ESCAPE RATE AND CONDITIONAL ESCAPE RATE FROM A PROBABILISTIC POINT OF VIEW

C DAVIS, N HAYDN, AND F YANG

ABSTRACT. We prove that for a sequence of nested sets $\{U_n\}$ with $\Lambda = \cap_n U_n$ a measure zero set, the localized escape rate converges to the extremal index of Λ, provided that the dynamical system is ϕ-mixing at polynomial speed. We also establish the general equivalence between the local escape rate for entry times and the local escape rate for returns.

CONTENTS

1. INTRODUCTION

In recent years there has been an increasing interest in open dynamical systems, which are dynamical systems with an invariant measure where one places a trap or hole in the phase space, and looks at the decay rate of the measure of points that are not caught by the trap up to some time (the survival set). This rate is known to be related to the rate of the correlations decay for the system (see [16]). When the correlations decay exponentially fast, the decay rate for the measure of the survival set is typically exponential , and depends on the location and size of the trap. We invite the readers to the review article [6] for a general overview on this topic.

When the decay rate for the measure of the survival set is normalised by the measure of the trap, one obtains the localized escape rate as the measure of the

C Davis, Department of Mathematics, University of Oklahoma, Norman, 73019. E-mail: `cdavis@math.ou.edu`.

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 Oklahoma, Norman, 73019. E-mail: `fan.yang-2@ou.edu`.

trap goes to zero. Such problems are related to the entry times and return times distribution but this similarity does not allow to deduce limiting statistics from each other since the limits are taken in different ways, and the rate of convergence for the entry times to its limiting distribution is usually insufficient for the study of escape rates.

In the past local escape rates have been associated with either metic holes centred at a point whose radius decreased to zero or, in the presence of partitions, with cylinder sets which decrease to a single point. In this case a dichotomy has been established for many systems which shows the local escape rate to be equal to one at non-periodic points and equal to the extremal index at periodic points. See the classical work [9] for conformal attractors, [2, 5] for the transfer operator approach for interval maps, and [15] for a probabilistic approach which applies to systems in higher dimension. This mirrors the behaviour of the limiting return times distributions that are Poisson at non-periodic points, and Pólya-Aeppli compound Poisson at periodic points in which case the compounded geometric distribution has the weights given by the extremal index $\theta \in (0, 1)$. See [13].

In this paper we will generalise the concept of localized escape rates to the cases when the limiting set of the shrinking neighbourhoods are not any longer points, periodic or non-periodic, but instead are allowed to be any null sets. Recently such a more general setting was used to obtain the limiting return times distribution at null sets which turn out to be compound Poisson in a more general sense where the associated parameters form a spectrum that is determined by the limiting cluster size distribution (for this see below the coefficients α_k). For more detail on this topic, see [14] and [18].

Unlike the conventional transfer operator method studied in [2] following the general setup in [16] where exponential decay of correlations is required, our approach in this paper only assumes the system to be ϕ-mixing at polynomial speed (surprisingly, this is enough to deduce that the escape rate is exponential!). This assumption may still not be optimal. However, we believe that the same results does not hold for α-mixing systems in general. See the counter examples in [2] for systems modeled by Young's towers with polynomial tail, and note that such systems are α-mixing at the same rate.

2. Statement of results

Throughout this paper, we will assume that $(\mathbf{M}, T, \mathcal{B}, \mu)$ is a measure preserving system on some compact metric space $\mathbf{M}$, with $\mathcal{B}$ the Borel σ-algebra. Unless otherwise specified, T is assumed to be non-invertible, although our result also holds in the invertible case, see Remark 2.3 and Theorem 4.12, 5.2 below. Let $\mathcal{A}$ be a measurable partition of $\mathbf{M}$ (finite or countable). Denote, by $\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j} \mathcal{A}$, its nth join (in the invertible case, see Remark 2.3). Then $\mathcal{A}^n$ is a partition of $\mathbf{M}$ and its elements are called n-cylinders. We assume that $\mathcal{A}$ is generating, that is $\bigcap_n A_n(x)$ consists of the singleton $\{x\}$.

Definition 1. (i) The measure μ is *left ϕ-mixing* with respect to $\mathcal{A}$ if

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

for all $A \in \sigma(\mathcal{A}^n)$, $n \in \mathbb{N}$ and $B \in \sigma(\bigcup_{j \geq 1} \mathcal{A}^j)$, where $\phi(k)$ is a decreasing function which converges to zero as $k \to \infty$. Here $\sigma(\mathcal{A}^n)$ is the σ-algebra generated by

n-cylinders.

(ii) The measure μ is *right ϕ-mixing* w.r.t. $\mathcal{A}$ if

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

for all $A \in \sigma(\mathcal{A}^n)$, $n \in \mathbb{N}$ and $B \in \sigma(\bigcup_j \mathcal{A}^j)$, where $\phi(k) \searrow 0$.

(iii) The measure μ is *ψ-mixing* w.r.t. $\mathcal{A}$ if

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

for all $A \in \sigma(\mathcal{A}^n)$, $n \in \mathbb{N}$ and $B \in \sigma(\bigcup_j \mathcal{A}^j)$, where $\psi(k) \searrow 0$. Clearly ψ-mixing implies both left and right ϕ-mixing with $\phi(k) = \psi(k)$.

Remark 2.1. If it is clear from context which type of mixing we are referring to (as is always the case in this paper), then we will suppress the subscripts R, L.

We write, for any subset $U \subset \mathbf{M}$,

$$\tau_U(x) = \min\{j \geq 1 : T^j(x) \in U\}$$

the first entry time to the set U. Then $\tau_U|_U$ is the first return time for points in U. We can now define the *escape rate* into U by

$$\rho(U) = \lim_{t \to \infty} \frac{1}{t}|\log \mathbb{P}(\tau_U > t)|$$

whenever the limit exists. It captures the exponential decay rate for the set of points whose orbits have not visited U before time t. Observe that if $U \subset U'$ then $\mathbb{P}(\tau_{U'} > t) \leq \mathbb{P}(\tau_U > t)$ and consequently $\rho(U) \leq \rho(U')$. We define the *conditional escape rate* as

$$\rho_U(U) = \lim_{t \to \infty} \frac{1}{t}|\log \mathbb{P}_U(\tau_U > t)|,$$

where $\mathbb{P}_U$ is the conditioned measure on U.

We are particularly interested in the asymptotic behavior of $\rho(U)$ along a sequence of positive measure sets $\{U_n\}$ whose measure goes zero. For this purpose, we call $\{U_n\}$ a *nested sequence of sets* if $U_{n+1} \subset U_n$, and if $\lim_n \mu(U_n) = 0$. For the measure zero set $\Lambda = \cap_n U_n$, we define the *localized escape rate at Λ* as:

$$(1) \quad \{\{\texttt{e.escaperate}\}\} \qquad \rho(\Lambda, \{U_n\}) = \lim_{n \to \infty} \frac{\rho(U_n)}{\mu(U_n)}.$$

provided that the limit exists. We will show that under certain conditions, the localized escape rate of Λ exists and does not depend on the choice of U_n.

2.1. **Local escape rate for unions of cylinders.** First, we consider with the case where each U_n is a union of κ_n-cylinders, for some non-decreasing sequence of integers $\{\kappa_n\}$.

We make some assumptions on the sizes of the nested sequence $\{U_n\}$. For each n and $j \geq 1$, we define $\mathcal{C}_j(U_n) = \{A \in \mathcal{A}^j, A \cap U_n \neq \emptyset\}$ the collection of all j-cylinders that have non-empty intersection with U_n. Then, we write

$$U_n^j = \bigcup_{A \in \mathcal{C}_j(U_n)} A$$

for the approximation of U_n by j-cylinders from outside. For each fixed j, $\{U_n^j\}_n$ is also nested, that is, $U_{n+1}^j \subset U_n^j$. Obviously we have $U_n \subset U_n^j$ for all j, and $U_n = U_n^j$ if $j \geq \kappa_n$.

Definition 2. A nested sequence $\{U_n \in \mathcal{A}^{\kappa_n}\}$ is called a *good neighborhood system*, if:

(1) $\kappa_n \nearrow \infty$ and $\kappa_n \mu(U_n)^\varepsilon \to 0$ for some $\varepsilon \in (0,1)$;
(2) there exists $C > 0$ and $p' > 1$ such that $\mu(U_n^j) \leq \mu(U_n) + Cj^{-p'}$ for all $j < \kappa_n$.

2.1.1. *Local escape rate and the extremal index.* We make the following definitions, following [14]

For a positive measure set $U \subset \mathbf{M}$, we define the higher order entry times recursively:

$$\tau_U^1 = \tau_U, \text{ and } \tau_U^j(x) = \tau_U^{j-1}(x) + \tau_U(T^{\tau_U^{j-1}}(x)).$$

For simplicity, we write $\tau_U^0 = 0$ on U.

For a nested sequence $\{U_n\}$, define

$$(2) \qquad \hat{\alpha}_\ell(K, U_n) = \mu_{U_n}(\tau_{U_n}^{\ell-1} \leq K),$$

i.e. $\hat{\alpha}_\ell(K, U_n)$ is the conditional probability of having at least $(\ell - 1)$ returns to U_n before time K. We shall assume that the limit $\hat{\alpha}_\ell(K) = \lim_{n \to \infty} \hat{\alpha}_\ell(K, U_n)$ exists for for $K \in \mathbb{N}$ large enough and every $\ell \geq 1$. By monotonicity the limits

$$(3) \qquad \hat{\alpha}_\ell = \lim_{K \to \infty} \hat{\alpha}_\ell(K)$$

exist as $\hat{\alpha}_\ell(K) \leq \hat{\alpha}_\ell(K') \leq 1$ for all $K \leq K'$. Since we put $\tau_U^0 = 0$, it follows that $\hat{\alpha}_1 = 1$. We will see later that the existence of the limits defining $\hat{\alpha}_\ell$ implies the existence of the following limits:

$$(4) \qquad \alpha_1 = \lim_{K \to \infty} \lim_{n \to \infty} \mu_{U_n}(\tau_{U_n} > K) = 1 - \hat{\alpha}_2.$$

$\alpha_1 \in [0, 1]$ is generally known as the *extremal index* (EI). See the discussion in Freitas et al [10].

The next theorem shows that the escape rate is indeed given by the extremal index.

Theorem A. *Assume that $T : \mathbf{M} \to \mathbf{M}$ preserves a probability measure μ that is right ϕ-mixing with $\phi(k) \leq Ck^{-p}$ for some $C > 0$ and $p > 1$, and $\{U_n\}$ is a good neighborhood system such that $\{\hat{\alpha}_\ell\}$ defined in (3) exists, and satisfies $\sum_\ell \ell \hat{\alpha}_\ell < \infty$.*

Then α_1 defined by (4) exists, and the localized escape rate at Λ exists and satisfies

$$\rho(\Lambda, \{U_n\}) = \alpha_1.$$

Remark 2.2. Theorem A has a similar formulation for left ϕ-mixing systems. See Remark 4.7 and Theorem 4.12, 5.2 for more detail.

For Gibbs-Markov systems (for the precise definition, see Section 3) the same result is true:

Theorem B. *Assume that $T : \mathbf{M} \to \mathbf{M}$ is a Gibbs-Markov system with respect to the partition $\mathcal{A}$. Let $\{U_n\}$ be a good neighborhood system such that $\{\hat{\alpha}_\ell\}$ defined in (3) exists, and satisfies $\sum_\ell \ell \hat{\alpha}_\ell < \infty$.*

Then α_1 defined by (4) exists. Furthermore, the localized escape rate at Λ exists and satisfies

$$\rho(\Lambda, \{U_n\}) = \alpha_1.$$

{r.1}

Remark 2.3. If T is invertible, then one has to define the n-join by

$$\mathcal{A}^n = \bigvee_{j=-n}^{n} T^{-j}\mathcal{A}.$$

In this case it is useful to write, for $m, n \in \mathbb{Z}$, $\mathcal{A}_m^n := \bigvee_{j=m}^{n} T^{-j}\mathcal{A}$. In particular we have $\mathcal{A}^n = \mathcal{A}_{-n}^n$. The ϕ and ψ-mixing properties are defined by the same formulas. For integers $m, m', n, n' > 0$, if $U \in \mathcal{A}_{-m}^n, V \in \mathcal{A}_{-m'}^{n'}$ then for $k > n + m'$ we have $\mu(U \cap T^{-k}V) = \mu(T^{-m}U \cap T^{-k-m}V)$ where $T^{-m}U \in \mathcal{A}_0^{m+n}, T^{-k-m}V \in \mathcal{A}_{k+m-m'}^{k+m+n'}$. Note that $k+m-m' > n+m > 0$, so the estimate can be treated in the same way as the non-invertible case with only minor adjustments. However, the approximation U_n^j need to be handled with care. See Remark 4.8 for a potential problem and the treatment.

2.1.2. In the absence of short returns. We will see below that when points in $\{U_n\}$ do not return to U_n "too soon", then $\alpha_1 = 1$. To formulate this, we define the *periodic of U* as:

$$\pi(U) = \min\{k > 0 : T^{-k}U \cap U \neq \emptyset\},$$

and the *essential periodic of U* by:

$$\pi_{\mathrm{ess}}(U) = \min\{k > 0 : \mu(T^{-k}U \cap U) > 0\}.$$

π and π_{ess} mark the shortest return of points in U. Note that $\pi(U) \leq \pi_{\mathrm{ess}}(U)$ for all measurable $U \in \mathbf{M}$, and equality holds if T is continuous, U is open and μ has full support.

{mc.0}

Corollary C. *Let $(\mathbf{M}, T, \mathcal{B}, \mu)$ be a measure preserving system. Assume that $\{U_n\}$ is a good neighborhood system with $\pi_{\mathrm{ess}}(U_n) \to \infty$, and $(T, \mu, \mathcal{A})$ satisfies one of the following two assumptions:*

(1) either μ is right ϕ-mixing with $\phi(k) \leq Ck^{-p}$ for some $p > 1$;
(2) or T is Gibbs-Markov;

then the localized escape rate at Λ exists and satisfies

$$\rho(\Lambda, \{U_n\}) = 1.$$

As an immediate corollary, we have:

{mc.1}

Corollary D. *The conclusion of Corollary C holds if the assumption "$\pi_{\mathrm{ess}}(U_n) \to \infty$" is replaced by the following assumptions:*

(1) T is continuous, $\Lambda = \cap_n U_n = \cap_n \overline{U}_n$;
(2) Λ intersects every forward orbit at most once, that is, for every $x \in \Lambda$ we have $\Lambda \cap \{T^k(x) : k \geq 1\} = \emptyset$.

2.2. From cylinders to open sets: exceedance rate for the extreme value process. Next, we deal with the case where $\{U_n\}$ consists of open sets. For this purpose, we consider a continuous observable

$$\varphi : \mathbf{M} \to \mathbb{R} \cup \{+\infty\},$$

such that the maximal value of φ, which could be positive infinite, is achieved on a μ measure zero closed set Λ. Then we consider the process generated by the dynamics of T and the observable φ:

$$X_0 = \varphi, \ X_1 = \varphi \circ T, \ \ldots, X_k = \varphi \circ T^k, \ldots.$$

Let $\{u_n\}$ be a non-decreasing sequence of real numbers. We will think of u_n as a sequence of thresholds, and the event $\{X_k > u_n\}$ marks an exceedance above the threshold u_n. Also denote by U_n the open set

(5) $$U_n := \{X_0 > u_n\}.$$

It is clear that $\{U_n\}$ is a nested sequence of sets.

We are interested in the set of points where $X_k(x)$ remains under the threshold u_n before time t. For this purpose, we put

$$M_n = \max\{X_k : k = 0, \ldots, n-1\},$$

and

$$\zeta(u_n) = \lim_{t \to \infty} \frac{1}{t} |\log \mathbb{P}(M_t < u_n)|.$$

Finally, define the *exceedance rate of φ along the thresholds $\{u_n\}$* as:

$$\zeta(\varphi, \{u_n\}) = \lim_{n \to \infty} \frac{\zeta(u_n)}{\mu(U_n)}.$$

We will make the following assumption on the shape of U_n. For each $r_n > 0$, we approximate U_n by two open sets ('o' and 'i' stand for 'outer' and 'inner'):

$$U_n^o = B_{r_n}(U_n) = \bigcup_{x \in U_n} B_{r_n}(x), \quad \text{and} \quad U_n^i = U_n \setminus \overline{B_{r_n}(U_n)} = U_n \setminus \left(\bigcup_{x \in \partial U_n} \overline{B_{r_n}(x)} \right).$$

It is easy to see that

$$\overline{U_n^i} \subset U_n \text{ and } \overline{U_n} \subset U_n^o,$$

The following assumption requires U_n to be *well approximable* by $U_n^{i/o}$.

{a.1}

Assumption 1. *There exists a positive, decreasing sequence of real numbers $\{r_n\}$ with $r_n \to 0$, such that*

(6) $$\mu\left(U_n^o \setminus U_n^i\right) = o(1)\mu(U_n).$$

Here $o(1)$ means the term goes to zero under the limit $n \to \infty$.

{m.3}

Theorem E. *Assume that*

(1) either μ is right ϕ-mixing with $\phi(k) \le Ck^{-p}$, $p > 1$;

(2) or $(T, \mu, \mathcal{A})$ is a Gibbs-Markov system.

Let $\varphi : \mathbf{M} \to \mathbb{R} \cup \{+\infty\}$ be a continuous function achieving its maximum on a measure zero set Λ. Let $\{u_n\}$ be a non-decreasing sequence of real numbers with $u_n \nearrow \sup \varphi$, such that the open sets U_n defined by (5) satisfy Assumption 1 for a sequence r_n that decreases to 0. Assume $\{\hat{\alpha}_\ell\}$, defined by (3), exist and satisfy $\sum_\ell \ell \hat{\alpha}_\ell < \infty$. Let κ_n be the smallest positive integer for which $\operatorname{diam} \mathcal{A}^{\kappa_n} \le r_n$ and assume:

(a) $\kappa_n \mu(U_n)^\varepsilon \to 0$ for some $\varepsilon \in (0, 1)$;

(b) U_n has small boundary: there exist $C > 0$ and $p' > 1$, such that

$$\mu\left(\bigcup_{A \in \mathcal{A}^j, A \cap B_{r_n}(\partial U_n) \neq \emptyset} A\right) \le Cj^{-p'} \text{ for all } n \text{ and } j \le \kappa_n.$$

Then the exceedance rate of φ along $\{u_n\}$ exists and satisfies

$$\zeta(\varphi, \{u_n\}) = \alpha_1.$$

2.3. **Conditional escape rate: a general theorem.** Our next theorem deals with the relation between the escape rate and the conditioned escape rate.

{m.4}

Theorem F. *For any measure preserving system* $(\mathbf{M}, T, \mathcal{B}, \mu)$ *and any positive measure set* $U \in \mathbf{M}$, *we have*

$$\rho(U) = \rho_U(U),$$

assuming one of them exists.

Note that this theorem does not rely on the mixing assumption nor any information on the shape of U. In particular, if one defines the *localized conditional escape rate at* Λ as

$$\rho_\Lambda(\Lambda, \{U_n\}) = \lim_{n \to \infty} \frac{\rho_{U_n}(U_n)}{\mu(U_n)},$$

then we immediately have $\rho(\Lambda, \{U_n\}) = \rho_\Lambda(\Lambda, \{U_n\}) - \alpha_1$ under the assumptions of Theorem A or B.

2.4. **Escape rate on Young's towers with exponential tail.** Young's towers, also known as the Gibbs-Markov-Young structure, is first introduced by Young in [19] and [20]. Young's tower can be viewed as a discrete time suspension over a Gibbs-Markov system $(\tilde{\Omega}, \tilde{T}, \tilde{\mu})$, such that the roof function R (in this case, it is usually call the *return time function*) is integrable with respect to the measure $\tilde{\mu}$.

{m.5}

Theorem G. *Assume that* T *is a* C^2 *map modeled by Young's tower, such that the return time function* R *has exponential tail: there exists* $\lambda \in (0, 1)$ *such that*

$$\tilde{\mu}(R > n) \lesssim \lambda^n.$$

Let $\{U_n \subset \tilde{\Omega}\}$ *be a nested sequence of sets satisfying the assumption of Theorem B in the cylinder case, or Theorem E in the open set case. Then the localized escape rate at* $\Lambda = \cap_n U_n$ *exists and satisfies*

$$\rho(\Lambda, \{U_n\}) = \alpha_1.$$

2.5. **Organization of the paper.** This paper is organized in the following way. In section 3 we then state some properties that surround the parameters of very short returns whose presence is unaffected by the Kac timescale. In section 4 we then prove the main results for cylinder approximations of the zero measure target set Γ. One crucial result here is Lemma 4.6 which yields the extremal index for the near zero time limiting distribution for entry times (as opposed to return times). Those results are then used in section 5 to extend them to the case when the approximating sets are metric neighbourhoods. In section 6 we then provide a general argument which shows that the local escape rate for entry times is the same as the local escape rate for returns. In section 7 we then show that the local escape rate persists for the induced map. Section 8 is dedicated to entertain us with examples.

3. Preliminaries

{s.3}

3.1. **Return and entry times along a nested sequence of sets.** In this section we recall the general results in [14] on the number of entries to an arbitrary null set Λ within a cluster.

Given a sequence of nested sets $U_n, n = 1, 2, \ldots$ with $U_{n+1} \subset U_n$, $\cap_n U_n = \Lambda$ and $\mu(U_n) \to 0$, we will fix a large integer $K > 0$ (which will be sent to infinity later), and assume that the limit

$$\hat{\alpha}_\ell(K) = \lim_{n \to \infty} \mu_{U_n}(\tau_{U_n}^{\ell-1} \leq K)$$

exists for K sufficiently large and for every $\ell \in \mathbb{N}$. By definition $\hat{\alpha}_\ell(K) \geq \hat{\alpha}_{\ell+1}(K)$ for all ℓ, and $\hat{\alpha}_1(K) = 1$ due to our choice of $\tau^0 = 0$ on U. Also note that $\hat{\alpha}_\ell(K)$ is non-decreasing in K for every ℓ. As a result, we have for every $\ell \geq 1$:

(7) $\hat{\alpha}_\ell = \lim_{K \to \infty} \hat{\alpha}_\ell(K)$ exists for every ℓ, and $\hat{\alpha}_1 = 1, \hat{\alpha}_\ell \geq \hat{\alpha}_{\ell+1}$.

Note that in the definition of $\hat{\alpha}$, the cut-off for the short return time K does not depend on the set U_n. Another way to study the short return properties for the nested sequence U_n is to look at

(8) $\hat{\beta}_\ell = \lim_{n \to \infty} \mu_{U_n}(\tau_{U_n}^{\ell-1} \leq s_n)$

for some increasing sequence of integers $\{s_n\}$, with $s_n \mu(U_n) \to 0$ as $n \to \infty$. This is the approach taken by Freitas et al in [12]. It is proven that for many systems (including Gibbs-Markov systems and Young's towers with polynomial tails), we have $\hat{\beta}_\ell = \hat{\alpha}_\ell$. See [18, Proposition 5.4 and 6.2]

To demonstrate the power of desynchronizing K from n, recall that for any set U, the essential periodic of U is given by:

$$\pi_{\text{ess}}(U) = \min\{k > 0 : \mu(T^{-k}U \cap U) > 0\}.$$

Then the following lemma can be easily verified using the definition of $\hat{\alpha}$:

{1.3.1}

Lemma 3.1. *Let $\{U_n\}$ be a sequence of nested sets. Assume that $\pi_{\text{ess}}(U_n) \to \infty$ as $n \to \infty$, then $\hat{\alpha}_\ell$ exists and equals zero for all $\ell \geq 2$.*

Proof. For each K, one can take n_0 large enough such that $\pi_{\text{ess}}(U_n) > K$ for all $n > n_0$. Then for $\ell \geq 2$,

$$\mu_{U_n}(\tau_{U_n}^{\ell-1} \leq K) \leq \mu_{U_n}\left(\bigcup_{k=0}^{K} T^{-k}U_n \cap U_n\right) = 0$$

since all the intersections have zero measure. $\square$

Note that $\hat{\alpha}_\ell(K)$ is the conditional probability to have at least $\ell - 1$ returns in a cluster with length K. If we consider the level set:

(9) $\alpha_\ell(K, U_n) = \mu_{U_n}(\tau_{U_n}^{\ell-1} \leq K < \tau_{U_n}^\ell)$,

and its limit

$$\alpha_\ell(K) = \lim_{n \to \infty} \alpha_\ell(K, U_n),$$

(10) $\alpha_\ell = \lim_{K \to \infty} \alpha_\ell(K)$,

then it is easy to see that

$$\alpha_\ell = \hat{\alpha}_\ell - \hat{\alpha}_{\ell+1}, \text{ and so } \hat{\alpha}_\ell = \sum_{j \geq \ell} \alpha_j$$

which, in particular, implies the existence of α_ℓ.

Next, following [14] we put for every integer $\ell > 0$ and $K > 0$,

$$\lambda_\ell(K, U_n) = \frac{\mathbb{P}(\sum_{i=1}^K \mathbb{I}_{U_n} \circ f^i = \ell)}{\mathbb{P}(\sum_{i=1}^K \mathbb{I}_{U_n} \circ f^i \geq 1)}. \tag{11}$$

In other words, $\lambda_\ell(K, U_n)$ is, conditioned on having an entry to the set U_n, the probability to have precisely ℓ entries in a cluster with length K. The next theorem provides the relation between $\hat{\alpha}_\ell$ and λ_ℓ:

Theorem 3.2. [14, Theorem 2] *Assume that U_n is a sequence of nested sets with* {t.1}
$\mu(U_n) \to 0$. Assume that the limits in (7) exist for K large enough and every $\ell \geq 1$. Also assume that $\sum_{\ell=1}^\infty \ell \hat{\alpha}_\ell < \infty$.
    Then

$$\lambda_\ell = \frac{\alpha_\ell - \alpha_{\ell+1}}{\alpha_1},$$

where $\alpha_\ell = \hat{\alpha}_\ell - \hat{\alpha}_{\ell+1}$. In particular, the limit defining λ_ℓ exists. Moreover, the average length of the cluster of entries satisfies

$$\sum_{\ell=1}^\infty \ell \lambda_\ell = \frac{1}{\alpha_1}.$$

For more properties on $\{\hat{\alpha}_\ell\}$, $\{\alpha_\ell\}$ and $\{\lambda_\ell\}$, we direct the readers to [14] and [18, Section 3].

3.2. **Gibbs-Markov systems.** A map $T : \mathbf{M} \to \mathbf{M}$ is called *Markov* if there is a countable measurable partition $\mathcal{A}$ on $\mathbf{M}$ with $\mu(A) > 0$ for all $A \in \mathcal{A}$, such that for all $A \in \mathcal{A}$, $T(A)$ is injective and can be written as a union of elements in $\mathcal{A}$. Write $\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j} \mathcal{A}$ as before, it is also assumed that $\mathcal{A}$ is (one-sided) generating.
    Fix any $\lambda \in (0, 1)$ and define the metric d_λ on $\mathbf{M}$ by $d_\lambda(x, y) = \lambda^{s(x,y)}$, where $s(x, y)$ is the largest positive integer n such that x, y lie in the same n-cylinder. Define the Jacobian $g = JT^{-1} = \frac{d\mu}{d\mu \circ T}$ and $g_k = g \cdot g \circ T \cdots g \circ T^{k-1}$.
    The map T is called *Gibbs-Markov* if it preserves the measure μ, and also satisfies the following two assumptions:
(i) The big image property: there exists $C > 0$ such that $\mu(T(A)) > C$ for all $A \in \mathcal{A}$.
(ii) Distortion: $\log g|_A$ is Lipschitz for all $A \in \mathcal{A}$.
    In view of (i) and (ii), there exists a constant $D > 1$ such that for all x, y in the same n-cylinder, we have the following distortion bound:

$$\left| \frac{g_n(x)}{g_n(y)} - 1 \right| \leq D d_\lambda(T^n x, T^n y),$$

and the Gibbs property:

$$D^{-1} \leq \frac{\mu(A_n(x))}{g_n(x)} \leq D.$$

It is well known (see, for example, Lemma 2.4(b) in [17]) that Gibbs-Markov systems are exponentially ϕ-mixing, that is, $\phi(k) \lesssim \eta^k$ for some $\eta \in (0, 1)$.[1]

---

[1]Here the notation $\lesssim$ means that the equality holds up to some constant $C > 0$.

4. Escape rate for unions of cylinders

This section contains the proof of Theorem A, B and Corollary C, D. We will suppress the dependence of ρ on $\{U_n\}$ and simply write $\rho(\Lambda)$ for the local escape rate at Λ.

4.1. **The block argument.** In this section, we will provide a general framework on the escape rate for polynomially ϕ-mixing systems. The main lemma, which is Lemma 4.3, allows us to reduce the escape rate (which is on the points that do not enter U in a *large time-scale*) to the probability of having *short entries*.

First we introduce the following standard result for systems that are either left or right ϕ-mixing. The proof can be found in [15, 1].

Lemma 4.1. [15, Lemma 4] *Assume that μ is either left or right ϕ-mixing for the partition $\mathcal{A}$. For $U \in \sigma(\mathcal{A}^{\kappa_n})$, let $s, t > 0$ and $\Delta < \frac{s}{2}$ then we have*

$$|\mathbb{P}(\tau_U > s + t) - \mathbb{P}(\tau_U > s)\mathbb{P}(\tau_U > t)| \leq 2(\Delta\mu(U) + \phi(\Delta - \kappa_n))\mathbb{P}(\tau_U > t - \Delta).$$

Iterating the previous lemma, we obtain:

Lemma 4.2. *Assume that μ is either left or right ϕ-mixing for the partition $\mathcal{A}$. Let $s > 0$ and $\Delta < \frac{s}{2}$. Define $q = \left\lfloor \frac{s}{\Delta} \right\rfloor$, $\eta = \frac{q}{q+1}$, and $\delta = 2(\Delta\mu(U) + \phi(\Delta - \kappa_n))$. Then for every $k \geq 3$ that is an integer multiple of q^{-1}, we have*

$$\mathbb{P}(\tau_U > ks) \leq (\mathbb{P}(\tau_U > s) + \delta^\eta)^{k-2}.$$

Proof. We follow the proof of Theorem 1 in [15] and observe that by the previous lemma, for $k \in [1, 3]$ that is an integer multiple of q^{-1}, we have

$$\mathbb{P}(\tau_U > ks) \leq \mathbb{P}(\tau_U > s)\mathbb{P}(\tau_U > (k-1)s) + \delta\mathbb{P}(\tau_U > (k-1-q^{-1})s)$$
$$\leq \mathbb{P}(\tau_U > s)^2 + \delta\mathbb{P}(\tau_U > s)$$
$$= \mathbb{P}(\tau_U > s)(\mathbb{P}(\tau_U > s) + \delta)$$
$$\leq (\mathbb{P}(\tau_U > s) + \delta^\eta)^{k-2},$$

where we used $\mathbb{P}(\tau_U > s) \leq \mathbb{P}(\tau_U > s) + \delta^\eta$, and $\mathbb{P}(\tau_U > s) \leq 1$.

For $k > 3$, we use induction on $m = k \cdot q \in \mathbb{N}$:

$$\mathbb{P}(\tau_U > ks) \leq \mathbb{P}(\tau_U > s)\mathbb{P}(\tau_U > (k-1)s) + \delta\mathbb{P}(\tau_U > (k-1-q^{-1})s)$$
$$\leq \mathbb{P}(\tau_U > s)(\mathbb{P}(\tau_U > s) + \delta^\eta)^{k-3} + \delta(\mathbb{P}(\tau_U > s) + \delta^\eta)^{k-3-q^{-1}}$$
$$= (\mathbb{P}(\tau_U > s) + \delta^\eta)^{k-3-q^{-1}}[\mathbb{P}(\tau_U > s)(\mathbb{P}(\tau_U > s) + \delta)^{q^{-1}} + \delta]$$
$$\leq (\mathbb{P}(\tau_U > s) + \delta^\eta)^{k-2}.$$

We justify the last inequality as follows. By definition of η, we have $\delta = \delta^\eta\delta^{\frac{\eta}{q}} \leq \delta^\eta(\mathbb{P}(\tau_U > s) + \delta^\eta)^{q^{-1}}$. Consider the bracketed term in the third line.

$$\mathbb{P}(\tau_U > s)(\mathbb{P}(\tau_U > s) + \delta)^{q^{-1}} + \delta$$
$$\leq \mathbb{P}(\tau_U > s)(\mathbb{P}(\tau_U > s) + \delta)^{q^{-1}} + \delta^\eta(\mathbb{P}(\tau_U > s) + \delta^\eta)^{q^{-1}}$$
$$= \mathbb{P}(\tau_U > s) + \delta^\eta)^{1+q^{-1}}.$$

By induction this completes the proof. □

The next lemma establishes the relation between the escape rate and the probability of short entries:

{l.mainlemma}

Lemma 4.3. *Assume that μ is either left or right ϕ-mixing for the partition $\mathcal{A}$, with $\phi(k) \leq Ck^{-p}$ for some $p > 0$. Let $\{U_n \in \sigma(\mathcal{A}^{\kappa_n})\}$ be a nested sequence of sets for some $\kappa_n \nearrow \infty$. Furthermore, assume that there exists $\varepsilon \in (0,1)$, such that $\kappa_n \mu(U_n)^\varepsilon \to 0$.*

Then we have

$$(12) \quad \{\text{e.mainlemma}\} \qquad \rho(\Lambda) = \lim_{n \to \infty} \frac{\mathbb{P}(\tau_{U_n} \leq s_n)}{s_n \mu(U_n)},$$

where $s_n = \lfloor \mu(U_n)^{-(1-a)} \rfloor$ for any fixed $a > 0$ small enough.

Remark 4.4. At first glance, the RHS of (12) is similar to the definition of the local escape rate in (1). However, since $s_n \ll \mu(U_n)^{-1}$ (where the latter is the average return time given by Kac's formula), $\mathbb{P}(\tau_{U_n} \leq s_n)$ concerns the probability of short entries to U.

Proof. Let $\{s_n\}, \{\Delta_n\}$ be increasing sequences of positive integers with $\Delta_n < s_n/2$, whose choice will be specified later. Write $q_n = \lfloor s_n/\Delta_n \rfloor$, $\eta_n = \frac{q_n}{q_n+1}$ and $\delta_n = 2(\Delta_n \mu(U_n) + \phi(\Delta_n - \kappa_n))$ as before. We again follow largely the proof of Theorem 1 in [15] and get by Lemma 4.2

$$\frac{1}{ks_n} |\log \mathbb{P}(\tau_{U_n} > ks_n)| = \frac{k-2}{ks_n} |\log (\mathbb{P}(\tau_{U_n} > s_n) + \delta_n^{\eta_n})|.$$

Take limit as $k \to \infty$ and note that $\mathbb{P}(\tau_{U_n} > s_n) = 1 - \mathbb{P}(\tau_{U_n} \leq s_n)$, we obtain

$$\rho(U_n) = \lim_{k \to \infty} \frac{1}{ks_n} |\log \mathbb{P}(\tau_{U_n} > ks_n)|$$

$$(13) \quad \{\text{e.8}\} \qquad = \frac{1}{s_n} \left(\mathbb{P}(\tau_{U_n} \leq s_n) + o(s_n \mu(U_n)) + \mathcal{O}(\delta_n^{\eta_n}) \right).$$

Here we used the trivial estimate

$$\mathbb{P}(\tau_{U_n} \leq s_n) \leq \mathbb{P}\left(\bigcup_{1 \leq k \leq s_n} T^{-k}(U_n) \right) \leq s_n \mu(U_n).$$

Divide (13) by $\mu(U_n)$ and let $n \to \infty$, we obtain

$$(14) \quad \{\text{e.9}\} \qquad \rho(\Lambda) = \lim_{n \to \infty} \left(\frac{\mathbb{P}(\tau_{U_n} \leq s_n)}{s_n \mu(U_n)} + \frac{\delta_n^{\eta_n}}{s_n \mu(U_n)} \right).$$

It remains to show that the second term converges to zero for some proper choice of $\{s_n\}$ and Δ_n. For this purpose, we fix some $a \in (0,1), b \in (\varepsilon, 1)$ and choose $s_n = \lfloor \mu(U_n)^{-(1-a)} \rfloor$, and $\Delta_n = \lfloor \mu(U_n)^{-b} \rfloor \gg \kappa_n = o(\mu(U_n)^{-\varepsilon})$. Then we have ($a_n \lesssim b_n$ means there exists constant C such that $a_n \leq C \cdot b_n$):

$$\frac{\delta_n^{\eta_n}}{s_n \mu(U_n)} \lesssim \frac{\Delta_n^{\eta_n} \mu(U_n)^{\eta_n}}{s_n \mu(U_n)} + \frac{\phi(\Delta_n - \kappa_n)^{\eta_n}}{s_n \mu(U_n)}$$

$$\leq \Delta_n \mu(U_n)^{\eta_n - a} + \Delta_n^{-p\eta_n} \mu(U_n)^{-a}$$

$$\leq \mu(U_n)^{\eta_n - a - b} + \mu(U_n)^{bp\eta_n - a}.$$

In order for both terms to go to zero, we need:
   (1) $1 - a > b$, which guarantees that $s_n \gg \Delta_n$, so $q_n \to \infty$ and consequently $\eta_n \nearrow 1$; then the first term will go to zero;
   (2) $bp > a$, so that the second term goes to zero.

Both requirements are satisfied if we take any $b \in (\varepsilon, 1)$, then choose $0 < a < \min\{1 - b, bp\}$. Combine this with (14) we conclude that

$$\rho(\Lambda) = \lim_{n \to \infty} \frac{\mathbb{P}(\tau_{U_n} \le s_n)}{s_n \mu(U_n)},$$

as desired. $\square$

In the remaining part of this section, we will prove that the RHS of (12) coincides with the extreme index defined by (4). But before we move on, let us state a direct corollary of the previous lemma, which is interesting in its own right.

{p.coarsebound}

Proposition 4.5. *Assume that μ is either left or right ϕ-mixing for the partition $\mathcal{A}$, with $\phi(k) \le Ck^{-p}$ for some $p > 0$. Let $\{U_n \in \sigma(\mathcal{A}^{\kappa_n})\}$ be a nested sequence of sets for some $\kappa_n \nearrow \infty$. Furthermore, assume that there exists $\varepsilon \in (0, 1)$, such that $\kappa_n \mu(U_n)^\varepsilon \to 0$.*
Then we have

$$(15) \qquad\qquad\qquad\qquad \rho(\Lambda) \in [0, 1],$$

provided that the local escape rate at Λ exists.

Proof. The lower bound is clear. For the upper bound, the trivial estimated which is already used in the proof of Lemma 4.3 yields

$$\frac{\mathbb{P}(\tau_{U_n} \le s_n)}{s_n \mu(U_n)} \le \frac{s_n \mu(U_n)}{s_n \mu(U_n)} = 1.$$

$\square$

{s.4.2}

4.2. **Proof of Theorem A and B.** First we prove Theorem A using the following lemma, which is stated for right ϕ-mixing systems. The proof can be adapted for left ϕ-mixing systems as well, with certain modification on the assumptions of U_n (in particular, on how U_n can be approximated by shorter cylinders). See Remark 4.7 below and the discussion in Section 4.3.

{1.3}

Lemma 4.6. *Let μ be right ϕ-mixing for the partition $\mathcal{A}$, with $\phi(k) \le Ck^{-p}$ for some $p > 1$. Assume that $\{U_n\}$ is a good neighborhood system, such that $\hat{\alpha}_\ell(K)$ exists for K large enough, and $\sum_\ell \hat{\alpha}_\ell < \infty$. Then we have*

$$\lim_{n \to \infty} \frac{\mathbb{P}(\tau_{U_n} \le s_n)}{s_n \mu(U_n)} = \alpha_1$$

for any increasing sequence $\{s_n\}$ for which $s_n \mu(U_n) \to 0$ as $n \to \infty$.

Proof. For an given integer s, write $Z_n^s = \sum_{j=1}^s \mathbb{I}_{U_n} \circ T^j$ which counts the number of entries to U_n before time s. Let K be a large integer, then by [14] Lemma 3 for every $\varepsilon > 0$ one has $\mathbb{P}(\tau_{U_n} \le K) = \alpha_1 K \mu(U_n)(1 + \mathcal{O}^*(\varepsilon))$ for all n large enough, where the notation $\mathcal{O}^*$ means that the implied constant is one (i.e. $x = \mathcal{O}^*(\varepsilon)$ if $|x| < \varepsilon$). For simplicity, assume $r = s_n/K$ is an integer and put

$$V_q = \{Z_n^K \circ T^{qK} \ge 1\},$$

$q = 0, 1, \ldots, r - 1$, and

$$D_q = \{V_q, Z_n^{(r-q-1)K} \circ T^{(q+1)K} = 0\}.$$

Then

$$\{Z_n^{s_n} \geq 1\} = \bigcup_{q=0}^{r-1} D_q$$

is a disjoint union. Let us now estimate

$$
\begin{aligned}
\mathbb{P}(Z_n^{(r-q-1)K} &\circ T^{(q+1)K} \geq 1, V_q) \\
&\leq \mathbb{P}(Z_n^{(r-q-1)K-2\sqrt{K}} \circ T^{(q+1)K+2\sqrt{K}} \geq 1, V_q) + 2\sqrt{K}\mu(U_n) \\
&\leq 2\sqrt{K}\mu(U_n) + \mu(V_q, Z_n^{s_n-(q+1)K-2\kappa_n} \circ T^{(q+1)K+2\kappa_n} \geq 1) \\
&\qquad + \sum_{j=qK}^{(q+1)K-1} \sum_{i=(q+1)K+2\sqrt{K}}^{(q+1)K+2\kappa_n} \mu(T^{-j}U_n \cap T^{-i}U_n)
\end{aligned}
$$

(16)

$$=: \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

To bound II, note that $\{Z_n^{s_n-(q+1)K-2\kappa_n} \circ T^{(q+1)K+2\kappa_n} \geq 1\}$ is the event of having a hit between $[(q+1)K+2\kappa_n, s_n]$. We cut this interval into $t_n = \left\lfloor \frac{s_n-(q+1)K-2\kappa_n}{K} \right\rfloor \geq 0$ (II is void when t_n is negative) many blocks with length K. This allow us to estimate:

$$
\begin{aligned}
\mathrm{II} &\leq \sum_{j=0}^{t_n+1} \mu(Z_n^K \circ T^{qK} \geq 1, Z_n^K \circ T^{(q+1+j)K+2\kappa_n} \geq 1) \\
&= \sum_{j=0}^{t_n+1} \sum_{k=1}^{K} \mu(T^{-qK-k}U_n, Z_n^K \circ T^{(q+1+j)K+2\kappa_n} \geq 1) \\
&\leq \sum_{j=0}^{t_n+1} \sum_{k=1}^{K} \mu(Z_n^K \geq 1)(\mu(U_n) + \phi((j+1)K+\kappa_n-k)) \\
&\leq \sum_{i=\kappa_n}^{s_n+K} \mu(V_q)(\mu(U_n) + \phi(i)),
\end{aligned}
$$

where we used (t_n+1) many blocks instead of t_n to cover the remaining $\leq K$ many hits at the end. The third inequality follows from right ϕ-mixing, and the last line is due to $\mu(Z_n^K \geq 1) = \mu(V_q)$.

For the third term in (16), we use right ϕ-mixing again to get (and recall that U_n^j is the outer-approximation of U_n by j-cylinders):

$$
\begin{aligned}
\text{III} &\leq \sum_{j=qK}^{(q+1)K-1} \sum_{i=(p+1)K+2\sqrt{K}}^{(q+1)K+2\kappa_n} \mu(U_n \cap T^{-(i-j)}U_n) \\
&\leq K \sum_{j=2\sqrt{K}}^{2\kappa_n} \mu(U_n^{j/2} \cap T^{-j}U_n) \\
&\leq K \sum_{j=2\sqrt{K}}^{2\kappa_n} \mu(U_n)(\mu(U_n^{j/2}) + \phi(j/2)) \\
&= \mathcal{O}(1)\mu(V_q) \sum_{j=2\sqrt{K}}^{2\kappa_n} (\mu(U_n^{j/2}) + \phi(j/2)),
\end{aligned}
$$

where the last equality follows from

$$
\mu(V_q) = \mathbb{P}(\tau_{U_n} \leq K) = \alpha_1 K \mu(U_n)(1 + \mathcal{O}^*(\varepsilon)).
$$

Combining the previous estimates, we get

$$
\begin{aligned}
\mathbb{P}(Z_n^{(r-q-1)K} &\circ T^{(q+1)K} \geq 1, V_q) \\
&\leq \ \mathbb{P}(Z_n^{(r-q-1)K-2\sqrt{K}} \circ T^{(q+1)K+2\sqrt{K}} \geq 1, V_q) + 2\sqrt{K}\mu(U_n) \\
&\leq \ 2\sqrt{K}\mu(U_n) + \mu(V_q) \sum_{i=\kappa_n}^{s_n+K} (\mu(U_n) + \phi(i)) \\
&\qquad\qquad\qquad + \mu(V_q)\mathcal{O}(1) \sum_{j=\sqrt{K}}^{\kappa_n} (\mu(U_n^j) + \phi(j)) \\
&\leq \ \mu(V_q)F,
\end{aligned}
$$

where

$$
F = \frac{2}{\sqrt{K}} + (s_n + K)\mu(U_n) + \mathcal{O}(1)(\phi^1(\sqrt{K}) + \sum_{j=\sqrt{K}}^{\kappa_n} \mu(U_n^i))
$$

and $\phi^1(u) = \sum_{i=u}^{\infty} \phi(i)$ is the tail-sum of ϕ which by assumption goes to zero as u goes to infinity.

If n is large enough so that $\max\{s_n\mu(U_n), \kappa_n\mu(U_n), \phi^1(\kappa_n)\} < \varepsilon$ then

$$
\begin{aligned}
F &\leq 2\varepsilon + \frac{2}{\sqrt{K}} + \mathcal{O}(1)\left(\phi^1(\sqrt{K}) + \kappa_n\mu(U_n) + \sum_{i=\sqrt{K}}^{\kappa_n} i^{-p'} \right) \\
&\lesssim \varepsilon + \frac{1}{\sqrt{K}} + \phi^1(\sqrt{K}) + K^{-\frac{p'-1}{2}},
\end{aligned}
$$

where we used the assumption that $\mu(U_n^i) \leq \mu(U_n) + Ci^{-p'}$ for some $p' > 1$. Consequently

$$
\mu(D_q) = \mu(V_q) - \mathbb{P}(V_q, Z_n^{(r-q-1)K} \circ T^{(q+1)K} \geq 1) = \mu(V_q)(1 + \mathcal{O}^*(F)),
$$

and since $\{Z_n^{qK} \geq 1, V_q\} = V_q$ and $\mu(V_q) = \mu(V_0)$ we get

$$\mathbb{P}(Z_n^{s_n} \geq 1) = \sum_{q=0}^{r-1} \mathbb{P}(D_q) = r\mu(V_0)(1 + \mathcal{O}^*(F)).$$

Since by [14] Lemma 3 $\mu(V_0) = \alpha_1 K \mu(U_n)(1 + \mathcal{O}^*(\varepsilon))$ we obtain

$$\mathbb{P}(\tau_{U_n} \leq s_n) = r\mu(V_0)(1 + \mathcal{O}^*(F)) = \alpha_1 s_n \mu(U_n)(1 + \mathcal{O}^*(\varepsilon + F)).$$

The statement of the lemma now follows if we let $\varepsilon \to 0$ and then $K \to \infty$. $\qquad\square$

{rem:bothsides}

Remark 4.7. Similar to the previous lemmas which hold for both left and right ϕ-mixing measures, Lemma 4.6 has a similar formulation in the left ϕ-mixing case. The estimate of II in (16) is mostly the same (see the proof of Lemma 4.9 below for more detail). However, this would require us to modify the definition of the approximated sets U_n^i as

$$\tilde{U}_n^i = T^{-(n-i)} A_i(T^{n-i} U_n),$$

with the assumption that the measure of $\tilde{U}_n^i$ is small (preferably summable in i, similar to (2) in Definition 2). This is indeed the treatment in [15, Lemma 3] when $\Lambda = \{x\}$. However, such an assumption may not hold when Λ is a non-singleton null set. The right ϕ-mixing property avoids this problem.

{r.invertible}

Remark 4.8. So far we have assumed that T is non-invertible. This is because in the invertible case, the approximation U_n^j and $\tilde{U}_n^j$ may become the entire space. As an example, take $\mathbf{M} = \Omega$ to be a full, two-sided shift space and $T = \sigma$ the left-shift. Let the sets U_n be n-approximation of an unstable leaf Γ through a non-periodic point $x \in \Omega$, e.g. $\Gamma = \{y \in \Omega : y_i = x_i \ \forall \ i \leq 0\}$. Obviously Γ is a null set but in this case we get that $\tilde{U}^j = \Omega$ the entire space whenever $i < n/2$. For a geometric example, let T be an Anosov diffeomorphisms on $\mathbb{T}^n$ with minimal unstable foliations and Λ be the local unstable manifold at some $x \in \mathbf{M}$. Then $T^j \Lambda$ eventually becomes ε-dense in $\mathbf{M}$, and the approximation $\tilde{U}_n^i$ (with respect to a Markov partition $\mathcal{A}$) is the entire space for i small. By symmetry and Remark 4.8, we see that if Λ is chosen to be a local stable manifold then $U_n^j = \mathbf{M}$ for j small.

On the other hand, in the proof of Lemma 4.6 the approximation U_n^j is only used to control III of (16). Later this observation will allow us to obtain a result for invertible systems where this term does not appear. See Theorem 4.12 and 5.2 below.

Below we state an alternate version of Lemma 4.6 where the right ϕ-mixing assumption is replaced by the Gibbs-Markov property. This allows us to bypass the issue stated in Remark 4.7 and keep the choice of U_n^i.

{1.3'}

Lemma 4.9. *Let $(T, \mu, \mathcal{A})$ be a Gibbs-Markov system. Assume that $\{U_n\}$ is a good neighborhood system, such that $\hat{\alpha}_\ell(K)$ exists for K large enough, and $\sum_\ell \hat{\alpha}_\ell < \infty$. Then we have*

$$\lim_{n \to \infty} \frac{\mathbb{P}(\tau_{U_n} \leq s_n)}{s_n \mu(U_n)} = \alpha_1$$

for any increasing sequence $\{s_n\}$ for which $s_n \mu(U_n) \to 0$ as $n \to \infty$.

Proof. Recall that Gibbs-Markov systems are left ϕ-mixing with exponential rate. The proof follows the lines of Lemma 4.6 up to equation (16), which is now estimated using the left ϕ-mixing as:

$$\text{II} = \mu(V_q, Z_n^{s_n-(q+1)K-2\kappa_n} \circ T^{(q+1)K+2\kappa_n} \geq 1)$$

$$\leq \sum_{i=(q+1)K+2\kappa_n}^{s_n} \mu(V_q \cap T^{-i}U_n)$$

$$\leq \sum_{i=\kappa_n}^{s_n} \mu(V_q)(\mu(U_n) + \phi(i)).$$

Note that the proof in this case is much short and the bound is almost the same as before.

For III, we first split the term into the summation over the intersections of U_n with $T^i U_n$:

$$\text{III} \leq \sum_{j=qK}^{(q+1)K-1} \sum_{i=(p+1)K+2\sqrt{K}}^{(q+1)K+2\kappa_n} \mu(U_n \cap T^{-(i-j)}U_n)$$

$$\leq K \sum_{j=2\sqrt{K}}^{2\kappa_n} \mu(U_n \cap T^{-j}U_n).$$

Each term in the summation can be bounded by:

$$\mu(U_n \cap T^{-j}U_n) \leq \sum_{A \in \mathcal{C}_j(U_n)} \mu(T^{-j}U_n \cap A)$$

$$= \sum_{A \in \mathcal{C}_j(U_n)} \frac{\mu(T^{-j}U_n \cap A)}{\mu(A)}\mu(A)$$

$$\lesssim \sum_{A \in \mathcal{C}_j(U_n)} \frac{\mu(T^j(T^{-j}U_n \cap A))}{\mu(T^j A)}\mu(A)$$

$$\lesssim \sum_{A \in \mathcal{C}_j(U_n)} \mu(U_n)\mu(A)$$

$$= \mu(U_n)\mu\left(\bigcup_{A \in \mathcal{C}_j(U_n)} A\right) = \mu(U_n)\mu(U_n^j),$$

where the third and forth inequality follow from the distortion and the big image property of Gibbs-Markov systems. See [18, Theorem D].

Then

$$\text{III} \leq c_1 K \mu(U_n) \sum_{j=2\sqrt{K}}^{\kappa_n} \mu(U_n^j) = \mathcal{O}(1)\mu(V_p) \sum_{j=2\sqrt{K}}^{2\kappa_n} \mu(U_n^j),$$

for some c_1 and the rest of the proof is identical to Lemma 4.6. $\square$

Now Theorem A and B are immediate consequences of Lemma 4.3, 4.6 and 4.9.

Proof of Corollary C. This corollary directly follows from Lemma 3.1. $\square$

Proof of Corollary D. We need the following proposition from [18]:

Proposition 4.10. [18, Proposition 6.3] *Let T be a continuous map on the compact metric space $\mathbf{M}$, and $\{U_n\}$ a nested sequence of sets such that $\cap_n U_n = \cap_n \overline{U}_n$. Then $\pi(U_n) \to \infty$ if and only if $\Lambda = \cap_n U_n$ intersects every forward orbit at most once.*

Since $\pi_{\mathrm{ess}}(U) \geq \pi(U)$, we have $\pi_{\mathrm{ess}}(U_n) \to \infty$. Combined with Corollary C, we obtain Corollary D. □

{s.4.3}

4.3. **Some remarks on the extremal index.** In the classic literature (for example, [10], [8] and [12]), the extremal index is defined as

$$(17) \quad \{\texttt{e.extremal}\} \qquad \theta = \lim_{n \to \infty} \mu_{U_n}(\tau_{U_n} > K_n),$$

where $K_n \to \infty$ is some increasing sequence of integers. It is shown in [18, Proposition 5.4] that under the assumption of Theorem B we have

$$\alpha_1 = \theta.$$

It is also straight forward to check that the proof of Lemma 4.6 and 4.9 remain true with α_1 replaced by θ. We state this as the following proposition:

{p.4.9}

Proposition 4.11. *Assume that one of the following assumptions holds:*

   (1) either μ is right ϕ-mixing with $\phi(k) \lesssim k^{-p}$, $p > 1$;
   (2) or $(T, \mu, \mathcal{A})$ is a Gibbs-Markov system.

Let θ be the extremal index defined by (17) for some sequence $\{K_n\}$. Then for any good neighborhood system $\{U_n\}$ and any increasing sequence $\{s_n\}$ with $s_n \mu(U_n) \to 0$ and $s_n/K_n \to \infty$, we have

$$\lim_{n \to \infty} \frac{\mathbb{P}(\tau_{U_n} \leq s_n)}{s_n \mu(U_n)} = \theta.$$

Furthermore, the local escape rate at $\Lambda = \bigcap_n U_n$ exists and satisfies

$$\rho(\Lambda) = \theta.$$

Note that in the proof of Lemma 4.6, the bound on II of (16) holds for both left and right ϕ-mixing systems, as already demonstrated in Lemma 4.9. On the other hand, for θ defined by (17), III of (16) does not exist when $K_n > \kappa_n^2$. By Remark 4.7 and 4.8, we can drop the right ϕ-mixing and the Gibbs-Markov assumption, obtaining the following theorem for left ϕ-mixing systems that is either invertible or non-invertible:

{t.leftmixing}

Theorem 4.12. *Assume that $T : \mathbf{M} \to \mathbf{M}$ is a dynamical system, either invertible or non-invertible, and preserves a measure μ that is left ϕ-mixing with $\phi(k) \leq Ck^p$ for some $C > 0$ and $p > 1$. Let $\{U_n \in \mathcal{A}^{\kappa_n}\}$ be a nested sequence of sets with $\kappa_n \mu(U_n)^\varepsilon \to 0$ for some $\varepsilon \in (0, 1)$.*

Assume that θ defined by (17) exists for some sequence $\{K_n\}$ with $K_n > \kappa_n^2$. Then the localized escape rate at Λ exists and satisfies

$$\rho(\Lambda) = \theta.$$

5. ESCAPE RATE FOR OPEN SETS: AN APPROXIMATION ARGUMENT

{s.5}

First, observe that

$$\{M_t < u_n\} = \{\tau_{U_n} > t\}.$$

As a result, we have

$$\zeta(u_n) = \lim_{t \to \infty} \frac{1}{t} |\log \mathbb{P}(M_t < u_n)| = \lim_{t \to \infty} \frac{1}{t} |\log \mathbb{P}(\tau_{U_n} > t)| = \rho(U_n),$$

therefore we have

$$\zeta(\varphi, \{u_n\}) = \rho(\Lambda, \{U_n\}).$$

The following proposition allows us to replace $\{U_n\}$ by its cylinder-approximation.

{p.approx}

Proposition 5.1. *Let $\{U_n\}$, $\{V_n\}$ and $\{W_n\}$ be sequences of nested sets with $V_n \subset U_n \subset W_n$ for each n, and $\Lambda = \bigcap_n U_n = \bigcap_n V_n = \cap_n W_n$. Assume that*

(18) $$\mu(W_n \setminus V_n) = o(1)\mu(V_n),$$

and $\rho(\Lambda, \{W_n\}) = \rho(\Lambda, \{V_n\}) = \alpha$.
Then

$$\rho(\Lambda, \{U_n\}) = \alpha.$$

Proof. $V_n \subset U_n \subset W_n$ implies that $\tau_{W_n} \geq \tau_{V_n} \geq \tau_{U_n}$ and therefore

$$\rho(W_n) \geq \rho(U_n) \geq \rho(V_n).$$

On the other hand, (18) means that $\mu(W_n)/\mu(V_n) \to 1$. We thus obtain

$$\rho(\Lambda, \{W_n\}) \geq \rho(\Lambda, \{U_n\}) \geq \rho(\Lambda, \{V_n\}),$$

and the proposition follows from the squeeze theorem. $\qquad\square$

Proof of Theorem E. For the sequence $\{r_n\}$ given in Assumption 1, we write κ_n for the smallest integer such that diam $\mathcal{A}^{\kappa_n} \leq r_n$. Then consider

$$V_n = \cup_{A \in \mathcal{A}^{\kappa_n}, A \subset U_n} A, \qquad W_n = \cup_{A \in \mathcal{A}^{\kappa_n}, A \cap U_n \neq \emptyset} A.$$

Clearly we have $V_n \subset U_n \subset W_n$ for each n. Moreover, the choice of κ_n gives

$$U_n^i \subset V_n, \qquad W_n \subset U_n^o.$$

Combine this with (6), we have $\mu(W_n \setminus V_n) = o(1)\mu(V_n)$.

Let us write $\hat{\alpha}_\ell^*$, $* = U, V, W$ for $\hat{\alpha}_\ell$ defined using $\{U_n\}, \{V_n\}, \{W_n\}$, respectively. Then it is proven in [18, Lemma 5.6] that

$$\hat{\alpha}_\ell^V = \hat{\alpha}_\ell^U = \hat{\alpha}_\ell^W.$$

In particular, $\sum_\ell \ell \hat{\alpha}_\ell^U < \infty$ implies that the same holds for $\hat{\alpha}_\ell^*$, $* = V, W$, and the value of α_1 defined by $\{V_n\}, \{U_n\}, \{W_n\}$ are equal.

It remains to show that $\{V_n\}$ and $\{W_n\}$ are good neighborhood systems. (1) of Definition 2 holds due to (a) in Theorem E. For (2) of Definition 2, observe that

$$\mu(V_n^j) = \mu\left(\bigcup_{A \in \mathcal{C}_j(V_n)} A\right) \leq \mu(V_n) + \mu\left(\bigcup_{A \in \mathcal{A}^j, A \cap B_{r_n}(\partial U_n) \neq \emptyset} A\right) \leq \mu(V_n) + Cj^{-p'},$$

thanks to (b) in Theorem E. A similar argument shows that $\{W_n\}$ is also a good neighborhood system.

Now we can apply Theorem A or B on $\{V_n\}$ and $\{W_n\}$ to obtain

$$\rho(\Lambda, \{W_n\}) = \rho(\Lambda, \{V_n\}) = \alpha_1.$$

It then follows from Proposition 5.1 that $\rho(\Lambda, \{U_n\}) = \alpha_1$. This concludes the proof of Theorem E. $\qquad\square$

Similar to Theorem 4.12, when the extremal index θ is defined as

$$\theta = \lim_{n \to \infty} \mu_{U_n}(\tau_{U_n} > K_n)$$

for some sequence $K_n > \kappa_n^2$, the conditions on the right ϕ-mixing and V_n^j can be dropped. We thus obtain the following version of Theorem 4.12 for open sets $\{U_n\}$:

{t.leftmixing.open}

Theorem 5.2. *Assume that $T : \mathbf{M} \to \mathbf{M}$ is a dynamical system, either invertible or non-invertible, and preserves a measure μ that is left ϕ-mixing with $\phi(k) \le Ck^p$ for some $C > 0$ and $p > 1$.*

Let $\varphi : \mathbf{M} \to \mathbb{R} \cup \{+\infty\}$ be a continuous function achieving its maximum on a measure zero set Λ. Let $\{u_n\}$ be a non-decreasing sequence of real numbers with $u_n \nearrow \sup \varphi$, such that the open sets U_n defined by (5) satisfy Assumption 1 for a sequence r_n that decreases to 0 as $n \to \infty$. Let κ_n be the smallest positive integer for which $\operatorname{diam} \mathcal{A}^{\kappa_n} \le r_n$ and assume that:

(i) $\kappa_n \mu(U_n)^\varepsilon \to 0$ for some $\varepsilon \in (0,1)$;
(ii) the extremal index θ defined by (17) exists for some sequence $K_n > \kappa_n^2$.

Then the exceedance rate of φ along $\{u_n\}$ exists and satisfies

$$\zeta(\varphi, \{u_n\}) = \rho(\Lambda, \{U_n\}) = \theta.$$

6. THE CONDITIONAL ESCAPE RATE

{s.6}

In this section we will prove Theorem F.

First we establish the following relation between the hitting times and return times.

Lemma 6.1. *For any set $U \subset M$ with $\mu(U) > 0$, let $A_k := \{x \in \mathbf{M} | \tau_U \ge k\}$, and $B_k := \{x \in U | \tau_U \ge k\} = A_k \cap U$. Then we have*

(19) $$\mu_U(A_k)\mu(U) = \mu(B_k) = \mu(A_k) - \mu(A_{k+1})$$

Proof. By definition we have $A_{k+1} \subset A_k$. Thus, we compute

$$\begin{aligned}
\mu(A_{k+1}) &= \mu(\cap_{j=1}^k T^{-j}U^c) \\
&= \mu(T^{-1}(\cap_{j=0}^{k-1} T^{-j}U^c)) \\
&= \mu(U^c \cap_{j=1}^{k-1} T^{-j}U^c) \\
&= \mu(U^c \cap A_k) \\
&= \mu(A_k) - \mu(U \cap A_k) \\
&= \mu(A_k) - \mu(B_k),
\end{aligned}$$

where the third equality follows from the invariance of μ. $\square$

Next, we need the following arithmetic lemma on the exponential decay rate for a sequence of real numbers $\{a_n\}$ and its difference sequence $\{b_n = a_n - a_{n+1}\}$.

Lemma 6.2. *Suppose that $\{a_n\}$ is a decreasing sequence of positive real numbers with $a_n \searrow 0$. Let $b_n = a_n - a_{n+1}$. Suppose, also, that b_n is monotonically decreasing. Then the following statements are equivalent:*

(1) $\lim_{n \to \infty} -\frac{\log a_n}{n} = \vartheta$ for some $\vartheta > 0$;
(2) $\lim_{n \to \infty} -\frac{\log b_n}{n} = \vartheta$ for some $\vartheta > 0$.

Remark 6.3. Note that there are counter-examples for which the statement of the lemma fails without the monotonicity assumption on the sequence $\{b_n\}$.

Proof. First note that $a_n = \sum_{k \geq n} b_k$; therefore (2) $\implies$ (1) is obvious. It thus remains to show that (1) $\implies$ (2). Let $\gamma > 1$ be fixed. Since by assumption the limit $\lim_{n \to \infty} \frac{1}{n} |\log a_n| = \vartheta$ exists and $\vartheta > 0$, there is an N so that

$$\left| \frac{\log a_n}{n} + \vartheta \right| < \varepsilon \quad \forall n \geq N,$$

for some positive $\varepsilon < (\gamma - 1)\vartheta/4$. Hence

$$\left| \frac{\log a_n}{n} - \frac{\log a_{\gamma n}}{\gamma n} \right| < 2\varepsilon \quad \forall n \geq N$$

which implies

$$a_{\gamma n} < \left(a_n e^{2\varepsilon n} \right)^\gamma < \frac{1}{2} a_n$$

for all n large enough (assuming $\varepsilon < \frac{\vartheta}{2}$). Since

$$a_n - a_{\gamma n} = \sum_{j=n}^{\gamma n - 1} b_j$$

we get by monotonicity of the sequence b_j

$$b_{\gamma n}(\gamma - 1)n \leq a_n - a_{\gamma n} \leq b_n(\gamma - 1)n$$

and consequently $a_n \leq 2 b_n(\gamma - 1)n$. Hence

$$\frac{\log b_{\gamma n}}{n} + \frac{\log(\gamma - 1)n}{n} \leq \frac{\log a_n}{n} \leq \frac{\log b_n}{n} + \frac{\log 2(\gamma - 1)n}{n}$$

which in the limit $n \to \infty$ yields

$$\limsup_{n \to \infty} \frac{\log b_n}{n} \leq -\frac{\vartheta}{\gamma}$$

and

$$-\vartheta \leq \liminf_{n \to \infty} \frac{\log b_n}{n}.$$

Since this applies to every $\gamma > 1$ we obtain that $\lim_{n \to \infty} \frac{1}{n} \log b_n = -\vartheta$. $\qquad\square$

Proof of Theorem F. The theorem now follows with $a_k = \mu(A_k) = \mathbb{P}(\tau_U \geq k)$, and $b_k = \mu(U)\mu(B_k) = \mu(U)\mathbb{P}_U(\tau_U \geq k)$.

$\qquad\square$

7. Escape rate under inducing

{s.7}

In this section, we will state a general theorem for the local escape rate under inducing. For this purpose, we consider a measure preserving dynamical system $(\tilde{\Omega}, \tilde{T}, \tilde{\mu})$ with $\tilde{\mu}$ being a probability measure. Given a measurable function $R : \tilde{\Omega} \to \mathbb{Z}^+$ consider the space $\Omega = \tilde{\Omega} \times \mathbb{Z}^+ / \sim$ with the equivalence relation $\sim$ given by

$$(x, R(x)) \sim (\tilde{T}(x), 0).$$

Define *the (discrete-time) suspension map over $\tilde{\Omega}$ with roof function R* as the measurable map T on the space Ω acting by

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

We will call Ω a *tower over* $\tilde{\Omega}$ and refer to the set $\Omega_k := \{(x,k) : x \in \tilde{\Omega}, k < R(x)\}$ as the *kth floor* where $\tilde{\Omega}$ can be naturally identified with the 0th floor called *the base of the tower.*

For $0 \le k < i$, set $\Omega_{k,i} = \{(x,k) : R(x) = i\}$. The map

$$\Pi : (x,k) \mapsto x$$

is naturally viewed as a projection from the tower Ω to the base $\tilde{\Omega}$ and for any given set $U \subset \Omega$ we will write

$$\tilde{U} = \Pi(U).$$

The measure $\tilde{\mu}$ can be lifted to a measure $\hat{\mu}$ on Ω by

$$\hat{\mu}(A) = \sum_{i=1}^{\infty} \sum_{k=0}^{i-1} \tilde{\mu}(\Pi(A \cap \Omega_{k,i})).$$

It is easy to verify that $\hat{\mu}$ is T-invariant and if $\tilde{\mu}(R) = \int R\,d\tilde{\mu} < \infty$ then $\hat{\mu}$ is a finite measure. In this case, the measure

$$\mu = \frac{\hat{\mu}}{\tilde{\mu}(R)}$$

is a T-invariant probability measure on Ω.

We write $\tilde{U} = \Pi(U) \subset \tilde{\Omega}$, $\tilde{\Lambda} = \cap_n \tilde{U}_n$ and define $\tilde{\rho}(\tilde{\Lambda}, \{\tilde{U}_n\})$ to be the localized escape rate at $\tilde{\Lambda}$ for the system $(\tilde{\Omega}, \tilde{T}, \tilde{\mu})$. The following theorem relates the escaped rate of the base system with that of the suspension. A similar result is obtained for continuous suspensions under the assumption that R is bounded, see [7].

{t.suspension}

Theorem 7.1. *Let (Ω, T, μ) be a discrete-time suspension over an ergodic measure preserving system $(\tilde{\Omega}, \tilde{T}, \tilde{\mu})$ with a roof function R satisfying the following assumptions:*

(1) R *has exponential tail: there exists $C, c > 0$ such that $\tilde{\mu}(R > n) \le Ce^{-cn}$;*
(2) *exponential large deviation estimate: for every $\varepsilon > 0$ small, there exists $C_\varepsilon, c_\varepsilon > 0$ such that the set*

$$B_{\varepsilon,k} = \left\{ y \in \tilde{\Omega} : \left| \frac{1}{n} \sum_{j=0}^{n-1} R(\tilde{T}^j y_0) - \frac{1}{\mu(\Omega_0)} \right| > \varepsilon \text{ for some } n \ge k \right\},$$

satisfies $\tilde{\mu}(B_{\varepsilon,k}) \le C_\varepsilon e^{-c_\varepsilon k}$.

Then for every nested sequence $\{U_n\}$, we have

$$\rho(\Lambda, \{U_n\}) = \tilde{\rho}(\tilde{\Lambda}, \{\tilde{U}_n\}).$$

Proof. The result of this theorem is in fact hidden in the proof of Theorem 4 of [15] and Theorem 3.2 (1) in [2]. We include the proof here for completeness.

For every $y = (x,m) \in \Omega$, we take $y_0 = x \in \tilde{\Omega}$. Then we have

(20)   {{e.0}}
$$\tau_{U_n}(y) = -m + \sum_{j=0}^{\tilde{\tau}_{\tilde{U}_n}(y_0)-1} R(\tilde{T}^j(y_0)),$$

where $\tilde{\tau}$ is the return times defined for the system $(\tilde{\Omega}, \tilde{T}, \tilde{\mu})$. By the Birkhoff ergodic theorem on $(\tilde{\Omega}, \tilde{T}, \tilde{\mu})$, we see that

$$\frac{1}{n} \sum_{j=0}^{n-1} R(\tilde{T}^j y_0) \to \int_{\tilde{\Omega}} R(y) \, d\tilde{\mu}(y) = \frac{1}{\mu(\Omega_0)},$$

where we apply the Kac's formula on the last equality and use the fact that μ is the lift of $\tilde{\mu}$.

On the other hand, since the return time function R has exponential tail, we get, for each $\varepsilon > 0$ and t large enough,

$$\mu((x,m) : m > \varepsilon t) \lesssim e^{-c\varepsilon t}.$$

To simplify notation, we introduce the set (n is fixed)

$$A_t = \left\{ y = (x,m) : m < \varepsilon t, \sum_{j=0}^{\tilde{\tau}_{\tilde{U}_n}(y_0)-1} R(\tilde{T}^j(y_0)) > (1+\varepsilon)t \right\} \cap B_{\varepsilon,k}^c.$$

Combine (20) with the previous estimates on $B_{\varepsilon,k}$, for $k = t(1+\varepsilon)$ we get

(21)
$$\left| \mu(\tau_{\tilde{U}_n} > t) - \mu(A_t) \right| \lesssim e^{-c\varepsilon t} + e^{-c_\varepsilon(1+\varepsilon)t}.$$

Note that A_t contains the set

$$A_t^- = \left\{ y : m < \varepsilon t, \tilde{\tau}_{\tilde{U}_n}(y_0) > \frac{(1+\varepsilon)t}{\mu^{-1}(\Omega_0) - \varepsilon} \right\},$$

and is contained in

$$A_t^+ = \left\{ y : m < \varepsilon t, \tilde{\tau}_{\tilde{U}_n}(y_0) > \frac{(1+\varepsilon)t}{\mu^{-1}(\Omega_0) + \varepsilon} \right\}.$$

Now we are left to estimate $\mu(A_t^\pm)$. Since μ is the lift of $\tilde{\mu}$, we have

(22)
$$\mu(A_t^\pm) = \frac{1}{\tilde{\mu}(R)} \sum_{j=0}^{\infty} \sum_{i=0}^{\min(\varepsilon t, R_j)} \tilde{\mu}(T^{-i} A_t^\pm \cap \Omega_{0,i})$$

$$= \mu(\Omega_0)(1 + \mathcal{O}(\varepsilon t))\tilde{\mu}(\tilde{A}_t^\pm),$$

where

$$\tilde{A}_t^\pm = \left\{ y_0 \in \Omega_0 : \tilde{\tau}_{\tilde{U}_n}(y_0) > \frac{(1+\varepsilon)t}{\mu^{-1}(\Omega_0) \pm \varepsilon} \right\}.$$

Let $\alpha = \tilde{\rho}(\tilde{\Lambda}, \{\tilde{U}_n\})$. Then we have (recall that $\tilde{\mu}(\tilde{U}_n)\mu(\Omega_0) = \mu(U_n)$)

$$\lim_{n\to\infty} \lim_{t\to\infty} \frac{1}{t\mu(U_n)} |\log \tilde{\mu}(\tilde{A}_t^\pm)| = \alpha \frac{(1+\varepsilon)}{1 \pm \varepsilon\mu(\Omega_0)}.$$

By (22), we get that

$$\lim_{n\to\infty} \lim_{t\to\infty} \frac{1}{t\mu(U_n)} |\log \mu(\tilde{A}_t^\pm)| = \alpha \frac{(1+\varepsilon)}{1 \pm \varepsilon\mu(\Omega_0)}.$$

For each $\varepsilon > 0$ we can take n_0 large enough, such that for $n > n_0$:

$$\alpha \frac{1+\varepsilon}{1 \pm \varepsilon\mu(\Omega_0)}\mu(U_n) < \min\{c\varepsilon, c_\varepsilon(1+\varepsilon)\}.$$

It then follows that the right-hand-side of (21) is of order $o(\mu(A_t^\pm))$. We thus obtain

$$\rho(\Lambda, \{U_n\}) = \lim_{n\to\infty} \lim_{t\to\infty} \frac{1}{t\mu(U_n)} |\log \mu(\tau_{U_n} > t)| \in \left(\alpha \frac{(1+\varepsilon)}{1+\varepsilon\mu(\Omega_0)}, \alpha \frac{(1+\varepsilon)}{1-\varepsilon\mu(\Omega_0)} \right)$$

for every $\varepsilon > 0$. This shows that $\rho(\Lambda, \{U_n\}) = \alpha = \tilde{\rho}(\tilde{\Lambda}, \{\tilde{U}_n\})$.   □

Proof of Theorem G. Young towers can be seen as discrete-time suspension over Gibbs-Markov maps. Moreover, the exponential tail of $\mu(R > n)$ implies the exponential large deviation estimate (see for example [2] Appendix B). Therefore Theorem G immediately follows from Theorem B, Theorem E and Theorem 7.1.   □

8. Examples

{s.8}

8.1. **Periodic and non-periodic points dichotomy.** First we consider the case where $\Lambda = \{x\}$ is a singleton, and $U_n = B_{\delta_n}(x)$ is a sequence of balls shrinking to x. Alternatively one could take $\varphi(y) = g(d(y,x))$ for some function $g(x) : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ achieving its maximum at 0 (for example, $g(y) = -\log y$) and let $u_n \nearrow \infty$ be a sequence of threshold tending to infinity. Then $U_n = \{y : \varphi(y) > u_n\}$ is a sequence of balls with diameter shrinking to zero.

This situation has been dealt with in [2] for certain interval maps, and in [15] for maps that are polynomially ϕ-mixing. A dichotomy is obtained: when x is non-periodic the local escape rate is 1; when x is periodic then $\rho(x) = 1 - \theta$ where

$$(23) \quad \{\{\text{e.theta}\}\} \qquad \theta = \theta(x) = \lim_{n\to\infty} \frac{\mu(U_n \cap T^{-m} U_n)}{\mu(U_n)},$$

where m is the periodic of x. When μ is an equilibrium state for some potential function $h(x)$ with zero pressure, one has $\theta = e^{S_m h(x)}$ where S_m is the Birkhoff sum. See [2].

Note that if x is non-periodic then one naturally deduces that $\pi(U_n) \nearrow \infty$ (see for example [15, Lemma 1]). When x is periodic, in [14, Section 8.3] it is shown that $\hat{\alpha}_\ell = \theta^{l-1}$ is a geometric distribution. In particular one has $\sum_\ell \ell\hat{\alpha}_\ell < \infty$ and $\alpha_1 = 1 - \theta$. This leads to the following theorem:

Theorem 8.1. *Assume that*

(1) *either μ is right ϕ-mixing with $\phi(k) \le Ck^{-p}$, $p > 1$;*
(2) *or $(T, \mu, \mathcal{A})$ is a Gibbs-Markov system.*

Assume that $0 < r_n < \delta_n$ satisfies

$$\mu(B_{\delta_n + r_n}(x) \setminus B_{\delta_n - r_n}(x)) = o(1)\mu(B_{\delta_n}(x)).$$

Write κ_n for the smallest positive integer with $\operatorname{diam} \mathcal{A}^{\kappa_n} \le r_n$. We assume that:

(a) *$\kappa_n \mu(U_n)^\varepsilon \to 0$ for some $\varepsilon \in (0,1)$;*
(b) *U_n has small boundary: there exists $C > 0$ and $p' > 1$, such that*
$$\mu\left(\bigcup_{A \in \mathcal{A}^j, A \cap B_{r_n}(\partial U_n) \ne \emptyset} A \right) \le Cj^{-p'} \text{ for all } n \text{ and } j \le \kappa_n.$$
(c) *when x is periodic with period m, θ defined by (23) exists.*

Then we have

$$\rho(\{x\}, \{B_{\delta_n}(x)\}) = \alpha_1 = \begin{cases} 1 & \text{if } x \text{ is non-periodic} \\ 1 - \theta & \text{if } x \text{ is periodic} \end{cases}.$$

This theorem improves [15, Theorem 2] by dropping the assumption $\theta < 1/2$. Also note that such results can be generalized to interval maps which can be modeled by Young's towers using Theorem G.

8.2. **Cantor sets for interval expanding maps.** For simplicity, below we will only consider the Cantor ternary set. However the argument below can be adapted to a large family of dynamically-defined Cantor set discussed in [12] with only minor modification.

Consider the uniformly expanding map $T(x) = 3x \mod 1$ defined on the unit interval $[0, 1]$. We take Λ to be the ternary Cantor set on $[0, 1]$, and define recursively: $U_0 = [0, 1]$; U_{n+1} is obtained by removing the middle third of each connected component of U_n. Then we have $\cap_n U_n = \Lambda$.

Theorem 8.2. *For the uniformly expanding map $T(x) = 3x \mod 1$ on $[0, 1]$, the Cantor ternary set Λ and the nested sets $\{U_n\}$, we have*

$$\rho(\Lambda, \{U_n\}) = \frac{1}{3}.$$

Proof. Let $\mathcal{A} = \{[0, 1/3), [1/3, 2/3), [2/3, 1]\}$ be a Markov partition of T, with respect to which the Lebesgue measure μ is exponentially ψ-mixing. Below we will verify the assumptions of Proposition 4.11.

It is easy to see that $U_n \in \mathcal{A}^n$, i.e., $\kappa_n = n$. On the other hand, $\mu(U_n) = 2^n/3^n$ which shows that item (1) of Definition 2 is satisfied for any $\varepsilon \in (0, 1)$. For item (2), note that $U_n^j = U_j$ which implies that

$$\mu(U_n^j) \le \mu(U_n) + \mu(U_j) = \mu(U_n) + \left(\frac{2}{3}\right)^j.$$

We conclude that $\{U_n\}$ is a good neighborhood system.

The extremal index can be found as follows. Let us write $U_n = \bigcup_{|\alpha|=n} J_\alpha$ where the disjoint union is over all n-words $\alpha = \alpha_1 \alpha_2 \ldots \alpha_n \in \{0, 2\}^n$ and

$$J_\alpha = \sum_{j=1}^n \frac{\alpha_j}{3^j} + 3^{-n} I,$$

where $I = [0, 1]$ is the unit interval. The length $|J_\alpha|$ is equal to 3^{-n}. For $j < n$

$$T^{-j} J_\alpha = \bigcup_{\beta \in \{0,1,2\}^j} J_{\beta\alpha}$$

(disjoint union), thus

$$U_n \cap T^{-j} U_n = \bigcup_{\alpha \in \{0,2\}^n} \bigcup_{\beta \in \{0,2\}^j} J_{\beta\alpha}$$

and therefore $U_n \cap T^{-j} U_n = U_{n+j}$. Consequently

$$\{\tau_{U_n} \le K\} \cap U_n = U_n \cap \bigcup_{j=1}^K T^{-j} U_n = U_{n+1}.$$

Since $\mu(U_{n+j}) = \left(\frac{2}{3}\right)^j \mu(U_n)$ this implies that $\hat{\alpha}_2(K, U_n) = \frac{\mu(U_{n+1})}{\mu(U_n)} = \frac{2}{3}$ and therefore $\alpha_1 = \frac{1}{3}$.

This result was recently shown in [12, Theorem 3.3] in more generality. By Proposition 4.11 we conclude that $\rho(\Lambda, \{U_n\}) = 1/3$. $\qquad\square$

{s.8.2}

8.3. **Submanifolds of Anosov maps.** In this section we consider the case where Λ is a submanifold for some Anosov map T. More importantly, we will show how our results can be applied to those cases where the extremal index θ is defined using time cut-off K_n that depends on U_n (see (17)).

Let $T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ be an Anosov system on $\mathbb{T}^2$ and μ be the Lebesgue measure. It is well known that μ is exponentially ψ-mixing with respect to its Markov partition $\mathcal{A}$. Also denote by $\lambda > 1$ the eigenvalue of T. Following [3] we take Λ to be a line segment with finite length $l(\Lambda)$. We will lift Λ to $\hat{\Lambda} \subset \mathbb{R}^2$ and parametrize $\hat{\Lambda}$ by $p_1 + tv$ for some $p_1 \in \mathbb{R}^2$ and $t \in [0, l(\Lambda)]$. Write p_2 for the other end point of $\hat{\Lambda}$, that is, $p_2 = p_1 + l(\Lambda)v$.

Consider the function $\varphi_\Lambda(y) = -\log d(x, \Lambda)$ which achieves its maximum $(+\infty)$ on Λ. Write $v^*, * = s, u$ for the unit vector along the stable and unstable direction respectively. Then we have:

Theorem 8.3. *For the sequence $\{u_n = \log n\}$,*

(1) *if Λ is not aligned with the stable direction v^s or the unstable direction v^u then $\zeta(\varphi_\Lambda, \{u_n\}) = 1$;*

(2) *if Λ is aligned with the unstable direction but $\{p_1 + tv^u, t \in \mathbb{R}\}$ has no periodic point, then $\zeta(\varphi_\Lambda, \{u_n\}) = 1$;*

(3) *if Λ is aligned with the stable direction but $\{p_1 + tv^s, t \in \mathbb{R}\}$ has no periodic point, then $\zeta(\varphi_\Lambda, \{u_n\}) = 1$;*

(4) *Λ is aligned with $v^*, * = s, u$ and L contains a periodic point with prime period q, then $\zeta(\varphi_\Lambda, \{u_n\}) = 1 - \lambda^{-q}$;*

(5) *Λ is aligned with the unstable direction v^u, Λ has no periodic points but $\{p_1 + tv^u, t \in \mathbb{R}\}$ contains a periodic point of prime period q; if $\Lambda \cap T^{-q}\Lambda = \emptyset$ then $\zeta(\varphi_\Lambda, \{u_n\}) = 1$; if $\Lambda \cap T^{-q}\Lambda \neq \emptyset$ then $\zeta(\varphi_\Lambda, \{u_n\}) = (1 - \lambda^{-q})\frac{|p_2|}{l(\Lambda)}$;*

(6) *Λ is aligned with the stable direction v^u, Λ has no periodic points but $\{p_1 + tv^u, t \in \mathbb{R}\}$ contains a periodic point of prime period q; if $\Lambda \cap T^{-q}\Lambda = \emptyset$ then $\zeta(\varphi_\Lambda, \{u_n\}) = 1$; if $\Lambda \cap T^{-q}\Lambda \neq \emptyset$ then $\zeta(\varphi_\Lambda, \{u_n\}) = (1 - \lambda^{-q})\frac{|p_2|}{l(\Lambda)}$;*

Proof. We will only prove case (1), in which we will need the result of [3, Theorem 2.1 (1)]. The other cases use similar arguments and correspond to case (2) to (6) of [3, Theorem 2.1].

Below we verify the assumptions of Theorem 5.2.

Put $\delta_n = e^{-u_n}$. Then we see that $U_n = \{y : \varphi_\Lambda(y) > u_n\} = B_{\delta_n}(\Lambda)$. Since μ is the Lebesgue measure, it is straight forward to verify that Assumption 1 is satisfied with $r_n = \delta_n^2 = e^{-2u_n}$.

By the hyperbolicity of T, there exists $C > 0$ such that $\operatorname{diam} \mathcal{A}^n < C\lambda^{-n}$. This invites us to take

$$\kappa_n = \left\lfloor \frac{\ln C + 2u_n}{\ln \lambda} \right\rfloor + 1 = \mathcal{O}(\log n)$$

which guarantees that $\operatorname{diam} \mathcal{A}^{\kappa_n} < r_n$. On the other hand, $\mu(U_n) \lesssim e^{-u_n}l(\Lambda) = \mathcal{O}(1/n)$, so item (i) of Theorem 5.2 is satisfied for any $\varepsilon \in (0, 1)$.

We are left with the extremal index θ defined by (17). For this purpose we choose $K_n = (\log n)^5 \gg \kappa_n^2$. Now we estimate:

$$
\begin{aligned}
\mu_{U_n}(\tau_{U_n} \le K_n) &\le \frac{1}{\mu(U_n)} \sum_{j=1}^{(\log n)^5} \mu(U_n \cap T^{-j}U_n) \\
&\lesssim n \sum_{j=1}^{(\log n)^5} \mu(U_n \cap T^{-j}U_n) \\
&= o(1)
\end{aligned}
$$

where the last inequality follows from [3, Section 3.3, page 16]. This shows that

$$
\theta = \lim_n \mu_{U_n}(\tau_{U_n} > K_n) = 1 - \lim_n \mu_{U_n}(\tau_{U_n} \le K_n) = 1,
$$

finishing the proof of (ii) of Theorem 5.2. We conclude that

$$
\zeta(\varphi, \{\log n\}) = \theta = 1.
$$

$\square$

References

[1] M. Abadi. Hitting, returning and the short correlation function. *Bull. Braz. Math. Soc.*, 37(4) (2006), 1–17.

[2] H Bruin, M Demers and M Todd: Hitting and escaping statistics: mixing, targets and holes; Adv. Math., **328** (2018), 1263–1298.

[3] M. Carney, M. Holland and M. Nicol Extremes and extremal indices for level set observables on hyperbolic systems Preprint. Available at arxiv.org/abs/1909.04748.

[4] (MR3009103) 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.

[5] M. Demers and M. Todd. Asymptotic escape rates and limiting distributions for multimodal maps. Preprint. Available at arxiv.org/abs/1905.05457.

[6] M. Demers and L-S. Young. Escape rates and conditionally invariant measures. *Nonlinearity*, 19 (2006), no. 2, 377-397.

[7] F. Dreher and M. Kessebhmer. Escape rates for special flows and their higher order asymptotics. *Ergodic Theory Dynam. Systems* 39 (2019), no. 6, 1501–1530.

[8] D. Faranda, A. C. Moreira Freitas, J. Milhazes Freitas, M. Holland, T. Kuna, V. Lucarini, M. Nicol, M. Todd, and S. Vaienti. *Extremes and Recurrence in Dynamical Systems*. Wiley, New York, 2016.

[9] A Ferguson and M Pollicott: Escape Rates for Gibbs Measures; Ergod. Th. & Dynam. Syst. **32(3)** (2012), 961–988.

[10] A. C. M. Freitas, J. M. Freitas and M. Todd. The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. *Comm. Math. Phys.*, **321** (2013), 483–527.

[11] A. C. M. Freitas, J. M. Freitas and M. Magalhães. Convergence of marked point processes of excesses for dynamical systems. *J. Eur. Math. Sic.*, 20 (2018), 2131–2179

[12] A. C. M. Freitas, J. M. Freitas, F. B. Rodrigues and J. V. Soares. Rare events for Cantor target sets. Preprint. Available at arxiv.org/abs/1903.07200.

[13] N. Haydn and S. Vaienti. The distribution of return times near periodic orbits. *Probab. Theory Related Fields*, 144 (2009), 517–542.

[14] N. Haydn and S. Vaienti. Limiting entry times distribution for arbitrary null sets. Preprint. Available at arxiv.org/abs/1904.08733.

[15] N. Haydn and F. Yang. Local Escape Rates for φ-mixing Dynamical Systems. To appear on *Ergodic Theory Dynam. Systems*. Preprint available at arxiv.org/abs/1806.07148.

[16] G. Keller and C. Liverani. Rare events, escape rates and quasistationarity: some exact formulae. *J. Stat. Phys.*, 135 (2009) 519–534.

[17] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Commun. Math. Phys.*, 260 (2005), 131–146.

[18] F. Yang. Rare event process and entry times distribution for arbitrary null sets on compact manifolds. Preprint. Available at arxiv.org/abs/1905.09956.

[19] (MR1637655) L-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Annals of Math.*, **7** (1998), 585–650.

[20] (MR1750438) L-S. Young. Recurrence time and rate of mixing. *Israel J. of Math.*, **110** (1999), 153–188.