists a finite family of pairwise disjoint sets $\mathbb{B} = \{B_i(\xi, r)\}_{i \in J_r} = \{B(\mathcal{W}_i(\xi), r)\}$ such that for every $i \in J_r$, $\mathcal{W}_i(\xi)$ satisfies*

$$(5.4) \qquad m^u_{\mathcal{W}_i(\xi)}(H) \leqslant \mu(H)^{1/4}.$$

*and*

$$(5.5) \qquad \mu\left(\bigcup_{i \in J_r} B_i(\xi, r)\right) \geqslant 1 - 10\mu(H)^{1/4}.$$

*Proof.* Let $\mathcal{K} = \{x \in P_\tau : m^u_{\mathcal{W}_{x,\xi}}(H) \leqslant \mu(H)^{1/4}\}$. By Lemma 4.4, $\mu(\mathcal{K}) > 1 - 4\mu(H)^{1/4}$. Divide $M$ into cubes $Q_j$ of size $R = \xi^{1-10\omega}$ so that each cube belongs to a single coordinate chart. In each cube which intersects $\mathcal{L}_{n,\tau}$, we choose a point $z_j \in \mathcal{L}_{n,\tau}$. Then divide each cube into bricks $Z_{ij}$ which are products of cubes of size $(1 + \hat{\varepsilon}^3)\xi$ in $E^u(z_j)$ direction and cubes of size $(1 + 2\hat{\varepsilon})r$ in $E^{cs}(z_j)$ direction. The centers $z_{ij}$ of $Z_{ij}$ lie on a lattice. Shifting that lattice if necessary we may assume that the proportion of $z_{ij}$ which is not in $\mathcal{K}$ is at most $2\dfrac{\mu(\mathcal{K}^c \cap Q_i)}{\mu(Q_i)}$. Now for each $z_{ij} \in \mathcal{K}$ we add $B_{ij} = B(W_{z_{ij}}(\xi), r)$ into our collection $\mathbb{B}$. By (2.10) it follows that for $\|a\| \leqslant \xi$ we have $\|\eta_z(a)\| \leqslant C\xi^{1+\alpha_4}$. Hence choosing $\omega < \alpha_4$ ensures that $B_{ij} \subset Z_{ij}$ and $\mu(B_{ij}) > (1 - \hat{\varepsilon}^2)\mu(Z_{ij})$.

Since by construction

$$\mu\left(\bigcup_{ij} Z_{ij}\right) \geqslant 1 - 2\mu(\mathcal{K}^c) \geqslant 1 - 8\mu(H)^{1/4}$$

we obtain (5.5) concluding the proof of the lemma. $\qquad \square$

We apply the above lemma with $\bar{\xi}_n := e^{-\varepsilon^3 n}$, $\bar{r} := \bar{\xi}_n^{1+\omega}$ and $H = \mathcal{L}^c_{n,\tau}$, to get a family $\mathbb{B}_1 = \{\bar{B}_s\} = \{B_s(\bar{\xi}_n, \bar{r})\}_{s \in J_r}$ as in the above lemma. Since $\mu(\mathcal{L}_{n,\tau}) \geqslant 1 - 3\varepsilon^{\mathbf{b}}$, (5.5) implies that

$$\mu(\bigcup_s \bar{B}_s) \geqslant 1 - 10(3\varepsilon^{\mathbf{b}})^{1/4} \geqslant 1 - 30\varepsilon^{\mathbf{b}/4}.$$

We apply Lemma 5.1 again with

$$(5.6) \qquad \xi_n = e^{\varepsilon^2 n - \eta_2 \varepsilon n}, \quad r \leqslant e^{-\eta_2 \varepsilon n},$$

and

$$(5.7) \qquad H' = \left(\mathcal{L}_{n,\tau} \cap \bigcup_s f^{-n}((1 - \varepsilon') \cdot \bar{B}_s)\right)^c,$$

where $\varepsilon' = \varepsilon^{100\mathbf{b}}$ to get a family $\mathbb{B}_2(r) = \{B_i(\xi_n, r)\}_{i \in J_r} = \{B(\mathcal{W}_i(\xi_n), r)\}$. Notice that $\mu(H') \leqslant 40\varepsilon^{\mathbf{b}/4}$ and so by the above lemma,

$$(5.8) \qquad \mu\left(\bigcup_{i \in J_r} B_i(\xi_n, r)\right) \geqslant 1 - 10 \cdot (40\varepsilon^{\mathbf{b}/4})^{1/4} \geqslant 1 - 100\varepsilon^{\mathbf{b}/16}.$$

Moreover, by (5.4) it follows that for every $i \in J_r$,

$$(5.9) \qquad m^u_{W_i(\xi_n)}\Big(\mathcal{L}_{n,\tau} \cap \bigcup_s f^{-n}\big((1-\varepsilon')\bar{B}_s\big)\Big) \geqslant 1 - \mu(H')^{1/4} \geqslant 1 - 100\varepsilon^{\mathbf{b}/16}.$$

We will now construct the fake center stable partitions of elements in $\mathbb{B}_2(r)$ by pulling back the $\mathbb{R}^{cs}$ partition from element in $\mathbb{B}_1$ (with time $n$). Since $\xi_n = e^{\varepsilon^2 n - \eta_2 \varepsilon n}$ is now fixed throughout the paper we denote

$$(5.10) \qquad \widehat{\mathcal{W}}_i = \mathcal{W}_i(\xi_n)$$

$$(5.11) \qquad B_i(r) = B_i(\xi_n, r)$$

where $r \leqslant e^{-\eta_2 \varepsilon n}$. For any $s$ consider the maximal connected components $\{R_{i,s,j}\}_j$ of the set $\widehat{\mathcal{W}}_i \cap f^{-n}\big((1-\varepsilon')\bar{B}_s\big)$. Notice that by (5.9) the union of $R_{i,s,j}$ (over $s, j$) for which $R_{i,s,j} \cap \mathcal{L}_{n,\tau} = \emptyset$ has measure $\leqslant 100\varepsilon^{\mathbf{b}/16}$ and so from now on we will restrict our attention to those $R_{i,s,j}$ for which $R_{i,s,j} \cap \mathcal{L}_{n,\tau} \neq \emptyset$. Let $z = z_{i,s,j} \in R_{i,s,j} \cap \mathcal{L}_{n,\tau}$ and let for $y \in R_{i,s,j}$

$$(5.12) \qquad \tilde{W}^{cs,n}_{i,s,j}(y) = (\tilde{f}^{(n)}_z)^{-1}(\mathbb{R}^{cs} + \tilde{f}^{(n)}_z(h_z^{-1} y))$$

where $\mathbb{R}^{cs} + y$ is the parallel to $\mathbb{R}^{cs}$ through $y$ and let

$$(5.13) \qquad \mathcal{F}_{i,s,j}(y) := W^{cs,n}_{i,s,j}(y) = h_z(\tilde{W}^{cs,n}_{i,s,j}(y)).$$

The following lemma summarizes the properties of the foliations $\{\mathcal{F}_{i,s,j}\}$ that will be important for our purposes.

**Lemma 5.2.** *For any $\hat{\varepsilon} > 0$ there exists $n_{\hat{\varepsilon}}$ such that for $n \geqslant n_{\hat{\varepsilon}}$, we have the following:*

  **F1.** *for any $i, s, j$ and any $k \leqslant n$,*
$$f^k(R_{i,s,j}) \subset h_{f^k z_{i,s,j}}(B^u(e^{\lambda(-n+k)} e^{-\varepsilon^5 n}));$$

  **F2.** *for any $i, s, j$ and for any $z, z' \in R_{i,s,j}$, we have*
$$(5.14) \qquad \left| \frac{\det(Df^n_z | E^u(z))}{\det(Df^n_{z'} | E^u(z'))} \right| \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon});$$

  **F3.** *for any $i$, $m^u_{\widehat{\mathcal{W}}_i}(\bigcup_{s,j} R_{i,s,j} \cap \mathcal{L}_{n,\tau}) \geqslant 1 - 100\varepsilon^{\mathbf{b}/16}$;*

  **F4.** *for any $i, s, j$, any $y' \in \mathcal{F}_{i,s,j}(y) \cap B_i(r)$ and any $0 \leqslant k \leqslant n$*
$$d(f^k y, f^k y') \leqslant e^{\varepsilon^{10} n} \cdot r.$$

  **F5.** *for any $i$, if*
$$B_i(r) \cap \mathcal{F}_{i,s,j}(y) \cap \mathcal{F}_{i,s',j'}(y') \neq \emptyset,$$
  *for some $y \in R_{i,s,j}, y' \in R_{i,s',j'}$, then $s = s'$, $j = j'$ and $y = y'$.*

*Proof.* We start by showing **F1**. By definition, $f^n(R_{i,s,j}) \subset (1-\hat{\varepsilon})\bar{B}_s$ and it is a connected component of $f^n(\widehat{\mathcal{W}}_i) \cap (1-\hat{\varepsilon})\bar{B}_s$. Moreover $z_n := f^n(z_{i,s,j}) \in f^n(R_{i,s,j}) \cap f^n(\mathbb{L}_{n,\tau})$. In particular, $z_n \in P_\tau \subset \tilde{P}_{\tau-1}$. Therefore, $f^n(R_{i,s,j}) \subset W^u_{z_n,\tau}$ and in particular $f^n(R_{i,s,j})$ its backward iterations satisfy (2.13). Since $f^n(R_{i,s,j}) \subset (1-\hat{\varepsilon})\bar{B}_s$ and $f^n(R_{i,s,j}) \cap P_\tau \neq \emptyset$

by construction, the Hölder continuity of $z \mapsto \eta_z$ on $P_\tau$ (see Lemma A.8) implies that that the unstable size of $f^n(R_{i,s,j})$ is at most

$$2(1+\hat{\varepsilon})\bar{\xi}_n = 2(1+\hat{\varepsilon})e^{-\varepsilon^3 n}.$$

This gives **F1.** by using $f^k(R_{i,s,j}) = f^{k-n}(f^n(R_{i,s,j}))$ and applying (2.13).

For **F2.** notice that

$$|\log \det(Df_z^n|E^u(z)) - \log \det(Df_{z'}^n|E^u(z'))| \leqslant$$

$$C_1 \sum_{k=0}^{n} \left[ d(f^k z, f^k z') + d(E^u(f^k z), E^u(f^k z')) \right] \leqslant C_2 n e^{-\varepsilon^5 n},$$

where in the last inequality we use **F1** to estimate the first term and Lemma A.10 with $L_1 = E^u(z), L_2 = E^u(z')$ to estimate the second term.

**F3** is immediate from (5.9).

We now show **F4.** In the coordinates $(a, b)$ introduced in (5.1), $\mathcal{F}_{i,s,j}(y)$ is given by the graph $a = A(b)$. Let $y = (A(b_0), b_0)$, $y' = (A(b_1), b_1)$. Consider a curve $\gamma(t) = \{h_z(A(b(t)), b(t))\}$ where $b(t) = b_0(1-t) + b_1(t)$. Recall that each component $R_{i,j,s}$ contains a point $z = z_{i,j,s} \in P_\tau$. Hence combining Lemma 2.6 with already established property **F1.** we conclude that[5] $\dot{\gamma}(t) \in \mathcal{C}_z^s$ for all $t$. Hence Lemma 2.10(2) tells us that $f^k$ increases distances by at most the factor $\tau^{-2}e^{2k\delta}$ proving **F4.** (since $\delta < \varepsilon^{100}$).

It remains to show **F5**. Assume first that $s = s'$. Then $y = y'$ since otherwise $\mathcal{F}(y)$ intersects $\widehat{\mathcal{W}}$ in two points, which is impossible by transversality. Namely if two points $x_1, x_2$ belong to $\mathcal{F}(y)$ then the segment (in $(a, b)$ coordinates introduced in (5.1)) joining $x_1$ and $x_2$ belongs to $\mathcal{C}_y^{cs}$. If the two points belong $\widehat{\mathcal{W}}$ then the segment joining them belongs to $\mathcal{C}^u$. Since $\mathcal{C}_y^u \cap \mathcal{C}_y^{cs} = \{0\}$, $\mathcal{F}(y) \cap \widehat{\mathcal{W}}$ is a single point. Since $R_{i,s,j} \cap R_{i,s,j'}$ contains $y$ we must have $j = j'$ completing the proof in the case $s = s'$.

Assume now that $s \neq s'$. We claim that for any $s$,

$$(5.15) \qquad\qquad B_i(r) \cap \mathcal{F}_{i,s,j}(y) \subset f^{-n}(\bar{B}_s)$$

The above claim immediately gives **F5.** as the sets $\{\bar{B}_s\}$ are pairwise disjoint. So it remains to show (5.15). By definition we have that $f^n(R_{i,s,j}) \subset (1-\hat{\varepsilon})\bar{B}_s$. In particular, $d(f^n y, \partial \bar{B}_s) \geqslant \hat{\varepsilon} \min(\bar{\xi}_n, \bar{r}) = \hat{\varepsilon} \cdot e^{-\varepsilon^3 n(1+\omega)}$. On the other hand by **F4** it follows that for every $y' \in B_i(r) \cap \mathcal{F}_{i,s,j}(y)$, $d(f^n y, f^n y') \leqslant e^{\varepsilon^{100} n} r \leqslant e^{\varepsilon^{100} n - \eta_2 \varepsilon n}$. Since for $n$ large enough, $e^{\varepsilon^{100} n - \eta_2 \varepsilon n} \ll \hat{\varepsilon} \cdot e^{-\varepsilon^3 n(1+\omega)}$, $f^n y' \in \bar{B}_s$. This finishes the proof. $\qquad\square$

---

[5]Lemma 2.6 tells us that the fake center stable leaves have tangent spaces in $\mathcal{C}_z^{cs}$, while **F1.** ensures that $\mathcal{F}(y)$ belongs to some fake center-stable leaf.

By enumerating (replacing the indices $(s, j)$ by $j$), we will denote the sets $\{R_{i,s,j}\}$ and the foliations $\mathcal{F}_{i,s,j}$ simply by $\{R_{i,j}\}$ and $\{\mathcal{F}_{i,j}\}$. Define

$$(5.16) \qquad \tilde{R}_{i,j}(r) := \bigcup_{y \in R_{i,j}} \mathcal{F}_{i,j}(y) \cap B_i(r).$$

## 6. Absolute continuity of fake center foliation.

In this section we continue assuming that $f$ is $C^{1+\alpha}$ diffeomorphism of a compact manifold $M$ preserving a smooth measure $\mu$ with some non-zero exponents. The goal of the section is to establish an absolute continuity at exponential scale of the foliations $\{\mathcal{F}_{i,j}\}$ which plays a crucial role in the proof of the main theorem. Let $\pi_{i,j}$ denote the holonomy along the foliation $\mathcal{F}_{i,j}$. Recall that $\widehat{\mathcal{W}}_i$ is the reference unstable manifold in $B_i(r)$ and $R_{i,j} \subset \widehat{\mathcal{W}}_i$.

**Definition 6.1.** *We say that an unstable manifold $\mathcal{W}$ of a point $z \in \mathcal{L}_{n,\tau}$, crosses $\tilde{R}_{i,j}$ completely if $\mathcal{W} \cap \pi_{i,j}(y) \neq \emptyset$ for every $y \in R_{i,j}$. $\mathcal{W}$ crosses $B_i(r)$ completely if for every $j$ it crosses $\tilde{R}_{i,j}$ completely.*

**Definition 6.2.** *Given $x$ a Lyapunov regular point, we say that $\mathcal{L}$ is an $(\beta, A)$-admissible manifold if*

*(1) $\mathcal{L} \subset Q_x$;*
*(2) $h_x^{-1}\mathcal{L} \subset \text{graph}(\eta_{\mathcal{L}})$, where $\eta_{\mathcal{L}} : \mathbb{R}^u \to \mathbb{R}^{cs}$ satisfies $\|\eta_{\mathcal{L}}\|_{C^{1+\beta}} \leqslant A$.*

**Remark 6.3.** *Note that $(\beta, A)$-admissibility depends on the reference point $x$. However, as long as the reference point is in $P_\tau$ we have a uniform control on the parameter $A$ when we change of charts. That is, there is a constant $K(\tau)$ so that if $x, y \in P_\tau$, $y \in Q_x$ and $\mathcal{L}$ is $(\beta, A)$-admissible for $x$ then it is $(\beta, K(\tau) \cdot A)$−admissible for $y$.*

Let $\beta > 0$ be a number that is small enough, one can take $\beta := \frac{1}{2}\left[\min\left(\alpha, \min_{i \leqslant 7} \alpha_i\right)\right]$, see Lemmas 2.4, 2.6, 2.8, 2.3, 2.11, A.10 for the definitions of the $\alpha_i$.

**Proposition 6.4.** *For every $\hat{\varepsilon} > 0$ there exists $n_{\hat{\varepsilon}}$ such that for every $n \geqslant n_{\hat{\varepsilon}}$, every $r \leqslant e^{-\eta_2 \varepsilon n}$ the following holds: let $x \in \mathcal{L}_{n,\tau}$ and let $\mathcal{L}$ be an $(\beta, 1)$-admissible manifold. Assume further that*

*(1) $\mathcal{L}$ crosses $\tilde{R}_{ij}(r)$ completely;*
*(2) there are points $x \in \hat{\mathcal{W}}_i$ and $x' \in \mathcal{L}$ such that if $\bar{x}' = h_x^{-1}(x')$, then $|\bar{x}'| \leqslant e^{-\eta_2 \varepsilon n}$ and $\|D_0 \eta_{\mathcal{L}}\| \leqslant e^{-\beta \eta_2 n}$.*

*Then the jacobian $J(\pi_{i,j})$ satisfies*

$$(6.1) \qquad J(\pi_{i,j})(z) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon}) \text{ for each } z \in R_{i,j}.$$

*In particular,*

$$(6.2) \qquad m_{\mathcal{W}}^u(\pi_{i,j}(R_{i,j})) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})m_{\hat{\mathcal{W}}_i}^u(R_{i,j}).$$

*Proof.* We will drop the index $i$ in the proof. Let $\bar{\mathcal{L}} = h_x^{-1}(\mathcal{L})$, $\overline{\hat{\mathcal{W}}} = h_x^{-1}(\hat{\mathcal{W}})$. We let $\bar{\pi}_j : \overline{\hat{\mathcal{W}}} \to \bar{\mathcal{L}}$ be $\bar{\pi}_j = h_x^{-1} \circ \pi_j \circ h_x$. For $\bar{z} \in \overline{\hat{\mathcal{W}}}$ let $\bar{z}' = \bar{\pi}_j(\bar{z}) \in \bar{\mathcal{L}}$. Set $\bar{\pi}_j^{(n)} : \tilde{f}_x^{(n)}\left(\overline{\hat{\mathcal{W}}}\right) \to$

$\tilde{f}_x^{(n)}(\bar{\mathcal{L}})$, the holonomy along the $\mathbb{R}^{cs}$ spaces. Then $\bar{\pi}_j = (\tilde{f}_x^{(n)})^{-1} \circ \bar{\pi}_j^{(n)} \circ \tilde{f}_x^{(n)}$. For a linear map $T : E \to F$ where $E, F \subset \mathbb{R}^D$, $\dim E = \dim F$, we denote $\bar{J}(T) = |\det(T)|$ where determinant is taken with respect to natural volume coming from euclidean distance in $\mathbb{R}^D$. We denote as well $\bar{J}(\bar{\pi}_j)(\bar{z}) = \bar{J}(D_{\bar{z}}\bar{\pi}_j)$.

So we have that

$$\bar{J}(\bar{\pi}_j)(\bar{z}) = \frac{\bar{J}(D_{\bar{z}}\tilde{f}_x^{(n)}|T_{\bar{z}}\overline{\widehat{\mathcal{W}}})}{\bar{J}(D_{\bar{z}'}\tilde{f}_x^{(n)}|T_{\bar{z}'}\bar{\mathcal{L}})} \bar{J}(\bar{\pi}_j^{(n)})(f_x^{(n)}(\bar{z})).$$

Thus the proposition is a consequence of the following two estimates:

$$(6.3) \qquad \forall z, z' \in \tilde{R}_{i,j} \quad \frac{\bar{J}(D_{\bar{z}}\tilde{f}_x^{(n)}|T_{\bar{z}}\overline{\widehat{\mathcal{W}}})}{\bar{J}(D_{\bar{z}'}\tilde{f}_x^{(n)}|T_{\bar{z}'}\bar{\mathcal{L}})} \in (1 - \hat{\varepsilon}^4, 1 + \hat{\varepsilon}^4);$$

$$(6.4) \qquad \forall z^* \in f^n(R_{i,j}) \quad \bar{J}(\bar{\pi}_j^{(n)})(\bar{z}^*) \in (1 - \hat{\varepsilon}^4, 1 + \hat{\varepsilon}^4)$$

We have

$$\frac{\bar{J}(D_{\bar{z}}\tilde{f}_x^{(n)}|T_{\bar{z}}\overline{\widehat{\mathcal{W}}})}{\bar{J}(D_{\bar{z}'}\tilde{f}_x^{(n)}|T_{\bar{z}'}\bar{\mathcal{L}})} = \prod_{k=0}^{n-1} \frac{\bar{J}\left(D_{\tilde{f}_x^{(k)}(\bar{z})}\tilde{f}_{f^k x}|T_{\tilde{f}_x^{(k)}(\bar{z})}\tilde{f}_x^{(k)}(\overline{\widehat{\mathcal{W}}})\right)}{\bar{J}\left(D_{\tilde{f}_x^{(k)}(\bar{z}')}\tilde{f}_{f^k x}|T_{\tilde{f}_x^{(k)}(\bar{z}')}\tilde{f}_x^{(k)}(\bar{\mathcal{L}})\right)}.$$

Taking logarithms we have to estimate:

$$\sum_{k=0}^{n-1} \left| \log \bar{J}\left(D_{\tilde{f}_x^{(k)}(\bar{z})}\tilde{f}_{f^k x}|T_{\tilde{f}_x^{(k)}(\bar{z})}\tilde{f}_x^{(k)}(\overline{\widehat{\mathcal{W}}})\right) - \log \bar{J}\left(D_{\tilde{f}_x^{(k)}(\bar{z}')}\tilde{f}_{f^k x}|T_{\tilde{f}_x^{(k)}(\bar{z}')}\tilde{f}_x^{(k)}(\bar{\mathcal{L}})\right) \right|.$$

By Lemma 2.4 we have that that for every $x \in LyapReg$, $Hol_{\alpha_2}(D\tilde{f}_x) \leqslant \delta$ and by Lemma 2.9, $\|D\tilde{f}_x\|, \|D\tilde{f}_x^{-1}\| \leqslant K'$. We can hence bound the above sum using Lemma 2.7 by

$$N(K') \sum_{k=0}^{n-1} \delta|\tilde{f}_x^{(k)}(\bar{z}) - \tilde{f}_x^{(k)}(\bar{z}')|^{\alpha_2} + d\left(T_{\tilde{f}_x^{(k)}(\bar{z})}\tilde{f}_x^{(k)}(\overline{\widehat{\mathcal{W}}}), T_{\tilde{f}_x^{(k)}(\bar{z}')}\tilde{f}_x^{(k)}(\bar{\mathcal{L}})\right)$$

By assumption (2) in the proposition, the fact that $\mathcal{L}$ is $(\beta, 1)$ admissible and Lemma A.8, we get using the transversality of $\mathcal{F}_{ij}$ to $\mathcal{L}$ and $\widehat{\mathcal{W}}$ that

$$|\bar{z} - \bar{z}'| \leqslant C e^{-\eta_2 \varepsilon n} \quad and \quad d\left(T_{\bar{z}}\overline{\widehat{\mathcal{W}}}, T_{\bar{z}'}\bar{\mathcal{L}}\right) \leqslant e^{-\beta\eta_2 n}.$$

Since $\bar{z}' \in W_x^{cs,n}(\bar{z})$, we get by Corollary A.7 that $|\tilde{f}_x^{(k)}(\bar{z}) - \tilde{f}_x^{(k)}(\bar{z}')| \leqslant e^{3k\delta}|\bar{z} - \bar{z}'|$ for every $k \in [0, n]$ and hence $|\tilde{f}_x^{(k)}(\bar{z}) - \tilde{f}_x^{(k)}(\bar{z}')| \leqslant C e^{-\eta_2 \varepsilon n} e^{3k\delta}$. We now apply Lemma A.10 with $L_1 = T\mathcal{L}$ and $L_2 = T\widehat{\mathcal{W}}$. Using Lemma A.8, the admissibility of $\mathcal{L}$ and the assumption $\|D_0 \eta_{\mathcal{L}}\| \leqslant e^{-\beta\eta_2 n}$ we conclude from (A.9) that

$$d\left(T_{\tilde{f}_x^{(k)}(\bar{z})}\tilde{f}_x^{(k)}(\overline{\widehat{\mathcal{W}}}), T_{\tilde{f}_x^{(k)}(\bar{z}')}\tilde{f}_x^{(k)}(\bar{\mathcal{L}})\right) \leqslant e^{k(-\lambda + \sqrt{\delta})}e^{-\beta\eta_2 n} + 6\delta C e^{-\alpha_7\eta_2\varepsilon n}e^{3k\alpha_7\delta}$$

proving (6.3) if $n$ is sufficiently large in terms of $C, \delta$ and $\hat{\varepsilon}$.

To prove (6.4) we notice that $\bar{\pi}_j^{(n)}$ is just holonomy along planes parallel to $\mathbb{R}^{cs}$ so there is a constant $C_3$ only depending on dimension so that

$$|\bar{J}(\bar{\pi}_j^{(n)})(\bar{z}^*)| \leqslant C_3 d\left(T_{\bar{z}^*}\tilde{f}_x^{(n)}(\widehat{\mathcal{W}}), T_{\bar{\pi}_j^{(n)}(\bar{z}^*)}\tilde{f}_x^{(n)}(\hat{\mathcal{L}})\right) \leqslant C_3\left(e^{n(-\lambda+4\delta-\beta\eta_2)} + 6\delta C e^{n(\alpha_7\delta-\alpha_7\eta_2\varepsilon)}\right).$$

Since $\delta = \varepsilon^{100}$, we get the result. $\qquad\square$

**Corollary 6.5.** *If for some $x \in \mathcal{L}_{n,\tau}$ and $r \leqslant e^{-\eta_2\varepsilon n}$, $\mathcal{W} = \mathcal{W}_{x,\xi}$ crosses $\tilde{R}_{i,j}(r)$ completely and $d(\mathcal{W}, \widehat{\mathcal{W}}_i) \leqslant e^{-\eta_2\varepsilon n}$ then (6.1) and (6.2) are satisfied.*

*Proof.* We need to verify the conditions of Proposition 6.4 with $\mathcal{L} = \mathcal{W}$. Note that (1) and (2) follow from the assumptions of the corollary and the fact that $\|D_0\eta_{\mathcal{L}}\| < e^{-\alpha_1\eta_2 n}$ follows from Lemma 2.3. Moreover, $\mathcal{L} = \mathcal{W}$ satisfies the assumptions of Definition 6.2 by Lemma A.8 (since $\beta < \alpha_4$). $\qquad\square$

Proposition 6.4 allows to establish the local product structure on the sets $\tilde{R}_{i,j}(r)$. Let $\boldsymbol{\rho}$ be the function defined in (5.2).

**Corollary 6.6.** *For every $\hat{\varepsilon}, \tau > 0$ there exists $n_{\hat{\varepsilon},\tau} \in \mathbb{N}$ such that for every $n \geqslant n_{\hat{\varepsilon},\tau}$ every $r \leqslant e^{-\eta_2\varepsilon n}$, we have*

$$\mu(\tilde{R}_{i,j}(r))) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})\boldsymbol{\rho}(x)m^u(R_{i,j}) \cdot \mathrm{mes}\left(B^{cs}(0, r)\right)$$

*where $x$ is an arbitrary point in $\tilde{R}_{i,j}(r) \cap P_\tau$.*

*Proof.* Recall that

$$\tilde{R}_{i,j}(r) = \bigcup_{y \in R_{i,j}} \left(\mathcal{F}_{i,j}(y) \cap B_i(r)\right).$$

In particular, we have coordinates $(a, b)$ on $B_i(r)$ given by (5.1). Let

$$\mathfrak{W}^{\mathfrak{b}} = \{x \in B_i(r) : b(x) = \eta_z(a(x)) + \mathfrak{b}\}.$$

Using uniform smoothness of $\widehat{\mathcal{W}}_i$ on $P_\tau$, the fact that $T\widehat{\mathcal{W}}_i(0) = \mathbb{R}^u \times \{0\}$ and the Hölder continuity of $\boldsymbol{\rho}(\cdot)$ on $P_\tau$, we conclude that the ratio of $\mu|_{B_i}$ to the the measure $\hat{\mu}$ given in our coordinates by $d\hat{\mu} = \boldsymbol{\rho}(x)dadb$ is between $1 - \hat{\varepsilon}^2$ and $1 + \hat{\varepsilon}^2$. Therefore

$$\mu(\tilde{R}_{i,j}(r))) \in \boldsymbol{\rho}(x)\left(1 - \hat{\varepsilon}^2, 1 + \hat{\varepsilon}^2\right)\int_{B_r^{cs}(0)} \mathrm{mes}(\mathfrak{R}^{\mathfrak{b}})d\mathfrak{b}$$

where $\mathfrak{R}^{\mathfrak{b}}$ is the image of $R_{i,j}$ under the $\mathcal{F}_{i,j}$–holonomy inside $\mathfrak{W}^{\mathfrak{b}}$. Since $R_{i,j} \cap \mathcal{L}_{n,\tau} \neq \emptyset$ (because we only consider the rectangles satisfying this condition), Proposition 6.4 applied to $\mathcal{L} = \mathfrak{W}^{\mathfrak{b}}$ (the assumptions of the proposition are verified since $\mathfrak{W}^{\mathfrak{b}}$ are just translations of $\widehat{\mathcal{W}}$ in the appropriate coordiantes) tells us that

$$\mathrm{mes}(\mathfrak{R}^{\mathfrak{b}}) \in \left(1 - \hat{\varepsilon}^2, 1 + \hat{\varepsilon}^2\right)m^u(R_{i,j})$$

for $n$ large enough and the result follows. $\qquad\square$

The above corollary implies also the following:

**Lemma 6.7.** *There exists $n_\varepsilon \in \mathbb{N}$ such that for every $n \geqslant n_\varepsilon$, for every $i$ we have for $r \leqslant e^{-\eta_2 \varepsilon n}$*

$$(6.5) \qquad \mu\left(\bigcup_j \tilde{R}_{i,j}(r)\right) \geqslant (1 - 200\varepsilon^{\mathbf{b}/16})\mu(B_i(r)).$$

*In particular,*

$$\mu\left(\bigcup_{i,j} \tilde{R}_{i,j}(r)\right) \geqslant 1 - 300\varepsilon^{\mathbf{b}/16}.$$

*Proof.* The first part is an immediate consequence of Corollary 6.6 and (5.2) (with $\hat{\varepsilon} = \varepsilon^{100b}$) by summing over $j$ and using **F3**. The second part follows from the first and (5.8). $\qquad\square$

Recall that for each $i$ we have the reference manifold $\widehat{\mathcal{W}}_i \subset B_i(r)$ of size $\xi_n = e^{\varepsilon^2 n - \eta_2 \varepsilon n}$.

**Proposition 6.8.** *For every $\hat{\varepsilon} > 0$ there exists $n_{\hat{\varepsilon}, \varepsilon} \in \mathbb{N}$ such that for every $n \geqslant n_{\hat{\varepsilon}, \varepsilon}$ the following holds. Let $z \in \mathcal{L}_{n,\tau}$. Suppose that $W^u_{z,\tau}$ crosses $B_i(r)$ completely and let $\mathcal{W}_z \subset B_i(r)$, $\mathcal{W}_z \subset W^u_{z,\tau}$ be a piece of the unstable manifold containing $z$ of size $\xi \geqslant (1 - \hat{\varepsilon}^2)\xi_n$. Let $\pi_{i,z} : \widehat{\mathcal{W}}_i \to \mathcal{W}_z$ be given by $\pi_z(y) := \pi_{i,j}(y)$ if $y \in R_{i,j}$. Then $\pi_z$ is $\hat{\varepsilon}$–measure preserving. In particular,*

$$m^u_{\mathcal{W}_z}\left(\pi_z(\bigcup_j R_{i,j})\right) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})m^u_{\widehat{\mathcal{W}}_i}(\bigcup_j R_{i,j})$$

*Proof.* Note that by Corollary 6.5 we have (since $\xi \geqslant (1 - \hat{\varepsilon}^2)\xi_n$)
$$m^u_{\mathcal{W}_z}\left(\pi_z(\bigcup_j R_{i,j})\right) = \sum_j m^u_{\mathcal{W}_z}(\pi_{i,j}(R_{i,j})) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})\sum_j m^u_{\widehat{\mathcal{W}}}(R_{i,j}). \qquad\square$$

## 7. Equidistribution of unstable leaves.

7.1. **Equidistribution criterion.** Let $O_z(r)$ denote a ball of radius $r > 0$ centered at $z \in M$. Denote

$$(7.1) \qquad\qquad \tilde{r}_n = e^{-\eta_2 \varepsilon n - \varepsilon^2 n}.$$

Recall Definition 6.1 and (5.11).

**Proposition 7.1.** *For every $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that for every $n \geqslant n_\varepsilon$ the following holds. Fix any*

$$B \in \left\{B_i(e^{-\eta_2 \varepsilon n})\right\} \cup \left\{(1 - \varepsilon^{4000})B_i(e^{-\eta_2 \varepsilon n})\right\} \cup \left\{O_z(r)\right\}_{z \in M},$$

*with $1 > r \geqslant e^{-\eta_2 \varepsilon n}$ and let $B_s \in \{B_i(\tilde{r}_n)\}$. Assume that $\mathcal{W} = \mathcal{W}^u_{x,\xi}$, with $e^{-\varepsilon^3 n} \leqslant \xi \leqslant \tau$ and $x \in \mathcal{L}_{n,\tau}$ crosses $B_s$ completely. Then*

$$m^u_{\mathcal{W}}\left(\left(\bigcup_j \tilde{R}_{s,j}(\tilde{r}_n)\right) \cap f^{-\varepsilon n}(B)\right) \leqslant (1 + \varepsilon^{20})\mu(B)m^u_{\mathcal{W}}\left(\bigcup_j \tilde{R}_{s,j}(\tilde{r}_n)\right).$$

*If moreover, for some $\zeta > 0$*

$$(7.2) \qquad \mu\left(\left(\bigcup_j \tilde{R}_{s,j}(\tilde{r}_n)\right) \cap f^{-\varepsilon n}((1-\zeta)B))\right) \geqslant (1 - 2\varepsilon^{40})\mu(B_s)\mu((1-\zeta)B),$$

*then we have*

$$m_{\mathcal{W}}^u\left(\left(\bigcup_j \tilde{R}_{s,j}(\tilde{r}_n)\right) \cap f^{-\varepsilon n}(B)\right) \geqslant (1 - \varepsilon^{20})\mu((1-\zeta)B)m_{\mathcal{W}}^u\left(\bigcup_j \tilde{R}_{s,j}(\tilde{r}_n)\right).$$

*Proof.* Denote $\tilde{B} = B_s$ and let $C_j = \tilde{R}_{s,j}(\tilde{r}_n) \cap \mathcal{W}$, $\tilde{C}_j = \tilde{R}_{s,j}(\tilde{r}_n)$, $\mathcal{C} = \bigcup_j \tilde{C}_j$.

We will first prove the upper bound. Call an index $j$ *nice* if $C_j \cap f^{-\varepsilon n}(B) \neq \emptyset$. Let $A^+ = \bigcup_{j \text{ nice}} C_j$. Since

$$m_{\mathcal{W}}^u\left(\mathcal{C} \cap f^{-\varepsilon n}B\right) \leqslant m_{\mathcal{W}}^u(A^+)$$

It is enough to show that

$$(7.3) \qquad m^u(A^+) \leqslant (1 + \varepsilon^{1000})m^u\left(\mathcal{W} \cap \mathcal{C}\right)\mu(B).$$

Let $\kappa = \kappa(\varepsilon)$ be small enough so that, we have

$$(7.4) \qquad \mu((1+\kappa)\cdot B) \leqslant (1 + \varepsilon^{2000})\mu(B) \quad \text{and} \quad \mu((1-\kappa)\cdot B) \geqslant (1 - \varepsilon^{2000})\mu(B).$$

We claim that

$$(7.5) \qquad \mathcal{C} \cap f^{-\varepsilon n}((1+\kappa)B) \supset \bigcup_{j \text{ nice}} \tilde{C}_j.$$

Indeed, let $z \in \tilde{C}_j$ with $j$ nice and let $z' = \pi_{s,j}(z) \in C_j$ and let $z'' \in C_j \cap f^{-\varepsilon n}(B)$. Then by **F4** and **F1** in Lemma 5.2 it follows that

$$d(f^{\varepsilon n}z'', f^{\varepsilon n}z) < d(f^{\varepsilon n}z'', f^{\varepsilon n}z') + d(f^{\varepsilon n}z', f^{\varepsilon n}z) \leqslant 2e^{\varepsilon^3 n} \cdot e^{-\eta_2 \varepsilon n - \varepsilon^2 n}.$$

By definition $f^{\varepsilon n}z'' \in B$. We claim that $f^{\varepsilon n}z \in (1+\varepsilon^{1000})B$ if $n$ is large enough. Indeed, if $B = O_z(r)$ with $r \geqslant e^{-\eta_2 \varepsilon n}$ is a ball, then the claim follows by triangle inequality as $2e^{-\eta_2 \varepsilon n - \varepsilon^2 n + \varepsilon^3 n} + r \leqslant (1+\kappa)r$. If $B = B_i(e^{-\eta_2 \varepsilon n})$ (or $B = (1 - \varepsilon^{4000}) \cdot B_i(e^{-\eta_2 \varepsilon n})$), then the claim again follows by triangle inequality, the definition of $B_i(\cdot, \cdot)$ (see (5.1)) and by again using that $2e^{\varepsilon^3 n} \cdot e^{-\eta_2 \varepsilon n - \varepsilon^2 n}$ is much less than $\kappa e^{-\eta_2 \varepsilon n}$ for $n$ large enough.

Since $z \in \tilde{C}_j$ is arbitrary, (7.5) follows.

By exponential mixing and (7.4), (7.5)

$$(7.6) \qquad \mu\left(\bigcup_{j \text{ nice}} \tilde{C}_j\right) \leqslant \mu\left(\mathcal{C} \cap f^{-\varepsilon n}((1+\kappa)B)\right) \leqslant$$

$$\mu(\tilde{B} \cap f^{-\varepsilon n}((1+\kappa)B)) \leqslant (1 + \varepsilon^{2000})\mu(\tilde{B})\mu((1+\kappa)B) \leqslant (1 + 3\varepsilon^{2000})\mu(\tilde{B})\mu(B).$$

By Corollary 6.6 it follows that if $\boldsymbol{\rho}_{\tilde{B}} = \boldsymbol{\rho}(x)$, for an arbitrary $x \in \tilde{B} \cap P_\tau$, then

$$(7.7) \qquad (1 - \varepsilon^{2000})\boldsymbol{\rho}_{\tilde{B}} \cdot m^u(A^+)\mathrm{mes}\left(B^{cs}(0, e^{-\eta_2 \varepsilon n - \varepsilon^2 n})\right) \leqslant \mu\left(\bigcup_{j \text{ nice}} \tilde{C}_j\right).$$

Also by Lemma 6.7 and Corollary 6.6

$$(7.8) \quad \mu(\tilde{B}) \leqslant (1 + \varepsilon^{2000})\mu(\mathcal{C}) \leqslant (1 + 4\varepsilon^{2000})\boldsymbol{\rho}_{\tilde{B}} \cdot m^u\left(\mathcal{W} \cap \mathcal{C}\right) \cdot \mathrm{mes}\left(B^{cs}(0, e^{-\eta_2\varepsilon n - \varepsilon^2 n})\right).$$

Combining (7.6)–(7.8) gives (7.3) finishing the proof of the upper bound.

We will now show the lower bound assuming additionally that (7.2) holds.

An index $j$ is called *good* if $C_j \subset f^{-\varepsilon n}(B)$. Let $A^- = \bigcup_{j \text{ good}} C_j$. Then by definition,

$$m^u_{\mathcal{W}}\left(f^{-\varepsilon n}B \cap \mathcal{C}\right) \geqslant m^u_{\mathcal{W}}(A^-).$$

It is therefore enough to show that

$$(7.9) \qquad\qquad m^u(A^-) \geqslant (1 - \varepsilon^{20})m^u\left(\mathcal{W} \cap \mathcal{C}\right)\mu((1 - \zeta)B).$$

We claim that

$$(7.10) \qquad\qquad \mathcal{C} \cap f^{-\varepsilon n}((1 - \zeta)B)) \subset \bigcup_{j \text{ good}} \tilde{C}_j$$

The proof of the above claim is analogous to the proof of (7.5). By (7.10) and (7.2),

$$\mu\Big(\bigcup_{j \text{ good}} \tilde{C}_j\Big) \geqslant \mu\left(\mathcal{C} \cap f^{-\varepsilon n}((1 - \zeta)B\right) \geqslant (1 - 2\varepsilon^{40})\mu(\tilde{B})\mu((1 - \zeta)B) \geqslant$$

$$(7.11) \qquad\qquad\qquad (1 - 3\varepsilon^{40})\mu(\tilde{B})\mu((1 - \zeta)B).$$

By Corollary 6.6 it follows that

$$(7.12) \qquad (1 + \varepsilon^{2000})\boldsymbol{\rho}_{\tilde{B}} \cdot m^u(A^-)\mathrm{mes}\left(B^{cs}(0, e^{-\eta_2\varepsilon n - \varepsilon^2 n})\right) \geqslant \mu\left(\bigcup_{j \text{ good}} \tilde{C}_j\right)$$

Moreover by Lemma 6.7 and Corollary 6.6,

$$(7.13) \quad \mu(\tilde{B}) \geqslant (1 - \varepsilon^{2000})\mu(\mathcal{C}) \geqslant (1 - \varepsilon^{2000})\boldsymbol{\rho}_{\tilde{B}} \cdot m^u(\mathcal{W} \cap \mathcal{C})\mathrm{mes}\left(B^{cs}(0, e^{-\eta_2\varepsilon n - \varepsilon^2 n})\right)$$

Combining (7.11)–(7.13) we get

$$m^u_{\mathcal{W}}(A^-) \geqslant (1 - 4\varepsilon^{40})m^u_{\mathcal{W}}(\mathcal{C})\mu((1 - \zeta)B)$$

finishing the proof of the lower bound.                                                $\square$

For $r > 0$ let $\{B_i(r)\}_{i \in J_r} = \{B_i(r)\}_{i \in J_r}$ be the family of parallelograms constructed in Lemma 5.1. The next result, proven in §A.3 will help us to verify the fully crossing assumption in Proposition 7.1.

**Lemma 7.2.** *For every $\hat{\varepsilon} > 0$ there exists $n_{\hat{\varepsilon}, \varepsilon} \in \mathbb{N}$ such that for every $n \geqslant n_{\hat{\varepsilon}, \varepsilon}$ any $\mathcal{W} \subset W^u_{\bar{z}, \tau}$, for some $\bar{z} \in \mathcal{L}_{n, \tau}$ any $i \in J_r$, for $r = e^{-\eta_2\varepsilon n}$ if $\partial\mathcal{W} \bigcap B_i(r) = \emptyset$ and $\mathcal{W} \cap (1 - \hat{\varepsilon}) \cdot B_i(r) \neq \emptyset$, then $\mathcal{W}$ crosses $B_i(r)$ completely.*

7.2. **The Main Proposition.** Recall that $\tilde{r}_n = e^{-\varepsilon^2 n - \varepsilon \eta_2 n}$ and let

(7.14)
$$\mathcal{R}_n := \bigcup_{i,j} \tilde{R}_{i,j}(\tilde{r}_n)$$

be the lamination defined by (5.16).

Let $\{B_i\,(e^{-\eta_2 \varepsilon n})\}_{i \in J}$ be the family given by Lemma 5.1.

**Proposition 7.3** (**Main Proposition**). *Assume that $f$ is exponentially mixing. For every $\varepsilon > 0$ there exists $\xi_0 < \tau$ and such that for each $\xi < \xi_0$ there is $\bar{n} = n(\varepsilon, \xi) \in \mathbb{N}$, such that for every $n \geqslant \bar{n}$, there exists a set $K_n \subset \mathcal{L}_{n,\tau}$, $\mu(K_n) \geqslant 1 - \varepsilon^{10}$ satisfying the following: let $\{B_i\}_{i \in J}$ be either $\{B_i(e^{-\eta_2 \varepsilon n})\}_{i \in J_{e^{-\eta_2 \varepsilon n}}}$ OR a family of disjoint balls $\{O_i(r)\}_{i \in P}$ with $r \in [e^{-\eta_2 \varepsilon n}, 1)$ and $\mu(\bigcup_{i \in P} O_i(r)) \geqslant 1 - \varepsilon^{10^5}$. Then for every $x \in K_n$, every unstable box $\mathcal{W}$ of size $\xi$ containing $x$ there is a subset $J'(x) \subset J$ such that $\mu(\bigcup_{i \in J'} B_i) > 1 - \varepsilon^{10}$ and we have*

(7.15)
$$m_{\mathcal{W}}^u(f^{-\varepsilon n}(\mathcal{L}_{n,\tau})) \geqslant 1 - \varepsilon^{10},$$

*and for every $i \in J'(x)$,*

(7.16)
$$m_{\mathcal{W}}^u\left(\mathcal{R}_n \cap f^{-\varepsilon n}(B_i)\right) \in (1 - \varepsilon^{10}, 1 + \varepsilon^{10})\mu(B_i),$$

*and for $\tilde{\varepsilon} = \varepsilon^{4000}$,*

(7.17)
$$m_{\mathcal{W}}^u\left(\mathcal{R}_n \cap f^{-\varepsilon n}((1 - \tilde{\varepsilon})B_i)\right) \in (1 - \varepsilon^{10}, 1 + \varepsilon^{10})\mu((1 - \tilde{\varepsilon})B_i).$$

7.3. **Proof of Main Proposition.** Fix $\varepsilon, \xi > 0$. Notice that it is enough to prove (7.16) for $\{B_i\}$ and then repeat the proof for $\{(1 - \tilde{\varepsilon})B_i\}$ and take intersections of the corresponding sets $K_n$. Let $\hat{\varepsilon} = \varepsilon^{10^4}$ and let $\bar{n} \geqslant n_{\hat{\varepsilon},\varepsilon}$, where $n_{\hat{\varepsilon},\varepsilon}$ is the maximum of the corresponding $n_{\hat{\varepsilon},\varepsilon}$ (or $n_\varepsilon$) coming from Proposition 7.1, Lemma 6.7, Corollary 6.6, Proposition 6.4, Lemma B.2, Lemma 5.1 and Lemma 7.2 are satisfied.

Let $\bar{\varepsilon} = \bar{\varepsilon}(\hat{\varepsilon}) < \varepsilon^{1000}$ be small enough so that for each $i \in J$, we have

(7.18)
$$\mu((1 + \bar{\varepsilon}) \cdot B_i) \leqslant (1 + \hat{\varepsilon}^2)\mu(B_i) \quad \text{and} \quad \mu((1 - \bar{\varepsilon}) \cdot B_i) \geqslant (1 - \hat{\varepsilon}^2)\mu(B_i).$$

We will now define the set $K_n$. Let $\{B_i(\tilde{r}_n)\}_{i \in J_{\tilde{r}_n}}$ be the family $\mathbb{B}_2(r)$ for $r = \tilde{r}_n$. To make the notation simpler we will denote the family $\{B_i(\tilde{r}_n)\}_{i \in J_{\tilde{r}_n}}$ by [6] $\{\tilde{B}_s\}_{s \in \bar{J}}$.

Let $\mathcal{B}_n(\hat{\varepsilon}) := \left(\bigcup_{s \in \bar{J}}(1 - \hat{\varepsilon}) \cdot \tilde{B}_s\right)$ and

$$K^\flat := \left\{x \in M \,:\, m_{W_{x,2\xi}^u}^u\left(\mathcal{B}_n(\hat{\varepsilon}) \cap f^{-\varepsilon n}(\mathcal{L}_{n,\tau}) \cap \mathcal{L}_{n,\tau}\right) \geqslant 1 - \varepsilon^{400}\right\}.$$

Define

$$K_n := K^\flat \cap \mathcal{L}_{n,\tau}.$$

Below we summarize the properties of $K_n$ needed in the proof. Let $x \in K_n$ and let $\mathcal{W}$ be an unstable box of size $\xi$ containing $x$, then

 **A.** $\mu(K_n) > 1 - \varepsilon^{10}$;

 **B.** $m_{\mathcal{W}}^u\left(f^{-\varepsilon n}(\mathcal{L}_{n,\tau}) \cap \mathcal{L}_{n,\tau} \cap \mathcal{B}_n(\hat{\varepsilon})\right) \geqslant 1 - \varepsilon^{300}$;

---

[6] The family $\{\tilde{B}_s\}_{s \in \bar{J}}$ should not be confused with $\{B_i(e^{-\eta_2 \varepsilon n})\}_{i \in J_{e^{-\eta_2 \varepsilon n}}}$ which is one of the possible targets in Proposition 7.3.

**C.** There exists a set $J' \subset J$ such that $\mu \left( \bigcup_{i \in J \setminus J'} B_i \right) < \varepsilon^{10}$ and for every $i \in J'$

$$(7.19) \qquad m^u_{\mathcal{W}} \left( \mathcal{R}_n \cap f^{-\varepsilon n}(B_i) \right) \in (1 - \varepsilon^{10}, 1 + \varepsilon^{10}) \mu(B_i).$$

Notice that **A**– **C** immediately imply the Main Proposition. Therefore, it remains to prove properties **A**– **C**.

*Proof of* **A***:* Notice that **A** follows by showing that $\mu(K^\flat) \geqslant 1 - \varepsilon^{20}$. But that estimate follows by applying Lemma 4.4 with $\mathfrak{B} = \mathcal{L}^c_{n,\tau} \cup f^{-\varepsilon n} \mathcal{L}^c_{n,\tau} \cup \mathcal{B}_n(\hat{\varepsilon})^c$ and $\varepsilon^{400}$ instead of $\varepsilon$ (note that $\mu(\mathfrak{B}) \leqslant \varepsilon^{4 \times 400}$ by the definition of $\hat{\varepsilon}$ and (5.8)).

*Proof of* **B***:* Since $\mathcal{W} \subset W^u_{x,2\xi}$ and $x \in K_1$

$$m^u \left( \mathcal{W} \setminus (f^{-\varepsilon n} \mathcal{L}_{n,\tau} \cap \mathcal{L}_{n,\tau}) \cap \mathcal{B}_n(\hat{\varepsilon}) \right) \leqslant$$

$$m^u \left( W^u_{x,2\xi} \setminus (f^{-\varepsilon n} L_{n,\tau} \cap \mathcal{L}_{n,\tau}) \cap \mathcal{B}_n(\hat{\varepsilon}) \right) \leqslant \varepsilon^{400} m^u(W^u_{x,2\xi}) \leqslant \varepsilon^{300} m^u(\mathcal{W}).$$

This finishes the proof of **B**.

*Proof of* **C***:* By **B** it follows that $m^u_{\mathcal{W}}(\mathcal{B}_n(\hat{\varepsilon})) \geqslant 1 - \varepsilon^{300}$. Let $S \subset \bar{J}$ be such that for $s \in S$, $\mathcal{W} \cap (1 - \varepsilon^{40}) \cdot \tilde{B}_s \neq \emptyset$ and $\partial \mathcal{W} \cap \tilde{B}_s = \emptyset$. Note that by Lemma 7.2, $\mathcal{W}$ crosses $\tilde{B}_s$ completely for every $s \in S$. Denote $\tilde{B}_{\mathcal{W}} = \bigcup_{s \in S} \tilde{B}_s$.

We claim that

$$(7.20) \qquad m^u_{\mathcal{W}}(\tilde{B}_{\mathcal{W}}) \geqslant 1 - 2\varepsilon^{300}.$$

Indeed, if for some $s' \in \bar{J}$, $\partial \mathcal{W} \cap \tilde{B}_{s'} \neq \emptyset$, then $\tilde{B}_{s'} \subset V_{n^{-2}}(\partial \mathcal{W})$ where $V_{n^{-2}}(\partial \mathcal{W})$ denotes $n^{-2}$ neighborhood of $\partial \mathcal{W}$. Therefore using **B** we get

$$m^u_{\mathcal{W}}(\tilde{B}_{\mathcal{W}}) \geqslant m^u_{\mathcal{W}}(\bigcup_{s \in \bar{J}} \tilde{B}_s) - m^u_{\mathcal{W}}(V_{n^{-2}}(\partial \mathcal{W})) \geqslant m^u_{\mathcal{W}}(\mathcal{B}_n(\varepsilon)) - C(\varepsilon, \xi) \cdot n^{-2} \geqslant 1 - 2\varepsilon^{300},$$

if $n$ is large enough.

Let $J_0 \subset J$ be such that for $i \in J_0$,

$$(7.21) \qquad m^u_{\mathcal{W}}(\mathcal{R}_n \cap \tilde{B}^c_{\mathcal{W}} \cap f^{-\varepsilon n} B_i) \geqslant \varepsilon^{150} \mu(B_i).$$

Then

$$2\varepsilon^{300} > m^u_{\mathcal{W}}(\tilde{B}^c_{\mathcal{W}}) \geqslant \sum_{i \in J_0} m^u_{\mathcal{W}}(\mathcal{R}_n \cap \tilde{B}^c_{\mathcal{W}} \cap f^{-\varepsilon n} B_i) \geqslant \varepsilon^{150} \mu(\bigcup_{i \in J_0} B_i),$$

and hence

$$\mu(\bigcup_{i \in J \setminus J_0} B_i) \geqslant 1 - 3\varepsilon^{150}.$$

We will therefore discard the set $J_0$ and, with a slight abuse of notation, still denote $J = J \setminus J_0$.

By Lemma B.2 applied to each $\tilde{B}_s$ and $(1 - \bar{\varepsilon}) B_i$, $i \in J$, we get

$$(7.22) \qquad \mu(\tilde{B}_{\mathcal{W}} \cap f^{-n}((1 - \bar{\varepsilon}) B_i) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon}) \mu(\tilde{B}_{\mathcal{W}}) \mu((1 - \bar{\varepsilon}) B_i).$$

By the right inequality in (7.18) and Lemma 5.1

$$\mu(\bigcup_{i\in J}(1-\bar{\varepsilon})B_i)) \geqslant (1-\hat{\varepsilon}^2)\mu(\bigcup_{i\in J}B_i) \geqslant (1-\hat{\varepsilon}^2)(1-3\varepsilon^{150}) \geqslant 1-4\varepsilon^{150}.$$

Therefore,

$$\mu\Big(\tilde{B}_{\mathcal{W}} \cap \Big(\bigcup_{i\in J}f^{-n}(1-\bar{\varepsilon})B_i\Big)\Big) \geqslant (1-\hat{\varepsilon})(1-4\varepsilon^{150})\mu(\tilde{B}_{\mathcal{W}}) \geqslant (1-5\varepsilon^{150})\mu(\tilde{B}_{\mathcal{W}})$$

Note that by Lemma 6.7 applied to each $s \in S$,

$$\mu(\mathcal{R}_n \cap \tilde{B}_{\mathcal{W}}) = \sum_{s\in S}\mu\Big(\bigcup_{j}\tilde{R}_{s,j}(e^{-\eta_2\varepsilon n-\varepsilon^2 n})\Big) \geqslant (1-\hat{\varepsilon})\sum_{s\in S}\mu(\tilde{B}_s) = (1-\hat{\varepsilon})\mu(\tilde{B}_{\mathcal{W}}).$$

The two inequalities above give

$$(7.23) \qquad \mu\Big(\mathcal{R}_n \cap \tilde{B}_{\mathcal{W}} \cap \Big(\bigcup_{i\in J}f^{-n}(1-\bar{\varepsilon})B_i\Big)\Big) \geqslant (1-6\varepsilon^{150})\mu(\tilde{B}_{\mathcal{W}}).$$

Let $J' \subset J$ be such that for every $i \in J'$,

$$(7.24) \qquad \mu\Big(\mathcal{R}_n \cap \tilde{B}_{\mathcal{W}} \cap f^{-n}\Big((1-\bar{\varepsilon})B_i\Big)\Big) \geqslant (1-\varepsilon^{100})\mu\Big(\tilde{B}_{\mathcal{W}} \cap f^{-n}((1-\bar{\varepsilon})B_i)\Big).$$

By (7.23), the definition of $J'$, and (7.22),

$$(1-6\varepsilon^{150})\mu(\tilde{B}_{\mathcal{W}}) \leqslant \mu\Big(\mathcal{R}_n \cap \tilde{B}_{\mathcal{W}} \cap \Big(\bigcup_{i\in J}f^{-n}((1-\bar{\varepsilon})B_i)\Big)\Big) =$$

$$\sum_{i\in J}\mu\Big(\mathcal{R}_n \cap \tilde{B}_{\mathcal{W}} \cap f^{-n}((1-\bar{\varepsilon})B_i)\Big) \leqslant$$

$$\sum_{i\in J'}\mu(\tilde{B}_{\mathcal{W}} \cap f^{-n}((1-\bar{\varepsilon})B_i)) + (1-\varepsilon^{100})\sum_{i\in J\setminus J'}\mu(\tilde{B}_{\mathcal{W}} \cap f^{-n}((1-\bar{\varepsilon})B_i)) \leqslant$$

$$(1+\hat{\varepsilon})\mu(\tilde{B}_{\mathcal{W}})\mu(\bigcup_{i\in J'}(1-\bar{\varepsilon})B_i) + (1+\hat{\varepsilon})(1-\varepsilon^{100})\mu(\tilde{B}_{\mathcal{W}})\mu(\bigcup_{i\in J\setminus J'}(1-\bar{\varepsilon})B_i) \leqslant$$

$$(1+\hat{\varepsilon})^2\mu(\tilde{B}_{\mathcal{W}}) - \varepsilon^{100}(1+\hat{\varepsilon})\mu(\tilde{B}_{\mathcal{W}})\mu(\bigcup_{i\in J\setminus J'}(1-\bar{\varepsilon})B_i).$$

From the above it follows that

$$\mu(\bigcup_{i\in J\setminus J'}B_i) \leqslant 2\mu(\bigcup_{i\in J\setminus J'}(1-\bar{\varepsilon})B_i) \leqslant \varepsilon^{10}.$$

It remains to show that (7.19) holds for $i \in J'$. Fix $i \in J'$. We will first show the upper bound, which is slightly easier. Notice that since $\mathcal{W}$ crosses $\tilde{B}_s$ completely, we have by the upper bound in Proposition 7.1 and (7.21),

$$m^u_{\mathcal{W}}(\mathcal{R}_n \cap f^{-\varepsilon n}B_i) \leqslant \varepsilon^{150}\mu(B_i) + m^u_{\mathcal{W}}\Big(\mathcal{R}_n \cap \tilde{B}_{\mathcal{W}} \cap f^{-\varepsilon n}(B_i)\Big) =$$

$$\varepsilon^{150}\mu(B_i) + \sum_{s\in S}m^u_{\mathcal{W}}\Big(\mathcal{R}_n \cap \tilde{B}_s \cap f^{-\varepsilon n}(B_i)\Big) \leqslant$$

$$\varepsilon^{150}\mu(B_i) + (1+\varepsilon^{20})\mu(B_i)\sum_{s\in S}m^u_{\mathcal{W}}(\mathcal{R}_n \cap \tilde{B}_s) \leqslant (1+\varepsilon^{10})\mu(B_i).$$

This finishes the proof of the upper bound.

We now proceed to show the lower bound in (7.19). Let $S' = S'(i) \subset S$ be such that for $s \in S'$, we have

(7.25) $$\mu(\mathcal{R}_n \cap \tilde{B}_s \cap f^{-n}(1 - \bar{\varepsilon})B_i) \geqslant (1 - \varepsilon^{40})\mu(\tilde{B}_s \cap f^{-n}(1 - \bar{\varepsilon})B_i).$$

Then by (7.24) and exponential mixing,

$$(1 - \varepsilon^{100})\mu(\tilde{B}_{\mathcal{W}})\mu(B_i) \leqslant \sum_{s \in S} \mu\Big(\mathcal{R}_n \cap \tilde{B}_s \cap f^{-n}((1 - \bar{\varepsilon})B_i)\Big) \leqslant$$

$$(1 + 2\hat{\varepsilon}) \sum_{s \in S'} \mu(\tilde{B}_s)\mu(B_i) + (1 + 2\hat{\varepsilon})(1 - \varepsilon^{40}) \sum_{s \in S \setminus S'} \mu(\tilde{B}_s)\mu(B_i) \leqslant$$

$$(1 + 2\hat{\varepsilon})\mu(B_i)\mu(\tilde{B}_{\mathcal{W}}) - \varepsilon^{40}\mu(B_i) \sum_{s \in S \setminus S'} \mu(\tilde{B}_s).$$

From this it follows that

(7.26) $$\sum_{s \in S'} \mu(\tilde{B}_s) \geqslant (1 - 2\varepsilon^{60})\mu(\tilde{B}_{\mathcal{W}}).$$

Take $s \in S'$. Then (7.25) holds. Now exponential mixing implies that (7.2) holds with $\zeta = \bar{\varepsilon}$. Since $\mathcal{W}$ crosses $\tilde{B}_s$ completely it follows from Proposition 7.1 that

(7.27) $$m_{\mathcal{W}}^u(\mathcal{R}_n \cap \tilde{B}_s \cap f^{-\varepsilon n}B_i) \geqslant (1 - 2\varepsilon^{20})\mu(B_i)m_{\mathcal{W}}^u(\mathcal{R}_n \cap \tilde{B}_s).$$

Summing the above over $s \in S'$, we get that

(7.28) $$m_{\mathcal{W}}^u(\mathcal{R}_n \cap \bigcup_{s \in S'} \tilde{B}_s \cap f^{-\varepsilon n}B_i) \geqslant (1 - 2\varepsilon^{20})\mu(B_i)m_{\mathcal{W}}^u(\mathcal{R}_n \cap \bigcup_{s \in S'} \tilde{B}_s).$$

Let $s \in S$. Since $\mathcal{W}$ crosses $\tilde{B}_s$ completely, we have by Proposition 6.4

$$m_{\mathcal{W}}^u(\mathcal{R}_n \cap \tilde{B}_s) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})m_{\widehat{\mathcal{W}}_s}^u(\mathcal{R}_n \cap \tilde{B}_s)$$

and

$$m_{\widehat{\mathcal{W}}_s}^u(\mathcal{R}_n \cap \tilde{B}_s) = m_{\widehat{\mathcal{W}}_s}^u(\bigcup_j R_{s,j}) \geqslant 1 - \hat{\varepsilon}.$$

From this it follows that for every $s, s' \in S$,

$$\frac{m_{\mathcal{W}}^u(\mathcal{R}_n \cap \tilde{B}_s)}{m_{\mathcal{W}}^u(\mathcal{R}_n \cap \tilde{B}_{s'})} \in (1 - 4\hat{\varepsilon}, 1 + 4\hat{\varepsilon}).$$

Combining this with (7.20) and Lemma 6.7 we get that

$$m_{\mathcal{W}}^u(\mathcal{R}_n \cap \bigcup_{s \in S'} \tilde{B}_s) \geqslant (1 - 2\varepsilon^{20})\frac{|S'|}{|S|}m_{\mathcal{W}}^u(\tilde{B}_{\mathcal{W}} \cap \mathcal{R}_n) \geqslant (1 - 4\varepsilon^{20})m_{\mathcal{W}}^u(\tilde{B}_{\mathcal{W}}) \geqslant 1 - 8\varepsilon^{20}.$$

Using (7.28), we get $m_{\mathcal{W}}^u(\mathcal{R}_n \cap f^{-\varepsilon n}(B_i)) \geqslant (1 - \varepsilon^{10})\mu(B_i)$. This finishes the proof of the lower bound and hence also the proof of **C**. The proof of Main Proposition is thus finished. $\qquad\square$

## 8. K-PROPERTY

In this section we will use Proposition 7.3 to show the $K$-property. In fact, the full strength of Proposition 7.3 is not needed for the $K$-property, i.e. we don't need a quantitative equidistribution. Since this result is of independent interest, we will present the proof in the more general case. For this we introduce the notion of equidistributed leaves.

**Definition 8.1.** Let $f \in C^{1+\alpha}(M, \mu)$. We say that the images of most unstable leaves become equidistributed under $f$ if for each ball $\mathbf{B} \subset M$ and $\varepsilon > 0$ there is $\xi > 0$ such that for each sufficiently large $n$ there is a set $K = K_n$ with $\mu(K) > 1 - \varepsilon$ such that for each $x \in K$ the size of unstable manifold of $x$ is greater than $\xi$ and moreover

$$(8.1) \qquad m^u_{W^u_{x,\xi}}\left(f^{-n}(\mathbf{B})\right) \in (1 - \varepsilon, 1 + \varepsilon)\mu(\mathbf{B}).$$

We note that this definition is equivalent to a stronger property, namely that (8.1) holds for each unstable box containing $x$ (rather than centered at $x$):

**Lemma 8.2.** If the images of most unstable leaves become equidistributed under $f$ then for each ball $\mathbf{B} \subset M$ and $\varepsilon > 0$ there is $\xi > 0$ such that for each sufficiently large $n$ there is a set $\hat{K} = \hat{K}_n$ with $\mu(\hat{K}) > 1 - \varepsilon$ such that for each $x \in \hat{K}$ for each $y$ such that $x \in W^u_{y,\xi}$ we have

$$(8.2) \qquad m^u_{W^u_{y,\xi}}\left(f^{-n}(\mathbf{B})\right) \in (1 - \varepsilon, 1 + \varepsilon)\mu(\mathbf{B}).$$

*Proof.* Take $\hat{\varepsilon} \ll \varepsilon$. Let $n$ be large and let $\tilde{K}_n$ be the set of points satisfying

$$m^u_{W^u_{x,\xi}}\left(f^{-n}(\mathbf{B})\right) \in (1 - \varepsilon, 1 + \varepsilon)\mu(\mathbf{B}).$$

Since the images of most unstable leaves become equidistributed under $f$, we have that $K_n \subset \tilde{K}_n$ and so $\mu(\tilde{K}_n) > 1 - \hat{\varepsilon}$ if $n$ is large enough. Let

$$\hat{K} = \hat{K}_n = \{x : r_u(x) \geqslant 3\xi \text{ and } m_{W^u_{x,3\xi}}(\tilde{K}) > 1 - \hat{\varepsilon}^{1/4}\}.$$

By Lemma 4.4, $\mu(\hat{K}) \geqslant 1 - 4\hat{\varepsilon}^{1/4}$. Next, take $x \in \hat{K}$ and $y \in M$ s.t. $x \in W^u_{y,\xi}$. Since $x \in \hat{K}_n$ we can find $z \in W^u_{x,3\xi} \cap \tilde{K}_n$ s.t. $d(y, z) \leqslant 3\hat{\varepsilon}^{(1/4D)}\xi$. Then

$$m^u_{W^u_{x,3\xi}}(W^u_{y,\xi} \cap W^u_{z,\xi}) \geqslant (1 - 4d\hat{\varepsilon}^{(1/4D)})m^u_{W^u_{x,3\xi}}(W^u_{z,\xi}).$$

Hence $y$ satisfies (8.2) provided that $4d\hat{\varepsilon}^{(1/4D)} + \hat{\varepsilon} < \varepsilon$. $\qquad \square$

**Proposition 8.3.** Let $f$ be a $C^{1+\alpha}$ diffeomorphism of a compact manifold $M$ preserving a smooth measure $\mu$. If the images of most unstable leaves become equidistributed under $f$, then $f$ has $K$ property.

*Proof.* We will use Lemma 2.17. Fix $\mathcal{D} \in \mathcal{B}$, $\varepsilon >$ and $\tau > \xi > 0$. Let $\{\mathbf{B}_i\}_{i \in J}$, be a family of balls with $r = r(\varepsilon, \xi)$ such that $\mu(\mathcal{D} \triangle \bigcup_{i \in J} \mathbf{B}_i) < \varepsilon^2$. It is enough show that (2.20) holds for $\mathbf{B}_i$ (and arbitrary small $\varepsilon$). Take $n$ sufficiently large and let $\hat{K} = \hat{K}_n$

be from Lemma 8.2. By Lemma 4.2 and Markov inequality, $\varepsilon$ almost every atom $A$ of $\bigvee_{i=N_1}^{N_2} f^i(\mathcal{P})$ is $\sqrt{\varepsilon}$ saturated and has the property that

$$(8.3) \qquad\qquad \mu|_A(\hat{K}) \geqslant 1 - \sqrt{\varepsilon}.$$

We claim that every atom with these two properties satisfies (2.20). Indeed by Lemma 4.3 we have

$$\mu_A(f^{-n}B) = \int_{\mathfrak{T}} \mu^u_{\mathcal{W}_t}(f^{-n}B) d\nu_A(t) \pm \sqrt{\varepsilon}.$$

Next, by Definition 8.1, if $\mathcal{W}_t \cap \hat{K} \neq \emptyset$ then $m^u_{\mathcal{W}_t}(f^{-n}B) = \mu(B) \pm \varepsilon$ while by our assumption (8.3), $\nu_A(t : \mathcal{W}_t \cap K = \emptyset) \leqslant \sqrt{\varepsilon}$. Combining the above estimates we obtain (2.20) proving the proposition. $\qquad\square$

As a consequence of the above result and the Main Proposition, we deduce:

**Corollary 8.4.** *If $f$ has exponential decay of correlation then $f$ is a $K$-automorphism.*

*Proof.* It is enough to show that Main Proposition implies that he images of most unstable leaves become equidistributed under $f$. For given $\varepsilon$, and **B**, let $\{O_i\}$be any family of balls containing **B**. Let $\bar{\varepsilon} = \min\left(\varepsilon^{100}, \min_{i \in \{1,\dots,m\}} \frac{\mu(O_i)}{100}\right)$. Let $n(\bar{\varepsilon}, \xi)$ come from the Main Proposition. Take $n \geqslant n(\bar{\varepsilon}, \xi)$ and let $K = K_n$ be as in the Main Proposition (for the family $\{O_i\}_{i \in J}$). Note that due to our choice of $\varepsilon$, $J'(x) = J$ for every $x \in K$ since removing even one ball from $J$ will decrease the measure by at least $100\bar{\varepsilon}$. In particular, the assertion of Main Proposition applies to the ball **B**. Note that $\bar{\varepsilon}n$ is large if $n$ is large. $\qquad\square$

## 9. VWB-property

9.1. **The main reduction.** The following lemma is the main step towards the proof of VWB–property.

**Proposition 9.1.** *For every $\varepsilon$ and $\xi > 0$ there exists $n_0 = n_0(\varepsilon, \xi)$ such that for every $n \geqslant n_0$, there exists a set $K = K_n \subset M$, $\mu(K) \geqslant 1 - \varepsilon$ so that for every $x, x' \in K$, and every unstable boxes $\mathcal{W}_1$, $\mathcal{W}_2$ of size $\xi$ containing $x$ and $x'$ respectively there exists an $\varepsilon$ measure preserving map $\theta : (\mathcal{W}_1, m^u_{\mathcal{W}_1}) \to (\mathcal{W}_2, m^u_{\mathcal{W}_2})$ such that*

$$(9.1) \qquad \frac{1}{n}\mathrm{Card}\left(\{i \in \{0, \dots, n-1\} : d(f^i x, f^i(\theta x)) < \varepsilon\}\right) > 1 - \varepsilon.$$

Before proving the proposition let us show how it implies our main result.

*Proof of Theorem 1.1.* Let $\mathcal{P}$ be a sufficiently fine partition with piecewise smooth boundary. We will verify the conditions of Corollary 2.22 with $\tilde{N} := n_0(\varepsilon, \xi)$ where $n_0$ comes from Proposition 9.1. Since $f$ enjoys the $K$-property by Corollary 8.4, we just need to verify (2.22). Moreover the partition $\mathcal{P}$ is regular since it has piecewise smooth boundary. Fix an arbitrary $\varepsilon > 0$. By Lemma 4.2 we can take a small $\xi$ and choose $N$

so large that for any $N' \geqslant N$, $\varepsilon^2$ almost every atom of the partition $\bigvee_N^{N'} T^i \mathcal{P}$ is $(\xi, \varepsilon^2)$ u-saturated.

Let $A$ be such an atom satisfying additionally $\mu_{|A}(K_n) \geqslant 1 - \varepsilon^2$ (where $K_n$ is from Proposition 9.1). Since $M$ is also $(\xi, \varepsilon^2)$ u-saturated Lemma 4.3 gives

$$\mu|A = \mu_{A,r} + \int_{\mathfrak{T}_A} \mu_{\mathcal{W}_t} d\nu_A(t), \quad \mu = \mu_{M,r} + \int_{\mathfrak{T}_M} \mu_{\mathcal{W}_t} d\nu_M(t).$$

By possibly removing from $\mathfrak{T}_A$ and $\mathfrak{T}_M$ sets of measure less than $\varepsilon^2$ we may assume that for each $t \in \mathfrak{T}_A \cup \mathfrak{T}_M$, $\mathcal{W}_t \cap K_n \neq \emptyset$. By Lemma 2.15 there is a measure preserving map

$$\Theta : \left( \mathfrak{T}_A, \frac{\nu_A}{\nu_A(\mathfrak{T}_A)} \right) \mapsto \left( \mathfrak{T}_M, \frac{\nu_M}{\nu_M(\mathfrak{T}_M)} \right).$$

Since both $\nu_A(\mathfrak{T}_A)$ and $\nu_M(\mathfrak{T}_M)$ are between $1 - \varepsilon$ and $1$, $\Theta$ is also an $\varepsilon$ measure preserving map $(\mathfrak{T}_A, \nu_A) \mapsto (\mathfrak{T}_M, \nu_M)$. By Proposition 9.1 for each $t \in \mathfrak{T}_A$ there exists a $\varepsilon$ measure preserving map $\theta_{t,S} : \mathcal{W}_t \mapsto \mathcal{W}_{\Theta(t)}$ satisfying (9.1) with $n = S$ (notice that $S \geqslant \tilde{N} = n_0$). Defining $\theta(x) = \theta_{t,S}(x)$ if $x \in \mathcal{W}_t$ for some $t \in \mathfrak{T}_A$ and defining it in an arbitrary way if $x \notin \bigcup_{t \in \mathfrak{T}_A} \mathcal{W}_t$ gives an $3\varepsilon$ preserving map between $(A, \mu|_A)$ and $(M, \mu)$ satisfying (2.22). Since $\varepsilon$ is arbitrary, $f$ is Bernoulli.    $\square$

9.2. **Plan of the proof of Proposition 9.1.** For $\varepsilon$ and $\xi$ let

$$n_0(\varepsilon, \xi) := \varepsilon^{-2} \max(n(\varepsilon^2, \xi), n_{\varepsilon^2})$$

where $n(\varepsilon, \xi)$ and $n_\varepsilon$ come from the Main Proposition and Proposition 6.8 respectively. Let $n \geqslant n_0$, then $n' := \varepsilon n \geqslant n(\varepsilon^2, \xi)$. Let $\{B_i\}_{i \in J}$ be the collection of parallelograms and $K_n$ be the set from the Main Proposition. For $x, x' \in K_n$ let $J'' = J'(x) \cap J'(x')$ where $J'(x), J'(x') \subset J$ are the subcollections given by the Main Proposition. Then $\mu \left( \bigcup_{i \in J''} B_i \right) \geqslant 1 - 2\varepsilon^{10}$, so it follows by (7.16) that for $\mathcal{W} \in \{\mathcal{W}_1, \mathcal{W}_2\}$

$$m_{\mathcal{W}}^u \left( \mathcal{R}_n \setminus \bigcup_{i \in J''} f^{-n'} B_i \right) \leqslant 3\varepsilon^{10}$$

From this and (7.15) it follows that

$$m_{\mathcal{W}}^u \left( \left[ \mathcal{R}_n \cap f^{-n'}(\mathcal{L}_{n,\tau}) \right] \setminus \bigcup_{i \in J''} f^{-n'} B_i \right) \leqslant 4\varepsilon^{10}.$$

The proof of Proposition 9.1 consists of three steps. On the first step we remove a small proportion of parallelograms from $J''$ so that for the remaining parallelograms most of the intersections of $f^{n'} \mathcal{W}_s \cap B_i$ happen inside the regular set and not too close to the boundary of $B_i$. On the second step we show that for good parallelograms constructed on the first step most of the components of $f^{n'} \mathcal{W}_s \cap B_i$ are Markov, i.e. they are fullly crossing $B_i$. On the third step we construct the coupling between the

Markov components using fake center stable holonomy. Steps 1–3 described above are curried out in §§9.3–9.5 respectively.

### 9.3. Prunning.

Let $J''' \subset J''$ be the set of $i$ such that for $s = 1, 2$, we have

$$(9.2) \quad (1 + \varepsilon^4) m^u_{\mathcal{W}_s}(\mathcal{R}_n \cap f^{-n'}(\mathcal{L}_{n,\tau}) \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)) \geqslant m^u_{\mathcal{W}_s}(\mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)).$$

**Lemma 9.2.** $m^u_{\mathcal{W}_s}\left( \bigcup_{i \in J'' \setminus J'''} f^{-n'}B_i \right) \leqslant \varepsilon^2.$

*Proof.* By definition $J'' \setminus J''' = \hat{J}_1 \cup \hat{J}_2$ where $\hat{J}_s$ is the set of indices such that (9.2) fails for $\mathcal{W}_s$. Note that for $i \in \hat{J}_s$

$$m^u_{\mathcal{W}_s}(f^{-n'}(\mathcal{L}^c_{n,\tau}) \cap \mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)) =$$

$$m^u_{\mathcal{W}_s}(\mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)) - m^u_{\mathcal{W}_s}(f^{-n'}(\mathcal{L}_{n,\tau}) \cap \mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)) \geqslant$$

$$\frac{\varepsilon^4}{2} m^u_{\mathcal{W}_s}(\mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)) \geqslant \frac{\varepsilon^4}{3} \cdot \mu(B_i),$$

the last inequality by Main Proposition and since $\tilde{\varepsilon} = \varepsilon^{4000}$. Therefore

$$\varepsilon^{10} \overset{(7.15)}{\geqslant} m^u_{W^s}(f^{-n'}(\mathcal{L}^c_{n,\tau})) \geqslant$$

$$m^u_{\mathcal{W}_s}(f^{-n'}(\mathcal{L}^c_{n,\tau}) \cap \mathcal{R}_n) \geqslant \sum_{i \in \hat{J}_s} m^u_{\mathcal{W}_s}(f^{-n'}(\mathcal{L}^c_{n,\tau}) \cap \mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B_i)) \geqslant \frac{\varepsilon^4}{3} \sum_{i \in \hat{J}_s} \mu(B_i).$$

Hence $\sum_{i \in \hat{J}_s} \mu(B_i) \leqslant 3\varepsilon^6$. Since $J'' \setminus J''' = \hat{J}_1 \cup \hat{J}_2$ it follows that

$$\sum_{i \in J'' \setminus J'''} \mu(B_i) \leqslant 6\varepsilon^6.$$

Now the Main Proposition gives

$$\sum_{i \in J'' \setminus J'''} m^u_{\mathcal{W}_s}(\mathcal{R}_n \cap f^{-n'}B_i) \leqslant 8\varepsilon^6.$$

Since

$$m^u_{\mathcal{W}_s}\left( \bigcup_{i \in J'' \setminus J'''} f^{-n'}B_i \right) \leqslant m^u_{\mathcal{W}_s}\left( \mathcal{R}_n \cap \left[ \bigcup_{i \in J'' \setminus J'''} f^{-n'}B_i \right] \right) + m^u_{\mathcal{W}_s}(\mathcal{R}^c_n)$$

the result follows. $\qquad \square$

9.4. **Abundance of Markov returns.** By Lemma 9.2 it is enough to construct for every $i \in J'''$, an $\varepsilon/10$- measure preserving map

$$(9.3) \quad \theta_i : \left( \mathcal{R}_n \cap f^{-n'}(\mathcal{L}_{n,\tau}) \cap f^{-n'}(B_i), m^u_{\mathcal{W}_1 \cap f^{-n'}(B_i)} \right) \rightarrow$$

$$\left( \mathcal{R}_n \cap f^{-n'}(\mathcal{L}_{n,\tau}) \cap f^{-n'}(B_i), m^u_{\mathcal{W}_2 \cap f^{-n'}(B_i)} \right)$$

such that (9.1) holds for $x$ and $\theta_i x$. From now, we fix $i \in J'''$ and drop it from the notation. Recall that $\tilde{\varepsilon} = \varepsilon^{4000}$. Fix $s = 1, 2$. For any $\bar{z} \in P_\tau \cap f^{n'}\mathcal{W}_s \cap B$ let $\mathcal{W}(\bar{z}) = W^u_{\bar{z},\tau} \cap B$. Notice that if $\bar{z}' \in \mathcal{W}(\bar{z}) \cap P_\tau \cap f^{n'}(\mathcal{W}_s)$, then $\mathcal{W}(\bar{z}') = \mathcal{W}(\bar{z})$. Moreover, if $\bar{z} \in P_\tau \cap f^{n'}\mathcal{W}_s \cap (1 - \tilde{\varepsilon})B$ and $d(f^{-n'}\bar{z}, \partial\mathcal{W}_s) \geqslant c^{-n'}$ (for some small $c > 0$), then the size [7] of $\mathcal{W}(\bar{z})$ is $\in (1 - \varepsilon^{10}, 1 + \varepsilon^{10})\xi_n$ (recall that $\widehat{\mathcal{W}_i} \subset B_i$ is an unstable cube of size $\xi_n$). From the above it follows that for some $\{z_{s,\ell}\}_{\ell=1}^{m_s} \subset \mathcal{L}_{n,\tau}$,

$$\{x \in \mathcal{W}_s \cap f^{-n'}(\mathcal{L}_{n,\tau}) \cap f^{-n'}((1 - \tilde{\varepsilon})B) \; : \; d(x, \partial\mathcal{W}_s) \geqslant c^{-n'}\} \subset \bigcup_{\ell=1}^{m_s} f^{-n'}(\mathcal{W}(z_{s,\ell}))$$

and

$$(9.4) \qquad \{x \in \mathcal{W}_s : d(x, \partial\mathcal{W}_s) \geqslant c^{-n'}\} \cap \bigcup_{\ell=1}^{m_s} f^{-n'}(\mathcal{W}(z_{s,\ell})) \subset \mathcal{W}_s \cap f^{-n'}(B),$$

where

$$(9.5) \qquad \text{the size of} \quad \mathcal{W}(z_{s,\ell}) \quad \text{is} \; \geqslant (1 - \varepsilon^{10})\xi_n$$

Let $W_{s,\ell} := \mathcal{W}(z_{s,\ell})$, $\widetilde{W}_{s,\ell} = W_{s,\ell} \cap \mathcal{R}_n$.

**Lemma 9.3.**

$$\frac{m^u_{\mathcal{W}_s}\left( \mathcal{R}_n \cap \left( \bigcup_{\ell=1}^{m_s} f^{-n'}\widetilde{W}_{s,\ell} \right) \right)}{\mu(B)} \in \left[ 1 - 10\varepsilon^2, 1 + 10\varepsilon^2 \right].$$

*Proof.* By (9.2) and Main Proposition for $(1 - \tilde{\varepsilon})B$ it follows that

$$m^u_{\mathcal{W}_s}\left( \mathcal{R}_n \cap f^{-n'}(\mathcal{L}_{n,\tau}) \cap f^{-n'}((1 - \tilde{\varepsilon})B) \right) \geqslant$$

$$(1 - 2\varepsilon^4)m^u_{\mathcal{W}_s}\left( \mathcal{R}_n \cap f^{-n'}((1 - \tilde{\varepsilon})B) \right) \geqslant$$

$$(1 - 3\varepsilon^4)\mu((1 - \tilde{\varepsilon})B) \geqslant (1 - \varepsilon^3)\mu(B).$$

Moreover, $m^u_{\mathcal{W}_s}(\{x \in \mathcal{W}_s \; : \; d(x, \partial\mathcal{W}_s) \leqslant c^{-n'}\}) \leqslant C(\varepsilon) \cdot c^{-n'} \leqslant \varepsilon^{10}\mu(B)$, for large enough $n$. This together with Main Proposition shows that

$$(9.6) \qquad \frac{m^u_{\mathcal{W}_s}\left( \mathcal{R}_n \cap \left( \bigcup_{\ell=1}^{m_s} f^{-n'}W_{s,\ell} \right) \right)}{\mu(B)} \in \left[ 1 - \varepsilon^2, 1 + \varepsilon^2 \right].$$

---

[7] This means that for some point $\tilde{z} \in \mathcal{W}(\bar{z})$, $\mathcal{W}(\bar{z})$ is inside an unstable cube of $\tilde{z}$ of side length $(1 + \varepsilon^{10})\xi_n$ and $\mathcal{W}(\bar{z})$ contains the unstable cube of $\tilde{z}$ of side length $(1 - \varepsilon^{10})\xi_n$.

Next let $z', z'' \in W_{s,\ell}$. Then

$$(9.7) \qquad \frac{J(f^{-n}(z'))}{J(f^{-n}(z''))} = \prod_{j=0}^{n'-1} \left[ \frac{J(f^{-1}(f^{-j}z'))}{J(f^{-1}(f^{-j}z''))} \right] \in (1 - \varepsilon^2, 1 + \varepsilon^2)$$

where the last step uses Corollary A.9 and Lemma A.10 (applied to $L^{(k)} = TW_{s,\ell} \circ f^{k-n'}$).

Let $\tilde{B} = B \cap \mathcal{R}_n$, $\overline{\mathcal{W}} = \widehat{\mathcal{W}} \cap \mathcal{R}_n$. We apply Proposition 6.8 for $\widehat{\mathcal{W}}$ and $\mathcal{W}_z = W_{s,\ell} \subset B$, with $z = z_{s,\ell}$. Notice that $z_{s,\ell} \in \mathcal{L}_{n,\tau}$ and by (9.5), $\zeta \geqslant (1 - \varepsilon^{10})\xi_n$. For $s = 1, 2$ let $\pi_{s,\ell} : \overline{\mathcal{W}} \to \widetilde{W}_{s,\ell}$ be given by Proposition 6.8. By construction $m_{\widehat{\mathcal{W}}}^u(\mathcal{R}_n^c) \leqslant 100\varepsilon^{\mathbf{b}/16}$ and since, by Proposition 6.8, $\pi_{s,\ell}$ can be considered as an $\varepsilon^2$ measure preserving map $\widehat{\mathcal{W}} \to W_{s,\ell}$ we conclude that $m_{W_{s,\ell}}^u(\mathcal{R}_n^c) \leqslant \varepsilon^2$. Combining this with (9.6) and (9.7) we obtain the result. $\qquad \square$

### 9.5. **Construction of coupling between unstable leaves.**

*Proof of Proposition 9.1.* By Lemma 9.3 it is enough to construct an $\varepsilon/100$ -measure preserving map

$$(9.8) \quad \theta' : \left( \mathcal{R}_n \bigcap \left( \bigcup_{\ell=1}^{m_1} f^{-n'} \left( \widetilde{W}_{1,\ell} \right) \right), \; m_{\bigcup_{\ell=1}^{m_1} f^{-n'}(\widetilde{W}_{1,\ell})}^u \right) \to$$

$$\left( \mathcal{R}_n \bigcap \left( \bigcup_{\ell=1}^{m_2} f^{-n'} \left( \widetilde{W}_{2,\ell} \right) \right), \; m_{\bigcup_{\ell=1}^{m_2} f^{-n'}(\widetilde{W}_{2,\ell})}^u \right)$$

so that (9.1) holds for $x$ and $\theta'x$.

Let $\tau_{s,\ell} = \pi_{s,\ell}^{-1} \circ f^{n'}$. Let $z', z''$ be two points in $\widehat{\mathcal{W}}$. Then

$$\frac{J(\tau_{s,\ell}^{-1}(z'))}{J(\tau_{s,\ell}^{-1}(z''))} = \frac{J(\pi_{s,\ell}(z'))}{J(\pi_{s,\ell}(z''))} \times \frac{J(f^{-n'}(\pi_{s,\ell}(z')))}{J(f^{-n}(\pi_{s,\ell}(z'')))} \in \left( 1 - \varepsilon^{3/2}, 1 + \varepsilon^{3/2} \right)$$

where the first factor is estimated by by Proposition 6.8 and the second is estimated by (9.7).

Thus for any subset $Q \subset \overline{\mathcal{W}}$ we have

$$m_{\mathcal{W}_s}^u \left( \bigcup_\ell \tau_{s,\ell}^{-1}(Q) \right) = \sum_\ell \int_Q J(\tau_{s,\ell}^{-1}z) dz$$

$$\in (1 - \varepsilon^{3/2}, 1 + \varepsilon^{3/2}) \frac{m_{\widehat{\mathcal{W}}}^u(Q)}{m_{\widehat{\mathcal{W}}}^u(\overline{\mathcal{W}} \cap f^{n'}\mathcal{R}_n)} \sum_\ell \int_{\overline{\mathcal{W}} \cap f^{n'}\mathcal{R}_n} J(\tau_{s,\ell}^{-1}z) dz$$

$$(9.9) \qquad = (1 - \varepsilon^{3/2}, 1 + \varepsilon^{3/2}) \frac{m_{\widehat{\mathcal{W}}}^u(Q)}{m_{\widehat{\mathcal{W}}}^u(\overline{\mathcal{W}} \cap f^{n'}\mathcal{R}_n)} m_{\mathcal{W}_s}^u \left( \mathcal{R}_n \cap \left[ \bigcup_\ell f^{-n'} \widetilde{W}_{s,\ell} \right] \right).$$

Now divide $\widehat{\mathcal{W}}$ into cubes $Q_k$ of size $\lambda^{-2n}$ with $\lambda = \|f\|_{C^1}$. By Lemma 9.3 and (9.9) (applied to $Q_k \cap f^{n'}\mathcal{R}_n$), for each $k$

$$(9.10) \qquad \frac{m_{\mathcal{W}_1}^u \left(\mathcal{R}_n \cap [\bigcup_\ell \tau_{1,\ell}^{-1}(Q_k)]\right)}{m_{\mathcal{W}_2}^u \left(\mathcal{R}_n \cap [\bigcup_\ell \tau_{2,\ell}^{-1}(Q_k)]\right)} \in \left(1 - 10\varepsilon^{3/2}, 1 + 10\varepsilon^{3/2}\right).$$

Let $\theta$ be any $10\varepsilon^{3/2}$ measure preserving map

$$\theta : \mathcal{R}_n \cap \bigcup_{\ell=1}^{m_1} f^{-n'}(\widetilde{W}_{1,\ell}) \to \mathcal{R}_n \cap \bigcup_{\ell=1}^{m_2} f^{-n'}(\widetilde{W}_{2,\ell})$$

which, for each $k$, maps

$$\mathcal{R}_n \cap \left[\bigcup_\ell \tau_{1,\ell}^{-1}(Q_k)\right] \to \mathcal{R}_n \cap \left[\bigcup_\ell \tau_{2,\ell}^{-1}(Q_k)\right]$$

(such a map exists by (9.10)). We will show that for $n' \leqslant t \leqslant n$

$$(9.11) \qquad d(f^t x, f^t(\theta(x))) < \varepsilon.$$

This will give (9.1) and finish the proof of the proposition. By construction there is a point $y$ such that $d(f^{n'}x, f^{n'}y) \leqslant \lambda^{-2n}$ and $f^{n'}y$ and $f^{n'}(\theta(x))$ are connected by $\pi_{s,\ell}$ and are in $B$. Since $\pi_{s,\ell}$ is the holonomy for the $\mathcal{F}_{i,j}$ foliation, **F4** in Lemma 5.2 gives that for $n' \leqslant t \leqslant n$.

$$d(f^t y, f^t \theta(x)) \leqslant \frac{\varepsilon}{2}.$$

On the other hand

$$d(f^t x f^t y) \leqslant \lambda^{t-n'} d(f^{n'}x, f^{n'}y) \leqslant \lambda^{-n}.$$

Now (9.11) follows by the triangle inequality. $\qquad \square$

## 10. Bernoulli property of skew products.

We consider a skew product $F$ acting on $X \times Y$ by the formula

$$F(x,y) = (fx, \tau(x)y)$$

where $f$ is a diffeo of $X$ preserving a measure $\mu_X$, $Y$ admits an action of a Lie group $G$ preserving a measure $\mu_Y$ and $\tau : X \to G$ is a smooth function. $F$ preserves the measure $\mu = \mu_X \times \mu_Y$.

**Theorem 10.1.** [19, Theorem 4.1(a) and Remark 4.2] *If $f$ and $G$ are exponentially mixing and if there are positive constants $C, \varepsilon_1, \varepsilon_2$ such that*

$$\mu_X(x : \|\tau_N(x)\| \leqslant \varepsilon_1 N) \leqslant Ce^{-\varepsilon_2 N}$$

*then $F$ is exponentially mixing.*

Here we consider the case where $f$ is a volume preserving Anosov diffeo, $G = SL_d(\mathbb{R})$ and $Y = G/\Gamma$ where $\Gamma$ is a cocompact lattice. We consider two types skewing functions.

(a) *Generalized $T, T^{-1}$ transformations.* Let $\tau(x) = \text{diag}\left(e^{\alpha_1(x)}, e^{\alpha_2(x)}, \ldots, e^{\alpha_d(x)}\right)$ where $\sum_{k=1}^{d} \alpha_k = 0$. In this case the conditions of Theorem 10.1 are satisfied provided there is $k$ such that $\mu(\alpha_k(x)) \neq 0$ (see [19, §8.2]). Thus if $\tau$ has non-zero drift then $F$ is Bernoulli. By contrast, in the zero drift case $F$ is **not** Bernoulli by [20, Example 4.3].

(b) $\tau$ is close to identity and it is s pinching and twisting in the sense of [4]. In this case the conditions of Theorem 10.1 are satisfied due to [24, Theorem 1.5]. In particular, if we assume that $d = 2$ then the analysis of [24, §3.7] shows that if the conditions of Theorem 10.1 fail than either $\tau$ preserves a Riemannian metric on $Y$ or there is $N$ such that $F^N$ preserves a line field. In the first case $F$ is not ergodic, while in the second case (a power of) $F$ is conjugated to the map

$$F(x, y) = \left( f(x), \begin{pmatrix} \alpha(x) & \beta(x) \\ 0 & \alpha^{-1}(x) \end{pmatrix} \right).$$

In case $\mu(\ln \alpha) \neq 0$ the conditions of Theorem 10.1 are still satisfied, so that $F$ is Bernoulli. On the other hand, in case $\mu(\ln \alpha) = 0$ it seems likely that combining the methods of [20] and [29] one can show that $F$ is **not** Bernoulli.

We see that in both cases (a) and (b) our main theorem gives optimal results.

## Appendix A. The Pesin Theory

A.1. **Estimates of norms.** We start with discussing L1. and L2. in Lemma 2.4:

**Lemma A.1.** *For every $\delta < \delta_0$, $L_{x,\delta}$ in Lemma 2.4 can be chosen so that $L_x : \mathbb{R}^D \to T_x M$ are isometries between $|\cdot|'_{x,\delta}$ and euclidean metric in $\mathbb{R}^D$ and so that for every $\tau > 0$, on $P_\tau$, $x \to L_{x,\delta}$ is $\alpha_2$- Hölder.*

*Proof.* By Lemma 2.3, $E^{cs}(x)$ and $E^u(x)$ are $\alpha_1$-Hölder continuous on $P_\tau$ (and hence also $\alpha_2$- Hölder continuous since $\alpha_2 \leqslant \alpha_1$). Taking a Hölder continuous varying basis for $E^{cs}(x)$ and $E^u(x)$ and applying Gram-Schmidt orthogonalization gives the Hölder continuity of $L_x$. $\qquad\square$

**Lemma A.2.** *Let $K(x) > 0$ be a measurable function. If $\sum_{n \in \mathbb{Z}} K(f^n(x))e^{-\varepsilon|n|} < \infty$ $x$ a.e. then there is positive measurable function $K(x; \varepsilon)$ such that $K(x) \leqslant K(x; \varepsilon)$ and $e^{-\varepsilon} K_\varepsilon(x) \leqslant K_\varepsilon(f(x)) \leqslant e^\varepsilon K_\varepsilon(x)$ for a.e. $x$.*

*Proof.* Define $K_\varepsilon(x) := \sum_{n \in \mathbb{Z}} K(f^n(x))e^{-\varepsilon|n|}$. $\qquad\square$

We shall use this lemma to assume that the function $\mathfrak{R}_\delta(x)$ from Lemma 2.2 can be taken sufficiently large. Namely given $\delta > 0$, $M > 0$ and recalling that by Theorem 2.2 a.e. $x$ is $(\lambda, \delta/2)$-regular, we set $\widetilde{\mathfrak{R}}_\delta(x) = \max(M, \mathfrak{R}_{\delta/2}(x))$ and $\mathfrak{R}'_\delta = \widetilde{\mathfrak{R}}_\delta(x; \delta)$. Then $\mathfrak{R}'_\delta$ satisfies all the properties of Definition 2.1 and additionally $\mathfrak{R}'_\delta(x) > M$. Thus we will assume in the arguments below that $\mathfrak{R}$ is sufficiently large (this will entail that the size $\mathfrak{r}_\delta$ of Pesin charts defined by (2.3) is sufficiently small).

In the considerations below we will omit $\delta$ in the notation for $|\cdot|'_{x,\delta}$ .

**Lemma A.3.** *Let $x$ be a $(\lambda, \delta/4)$-Lyapunov regular point, where $\delta$ is sufficiently small. For $w \in T_x M$ we have*

$$\frac{1}{\sqrt{2}} |w|_x \leqslant |w|'_x \leqslant \frac{\sqrt{2}(\mathfrak{R}_{\delta/4}(x))^2}{\sqrt{1 - e^{-\delta}}} |w|_x,$$

(A.1) $\qquad |D_x f u|'_{fx} \geqslant e^{\lambda - \delta} |u|'_x$ *for* $u \in E^u(x)$, $\qquad |D_x f v|'_{fx} \leqslant e^{\delta} |v|'_x$ *for* $v \in E^{cs}(x)$

*and*

$$\|D_x f\|'_{x \to fx} \leqslant \sqrt{2} \|Df\|_{C^0}, \qquad \|(D_x f)^{-1}\|'_{fx \to x} \leqslant \sqrt{2} \|Df^{-1}\|_{C^0}.$$

*Proof.* Clearly $|v|_x \leqslant |v|_x$ for $v \in E^{cs}$ and $v \in E^u$. Now, if $w = w^{cs} + w^u$, then

$$|w|_x \leqslant |w^{cs}|_x + |w^{cs}|_x \leqslant |w^{cs}|'_x + |w^{cs}|'_x \leqslant \sqrt{2} |w|'_x.$$

Since $x$ is $(\lambda, \delta/4)$- Lyapunov regular, $\mathfrak{R}_{\delta/4}(f^k x) \leqslant e^{\delta/4|k|} \mathfrak{R}(x)$ and for $v \in E^u(x)$, $k \leqslant 0$,

$$|D_x f^k v|_{f^k x} \leqslant \mathfrak{R}_{\delta/4}(f^k(x)) e^{k\lambda - k\delta/4} |v|_x.$$

Hence $|D_x f^k v|_{f^k x} \leqslant \mathfrak{R}_{\delta/4}(x) e^{k\lambda - k\delta/2} |v|_x$ and

$$|v|_x'^2 \leqslant \sum_{m \leqslant 0} (\mathfrak{R}_{\delta/4}(x))^2 e^{2m(\lambda - \delta/2)} e^{-2\lambda m + 2\delta m} |v|_x^2 = \frac{(\mathfrak{R}_{\delta/4}(x))^2}{1 - e^{-\delta}} |v|_x^2.$$

A similar computation gives that for $v \in E^{cs}$, $|v|_x'^2 \leqslant \frac{(\mathfrak{R}_{\delta/4}(x))^2}{1 - e^{-\delta}} |v|_x^2$.

Finally, if we write $w = v^{cs} + v^u$, then

(A.2) $\qquad |w|_x'^2 = |v^{cs}|_x'^2 + |v^u|_x'^2 \leqslant \frac{(\mathfrak{R}_{\delta/4}(x))^2}{1 - e^{-\delta}} (|v^{cs}|_x^2 + |v^u|_x^2).$

Next denoting by $\theta$ the angle between $v^{cs}$ and $v^u$ and recalling that $\theta \geqslant \mathfrak{R}_{\delta/4}(x)$ since $x$ is $(\lambda, \delta/4)$–Lyapunov regular, we obtain

$$|v^{cs} + v^u|_x^2 = |v^{cs}|_x^2 + |v^u|_x^2 + 2|v^{cs}|_x |v^u|_x \cos\theta \geqslant \left(|v^{cs}|_x^2 + |v^u|_x^2\right)(1 - \cos\theta) + \cos\theta(|v^{cs}|_x + |v^u|_x)^2$$

(A.3) $\qquad \geqslant \left(|v^{cs}|_x^2 + |v^u|_x^2\right)(1 - \cos\theta) \geqslant \frac{|v^{cs}|_x^2 + |v^u|_x^2}{2(\mathfrak{R}_{\delta/4}(x))^2}.$

Combining (A.2) and (A.3) we get

$$|w|_x'^2 \leqslant \frac{(\mathfrak{R}_{\delta/4}(x))^2}{1 - e^{-\delta}} 2(\mathfrak{R}_{\delta/4}(x))^2 (|v^{cs} + v^u|_x^2) = \frac{2(\mathfrak{R}_{\delta/4}(x))^4}{1 - e^{-\delta}} |v|_x^2.$$

Now let us bound $\|D_x f\|'_{x \to fx}$ from above, all the other inequalities follow essentially the same approach and can be recovered from the computation that follows.

Notice that since $E^{cs}(x)$ and $E^u(x)$ are orthogonal w.r.t. the $|\cdot|'_x$ metric and $E^{cs}(fx)$ and $E^u(fx)$ are orthogonal w.r.t. the $|\cdot|'_{fx}$ metric and $D_x f$ maps $E^{cs}(x)$ to $E^{cs}(fx)$ and $E^u(x)$ to $E^u(fx)$, once we control the norms of $D_x f|_{E^{cs}(x)}$ and $D_x f|_{E^u(x)}$, we can easily bound $D_x f$. Indeed, if $\|D_x f|_{E^{cs}(x)}\|'_{x \to fx} \leqslant A$, $\quad \|D_x f|_{E^u(x)}\|'_{x \to fx} \leqslant A$ and $v \in T_x M$, $v = v^{cs} + v^u$, then

$$|D_x f v|_{fx}'^2 = |D_x f v^{cs}|_{fx}'^2 + |D_x f v^u|_{fx}'^2 \leqslant A^2 |v^{cs}|_x'^2 + A^2 |v^u|_x'^2 = A^2 |v|_x'^2.$$

So, take $v \in E^u(x)$. Then

$$
\begin{aligned}
|D_x f v|_{fx}'^2 &= \sum_{m \leqslant 0} |D_{fx} f^m (D_x f v)|_{f^{m+1}x}^2 e^{-2\lambda m + 2\delta m} \\
&= \sum_{m \leqslant 0} |D_x f^{m+1} v|_{f^{m+1}x}^2 e^{-2\lambda(m+1) + 2\delta(m+1)} e^{2\lambda - 2\delta} \\
&= e^{2\lambda - 2\delta} \left( |D_x f v|_{fx}^2 e^{-2\lambda + 2\delta} + \sum_{m \leqslant 0} |D_x f^m v|_{f^m x}^2 e^{-2\lambda m + 2\delta m} \right) \\
&= |D_x f v|_{fx}^2 + e^{2\lambda - 2\delta} |v|_x'^2 \leqslant \|Df\|_{C^0}^2 |v|_x^2 + e^{2\lambda - 2\delta} |v|_x'^2 \\
&\leqslant \left( \|Df\|_{C^0}^2 + e^{2\lambda - 2\delta} \right) |v|_x'^2 \leqslant 2\|Df\|_{C^0}^2 |v|_x'^2
\end{aligned}
$$

where we are using that $e^{\lambda - \delta} \leqslant \|Df\|_{C^0}$ since otherwise all Lyapunov exponents of $f$ would be smaller than $\lambda - \delta$.

Take $v \in E^{cs}(x)$ now and let us bound $|D_x f v|_{fx}'^2$,

$$
|D_x f v|_{fx}'^2 = \sum_{m \geqslant 0} |D_{fx} f^m (D_x f v)|_{f^{m+1}x}^2 e^{-2\delta m} = \sum_{m \geqslant 0} |D_x f^{m+1} v|_{f^{m+1}x}^2 e^{-2\delta(m+1)} e^{2\delta}
$$

$$
= e^{2\delta} \left( \sum_{m \geqslant 0} |D_x f^m v|_{f^m x}^2 e^{-2\delta m} - |v|_x^2 \right) = e^{2\delta} (|v|_x'^2 - |v|_x^2) \leqslant e^{2\delta} |v|_x'^2 \leqslant \|Df\|_{C^0}^2 |v|_x'^2
$$

where we are using that $e^{\delta} \leqslant \|Df\|_{C^0}$, since we are supposing that $\delta$ is sufficiently small, in particular $\delta < \lambda$. Observe that the bound for vectors on $E^u$ and $E^{cs}$ is not completely symmetric, that is why we wrote both. The bound for $\|(D_x f)^{-1}\|_{fx \to x}'$ is symmetric to $\|D_x f\|_{x \to fx}'$ and hence we omit it. $\qquad \square$

A.2. **Proof of Lemma 2.6.** For simplicity we will omit the $\delta$ from the notation for $h_x, \tilde{f}_x$ etc. We start with the following lemma:

**Lemma A.4.** *For $p \in Lyapreg$ and $\bar{z} \in \mathbb{R}^D$ let us write*

$$
\left( D_{\bar{z}} \tilde{f}_p \right)^{-1} = \begin{pmatrix} (A_p^{uu})^{-1} + \hat{T}_{\bar{z},p}^{uu} & \hat{T}_{\bar{z},p}^{cu} \\ \hat{T}_{\bar{z},p}^{uc} & (A_p^{cc})^{-1} + \hat{T}_{\bar{z},p}^{cc} \end{pmatrix}
$$

*w.r.t. the splitting $\mathbb{R}^u \times \mathbb{R}^{cs}$. Let*

(A.4) $$\mathbf{K} := 2 \max(\|Df\|_{C^0}, \|Df^{-1}\|_{C^0}).$$

*We have*

*(1)* $\|(A_p^{uu})^{-1}\| \leqslant e^{-\lambda + \delta}$, $\|D_{\bar{y}} \tilde{f}_p\| \leqslant \mathbf{K}$ *for every $\bar{y} \in \mathbb{R}^D$;*

*(2)* $\|A_p^{cc}\| \leqslant e^{\delta}$, $\|(D_{\bar{y}} \tilde{f}_p)^{-1}\| \leqslant \mathbf{K}$ *for every $\bar{y} \in \mathbb{R}^D$;*

*(3)* $H\ddot{o}l_{\alpha_2}(\hat{T}_{\bar{z},p}^{ab}) \leqslant \delta$ *for $a, b = u, c$ in all its combinations, i.e.*

$$
\|\hat{T}_{\bar{z},p}^{ab} - T_{\bar{z}',p}^{ab}\| \leqslant \delta \|\bar{z} - \bar{z}'\|^{\alpha_2},
$$

*where*

(A.5) $$\alpha_2 = \min(\alpha/2, \alpha_1);$$

*(4)* $\hat{T}_{\bar{z},p}^{ab} = 0$ *for $\bar{z} = 0$ and for $|\bar{z}| \geqslant 1$ for $a, b = u, c$ in all its combinations.*

*Proof of Lemma A.4.* Let $\hat{f}_p : T_pM \to T_{fp}M$ be so that $\tilde{f}_p = L_{fp}^{-1} \circ \hat{f}_p \circ L_p$. Thus $\hat{f}_p(z) = \exp_{fp}^{-1} \circ f \circ \exp_p(z)$ if $\|z\| \leqslant \mathfrak{r}_\delta(p)$ and $\hat{f}_p(z) = D_p f(z)$ if $\|z\| > 2\mathfrak{r}_\delta(p)$. Since $L_x$ is a linear isometry from $\mathbb{R}^D$ with euclidean metric and $T_xM$ with $|\cdot|_x'$ metric, we can prove the statements for $\hat{f}_p$ w.r.t. $|\cdot|_p'$ and $|\cdot|_{fp}'$. Let us write

$$\left(D_{\hat{z}}\hat{f}_p\right)^{-1} = \begin{pmatrix} (\hat{A}_p^{uu})^{-1} + \hat{\hat{T}}_{\hat{z},p}^{uu} & \hat{\hat{T}}_{\hat{z},p}^{cu} \\ \hat{\hat{T}}_{\hat{z},p}^{uc} & (\hat{A}_p^{cc})^{-1} + \hat{\hat{T}}_{\hat{z},p}^{cc} \end{pmatrix}$$

w.r.t. the splitting $E^u(fp) \oplus E^{cs}(fp) \to E^u(p) \oplus E^{cs}(p)$. In particular

$$\hat{A}_p^{uu} = D_p f | E^u(p) \quad \text{and} \quad \hat{A}_p^{cc} = D_p f | E^{cs}(p).$$

Now (A.1) gives the bound on $(A_p^{uu})^{-1}$ and $A_p^{cc}$.

Now we bound the Hölder norm. Combining (2.2) with Lemma A.3 we get that

$$\|(D_{\hat{z}_1}\hat{f}_p)^{-1} - (D_{\hat{z}_2}\hat{f}_p)^{-1}\|_{T_{fp}M \to T_pM}' \leqslant \frac{(\mathfrak{R}_{\delta/4}(p))^2}{\sqrt{1-e^{-\delta}}}\sqrt{2}C_1|\hat{z}_1 - \hat{z}_2|_p^\alpha \leqslant$$

$$\frac{(\mathfrak{R}_{\delta/4}(p))^2}{\sqrt{1-e^{-\delta}}}(\sqrt{2})^{1+\alpha}C_1|\hat{z}_1 - \hat{z}_2|_p'^\alpha \leqslant \frac{(\mathfrak{R}_{\delta/4}(p))^2}{\sqrt{1-e^{-\delta}}}(\sqrt{2})^{1+\alpha}C_1(\max_{i=1,2}|\hat{z}_i|'^{\alpha/2})|\hat{z}_1 - \hat{z}_2|_p'^{\alpha/2}$$

if $|\hat{z}_1|_p, |\hat{z}_2|_p \leqslant \frac{1}{C_1}$.

Finally (see (2.3)), if $|\hat{z}_1|_p', |\hat{z}_2|_p' \leqslant \mathfrak{r}_\delta(p)$ then

$$\|(D_{\hat{z}_1}\hat{f}_p)^{-1} - (D_{\hat{z}_2}\hat{f}_p)^{-1}\|_{T_{fp}M \to T_pM}' \leqslant \delta|\hat{z}_1 - \hat{z}_2|_p'^{\alpha/2}.$$

Since $E^{cs}(p)$ and $E^u(p)$ are orthogonal w.r.t. $|\cdot|_p'$ this gives the bound for $\hat{\hat{T}}_{\hat{z},p}^{ab}$.
Now we can bound also

$$\begin{aligned} \|(D_{\hat{z}}f_p)^{-1}\|_{fp \to p}' &\leqslant \|(D_0 f_p)^{-1}\|_{fp \to p}' + \|(D_0 f_p)^{-1} - (D_{\hat{z}}f_p)^{-1}\|_{fp \to p}' \\ &\leqslant \sqrt{2}\|Df\|_{C^0} + 2^{\alpha/2}\delta\mathfrak{r}_\delta^{\alpha/2}(p) \leqslant 2\|Df\|_{C^0}, \end{aligned}$$

by taking $\delta$ small enough. Similarly we can bound Hölder constants for $D_{\bar{z}}f_p$ and its norm. $\square$

We shall assume that $\delta > 0$ is small enough in terms of $\mathbf{K}$ given by (A.4) and $\lambda$ such that the following inequalities are satisfied:

$$\frac{3\delta}{1-e^{-\lambda+\sqrt{\delta}}} \leqslant \frac{1}{2\mathbf{K}^2}, \quad (e^{-\lambda+\delta} + \delta)e^{2\delta} \leqslant e^{-\lambda+\sqrt{\delta}},$$

and

$$e^{-\lambda+\sqrt{\delta}} + 2\delta < 1, \quad \left(1 + e^{-\lambda+\sqrt{\delta}}\right)e^{3\delta} < 2, \quad \frac{1}{1-2\delta e^\delta} \leqslant e^{3\delta}, \quad \delta < \frac{1}{100}.$$

We assume also that the number $\alpha_6 > 0$ is small so that

$$\mathbf{K}^{\alpha_6}e^{-\lambda+\sqrt{\delta}} < 1.$$

**Lemma A.5.** *Let $p \in LyapReg(\delta)$ where $\delta$ is sufficiently small. Given $\bar{z} \in \mathbb{R}^D$ and a map $L : \mathbb{R}^{cs} \to \mathbb{R}^u$ with $\|L\|_{C^0} \leqslant 1$, we can define $\Gamma_{cs,p,\bar{z}}(L) : \mathbb{R}^{cs} \to \mathbb{R}^u$ so that*

$$\mathrm{graph}\Big(\Gamma_{cs,p,\bar{z}}(L)\Big) = \Big(D_{\bar{z}}\tilde{f}_p\Big)^{-1}(\mathrm{graph}(L)).$$

*Moreover, for every $\bar{z}, \bar{z}_i \in \mathbb{R}^D$ and $\|L\|_{C^0}, \|L_i\|_{C^0} \leqslant 1$, $i = 1, 2$, then*

$$(A.6) \qquad \|\Gamma_{cs,p,\bar{z}}(L)\|_{C^0} \leqslant e^{-\lambda+\sqrt{\delta}}\|L\|_{C^0} + 2\delta \min\{1, |\bar{z}|^{\alpha_6}\}$$

*and*

$$(A.7) \quad \|\Gamma_{cs,p,\bar{z}_1}(L_1) - \Gamma_{cs,p,\bar{z}_2}(L_2)\|_{C^0} \leqslant e^{-\lambda+\sqrt{\delta}}\|L_1 - L_2\|_{C^0} + 6\delta|\bar{z}_1 - \bar{z}_2|^{\alpha_6}.$$

*Proof.* Denote

$$A_{\bar{z},p,L}^{cc} = \left(A_p^{cc}\right)^{-1} + \hat{T}_{\bar{z},p}^{cc} + \hat{T}_{\bar{z},p}^{uc} \circ L,$$

and let

$$\Gamma_{cs,p,\bar{z}}(L) := \left(\left(\left(A_p^{uu}\right)^{-1} + \hat{T}_{\bar{z},p}^{uu}\right) \circ L + \hat{T}_{\bar{z},p}^{cu}\right) \circ (A_{\bar{z},p,L}^{cc})^{-1}.$$

For simplicity of notation we denote $\hat{L} = \Gamma_{cs,p,\bar{z}}(L)$, $\|\cdot\| = \|\cdot\|_{C^0}$. Notice that $(D_{\bar{z}}\tilde{f}_p)^{-1}(\mathrm{graph}(L)) = \mathrm{graph}(\hat{L})$.

By Lemma A.4(2)

$$(A.8) \quad \|(A_{\bar{z},p,L}^{cc})^{-1}\| \leqslant \frac{e^\delta}{1 - e^\delta(\|\hat{T}_{\bar{z},p}^{cc}\| + \|\hat{T}_{\bar{z},p}^{cc}\|\|L\|)} \leqslant \frac{e^\delta}{1 - e^\delta\delta \min\{1, |\bar{z}|^{\alpha_2}\}(1 + \|L\|)}.$$

Since $\|L\| \leqslant 1$ and $\delta > 0$ is small enough, it follows that $\|(A_{\bar{z},p,L}^{cc})^{-1}\| \leqslant e^{4\delta}$.

Also since $\|L_i\| \leqslant 1$, by Lemma A.4(3) it follows that

$$\|A_{\bar{z}_1,p,L_1}^{cc} - A_{\bar{z}_2,p,L_2}^{cc}\| \leqslant \delta|\bar{z}_1 - \bar{z}_2|^{\alpha_2}(1 + \|L_1 - L_2\|).$$

Hence using the identity $(L')^{-1} - (L'')^{-1} = (L')^{-1}(L'' - L')(L'')^{-1}$ valid for arbitrary invertible linear maps $L', L''$ we get

$$\| \left(A_{\bar{z}_1,p,L_1}^{cc}\right)^{-1} - \left(A_{\bar{z}_2,p,L_2}^{cc}\right)^{-1} \| \leqslant e^{8\delta}\delta|\bar{z}_1 - \bar{z}_2|^{\alpha_2}(1 + \|L_1 - L_2\|).$$

Combining this estimate with (A.8) and Lemma A.4(2) we get

$$\|\hat{L}\| \leqslant \left((e^{-\lambda+\delta} + \delta)\|L\| + \delta \min\{1, |\bar{z}|^{\alpha_2}\}\right)e^{2\delta} \leqslant e^{-\lambda+\sqrt{\delta}}\|L\| + 2\delta \min\{1, |\bar{z}|^{\alpha_2}\}$$

which gives (A.6).

Moreover,

$$\|\Gamma_{cs,p,\bar{z}_1}(L_1) - \Gamma_{cs,p,\bar{z}_2}(L_2)\|$$
$$= \| \left(\left(\left(A_p^{uu}\right)^{-1} + \hat{T}_{\bar{z}_1,p}^{uu}\right) \circ L_1 + \hat{T}_{\bar{z}_1,p}^{cu}\right) \circ (A_{\bar{z}_1,p,L_1}^{cc})^{-1}$$
$$- \left(\left(\left(A_p^{uu}\right)^{-1} + \hat{T}_{\bar{z}_2,p}^{uu}\right) \circ L_2 + \hat{T}_{\bar{z}_2,p}^{cu}\right) \circ (A_{\bar{z}_2,p,L_2}^{cc})^{-1}\|$$
$$\leqslant \| \left(\left(\left(A_p^{uu}\right)^{-1} + \hat{T}_{\bar{z}_1,p}^{uu}\right) \circ L_1 + \hat{T}_{\bar{z}_1,p}^{cu}\right) \circ \left((A_{\bar{z}_1,p,L_1}^{cc})^{-1} - (A_{\bar{z}_2,p,L_2}^{cc})^{-1}\right) \|$$
$$+ \| \left(\left(\left(A_p^{uu}\right)^{-1} + \hat{T}_{\bar{z}_1,p}^{uu}\right) \circ L_1 + \hat{T}_{\bar{z}_1,p}^{cu}\right) - \left(\left(\left(A_p^{uu}\right)^{-1} + \hat{T}_{\bar{z}_2,p}^{uu}\right) \circ L_2 + \hat{T}_{\bar{z}_2,p}^{cu}\right)\right) \circ (A_{\bar{z}_2,p,L_2}^{cc})^{-1}\|$$

$$\leqslant \quad \left\| \left( (A_p^{uu})^{-1} + \hat{T}_{\bar{z}_1,p}^{uu} \right) \circ L_1 + \hat{T}_{\bar{z}_1,p}^{cu} \right\| \left\| (A_{\bar{z}_1,p,L_1}^{cc})^{-1} - (A_{\bar{z}_2,p,L_2}^{cc})^{-1} \right\|$$

$$+ \quad \left\| \left( \left( (A_p^{uu})^{-1} + \hat{T}_{\bar{z}_1,p}^{uu} \right) \circ L_1 + \hat{T}_{\bar{z}_1,p}^{cu} \right) - \left( \left( (A_p^{uu})^{-1} + \hat{T}_{\bar{z}_2,p}^{uu} \right) \circ L_2 + \hat{T}_{\bar{z}_2,p}^{cu} \right) \right\| \left\| (A_{\bar{z}_2,p,L_2}^{cc})^{-1} \right\|$$

$$\leqslant \quad e^{8\delta} \delta |\bar{z}_1 - \bar{z}_2|^{\alpha_2} (1 + \|L_1 - L_2\|) + \left( (e^{-\lambda+\delta} + \delta)\|L_1 - L_2\| + 2\delta|\bar{z}_1 - \bar{z}_2|^{\alpha_2} \right) e^{2\delta}$$

$$\leqslant e^{-\lambda+\sqrt{\delta}} \|L_1 - L_2\| + 6\delta |\bar{z}_1 - \bar{z}_2|^{\alpha_2} \text{ proving (A.7).} \qquad \square$$

**Lemma A.6.** *Let* $p \in LyapReg$ *and let* $\bar{z}_1, \bar{z}_2 \in \mathbb{R}^D$, *then*

$$|\tilde{f}_p(\bar{z}_1) - \tilde{f}_p(\bar{z}_2)| \leqslant 2\mathbf{K}|\bar{z}_1^u - \bar{z}_2^u| + e^{2\delta}|\bar{z}_1^{cs} - \bar{z}_2^{cs}|,$$

*where* $\mathbf{K} = 2\max(\|Df\|_{C^0}, \|Df^{-1}\|_{C^0})$ *(see (A.4)).*
*Moreover, if* $|\bar{z}_1^u - \bar{z}_2^u| \leqslant \frac{1}{2\mathbf{K}^2}|\bar{z}_1^{cs} - \bar{z}_2^{cs}|$ *then*

$$|\tilde{f}_p(\bar{z}_1) - \tilde{f}_p(\bar{z}_2)| \leqslant e^{3\delta}|\bar{z}_1 - \bar{z}_2|.$$

*Proof.* The proof is a simple consequence of the mean value theorem and Lemma A.4.
$\square$

For $n \in \mathbb{N}$ and $k \leqslant n$ let $L_k^{cs} : \mathbb{R}^{cs} \to \mathbb{R}^u$ be such that

$$\left( D_{\tilde{f}_{f^k x}^{(n-k)}(\bar{z})} \tilde{f}_{f^{n-k}x}^{(k)} \right)^{-1} (\mathbb{R}^{cs}) = \mathrm{graph}(L_k).$$

*Proof of Lemma 2.6.* We define inductively $\eta_k : \mathbb{R}^{cs} \to \mathbb{R}^u$ by letting $\eta_0$ to be the constant map $\eta_0(\bar{z}) \equiv \left( \tilde{f}_x^{(n)}(\bar{y}) \right)^u$ and setting

$$\mathrm{graph}(\eta_k) = \left( \tilde{f}_{f^{n-k}x} \right)^{-1} (\mathrm{graph}(\eta_{k-1}))$$

for $1 \leqslant k \leqslant n$. Then we define $\tilde{\eta}_{x,\bar{y}}^{cs,n} = \eta_n$. We have that $\mathrm{graph}(L_k)$ is the tangent plane to the $\mathrm{graph}(\eta_k)$ at the point $\bar{z}_k = \tilde{f}_{f^n x}^{-k}(\bar{z})$,

$$(\bar{z}^{cs}, \eta_k(\bar{z}^{cs})) = \left( \tilde{f}_{f^k x}^{(n-k)} \right) (\bar{z}, \eta_0(\bar{z})).$$

Applying (A.6) inductively to $L_k$ (with $L_0 = 0$), we get

$$\|D\eta_k\|_{C^0} \leqslant e^{-\lambda+\sqrt{\delta}} \|D\eta_{k-1}\|_{C^0} + 3\delta$$

and hence

$$\|D\eta_k\|_{C^0} \leqslant \frac{3\delta}{1 - e^{-\lambda+\sqrt{\delta}}}$$

for all $k \in [0, n]$.

Recall that $\delta$ is such that $\frac{3\delta}{1-e^{-\lambda+\sqrt{\delta}}} \leqslant \frac{1}{2\mathbf{K}^2}$. Note that for $\bar{z}_1, \bar{z}_2 \in \mathrm{graph}(\eta_n) = W_x^{cs,n}(\bar{y})$, we have $|\bar{z}_1^u - \bar{z}_2^u| \leqslant \frac{|\bar{z}_1^{cs} - \bar{z}_2^{cs}|}{2\mathbf{K}^2}$. Hence Lemma A.6 gives that for $k \in [0, n]$,

$$|\tilde{f}_x^{(n-k+1)}(\bar{z}_1) - \tilde{f}_x^{(n-k+1)}(\bar{z}_2)| \leqslant e^{3\delta} |\tilde{f}_x^{(n-k)}(\bar{z}_1) - \tilde{f}_x^{(n-k)}(\bar{z}_2)|.$$

Let $L_i^{(k)}$ satisfy

$$\left( D_{\tilde{f}_{f^k x}^{(n-k)}(\bar{z}_i)} \tilde{f}_{f^{n-k}x}^{(k)} \right)^{-1} (\mathbb{R}^{cs}) = graph(L_i^{(k)}).$$

Assuming inductively that (2.5) holds for $k-1$, that is

$$\|L_1^{(k-1)}(\bar{z}_1) - L_1^{(k-1)}(\bar{z}_2)\| \leqslant \frac{12\delta}{1-e^{-\lambda+\sqrt{\delta}}}|\tilde{f}_x^{(n-k+1)}(\bar{z}_1) - \tilde{f}_x^{(n-k+1)}(\bar{z}_2)|^{\alpha_6},$$

(note that (2.5) trivially holds for $k=0$) and applying (A.7) we get

$$
\begin{aligned}
\|L_1^{(k)}(\bar{z}_1) - L_1^{(k)}(\bar{z}_2)\| &\leqslant e^{-\lambda+\sqrt{\delta}}\|L_1^{(k-1)}(\bar{z}_1) - L_1^{(k-1)}(\bar{z}_2)\| + 6\delta|\tilde{f}_x^{(n-k+1)}(\bar{z}_1) - \tilde{f}_x^{(n-k+1)}(\bar{z}_2)|^{\alpha_6} \\
&\leqslant \left(\frac{12\delta e^{-\lambda+\sqrt{\delta}}}{1-e^{-\lambda+\sqrt{\delta}}} + 6\delta\right)|\tilde{f}_x^{(n-k+1)}(\bar{z}_1) - \tilde{f}_x^{(n-k+1)}(\bar{z}_2)|^{\alpha_6} \\
&\leqslant \left(\frac{2e^{-\lambda+\sqrt{\delta}}}{1-e^{-\lambda+\sqrt{\delta}}} + 1\right)6\delta e^{3\alpha\delta}|\tilde{f}_x^{(n-k)}(\bar{z}_1) - \tilde{f}_x^{(n-k)}(\bar{z}_2)|^{\alpha_6} \\
&= \left(1 + e^{-\lambda+\sqrt{\delta}}\right)e^{3\alpha_6\delta}\frac{6\delta}{1-e^{-\lambda+\sqrt{\delta}}}|\tilde{f}_x^{(n-k)}(\bar{z}_1) - \tilde{f}_x^{(n-k)}(\bar{z}_2)|^{\alpha_6} \\
&\leqslant \frac{12\delta}{1-e^{-\lambda+\sqrt{\delta}}}|\tilde{f}_x^{(n-k)}(\bar{z}_1) - \tilde{f}_x^{(n-k)}(\bar{z}_2)|^{\alpha_6} \qquad \qquad \square
\end{aligned}
$$

**Corollary A.7.** *Let $x \in LyapReg$. If $\bar{z}_1, \bar{z}_2$ belong to the same $\widetilde{W}_x^{cs,n}$ leaf then*

$$|\tilde{f}_x^{(k)}(\bar{z}_1) - \tilde{f}_x^{(k)}(\bar{z}_2)| \leqslant e^{3k\delta}|\bar{z}_1 - \bar{z}_2|$$

*for $k \in [0, n]$.*

*Proof.* This follows by applying inductively the second inequality of Lemma A.6. $\square$

### A.3. **Hölder foliations.**

*Proof of Lemma 7.2.* Recall the coordinates $(a, b)$ introduced in (5.1). Let $\tilde{b} = b - \eta_z(a)$. In coordinates $(a, \tilde{b})$ $\mathcal{W}$ is a graph of a function $\tilde{b} = \tilde{\eta}(a)$ where $\tilde{\eta}$ is defined for $\|a\| \leqslant \xi$ (since $\partial\mathcal{W} \bigcap B_i(r) = \emptyset$) and moreover there is $\bar{a}$ with $\|\bar{a}\| \leqslant (1-\hat{\varepsilon})\xi$ such that $\|\tilde{\eta}(\bar{a})\| \leqslant (1-\hat{\varepsilon})r$. By Hölder continuity of $E^u$ on $P_\tau$ (Lemma 2.3) it follows that on the ball of radius $\xi$

$$\|\tilde{\eta}(a) - \tilde{\eta}(\bar{a})\| \leqslant K(\tau)\xi^{1+\alpha_2},$$

whence

$$\|\tilde{\eta}(a)\| \leqslant (1-\hat{\varepsilon})r + K(\tau)\xi^{1+\alpha_2} \leqslant r.$$

Next, pick $y \in \widehat{\mathcal{W}}$. By Lemma 2.6, $\mathcal{F}(y)$ is given by a graph of a function $a = \phi(b)$ with $\nabla\phi$ small. Applying the Implicit Function Theorem to the equation $a = \phi(\tilde{b} + \eta_z(a))$ we see that $\mathcal{F}(y)$ can also be given by the equation $a = \tilde{\phi}(\tilde{b})$ with $\nabla\tilde{\phi}$ small. Since $y \in (1-\hat{\varepsilon})B_i$ we have $\|\tilde{\phi}(0)\| \leqslant (1-\varepsilon)\xi$. Therefore on the ball of radius $r$

$$\|\tilde{\phi}(\tilde{b})\| \leqslant (1-\varepsilon)\xi + Cr \leqslant \xi.$$

Now the fact that $\mathcal{W}$ and $\mathcal{F}(y)$ intersect follows from transversality. More precisely, we need to show that the system

$$\tilde{b} = \tilde{\eta}(a), \quad a = \tilde{\phi}(\tilde{b})$$

has the unique solution. The existence of the solution follows from the Fixed Point Theorem applied to map of $\{\|a\| \leqslant \xi\} \times \{\|\tilde{b}\| \leqslant r\}$ defined by $\Phi(a, b) = (\tilde{\phi}(b), \tilde{\eta}(a))$.

The uniqueness follows since if there were two intersection points, then the line joining them would belong to both $\mathcal{C}_{\mathfrak{b}}^u$ and $\mathcal{C}_{\mathfrak{b}}^{cs}$. Alternatively, the Hölder regularity of $E^u$ and $E^{cs}$ (see (2.5) and A.8) implies that $\Phi$ is a contraction if both $r$ and $\xi$ are small. $\qquad\square$

A.4. **Hadamard Perron, unstable foliation.** In this subsection we state the results analogous to ones proven in §A.2 but with the unstable direction instead of the center stable directions. The proofs for the unstable direction are analogous to the proofs for the center stable direction and hence we will omit them.

**Lemma A.8.** *There exists* $\alpha_4, \tilde{\delta}_f > 0$ *such that the following holds. Fix* $\delta \in (0, \tilde{\delta}_f)$, *and* $x \in LyapReg$. *Then for every* $\bar{y} \in \mathbb{R}^D$, *there is* $\tilde{\eta}_{x,\bar{y}}^u : \mathbb{R}^u \to \mathbb{R}^{cs}$ *such that*

$$\widetilde{W}_x^u(\bar{y}) = graph\left(\tilde{\eta}_{x,\bar{y}}^u\right),$$

$$\|D\tilde{\eta}_{x,\bar{y}}^u\|_{C^0} \leqslant \frac{3\delta}{1 - e^{-\lambda + \sqrt{\delta}}}$$

*and*

$$[D\tilde{\eta}_{x,\bar{y}}^u]_{C^{\alpha_4}} \leqslant \frac{12\delta}{1 - e^{-\lambda + \sqrt{\delta}}}.$$

**Corollary A.9.** *If* $\bar{z}_1, \bar{z}_2$ *belong to the same* $\widetilde{W}_{f^k x}^u$ *leaf then*

$$|(\tilde{f}_x^{(k)})^{-1}(\bar{z}_1) - (\tilde{f}_x^{(k)})^{-1}(\bar{z}_2)| \leqslant e^{k(-\lambda + 2\delta)}|\bar{z}_1 - \bar{z}_2|$$

*for* $k \in [0, n]$.

We finish with the following lemma:

**Lemma A.10.** *There exists* $\alpha_7 > 0$ *such that the following holds: Let* $\bar{z} \in \mathbb{R}^D$ *and* $E = graph(L)$ *where* $L : \mathbb{R}^u \to \mathbb{R}^{cs}$ *and* $\|L\| \leqslant \mathbf{q} \leqslant 1$. *Then* $D_{\bar{z}}\tilde{f}_x^{(k)}(E) = graph(L_k)$ *where* $L_k : \mathbb{R}^u \to \mathbb{R}^{cs}$ *and* $\|L_k\| \leqslant 2\mathbf{q}$. *Moreover, if we call* $\bar{z}_k = \tilde{f}_x^{(k)}(\bar{z})$ *then*

$$\|L_k\| \leqslant e^{k(-\lambda + \sqrt{\delta})}\|L\| + 2\delta \sum_{i=1}^{k} e^{(k-i)(-\lambda + \sqrt{\delta})} \min\{1, |\bar{z}_k|^{\alpha_7}\}.$$

*Finally if* $L_i$ *and* $\bar{z}_i$ *are as above,* $\|L_i\| \leqslant 1$, *and we define* $L_i^{(k)}$ *as above we get that*

$$(A.9) \quad \|L_1^{(k)} - L_2^{(k)}\| \leqslant e^{k(-\lambda + \sqrt{\delta})}\|L_1 - L_2\| + 6\delta \sum_{i=1}^{k} e^{(k-i)(-\lambda + \sqrt{\delta})}|\tilde{f}_x^{(i)}(\bar{z}_1) - \tilde{f}_x^{(i)}(\bar{z}_2)|^{\alpha_7}.$$

*Proof.* The proof is analogous to the proof of Lemma 2.6. $\qquad\square$

## Appendix B. Exponential mixing.

Recall that a $C^{1+\alpha}$ map $f$ on $M$ preserving a measure $\mu$ is exponentially mixing if there are constants $C > 0, \mathbf{r} > 0$ and $\eta_{\mathbf{r}} > 0$ such that

$$(B.1) \qquad \left| \int_M \phi(x)\psi(f^n x)d\mu - \int_M \phi\, d\mu \int_M \psi\, d\mu \right| \leqslant Ce^{-\eta_{\mathbf{r}} n}\|\phi\|_{\mathbf{r}}\|\psi\|_{\mathbf{r}},$$

for all $\phi, \psi \in C^{\mathbf{r}}(M)$. Examples of exponentially mixing maps include volume preserving Anosov diffeomorphsims [9, 38], time one maps of contact Anosov flows [35],

mostly contracting systems [14, 17] (including the examples constructed in [49], and [18]) partially hyperbolic translations on homogeneous spaces [33], and partially hyperbolic automorphisms of nilmanifolds [23]. We also note that a product of exponentially mixing diffeomorphisms is exponentially mixing (see e.g. [21, Theorem A.4]) and, more generally, sufficient conditions for mixing of skew products with exponentially mixing base and fibers are discussed in [19] (see Theorem 10.1 of the present paper).

Here we discuss several properties of exponential mixing maps used in our proof.

**Lemma B.1.** *Suppose that $\mu$ is a smooth measure. Then if* (B.1) *holds for some $\mathbf{r} > 0$ then it holds for all $\tilde{\mathbf{r}} > 0$ (with possibly a different exponent $\eta_{\tilde{\mathbf{r}}}$).*

*Proof.* Note that if (B.1) holds for some $\mathbf{r} > 0$ then it holds for all larger $\tilde{\mathbf{r}}$ with $\eta_{\tilde{\mathbf{r}}} = \eta_{\mathbf{r}}$. Therefore it suffices to prove that (B.1) holds for all sufficiently small $\tilde{\mathbf{r}}$. So we will assume below that $\tilde{\mathbf{r}} \leqslant 1$. Note that given a function $\phi$ and $\varepsilon > 0$ we can find a function $\phi_\varepsilon$ such that $\|\phi_\varepsilon\|_{\mathbf{r}} \leqslant C_1 \varepsilon^{-\mathbf{r}} \|\phi\|_{\tilde{\mathbf{r}}}$ and $\|\phi - \phi_\varepsilon\|_{C^0} \leqslant C_1 \varepsilon^{\tilde{\mathbf{r}}} \|\phi\|_{\tilde{\mathbf{r}}}$. Indeed, applying a partition of unity we can assume that $\phi$ is supported in a single coordinate chart. Using the coordinates coming from that chart it suffices to prove the result for functions supported on compact domain in $\mathbb{R}^d$, in which case we can just take

$$\phi_\varepsilon(x) = \frac{1}{\varepsilon^d} \int_{\mathbb{R}^d} \phi(y) p\left(\frac{y - x}{\varepsilon}\right) dy$$

where $p(\cdot)$ is a probability density supported on a unit ball in $\mathbb{R}^d$.

Now given arbitrary $\phi, \psi \in C^{\tilde{\mathbf{r}}}$ let $\phi_\varepsilon$ and $\psi_\varepsilon$ denote the approximations described above. Then

$$\int_M \phi(x)\psi(f^n x)d\mu = \int_M \phi_\varepsilon(x)\int_M \psi_\varepsilon(f^n x) + O\left(\varepsilon^{\tilde{\mathbf{r}}}\|\phi\|_{\tilde{\mathbf{r}}}\|\psi\|_{\tilde{\mathbf{r}}}\right)$$

$$= \int_M \phi_\varepsilon \, d\mu \int_M \psi_\varepsilon \, d\mu + O\left(\varepsilon^{\tilde{\mathbf{r}}}\|\phi\|_{\tilde{\mathbf{r}}}\|\psi\|_{\tilde{\mathbf{r}}}\right) + O\left(\varepsilon^{-2\mathbf{r}}e^{-\eta_{\mathbf{r}} n}\|\phi\|_{\tilde{\mathbf{r}}}\|\psi\|_{\tilde{\mathbf{r}}}\right)$$

$$= \int_M \phi \, d\mu \int_M \psi \, d\mu + O\left(\varepsilon^{\tilde{\mathbf{r}}}\|\phi\|_{\tilde{\mathbf{r}}}\|\psi\|_{\tilde{\mathbf{r}}}\right) + O\left(\varepsilon^{-2\mathbf{r}}e^{-\eta_{\mathbf{r}} n}\|\phi\|_{\tilde{\mathbf{r}}}\|\psi\|_{\tilde{\mathbf{r}}}\right)$$

Taking $\varepsilon = e^{-\eta_{\mathbf{r}} n/(2\mathbf{r}+\tilde{\mathbf{r}})}$ we obtain that $f$ is exponentially mixing on $C^{\tilde{\mathbf{r}}}$ with

(B.2) $$\eta_{\tilde{\mathbf{r}}} = \frac{\tilde{\mathbf{r}}\eta_{\mathbf{r}}}{2\mathbf{r} + \tilde{\mathbf{r}}}$$

as claimed. □

Recall that parallelograms $B(\xi, r)$ are defined by (5.1).

**Lemma B.2.** *There exists $\eta_2 > 0$ such that for every $\hat{\varepsilon} > 0$ there exists $n_{\hat{\varepsilon}} \in \mathbb{N}$ such that for every $n \geqslant n_{\varepsilon'}$, every $\xi, r, \xi', r' \geqslant e^{-3\eta_2 n}$ we have*

$$\mu(f^n(B(\xi, r)) \cap B(\xi', r')) \in (1 - \hat{\varepsilon}, 1 + \hat{\varepsilon})\mu(B(\xi, r)\mu(B(\xi', r')).$$

*Proof.* The proof follows a standard argument where we approximate the characteristic function of $B(\xi, r)$ and $B(\xi', r')$ by smooth functions with a controlled error. We then deduce the result from exponential mixing for smooth functions. We provide the details to obtain an explicit expression for $\eta_2$.

Let $\psi_{\rho,\varepsilon}$ denote a function which is 1 on the cube of size $\rho$ centered at the origin, is 0 outside of the cube of size $\rho + \varepsilon$ and has Lipshitz norm of order $O(1/\varepsilon)$. Recalling coordinates $(a, b)$ introduced in (5.1) and let $\tilde{b} = b - \eta_z(a)$. Consider the function $\Psi_\varepsilon(x) = \psi_{\xi,\varepsilon}(a(x))\psi_{r,\varepsilon}(b(x))$ and let $K(\tau) := \max_{x \in P_\tau} \boldsymbol{\rho}(x)$ (see (5.2)).

Then

$$\|\Psi_\varepsilon\|_{Lip} = O(\varepsilon), \quad \|\Psi_\varepsilon - 1_{B(\xi,r)}\|_{L^2} = O\left(K(\tau)\sqrt{\varepsilon\Delta}\right),$$

where $\Delta = \xi^{d^u-1}r^{d_{cs}} + \xi^{d^u}r^{d_{cs}-1}$ and the second equality relies on (5.2). Thus letting $\Psi'_\varepsilon$ be the similar approximation for $1_{B(\xi',r')}$ we get

$$\mu(f^n(B(\xi,r)) \cap B(\xi',r')) = \mu\left(\Psi_\varepsilon\left(\Psi'_\varepsilon \circ f^n\right)\right) + O\left(K(\tau)\sqrt{\varepsilon\Delta}\right)$$

$$= \mu(\Psi_\varepsilon)\mu(\Psi'_\varepsilon) + O\left(K(\tau)\sqrt{\varepsilon\Delta}\right) + O\left(\varepsilon^{-2}e^{-\eta_1 n}\right) =$$

$$\mu(B(\xi,r)\mu(B(\xi',r')) + O\left(K(\tau)\sqrt{\varepsilon\Delta}\right) + O\left(\varepsilon^{-2}e^{-\eta_1 n}\right)$$

where $\eta_1$ is the mixing exponent for Lipschitz functions. Choosing $\varepsilon = e^{-\eta_1 n/5}/\Delta^{1/10}$ we get both $O(\cdot)$ terms are of the same order, whence

$$(B.3) \qquad \mu(f^n(B(\xi,r)) \cap B(\xi',r')) = \mu(B(\xi,r)\mu(B(\xi',r')) + O\left(K(\tau)\Delta^{2/5}e^{-\eta_1 n/5}\right).$$

We want that the first term on the RHS to be much larger than the second. Notice that $K(\tau)$ only depends on $\tau$ which depends on the $\varepsilon$, so this term will be absorbed simply by taking large enough $n$. By (5.2) the first term is of order $\xi^{2d_u}r^{2d_{cs}}$. Now a direct computation using the bound $\xi, r, \xi', r' \geqslant e^{-3\eta_2 n}$ shows that the first term in (B.3) dominates provided that

$$(B.4) \qquad \eta_2 < \eta_1 \left(\frac{D}{5} - \frac{2}{15}\right).$$

where $D = \dim(M)$ and $\eta_1$ is given by (B.2), with $\tilde{r} = 1$, that is, $\eta_1 = \begin{cases} \eta & \text{if } \mathbf{r} < 1, \\ \frac{\eta}{1+2\mathbf{r}} & \text{if } \mathbf{r} \geqslant 1. \end{cases}$

$\square$

We will finish this section with proving Lemma 3.1:

*Proof of Lemma 3.1.* Since $\#\text{supp}(\mu) > 1$, there exists $\kappa > 0$ and $p, p' \in supp(\mu)$ with $d(p, p') > \kappa$. Let $A_1, A_2$ be balls of radius $\kappa/10$ centered at $p, p'$ respectively. By definition $c' = \min(\mu(A_1), \mu(A_2)) > 0$. Let $B \subset M$ be a set with $\mu(B) \geqslant 1 - c'/2$.

We have the following claim:

**CLAIM:** There exist $\bar{\kappa} > 0$ such that for every $r > 0$, $i = 1, 2$ there exists $z_r^i \in A_i \cap B$ with

$$\mu(O_r(z_r^i)) > \bar{\kappa} \cdot vol(O_r(z_r^i)).$$

*Proof of* **CLAIM**. Consider the Besicovitch cover $\{O_r(z)\}_{z \in A_i \cap B}$ of the set $A_i \cap B$. By the Besicovitch covering theorem, $A_i \cap B \subset \bigcup_{i=1}^{c_{dimM}} \bigcup_{j \leqslant m_i} O_r(z_{ij})$, where for every $i \leqslant c_{dimM}$

and every $j, j' \leqslant m_i$, $j \neq j'$, $O_r(z_{ij}) \cap O_r(z_{ij'}) = \emptyset$. Then

$$c'/2 \leqslant \mu(A_i \cap B) \leqslant \sum_{i \leqslant c_{dimM}} \sum_{j \leqslant m_i} \mu(O_r(z_{ij})) \leqslant \bar{\kappa} \sum_{i \leqslant c_{dimM}} \sum_{j \leqslant m_i} vol(O_r(z_{ij})) \leqslant \bar{\kappa} c_{dimM} vol(M),$$

which is a contradiction if $\bar{\kappa}$ is small enough. $\hfill\square$

We will use the above claim for $r = e^{-\bar{\eta}m}$, for $\bar{\eta}$ to be determined in what follows. Let $\phi_i \in C^\infty(M)$ be a function such that $0 \leqslant \phi_i \leqslant 1$, $\phi_i \equiv 1$ on $O_r(z_r^i)$ and $\phi_i \equiv 0$ outside $O_{2r}(z_r^i)$. Suppose that $\bar{\eta}\mathbf{r} \leqslant \eta/4$ so that $\|\phi_i\|_{\mathbf{r}} \leqslant C' \cdot e^{\eta n/4}$. Using exponential mixing for $\phi_1$ and $\phi_2$, by the above bounds on $C^{\mathbf{r}}$ norms, we get for any $i, j \in \{1, 2\}$,

$$\left| \int_M \phi_i \circ f^n \cdot \phi_j d\mu - \int_M \phi_i d\mu \int_M \phi_j d\mu \right| \leqslant CC' \cdot e^{-\eta n/2}.$$

Suppose now that $\bar{\eta}d < \eta/5$. Since $\phi_i$ equals 1 on $O_r(z_r^i)$, by the **CLAIM** we get that:

$$\int_M \phi_i \circ f^n \cdot \phi_j d\mu \geqslant \bar{\kappa} \cdot vol(O_r(z_r^i)) \cdot vol(O_r(z_r^i)) - CC' e^{-\eta n/2} > 0,$$

Since $\phi_i$ is 0 outside $O_{2r}(z_r^i)$, the above for $i = 1$ and $j = 1, 2$ gives that $f^n(O_{2r}(z_r^1))$ intersects both $O_{2r}(z_r^1)$ and $O_{2r}(z_r^2)$. Since $z_r^i \in A_i$ and $d(A_1, A_2) \geqslant \kappa/2$ it follows that

$$\mathrm{diam}(f^n(O_{2r}(z_r^1)) \geqslant \kappa/4.$$

This finishes the proof by taking

(B.5) $$\hat{\eta} = \bar{\eta}/2 = \eta \min \left( \frac{1}{10D}, \frac{1}{8\mathbf{r}} \right).$$

and $c = \min \left( \dfrac{\kappa}{4}, \dfrac{c'}{2} \right).$ $\hfill\square$

## Appendix C. Covering.

**Lemma C.1.** *For each $K$ there exists $\varepsilon_0$ such that for $\varepsilon \leqslant \varepsilon_0$ the following holds. Let $(\Omega, \nu)$ be a measure space, $B \subset D \subset \Omega$ be sets such that $\nu(B) \leqslant \varepsilon^3 \nu(D)$. Let $\{Q_x\}_{x \in D}$ be a measurable family of sets such that if $R_y = \{x : y \in Q_x\}$ then there is $v$ such that for all $x, y$*

(C.1) $$\frac{v}{K} \leqslant \nu(Q_x) \leqslant Kv, \quad \nu(R_y) \leqslant Kv.$$

*Then*

$$\nu(x \in D : \nu(Q_x \cap B) \geqslant \varepsilon \nu(Q_x)) \leqslant \varepsilon \nu(D).$$

*Proof.* Let $Y$ be a random point in $\Omega$ obtained as follows. First choose $X$ uniformly from $D$ and then choose $Y$ uniformly in $Q_X$. Note that $Y$ has bounded density with respect to $\nu$, namely the density is

$$p(y) = \frac{\nu(R_y)}{\int_D \nu(Q_x) d\nu(x)} \leqslant \frac{K^2}{\nu(D)}.$$

Thus

$$\mathbb{P}(Y \in B) = \int_B p(y) d\nu(y) \leqslant K^2 \varepsilon^3.$$

On the other hand

$$\mathbb{P}(Y \in B) = \frac{1}{\nu(D)} \int_D \frac{\nu(Q_x \cap B)}{\nu(Q_x)} dx.$$

So by Markov inequality the set of $x$ where the integrand is greater than $\varepsilon$ has measure smaller than $\frac{\varepsilon^2}{K^2}\nu(D)$. □

## References

[1] Alves J. F., Luzzatto S., Pinheiro, V. *Markov structures and decay of correlations for non-uniformly expanding dynamical systems,* Ann. Inst. H. Poincare **22** (2005) 817–839.

[2] Alves J. F., Pinheiro, V. *Gibbs-Markov structures and limit laws for partially hyperbolic attractors with mostly expanding central direction,* Adv. Math. **223** (2010) 1706–1730.

[3] Anosov D. V., Sinai Ya. G. *Certain smooth ergodic systems,* Russ. Math. Surv. **22** (1967) 103–167.

[4] Avila A., Viana M. *Simplicity of Lyapunov spectra: a sufficient criterion,* Port. Math. **64** (2007) 311–376.

[5] Avila A.,Viana M., Wilkinson A. *Absolute continuity, Lyapunov exponents and rigidity I: geodesic flows,* JEMS **17** (2015) 1435–1462.

[6] L. Barreira, Y. Pesin *Non Uniform Hyperbolicity. Dynamics of systems with nonzero Lyapunov exponents.* Encyclopedia of Math. and Appl. **115** (2007) Cambridge Univ. Press, Cambridge, xiv+513 pp.

[7] Björklund M., Gorodnik A. *Central limit theorems for group actions which are exponentially mixing of all orders,* J. Anal. Math. **141** (2020) 457–482.

[8] Bowen R. *Bernoulli equilibrium states for Axiom A diffeomorphisms,* Math. Systems Theory **8** (1974/75) 289–294.

[9] Bowen, R.: *Equilibrium states and the ergodic theory of Anosov diffeomorphisms. 2d revised edition.* Springer Lecture Notes in Math. **470** (2008), viii+75 pp.

[10] Brin M. I., Pesin Ya. B. *Partially hyperbolic dynamical systems,* (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **38** (1974) 170–212.

[11] A. Brown, F. Rodriguez Hertz *Measure rigidity for random dynamics on surfaces and related skew products.* arXiv:1506.06826

[12] A. Brown, F. Rodriguez Hertz, Z. Wang *Smooth ergodic theory of $\mathbb{Z}^d$-actions.* arXiv:1610.09997

[13] Burns K., Wilkinson A. *On the ergodicity of partially hyperbolic systems,* Ann. of Math. **171** (2010) 451–489.

[14] Castro A. *Fast mixing for attractors with a mostly contracting central direction,* Ergodic Th. Dynam. Systems **24** (2004) 17–44.

[15] Chernov N. I., Haskell C. *Nonuniformly hyperbolic K-systems are Bernoulli,* Ergodic Theory Dynam. Systems **16** (1996) 19–44.

[16] Cornfeld I. P., Fomin S. V., Sinai Ya. G. *Ergodic theory,* Grundlehren der Mathematischen Wissenschaften **245** (1982) Springer, New York, x+486 pp.

[17] Dolgopyat D. *On dynamics of mostly contracting diffeomorphisms,* Comm. Math. Phys. **213** (2000) 181–201.

[18] Dolgopyat D. *On differentiability of SRB states for partially hyperbolic systems,* Invent. Math. **155** (2004) 389–449.

[19] Dolgopyat D., Dong C., Kanigowski A., Nandori P. *Mixing properties of generalized $T, T^{-1}$ transformations,* arXiv:2004.07298, to appear in Israel J. Math.

[20] Dolgopyat D., Dong C., Kanigowski A., Nandori P. *Flexibility of statistical properties for smooth systems satisfying the central limit theorem,* arXiv:2006.02191.

[21] Dolgopyat D., Fayad B., Liu S. *Multiple Borel Cantelli Lemma in dynamics and MultiLog law for recurrence,* arXiv:2103.08382.

[22] Dong C., Kanigowski A. *Bernoulli property for certain skew products over hyperbolic systems,* arXiv:1912.08132.

[23] Gorodnik A., Spatzier R. *Exponential mixing of nilmanifold automorphisms,* J. Anal. Math. **123** (2014) 355–396.

[24] Gouezel S., Stoyanov L. *Quantitative Pesin theory for Anosov diffeomorphisms and flows,* Ergodic Th. Dynam. Systems **39** (2019) 159–200.

[25] Grayson M., Pugh C., Shub M. *Stably ergodic diffeomorphisms,* Ann. of Math. **140** (1994) 295–329.

[26] Hirsch M. W., Pugh C. C., Shub M. *Invariant manifolds,* Lecture Notes in Math. **583** (1977) Springer, Berlin-New York. ii+149 pp.

[27] Host B. *Mixing of all orders and pairwise independent joinings of systems with singular spectrum,* Israel J. Math. **76** (1991) 289–298.

[28] Kanigowski A. *Bernoulli property for homogeneous systems,* arXiv:1812.03209.

[29] Kanigowski A., Rodriguez Hertz F., Vinhage K. *On the non-equivalence of the Bernoulli and K properties in dimension four,* J. Mod. Dyn. **13** (2018) 221–250.

[30] Katok A. *Smooth non-Bernoulli K-automorphisms,* Invent. Math. **61** (1980) 291–299.

[31] Katok A., Hasselblatt B. *Introduction to the modern theory of dynamical systems,* With a supplementary chapter by Katok and L. Mendoza. Encyclopedia of Math and Appl. **54** (1995) Cambridge Univ. Press, Cambridge, xviii+802 pp.

[32] Katznelson Y. *Ergodic automorphisms of $\mathbb{T}^n$ are Bernoulli shifts,* Israel J. Math. **10** (1971) 186–195.

[33] Kleinbock, D. Y., Margulis, G. A.: *Bounded orbits of nonquasiunipotent flows on homogeneous spaces, AMS Transl. Ser. 2* **171** (1996), 141–172.

[34] Ledrappier F. *Some properties of absolutely continuous invariant measures on an interval,* Ergodic Theory Dynam. Systems **1** (1981) 77–93.

[35] Liverani, C.,*On contact Anosov flows,* Ann. of Math. (2), **159** (2004), no. 3, 1275–1312.

[36] Ornstein D. S., Weiss B. *Geodesic flows are Bernoullian,* Israel J. Math. **14** (1973) 184–198.

[37] Ornstein D. S., Weiss B. *On the Bernoulli nature of systems with some hyperbolic structure,* Ergod. Theory Dyn. Sys. **18**(2), (1998), 441-456.

[38] Parry, W., Pollicott, M. *Zeta functions and the periodic orbit structure of hyperbolic dynamics,* Astérisque **187-188** (1990) 268 pp.

[39] Pesin Ya. B. *Families of invariant manifolds that correspond to nonzero characteristic exponents,* Math. USSR-Izv. **40** (1976) 1261–1305.

[40] Pesin Ya. B. *Geodesic flows in closed Riemannian manifolds without focal points,* (Russian) Izv. Akad. Nauk SSSR **41** (1977) 1252–1288.

[41] Pugh C. C., Shub M. *Stably ergodic dynamical systems and partial hyperbolicity,* J. Complexity **13** (1997) 125–179.

[42] Pugh C. C. Shub M. *Stable ergodicity and julienne quasi-conformality,* JEMS **2** (2000) 1–52.

[43] Pugh C. C., Shub M., Wilkinson A. *Hölder foliations,* Duke Math. J. **86** (1997) 517–546.

[44] Ratner M. *Anosov flows with Gibbs measures are also Bernoullian,* Israel J. Math. **17** (1974) 380–391.

[45] Rohlin, V. A. *On endomorphisms of compact commutative groups,* Izvestiya Akad. Nauk SSSR. **13** (1949) 329–340.

[46] Rohlin V. A. *On the fundamental ideas of measure theory,* AMS Translation **71** (1952) 55 pp.

[47] Rohlin V. A., Sinai Ya. G. *The structure and properties of invariant measurable partitions,* (Russian) Dokl. Akad. Nauk SSSR **141** (1961) 1038–1041.

[48] Rudolph D. *Asymptotically Brownian skew products give non-loosely Bernoulli K-automorphisms.* Invent. Math., **91** (1988) 105–128.

[49] Shub M., Wilkinson A. *Pathological foliations and removable zero exponents,* Invent. Math. **139** (2000) 495–508.

[50] Sinai Ya. G. *Classical dynamic systems with countably-multiple Lebesgue spectrum-II,* (Russian) Izv. Akad. Nauk SSSR **30** (1966) 15–68.

[51] Sinai Ya. G. *The hierarchy of stochastic properties of deterministic systems,* Encyclopaedia Math. Sciences **100** (2000) 106–108.