LIMITING DISTRIBUTION AND ERROR TERMS FOR THE NUMBER OF VISITS TO BALLS IN NON-UNIFORMLY HYPERBOLIC DYNAMICAL SYSTEMS

Nicolai T. A. Haydn

Department of Mathematics
University of Southern California
Los Angeles, CA 90089, USA

Kasia Wasilewska

Department of Mathematics
Harvard-Westlake School
Studio City, CA 91604, USA

(Communicated by Carlangelo Liverani)

Abstract. We show that for systems that allow a Young tower construction with polynomially decaying correlations the return times to metric balls are in the limit Poisson distributed. We also provide error terms which are powers of logarithm of the radius. In order to get those uniform rates of convergence the balls centres have to avoid a set whose size is estimated to be of similar order. This result can be applied to non-uniformly hyperbolic maps and to any invariant measure that satisfies a weak regularity condition. In particular it shows that the return times to balls is Poissonian for SRB measures on attractors.

1. **Introduction.** Poincaré's recurrence theorem [24] established that for measure preserving maps points return to neighbourhoods arbitrarily often almost surely. The return time for the first return was quantified by Kac [19] in 1947 and since then there have been efforts to describe the return statistics for shrinking neighbourhoods. One looks at returns for orbit segments whose length is given by a parameter t scaled by the size of the target set. The scaling factor is suggested by Kac's theorem. In [21, 20] it was shown that for ergodic maps one can achieve any limiting statistics if one chooses the shrinking target sets suitably. For a generating partition the natural neighbourhoods are cylinder sets and for those the limiting distributions for entry and return times were shown under various mixing conditions to be exponential with parameter 1 (see for instance [17, 13, 7, 1]). However Kupsa constructed an example which has a limiting hitting time distribution almost everywhere and which is not the exponential distribution with parameter 1. For multiple returns it has been established under various mixing conditions that the limiting distribution is Poissonian almost surely. The first such result is due to Doeblin [10] for the Gauss map for which he showed that at the origin multiple returns to cylinder sets are in the limit Poisson distributed. Pitskel [23] (see also [8]) used the moment method to prove that the return times are in the limit Poissonian for

equilibrium states for Axiom A maps using Markov partitions. Similar results have been shown for some non-uniformly expanding maps [18, 11], for rational maps [14] and for ϕ-mixing maps [3, 2]. Using the Chen-Stein method, it was also shown for toral automorphism (by way of harmonic analysis) [9] and for unions of cylinders and ϕ-mixing measures in [16]. More results are mentioned in the review [15].

For the return times to metric balls B_ρ on manifolds [6] proves the limiting distribution to be Poissonian for the SRB measure on a one-dimensional attractor which allows the construction of a Young tower [26, 27] with exponentially decaying correlations. The speed of convergence turns out to be a (positive) power of the radius ρ of the target ball. Here we prove a similar result in the case when the correlations decay at a polynomial rate. Also the attractor is not required to be one-dimensional. In fact, the result here applies to any invariant measure that can be constructed using the tower construction whether it be absolutely continuous or not. The speed of convergence in this case is a negative power of $|\log \rho|$.

Let us note that recently Pène and Saussol [22] have obtain the limiting Poisson distribution for return times for the SRB measure assuming some geometric regularity. They moreover give the examples of the intermittend solenoid which was introduced in [4] and the stadium billiard both of which have polynomially decaying correlations and are thus covered by the results of this paper. A large part of [22] is devoted to prove the annulus condition which is implicit in our Theorem 3 below. In the case of dispersive billiards with polynomially decaying correlations the limiting Poisson distribution for return times was shown in [12] using the induced map though no rate of convergence was provided.

The results of this paper are taken from [25] and are organised as follows: In Section 2 we state the results where the exact conditions are given in later sections (indicated within the theorems). In Section 3 we give the description of Young's tower construction which is central to the result. In Section 4 we prove the main result Theorem 1 which proceeds in several steps which are outlined at the beginning of the section. In Section 5 we consider the case of a diffeomorphism on a manifold and show that very short returns i.e. those that are of order $|\log \rho|$ (where ρ is the radius of the target ball) constitute a very small portion of the manifold whose measure can be bounded. We use and adapt an argument from [6] Lemma 4.1.

2. **Results.** Let (M, T) be a dynamical system on the compact metric space M. For a positive parameter $\mathfrak{a}$ define the set

$$\mathcal{V}_\rho(\mathfrak{a}) = \{\mathsf{x} \in M : B_\rho(\mathsf{x}) \cap T^n B_\rho(\mathsf{x}) \neq \varnothing \text{ for some } 1 \leq n < \mathfrak{a}|\log \rho|\}, \qquad (1)$$

where $\rho > 0$. The set $\mathcal{V}_\rho$ represents the points within M with very short return times.

For a ball $B_\rho(\mathsf{x}) \subset M$ we define the counting function

$$S_{\rho,\mathsf{x}}^t(x) = \sum_{n=0}^{\lfloor t/\mu(B_\rho(\mathsf{x})) \rfloor - 1} \mathbb{1}_{B_\rho(\mathsf{x})} \circ T^n(x).$$

which tracks the number of visits a trajectory of the point $x \in M$ makes to the ball $B_\rho(\mathsf{x})$ on an orbit segment of length $N = N_{\rho,\mathsf{x}}^t = \lfloor t/\mu(B_\rho(\mathsf{x})) \rfloor$. (We often omit the sub- and superscripts and simply use $S(x)$.)

Let us now state our main results. Definitions and background material are in Section 3. In the first theorem, M could be a general metric space and the reference measure used in the tower construction need not be absolutely continuous with respect to the Lebesgue measure.

Theorem 1. *Let* (M, T) *be a dynamical system which can be modeled by a Young tower. Suppose that the tail of the tower's return time function decays polynomially with degree* $\lambda > 7$. *Let* μ *be the SRB measure admitted by the system and assume that* μ *is geometrically and* ξ-*regular. Exact assumptions are in Section 4.1.*

Then there exist constants $\kappa, \kappa' \in (0, \frac{\lambda-7}{4})$ *and* $\mathfrak{a}, C, C' > 0$ *such that for* ρ *sufficiently small there exists a set* $\mathcal{X}_\rho \subset M$ *with* $\mu(\mathcal{X}_\rho) \leq C' |\log \rho|^{-\kappa'}$ *such that for all* ρ-*balls with centers* $\mathsf{x} \notin \mathcal{X}_\rho \cup \mathcal{V}_\rho(\mathfrak{a})$ *we have*

$$\left| \mathbb{P}(S = k) - e^{-t} \frac{t^k}{k!} \right| \leq C |\log \rho|^{-\kappa} \qquad \text{for all } k \in \mathbb{N}_0. \tag{2}$$

Theorem 1 establishes the limiting statistics of returns outside the set $\mathcal{V}_\rho$ of very short returns. If M is a compact manifold and the map T is sufficiently regular then one can say something about the size of $\mathcal{V}_\rho$. This is done in the following result where we obtain for smooth maps on manifolds the following limiting result that at the same time controls the size of the forbidden set.

Theorem 2. *Let* (M, T) *be a dynamical system satisfying the assumptions of Theorem 1 and where* T *is a* C^2-*diffeomorphism the compact manifold* M *and* $\lambda > 9$. *Let* μ *be the invariant measure. For all* ρ *sufficiently small there exist sets* $\mathcal{X}_\rho, \mathcal{V}_\rho(\mathfrak{a})$ *so that for* $\mathsf{x} \notin \mathcal{X}_\rho \cup \mathcal{V}_\rho(\mathfrak{a})$ *the function* $S^t_{\rho, \mathsf{x}}$ *counting the number of visits to the ball* $B_\rho(\mathsf{x})$ *satisfies* (2) *where the constants* $\kappa \in (0, \frac{\lambda-7}{4})$ *and* $C > 0$ *are independent of* ρ *and* x.

Further there exists $\kappa' \in (0, \frac{\lambda-9}{4}]$ *and* $C'' > 0$ *so that*

$$\mu(\mathcal{X}_\rho \cup \mathcal{V}_\rho(\mathfrak{a})) \leq C'' |\log \rho|^{-\kappa'}$$

for all ρ *small enough* $(\mathfrak{a} = [4(\|DT\|_{\mathscr{L}^\infty} + \|DT^{-1}\|_{\mathscr{L}^\infty})]^{-1})$.

Under some more favourable conditions one can conclude that the convergence is almost sure. For this see the remark following Proposition 5.1.

The assumptions in Theorems 1 and 2 imply that the invariant measure is absolutely continuous with respect to a given reference measure on unstable leaves (see precise assumptions) and can of course be singular with respect to the Lebesgue measure on unstable leaves. In the classical case when the reference measure is the Lebesgue measure on unstable leaves, then the invariant measure is the SRB measure and we obtain the following result which generalises the main theorem of [6] from exponentially decaying correlation to polynomially decaying correlations. Technically, the main part of Theorem 3 is to verify the annulus condition (A4)(a). Moreover, here we don't require the attractor to have one-dimensional unstable manifolds.

Theorem 3. *Let* (M, T) *be a dynamical system on a compact manifold* M *of dimension* D *where* $T : M \to M$ *is a* C^2 *diffeomorphism with attractor* $\mathscr{A}$. *Suppose the system can be modeled by a Young tower whose return time function decays polynomially with degree* $\lambda > \max\{9, \frac{D+2}{u}\}$, *where* u *is the dimension of the unstable leaves. Let* μ *be the SRB measure admitted by the system. Exact assumptions are in Section 6.*

For every $\kappa < \frac{\lambda-7}{4}$ *there exists* $\hat{C} > 0$, *and sets* $\mathcal{Z}_\rho \subset M$ *with*

$$\mu(\mathcal{Z}_\rho) \leq \hat{C} |\log \rho|^{-\frac{\lambda-9}{4}}$$

so that for $x \notin \mathcal{Z}_\rho$, the function $S^t_{\rho,x}$ counting the number of visits to the ball $B_\rho(x)$ satisfies

$$\left| \mathbb{P}(S = k) - e^{-t}\frac{t^k}{k!} \right| \leq \hat{C}\,|\log\rho|^{-\kappa} \qquad \text{for all } k \in \mathbb{N}_0.$$

Throughout the paper $C_1, C_0, \ldots$ and $\alpha, \beta, \ldots$ denote global constants while $c_0, c_1, \ldots$ are locally defined constants.

3. **Young tower.** We shall use the tower method which was developed by L-S Young in [26, 27] to construct invariant measures and to obtain decay rates for correlations.

(I) Foliation, return time function and Markov partition: Let Γ^u be a collection of unstable leaves γ^u and Γ^s a collection of stable leaves γ^s. We assume that $\gamma^u \cap \gamma^s$ consists of a single point for all $(\gamma^u, \gamma^s) \in \Gamma^u \times \Gamma^s$. We also assume that the base set $\Lambda = \left(\bigcup_{\gamma^u \in \Gamma^u} \gamma^u\right) \cap \left(\bigcup_{\gamma^s \in \Gamma^s} \gamma^s\right)$ is, up to null sets with respect to a given 'reference measure' $\hat{m}$, the disjoint union of rectangles $\Lambda_i = \left(\bigcup_{\gamma^u \in \Gamma^u} \gamma^u\right) \cap \left(\bigcup_{\gamma^s \in \Gamma^s_i} \gamma^s\right)$, $i \in \mathbb{N}$, where Γ^s_i form a disjoint decomposition of Γ^s. The map T contracts along the stable leaves and similarly T^{-1} contracts along the unstable leaves as decribed in Assumption (A1) below.

There exists a return time function $R : \Lambda \to \mathbb{N}$ constant on each Λ_i such that $T^R \Lambda_i \subset \Lambda$. We put $R_i = R|_{\Lambda_i}$. Since any point x from $\Lambda_i \subset \Lambda$ will return to Λ after R_i iterations we can define the return transform $\hat{T} : \Lambda \to \Lambda$ by

$$\hat{T}x := T^{R_i}x \quad \text{for all } x \in \Lambda_i \quad \text{and } i \in \mathbb{N}$$

For convenience we extend the return time function $\hat{T}$ to all $x \in M$ by putting $\hat{T}(x) = T^j(x)$ where $j \geq 0$ is the smallest integer so that $T^j x \in \Lambda$.

The pairwise disjoint subsets $\{\Lambda_i\}_{i \in N}$ are assumed to satisfy the Markov property under the return transform as follows: $\hat{T}(\gamma^u \cap \Lambda_i) = \hat{\gamma}^u \cap \Lambda$ for some $\hat{\gamma}^u \in \Gamma^u$ and similarly if $\gamma^s \in \Gamma^s_i$ then $\hat{T}(\gamma^s) \subset \hat{\gamma}^s$ for some $\hat{\gamma}^s \in \Gamma^s$ (see [26]).

If $\gamma^u, \hat{\gamma}^u \in \Gamma^u$ are two unstable leaves then the holonomy map $\Theta : \gamma^u \cap \Lambda \to \hat{\gamma}^u \cap \Lambda$ is defined by $\Theta(x) = \hat{\gamma}^u \cap \gamma^s(x)$ for $x \in \gamma^u \cap \Lambda$, where $\gamma^u(x)$ be the local unstable leaf through x.

Without loss of generality we will assume that the greatest common divisor of all of the values R_i of the return times function is equal to one. The function R is assumed to be integrable on each unstable leaf γ^u with respect to a 'reference measure' $\hat{m}$. That is

$$\sum_i R_i \hat{m}_{\gamma^u}(\Lambda_i) = \int_\Lambda R\,d\hat{m}_{\gamma^u} < \infty$$

for all local unstable leaves γ^u, where $\hat{m}_{\gamma^u}$ is the conditional measure on γ^u. Precise assumptions are formulated in Assumptions (A1) and (A2).

(II) Separation Time: Based on the flight time R and the partition of the base, we define the *separation time*

$$s(x, y) = \min\{k \geq 0 : \hat{T}^k x \text{ and } \hat{T}^k y \text{ lie in distinct } \Lambda_i\},$$

so that for $x, y \in \Lambda_i$, $s(x, y) = 1 + s(\hat{T}x, \hat{T}y)$ and in particular $s(x, y) \geq 1$. If $F \subset \Lambda$ then we also put

$$s(F) = \min_{x,y \in F} s(x, y).$$

We extend the separation function to points outside Λ as follows. If for points x, y there exists an integer $k \geq 0$ (smallest) so that $T^k x, T^k y \in \Lambda_i$ for some i, then $s(x, y) = s(T^k x, T^k y)$.

This definition of separation time is in accordance with [26] where the case of polynomially decaying correlations was proven. It differs from the usage in [27] which covers the exponentially decaying case. The separation time considered there is R^s (ergodic sum) in our terminoloy.

(III) The Jacobian: Even though the original system $(M, \mathscr{B}, T, \mu)$ is not necessarily differentiable in the ordinary sense, one uses the Radon-Nikodym derivative of T with respect to the 'reference measure' $\hat{m}$ (following [26], p. 596). The derivative exists and is well defined because every $T^{R_i}|_{\Lambda_i}$ and its inverse are non-singular with respect to the conditional probability measure $\hat{m}_{\gamma^u}$ on unstable leaves $\gamma^u \in \Gamma^u$. Let

$$JT = \frac{d(T_*^{-1} \hat{m}_{\gamma^u})}{d\hat{m}_{\gamma^u}}.$$

The requirements on JT will be spelled out in Assumption (A1).

(IV) The SRB measure: According to [26, 27] (M, T) has a generalised SRB measure μ given by

$$\mu(S) = \sum_{i=1}^{\infty} \sum_{j=0}^{R_i - 1} m(T^{-j}S \cap \Lambda_i) \quad \text{for sets } S \subset M.$$

where m is the generalised SRB measure for the uniformly expanding system $(\Lambda, \hat{T})$.

If we denote by m_{γ^u} the conditional measure on unstable leaves γ^u, then $dm = dm_{\gamma^u} d\nu(\gamma^u)$ where $d\nu$ is the transversal measure. We will refer to the portions of the tower above each Λ_i as *beams* and assume that for $n \in \mathbb{N}$, there are only finitely many i's for which $R_i = n$, i.e. for every n there are only finitely many beams with that height.

The map T maps each level bijectively onto the next and at the last level unstable leaves inside the rectangles map bijectively to entire full length unstable leaves. (Note that R_i may not be the first time Λ_i returns to Λ.)

4. **Proof of Theorem 1.** In this section we prove Theorem 1. We begin by stating the precise assumptions necessary for the result and derive some of the consequences that follow with minimal work. In Section 4.3 we introduce cylinder sets. In order to approximate the metric balls we will restrict to those cylinder set that have only short returns. We then provide several results on the behaviour of cylinder sets under suitable applications of the return map $\hat{T}$. The succeeding Section 4.4 contains estimates concerning the portions of the space M which have to be omitted in order to obtain good asymptotic behaviour for the long returns. This is the 'forbidden set'. Section 4.5 utilizes the Poisson approximation theorem from Section 7 to establish a splitting of the error term to the Poisson distribution into two parts $\mathcal{R}_1$ and $\mathcal{R}_2$. The remainder of the section is devoted to estimating these error terms one by one. In Section 4.6 we estimate the error $\mathcal{R}_1$ which comes from long term interactions and uses decay of correlations. Sections 4.7 and 4.8 are

devoted to bounding the term $\mathcal{R}_2$ which comes from short time (but not very short time) interactions. This is the place where the Young tower construction comes to play and where we have to use approximations by cylinder sets in order to make careful distinctions between short returns and long returns to balance out different contributions to the error term. In Section 4.9 the different error terms are brought together and the various parameters are optimised.

4.1. **Assumptions.** Let (M, T) be a dynamical system equipped with a metric d and let μ be the SRB measure associated to the system (whose existence follows by [26, 27] from Assumptions (A1) and (A2)). We will require the following:

(A1) *Regularity of the Jacobian and the metric on the leaves*
Recall that the Jacobian $J\hat{T}$ measures the expansion rate of the reference measure in the unstable direction. We assume there exists a constant $C_0 > 0$ and $\alpha \in (0, 1)$ such that for any x, y in Λ with $s(x, y) \geq 1$

(a) $\qquad \left| \log \dfrac{J\hat{T}x}{J\hat{T}y} \right| \leq C_0\, \alpha^{s(\hat{T}x, \hat{T}y)}$ if $\gamma^u(x) = \gamma^u(y)$;

(b) $\qquad d(\hat{T}^k x, \hat{T}^k y) \leq C_0\, \alpha^{s(x,y)-k}$ for $0 \leq k < s(x, y)$ for $x, y \in \gamma^u$;

(c) $\qquad \log \displaystyle\prod_{k=n}^{\infty} \dfrac{J\hat{T}(\hat{T}^k x)}{J\hat{T}(\hat{T}^k y)} \leq C_0 \alpha^n$ if $\gamma^s(x) = \gamma^s(y)$;

(d) $\qquad \dfrac{d\Theta^{-1}\hat{m}_{\Theta\gamma^u}}{d\hat{m}_{\gamma^u}}(x) = \log \displaystyle\prod_{k=0}^{\infty} \dfrac{J\hat{T}(\hat{T}^k x)}{J\hat{T}(\hat{T}^k \Theta x)}$;

(e) $\qquad d(\hat{T}^n x, \hat{T}^n y) \leq C_0 \alpha^n$ for $n \in \mathbb{N}$ if $\gamma^s(x) = \gamma^s(y)$.

(A2) *Polynomial Decay of the Tail*
There exist constants C_1 and $\lambda > 7$ such that

$$\hat{m}_{\gamma^u}(R > k) \leq C_1\, k^{-\lambda} \tag{3}$$

for every unstable leaf γ^u. With regard to Assumption (A1)(d) this condition is satisfied for all unstable leaves if it can be verified for one γ^u.

(A3) *Additional assumption on λ*
Let ς be the dimension of the measure m_{γ^u} and $\hat{\varsigma}$ the dimension of μ (that is, by ergodicity $\mu = \lim_{\rho \to 0} \frac{\log \mu(B_\rho(x))}{\log \rho}$ for μ-almost every x). We will require that

$$\xi = \varsigma(\lambda - 1) - \hat{\varsigma} > 1. \tag{4}$$

(A4) *Regularity of the invariant measure*
Let $\xi = \varsigma(\lambda - 1) - \hat{\varsigma} > 1$ by (A3) and suppose that the positive constant $\varsigma' < \varsigma$ is fixed. There exist a set $\mathcal{E}_\rho \subset M$ satisfying $\mu(\mathcal{E}_\rho) \leq |\log \rho|^{-\frac{\lambda-4}{3}}$ so that for ρ small enough:

(a) (ξ-regularity) There exists $w_0 \in (1, \xi)$ and $a > 0$ so that

$$\frac{\mu(B_{\rho+\rho^w}(\mathsf{x}) \setminus B_{\rho-\rho^w}(\mathsf{x}))}{\mu(B_\rho(\mathsf{x}))} \leq \frac{1}{g(w)|\log \rho|^a}$$

for all $\mathsf{x} \notin \mathcal{E}_\rho$ and $w > w_0$ where the function $g(w)$ is so that $\sum_n^\infty g(n^\beta)^{-1} < \infty$ for some $\frac{1}{2} < \beta < 1 - \frac{3}{\lambda - 1}$. We say μ is ξ-regular.

(b) (geometric regularity) There exists a $\varsigma' < \varsigma$, $\hat{\varsigma}'' > \hat{\varsigma}$ satisfying (4) and $C_2 > 0$ such that

$$m_{\gamma^u}(B_\rho(\mathsf{x})) \leq C_2\,\rho^{\varsigma'}, \qquad \mu(B_\rho(\mathsf{x})) \geq C_2\rho^{\hat{\varsigma}''}$$

for all $B_\rho(\mathsf{x}) \subset M$ for which $\mathsf{x} \notin \mathcal{E}_\rho$ and all unstable leaves γ^u.

For simplicity sake we will from now on write B_ρ for $B_\rho(\mathsf{x})$.

Remark. Also let us note that sometimes the following annulus condition is used (see e.g. [22]):

$$\frac{\mu(B_{\rho + \delta\rho}(\mathsf{x}) \setminus B_{\rho - \delta\rho}(\mathsf{x}))}{\mu(B_\rho(\mathsf{x}))} \leq c_0 \frac{\delta\rho^g}{\rho^h}. \tag{5}$$

If $\xi g > h$ then (5) implies the regularity condition (A4)(a), for if one puts $\delta\rho = \rho^w$ then $\frac{\delta\rho^g}{\rho^h} = \rho^{gw - h}$ and the exponent can be made positive with a $w \in (1, \xi)$. The summability condition follows from the fact that $\sum_{n=p}^\infty \rho^{gn^\beta - h} = \mathcal{O}(\rho^{gp^\beta - h})$ decays to zero faster than any negative power of $|\log \rho|$ as $\rho \to 0$ considering that in Section 4.6 p is proportional to a negative power of $\mu(B_\rho)$.

4.2. **Immediate consequences of the assumptions.** In this section we list some basic results which will be needed in the proof of the main results.

Lemma 4.1 (Distortion). *There exists a constant $C_3 > 1$ such that*
(i) for any x and y in Λ, $\gamma^u(x) = \gamma^u(y)$ with separation time $s(x, y) \geq q$:

$$J\hat{T}^q x \in J\hat{T}^q y \left[\frac{1}{C_3}, C_3\right]. \tag{6}$$

(ii) For any $F \subset F' \subset \Lambda_i \cap \gamma^u$ (for some i) and for any $q \leq s(F') = \inf_{x, x' \in F'} s(x, x')$

$$\frac{1}{C_3} \frac{m_{\hat{\gamma}^u}(\hat{T}^q F)}{m_{\hat{\gamma}^u}(\hat{T}^q F')} \leq \frac{m_{\gamma^u}(F)}{m_{\gamma^u}(F')} \leq C_3 \frac{m_{\hat{\gamma}^u}(\hat{T}^q F)}{m_{\hat{\gamma}^u}(\hat{T}^q F')}, \tag{7}$$

where $\hat{\gamma}^u = \gamma^u(\hat{T}^q(F))$.

Proof. (i) Let $x, y \in \Lambda$, $\gamma^u(x) = \gamma^u(y)$ and let $q \geq 1$ be an integer less than or equal to $s(x, y)$. Then by the chain rule and (A3)(a)

$$\begin{aligned}
\left|\log \frac{J\hat{T}^q x}{J\hat{T}^q y}\right| &\leq \sum_{j=0}^{q-1}\left|\log \frac{J\hat{T}(\hat{T}^j x)}{J\hat{T}(\hat{T}^j y)}\right| \\
&\leq \sum_{j=0}^{q-1} C_0 \alpha^{s(\hat{T}(\hat{T}^j x), \hat{T}(\hat{T}^j x))} \\
&= C_0 \sum_{j=0}^{q-1} \alpha^{q - (j+1)} \\
&\leq \frac{C_0}{1 - \alpha}.
\end{aligned}$$

as $\alpha^{s(\hat{T}^{j+1} x, \hat{T}^{j+1} y)} \leq \alpha^{q - (j+1)}$ which implies the statement (i) with $C_3 \geq c_1 = e^{\frac{C_0}{1 - \alpha}}$.

(ii) This follows from part (i) since

$$\frac{c_1^{-1} m_{\hat{\gamma}^u}(\hat{T}^q F)}{c_1 m_{\hat{\gamma}^u}(\hat{T}^q F')} \leq \frac{m_{\gamma^u}(F)}{m_{\gamma^u}(F')} \leq \frac{c_1 m_{\hat{\gamma}^u}(\hat{T}^q F)}{c_1^{-1} m_{\hat{\gamma}^u}(\hat{T}^q F')}.$$

The result now follows with $C_3 = c_1^2$.   □

According to [27](Theorem 3) the decay of correlations is polynomial: Let ϕ be a Lipschitz continuous function and $\psi \in \mathscr{L}^\infty$ constant on local stable leaves, that is $\psi(x) = \psi(y)$ if there exists a k and $j < R_k$ so that $T^{-j}x, T^{-j}y \in \gamma^s \cap \Lambda_k$ for some $\gamma^s \in \Gamma^s$. Then one has

$$\left| \int_M \phi\, \psi \circ T^n\, d\mu - \int_M \phi\, d\mu \int_M \psi\, d\mu \right| \leq \varphi_n \|\phi\|_{Lip} \|\psi\|_{\mathscr{L}^\infty} \tag{8}$$

where the decay function $\varphi_n = \mathcal{O}(1) \sum_{k>n} m(R > k) \leq C_4\, n^{-\lambda+1}$, for some $C_4 > 0$ where $\lambda > 0$ is the tail decay exponent from assumption (A2). Note that in general for functions ψ which are not constant on local stable leaves, the supremum norm on the RHS of (8) has to be replaced by the Lipschitz norm (see [26]).

We will also need the following function of $s \in \mathbb{R}^+$:

$$\Omega(s) := \sqrt{\sum_{i:R_i>s} R_i\, m(\Lambda_i)}.$$

Since the return time R is integrable $\Omega(s) \to 0$ as $s \to \infty$.

Lemma 4.2 (Decay of Ω). *There exists a constant C_5 such that for $s \geq 4$*

$$\Omega(s) \leq C_5 s^{-\theta}$$

where $\theta = (\lambda - 1)/2$.

Proof. By definition

$$\Omega(s)^2 = \sum_{i:R_i>s} R_i\, m(\Lambda_i) \leq \sum_{k=s}^\infty m(R > k) + s\, m(R > s)$$

$$\leq \sum_{k=s}^\infty C_1 k^{-\lambda} + sC_1 s^{-\lambda} \leq c_1 s^{2\theta}$$

using the tail decay, where $c_1 < \infty$. We complete the proof by setting $C_5 = \sqrt{c_1}$.   □

4.3. **Cylinder sets.** Let s be a given integer. We shall separate the beams with return times greater than s and also portions of the base that visit those beams during the "flight". The beams with heights less than s will be referred to as "short" and constitute the principal part. The "tall" beams (i.e. when the returns are $> s$) will be treated like error terms and contribute to the "forbidden" set $\mathcal{X}_\rho$ whose size is small and estimated in Section 4.4. Let us introduce several quantities that will be needed to deal with the long return times.

(I) For indices $(i_0, \ldots, i_l) \in \mathbb{N}^{l+1}$ we define the *l-cylinder* (w.r.t. the map $\hat{T}$) by

$$\zeta_{i_0,\ldots,i_l} = \Lambda_{i_0} \cap \hat{T}^{-1}\Lambda_{i_1} \cap \hat{T}^{-2}\Lambda_{i_2} \cap \ldots \cap \hat{T}^{-l}\Lambda_{i_l},$$

and denote by $\mathfrak{I}$ the collection of indices $(i_0, \ldots, i_l)$ such that the associated cylinder $\zeta_{i_0,\ldots,i_l}$ is non-empty.

(II) For every Λ_i let's define the subset consisting exclusively of points that only visit short beams:

$$\tilde{\Lambda}_i = \{x \in \Lambda_i : \forall l \leq n,\ R(\hat{T}^l x) \leq s\}.$$

Let us note that if the original beam Λ_i happens to be tall (i.e. $R_i > s$) then the corresponding $\tilde{\Lambda}_i$ will be empty. With $\Lambda = \bigcup_i \Lambda_i$ and $\tilde{\Lambda} = \bigcup_i \tilde{\Lambda}_i$ we obtain in particular $\Lambda \setminus \tilde{\Lambda} = \{x \in \Lambda : \exists l \in [0,n] \text{ s.t. } R(\hat{T}^l x) \leq s\}$. Similarly we define the restriction of a cylinder $\zeta_{i_0,\dots,i_l}$ to short returns by

$$\tilde{\zeta}_{i_0,\dots,i_l} = \left\{ x \in \zeta_{i_0,\dots,i_l} : R(\hat{T}^j(x)) \leq s \; \forall \, j = 0, \dots, l \right\}.$$

In other words

$$\tilde{\zeta}_{i_0,\dots,i_l} = \begin{cases} \zeta_{i_0,\dots,i_l} & \text{if } R_{i_0}, \dots, R_{i_l} \leq s \\ \varnothing & \text{otherwise.} \end{cases} \tag{9}$$

In particular we see that if one of the beams on the cylinder's path is tall, i.e. $R_{i_j} > s$ for a $j \in [0,l]$, then it must have originated inside $\Lambda_{i_0} \setminus \tilde{\Lambda}_{i_0}$. Thus

$$\bigcup_{i_0,\dots,i_l} \zeta_{i_0,\dots,i_l} \setminus \tilde{\zeta}_{i_0,\dots,i_l} \subset \bigcup_{i_0} \Lambda_{i_0} \setminus \tilde{\Lambda}_{i_0} \qquad \text{and} \qquad \bigcup_i \tilde{\Lambda}_i \subset \bigcup_{i_0,\dots,i_l} \tilde{\zeta}_{i_0,\dots,i_l} \tag{10}$$

as long as $l \leq n$.

(III) For given n, j and $i_0 \in \mathbb{N}$, $j < R_{i_0} \leq s$ we define the set of 'suitable' symbols by

$$I_{i_0,j,n} = \left\{ (i_0, \dots, i_l) \in \mathfrak{I} : \sum_{k=0}^{l-1} R_{i_k} \leq n + j < \sum_{k=0}^{l} R_{i_k} \right\}. \tag{11}$$

Accordingly a string of symbols $(i_0, \dots, i_l)$ is called (n,j)-*minimal* if it satisfies the property $R^l \leq n + j < R^{l+1}$, where $R^l = \sum_{k=0}^{l-1} R_{i_k}$. Note that for any given values j and n the cylinders indexed by $I_{i,j,n}$ partition the beam base Λ_i, up to set of measure zero, i.e.

$$\Lambda_i = \bigsqcup_{\tau \in I_{i,j,n}} \zeta_\tau.$$

In what follows we shall often write I, instead of $I_{i,j,n}$.

If $\tau = (i_0, \dots, i_l)$ then let $\tau' = (i_0, \dots, i_{l-1})$. Since $\zeta_\tau = \zeta_{\tau'} \cap \hat{T}^{-l} \Lambda_{i_l} \subset \zeta_{\tau'}$, we will write $\tau \subset \tau'$ to reflect the relationship between the cylinders.

The remainder of this section is taken up by providing some essential estimates involving the quantities introduced.

Lemma 4.3. *The diameter of the cylinder set $\zeta_{i_0,\dots,i_l}$ restricted to unstable leaves is exponentially small:*

$$|\zeta_{i_0,\dots,i_l} \cap \gamma^u| \leq C_0 \, \alpha^{l+1}. \tag{12}$$

Proof. Let x and y be two points in $\zeta_{i_0,\dots,i_l}$, then by definition we have

$$\hat{T}^k x, \hat{T}^k y \in \Lambda_{i_k} \cap \gamma^u \quad \text{for} \quad 0 \leq k \leq l.$$

It follows that $s(x,y) \geq l + 1$. Therefore by Assumption (A1b):

$$d(x,y) \leq C_0 \, \alpha^{s(x,y)} \leq C_0 \, \alpha^{l+1}$$

and, since the points x, y were arbitrary, we conclude that

$$|\zeta_{i_0,\dots,i_l} \cap \gamma^u| \leq C_0 \, \alpha^{l+1}.$$

$\square$

Lemma 4.4. *Let $\tau = (i_0, \ldots, i_l)$ be an index in $I_{i_0, j, n}$ and put $\tau' = (i_0, \ldots, i_{l-1})$. Then*

$$\frac{m_{\gamma^u}(T^{-j}\mathcal{B} \cap \tilde{\zeta}_{\tau'})}{m_{\gamma^u}(\tilde{\zeta}_{\tau'})} \leq C_6 \, \mu(B_\rho)$$

for any set $\mathcal{B} \subset T^{-n}B_\rho$ and for all $\tilde{\zeta}_{\tau'} \neq \varnothing$.

Proof. By inclusion we have

$$\frac{m_{\gamma^u}(T^{-j}\mathcal{B} \cap \tilde{\zeta}_{\tau'})}{m_{\gamma^u}(\tilde{\zeta}_{\tau'})} \leq \frac{m_{\gamma^u}(T^{-(n+j)}B_\rho \cap \tilde{\zeta}_{\tau'})}{m_{\gamma^u}(\tilde{\zeta}_{\tau'})}.$$

Let the number b be such that $n + j - b = R^l$. Recall that $\tau \in I$ means that it is (n, j)-minimal, i.e. that $n + j$ lies between R^l and R^{l+1}, we have $0 \leq b < R_{i_l}$. Further $T^{n+j-b} = \hat{T}^l$. Recall that $s(\zeta_{\tau'}) = l$, thus we can push both the numerator and the denominator forward by $\hat{T}^l$ and use distortion, Lemma 4.1 and (A1)(a), to obtain for $\tilde{\zeta}_{\tau'} \neq \varnothing$:

$$\frac{m_{\gamma^u}(T^{-(n+j)}B_\rho \cap \tilde{\zeta}_{\tau'})}{m_{\gamma^u}(\tilde{\zeta}_{\tau'})} \leq C_3 \frac{m_{\hat{\gamma}^u}(T^{-b}B_\rho \cap \hat{T}^l\zeta_{\tau'})}{m_{\hat{\gamma}^u}(\hat{T}^l\zeta_{\tau'})} \tag{13}$$

since by assumption and (9) $\tilde{\zeta}_{\tau'} = \zeta_{\tau'}$. Here we put again $\hat{\gamma}^u = \gamma^u(\hat{T}^l x)$ for $x \in \zeta_{\tau'} \cap \gamma^u$. Now $\hat{T}^l(\zeta_{\tau'} \cap \gamma^u) = \Lambda \cap \hat{\gamma}^u$, because $s(\zeta_{\tau'}) = l$. For the numerator we obtain

$$m_{\hat{\gamma}^u}(T^{-b}B_\rho \cap \Lambda) \leq c_1 \int m_{\hat{\gamma}^u}(T^{-b}B_\rho \cap \Lambda) \, d\nu(\hat{\gamma}^u) \leq c_1 \mu(T^{-b}B_\rho) = c_1\mu(B_\rho)$$

for some c_1, and for the denominator we use that $\mu_{\hat{\gamma}^u}(\Lambda) \geq c_2$ for some $c_2 > 0$. The lemma now follows with $C_6 = C_3 c_1 / c_2$. $\qquad\square$

Lemma 4.5. *Consider a collection of cylinders $\zeta_{\tau'} = \zeta_{i_0, \ldots, i_{l-1}} \in \Lambda_{i_0}$ such that*

$$i) \quad \exists \tau \subset \tau' \text{ with } \tau \in I_{i_0, j, n},$$
$$ii) \quad \zeta_{\tau'} \cap T^{-j}B_\rho \neq \varnothing.$$

Then there exists a constant C_7 such that

$$\sum_{\substack{\tau' \,|\, \exists \tau \subset \tau' \\ \tau \in I \\ T^{-j}B_\rho \cap \zeta_{\tau'} \neq \varnothing}} m(\tilde{\zeta}_{\tau'}) \leq m(\gamma^s(T^{-j}(B_{\rho + C_7\,\alpha^{n/s}}) \cap \Lambda_{i_0})),$$

where we use the notation $\gamma^s(\mathcal{B}) = \bigcup_{\gamma^s \in \Gamma^s : \gamma^s \cap \mathcal{B}} \gamma^s$ (note that here $\mathcal{B} \subset \Lambda$).

Proof. All the unions, sums and maxima in this proof are subscripted with "$\tau' \,|\, \exists \tau \subset \tau', \tau \in I_{i_0, j, n}, T^{-j}B_\rho \cap \zeta_{\tau'} \neq \varnothing$," unless otherwise specified. From Lemma 4.3 we know that

$$|\zeta_{\tau'} \cap \gamma^u| \leq C_0 \, \alpha^{s(\zeta_{\tau'})} = C_0 \, \alpha^l$$

for unstable leaves γ^u. Without loss of generality we can assume that $\tilde{\zeta}_{\tau'} = \zeta_{\tau'}$. Since $\zeta_{\tau'}$ contains a cylinder τ satisfying (n, j)-minimality, we deduce from $n + j < \sum_{k=0}^l R_{i_k} \leq (l+1)\, s$ a lower bound on l:

$$l \geq \frac{n+j}{s} - 1 \geq \frac{n}{s} - 1.$$

Thus

$$|\zeta_{\tau'} \cap \gamma^u| \leq C_0 \, \alpha^{\frac{n}{s} - 1}$$

and for $j < R_{i_0}$ we can further say by Assumption (A1)(b) that

$$|T^j(\zeta_{\tau'} \cap \gamma^u)| \le C_0\, \alpha^{\frac{n}{s}-2}$$

as $s(T^j\zeta_{\tau'}) = s(\zeta_{\tau'}) - 1$ for $1 \le j < R_{i_0}$. Now put $C_7 = c_0\alpha^{-2}$. Since $\zeta_{\tau'} \cap T^{-j}B_\rho \ne \varnothing$ and therefore $T^j\zeta_{\tau'} \cap B_\rho \ne \varnothing$ we obtain

$$T^j\zeta_{\tau'} \subset \bigcup_{\gamma^s \in \Gamma^s:\gamma^s \cap T^{-j}B_{\rho+C_7\,\alpha^{n/s}} \cap \Lambda \ne \varnothing} T^j\gamma^s.$$

As the above estimate works for any cylinder whose subscript τ' satisfies "$\tau' \,|\, \exists \tau \subset \tau', \tau \in I_{i_0,j,n}, T^{-j}B_\rho \cap \zeta_{\tau'} \ne \varnothing$," we have

$$\bigcup \zeta_{\tau'} \subset \gamma^s(T^{-j}(B_{\rho+C_7\,\alpha^{n/s}}) \cap \Lambda).$$

Moreover, since $\bigcup \zeta_{\tau'} \subset \Lambda_{i_0}$ we in fact have

$$\bigcup \zeta_{\tau'} \subset \gamma^s(T^{-j}(B_{\rho+C_7\,\alpha^{n/s}}) \cap \Lambda_{i_0}),$$

and we can conclude

$$\sum m(\zeta_{\tau'}) = m\left(\bigcup \zeta_{\tau'}\right) \le m(\gamma^s(T^{-j}(B_{\rho+C_7\,\alpha^{n/s}}) \cap \Lambda_{i_0})).$$

$\square$

4.4. Measure of the forbidden set $\mathcal{X}_\rho$.

We will need the following lemma.

Lemma 4.6. ([6] Lemma A.3) Let ℓ_0 and ℓ_1 be two finite positive measures on a D-dimensional Riemannian manifold M. For $\omega \in (0,1)$ and $\rho \in (0,1)$ define

$$\mathcal{D} = \{x \in M : \ell_1(B_\rho(x)) \ge \omega\ell_0(B_\rho(x))\}.$$

There exists an integer $p(D)$ such that

$$\ell_0(\mathcal{D}) \le p(D)\,\omega^{-1}\ell_1(M).$$

We will also need the following result on the size of the set where tall towers dominate, that is where R_i is larger than s. Recall that $\Omega(s) = [\sum_{i,R_i>s} R_i\, m(\Lambda_i)]^{\frac{1}{2}}$.

Lemma 4.7. For $n,s \ge 1$ there exist sets $\mathcal{D}_{n,s} \subset M$ such that the non-principal part contributions are estimated as

$$\sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B} \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) < \sqrt{n+2}\,\Omega(s)\mu(B_\rho)$$

for any $\mathcal{B} \subset B_\rho(x)$ and $x \notin \mathcal{D}_{n,s}$ where ($p(D)$ as above)

$$\mu(\mathcal{D}_{n,s}) \le p(D)\sqrt{n+2}\,\Omega(s).$$

Proof. We employ Lemma 4.6, with $\ell_0 = \mu$ and $\ell_1(\cdot) = \sum_i \sum_{j=0}^{R_i-1} m(T^{-j}(\cdot) \cap (\Lambda_i \setminus \tilde{\Lambda}_i))$. Define

$$\mathcal{D}_{n,s} = \{x \in M : \ell_1(B_\rho(x)) \ge \sqrt{n+2}\,\Omega(s)\mu(B_\rho(x))\};$$

from Lemma 4.6 we know that $\mu(\mathcal{D}_{n,s}) \le p(D)\,(\sqrt{n+2}\,\Omega(s))^{-1}\ell_1(M)$. Since $R(\hat{T}^l x) > s$ exactly if $x \in \hat{T}^{-l}\{R > s\}$ we get

$$\Lambda_i \setminus \tilde{\Lambda}_i = \{x \in \Lambda_i : \exists l \le n, \text{ such that } R(\hat{T}^l x) \ge s \} = \bigcup_{l=0}^{n} \hat{T}^{-l}\{R \ge s\} \cap \Lambda_i.$$

Since $\tilde{\Lambda}_i = \varnothing$ for $R_i > i$ we bound the measure $\ell_1(M)$ as follows

$$
\begin{aligned}
\ell_1(M) &\leq \sum_{i, R_i \leq s} \sum_{j=0}^{R_i-1} m(\Lambda_i \setminus \tilde{\Lambda}_i) + \sum_{i, R_i > s} \sum_{j=0}^{R_i-1} m(\Lambda_i) \\
&\leq \sum_{i, R_i \leq s} s\, m(\Lambda_i \setminus \tilde{\Lambda}_i) + \sum_{i, R_i > s} R_i\, m(\Lambda_i) \\
&\leq s \sum_{i, R_i \leq s} m\left(\bigcup_{l=0}^{n} \hat{T}^{-l}\{R > s\} \cap \Lambda_i\right) + \Omega(s)^2 \\
&\leq s\, m\left(\bigcup_{l=0}^{n} \hat{T}^{-l}\{R > s\}\right) + \Omega(s)^2 \\
&\leq s(n+1)m(\{R > s\}) + \Omega(s)^2 \\
&\leq (n+2)\Omega(s)^2
\end{aligned}
$$

as $s\, m(\{R > s\}) \leq \sum_{j=s+1}^{\infty} j\, m(\{R = j\}) = \Omega(s)^2$. Hence

$$\mu(\mathcal{D}_{n,s}) \leq p(D)(\sqrt{n+2}\,\Omega(s))^{-1}\ell_1(M) = p(D)\sqrt{n+2}\,\Omega(s).$$

Outside the set $\mathcal{D}_{n,s}$ we have

$$
\begin{aligned}
\sum_{i, R_i \leq s} \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B} \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) &\leq \sum_{i, R_i \leq s} \sum_{j=0}^{R_i-1} m(T^{-j}(B_\rho) \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) \\
&= \ell_1(B_\rho) \\
&\leq \sqrt{n+2}\,\Omega(s)\mu(B_\rho).
\end{aligned}
$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Let $\eta \in (\frac{3}{\lambda-1}, 1)$ and $\beta < 1 - \frac{3}{\lambda-1}$ be according to Assumption (A4). Put $\hat{\sigma} = \frac{\lambda-1}{2}\min\{\eta, (1-\beta)\} - \frac{3}{2}$. By assumption $\hat{\sigma} > 0$ and we can finally estimate the size of the forbidden set defined by

$$\mathcal{X}_\rho = \bigcup_{n=J}^{p-1} \left(\mathcal{D}_{n,n^\eta} \cup \mathcal{D}_{n,n^{1-\beta}}\right) \cup \mathcal{E}_\rho$$

whose parts have essentially been estimated in the previous lemma and assumption (A4).

Proposition 4.1. *There exist a constant C_8 such that*

$$\mu(\mathcal{X}_\rho) \leq C_8|\log\rho|^{-\hat{\sigma}}.$$

Proof. We estimate the contributions to $\mathcal{X}_\rho$ separately in the three following paragraphs.
(I) By Lemma 4.7 and since $\Omega(s) \lesssim s^{-\frac{\lambda-1}{2}}$

$$\mu\left(\bigcup_{n=J}^{p-1} \mathcal{D}_{n,n^\eta}\right) \leq p(D) \sum_{n=J}^{p-1} \sqrt{n+2}\,\Omega(n^\eta) \leq c_1 \sum_{n=J}^{\infty} n^{\frac{1}{2}}(n^\eta)^{-\frac{\lambda-1}{2}} \leq c_2 J^{\frac{\eta+3-\eta\lambda}{2}}$$

as $\eta + 1 - \eta\lambda < 0$ and $J = \lfloor \mathfrak{a}|\log\rho|\rfloor$. The term with $s = n^{1-\beta}$ is estimated similarly. Therefore

$$\mu\left(\bigcup_{n=J}^{p-1}\mathcal{D}_{n,n^\eta}\right) + \mu\left(\bigcup_{n=J}^{p-1}\mathcal{D}_{n,n^{1-\beta}}\right) \leq c_3|\log\rho|^{-\hat{\sigma}}.$$

(II) By Assumption (A4) $\mu(\mathcal{E}_\rho) \leq |\log\rho|^{-\lambda/2}$.
Combining the estimates from (I) and (II) results in

$$\mu(\mathcal{X}_\rho) \leq \mu\left(\bigcup_{n=J}^{p-1}\mathcal{D}_{n,s}\right) + \mu\left(\bigcup_{n=J}^{p-1}\mathcal{D}_{n,n^{1-\beta}}\right) + \mu(\mathcal{E}_\rho) \leq C_8|\log\rho|^{-\hat{\sigma}}$$

for some C_8 as $\hat{\sigma} < \frac{\lambda}{2}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

4.5. **Poisson approximation of the return times distribution.** To prove Theorem 1 we will employ the Poisson approximation theorem from Section 7. Let x be a point in the phase space and $B_\rho := B_\rho(\mathsf{x})$ for $\rho > 0$. Let $X_n = \mathbb{1}_{B_\rho} \circ T^{n-1}$, then we put $N = \lfloor t/\mu(B_\rho)\rfloor$, where t is a positive parameter. We write $S_a^b = \sum_{n=a}^b X_n$ (and $S = S_1^N$). Then for any $2 \leq p \leq N$ (C_{12} from Section 7)

$$\left| \mathbb{P}(S = k) - \frac{t^k}{k!}e^{-t}\right| \leq C_{12}(N(\mathcal{R}_1 + \mathcal{R}_2) + p\,\mu(B_\rho)), \qquad (14)$$

where

$$\mathcal{R}_1 = \sup_{\substack{0<j<N-p \\ 0<q<N-p-j}} \left| \mathbb{E}(\mathbb{1}_{B_\rho}\mathbb{1}_{S_{p+1}^{N-j}=q}) - \mu(B_\rho)\,\mathbb{E}(\mathbb{1}_{S_{p+1}^{N-j}=q})\right|$$

$$\mathcal{R}_2 = \sum_{n=1}^{p-1} \mathbb{E}(\mathbb{1}_{B_\rho}\,\mathbb{1}_{B_\rho}\circ T^n).$$

Since we restrict to the complement of the set $\mathcal{V}_\rho$ (cf. (1)) we have from now on

$$\mathcal{R}_2 = \sum_{n=J}^{p-1} \mu(B_\rho \cap T^{-n}B_\rho),$$

where $J = \lfloor \mathfrak{a}|\log\rho|\rfloor$. Note that if $k > N$ then $\mathbb{P}(S = k) = 0$ and

$$\left|\mathbb{P}(S = k) - \frac{t^k}{k!}e^{-t}\right| = \frac{t^k}{k!}e^{-t} \leq |\log\rho|^{-\frac{\lambda-4}{2}} \qquad \forall k > N \qquad (15)$$

using the fact that $\mu(B_\rho) \lesssim \rho^{\varsigma'}$ and for ρ sufficiently small.

We now proceed to estimate the error between the distribution of S and a Poissonian for $k \leq N$ based on Theorem 7.1.

4.6. **Estimating $\mathcal{R}_1$.** By invariance of the measure μ we can also write

$$\mathcal{R}_1 = \sup_{\substack{0<j<N-p \\ 0<q<N-p-j}} \left| \mu(B_\rho \cap T^{-p}\{S_1^{N-j-p} = q\}) - \mu(B_\rho)\,\mu(\{S_1^{N-j-p} = q\})\right|.$$

We now use the decay of correlations (8) to obtain an estimate for $\mathcal{R}_1$. Approximate $\mathbb{1}_{B_\rho}$ by Lipschitz functions from above and below as follows:

$$\phi(x) = \begin{cases} 1 & \text{on } B_\rho \\ 0 & \text{outside } B_{\rho+\delta\rho}\end{cases} \qquad \text{and} \qquad \tilde{\phi}(x) = \begin{cases} 1 & \text{on } B_{\rho-\delta\rho} \\ 0 & \text{outside } B_\rho\end{cases}$$

with both functions linear within the annuli. The Lipschitz norms of both ϕ and $\tilde{\phi}$ are equal to $1/\delta\rho$ and $\tilde{\phi} \le \mathbb{1}_{B_\rho} \le \phi$.

We obtain

$$\mu(B_\rho \cap \{S_p^{N-j} = q\}) - \mu(B_\rho)\,\mu(\{S_1^{N-j-p} = q\})$$

$$\le \int_M \phi\,(\mathbb{1}_{S_p^{N-j}=q})\,d\mu - \int_M \mathbb{1}_{B_\rho}\,d\mu \int_M \mathbb{1}_{S_1^{N-j-p}=q}\,d\mu$$

$$= X + Y$$

where

$$X = \left(\int_M \phi\,d\mu - \int_M \mathbb{1}_{B_\rho}\,d\mu \right) \int_M \mathbb{1}_{S_1^{N-j-p}=q}\,d\mu$$

$$Y = \int_M \phi\,(\mathbb{1}_{S_p^{N-j}=q})\,d\mu - \int_M \phi\,d\mu \int_M \mathbb{1}_{S_1^{N-j-p}=q}\,d\mu.$$

The two terms X and Y are estimated separately. The first term is estimated as follows:

$$X \le \int_M \mathbb{1}_{S_1^{N-j-p}=q}\,d\mu \int_M (\phi - \mathbb{1}_{B_\rho})\,d\mu \le \mu(B_{\rho+\delta\rho} \setminus B_\rho).$$

In order to estimate the second term Y we use the decay of correlations and have to approximate $\mathbb{1}_{S_1^{N-j-p}=q}$ by a function which is constant on local stable leaves. For that purpose put

$$\mathcal{S}_n = \bigcup_k \bigcup_{j=0}^{R_k-1} \bigcup_{\substack{\gamma^s \in \Gamma^s \\ \gamma^s \subset T^{-n-j}B_\rho \cap \Lambda_k}} T^{n+j}\gamma^s, \qquad \partial\mathcal{S}_n = \bigcup_k \bigcup_{j=0}^{R_k-1} \bigcup_{\substack{\gamma^s \in \Gamma^s \\ \gamma^s \cap T^{-n-j}B_\rho \cap \Lambda_k \neq \varnothing}} T^{n+j}\gamma^s$$

and

$$\mathscr{S}_p^{N-j} = \bigcup_{n=p}^{N-j} \mathcal{S}_n, \qquad \partial\mathscr{S}_p^{n-j} = \bigcup_{n=p}^{N-j} \partial\mathcal{S}_n.$$

The set

$$\mathscr{S}_p^{N-j}(q) = \{S_p^{N-j} = q\} \cap \mathscr{S}_p^{N-j}$$

is then a union of local stable leaves. This follows from the fact that by construction $T^{n+j}y \in B_\rho$ if and only if $T^{n+j}\gamma^s(y) \subset B_\rho$ for a point $y \in T^{-j}B_\rho \cap \Lambda_k$. We also have $\{S_p^{N-j} = q\} \subset \tilde{\mathscr{S}}_p^{N-j}(q)$ where the set $\tilde{\mathscr{S}}_p^{N-j}(q) = \mathscr{S}_p^{N-j}(q) \cup \partial\mathscr{S}_p^{N-j}$ is a union of local stable leaves.

Denote by ψ_p^{N-j} the characteristic function of $\mathscr{S}_p^{N-j}(q)$ and by $\tilde{\psi}_p^{N-j}$ the characteristic function of $\tilde{\mathscr{S}}_p^{N-j}(q)$. Then ψ_p^{N-j} and $\tilde{\psi}_p^{N-j}$ are constant on local stable leaves and satisfy

$$\psi_p^{N-j} \le \mathbb{1}_{S_p^{N-j}=q} \le \tilde{\psi}_p^{N-j}.$$

Since $\{y : \psi_p^{N-j}(y) \neq \tilde{\psi}_p^{N-j}(y)\} \subset \partial\mathscr{S}_p^{N-j}$ we need to estimate the measure of $\partial\mathscr{S}_p^{N-j}$.

For integers n and s let $\tilde{\Lambda}_i$ be as before, then by Lemma 4.7 for $\mathsf{x} \notin \mathcal{D}_{n,s}$ we have

$$\sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B} \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) \le \sqrt{n+2}\,\Omega(s)\mu(B_{\rho+\alpha^l}),$$

where $\mathcal{B} = B_{\rho+\alpha^l} \setminus B_{\rho-\alpha^l}$. For points $y \in \Lambda \setminus \tilde{\Lambda}$ we let l be so that $R^l(y) \le n < R^{l+1}(y)$ then we get $l \ge n/s \ge n^\beta$ where we choose $s(n) = \lfloor n^{1-\beta} \rfloor$ and $\beta < 1$ so

that $\sum_n^\infty g(n^\beta)^{-1} < \infty$ in accordance with Assumption (A4). By the contraction property $\operatorname{diam}(T^n\gamma^s(y)) \le A^s\alpha^{n/s} \le A^{n^{1-\beta}}\alpha^{n^\beta}$ for all $y \in \Lambda \setminus \tilde\Lambda$, where the factor A^s accounts for the lack of control on the contraction between returns to the base Λ. If $\beta > \frac{1}{2}$ then $\operatorname{diam}(T^n\gamma^s(y)) \le 2\alpha^{n^\beta}$ for all n large enough. Consequently

$$\bigcup_{\substack{\gamma^s \subset \Lambda\setminus\tilde\Lambda \\ T^n\gamma^s \subset B_\rho}} T^n\gamma^s \subset B_{\rho+2\alpha^l} \setminus B_{\rho-2\alpha^l}$$

and therefore

$$
\begin{aligned}
\mu(\partial\mathscr{S}_p^{N-j}) &\le \mu\left(\bigcup_{n=p}^{N-j} T^{-n}\left(B_{\rho+2\alpha^l} \setminus B_{\rho-2\alpha^l}\right)\right) \\
&\le \sum_{n=p}^{N-j} \mu(B_{\rho+2\alpha^l} \setminus B_{\rho-2\alpha^l}) \\
&\le \sum_{n=p}^{\infty} \mu(B_\rho)\frac{1}{g(w)|\log\rho|^a} \\
&\le c_1\mu(B_\rho)\frac{1}{|\log\rho|^a}
\end{aligned}
$$

where we used $w(n) = n^\beta \frac{\log\alpha}{\log\rho}$. If we split $p = p' + p''$ then we can estimate as follows:

$$
\begin{aligned}
Y &= \left|\int_M \phi\, T^{-p'}(\mathbb{1}_{S_{p''}^{N-j-p'}=q})\,d\mu - \int_M \phi\,d\mu \int_M \mathbb{1}_{S_1^{N-j-p}=q}\,d\mu\right| \\
&\le \varphi_{p'}\|\phi\|_{Lip}\|\mathbb{1}_{\tilde{\mathscr{S}}_{p''}^{N-j-p'}}\|_{\mathscr{L}^\infty} \\
&\quad + 2\sum_{n=p''}^{\infty}\sum_i\sum_{k=0}^{R_i-1} m(T^{-k}(B_{\rho+\alpha^l} \setminus B_{\rho-\alpha^l}) \cap (\Lambda_i \setminus \tilde\Lambda_i)) + 2\mu(\partial\mathscr{S}_{p''}^{N-j})
\end{aligned}
$$

where the triple sum on the RHS is by Lemma 4.7 bounded by

$$2\sum_{n=p''}^{\infty} \sqrt{n+2}\,\Omega(s)\mu(B_{\rho+\alpha^l}) \le c_2\mu(B_\rho)\sum_{n=p''}^{\infty} n^{\frac{1}{2}-(1-\beta)\frac{\lambda-1}{2}} \le c_3\mu(B_\rho)p^{\frac{3}{2}-(1-\beta)\frac{\lambda-1}{2}}$$

assuming $\frac{3}{2} - (1-\beta)\frac{\lambda-1}{2} < 0$ where in the last estimate we put $p' = p/2$. Since $\beta > \frac{1}{2}$ this requires $\lambda > 7$. Now let $\delta\rho = \rho^w$ where $w \in (1,\xi)$ is chosen in accordance with Assumption (A4) (this is possible since $\xi = \varsigma(\lambda-2) > 1$). Hence

$$
\begin{aligned}
\mu(B_\rho \cap T^{-p}\{S_1^{N-j-p} = q\}) &- \mu(B_\rho)\,\mu(\{S_1^{N-j-p} = q\}) \\
&\le \varphi_{p/2}/\delta\rho + \mu(B_\rho \setminus B_{\rho-\delta\rho}) + c_3\mu(B_\rho)\left(p^{\frac{3}{2}-(1-\beta)\frac{\lambda-1}{2}} + |\log\rho|^{-a}\right) \\
&\le \varphi_{p/2}\rho^{-w} + c_4\,\mu(B_\rho)\left(p^{\frac{3}{2}-(1-\beta)\frac{\lambda-1}{2}} + |\log\rho|^{-a}\right)
\end{aligned}
$$

In the same way we obtain a lower estimate. Since $\mathcal{D}_{n,n^{1-\beta}} \subset \mathcal{X}_\rho$ for $n = J, J+1, \ldots$ we conclude that for $\mathsf{x} \notin \mathcal{X}_\rho$ one has:

$$\mathcal{R}_1 \le \varphi_{p/2}\rho^{-w} + c_4\,\mu(B_\rho)\left(p^{\frac{3}{2}-(1-\beta)\frac{\lambda-1}{2}} + |\log\rho|^{-a}\right). \tag{16}$$

Remark. If the contraction in the stable direction is known to be monotone between returns to the base, then the factor A^s above won't be needed and the lower

bound for β would thus be 0 and not $\frac{1}{2}$. In that case λ in Theorem 1 would be required to be greater than 4 and thus we obtain a convergence rate (in all three theorems) given by any κ less than $\frac{\lambda-4}{2}$.

4.7. **Estimating the individual terms of $\mathcal{R}_2$ (for n fixed).** We will estimate the measure of each of the summands comprising $\mathcal{R}_2$ individually with the help of the Young tower. Fix n and for the sake of simplicity we will denote $B_\rho \cap T^{-n}B_\rho$ by $\mathcal{B}_n$. Then

$$\mu(B_\rho \cap T^{-n}B_\rho) = \mu(\mathcal{B}_n) = \sum_{i=1}^{\infty} \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B}_n \cap \Lambda_i). \tag{17}$$

With $s = \lfloor n^\eta \rfloor$ let $\tilde{\Lambda}_i$ be as in Section 4.3 (II), then

$$\begin{aligned}
\mu(\mathcal{B}_n) &= \sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B}_n \cap \tilde{\Lambda}_i) + \sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B}_n \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) \\
&\leq \sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B}_n \cap \tilde{\Lambda}_i) + \sqrt{n+2}\,\Omega(s)\mu(B_\rho) \tag{18}
\end{aligned}$$

using Lemma 4.7 for the second term on the RHS to the complement of the set $\mathcal{D}_{n,s}$. Since ($\tilde{\zeta}$ as in Section 4.3)

$$m(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_\tau) \leq \sum_{\tau' \mid \exists \tau \subset \tau'} m(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_{\tau'}) = \sum_{\substack{\tau' \mid \exists \tau \subset \tau', \tau \in I \\ T^{-j}B_\rho \cap \zeta_{\tau'} \neq \varnothing}} m(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_{\tau'})$$

we get, since by (10) $\bigcup_i \tilde{\Lambda}_i \subset \bigcup_{\tau \in I} \tilde{\zeta}_\tau$, for each of the summands in the principal term on the RHS of (18)

$$\begin{aligned}
m_{\gamma^u}(T^{-j}\mathcal{B}_n \cap \tilde{\Lambda}_i) &\leq \sum_{\tau \in I_{i,j,n}} m_{\gamma^u}(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_\tau) \\
&\leq \sum_{\substack{\tau' \mid \exists \tau \subset \tau', \tau \in I \\ T^{-j}B_\rho \cap \zeta_{\tau'} \neq \varnothing}} m_{\gamma^u}(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_{\tau'}) \\
&= \sum_{\substack{\tau' \mid \exists \tau \subset \tau', \tau \in I \\ T^{-j}B_\rho \cap \zeta_{\tau'} \neq \varnothing}} \frac{m_{\gamma^u}(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_{\tau'})}{m_{\gamma^u}(\tilde{\zeta}_{\tau'})} m_{\gamma^u}(\tilde{\zeta}_{\tau'}) \\
&\leq C_6\,\mu(B_\rho) \sum_{\substack{\tau' \mid \exists \tau \subset \tau', \tau \in I \\ T^{-j}B_\rho \cap \zeta_{\tau'} \neq \varnothing}} m_{\gamma^u}(\tilde{\zeta}_{\tau'})
\end{aligned}$$

where we used Lemma 4.4 in the last step. From Lemma 4.5 we obtain a bound for the sum of measures of the cylinders, whence

$$\sum_{\tau \in I} m(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_\tau) \leq C_6\,\mu(B_\rho)\,m(\gamma^s(T^{-j}(B_{\rho+C_7\,\alpha^{n/s}}) \cap \Lambda_i))$$

using the product structure of the measure m. Note that since $B_{\rho+C_7\,\alpha^{n/s}}(\mathsf{x}) \subset B_{2\rho}(\mathsf{x}) \cup B_{2C_7\alpha^{n/s}}(\mathsf{x})$ we obtain on unstable leaves by Assumption (A1)(b) that

$$m_{\gamma^u}(\gamma^s(T^{-j}(B_{\rho+C_7\,\alpha^{n/s}}) \cap \Lambda)) \leq m_{\gamma^u}(B_{2\rho}) + m_{\gamma^u}(B_{2C_7\alpha^{n/s}}) \leq (2\rho)^{\varsigma'} + (2C_7\alpha^{n/s})^{\varsigma'}$$

using the geometric regularity (A4)(b) provided that the radius $2C_7\,\alpha^{n/s}$ is small enough. Since $n \geq J$ and s depends on n we can guarantee the above radius to be sufficiently small provided that ρ is small enough. Therefore

$$m(\gamma^s(T^{-j}(B_{\rho + C_7\,\alpha^{n/s}}) \cap \Lambda_i)) \leq c_1 \left((2\rho)^{\varsigma'} + (2C_7\alpha^{n/s})^{\varsigma'}\right)$$

for some c_1. Thus the first term (principal term) on the RHS of (18) can be bounded as follows

$$\sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{B}_n \cap \tilde{\Lambda}_i) \leq \sum_i \sum_{j=0}^{R_i-1} \sum_{\tau \in I} m(T^{-j}\mathcal{B}_n \cap \tilde{\zeta}_\tau)$$

$$\leq \sum_i \sum_{j=0}^{R_i-1} C_6\,\mu(B_\rho)\,m(\gamma^s(T^{-j}(B_{\rho + C_7\,\alpha^{n/s}}) \cap \Lambda_i))$$

$$\leq C_6'\,\mu(B_\rho)\left(\rho^{\varsigma'} + (\alpha^{n/s})^{\varsigma'}\right) \tag{19}$$

for a constant C_6'.

4.8. Estimating $\mathcal{R}_2$. Combining inequalities (18) and (19) results in

$$\mu(B_\rho \cap T^{-n}B_\rho) = \mu(\mathcal{B}_n) \leq C_6'\,\mu(B_\rho)\left(\rho^{\varsigma'} + (\alpha^{n/s})^{\varsigma'}\right) + \Omega(s)\sqrt{n+2}\mu(B_\rho),$$

provided that $n \geq J$ and the center x of the ball B_ρ lies outside the set $\mathcal{D}_{n,s}$.

As before let $s = \lfloor n^\eta \rfloor$ where $\eta \in (\frac{3}{\lambda-1}, 1)$. Summing up the $\mathcal{B}_n$ terms over $n = J, \ldots, p-1$, we see that outside the set of forbidden ball centers $\mathcal{V}_\rho \cup \mathcal{X}_\rho$ we get with $\tilde{\alpha} = \alpha^{\varsigma'}$

$$\mathcal{R}_2 = \sum_{n=J}^{p-1} \mu(B_\rho \cap T^{-n}B_\rho) \leq \mu(B_\rho) \sum_{n=J}^{p-1} \left(C_6'\left(\rho^{\varsigma'} + \tilde{\alpha}^{n/s}\right) + 2\sqrt{n}\,\Omega(s)\right). \tag{20}$$

For ρ small and by Lemma 4.2 $\Omega(s) \leq C_5 s^{-\theta}$ where $\theta = \frac{\lambda-1}{2}$ we obtain

$$\mathcal{R}_2 \leq c_1\mu(B_\rho)\left(p\,\rho^{\varsigma'} + \sum_{n=J}^{p-1} \tilde{\alpha}^{n^{1-\eta}} + \sum_{n=J}^{\infty} n^{\frac{1}{2}-\eta\theta}\right)$$

$$\leq c_2\mu(B_\rho)\left(p\,\rho^{\varsigma'} + \tilde{\alpha}^{\frac{1}{2}J^{1-\eta}} + J^{\frac{3}{2}-\eta\theta}\right)$$

since $\sum_{n=J}^{\infty} \tilde{\alpha}^{n^{1-\eta}} \leq c_3\,\tilde{\alpha}^{\frac{1}{2}J^{1-\eta}}$ for some c_3 (and ρ small enough). As $\tilde{\alpha}^{\frac{1}{2}J^{1-\eta}} \leq J^{-\sigma}$, $\sigma = \eta\theta - \frac{3}{2}$, for ρ small, we get

$$\mathcal{R}_2 \leq C_9\,\mu(B_\rho)\left(p\,\rho^{\varsigma'} + J^{-\sigma}\right)$$

for some C_9. Note that the above it true provided ρ is sufficiently small and the center x of the ball B_ρ is not in $\mathcal{X}_\rho \cup \mathcal{V}_\rho$.

4.9. Estimate of the total error. Now we want to bound all of the error components from inequality (14) with terms of the order of $J^{-\sigma}$ or a negative power of $|\log \rho|$, where $\sigma = \eta\frac{\lambda-1}{2} - \frac{3}{2} < \frac{\lambda-4}{2}$. To that end we choose the length of the gap p to be

$$p = \left\lfloor J^{-\sigma}\rho^{-\varsigma'} \right\rfloor,$$

and estimate the three error terms on the RHS of (14) separately.

(I) The last summand is immediately estimated as

$$p\mu(B_\rho) \leq p\, c_1 \rho^{\varsigma'} \leq J^{-\sigma}.$$

(II) For the term involving $\mathcal{R}_1$ we obtain

$$N\mathcal{R}_1 \leq \frac{t}{\mu(B_\rho)}\left(\varphi_{p/2}\rho^{-w} + c_2\,\mu(B_\rho)\left(p^{\frac{3}{2}-(1-\beta)\frac{\lambda-1}{2}} + |\log\rho|^{-a}\right)\right).$$

Let $w < \xi$ and $\varsigma' < \varsigma$, $\hat{\varsigma}'' > \hat{\varsigma}$ so that $\varsigma'(\lambda-1) - \hat{\varsigma}'' - w > 0$. Since $\mu(B_\rho) \geq \rho^{\hat{\varsigma}''}$ for ρ small enough we get for the first term on the RHS for ρ small:

$$\frac{\varphi_{p/2}\rho^{-w}}{\mu(B_\rho)} \leq c_3\frac{p^{1-\lambda}}{\mu(B_\rho)\,\rho^w} \leq c_4\,\rho^{\varsigma'(\lambda-1)-\hat{\varsigma}''-w}\,J^{\sigma(\lambda-1)} \leq J^{-\sigma}$$

and so (with some c_5, c_6) since by Assumption (A4)(a) $\frac{3}{2} - (1-\beta)\frac{\lambda-1}{2} < 0$ we conclude

$$N\mathcal{R}_1 \leq c_5\left(J^{-2\sigma} + |\log\rho|^{-a}\right) \leq c_6|\log\rho|^{-\min\{a,\sigma\}}.$$

(III) Utilizing the estimate from Section 4.8 and using the fact that $N = \lfloor t/\mu(B_\rho)\rfloor$ yield

$$N\mathcal{R}_2 \leq t\,C_9\left(p\,\rho^{\varsigma'} + J^{-\sigma}\right) \leq c_7 t|\log\rho|^{-\sigma}$$

for some c_7.

Combining the results of estimates (I), (II) and (III) above we obtain for ρ sufficiently small the RHS of (14) as follows ($\mathsf{x} \notin \mathcal{X}_\rho$)

$$N(\mathcal{R}_1 + \mathcal{R}_2) + p\mu(B_\rho) \leq \frac{c_6}{|\log\rho|^{\min\{a,\sigma\}}} + \frac{c_7 t}{|\log\rho|^\sigma} + \frac{1}{J^\sigma} \leq \frac{c_8(1+t)}{|\log\rho|^\kappa}.$$

for some c_8, where $\kappa = \min\{\sigma, a\}$ is positive. This now concludes the proof of Theorem 1 as it shows that for ρ small enough and for any $\mathsf{x} \notin \mathcal{X}_\rho$ we have for $k \leq N$

$$\left|\mathbb{P}(S = k) - \frac{t^k}{k!}\,e^{-t}\right| \leq C(1+t^2)|\log\rho|^{-\kappa}.$$

By (15) this estimate also extends to $k > N$. By Proposition 4.1 the size of the forbidden set is then $\mu(\mathcal{X}_\rho) = \mathcal{O}(|\log\rho|^{-\kappa'})$ where $\kappa' = \hat{\sigma}$. The proof of Theorem 1 is thus complete. $\qquad\square$

Remark. In the case when $a \geq \sigma$ and $\beta > \frac{1}{2}$ arbitrarily close to $\frac{1}{2}$ then $\kappa = \kappa' = \frac{\lambda-7}{4}$.

5. **Very short returns and proof of Theorem 2.** In this section we prove Theorem 2 which is a direct consequence of Theorem 1 and Proposition 5.1 below, which estimates the measure of the set with short return times. We first state the precise assumptions. Again we use the Young tower construction.

5.1. **Assumptions.** Let (M, T) be a dynamical system equipped with a metric d. Assume that the map $T : M \to M$ is a C^2-diffeomorphism. As at the start of the paper the set $\mathcal{V}_\rho \subset M$ is given by

$$\mathcal{V}_\rho = \{\mathsf{x} \in M : B_\rho(\mathsf{x}) \cap T^n B_\rho(\mathsf{x}) \neq \varnothing \text{ for some } 1 \leq n < J\},$$

where $J = \lfloor \mathfrak{a}\,|\log\rho|\rfloor$ and $\mathfrak{a} = (4\log A)^{-1}$ with

$$A = \|DT\|_{\mathscr{L}^\infty} + \|DT^{-1}\|_{\mathscr{L}^\infty}$$

$(A \geq 2)$. Suppose the system can be modeled by a Young tower possessing a reference measure m and that the greatest common divisor of the return times R_i is equal to one. Let μ be the SRB measure associated to the system. We require the following in Proposition 5.1:

(B1) *Regularity of the Jacobian and the metric on the Tower*
The same as Assumption (A1)

(B2) *Polynomial Decay of the Tail*
There exist constants C_1 and $\lambda > 9$ such that

$$\hat{m}_{\gamma^u}(R > k) \leq C_1 \, k^{-\lambda}$$

for unstable leaves γ^u.

(B3) *Geometric regularity of the measure*
Let ς be the dimension of the measure μ. Suppose that the positive constant $\varsigma' < \varsigma$ is fixed. There exists a constant $C_2 > 0$ and a set $\mathcal{E}_\rho \subset M$ satisfying $\mu(\mathcal{E}_\rho) \leq C_2 |\log \rho|^{-\lambda/2}$ such that

$$m_{\gamma^u}(B_\rho(\mathsf{x})) \leq C_2 \, \rho^{\varsigma'}$$

for all $\mathsf{x} \notin \mathcal{E}_\rho$ and ρ small.

Note that in Proposition 5.1 we don't require the measure to be ξ-regular.

5.2. **Estimate on the measure of $\mathcal{V}_\rho$.** Before we prove the main result of this section we shall present a lemma which will be needed in the proof of Proposition 5.1.

Lemma 5.1. *Let $\hat{C}_1 > 0$ be a constant and $\hat{\alpha} \in (0,1)$. Then for all sufficiently small ρ*

$$4 \, A^{n+b}\rho + A^{b+j}\hat{C}_1\hat{\alpha}^{\,|\log \rho|^{3/4}} \leq e^{-|\log \rho|^{1/4}}.$$

for any $n \leq J$ and $b, j \leq J^{\frac{1}{4}}$ ($J = \lfloor \mathfrak{a}|\log \rho| \rfloor$).

Proof. By assumption

$$A^{n+b}\rho \leq A^{2J}\rho \leq 2^{-1}A^{2\mathfrak{a}|\log \rho|}\rho = 2^{-1}\rho^{1-2\mathfrak{a}\log A} = 2^{-1}\rho^{1/2}$$

and also

$$A^{b+j}\hat{\alpha}^{|\log \rho|^{\frac{3}{4}}} \leq A^{1\mathfrak{a}|\log \rho|^{\frac{1}{4}}}\hat{\alpha}^{|\log \rho|^{\frac{3}{4}}} \leq (2\hat{C}_1)^{-1}e^{-|\log \rho|^{\frac{1}{4}}}.$$

Since $\rho^{\frac{1}{2}} = e^{-\frac{1}{2}|\log \rho|}$ the statement of the lemma follows for ρ small enough. $\qquad\square$

Now we can show that the set of centres where small balls have very short returns is small. To be precise we have the following result:

Proposition 5.1. *There exist constants $C_{10} > 0$ such that for all ρ small enough*

$$\mu(\mathcal{V}_\rho) \leq \frac{C_{10}}{|\log \rho|^{\frac{\lambda-9}{4}}}.$$

Proof. We largely follow the proof of Lemma 4.1 of [6]. Let us note that since T is a diffeomorphism one has

$$B_\rho(\mathsf{x}) \cap T^n B_\rho(\mathsf{x}) \neq \varnothing \qquad \Longleftrightarrow \qquad B_\rho(\mathsf{x}) \cap T^{-n}B_\rho(\mathsf{x}) \neq \varnothing.$$

We partition $\mathcal{V}_\rho$ into level sets $\mathcal{N}_\rho(n)$ as follows

$$\mathcal{V}_\rho = \{\mathsf{x} \in M : B_\rho(\mathsf{x}) \cap T^{-n}B_\rho(\mathsf{x}) \neq \varnothing \text{ for some } 1 \leq n < J\} = \bigcup_{n=1}^{J-1} \mathcal{N}_\rho(n)$$

where
$$\mathcal{N}_\rho(n) = \{x \in M : B_\rho(x) \cap T^{-n}B_\rho(x) \neq \varnothing\}.$$
The above union is split into two collections $\mathcal{V}_\rho^1$ and $\mathcal{V}_\rho^2$, where
$$\mathcal{V}_\rho^1 = \bigcup_{n=1}^{\lfloor \mathfrak{b}J \rfloor} \mathcal{N}_\rho(n) \quad \text{and} \quad \mathcal{V}_\rho^2 = \bigcup_{n=\lceil \mathfrak{b}J \rceil}^{J} \mathcal{N}_\rho(n).$$

and where the constant $\mathfrak{b} \in (0,1)$ will be chosen below. In order to find the measure of the total set we will estimate the measures of the two parts separately.

(I) Estimate of $\mathcal{V}_\rho^2$

We will derive a uniform estimate for the measure of the level sets $\mathcal{N}_\rho(n)$ when $n > \mathfrak{b}J$

$$\mu(\mathcal{N}_\rho(n)) = \sum_{i=1}^{\infty} \sum_{j=0}^{R_i-1} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \Lambda_i)$$

$$= \sum_i \sum_{j=0}^{R_i-1} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\Lambda}_i) + \sum_i \sum_{j=0}^{R_i-1} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) \quad (21)$$

as $\mu(\mathcal{N}_\rho(n)) = \mu(T^{-n}\mathcal{N}_\rho(n))$ and where here we put $s = J^{\frac{1}{4}}$ which means
$$\tilde{\Lambda}_i = \{x \in \Lambda_i : \forall l \leq n, \ R(\hat{T}^l x) \leq J^{\frac{1}{4}}\}.$$

On the RHS the first term is the principal term, and the other two terms will be treated as error terms. Let us first estimate the error term. Note that like in the proof of Lemma 4.7 and assumption (B2) the second sum on the RHS can be bounded by

$$\sum_i \sum_{j=0}^{R_i-1} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap (\Lambda_i \setminus \tilde{\Lambda}_i)) \leq \sum_i \sum_{j=0}^{R_i-1} m(\Lambda_i \setminus \tilde{\Lambda}_i)$$

$$\leq \sum_i R_i \, m(\Lambda_i \setminus \tilde{\Lambda}_i)$$

$$\leq J^{\frac{1}{4}}(n+1) \, m(R > J^{\frac{1}{4}})$$

$$\leq c_1 n J^{-\frac{\lambda-1}{4}}.$$

As for the first sum (principal term) on the RHS in (21), we can decompose the beam base $\tilde{\Lambda}_i$ into cylinder sets, as in Section 4.3, with the index set I defined just as in (11)
$$\tilde{\Lambda}_i \subset \bigsqcup_{\tau \in I_{i,j,n}} \tilde{\zeta}_\tau.$$
Then
$$m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\Lambda}_i) \leq \sum_{\tau \in I} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_\tau) \leq \sum_{\substack{\tau'|\exists \tau \in I \\ \tau \subset \tau'}} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'}).$$

Incorporating the above estimates and decomposition into (21), we obtain ($c_2 = c_1 + C_5^2$)

$$\mu(\mathcal{N}_\rho(n)) \leq \sum_i \sum_{j=0}^{R_i-1} \sum_{\tau'|\exists \tau \in I} m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'}) + c_2 n J^{-\frac{\lambda-1}{4}}. \qquad (22)$$

We will consider each of the measures $m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'})$ separately. Let $\tau = (i_0, \ldots, i_l) \in I_{i_0,j,n}$ and $\tau' = (i_0, \ldots, i_{l-1})$ as in Section 4.3. By distortion of the Jacobian, Lemma 4.1, we obtain for $\tilde{\zeta}_{\tau'} \neq \varnothing$ (which implies $\tilde{\zeta}_{\tau'} = \zeta_{\tau'}$):

$$m_{\gamma^u}(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'}) = \frac{m_{\gamma^u}(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'})}{m_{\gamma^u}(\tilde{\zeta}_{\tau'})} m_{\gamma^u}(\tilde{\zeta}_{\tau'})$$

$$\leq C_3 \frac{m_{\hat{\gamma}^u}(\hat{T}^l(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'}))}{m_{\hat{\gamma}^u}(\Lambda)} m_{\gamma^u}(\tilde{\zeta}_{\tau'}), \quad (23)$$

where, as before, $\hat{\gamma}^u = \gamma^u(\hat{T}^l x)$ for $x \in \zeta_{\tau'} \cap \gamma^u$. For the last line compare with (13). We estimate the numerator by finding a bound for the diameter of the set. Let the points x and z in $T^{-n}\mathcal{N}_\rho(n)$ be such that $T^{-j}x, T^{-j}z \in T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'} \cap \gamma^u$ for an unstable leaf γ^u. From the (n,j)-minimality of τ $(R^l \leq n + j < R^{l+1})$ we know that

$$\hat{T}^l = T^{n+j-b}, \quad \text{where} \quad b = n + j - R^l < R_{i_l},$$

therefore

$$d(\hat{T}^l T^{-j}x, \hat{T}^l T^{-j}z) = d(T^{n-b}x, T^{n-b}z).$$

Incorporating the definition of the constant A

$$d(T^{n-b}x, T^{n-b}z) \leq A^b \, d(T^n x, T^n z).$$

Note that $T^n x, T^n z \in \mathcal{N}_\rho(n)$ and so

$$d(T^n x, T^n z) \leq d(T^n x, x) + d(x, z) + d(z, T^n z) \leq 4A^n \rho + d(x, z).$$

Further,

$$d(x, z) \leq A^j \, d(T^{-j}x, T^{-j}z).$$

Now, $T^{-j}x, T^{-j}z$ are both elements of $\zeta_{\tau'}$ and $s(\zeta_{\tau'}) = l$. Thus

$$s(T^{-j}x, T^{-j}z) \geq s(\zeta_{\tau'}) = l.$$

We derive the lower bound on l from (n,j)-minimality which yields

$$n \leq n + j < R^{l+1} \leq (l+1)J^{1/4} \implies l > nJ^{-1/4} - 1.$$

Then employing (A1)(b) and keeping in mind that $n > \mathfrak{b}J$

$$\begin{aligned}
d(T^{-j}x, T^{-j}z) &\leq C_0 \, \alpha^{s(T^{-j}x, T^{-j}z)} \\
&\leq (C_0/\alpha) \, \alpha^{nJ^{-1/4}} \\
&\leq (C_0/\alpha) \, \alpha^{\mathfrak{b}JJ^{-1/4}} \\
&\leq (C_0/\alpha) \, \hat{\alpha}^{|\log \rho|^{3/4}},
\end{aligned}$$

where $\hat{\alpha} = \alpha^{2^{-3/4} \mathfrak{a}^{3/4} \mathfrak{b}} < 1$. Therefore

$$\begin{aligned}
d(\hat{T}^l T^{-j}x, \hat{T}^l T^{-j}z) &= d(T^{n-b}x, T^{n-b}z) \\
&\leq A^b \, d(T^n x, T^n z) \\
&\leq A^b \, (4A^n \rho + d(x, z)) \\
&\leq 4 \, A^{n+b}\rho + A^{b+j} \, d(T^{-j}x, T^{-j}z) \\
&\leq 4 \, A^{n+b}\rho + A^{b+j} \, (C_0/\alpha) \, \hat{\alpha}^{|\log \rho|^{3/4}},
\end{aligned}$$

and by Lemma 5.1 ($\hat{C}_1 = C_0/\alpha$) since $\mathfrak{b}J < n \leq J$ and $b, j < R_{i_l} < J^{\frac{1}{4}} \leq J$:

$$d(\hat{T}^l T^{-j}x, \hat{T}^l T^{-j}z) \leq e^{-|\log \rho|^{1/4}}.$$

Taking the supremum over all points x and z yields

$$|\hat{T}^l(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \zeta_{\tau'} \cap \gamma^u)| \leq e^{-|\log \rho|^{1/4}}.$$

By assumption (B3) on the relationship between the measure and the metric

$$m_{\hat{\gamma}^u}(\hat{T}^l(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \zeta_{\tau'})) \leq C_2 e^{-\varsigma'|\log \rho|^{1/4}}$$

for $\varsigma' < \varsigma$, which implies by the product structure of m that

$$m(\hat{T}^l(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \zeta_{\tau'})) \leq c_3 e^{-\varsigma'|\log \rho|^{1/4}}.$$

Incorporating the estimate into (23) yields

$$m(T^{-(n+j)}\mathcal{N}_\rho(n) \cap \tilde{\zeta}_{\tau'}) \leq c_4 \, e^{-\varsigma'|\log \rho|^{1/4}} m(\tilde{\zeta}_{\tau'}),$$

where $c_4 \leq \frac{2c_2 C_3}{m(\Lambda)}$. Substituting this into estimate (22) we see that $(n \leq J)$

$$\mu(\mathcal{N}_\rho(n)) \leq c_4 \, e^{-\varsigma'|\log \rho|^{1/4}} \sum_i \sum_{j=0}^{R_i-1} \sum_{\substack{\tau'|\exists \tau \in I \\ T^{-(n+j)}\mathcal{N}_\rho(n) \cap \zeta_{\tau'} \neq \varnothing}} m(\tilde{\zeta}_{\tau'}) + c_2 J^{-\frac{\lambda-5}{4}}.$$

Next we have to bound the triple sum on the RHS. As we showed before all of the $\tilde{\zeta}_{\tau'}$ with $\tau'|\exists \tau \in I$ are disjoint and are all subsets of $\tilde{\Lambda}_i$, therefore

$$\sum_i \sum_{j=0}^{R_i-1} \sum_{\substack{\tau'|\exists \tau \in I \\ T^{-(n+j)}\mathcal{N}_\rho(n) \cap \zeta_{\tau'} \neq \varnothing}} m(\tilde{\zeta}_{\tau'}) \leq \sum_i \sum_{j=0}^{R_i-1} m(\tilde{\Lambda}_i) \leq \mu(M) = 1$$

Hence for $n = \lceil \mathfrak{b}J \rceil, \ldots, J$ we obtain

$$\mu(\mathcal{N}_\rho(n)) \leq c_4 \, e^{-\varsigma'|\log \rho|^{1/4}} + c_2 J^{-\frac{\lambda-5}{4}} \leq c_5 J^{-\frac{\lambda-5}{4}}$$

for some c_5 and ρ small enough, and consequently

$$\mu(\mathcal{V}_\rho^2) \leq \sum_{n=\lceil \mathfrak{b}J \rceil}^{J} \mu(\mathcal{N}_\rho(n)) \leq \sum_{n=\lceil \mathfrak{b}J \rceil}^{J} c_5 J^{-\frac{\lambda-5}{4}} \leq c_6 |\log \rho|^{-\frac{\lambda-9}{4}} \tag{24}$$

for some constant c_6 (and ρ small enough) as $J = \lfloor \mathfrak{a}|\log \rho|\rfloor$.

(II) Estimate of $\mathcal{V}_\rho^1$

Here we consider the case $1 \leq n \leq \lfloor \mathfrak{b}J \rfloor$. Following [6] we put

$$s_p = 2^p \frac{A^{n \, 2^p} - 1}{A^n - 1}.$$

By [6] Lemma B.3 one has $\mathcal{N}_\rho(n) \subset \mathcal{N}_{s_p \rho}(2^p n)$ for any $p \geq 1$, and in particular for $p(n) = \lfloor \lg \mathfrak{b}J - \lg n \rfloor + 1$. Therefore

$$\bigcup_{n=1}^{\lfloor \mathfrak{b}J \rfloor} \mathcal{N}_\rho(n) \subset \bigcup_{n=1}^{\lfloor \mathfrak{b}J \rfloor} \mathcal{N}_{s_{p(n)}\rho}(2^{p(n)}n).$$

Now define

$$n' = n2^{p(n)} \qquad \text{and} \qquad \rho' = s_{p(n)}\rho.$$

A direct computation shows that $1 \leq n \leq \lfloor \mathfrak{b}J \rfloor$ implies $\lceil \mathfrak{b}J \rceil \leq n' \leq 2\mathfrak{b}J$ and so

$$\mathcal{V}_\rho^1 = \bigcup_{n=1}^{\lfloor \mathfrak{b}J \rfloor} \mathcal{N}_\rho(n) \subset \bigcup_{n=1}^{\lfloor \mathfrak{b}J \rfloor} \mathcal{N}_{s_{p(n)}\rho}(2^{p(n)}n) \subset \bigcup_{n'=\lceil \mathfrak{b}J \rceil}^{2\mathfrak{b}J} \mathcal{N}_{\rho'}(n').$$

Therefore to estimate the measure of $\mathcal{V}_\rho^1$ it suffices to find a bound for $\mathcal{N}_{\rho'}(n')$ when $n' \geq \mathfrak{b}J$. This is accomplished by using an argument analogous to the first part of the proof. We replace all the n with n' and ρ with ρ'. The cutoff $J^{1/4} = (\lfloor \mathfrak{a} \, | \log \rho | \rfloor)^{1/4}$ remains unchanged. We get for $\mathfrak{b} < 1/3$

$$\mu(\mathcal{N}_{\rho'}(n')) \leq c_4 \, e^{-\varsigma' |\log \rho'|^{1/4}} + c_2 n' J^{-\frac{\lambda-1}{4}} \leq c_7 J^{-\frac{\lambda-5}{4}}$$

and thus obtain an estimate similar to (24):

$$\mu(\mathcal{V}_\rho^1) \leq \sum_{n'=\lceil \mathfrak{b}J \rceil}^{2\mathfrak{b}J} \mu(\mathcal{N}_{\rho'}(n')) \leq \sum_{n'=\lceil \mathfrak{b}J \rceil}^{J} c_7 J^{-\frac{\lambda-5}{4}} \leq c_8 |\log \rho|^{-\frac{\lambda-9}{4}}.$$

(III) Final estimate

Overall we obtain for all ρ sufficiently small

$$\mu(\mathcal{V}_\rho) \leq \mu(\mathcal{V}_\rho^1) + \mu(\mathcal{V}_\rho^2) \leq C_{10} \, |\log \rho|^{-\frac{\lambda-9}{4}},$$

where $C_{10} = c_6 + c_8$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Remark. If condition (5) is satisfied for some $g \geq h$ and if $\lambda > 13$ then $\mathbb{P}(S_{\rho,\mathsf{x}}^t = k)$ converges almost surely to $e^{-t}\frac{t^k}{k!}$ as $\rho \to 0$.

Let $\rho_j = e^{-j^\alpha}$, $j = 1, 2, \ldots$, for some $\alpha \in (0,1)$ so that $\alpha\frac{\lambda-9}{4} > 1$. Then by Proposition 5.1

$$\sum_j \mu(\mathcal{V}_{\rho_j}) \leq C_{10} \sum_j j^{-\alpha\frac{\lambda-9}{4}} < \infty$$

and by the Borel-Cantelli lemma, μ-almost every x lies in at most finitely many of the sets $\mathcal{V}_{\rho_j}$. For any small $\rho > 0$ there is a j so that $\rho \in [\rho_j, \rho_{j-1})$. Then

$$\left| \mathbb{P}(S_{\rho,\mathsf{x}}^t = k) - \mathbb{P}(S_{\rho_j,\mathsf{x}}^t = k) \right| \leq N_{\rho_j,\mathsf{x}}^t \mu(B_\rho(\mathsf{x}) \setminus B_{\rho_j}(\mathsf{x})) \leq c_1 \frac{\mu(B_{\rho_{j-1}}(\mathsf{x}) \setminus B_{\rho_j}(\mathsf{x}))}{\mu(B_{\rho_j}(\mathsf{x}))}$$

and using (5) with $g \geq h$ we conclude that

$$\left| \mathbb{P}(S_{\rho,\mathsf{x}}^t = k) - \mathbb{P}(S_{\rho_j,\mathsf{x}}^t = k) \right| \leq c_2 \rho_j^{g-h} j^{g(1-\alpha)} \to 0$$

as $\rho \to 0$ (and thus $j \to \infty$) where we used that $\rho_j - \rho_{j-1} = \rho_j \mathcal{O}(j^{-(1-\alpha)})$. Hence $\mathbb{P}(S_{\rho,\mathsf{x}}^t = k)$ converges for μ-almost every x to the limit $e^{-t}\frac{t^k}{k!}$ as $\rho \to 0$.

6. **Recurrence under an absolutely continuous measure.** Here we consider measures that are absolutely continuous with respect to Lebesgue measure on the unstable leaves.

Let d be the metric on M and $T : M \to M$ a C^2-diffeomorphism with attractor $\mathscr{A}$. We assume that the system can be modeled by a Young tower possessing a reference measure $\hat{m}$, and that the greatest common divisor of the return times R_i is equal to one. Let μ be the SRB measure on the attractor, that is μ_{γ^u} is absolutely continuous with respect to Lebesgue ℓ on the unstable leaves γ^u where its density function f is regular and bounded on unstable leaves γ^u: $\mu_{\gamma^u}(F) = \int_F f \, d\ell$ for F on an unstable leaf γ^u. By [27] $\frac{1}{\mathfrak{c}} \leq f(\mathsf{x}) \leq \mathfrak{c}$ for a.e. $\mathsf{x} \in \mathscr{A}$ and some $\mathfrak{c} > 1$. We require Assumptions (A1), (A2) and (A3) to be satisfied with λ larger than 9. Note that the attractor $\mathscr{A} \subset M$ is given by

$$\mathscr{A} = \bigcup_{i=1}^{\infty} \bigcup_{j=0}^{R_i-1} T^j \Lambda_i.$$

6.1. **Regularity of the SRB measure μ.** Let us note that by [6] Lemma B.2 there exists a $\varsigma > 0$ and a set $\mathcal{U}'_\rho \subset M$ such that $\mu(\mathcal{U}'_\rho) \leq c_1 \Omega(|\log \rho|)$ for some c_1 and so that $\mu(B_\rho(\mathsf{x})) \leq \rho^\varsigma$ for all $\rho \in (0, 1]$ and all $\mathsf{x} \notin \mathcal{U}'_\rho$. Hence μ is geometrically regular for ς.

The following proposition shows that the SRB measure μ is ξ-regular for any $\xi > \frac{2}{u}(D+1) - 1$ where $u \geq 1$ is the dimension of the unstable manifolds and D is the dimension of M.

Proposition 6.1. *Let u be the dimension of the unstable leaves and let $w_0 > D + 2 - u$. Then for every $a < \frac{\lambda - 2}{2}$ and $w > w_0$ there exists a constant C_{11} and a set $\mathcal{U}''_\rho(w) \subset M$ satisfying $\mu(\mathcal{U}''_\rho) = \mathcal{O}(((w - w_0)|\log \rho|)^{-a})$ such that*

$$\mu(B_{\rho+\rho^w}(\mathsf{x}) \setminus B_\rho(\mathsf{x})) \leq C_{11}\mu(B_\rho(\mathsf{x}))((w - w_0)|\log \rho|)^{-a}$$

for every $\rho > 0$ and for every $\mathsf{x} \notin \mathcal{U}''_\rho$.

Proof. Put $\mathcal{A} = B_{\rho+\rho^w}(\mathsf{x}) \setminus B_\rho(\mathsf{x})$. For $s, l > 0$ we let as in Section 4.3 $\tilde{\Lambda}_i = \{x \in \Lambda_i : R(\hat{T}^k x) \leq s \ \forall k \leq l\}$. Then as in (18) we use Lemma 4.7 to obtain for $\mathsf{x} \notin \mathcal{D}_{sl,s}$:

$$
\begin{aligned}
\mu(\mathcal{A}) &\leq \sum_i \sum_{j=0}^{R_i-1} m(T^{-j}\mathcal{A} \cap \tilde{\Lambda}_i) + 2\sqrt{sl}\,\Omega(s)\mu(B_\rho) \\
&\leq \sum_i \sum_{j=0}^{R_i-1} \sum_{\tau \in I_{i,j,n}} m(T^{-j}\mathcal{A} \cap \tilde{\zeta}_\tau) + 2\sqrt{sl}\,\Omega(s)\mu(B_\rho) \\
&\leq \sum_i \sum_{j=0}^{R_i-1} \sum_{\substack{\tau' \,|\, \exists \tau \subset \tau' \\ \tau \in I \\ T^{-j}\mathcal{A} \cap \zeta_{\tau'} \neq \varnothing}} m(T^{-j}\mathcal{A} \cap \tilde{\zeta}_{\tau'}) + 2\sqrt{sl}\,\Omega(s)\mu(B_\rho) \\
&\leq C'_3 \sum_i \sum_{j=0}^{R_i-1} \sum_{\substack{\tau' \,|\, \exists \tau \subset \tau' \\ \tau \in I \\ T^{-j}\mathcal{A} \cap \zeta_{\tau'} \neq \varnothing}} \int m_{\hat{\gamma}^u}(\hat{T}^l(T^{-j}\mathcal{A} \cap \tilde{\zeta}_{\tau'}))m_{\gamma^u}(\tilde{\zeta}_{\tau'})\,d\nu(\gamma^u) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + 2\sqrt{sl}\,\Omega(s)\mu(B_\rho),
\end{aligned}
$$

where $|\tilde{\zeta}_{\tau'}| \leq \alpha^l$. Here we proceeded as in (19) and put $\hat{\gamma}^u = \gamma^u(\hat{T}^l x)$ for some $x \in \zeta_{\tau'} \cap \gamma^u$. As before, the second term on the RHS estimates the contributions from terms $m(T^{-j}\mathcal{A} \cap (\Lambda_i \setminus \tilde{\Lambda}_i))$ and from the tall beams where $R_i > s$.

Since the map T is C^2 on M we get that $\tilde{\zeta}_{\tau'} \cap \gamma^u$ are on nearly flat segment of an unstable leaf γ^u, which allows us to estimate the measure of the intersection $\mathcal{A} \cap \tilde{\zeta}_{\tau'} \cap \gamma^u$. Since the density of m_{γ^u} is regular and bounded on γ^u we obtain

$$m_{\gamma^u}(\mathcal{A} \cap \tilde{\zeta}_{\tau'}) \leq c_1 \rho^{u-1}\delta\rho = c_1 \rho^{u-1+w}$$

as $\delta\rho = \rho^w$ is the thickness of the annulus $\mathcal{A}$. Here u denotes the dimension of the unstable manifolds and distances are measured inside the unstable leaf. Moreover, as $\hat{T}^l$ is one-to-one on $\tilde{\zeta}_{\tau'}$, and as the map T expands distances on M by at most a factor A at each iteration (i.e. the Jacobian is bounded by $(A^{sl})^u$) we conclude that

$$m_{\hat{\gamma}^u}(\hat{T}^l(T^{-j}\mathcal{A} \cap \tilde{\zeta}_{\tau'})) \leq m_{\gamma^u}(\hat{T}^l T^{-j}\mathcal{A} \cap \Lambda) \leq c_2(A^{sl})^u m_{\gamma^u}(\mathcal{A} \cap \tilde{\zeta}_{\tau'}) \leq c_3 A^{nu}\rho^{u-1+w},$$

where we put $n = sl$.

Since $\hat{T}^l \tilde{\zeta}_{\tau'} = \Lambda$ if $\tilde{\zeta}_{\tau'} \neq \varnothing$ (see Section 4.3 (II)) we obtain

$$
\begin{aligned}
\mu(\mathcal{A}) &\leq A^{nu} \rho^{u-1+w} \sum_i \sum_{j=0}^{R_i-1} \sum_{\substack{\tau' \mid \exists \tau \subset \tau' \\ \tau \in I \\ T^{-j}\mathcal{A} \cap \zeta_{\tau'} \neq \varnothing}} m(\tilde{\zeta}_{\tau'}) + 2\sqrt{n}\,\Omega(s)\mu(B_\rho) \\
&\leq c_4 A^{nu} \rho^{u-1+w}(\rho + \rho^w + \alpha^{\frac{n}{s}})^u + 2\sqrt{n}\,\Omega(s)\mu(B_\rho)
\end{aligned}
$$

Let $w_0 > D + 2 - u$ and define $k(w) = \frac{w-w_0}{2u \log A}$ for $w > w_0$. Then put $s = \lceil k(w)|\log\rho|^{\tilde{\eta}} \rceil$ and $l = \lfloor |\log\rho|^{1-\tilde{\eta}} \rfloor$, where $\tilde{\eta} \in (0,1)$. Thus $n = sl \leq 2k(w)|\log\rho|$ and we obtain $A^{nu}\rho^{u-1+w} = \rho^{-2\tilde{u}k(w)\log A + u - 1 + w} < \rho^{D+1}$. We have $\rho^{D+1} \leq \mu(B_\rho(\mathsf{x}))$ for all $\mathsf{x} \notin \hat{\mathcal{U}}''_\rho$ where $\hat{\mathcal{U}}''_\rho \subset M$ is a small set whose measure is by [6] Lemma A.1 bounded by $\mathcal{O}(\rho)$ for some constant C_{11}. This implies

$$
\begin{aligned}
\mu(\mathcal{A}) &\leq c_5 \mu(B_\rho)\left(\rho^{\varsigma'} + \alpha^{\varsigma' l} + \sqrt{n}\,\Omega(s)\right) \\
&\leq c_6 \mu(B_\rho)\left(\rho^{\varsigma'} + \alpha^{\varsigma'|\log\rho|^{1-\tilde{\eta}}} + (k(w)|\log\rho|)^{-\tilde{\eta}\frac{\lambda-1}{2}}(k(w)|\log\rho|)^{\frac{1}{2}}\right) \\
&\leq \frac{C_{11}}{(w-w_0)^a |\log\rho|^a}\mu(B_\rho)
\end{aligned}
$$

for a constant C_{11} and for all ρ small enough. Since we can choose $\tilde{\eta}$ arbitrarily close to 1, we obtain any exponent $a < \frac{\lambda-2}{2}$. This applies for points $\mathsf{x} \notin \mathcal{U}''_\rho$, where the measure of the forbidden set $\mathcal{U}''_\rho = \hat{\mathcal{U}}''_\rho \cup \mathcal{D}_{sl,s}$ is bounded by $\mathcal{O}(1)((w-w_0)|\log\rho|)^{-a}$. □

6.2. **Proof of Theorem 3.** To prove this result we combine Theorems 1 and 2. Let $J = \lfloor \mathfrak{a}|\log\rho| \rfloor$ as before, and take $\mathfrak{a} = [4\log(\|DT\|_{\mathscr{L}^\infty} + \|DT^{-1}\|_{\mathscr{L}^\infty})]^{-1}$. Clearly the dimension ς of the measure m_{γ^u} is equal to the dimension u and the dimension $\hat{\varsigma}$ is less or equal to u. Thus $\xi = \varsigma(\lambda-1) - \hat{\varsigma} \geq u(\lambda-2)$.

Since by assumption $\frac{D+2}{u} < \lambda$ we obtain $D + 2 - u < \xi$ and therefore by the previous sub-section μ is geometric regular and ξ-regular provided one restricts to the set $M \setminus \mathcal{U}_\rho$, where $\mathcal{U}_\rho(w) = \mathcal{U}'_\rho \cup \mathcal{U}''_\rho(w)$ and $\mu(\mathcal{U}_\rho) \leq c_1((w-w_0)|\log\rho|)^{-a}$ for some c_1 and any $a < \frac{\lambda-2}{2}$. Let $g(w) = C_{11}(w-w_0)^{-a}$, then $g(w)$ satisfies the summability condition in Assumption (A4) for any $\beta > \frac{1}{a} > \frac{2}{\lambda-2}$, and so (A4) is satisfied outside the set $\mathcal{E}_\rho = \bigcup_{n=J}^\infty \mathcal{U}_\rho(n^\beta)$ whose measure is bounded by $\mathcal{O}(|\log\rho|^{-a})$. The measure of the very short return set

$$
\mathcal{V}_\rho = \{\mathsf{x} \in \mathscr{A} : B_\rho(\mathsf{x}) \cap T^n B_\rho(\mathsf{x}) \neq \varnothing \text{ for some } 1 \leq n < J\}
$$

has been estimated in Proposition 5.1.

With the forbidden set $\mathcal{Z}_\rho = \mathcal{X}_\rho \cup \mathcal{V}_\rho \cup \mathcal{E}_\rho$ whose size is bounded by

$$
\mu(\mathcal{Z}_\rho) \leq C'' |\log\rho|^{-\frac{\lambda-9}{4}}
$$

we get that for any point $\mathsf{x} \in \mathscr{A} \setminus \mathcal{Z}_\rho$ the function S counting the number of visits to the ball $B_\rho(\mathsf{x})$ can be approximated by a Poissonian, that is

$$
\left| \mathbb{P}(S = k) - e^{-t}\frac{t^k}{k!} \right| \leq C |\log\rho|^{-\kappa} \qquad \text{for all } k \in \mathbb{N}_0,
$$

for any $\kappa < \frac{\lambda-7}{4}$ by the remark following the proof of Theorem 1. Now we set $\hat{C} = \max(C, C'')$. □

7. **Poisson approximation theorem.** This section contains the abstract Poisson approximation theorem which establishes the distance between sums of $\{0,1\}$-valued dependent random variables X_n and a random variable that is Poisson distributed. It is used in Section 4.5 in the proof of Theorem 1 and compares the number of occurrences in a finite time interval with the number of occurrences in the same interval for a Bernoulli process $\{\tilde{X}_n : n\}$.

Theorem 7.1. [6] *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary $\{0,1\}$-valued process and t a positive parameter. Let $S_a^b = \sum_{n=a}^b X_n$ and define $S := S_1^N$ for convenience's sake where $N = \lfloor t/\epsilon \rfloor$ and $\epsilon = \mathbb{P}(X_1 = 1)$. Additionally, let ν be the Poisson distribution measure with mean $t > 0$. Finally, assume that $\epsilon < \frac{t}{2}$. Then there exists a constant C_{12} such that for any $E \subset \mathbb{N}_0$, and $2 \le p < N$ we have*

$$|\mathbb{P}(S \in E) - \nu(E)| \le C_{12}\#\{E \cap [0,N]\}\,(N(\mathcal{R}_1 + \mathcal{R}_2) + p\epsilon)$$

where,

$$\mathcal{R}_1 = \sup_{\substack{0<j<N-p \\ 0<q<N-p-j}} \{|\mathbb{P}(X_1 = 1 \wedge S_{p+1}^{N-j} = q) - \epsilon\,\mathbb{P}(S_{p+1}^{N-j} = q)|\}$$

$$\mathcal{R}_2 = \sum_{n=2}^p \mathbb{P}(X_1 = 1 \wedge X_n = 1).$$

Proof. Let $(\tilde{X}_n)_{n\in\mathbb{N}}$ be a sequence of independent, identically distributed random variables taking values in $\{0,1\}$, constructed so that $\mathbb{P}(\tilde{X}_1 = 1) = \epsilon$. Further assume that the $\tilde{X}_n$'s are independent of the X_n's. Let $\tilde{S} = \sum_{n=1}^N \tilde{X}_n$. Then

$$|\mathbb{P}(S \in E) - \nu(E)| \le |\mathbb{P}(S \in E) - \mathbb{P}(\tilde{S} \in E)| + |\mathbb{P}(\tilde{S} \in E) - \nu(E)|$$

$$\le \sum_{k\in E\cap[0,N]} |\mathbb{P}(S = k) - \mathbb{P}(\tilde{S} = k)| + \sum_{k=0}^\infty \left|\mathbb{P}(\tilde{S} = k) - \frac{t^k}{k!}e^{-t}\right|$$

Thanks to [5] we can bound the second sum using the estimate

$$\sum_{k=0}^\infty \left|\mathbb{P}(\tilde{S} = k) - \frac{t^k}{k!}e^{-t}\right| \le \frac{2t^2}{N}. \tag{25}$$

For summands of the remaining term we utilize the proof of Theorem 2.1 from [6] according to which for every $k \le N$,

$$|\mathbb{P}(S = k) - \mathbb{P}(\tilde{S} = k)| \le 2N(\mathcal{R}_1 + \mathcal{R}_2 + p\epsilon^2) + 4p\epsilon.$$

As $N \le t/\epsilon$ this becomes

$$|\mathbb{P}(S = k) - \mathbb{P}(\tilde{S} = k)| \le 6\,t\,(N(\mathcal{R}_1 + \mathcal{R}_2) + p\epsilon). \tag{26}$$

Combining (25) and (26) yields

$$|\mathbb{P}(S \in E) - \nu(E)| \le \sum_{k\in E\cap[0,N]} |\mathbb{P}(S = k) - \mathbb{P}(\tilde{S} = k)| + \frac{2t^2}{N}$$

$$\le \sum_{k\in E\cap[0,N]} 6\,t\,(N(\mathcal{R}_1 + \mathcal{R}_2) + p\epsilon) + \frac{2t^2}{t/\epsilon - 1}$$

$$\le 6\,t\,\#\{E \cap [0,N]\}\,(N(\mathcal{R}_1 + \mathcal{R}_2) + p\epsilon) + 4\,t\epsilon$$

$$\le C_{12}\#\{E \cap [0,N]\}\,(N(\mathcal{R}_1 + \mathcal{R}_2) + p\epsilon)$$

for some $C_{12} < \infty$. □

REFERENCES

[1] M. Abadi, Hitting, returning and the short correlation function, *Bull. Braz. Math. Soc.*, **37** (2006), 593–609.

[2] M. Abadi, Poisson approximations via Chen-Stein for non-Markov processes, in *In and Out of Equilibrium 2* (eds. V. Sidoravicius and M. E. Vares), Progr. Probab., 60, Birkhäuser, Basel, 2008, 1–19.

[3] M. Abadi and N. Vergne, Sharp errors for point-wise Poisson approximations in mixing processes, *Nonlinearity*, **21** (2008), 2871–2885.

[4] J. F. Alves and V. Pinheiro, Slow rates of mixing for dynamical systems with hyperbolic structure, *J. Stat. Phys.*, **131** (2008), 505–534.

[5] R. Arratia, L. Goldstein and L. Gordon, Two moments suffice for Poisson approximations: The Chen-Stein method, *Ann. Probab.*, **17** (1989), 9–25.

[6] J.-R. Chazottes and P. Collet, Poisson approximation for the number of visits to balls in nonuniformly hyperbolic dynamical systems, *Ergod. Th. & Dynam. Sys.*, **33** (2013), 49–80.

[7] P. Collet, Statistics of closest return for some non-uniformly hyperbolic systems, *Ergod. Th. & Dynam. Sys.*, **21** (2001), 401–420.

[8] M. Denker, Remarks on weak limit laws for fractal sets, in *Fractal Geometry and Stochastics* (Finsterbergen, 1994), Progress in Probability, 37, Birkhäuser, Basel, 1995, 167–178.

[9] M. Denker, M. Gordin and A. Sharova, A Poisson limit theorem for toral automorphisms, *Illinois J. Math.*, **48** (2004), 1–20.

[10] W. Doeblin, Remarques sur la théorie métrique des fraction continues, *Compositio Mathematica*, **7** (1940), 353–371.

[11] D. Dolgopyat, Limit theorems for partially hyperbolic systems, *Trans. Am. Math. Soc.*, **356** (2004), 1637–1689.

[12] J. M. Freitas, N. Haydn and M. Nicol, Convergence of rare event point processes to the Poisson process for planar billiards, *Nonlinearity*, **27** (2014), 1669–1687.

[13] A. Galves and B. Schmitt, Inequalities for hitting time in mixing dynamical systems, *Random Comput. Dynam.*, **5** (1997), 337–347.

[14] N. T. A. Haydn, Statistical properties of equilibrium states for rational maps, *Ergod. Th. & Dynam. Sys.*, **20** (2000), 1371–1390.

[15] N. T. A. Haydn, Entry and return times distribution, *Dynamical Systems: An International Journal*, **28** (2013), 333–353.

[16] N. T. A. Haydn and Y. Psiloyenis, Return times distribution for Markov towers with decay of correlations, *Nonlinearity*, **27** (2014), 1323–1349.

[17] M. Hirata, Poisson law for Axiom A diffeomorphisms, *Ergod. Th. & Dynam. Syst.*, **13** (1993), 533–556.

[18] M. Hirata, B. Saussol and S. Vaienti, Statistics of return times: A general framework and new applications, *Comm. Math. Phys.*, **206** (1999), 33–55.

[19] M. Kač, On the notion of recurrence in discrete stochastic processes, *Bull. Amer. Math. Soc.*, **53** (1947), 1002–1010.

[20] M. Kupsa and Y. Lacroix, Asymptotics for hitting times, *Ann. of Probab.*, **33** (2005), 610–619.

[21] Y. Lacroix, Possible limit laws for entrance times of an ergodic aperiodic dynamical system, *Israel J. Math.*, **132** (2002), 253–263.

[22] F. Pène and B. Saussol, Poisson law for some nonuniformly hyperbolic dynamical systems with polynomial rate of mixing, preprint, Université de Bretagne Occidentale.

[23] B. Pitskel, Poisson law for Markov chains, *Ergod. Th. & Dynam. Syst.*, **11** (1991), 501–513.

[24] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, Vol. 3, Gauthiers-Villars, Paris 1899.

[25] K. Wasilewska, *Limiting Distribution and Error Terms for the Number of Visits to balls in Mixing Dynamical Systems*, Ph.D. thesis, USC 2013.

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

[27] L.-S. Young, Recurrence times and rates of mixing, *Israel J. Math.*, **110** (1999), 153–188.

E-mail address: nhaydn@usc.edu

E-mail address: kwilliams@hw.com