ENTRY TIMES DISTRIBUTION FOR DYNAMICAL BALLS ON METRIC SPACES

N HAYDN AND F YANG

Abstract. We show that the entry and return times for dynamical balls (Bowen balls) is exponential for systems that have an α-mixing invariant measure with certain regularities. We also show that systems modeled by Young's tower has exponential entry time distribution for dynamical balls. We also apply the results to conformal maps and expanding maps on the interval.

1. Introduction

In this paper the distribution of entry and return times are studied for continuous maps on metric spaces. There are many results on the distribution of return times to cylinder sets for instance from 1990 by Pitskel [25] (see also [10]) for Axiom A maps using Markov partitions and also by Hirata. For ψ-mixing maps, Galves and Schmitt [12] have introduced a method to obtain the limiting distribution for the first entry time to cylinder sets. Abadi then generalised that approach to ϕ and also α-mixing measures. The nature of the return set however is critical for the longtime statistics of return and Lacroix and Kupsa [19, 18] have given examples where for an ergodic system any return time distribution can be realised by taking a limit along a suitably chosen sequence of return sets. For entry and returns to balls, Pitskel's result shows using an approximation argument that for Axiom A maps on the two-dimensional torus return times are Poisson distributed. Recently, Chazottes and Collet [9] showed a similar result for attractors in the case of exponentially decaying correlations. This was in [17] extended to polynomial decay of correlations where the error terms are logarithmic. A similar result without error terms and requiring sufficient regularity of the invariant measure was proven in [23] (see also [14]). In this paper we consider another class of entry sets, namely dynamical balls on metric spaces. It has been shown elsewhere that dynamical balls exhibit good limiting

Date: January 3, 2017.

N Haydn, Department of Mathematics, University of Southern California, Los Angeles, 90089-2532. E-mail: nhaydn@usc.edu.
F Yang, Department of Mathematics, University of Southern California, Los Angeles, 90089-2532. E-mail: yang617@usc.edu.

statistic and in particular have the equipartition property which for partitions is associated with the theorem of Shannon-McMillan-Breiman [20]. Similarly, a theorem of Ornstein and Weiss [21, 22] has a counterpart for metric balls [27]: It was shown that the exponential growth rate of recurrence times is equal to the metric entropy.

In this paper we first study the distribution of the first entry time assuming that there is a generating partition which is α-mixing. This requirement is satisfied by a large number of systems, in particular by those which allow Young's tower construction with an exponential or polynomial tail. This is shown in section 6. In section 3 we deduce the first return time distribution.

2. Main results

Let (X, d) be a compact metric space, $T : X \circlearrowleft$ and (X, T, μ) be a continuous transformation and μ a T-invariant Borel probability measure on X. For a set $B \subset X$ denote by $\tau_B(y)$ the first time when the orbit of y enters B, i.e.

$$\tau_B(y) = \min\{j > 0 : T^j y \in B\} \in \mathbb{N} \cup \{\infty\}.$$

If the domain of τ_B is the entire space X, then τ_B is the *first entry time* and if one considers the restriction of τ_B to the set B itself, then τ_B is the *first return time*. By Poincaré's recurrence theorem the first return time $\tau_B|_B$ is finite μ-almost everywhere for T-invariant probability measures μ. If moreover μ is ergodic the one has $\int_B \tau_B \, d\mu = 1$ by Kac's theorem if $\mu(B) > 0$. A large number of results on the limiting distribution have been proven in the case when B are cylinder sets, most notably by Galves and Schmitt [12] for ψ-mixing measures where they introduced a method which later was in particular by Abadi [1, 2] extended to the first entries and returns for ϕ-mixing and α-mixing measures.

If T is a map on a metric space Ω with metric d, then the nth *Bowen ball* is given by

$$B_{\varepsilon,n}(x) = \{y \in \Omega : d(T^j x, T^j y) < \varepsilon, 0 \le j < n\}.$$

Bowen balls are used to define the metric entropy and also the pressure for potentials (see e.g. [28]). Then in [8] the equivalent for the theorem of Shannon-McMillan-Breiman was proven for Bowen balls. Namely, for every ergodic T-invariant probability measure μ the limit

$$h(\mu) = \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} |\log \mu(B_{\epsilon,n}(x))|$$

exists for μ-almost every x, where the limit $h(\mu)$ is the measure theoretic entropy of μ. Moreover, in [27] (see also [11]) it was shown that for an ergodic T-invariant probability measures μ the limit

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \log R_{\varepsilon,n}(x) = h(\mu)$$

exists almost everywhere, where $R_{\varepsilon,n}(x) = \tau_{B_{\varepsilon,n}(x)}(x)$ is the *recurrence time* to the Bowen ball (the limit in n is $\limsup$ or $\liminf$). The limit is the dynamical ball equivalent of Ornstein-Weiss's formula [21, 22].

In this paper we want to address the distribution of the entry time function $\tau_{B_{\varepsilon,n}(x)}$ and in order to get meaningful results we have to assume some mixing property. We will require the measure to be α-mixing with respect to a finite or countably infinite generating partition. For that purpose let $\mathcal{A}$ be a measurable partition of X which is generating and either finite or countably infinite ($\mathcal{A}$ is generating if the elements $\mathcal{A}^\infty$ are single points). Denote by $\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j}\mathcal{A}$ its n-th join and write

$$\gamma_n = \operatorname{diam}(\mathcal{A}^n)$$

for its diameter. Since $\mathcal{A}$ is generating γ_n decreases to 0 as $n \to \infty$.

The following two properties of the invariant measure μ play a central role:

(I) Mixing property: We say the measure μ is α-*mixing* with respect to the partition $\mathcal{A}$ if

$$|\mu(A \cap T^{-n-k}B) - \mu(A)\mu(B)| \leq \alpha(k)$$

for all $A \in \sigma(\mathcal{A}^n)$, $B \in \sigma(\bigcup_j \mathcal{A}^j)$, where $\alpha(k)$ is a decreasing function which converges to zero as $k \to \infty$.

(II) Regularity: In order to control the measure of an annulus compared to the metric ball of the same size we put

$$\varphi(\epsilon, \delta, x) = \frac{\mu(B(x, \epsilon + \delta)) - \mu(B(x, \epsilon - \delta))}{\mu(B(x, \epsilon))}$$

for $0 < \delta < \epsilon$ and $x \in X$. The function φ measures the proportion of the measure of the annulus $B(x, \epsilon + \delta) \setminus B(x, \epsilon - \delta)$ to the ball $B(x, \epsilon)$. This is needed in order to control the approximation of balls by cylinder sets below. For instance, if μ is a Riemannian measure on a manifold X of dimension d, then $\varphi(\epsilon, \delta, x) = \mathcal{O}(\delta/\epsilon)$.

We now can formulate our main results which we take care to formulate in two versions which differ by the assumption made on the "annulus function" φ. The first pair of theorem is for the limiting entry times distribution and the second pair of theorems for the limiting return times distribution.

2.1. Entry times distribution.

Theorem 1. *Let μ an α-mixing T-invariant probability measure on Ω. Assume that there exist constants $0 < \gamma < 1$ and $\zeta, \xi, \kappa > 0$ such that $\gamma_n = \mathcal{O}(\gamma^{n^\xi})$, $\alpha(n) = \mathcal{O}(n^{-(2+\kappa)})$. Also assume that for every $\epsilon > 0$ there exist $C_\epsilon > 0$ independent of x, such that*

$$\varphi(\epsilon, \delta, x) \leq \frac{C_\epsilon}{|\log \delta|^\zeta}$$

for every x and all $\delta < \epsilon$ small enough. Furthermore we assume that $\xi \cdot \zeta > 3$

Then there exists an $\omega > 0$*, a constant* C_1 *and* $0 < \lambda_{B_{\epsilon,n}(x)} < 2$ *so that*

$$\left|\mathbb{P}\left(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}\right) - e^{-t}\right| \leq C_1\mu(B_{\epsilon,n}(x))^{\omega}.$$

Theorem 3 below can be seen as an application of this theorem.

If, as in the next result, the measure has good regularity then we relax the condition on the decrease of the diameter of cylinders considerably.

In the next several theorems we will always assume that C_ϵ is a constant that depends only on ϵ.

Theorem 2. *Assume that there exist constants,* $a, \kappa, \zeta > 0$ *satisfying* $a\zeta > 3$*, such that* $\gamma_n = \mathcal{O}(n^{-a})$*,* $\alpha(n) = \mathcal{O}(n^{-(2+\kappa)})$ *and*

$$\varphi(\epsilon, \delta, x) \leq C_\epsilon \delta^\zeta$$

for some constant C_ϵ*.*

Then there exists an $\omega > 0$*, a constant* C_1 *and* $0 < \lambda_{B_{\epsilon,n}(x)} < 2$ *so that*

$$\left|\mathbb{P}\left(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}\right) - e^{-t}\right| \leq C_2\mu(B_{\epsilon,n}(x))^{\omega}.$$

The annulus condition of the last theorem is for instance satisfied in the case when μ is absolutely continuous with respect to the Lebesgue measure, or Riemannian volume if X is a manifold. Then $\zeta = 1$.

The clustering factor λ_B has been more closely examined in [4]. The amount by which the clustering factor deviates from the value 1 captures essentially the periodic behaviour of the approximating set. As shown there, for cylinder sets for instance the factor goes in the limit to 1 almost surely.

For diffeormorphism on compact manifolds that allow a Young tower construction we obtain the following result, where for the details of the terms we refer to Section 6.

Theorem 3. *Let* μ *be the SRB measure for a differentiable map* T *on a manifold* X *which is modelled by a Young's tower. If the tail of the tower decays at least polynomially with power* $\lambda > 5 + \sqrt{15}$*, then*

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\mathbb{P}\left(\tau_{B_{\epsilon,n}(x)} \geq \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}\right) = e^{-t}$$

for $t > 0$ *and almost every* $x \in X$*.*

2.2. **Return times distribution.** While the previous three theorems give us limiting results for the entry times distribution, the next two theorems establish equivalent results for the return times.

For all set $A \subset \Omega$ define the *period* of A by $\tau(A) = \min\{k > 0 : T^{-k}A \cap A \neq \emptyset\}$ and put for any $\Delta < 1/\mu(A)$

$$a_A = \mathbb{P}_A(\tau_A > \tau(A) + \Delta).$$

In our setting we will choose $N(n) \leq \Delta \leq 1/\mu(B_{\epsilon,n}(x))$, where $N(n)$ will be determined later. We write for simplicity $B = B_{\epsilon,n}(x)$.

Theorem 4. *Let μ be α-mixing with $\alpha(n) = \mathcal{O}(n^{-(2+\kappa)})$ for some $\kappa > 0$. Assume that $\xi\zeta > 5$ with the remaining conditions same as in Theorem 1.*

Then there exists an $\omega > 0$ so that for all $t > 0$ and $x \in X$:

$$\left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - a_B e^{-t}\right| \leq C_3\mu(B)^\omega$$

for some constant C_3 and a parameter λ_B which is bounded as in Lemma 4.

Theorem 5. *Assume there are $a, \kappa, \zeta > 0$ satisfying $a\zeta > 5$, such that $\gamma_n = \mathcal{O}(n^{-a})$, $\alpha(n) = \mathcal{O}(n^{-(2+\kappa)})$ and*

$$\varphi(\epsilon, \delta, x) \leq C_\epsilon \delta^\zeta.$$

The remaining conditions are as in Theorem 2.

Then there exists an $\omega > 0$ so that

$$\left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - a_B e^{-t}\right| \leq C_4\mu(B)^\omega$$

for some constant C_4.

Observe that in Theorems 1 and 5 the conditions $a\zeta > 3$ and $a\zeta > 5$ respectively are vacuous in the case when the diameter γ_n of the joins $\mathcal{A}^n$ decay at an exponential rate (or even super polynomially).

The parameter $a_B \leq 1$ captures the periodic behaviour of the set B. For this see [3] where is also shown that if $a_B \to 1$ as $n \to \infty$ then the clustering factor λ_B converges to 1.

3. Examples

3.1. **Gibbs states for conformal repeller.** Let T be a C^1-map on a Riemannian manifold M. A conformal repeller is then a maximal compact set $\Omega \subset M$ so that T acts conformally on Ω and is expanding, that is there exists a $\beta > 1$ so that $|DT^k v| \geq \beta^k$ for all large enough k and all $v \in T_xM, \forall x \in \Omega$. For simplicity we shall assume that $T : \Omega \circlearrowleft$ is topologically transitive.

For a Hölder continuous potential function $f : \Omega \to \mathbb{R}$ one then has a unique Gibbs state μ which has the property that for every generating measurable partition $\mathcal{A}$ the following estimate applies for all $n \in \mathbb{N}$ and all $x \in \Omega$

$$\mu(A_n(x)) \in \left(\frac{1}{c_0}, c_0\right) e^{f^n(x)-nP(f)},$$

for some constant $c_0 > 1$, where $P(f)$ is the pressure of f and $f^n = f + f \circ T + \cdots + f \circ T^{n-1}$ is the nth ergodic sum of f. The Gibbs state μ is usually constructed using the transfer operator (see e.g. [26]). We then obtain the following result:

Theorem 6. *Let $\Omega \subset M$ be a conformal repeller for the C^1-map $T : M \circlearrowleft$ and let μ be a Gibbs state for a Hölder continuous potential function $f : \Omega \to \mathbb{R}$.*

Then there exists an $\omega > 0$ so that (with $B = B_{\epsilon,n}(x)$)
(i)

$$\left|\mathbb{P}\left(\tau_B > \frac{t}{\lambda_B \mu(B)}\right) - e^{-t}\right| \le C_1 \mu(B)^\omega.$$

for $t > 0$ and almost every $x \in X$.
(ii)

$$\left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B \mu(B)}\right) - a_B e^{-t}\right| \le C_4 \mu(B)^\omega$$

for $t > 0$ and almost every $x \in X$ (recall $a_B = \mathbb{P}_B(\tau_B > \tau(B) + \Delta)$).

Proof. A key property of a conformal repeller is that it allows Markov partitions of arbitrarily small diameter. Let $\mathcal{A}$ be a generating partition, that is the elements of $\mathcal{A}^\infty$ consist of singletons. We will verify the assumptions of Theorems 1 and 5:
(i) Because of expansiveness of the map T it follows immediately that diam $\mathcal{A}^n \le \gamma^n$ where $\gamma \le \frac{1}{\beta} < 1$.
(ii) The measure μ is in fact ψ-mixing where ψ decays exponentially fast, therefore a fortiori also α-mixing with exponentially decaying α.
(iii) Is satisfied for any $w > 1$ as μ is diametrically regular [24] and thus also has the annular decay property [7]. This yields

$$\frac{\mu(B_{\epsilon+\delta}(x) \setminus B_{\epsilon-\delta}(x))}{\mu(B_\epsilon(x))} \le c_1 \left(\frac{\delta}{\epsilon}\right)^\zeta$$

for some $\zeta > 0$, a constant c_1 and for every point $x \in \Omega$. Therefore $\varphi(\epsilon, \delta, x) = C_\epsilon \delta^\zeta$, where $C_\epsilon \le c_1 \epsilon^{-1}$.

The result now follows from Theorems 1 and 5 respectively. □

3.2. **Interval maps.** As an example we consider expanding maps on the interval. If $T : I \circlearrowleft$ is a uniformly expanding piecewise C^2-map of the unit interval $I = [0, 1]$ and satisfies the Markov property, then it has an absolutely continous invariant measure μ which has a strictly positive density h with respect to Lebesgue measure λ (see e.g. [13]).

Theorem 7. *Let T be a uniformly expanding piecewise C^2 map of the interval which satisfies the Markov property. Let μ be absolutely continuous invariant measure.*

Then there exists an $\omega > 0$ so that (with $B = B_{\epsilon,n}(x)$)
(i)

$$\left|\mathbb{P}\left(\tau_B > \frac{t}{\lambda_B \mu(B)}\right) - e^{-t}\right| \le C_1 \mu(B)^{\omega}.$$

for $t > 0$ and almost every $x \in X$.
(ii)

$$\left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B \mu(B)}\right) - a_B e^{-t}\right| \le C_4 \mu(B)^{\omega}$$

for $t > 0$ and almost every $x \in X$ (recall $a_B = \mathbb{P}_B(\tau_B > \tau(B) + \Delta)$).

Note that the absolutely continuous invariant measure is a Gibbs state for the Hölder continuous potential function $f = -\log|T'|$ which has zero pressure, i.e. $P(f) = 0$.

Proof. We verify the assumptions of Theorems 1 and 5.
(i) μ is in fact exponentially ψ-mixing which implies the α-mixing property with an exponentially decaying α.
(ii) One has $\operatorname{diam} \mathcal{A}^n \le \gamma^n$ with $\gamma = (\inf|T'|)^{-1} < 1$.
(iii) The annulus property is satisfied because μ is equivalent to the Legesgue measure. Therefore

$$\frac{\mu(B_{\epsilon+\delta}(x) \setminus B_{\epsilon-\delta}(x))}{\mu(B_\epsilon(x))} \le c_1 \delta$$

for some constant c_1. Hence $\varphi(\epsilon, \delta, x) \le C_\epsilon \delta$ which implies $\zeta = 1$ and $C_\epsilon = c_1$.
The result now follows from Theorems 1 and 5 respectively. □

4. First entry times distribution for Bowen balls

4.1. **Approximation by unions of cylinder sets.** The next result will be our principal technical result on which all the other theorems are based. For that purpose let $N(n)$ be an increasing sequence in n. In practice our choice of $N(n)$ depends on the point x and the dynamical ball $B_{\epsilon,n}(x)$; however to simplify notations we omit the dependence on x and ϵ. We want to approximate the Bowen balls $B_{\epsilon,n}(x)$ by a unions of $N(n)$-cylinders from the inside. For this purpose put

$$\widetilde{B}_{\epsilon,n}(x) = \bigcup_{A^{N(n)} \in \mathcal{A}^{N(n)}, A^{N(n)} \subset B_{\epsilon,n}(x)} A$$

which is the largest union of all $N(n)$-cylinders contained in $B_{\epsilon,n}(x)$. Before stating the theorem, let us define, for subsets $B \subset X$ and $f > 0$,

$$\lambda_{B,f} = \frac{-\log \mathbb{P}(\tau_B > f)}{f \mu(B)}.$$

To simplify notation we will write $\lambda_{B_{\epsilon,n}(x)} = \lambda_{B_{\epsilon,n}(x),f}$, dropping the dependence on f. The choice of f will be made clear in the next section. Again we put $\gamma_k = \operatorname{diam} \mathcal{A}^k$ and formulate the following technical result:

Theorem 8. *Let μ be an α-mixing T-invariant probability measure on Ω. Let $x \in X$ and assume there exist $\epsilon_0 > 0$ and an increasing sequence $\{N(n)\}_{n=1}^\infty$ satisfying $n < N(n) < \frac{1}{4}\mu(B_{\epsilon,n}(x))^{-1}$ such that*

$$\varphi(\epsilon, \gamma_{N(n)-k}, T^k x) \le \vartheta_n(\epsilon) \cdot \frac{\mu(B_{\epsilon,n}(x))}{ns} \tag{1}$$

for all $\epsilon < \epsilon_0$, $0 \le k \le n-1$, where $s = \alpha^{-1}(C'\mu(\widetilde{B})) + N(n)$ for some $0 < C' < 1$ and a sequence of real numbers $\vartheta_n(\epsilon) \to 0$ as $n \to \infty$ for every ϵ.

Then there exist $\lambda_{B_{\epsilon,n}(x)} \in \left(\frac{C}{s}, 2\right)$ and constants C_5, C_6 such that

$$\begin{aligned}
&\left|\mathbb{P}\left(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}\right) - e^{-t}\right| \\
&\qquad\qquad \le \quad \vartheta_n(\epsilon)\frac{t}{s\lambda_{B_{\epsilon,n}(x)}} + 3f\mu(B_{\epsilon,n}(x)) + C_5\frac{sN(n)}{f} + C_6 s\frac{\alpha(N(n))}{f\mu(\widetilde{B})}
\end{aligned}$$

for all $f \in \left(2N(n), \frac{1}{2}\mu(B_{\epsilon,n}(x))^{-1}\right)$.

4.2. **First entry time for the cylinder approximation $\widetilde{B}_{\epsilon,n}(x)$.** We begin with several lemmata, the first one of which is evident.

Lemma 1. *For all n and x we have $\sum_{k=1}^{N(n)} \mu(\widetilde{B}_{\epsilon,n}(x) \cap T^{-k}\widetilde{B}_{\epsilon,n}(x)) \le N(n)\mu(\widetilde{B}_{\epsilon,n}(x))$ and $\mathbb{P}(\tau_{\widetilde{B}_{\epsilon,n}(x)} \le t) \le t\mu(\widetilde{B}_{\epsilon,n}(x))$.*

To simply notation, we fix ϵ and n for a moment and write $B = B_{\epsilon,n}(x)$ and $\widetilde{B} = \widetilde{B}_{\epsilon,n}(x)$. The following lemma establishes an approximate exponential identity of the entry times distribution. For this see also e.g. [12, 1, 3].

Lemma 2. *For all Δ, f such that $f \ge \Delta > N(n)$ and $g \in \mathbb{N}$ we have*

$$\left|\mathbb{P}(\tau_{\widetilde{B}} > f+g) - \mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f)\right| \le 2\Delta\mu(\widetilde{B}_{\epsilon,n}(x)) + \alpha(\Delta - N(n))$$

Proof. We proceed in the traditional way splitting the difference into three parts:

$$\begin{aligned}
&\left|\mathbb{P}(\tau_{\widetilde{B}} > g+f) - \mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f)\right| \\
&\qquad \le \quad \left|\mathbb{P}(\tau_{\widetilde{B}} > g+f) - \mathbb{P}(\tau_{\widetilde{B}} > g \cap \tau_{\widetilde{B}} \circ T^{g+\Delta} > f-\Delta)\right| \\
&\qquad \quad + \left|\mathbb{P}(\tau_{\widetilde{B}} > g \cap \tau_{\widetilde{B}} \circ T^{g+\Delta} > f-\Delta) - \mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f-\Delta)\right| \\
&\qquad \quad + \left|\mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f-\Delta) - \mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f)\right| \\
&\qquad = \quad I + II + III.
\end{aligned}$$

In the first term we open up a gap of size Δ. It is estimated as follows

$$I = \mathbb{P}(\tau_{\widetilde{B}} > g \cap \tau_{\widetilde{B}} \circ T^{g+\Delta} > f - \Delta \cap \tau_{\widetilde{B}} \circ T^{g} \leq \Delta) \leq \mathbb{P}(\tau_{\widetilde{B}} \leq \Delta) \leq \Delta\mu(\widetilde{B}).$$

Similarly for the third term in which we close the gap:

$$III = \mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(f - \Delta < \tau_{\widetilde{B}} \leq f) \leq \mathbb{P}(\tau_{\widetilde{B}} > g)\Delta\mu(\widetilde{B}) \leq \Delta\mu(\widetilde{B}).$$

For the second term we use the α-mixing property. Notice that $\widetilde{B} = \widetilde{B}_{\epsilon,n}(x)$ is a union of $N(n)$-cylinders which implies $\{\tau_{\widetilde{B}} > g\} \in \sigma(\mathcal{A}^{N(n)+g})$ and consequently

$$II = \left|\mathbb{P}(\tau_{\widetilde{B}} > g \cap \tau_{\widetilde{B}} \circ T^{g+\Delta} > f - \Delta) - \mathbb{P}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f - \Delta)\right| \leq \alpha(\Delta - N(n))$$

where $\Delta - N(n)$ is the size of the gap.

The three parts combined now prove the lemma. □

Let us now put $\theta = \theta(f) = -\log \mathbb{P}(\tau_{\widetilde{B}} > f)$ where $f > 0$. We then have the following estimate.

Lemma 3. *Let $f > \Delta > N(n)$ then for all $k \geq 1$ we have*

$$\left|\mathbb{P}(\tau_{\widetilde{B}} > kf) - e^{-\theta k}\right| \leq \frac{2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n))}{\mathbb{P}(\tau_{\widetilde{B}} \leq f)}.$$

Proof. First we use induction to prove that

$$\left|\mathbb{P}(\tau_{\widetilde{B}} > kf) - e^{-\theta k}\right| \leq 2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n))) \cdot (1 + e^{-\theta} + \cdots + e^{-\theta(k-2)}).$$

Clearly it holds for $k = 1$ by the definition of θ. For $k > 1$ we use induction. Assuming that it holds for k, we obtain:

$$\begin{aligned}
&\left|\mathbb{P}(\tau_{\widetilde{B}} > (k+1)f) - e^{-\theta(k+1)}\right| \\
&\quad\leq \left|\mathbb{P}(\tau_{\widetilde{B}} > (k+1)f) - \mathbb{P}(\tau_{\widetilde{B}} > kf) \cdot e^{-\theta}\right| + \left|\mathbb{P}(\tau_{\widetilde{B}} > kf) \cdot e^{-\theta} - e^{-\theta(k+1)}\right| \\
&\quad\leq 2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n)) + e^{-\theta}\left|\mathbb{P}(\tau_{\widetilde{B}} > kf) - e^{-\theta k}\right| \\
&\quad\leq 2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n)) + e^{-\theta}(2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n))) \cdot (1 + e^{-\theta} + \cdots + e^{-\theta(k-2)}) \\
&\quad= (2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n))) \cdot (1 + e^{-\theta} + \cdots + e^{-\theta(k-1)}).
\end{aligned}$$

Hence

$$\left|\mathbb{P}(\tau_{\widetilde{B}} > kf) - e^{-\theta k}\right| \leq (2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n)))\frac{1}{1 - e^{-\theta}}$$

for all $k \in \mathbb{N}$ and the lemma follows since $\frac{1}{1-e^{-\theta}} = \frac{1}{\mathbb{P}(\tau_{\widetilde{B}} \leq f)}$. □

For subset $B \subset X$ let us define

$$\lambda_{B,f} = \frac{-\log \mathbb{P}(\tau_B > f)}{f\mu(B)}.$$

For the approximations $\widetilde{B}_{\epsilon,n}(x)$ we then obtain the following estimate.

Lemma 4. *Let $f \in \mathbb{N}$ be such that $f\mu(\widetilde{B}_{\epsilon,n}(x)) \le \frac{1}{2}$. Then there exist $C_7 > 0$ such that*

$$\frac{C_7}{s} \le \lambda_{\widetilde{B}_{\epsilon,n}(x),f} \le 2$$

where, as before, $s = \alpha^{-1}(C'\mu(\widetilde{B}_{\epsilon,n}(x))) + N(n)$ for some $0 < C' < 1$

Proof. We follow the proof in Galves-Schmitt [12] and Abadi [3]. To estimate $\lambda_{\widetilde{B}_{\epsilon,n}(x),f}$ we use the simple estimate

$$\frac{\theta}{2} \le 1 - e^{-\theta} \le \theta$$

for all $\theta \in [0,1]$. Let us write $\widetilde{B}$ for $\widetilde{B}_{\epsilon,n}(x)$ and note that $\mathbb{P}\{\tau_{\widetilde{B}} \le f\} = 1 - e^{-\theta}$ as $\theta = -\log \mathbb{P}\{\tau_{\widetilde{B}} > f\}$. By Lemma 1,

$$\lambda_{\widetilde{B},f} = \frac{\theta}{f\mu(\widetilde{B})} \le \frac{2\mathbb{P}\{\tau_{\widetilde{B}} \le f\}}{f\mu(\widetilde{B})} \le 2$$

For the lower bound, notice that $\{\tau_{\widetilde{B}} > f\} = \bigcap_{j=0}^{[f]} \left(T^{-js}(\widetilde{B})\right)^c \subset \bigcap_{j=0}^{[\frac{f}{s}]} \left(T^{-js}(\widetilde{B})\right)^c$. As in the proof of Lemma 3 we obtain first by the mixing property

$$\begin{aligned}\mu\left(\bigcap_{j=0}^{k+1}\left(T^{-js+1}(\widetilde{B})\right)^c\right) &= \mu\left(\bigcap_{j=0}^{k}\left(T^{-js+1}(\widetilde{B})\right)^c \cap T^{-(k+1)s+1}(\widetilde{B})\right)\\ &\le \mu\left(\bigcap_{j=0}^{k}\left(T^{-js+1}(\widetilde{B})\right)^c\right)\mu(\widetilde{B}^c) + \alpha(s - N(n))\end{aligned}$$

which yields then by repeated application

$$\mu\left(\bigcap_{j=0}^{[\frac{f}{s}]}\left(T^{-js+1}(\widetilde{B})\right)^c\right) \le \mu(\widetilde{B}^c)^{[\frac{f}{s}]} + \alpha(s - N(n))\frac{1}{1 - \mu(\widetilde{B}^c)}.$$

Consequently

$$\mathbb{P}(\tau_{\widetilde{B}} > f) \le (1 - \mu(\widetilde{B}))^{f/s} + \alpha(s - N(n))\frac{1 - (1 - \mu(\widetilde{B}))^{f/s}}{\mu(\widetilde{B})},$$

and therefore

$$\mathbb{P}(\tau_{\widetilde{B}} \le f) \quad \ge \quad (1 - (1 - \mu(\widetilde{B}))^{f/s})\big(1 - \frac{\alpha(s - N(n))}{\mu(\widetilde{B})}\big)$$

$$\geq \frac{f}{s}\mu(\widetilde{B})\left(1-\frac{\alpha(s-N(n))}{\mu(\widetilde{B})}\right).$$

Thus

$$\lambda_{\widetilde{B},f}=\frac{\theta}{f\mu(\widetilde{B})}\geq\frac{\mathbb{P}\{\tau_{\widetilde{B}}\leq f\}}{f\mu(\widetilde{B})}\geq\frac{1}{s}\left(1-\frac{\alpha(s-N(n))}{\mu(\widetilde{B})}\right).$$

In particular since $s=\alpha^{-1}(C'\mu(\widetilde{B}))+N(n)$ we get $\lambda_{\widetilde{B},f}\geq\frac{C_7}{s}$ and $\mathbb{P}(\tau_{\widetilde{B}}\leq f)\geq\frac{C_7 f\mu(B)}{s}$ for some constant C_7. □

4.3. **Proof of Theorem 8.** In the previous section the Bowen ball $B_{\epsilon,n}(x)$ was estimated from the inside. To proof Theorem 8 we thus have to show that the contribution of the 'annulus'

$$\widetilde{\partial B}_{\epsilon,n}(x)=\bigcup_{A\in\mathcal{A}^{N(n)},\,A\cap\partial B_{\epsilon,n}(x)\neq\emptyset}A$$

goes to zero as $n\to\infty$ and then $\epsilon\to 0$. We clearly have $B_{\epsilon,n}(x)\setminus\widetilde{B}_{\epsilon,n}(x)\subset\widetilde{\partial B}_{\epsilon,n}(x)$ and also $\tau_{B_{\epsilon,n}(x)}\geq\tau_{\widetilde{B}_{\epsilon,n}(x)}$ since $\widetilde{B}_{\epsilon,n}(x)\subset B_{\epsilon,n}(x)$. The following lemma estimates the size of the annulus.

Lemma 5. *Assume that (1) holds for some sequence of real numbers $\vartheta_n(\epsilon)\to 0$ as $n\to\infty$. With the notation as above (and in particular with $s=\alpha^{-1}(C'\mu(\widetilde{B}))+N(n)$) we obtain*

$$\mu(\widetilde{\partial B}_{\epsilon,n}(x))=\vartheta_n(\epsilon)\frac{\mu(B_{\epsilon,n}(x))}{s}.$$

Proof. Since T is continuous, $\partial B_{\epsilon,n}(x)\subset\bigcup_{k=0}^{n-1}T^{-k}\partial B(T^kx,\epsilon)$. Hence if $A^{N(n)}\cap\partial B_{\epsilon,n}(x)\neq\emptyset$ then we must have $A^{N(n)-k}(T^ky)\cap\partial B(T^kx,\epsilon)\neq\emptyset$ for some $0\leq k\leq n-1$, $y\in A^{N(n)}$. Notice that $\operatorname{diam}(A^{N(n)-k}(T^ky))\leq\gamma_{N(n)-k}$, we have

$$\begin{aligned}\widetilde{\partial B}_{\epsilon,n}(x)&\subset\bigcup_{k=0}^{n-1}T^{-k}(B(\partial B(T^kx,\epsilon),\gamma_{N(n)-k}))\\&\subset\bigcup_{k=0}^{n-1}T^{-k}(B(T^kx,\epsilon+\gamma_{N(n)-k})\setminus B(T^kx,\epsilon-\gamma_{N(n)-k})),\end{aligned}$$

hence

$$\mu(\widetilde{\partial B}_{\epsilon,n}(x))\leq n\cdot\sup_{0\leq k\leq n-1}\mu(B(T^kx,\epsilon+\gamma_{N(n)-k})\setminus B(T^kx,\epsilon-\gamma_{N(n)-k}))$$

$$\begin{aligned}
&= n \cdot \sup_{0\le k\le n-1} \{\varphi(\epsilon, \gamma_{N(n)-k}, T^k x)\cdot \mu(B(T^k x,\epsilon))\}\\
&= \vartheta_n(\epsilon)\frac{\mu(B_{\epsilon,n}(x))}{s} \sup_{0\le k\le n-1} \mu(B(T^k x,\epsilon))\\
&\le \vartheta_n(\epsilon)\frac{\mu(B_{\epsilon,n}(x))}{s}.
\end{aligned}$$

In particular we have $\mu(B_{\epsilon,n}(x))/\mu(\widetilde{B}_{\epsilon,n}(x)) = \mathcal{O}(1)$ as $n\to\infty$. □

From now on we set $\lambda_B = \lambda_{\widetilde{B},f}$.

Lemma 6. *For $f \le \frac{1}{2}\mu(\widetilde{B}_{\epsilon,n}(x))^{-1}$ one has*

$$\left|\mathbb{P}\left(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_B \mu(B_{\epsilon,n}(x))}\right) - \mathbb{P}\left(\tau_{\widetilde{B}_{\epsilon,n}(x)} > \frac{t}{\lambda_B \mu(\widetilde{B}_{\epsilon,n}(x))}\right)\right| \le 3\frac{\vartheta_n(\epsilon)t}{s\lambda_{\widetilde{B},f}}$$

for all $t>0$.

Proof. We have

$$\begin{aligned}
&\left|\mathbb{P}\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - \mathbb{P}\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right|\\
\le\ &\left|\mathbb{P}\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - \mathbb{P}\left(\tau_B > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right|\\
&\qquad + \left|\mathbb{P}\left(\tau_B > \frac{t}{\lambda_B\mu(\widetilde{B})}\right) - \mathbb{P}\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right|\\
=\ & I + II.
\end{aligned}$$

We estimate the two terms separately.

For the term I first notice that $\widetilde{B}_{\epsilon,n}(x)\subset B_{\epsilon,n}(x)$ which implies $\tau_{\widetilde{B}_{\epsilon,n}(x)} \ge \tau_{B_{\epsilon,n}(x)}$ and $\frac{t}{\lambda_B\mu(B)} \le \frac{t}{\lambda_B\mu(\widetilde{B})}$. Therefore

$$\begin{aligned}
I &\le \mathbb{P}\left(\frac{t}{\lambda_B\mu(B)} \le \tau_B \le \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\\
&\le \frac{t}{\lambda_B}\left(\frac{1}{\mu(\widetilde{B})} - \frac{1}{\mu(B)}\right)\mu(B)\\
&\le 2\frac{\vartheta_n(\epsilon)t}{s\lambda_B}
\end{aligned}$$

as $\frac{\mu(B)}{\mu(\widetilde{B})} \leq 2$ and where we used that

$$\frac{1}{\mu(\widetilde{B})} - \frac{1}{\mu(B)} = \frac{\mu(B \setminus \widetilde{B})}{\mu(B)\mu(\widetilde{B})} \leq \frac{\mu(\widetilde{\partial B}_{\epsilon,n}(x))}{\mu(B)\mu(\widetilde{B})} \leq \vartheta_n(\epsilon)\frac{1}{s\mu(\widetilde{B})} \tag{2}$$

by Lemma 5.

The term II we estimate as follows:

$$\begin{aligned}
\mathrm{II} &= \mathbb{P}\left(\left\{\tau_B \leq \frac{t}{\lambda_B \mu(\widetilde{B})}\right\} \cap \left\{\tau_{\widetilde{B}} > \frac{t}{\lambda_B \mu(\widetilde{B})}\right\}\right) \\
&\leq \mathbb{P}\left(\tau_{\widetilde{\partial B}_{\epsilon,n}(x)} \leq \frac{t}{\lambda_B \mu(\widetilde{B})}\right) \\
&\leq \frac{t}{\lambda_B}\frac{\mu(\widetilde{\partial B}_{\epsilon,n}(x))}{\mu(\widetilde{B})} \\
&= \frac{\vartheta_n(\epsilon)t}{s\lambda_B},
\end{aligned}$$

where in the last line we proceeded as for the term I above. Since by Lemma 4, $s\lambda_B > C_7/2$ the result follows. □

Proof of Theorem 8. We have to estimate $|\mathbb{P}(\tau_{\widetilde{B}_{\epsilon,n}(x)} > \frac{t}{\lambda_B \mu(\widetilde{B}_{\epsilon,n}(x))}) - e^{-t}|$, where, as before, $\lambda_B = \lambda_{\widetilde{B},f}$. We put $\Delta = 2N(n)$ and pick $f > \Delta = 2N(n)$ with $f \leq \frac{1}{2}\mu(\widetilde{B}_{\epsilon,n}(x))^{-1}$. Then $t > 0$ can be written as $t = kf + r$ with $0 \leq r < f$ and k integer. Set $t' = t - r = kf$, then

$$\begin{aligned}
&|\mathbb{P}(\tau_{\widetilde{B}} > t) - e^{-\lambda_B \mu(\widetilde{B})t}| \\
&\quad\leq |\mathbb{P}(\tau_{\widetilde{B}} > t) - \mathbb{P}(\tau_{\widetilde{B}} > t')| + |\mathbb{P}(\tau_{\widetilde{B}} > t') - e^{-\lambda_B \mu(\widetilde{B})t'}| + |e^{-\lambda_B \mu(\widetilde{B})t'} - e^{-\lambda_B \mu(\widetilde{B})t}| \\
&\quad= I + II + III.
\end{aligned}$$

The first term is easily estimated by

$$I = \mathbb{P}(t' < \tau_{\widetilde{B}} \leq t) \leq r\mu(\widetilde{B}) < f\mu(\widetilde{B}).$$

For the third term we use the mean value theorem according to which there exist $t_0 \in [\lambda_B \mu(\widetilde{B})t', \lambda_B \mu(\widetilde{B})t]$ such that

$$III = e^{-t_0}\lambda_B \mu(\widetilde{B})r \leq 2f\mu(\widetilde{B})$$

using Lemma 4 in the last estimate.

To the second term, II, we apply Lemma 3 and obtain

$$\begin{aligned} II \leq& \frac{2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n))}{\mathbb{P}(\tau_{\widetilde{B}} \leq f)} \\ \leq& \frac{s(2\Delta\mu(\widetilde{B}) + \alpha(\Delta - N(n)))}{Cf\mu(\widetilde{B})} \\ =& C_5\frac{sN(n)}{f} + C_6 s\frac{\alpha(N(n))}{f\mu(\widetilde{B})}. \end{aligned}$$

All three estimates combined yield

$$|\mathbb{P}(\tau_{\widetilde{B}_{\epsilon,n}(x)} > \frac{t}{\lambda_B\mu(\widetilde{B}_{\epsilon,n}(x))}) - e^{-t}| \leq 3f\mu(B_{\epsilon,n}(x)) + C_5\frac{sN(n)}{f} + C_6 s\frac{\alpha(N(n))}{f\mu(\widetilde{B})}.$$

□

4.4. **Proof of Theorem 1 and 2.** Now we can prove the first pair of theorems. For that purpose let us establish the following notation. For some $0 < \eta < \frac{1}{2}, \beta \in (\eta, 1)$, which will be determined later, we set $N(n) = \mu(B_{\epsilon,n}(x))^{-\eta}$ (length of cylinders), $f = \mu(B)^{-\beta}$, $\widetilde{B}_{\epsilon,n}(x) = \bigcup_{A\in\mathcal{A}^{N(n)}, A\subset B_{\epsilon,n}(x)} A$ (inner approximation) and

$$\lambda_{B_{\epsilon,n}(x)} = \frac{-\log\mathbb{P}(\tau_{\widetilde{B}_{\epsilon,n}(x)} > f)}{f\mu(\widetilde{B}_{\epsilon,n}(x))}.$$

as before.

Proof of Theorem 1. In order to apply Theorem 8 we first verify (1) with $\gamma_n = \mathcal{O}(\gamma^{n^\xi})$. Fix some $0 < \eta < \frac{1}{2}$ and set $N(n) = [\mu(B_{\epsilon,n}(x))^{-\eta}]$. Then

$$\varphi(\epsilon, \gamma_{N(n)-k}, T^k x) \leq \frac{C_\epsilon}{|\log\gamma^{(N(n)-k)^\xi}|^\zeta} \leq \frac{C_\epsilon}{N(n)^{\xi\zeta}} = C_\epsilon\mu(B)^{(\xi\zeta)\eta}.$$

Here we simply replace $N(n) - k$ by $N(n)$ since $N(n)$ is much larger than k (replace C_ϵ by a somewhat larger constant).

Since

$$s = \alpha^{-1}(C'\mu(\widetilde{B})) + N(n) \leq c_1\mu(B)^{-\frac{1}{2+\kappa}} + \mu(B)^{-\eta} \tag{3}$$

for some constant c_1, we have

$$\frac{ns\varphi(\epsilon, \gamma_{N(n)-k}, T^k x)}{\mu(B)} \leq \vartheta_n(\epsilon),$$

where

$$\vartheta_n(\epsilon) \leq C_\epsilon ns\mu(B)^{\xi\zeta\eta-1} \leq c_2 n\mu(B)^{\xi\zeta\eta-1}\left(\mu(B)^{-\frac{1}{2+\kappa}} + \mu(B)^{-\eta}\right),$$

which converges to 0 if $\eta > \eta_0 = \max\{\frac{1}{\xi\zeta-1}, \frac{1}{\xi\zeta}\frac{3+\kappa}{2+\kappa}\}$.

Applying Theorem 8 yields as $s\lambda_{B_{\epsilon,n}(x)} \ge C_7$:

$$\left|\mathbb{P}(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}) - e^{-t}\right| \le \vartheta_n(\epsilon)t + 2f\mu(B_{\epsilon,n}(x)) + C_5\frac{sN(n)}{f} + C_6 s\frac{\alpha(N(n))}{f\mu(\widetilde{B})}.$$

Since $f = \mu(B)^{-\beta}$ for some $\beta < 1$, the second term on the RHS converges to 0. The last two terms then are bounded as follows:

$$C_5\frac{sN(n)}{f} + C_6 s\frac{\alpha(N(n))}{f\mu(\widetilde{B})} \le c_3\mu(B)^\beta\left(\mu(B)^{-\eta} + \mu(B)^{\eta(2+\kappa)-1}\right)\left(\mu(B)^{-\frac{1}{2+\kappa}} + \mu(B)^{-\eta}\right).$$

In order that all terms converge to 0 we need $\eta > \eta_0$ so that

$$\omega_1 = \beta - \max\left\{\eta, \frac{1}{2+\kappa}\right\} - \max\left\{\eta, 1 - \eta(2+\kappa)\right\}$$

is positive. This can be achieved by picking $\eta \in (\eta_1, \frac{1}{2})$, where $\eta_1 = \max\{\frac{1}{\xi\zeta-1}, \frac{1}{2+\kappa}\}$ is less than $\frac{1}{2}$ (note that $\eta_1 \ge \eta_0$). This also implies that $\omega_2 = \xi\zeta\eta - \max\{\frac{1}{2+\kappa}, \eta\}$ is positive. Now put $\omega = \min\{\omega_1, \omega_2, 1-\beta\}$. □

Proof of Theorem 2. In order to apply Theorem 8 we verify that condition (1) holds:

$$\varphi(\epsilon, \gamma_{N(n)-n}, T^k x) \le C_\epsilon \gamma_{N(n)-k}^\zeta \le c_1(N(n)-k)^{-a\zeta} \le c_2 N(n)^{-a\zeta}$$

(for some c_1, c_2) and therefore $ns\varphi(\epsilon, \gamma_{N(n)-k}, T^k x)/\mu(B) \le \vartheta_n(\epsilon)$, where

$$\vartheta_n(\epsilon) \le c_3 ns\mu(B)^{a\zeta\eta-1} \le c_4 n\mu(B)^{a\zeta\eta-1}\left(\mu(B)^{-\frac{1}{2+\kappa}} + \mu(B)^{-\eta}\right).$$

The last expression converges (exponentially fast) to zero if $a\zeta\eta - 1 - \max\{\frac{1}{2+\kappa}, \eta\}$ is positive which can be achieved by picking $\eta < \frac{1}{2}$ close enough to $\frac{1}{2}$.

The remainder of the proof is identical to the proof of Theorem 1. □

5. First return time distribution

In this section we will prove Theorems 4 and 5 which establish the limiting distribution of the first return time to Bowen balls and provide rates of convergence. We use the same notation as in the previous section. Let $\tau(A)$ be the period of A and as in [3] denote by $a_A = \mathbb{P}_A(\tau_A > \tau(A) + \Delta)$ the relative size of the set of long returns, where $\Delta < 1/\mu(A)$. Again we put $\widetilde{B} = \widetilde{B}_{\epsilon,n}(x)$ and $B = B_{\epsilon,n}(x)$ and let us first prove the following two lemmata.

5.1. **Replacing B by its cylinder approximation $\tilde{B}$.** To prepare for the proof of Theorem 4 and 5 we need two lemmata. The first, Lemma 7, estimates the error that is made in the return times distribution when replacing the Bowen ball B by its N-cylinder approximation $\tilde{B}$. The second, Lemma 8 establishes an exponential-like property for the approximation which then leads to the exponential limiting return times distribution for the cylinder approximations $\tilde{B}$.

Lemma 7. *There exists a constant C_8 so that*

$$\left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - \mathbb{P}_{\widetilde{B}}\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right| \le C_8\frac{t\vartheta_n(\epsilon)}{\mu(B)}.$$

Proof. Let us first estimate the following term:

$$I = \left|\mathbb{P}_{\widetilde{B}}\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right) - \mathbb{P}_B\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right|$$

which is split into two parts $I \le I_1 + I_2$. For for the first part we obtain by Lemma 5

$$\begin{aligned} I_1 &= \frac{1}{\mu(B)}\left|\mathbb{P}\left(\left\{\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right\}\cap\widetilde{B}\right) - \mathbb{P}\left(\left\{\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right\}\cap B\right)\right| \\ &\le \frac{\mu(\widetilde{\partial B_{\epsilon,n}}(x))}{\mu(B)} \\ &\le \frac{\vartheta_n(\epsilon)}{s}. \end{aligned}$$

The second part is by (2)

$$I_2 = \mathbb{P}\left(\left\{\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right\}\cap B\right)\left|\frac{1}{\mu(\widetilde{B})} - \frac{1}{\mu(B)}\right| \le \frac{\vartheta_n(\epsilon)}{s\mu(\widetilde{B})}.$$

Hence

$$I \le 2\frac{\vartheta_n(\epsilon)}{s\mu(\widetilde{B})}.$$

Let us next estimate the term

$$II = \left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - \mathbb{P}_B\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right|,$$

which again splits into two parts $II = II_1 + II_2$ as follows. The first part is

$$II_1 = \left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(B)}\right) - \mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right|$$

$$
\begin{aligned}
&\leq \mathbb{P}_B\left(\frac{t}{\lambda_B\mu(B)} < t < \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\\
&\leq \frac{1}{\mu(B)}\left(\frac{t}{\lambda_B\mu(\widetilde{B})} - \frac{t}{\lambda_B\mu(B)}\right)\mu(B)\\
&\leq \vartheta_n(\epsilon)\frac{t}{s\lambda_B\mu(\widetilde{B})}
\end{aligned}
$$

by (2). For the second part we obtain

$$
II_2 = \left|\mathbb{P}_B\left(\tau_B > \frac{t}{\lambda_B\mu(\widetilde{B})}\right) - \mathbb{P}_B\left(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})}\right)\right| \leq \mathbb{P}_B\left(\tau_{B\setminus\widetilde{B}} < \frac{t}{\lambda_B\mu(\widetilde{B})}\right) \leq \frac{t}{\lambda_B}\frac{\mu(B\setminus\widetilde{B})}{\mu(B)\mu(\widetilde{B})}
$$

and therefore by Lemma 5

$$
II_2 \leq \frac{t}{\lambda_B}\frac{\mu(\widetilde{\partial B}_{\epsilon,n}(x))}{\mu(B)\mu(\widetilde{B})} \leq \frac{t\vartheta_n(\epsilon)}{s\lambda_B\mu(\widetilde{B})}.
$$

Finally we obtain for some constant C_8 that

$$
I + II_1 + II_2 \leq C_8\frac{t\vartheta_n(\epsilon)}{\mu(B)}
$$

where we used that $s\lambda_B \geq C_7$ by Lemma 4 and $\frac{\mu(B)}{\mu(\widetilde{B})} = \mathcal{O}(1)$. □

The following lemma which is similar to Lemma 2 establishes an approximate exponential-like identity.

Lemma 8. *For all Δ, f, g such that $f \geq \Delta > N(n)$, $g \geq \Delta + \tau(\widetilde{B})$ we have*

$$
\left|\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > f+g) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > g)\mathbb{P}(\tau_{\widetilde{B}} > f)\right| \leq 2\Delta\mu(\widetilde{B}_{\epsilon,n}(x)) + 2\frac{\alpha(\Delta - N(n))}{\mu(\widetilde{B})}
$$

Proof. We proceed as in the proof of Lemma 2 to write the left-hand-side as I + II + III. The only difference is in I:

$$
\begin{aligned}
I =&|\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > f+g) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > g \cap \tau_{\widetilde{B}}\circ T^{g+\Delta} > f-\Delta)|\\
\leq&\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}}\circ T^g \leq \Delta)\\
=&\frac{1}{\mu(\widetilde{B})}\mathbb{P}(\widetilde{B}\cap\{\tau_{\widetilde{B}}\circ T^g \leq \Delta\})\\
\leq&\mathbb{P}(\tau_{\widetilde{B}} \leq \Delta) + \frac{\alpha(\Delta - N(n))}{\mu(\widetilde{B})} \leq \Delta\mu(\widetilde{B}) + \frac{\alpha(\Delta - N(n))}{\mu(\widetilde{B})}
\end{aligned}
$$

by the α-mixing property. The estimates of the terms II and III are identical to the proof of Lemma 2. □

5.2. **Proof of Theorem 4 and 5.**

Proof of Theorem 4. We will first show that $\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})})$ satisfies the exponential law. For all $t > (\tau(\widetilde{B}) + 2\Delta)\lambda_B\mu(\widetilde{B})$, let $u = \frac{t}{\lambda_B\mu\widetilde{B}}$ and we have

$$\begin{aligned}
&\left|\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > u) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \tau(\widetilde{B}) + \Delta)e^{-t}\right| \\
&\quad\leq \left|\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > u) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \tau(\widetilde{B}) + \Delta)\mathbb{P}(\tau_{\widetilde{B}} > u - (\tau(\widetilde{B}) + \Delta))\right| \\
&\qquad + \left|\mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \tau(\widetilde{B}) + \Delta)\mathbb{P}(\tau_{\widetilde{B}} > u - (\tau(\widetilde{B}) + \Delta)) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \tau(\widetilde{B}) + \Delta)e^{-t}\right| \\
&\quad= I + II
\end{aligned}$$

where

$$I \leq 2\Delta\mu(\widetilde{B}) + 2\frac{\alpha(\Delta - N(n))}{\mu(\widetilde{B})}$$

by Lemma 8. For the second term II we have

$$\begin{aligned}
II =& a_B\left|\mathbb{P}(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})} - (\tau(\widetilde{B}) + \Delta)) - e^{-t}\right| \\
\leq& a_B\left|\mathbb{P}(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})} - (\tau(\widetilde{B}) + \Delta)) - e^{-t+(\tau(\widetilde{B})+\Delta)\lambda_B\mu(\widetilde{B})}\right| + a_B\left|e^{-t} - e^{-t+(\tau(\widetilde{B})+\Delta)\lambda_B\mu(\widetilde{B})}\right|.
\end{aligned}$$

To the first term we apply Theorem 1 with the parameter value $t' = t - (\tau(\widetilde{B}) + \Delta)\lambda_B\mu(\widetilde{B})$ and to the second term we apply the Mean Value Theorem. Hence $II \leq c_1\mu(\widetilde{B})^{\omega_1}$ for some $\omega_1 > 0$ from Theorem 1. This proves the statement of Theorem 4 for the set $\widetilde{B}$. To prove the theorem for the set B we use Lemma 7 and put $N(n) = \mu(B)^{-\eta}$ for some $\eta \in (0, 1/2)$. Then

$$\begin{aligned}
&\left|\mathbb{P}_B(\tau_B > \frac{t}{\lambda_B\mu(B)}) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})})\right| \\
&\qquad\leq C_8\frac{t\vartheta_n(\epsilon)}{\mu(B)} \\
&\qquad= \frac{\mathcal{O}(t)ns}{|\log\gamma^{N^\xi}|^\zeta\mu(B)^2} \\
&\qquad= \mathcal{O}(t)ns\mu(B)^{\xi\zeta\eta-2} \\
&\qquad= \mathcal{O}(t)n\mu(B)^{\xi\zeta\eta-2}\left(\mu(B)^{-\frac{1}{2+\kappa}} + \mu(B)^{-\eta}\right).
\end{aligned}$$

by (3) and as $s\lambda_B = \mathcal{O}(1)$ and $s = \alpha^{-1}(C'\mu(\widetilde{B})) + N(n)$ by Lemma 4. Choosing η close to $\frac{1}{2}$ will achieve that $\omega_2 = \xi\zeta\eta - 2 - \max\{\frac{1}{2+\kappa}, \eta\}$ is positive. Now put $\omega = \min\{\omega_1, \omega_2\}$. □

Proof of Theorem 5. The first part of the proof is identical to the proof of Theorem 4. For the second part we get different estimates as $\gamma_n = \text{diam}(\mathcal{A}^n) = \mathcal{O}(n^{-a})$ for some $a > 0$. To prove the theorem for the set B we use Lemma 7 and put $N(n) = \mu(B)^{-\eta}$ for some $\eta \in (0, 1/2)$. Thus

$$\begin{aligned}\left|\mathbb{P}_B(\tau_B > \frac{t}{\lambda_B\mu(B)}) - \mathbb{P}_{\widetilde{B}}(\tau_{\widetilde{B}} > \frac{t}{\lambda_B\mu(\widetilde{B})})\right| \\ &\le C_8\frac{t\vartheta_n(\epsilon)}{\mu(B)} \\ &= \frac{\mathcal{O}(t)ns}{N(n)^{a\zeta}\mu(B)^2} \\ &= \mathcal{O}(t)n\mu(B)^{a\zeta\eta-2}\left(\mu(B)^{-\frac{1}{2+\kappa}} + \mu(B)^{-\eta}\right).\end{aligned}$$

by (3). A choice of η close to $\frac{1}{2}$ will achieve that $\omega_2 = a\zeta\eta - 2 - \max\{\frac{1}{2+\kappa}, \eta\}$ is positive. Now put $\omega = \min\{\omega_1, \omega_2\}$ (ω_1 from the proof of Theorem 4). □

6. Maps with Young's tower

In this section we show how the results of the previous section can be applied to dynamical systems that can be modelled by a Markov tower as Young constructed in [29, 30].

We assume that T is a differentiable map on the manifold X. Then one assumes there is a subset $\Omega_0 \subset X$ with the following properties:
(i) Ω_0 is partitioned into disjoint sets $\Omega_{0,i}$, $i = 1, 2, \dots$ and there is a *return time function* $R : \Omega_0 \to \mathbb{N}$, constant on the partition elements $\Omega_{0,i}$, such that T^R maps $\Omega_{0,i}$ bijectively to the entire set Ω_0. We write $R_i = R|_{\Omega_{0,i}}$. Moreover, it is assumed that the $\Omega_{0,i}$ are rectangles, that is, if $\gamma^u(x)$ denotes the unstable leaf through $x \in \Omega_{0,i}$ and $\gamma^s(y)$ the stable leaf at $y \in \Omega_{0,i}$, then there is a unique interestion $\gamma^u(x) \cap \gamma^s(y)$ which also lies in $\Omega_{0,i}$. It is also assumed that the $\Omega_{0,i}$ satisfy the Markov property. If γ^u and $\hat{\gamma}^u$ are two unstable leaves (in some $\Omega_{i,0}$), then the holonomy $\Theta : \gamma^u \to \hat{\gamma}^u$ is given by $\Theta(x) = \hat{\gamma}^u \cap \gamma^s(x)$, $x \in \gamma^u$.
(ii) For $j = 0, 1, \dots, R_i - 1$ put $\Omega_{j,i} = \{(x, j) : x \in \Omega_{0,i}\}$ and define $\Omega = \bigcup_{i=1}^\infty \bigcup_{j=0}^{R_i-1} \Omega_{j,i}$. Ω is the *Markov tower* for the map T. It has the associated partition $\mathcal{A} = \{\Omega_{j,i} : 0 \le j < R_i, i = 1, 2, \dots\}$ which typically is countably infinite. The map $F : \Omega \to \Omega$

is given by

$$F(x,j) = \begin{cases} (x, j+1) & \text{if } j < R_i - 1 \\ (\hat{F}x, 0) & \text{if } j = R_i - 1 \end{cases}$$

where we put $\hat{F} = F^R$ for the induced map on Ω_0.

(iii) The *separation function* $s(x,y)$, $x, y \in \Omega_0$, is defined as the largest positive n so that $(T^R)^j x$ and $(T^R)^j y$ lie in the same sub-partition elements for $0 \le j < n$. That is $(T^R)^j x, (T^R)^j y \in \Omega_{0,i_j}$ for some $i_0, i_1, \dots, n-1$. We extend the separation function to all of Ω by putting $s(x,y) = s(T^{R-j}x, T^{R-j}y)$ for $s, y \in \Omega_{j,i}$.

(iv) Let ν be a finite given 'reference' measure on Ω and let ν_{γ^u} be the conditional measure on the unstable leaves. We assume that the Jacobian $JF = \frac{d(F_*^{-1}\nu_{\gamma^u})}{d\nu_{\gamma^u}}$ is Hölder continuous in the following sense: there exists a $\gamma \in (0,1)$ so that

$$\left| \frac{JF^R x}{JF^R y} - 1 \right| \le \text{const}\gamma^{s(\hat{F}x, \hat{F}y)}$$

for all $x, y \in \Omega_{0,i}$, $i = 1, 2, \dots$.

If the return time R is integrable with respect to ν then by [30] Theorem 1 there exists an F-invariant probability measure μ (SRB measure) on Ω which is absolutely continuous with respect to ν.

For each $n \in \mathbb{N}$ the elements of the nth join $\mathcal{A}^n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{A}$ of the partition $\mathcal{A} = \{\Omega_{i,j}\}$ are called n-cylinders and form a new partition of Ω, a refinement of the original partition. The σ-algebra generated by all n-cylinders $\mathcal{A}^\ell$, for all $\ell \ge 1$, is the σ-algebra of the system (Ω, μ).

In order to prove Theorem 3 we verify the conditions in Theorem 8.

Lemma 9. *Assume that $\nu(R > k) \le p(k)$ where $p(k)$ is a decreasing sequence in k that is at least polynomial with power > 1. Then the invariant measure μ is α-mixing with respect to the partition $\mathcal{A}$, with $\alpha(k) \sim p(k)$.*

Proof. Denote by $\mathcal{C}_\gamma$ the space of Hölder continuous functions φ on Ω for which $|\varphi(x) - \varphi(y)| \le C_\varphi \gamma^{s(x,y)}$. If C_φ is smallest then $\|\varphi\|_\gamma = |\varphi|_\infty + C_\varphi$ defines a norm and $\mathcal{C}_\gamma = \{\varphi : \ \|\varphi\|_\gamma < \infty\}$. Let $\mathcal{L} : \mathcal{C}_\gamma \to \mathcal{C}_\gamma$ be the transfer operator defined by $\mathcal{L}\varphi(x) = \sum_{x' \in T^{-1}x} \frac{\varphi(x')}{JT(x')}$, $\varphi \in \mathcal{C}_\gamma$. Then ν is a fixed point of its adjoint, i.e. $\mathcal{L}^*\nu = \nu$ and $h = \frac{d\mu}{d\nu} = \lim_{n\to\infty} \mathcal{L}^n \lambda$ is Hölder continuous, where λ can be any initial density distribution in $\mathcal{C}_\gamma$. In fact, by [30] Theorem 2(II) one has

$$\|\mathcal{L}^k \lambda - h\|_{\mathscr{L}^1} \le p(k) \|\lambda\|_\gamma \tag{4}$$

where the 'decay function' $p(k) = \mathcal{O}(k^{-\beta})$ if the return times decay polynomially with power β, that is if $\nu(R > j) \le \text{const.} j^{-\beta}$. If the return times decay exponentially, i.e. if $\nu(R > j) \le \text{const.}\vartheta^j$ for some $\vartheta \in (0,1)$, then there is a $\tilde{\vartheta} \in (0,1)$ so that $p(k) \le \text{const.}\tilde{\vartheta}^k$.

As in the proof of [30] Theorem 3 we put $\lambda = \mathcal{L}^n h\chi_A$ which is a strictly positive function. Then $\eta = \frac{\lambda}{\mu(A)}$ is a density function as $\nu(\lambda) = \nu(\mathcal{L}^n h\chi_A) = \nu(h\chi_A) = \mu(A)$. Since by [16] there exists a constant c_1 so that $\|\mathcal{L}^n\chi_A\|_\gamma \le c_1$ for all $A \in \sigma(\mathcal{A}^n)$ and n we see that $\|\lambda\|_\gamma \le c_1$ uniformly in n and $A \in \sigma(\mathcal{A}^n)$. Hence

$$\begin{aligned}
\mu(A\cap T^{-k-n}B) - \mu(A)\mu(B) &= \nu(h\chi_A(\chi_B\circ T^{k+n})) - \nu(h\chi_A)\nu(h\chi_B)\\
&= \mu(A)(\nu(\chi_B\mathcal{L}^k\eta) - \nu(h\chi_B))\\
&= \mu(A)\int \chi_B(\mathcal{L}^k\eta - h)\,d\nu\\
&= \int_B (\mathcal{L}^k\lambda - \mu(A)h)\,d\nu.
\end{aligned}$$

Using the estimates from the $\mathscr{L}^1$-convergence of $\mathcal{L}^k\eta - h$ from (4) yields

$$\begin{aligned}
\left|\mu(A\cap T^{-k-n}B) - \mu(A)\mu(B)\right| &\le \mu(A)\int \chi_B|\mathcal{L}^k\lambda - h|\,d\nu\\
&\le \mu(A)c_1\|\eta\|_\gamma p(k)\\
&\le c_3 p(k)
\end{aligned}$$

as $\|\eta\|_\gamma = \frac{1}{\mu(A)}\|\lambda\|_\gamma \le \frac{C_3}{\mu(A)}$. In particular we can write

$$\left|\mu(A\cap T^{-k-n}B) - \mu(A)\mu(B)\right| \le \alpha(k)$$

for all $A \in \sigma(\mathcal{A}^n)$, $B \in \sigma(\bigcup_{j\ge1}\mathcal{A}^j)$, where $\alpha(k) = c_3p(k)$. $\square$

Lemma 10. *Let $\zeta < \frac{\lambda}{2} - 1$. Then there exists an ϵ_0 so that for every $\delta < \epsilon_0$ there exists a set $\mathcal{U}_\delta \subset X$, of measure $\mathcal{O}(|\log\delta|^{-\zeta})$ so that $\varphi(\epsilon,\delta,x) = \mathcal{O}(|\log\delta|^{-\zeta})$ uniformly in $x \notin \mathcal{U}_\delta$.*

Proof. It was shown in [17] Proposition 6.1 that for all w large enough there exists a set $\mathcal{U} \subset X$ such that $\mu(\mathcal{U}) = \mathcal{O}((w|\log\epsilon|)^{-\zeta})$ and $\varphi(\epsilon,\epsilon^w,x) = \mathcal{O}((w|\log\epsilon|)^{-\zeta})$ uniformly in $x \notin \mathcal{U}$ where ζ is any number less than $\frac{\lambda}{2} - 1$. Hence there exists an $\epsilon_0 > 0$ so that we can write $\delta = \epsilon^w$ with w large enough (larger than $\frac{2}{u}(D+1) - 1$ where D is the dimension of the manifold X and u is the dimension of the unstable leaves) for all $\delta < \epsilon_0$. Since $\log\delta = w\log\epsilon$ we obtain the statement of the lemma. $\square$

Let us denote by $\tilde\Omega_{j,i}$ the principal parts of $\Omega_{j,i}$. For integers N, m ($N \gg m$) we put $\tilde\Omega_{j,i} = \{x \in \Omega_{j,i} : R(\hat F^j x) \le s\ \forall\ j = 0,\dots,[N/m]\}$. In this way we pick out the return times that are not too long. In particular $\tilde\Omega_{0,i} = \varnothing$ if $R_i > m$. Let us put $\tilde\Omega = \bigcup_i \bigcup_{j=0}^{R_i-1}\tilde\Omega_{j,i}$ (disjoint unions).

We also define $\omega(m) = \sqrt{\sum_{i:R_i>m} R_i\,\nu(\Omega_{0,i})}$ and note that $\omega(m) = \mathcal{O}(m^{-\frac{\lambda-1}{2}})$.

Lemma 11. [17] *There exists a constant C_9 and for $N, n, m \geq 1$ $(N > n, m)$ there exist sets $\mathcal{V}_{N,m} \subset M$ such that the non-principal part contributions are estimated as*

$$\mu(\mathcal{B} \cap (\Omega \setminus \tilde{\Omega})) < \sqrt{n+2}\,\omega(m)\mu(B_{\epsilon,n})$$

for any $\mathcal{B} \subset B_{\epsilon,n}(x)$ and $x \notin \mathcal{V}_{N,m}$ where

$$\mu(\mathcal{V}_{N,m}) \leq C_9\sqrt{n+2}\,\omega(m).$$

Proof of Theorem 3. In order to apply Theorem 8 to $B_{\epsilon,n}(x)\cap\tilde{\Omega}$ we will pick $\eta \in (0, \frac{1}{2})$ below and put $N(n) = [\mu(B_{\epsilon,n}(x))^{-\eta}]$. We then choose $m = N^\alpha$ for some $\alpha \in (\frac{1}{\lambda-1}, 1)$ (see estimate of $\mathcal{F}$ below). Then according to Lemma 11 $\mathrm{diam}(A) \leq \gamma^{\frac{N}{m}}$ for some $\gamma < 1$ for all n-cylinders A which belong to $\tilde{\Omega}$. As in the proof of Theorem 1 we then conclude that

$$\varphi(\epsilon, \gamma_{N(n)-k}, T^k x) \leq \frac{c_1 m^\zeta}{N(n)^\zeta} \leq c_1 m^\zeta \mu(B)^{\zeta\eta}$$

(for some constant c_1) provided $T^k x \notin \mathcal{U}_{\gamma_{N(n)-k}}$ for $k = 0, \dots, n-1$. Since $s = \alpha^{-1}(C'\mu(B)) + N(n) \leq c_2\mu(B)^{-\frac{1}{\lambda}} + \mu(B)^{-\eta}$, we obtain

$$\frac{ns \cdot \varphi(\epsilon, \gamma_{N(n)-k}, T^k x)}{\mu(B)} \leq c_3 n \cdot \mu(B)^{\eta\zeta-1-\eta\alpha\zeta}\left(\mu(B)^{-\frac{1}{\lambda}} + \mu(B)^{-\eta}\right).$$

The RHS converges to zero if $\eta\zeta - 1 - \eta\alpha\zeta - \max\{\frac{1}{\lambda}, \eta\}$ is positive. To satisfy Lemma 10 it is required that $\zeta < \frac{\lambda}{2} - 1$. Then can choose $\alpha \in (\frac{1}{\lambda-1}, 1)$ in such a way that the above expression is positive for an $\eta < \frac{1}{2}$ close to $\frac{1}{2}$. This can be done if $\lambda > 5 + \sqrt{15}$.

We now proceed as in the proof Theorem 1 to estimate the contribution to the error made by $(B \setminus \widetilde{B}) \cap \tilde{\Omega}$. For the portion that lies in $\Omega \setminus \tilde{\Omega}$ we use Lemma 11 and thus obtain combining the two contributions:

$$\left|\mathbb{P}(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}) - e^{-t}\right| \leq c_4\left(t\mu(B)^a + \mu(B)^b + \sqrt{N}\,\omega(m)\mu(B)\right)$$

provided x does not lie in the forbidden set $\mathcal{F} = \mathcal{V}_{N,m} \cup \bigcup_{k=0}^{n-1} T^{-k}\mathcal{U}_{\gamma_{N(n)-k}}$ whose measure is by Lemmata 11 and 10 bounded by

$$\mu(\mathcal{F}) \leq c_5\left(\sqrt{N+2}\,\omega(m) + n|\log\gamma_{N(n)}|^{-\zeta}\right) \leq c_6\left(\sqrt{N}\,m^{-\frac{\lambda-1}{2}} + nN^{-\zeta}m^\zeta\right)$$

which goes to zero as $n \to \infty$ since $m = N^\alpha$ and $\frac{1}{\lambda-1} < \alpha < 1$. Thus $\mu(\mathcal{F}) \to 0$ as $n \to \infty$ and therefore $\mathbb{P}(\tau_{B_{\epsilon,n}(x)} > \frac{t}{\lambda_{B_{\epsilon,n}(x)}\mu(B_{\epsilon,n}(x))}) \longrightarrow e^{-t}$ as $n \to \infty$, $\epsilon \to 0$ for every $x \notin \liminf_{n\to\infty,\epsilon\to 0}\mathcal{F}_{\epsilon,n}$.

□

References

[1] M. Abadi: Exponential Approximation for Hitting Times in Mixing Processes; Math. Phys. Electron. J. **7** (2001), Paper 2, 19 pp.

[2] M. Abadi: Sharp Error Terms and Neccessary Conditions for Exponential Hitting Times in Mixing Processes; Ann. of Probab. **32** (2004), 243–264.

[3] M. Abadi: Hitting, Returning and the Short Correlation Function; Bull. Braz. Math. Soc. (N.S.) **37** (2006), 593–609.

[4] M Abadi and B Saussol: Almost sure convergence of the clustering factor in α-mixing processes; Stoch. Dynam. **16(3)** (2016).

[5] A D Barbour and L H Y Chen: *An Introduction to Stein's Method*; Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, Vol. **4** 2005.

[6] R Bowen: *Equilibrium States for Anosov Diffeomorphism*; Springer Lecture Notes 470, Springer, New York/Berlin 1975. Second revised edition Springer-Verlag, Berlin, 2008.

[7] S M Buckley: Is the maximal function of a Lipschitz function continuous? Ann. Acad. Scient. Fenn. Math. **24** (1999), 519–528.

[8] M Brin and A Katok: On local entropy; Proceedings, Springer

[9] 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.

[10] M Denker: Remarks on weak limit laws for fractal sets; *Progress in Probability* Vol. 37, Birkhäuser 1995, 167–178.

[11] T Downarowicz and B Weiss: Entropy theorem along times when x visits a set; Ill. J. Math. **48** (2004), 59–69.

[12] A. Galves and B. Schmitt: Inequalities for Hitting Times in Mixing Dynamical Systems; Random Comput. Dynam. **5** (1997), 337–347.

[13] Gora and Boyarski:

[14] N T A Haydn: Entry and return times distribution; Dynamical Systems: An International Journal **28(3)** (2013), 333–353.

[15] N. Haydn, Y. Lacroix and S. Vaienti: Hitting and Return Times in Ergodic Dynamical Systems; Ann. of Probab. **33** (2005), 2043–2050.

[16] N Haydn and Y Psiloyenis: Return times distribution for Markov towers with decay of correlations; to appear in Nonlinearity.

[17] N Haydn and K Wasilewska: Limiting distribution and error terms for the number of visits to balls in non-uniformly hyperbolic dynamical systems; available at `http://arxiv.org/abs/1402.2990`.

[18] M Kupsa and Y Lacroix: Asymptotics for hitting times, Ann. of Probab. **33(3)** (2005), 610–614.

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

[20] R. Mañé: *Ergodic Theory and Differentiable Dynamics*; Springer 1987.

[21] Ornstein and Weiss; Entropy and Data Compression Schemes; IEEE Transactions on Information Theory 39 (1993), 78–83

[22] Ornstein and Weiss; Entropy and Recurrence Rates for Stationary Random Fields; IEEE Transactions on Information Theory 48 (6) (2002), 1694–97

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

[24] Y Pesin and H Weiss: A Multifractal Analysis of Equilibrium Measures for Conformal Expanding Maps and Moran-like Geometric Constructions; J. Stat. Phys. **86(1/2)** (1997), 233–275.

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

[26] D Ruelle: *Thermodynamic formalism*; Encyclopedia of Mathematics and its Applications Vol. 5, Addison-Wesley (1978).

[27] P Varandas: Entropy and Poincaré recurrence from a geometrical viewpoint; Nonlinearity **22(10)** (2009), 2365–2375.

[28] P. Walters: An Introduction to Ergodic Theory; Springer 1982.

[29] L-S Young: Statistical properties of dynamical systems with some hyperbolicity; Annals of Math. **7** (1998), 585–650.

[30] L-S Young: Recurrence time and rate of mixing; Israel J. of Math. **110** (1999), 153–188.