

Pressure and escape rates for random subshifts of finite type

Kevin McGoff

ABSTRACT. In this work we consider several aspects of the thermodynamic formalism in a randomized setting. Let X be a non-trivial mixing shift of finite type, and let $f : X \rightarrow \mathbb{R}$ be a Hölder continuous potential with associated Gibbs measure μ . Further, fix a parameter $\alpha \in (0, 1)$. For each $n \geq 1$, let \mathcal{F}_n be a random subset of words of length n , where each word of length n that appears in X is included in \mathcal{F}_n with probability $1 - \alpha$ (and excluded with probability α), independently of all other words. Then let $Y_n = Y(\mathcal{F}_n)$ be the random subshift of finite type obtained by forbidding the words in \mathcal{F}_n from X . In our first main result, for α sufficiently close to 1 and n tending to infinity, we show that the pressure of f on Y_n converges in probability to the value $P_X(f) + \log(\alpha)$, where $P_X(f)$ is the pressure of f on X . Additionally, let $H_n = H(\mathcal{F}_n)$ be the random hole in X consisting of the union of the cylinder sets of the words in \mathcal{F}_n . For our second main result, for α sufficiently close to one and n tending to infinity, we show that the escape rate of μ -mass through H_n converges in probability to the value $-\log(\alpha)$ as n tends to infinity.

1. Introduction

Random subshifts of finite type were introduced in [29], and they have subsequently been studied in [3, 30, 31]. Let us quickly recall their definition. Let X be a non-trivial mixing subshift of finite type (SFT), and let $B_n(X)$ be the set of words of length n that appear in X . For a fixed parameter $\alpha \in (0, 1)$ and $n \in \mathbb{N}$, let \mathcal{F}_n be the randomly selected subset of $B_n(X)$ formed by including each word from $B_n(X)$ in \mathcal{F}_n with probability $1 - \alpha$ (and excluding it from \mathcal{F}_n with probability α), independently of all other words. Then let $Y_n = Y(\mathcal{F}_n)$ be the set of points in X that do not contain any word from \mathcal{F}_n . We refer to Y_n as a random SFT. In order to study random SFTs, we fix the ambient system X and the parameter α . Then we seek to describe the properties of Y_n that have probability tending to one as n tends to infinity. This framework gives a precise way to describe the behavior of “typical” SFTs within the ambient system X .

Previous work on random SFTs [3, 29–31] has established the existence of at least one critical value α_c such that the typical behavior of Y_n changes abruptly as α crosses this value. Indeed, when $\alpha < \alpha_c$, there is a positive limit for the probability that Y_n is empty, and Y_n has zero entropy with probability tending to one. On the other hand, for $\alpha > \alpha_c$, the probability that Y_n is empty tends to zero, and the entropy of Y_n converges in probability to the value $h(X) + \log(\alpha)$, where $h(X)$

2010 *Mathematics Subject Classification*. Primary: 37B10.

is the entropy of X . (Note that this value is positive for $\alpha > \alpha_c$.) Furthermore, for α close enough to one, Y_n contains a unique “giant component,” which is itself a mixing SFT with full entropy, as well as a random number of isolated periodic orbits. See [29] for details.

In the present work, we study some aspects of the thermodynamic formalism for random SFTs in the super-critical regime ($\alpha > \alpha_c$). In our first main result (Theorem 1.1), we describe the distribution of the pressure of random SFTs for a fixed potential function, and in our second main result (Theorem 1.4), we describe the distribution of the escape rate of mass of Gibbs measures through random holes. Although these two topics may not at first appear to be related, they are in fact quite closely connected, as demonstrated by Proposition 7.1.

1.1. Pressure of random SFTs. Suppose that $f : X \rightarrow \mathbb{R}$ is a fixed Hölder continuous potential function. We seek to identify the limiting behavior of the pressure of f restricted to the random SFT Y_n in the limit as n tends to infinity. For notation, let $P_Y(f)$ denote the topological pressure of f restricted to any subshift $Y \subset X$ (see Section 2 for definitions).

THEOREM 1.1. *Let X be a non-trivial mixing SFT and $f : X \rightarrow \mathbb{R}$ Hölder continuous. Then there exists $\gamma_0 \in (0, 1)$ such that for each $\alpha \in (\gamma_0, 1]$ and for each $\epsilon > 0$, there exists $\rho > 0$ such that for all large enough n ,*

$$\mathbb{P}_\alpha \left(\left| P_{Y_n}(f) - (P_X(f) + \log(\alpha)) \right| \geq \epsilon \right) < e^{-\rho n}.$$

In other words, the pressure of f on the random SFT Y_n converges in probability to the value $P_X(f) + \log(\alpha)$. Note that when $f \equiv 0$, we recover the result of [29] regarding the entropy of random SFTs.

REMARK 1.2. The particular value of γ_0 that we use in our proof is given in Definition 2.1. It is possible that this value of γ_0 is not optimal, in the sense that for some choices of X and f , the statement might remain true with a smaller value of γ_0 . However, one may check that for $f \equiv 0$, our definition of γ_0 is equal to α_c , which is optimal.

REMARK 1.3. The proof of Theorem 1.1 appears in Section 3. The broad outline of this proof is similar to the outline of the proof of [29, Theorem 1.3] concerning the entropy of random SFTs. However, the core technical results in the proof, which appear in Section 4, require new ideas to handle the fact that f may not be zero. In particular, we must estimate the μ measure of the appearance of certain types of repeated patterns, where μ is a Gibbs measure but not necessarily a measure of maximal entropy, and this generalization requires substantially new ideas.

1.2. Escape rate through random holes. Our second main result involves thinking of the random set of forbidden words \mathcal{F}_n as a hole in the ambient system X , creating an open dynamical system. For an introduction to open systems, see [14] and references therein. Previous work on open systems has focused largely on the existence and properties of the escape rate of mass through the hole, as well as the existence and properties of conditionally invariant distributions (conditional on avoiding the hole); for an incomplete sampling of the literature on open systems, see [4–13, 15–25, 27, 28, 33]. In this work we focus on the escape rate of mass

through the hole, which we define below. Let $\sigma : X \rightarrow X$ denote the left-shift map on X . For a Borel probability measure μ on X and a hole $H \subset X$, the escape rate of μ through H is defined to be $-\varrho(\mu : H)$, where

$$\varrho(\mu : H) = \lim_m \frac{1}{m} \log \mu \left(\left\{ x \in X : \forall k \in \{0, \dots, m-1\}, \sigma^k(x) \notin H \right\} \right),$$

whenever the limit exists.

In this work we consider random holes H_n , constructed from the set of forbidden words \mathcal{F}_n as follows:

$$H_n = \bigcup_{w \in \mathcal{F}_n} \{x \in X : x_0 \dots x_{n-1} = w\}.$$

Our goal is to describe the escape rate of mass of Gibbs measures through H_n .

THEOREM 1.4. *Let X be a non-trivial mixing SFT and let μ be the Gibbs measure associated to the Hölder continuous potential function $f : X \rightarrow \mathbb{R}$. Then there exists $\gamma_0 \in (0, 1)$ such that for each $\alpha \in (\gamma_0, 1]$ and for each $\epsilon > 0$, there exists $\rho > 0$ such that for all large enough n ,*

$$\mathbb{P} \left(\left| \varrho(\mu : H_n) - \log(\alpha) \right| \geq \epsilon \right) < e^{-\rho n}.$$

Thus the escape rate of μ through the randomly selected hole H_n converges in probability to $-\log(\alpha)$.

REMARK 1.5. With probability tending to one, the hole H_n consists of the union of approximately $(1 - \alpha)|B_n(X)|$ cylinder sets of length n . One may think of H_n as typically consisting of the union of many small holes spread randomly throughout the state space. Furthermore, the expected value of the μ -measure of H_n is $1 - \alpha$, which remains bounded away from zero as n tends to infinity. In this sense, the holes considered here differ substantially from the “small holes” studied in some previous work [5, 16, 19, 23].

REMARK 1.6. In [5], the authors prove that in the deterministic setting, the escape rate depends on both the size (measure) of the hole and its precise location in state space. In contrast, Theorem 1.4 shows that for random holes, the escape rate is well approximated by a function that depends only on the expected measure of the hole. Indeed, the expected measure of the hole is $(1 - \alpha)$ (see Remark 2.3), so the expected measure of its complement is α . Then Theorem 1.4 yields that the escape rate converges in probability to the value $-\log(\alpha)$.

REMARK 1.7. A naive approximation of the hitting time for the hole is given by a geometrically distributed random variable τ with probability of success $p = 1 - \alpha$. Since $\mathbb{P}(\tau = k) = \alpha^{k-1}(1 - \alpha)$, we have that

$$\lim_k \frac{1}{k} \log \mathbb{P}(\tau > k) = \log \alpha.$$

From this perspective, Theorem 1.4 may be interpreted as giving precise meaning to the statement that the escape rate through the random hole is approximately the same as one would obtain if the hitting time of the hole were geometrically distributed with probability of success equal to the measure of the hole.

As a consequence of Theorem 1.4, we can also estimate the escape rate of mass of Gibbs measures for Axiom A diffeomorphisms through randomly selected Markov holes. As the proof relies solely on Theorem 1.4 and the well-known relationship between Markov partitions for Axiom A diffeomorphisms and SFTs (see [2]), we omit the proof.

COROLLARY 1.8. *Let $T : M \rightarrow M$ be an Axiom A diffeomorphism such that the restriction of T to its non-wandering set is topologically mixing, and let $f : M \rightarrow \mathbb{R}$ be a Hölder continuous potential with Gibbs measure ν . Let ξ be a finite Markov partition of M with diameter small enough that the symbolic dynamics is well-defined, and let $\xi^n = \bigvee_{k=0}^{n-1} T^{-k}\xi$. Then there exists $\gamma_0 \in (0, 1)$ such that for each $\alpha \in (\gamma_0, 1]$, the following holds. Let H_n be the randomly selected hole obtained by including each cell of ξ^n independently with probability $1 - \alpha$. Then for each $\epsilon > 0$, there exists $\rho > 0$ such that for large enough n ,*

$$\mathbb{P}\left(\left|\varrho(\nu : H_n) - \log(\alpha)\right| \geq \epsilon\right) < e^{-\rho n}.$$

1.3. Outline of the paper. The following section collects some background definitions and results that are used elsewhere in the paper. In Section 3, we present the proof of Theorem 1.1 with the help of several technical lemmas. These technical lemmas are then proved in Sections 4, 5, and 6. Finally, in Section 7, we establish Proposition 7.1, which relates escape rates to pressure, and then we prove Theorem 1.4.

2. Background and notation

2.1. Symbolic dynamics. Let \mathcal{A} be a finite set, which we call the alphabet. We let $\Sigma = \mathcal{A}^{\mathbb{Z}}$ denote the full-shift, and we let $\sigma : \Sigma \rightarrow \Sigma$ be the left-shift map, $\sigma(x)_n = x_{n+1}$. We endow Σ with the product topology from the discrete topology on \mathcal{A} , which makes σ a homeomorphism. We define the metric $d(\cdot, \cdot)$ on Σ by the rule $d(x, y) = 2^{-n(x, y)}$, where $n(x, y)$ is the infimum of all $|m|$ such that $x_m \neq y_m$.

A subset $X \subset \Sigma$ is a *subshift* if it is closed and $\sigma(X) = X$. In the context of a subshift X , we also use the symbol σ to denote the restriction of the left-shift to X . A *word* on \mathcal{A} is an element of \mathcal{A}^m for some $m \geq 1$. We also refer to the empty word as a word. If $u = u_1 \dots u_m$ is in \mathcal{A}^m , then we say that u has length m , and we let u_i^j denote the subword $u_i \dots u_j$. Further, we let $B_m(X)$ denote the set of words of length m that appear in some point in X . For any word w in $B_m(X)$, we let $[w]$ denote the set of points $x \in X$ such that $x_0 \dots x_{m-1} = w$. Also, for $x \in X$ and $i \leq j$, we let $x[i, j]$ denote the set of points $y \in X$ such that $y_i \dots y_j = x_i \dots x_j$.

A subset $X \subset \Sigma$ is a *subshift of finite type* (SFT) if there exists a natural number m and a collection of words $\mathcal{F} \subset \mathcal{A}^m$ such that X is exactly the set of points in Σ that contain no words from \mathcal{F} . We say that an SFT is non-trivial if it contains at least two points. The SFT X is mixing if there exists N such that for all points $x, y \in X$, there exists a point $z \in X$ such that $x(-\infty, 0] = z(-\infty, 0]$ and $y[N, \infty) = z[N, \infty)$.

For any subshift $X \subset \Sigma$, we let $M(X, \sigma)$ denote the set of Borel probability measures μ on X such that $\mu(\sigma^{-1}A) = \mu(A)$ for all Borel sets $A \subset X$. Suppose $\mu \in M(X, \sigma)$. When it will not cause confusion, we write $\mu(w)$ to denote the measure of the cylinder set $[w]$, where $w \in B_m(X)$ for some $m \geq 1$.

For any measure $\mu \in M(X, \sigma)$, one may define the entropy of μ as

$$h(\mu) = \lim_m \frac{1}{m} \sum_{w \in B_m(X)} -\mu(w) \log \mu(w),$$

where the limit exists by subadditivity.

2.2. Pressure and equilibrium states. Let Y be a subshift, and let $f : Y \rightarrow \mathbb{R}$ be continuous. For $m \geq 1$ and w in $B_m(Y)$, let

$$S_m f(w) = \sup_{x \in Y \cap [w]} \sum_{k=0}^{m-1} f \circ \sigma^k(x).$$

Then let

$$\Lambda_m(Y) = \sum_{w \in B_m(Y)} e^{S_m f(w)}.$$

Finally, the (topological) pressure of f on Y is defined as

$$P_Y(f) = \lim_{m \rightarrow \infty} \frac{1}{m} \log \Lambda_m(Y),$$

where the limit exists by subadditivity.

The well-known Variational Principle (see [34]) states that

$$P_Y(f) = \sup \left\{ \int f d\mu + h(\mu) : \mu \in M(Y, \sigma) \right\}.$$

For a subshift Y , this supremum must be realized, and any measure that attains the supremum is known as an equilibrium state for f on Y .

Now suppose that X is a mixing SFT and $f : X \rightarrow \mathbb{R}$ is Hölder continuous. In this case, it is known that there is a unique equilibrium state $\mu \in M(X, \sigma)$ for f , and furthermore μ satisfies the following Gibbs property: there exists $K > 1$ such that for all $n \geq 1$ and $x \in X$,

$$(2.1) \quad K^{-1} \leq \frac{\mu(x[0, n-1])}{\exp(-P_X(f) \cdot n + \sum_{k=0}^{n-1} f \circ \sigma^k(x))} \leq K.$$

We may now give a definition for the parameter γ_0 that appears in Theorems 1.1 and 1.4.

DEFINITION 2.1. Let X be a non-empty mixing SFT, and let $f : X \rightarrow \mathbb{R}$ be a Hölder continuous potential with associated Gibbs measure μ . Then let $\gamma_0 = \gamma_0(X, f)$ be defined by

$$\gamma_0 = \inf \left\{ \gamma > 0 : \exists n_0, \forall m \geq n_0, \forall u \in B_m(X), \mu(u) \leq \gamma^m \right\}.$$

Note that by [1, Lemma 5], if X is non-trivial, then $\gamma_0 < 1$. We also make the following remark. Suppose $1 \geq \gamma > \gamma_0$, and fix n_0 such that $\mu(u) \leq \gamma^{|u|}$ whenever $|u| \geq n_0$. Then for any word u we have $\mu(u) \leq \gamma^{|u|-n_0}$; indeed, if $|u| \geq n_0$, then it follows from the choice of n_0 , and if $|u| \leq n_0$, then $\mu(u) \leq 1 \leq \gamma^{|u|-n_0}$.

It is well-known (see, e.g., [2, Proof of Proposition 1.14]) that μ satisfies a mixing property called ψ -mixing, from which a variety of mixing-type estimates may be deduced. The bounds required for the present work are summarized in the following lemma, which we state without proof.

LEMMA 2.2. *Let X be a non-trivial mixing SFT with Hölder continuous potential $f : X \rightarrow \mathbb{R}$ and associated Gibbs measure μ . Then there exist constants $K > 0$ and $g_0 \geq 1$ such that:*

- *the Gibbs property (2.1) holds;*
- *for all $m, n \geq 1$ and for all $u \in B_m(X)$ and $v \in B_n(X)$ such that $uv \in B_{m+n}(X)$, we have*

$$\mu(uv) \leq K\mu(u)\mu(v);$$

- *for all $m, n \geq 1$ and for all $u \in B_m(X)$ and $v \in B_n(X)$ such that $uv \in B_{m+n}(X)$, we have*

$$\mu(\sigma^{-m}[v] \mid [u]) \leq K\mu([v]);$$

- *for $g \geq g_0$, for all $m, n \geq 1$ and for all $u \in B_m(X)$ and $v \in B_n(X)$, we have*

$$\mu\left([u] \cap \sigma^{-g+m}[v]\right) \geq K^{-1}\mu([u])\mu([v]).$$

2.3. Basics of random SFTs. Let X be a non-trivial mixing SFT. Fix $\alpha \in (0, 1)$. Recall that \mathcal{F}_n denotes the random subset of $B_n(X)$ formed by including each word with probability $1 - \alpha$, independently of all other words, and $Y_n = Y(\mathcal{F}_n)$ is the random SFT formed by forbidding the words \mathcal{F}_n from X . Here we establish some notation and basic facts for random SFTs.

Let $u \in B_k(X)$ for some $k \geq n$. We let $W_n(u)$ denote the set of all words of length n that appear in u :

$$W_n(u) = \{v \in B_n(X) : \exists j \in \{1, \dots, k - n + 1\}, u_j^{j+n-1} = v\}.$$

Then let ξ_u denote the indicator function of the event that u contains no words from \mathcal{F}_n , *i.e.*,

$$\xi_u = \begin{cases} 1, & \text{if } W_n(u) \cap \mathcal{F}_n = \emptyset \\ 0, & \text{otherwise.} \end{cases}$$

Since each word in $W_n(u)$ is excluded from \mathcal{F}_n with probability α , independently of all other words, we have that

$$(2.2) \quad \mathbb{E}[\xi_u] = \mathbb{P}(W_n(u) \cap \mathcal{F}_n = \emptyset) = \alpha^{|W_n(u)|}.$$

Furthermore, for $u, v \in B_k(X)$, the covariance of ξ_u and ξ_v is given by

$$(2.3) \quad \begin{aligned} \mathbb{E}\left[(\xi_u - \mathbb{E}[\xi_u])(\xi_v - \mathbb{E}[\xi_v])\right] &= \mathbb{E}[\xi_u \xi_v] - \mathbb{E}[\xi_u]\mathbb{E}[\xi_v] \\ &= \alpha^{|W_n(u) \cup W_n(v)|} - \alpha^{|W_n(u)| + |W_n(v)|} \\ &= \alpha^{|W_n(u) \cup W_n(v)|} (1 - \alpha^{|W_n(u) \cap W_n(v)|}). \end{aligned}$$

REMARK 2.3. Let X be a non-trivial mixing SFT, and let $\mu \in M(X, \sigma)$. Then the expected value of the μ -measure of the hole H_n is $1 - \alpha$, since

$$\begin{aligned} \mathbb{E}[\mu(H_n)] &= \mathbb{E}\left[\sum_{u \in B_n(X)} \mu(u)(1 - \xi_u)\right] \\ &= \sum_{u \in B_n(X)} \mu(u)(1 - \mathbb{E}[\xi_u]) = (1 - \alpha)\mu(B_n(X)) = 1 - \alpha. \end{aligned}$$

2.4. Repeats and repeat covers. We use interval notation to denote intervals in \mathbb{Z} . For example, $[1, 3] = \{1, 2, 3\}$ and $[0, 5) = \{0, 1, 2, 3, 4\}$. Furthermore, for a set $F \subset \mathbb{Z}$ and $t \in \mathbb{Z}$, we let $t + F = \{t + s : s \in F\}$. For n in \mathbb{N} and $F \subset \mathbb{Z}$, we let $\mathcal{C}_n(F)$ denote the set of intervals of length n contained in F :

$$\mathcal{C}_n(F) = \{t + [0, n) : t + [0, n) \subset F\}.$$

We will also consider sets of pairs of intervals; that is, we consider sets $\mathcal{R} \subset \mathcal{C}_n(F) \times \mathcal{C}_n(F)$. We let $\text{Power}(\mathcal{C}_n(F) \times \mathcal{C}_n(F))$ denote the power set of $\mathcal{C}_n(F) \times \mathcal{C}_n(F)$. For such \mathcal{R} , we let $|\mathcal{R}|$ denote the number of pairs in \mathcal{R} , and we let

$$A(\mathcal{R}) = \bigcup_{(I_1, I_2) \in \mathcal{R}} I_2.$$

Now we define *repeats* and *repeat covers*, which were used implicitly in [29] and then defined explicitly in [31].

DEFINITION 2.4. Let \mathcal{A} be a finite set. Let $F \subset \mathbb{Z}$, and let $u \in \mathcal{A}^F$. A pair (I_1, I_2) in $\mathcal{C}_n(F) \times \mathcal{C}_n(F)$ is an *n-repeat* (or just a *repeat*) for u if $u_{I_1} = u_{I_2}$ and I_1 is the lexicographically minimal occurrence of the word u_{I_1} in u . In that case, the word u_{I_1} is called a repeated word for u . Furthermore, a set $\mathcal{R} \subset \mathcal{C}_n(F) \times \mathcal{C}_n(F)$ is a *repeat cover* for u if

- (1) each pair $(I_1, I_2) \in \mathcal{R}$ is a repeat for u , and
- (2) for each repeat (I_1, I_2) for u , we have $I_2 \subset A(\mathcal{R})$.

Note that every pattern $u \in \mathcal{A}^F$ has a repeat cover, which contains all repeats for u . When \mathcal{R} is a repeat cover for u , we may refer to $A(\mathcal{R})$ as the repeat area for u (and we note that it is independent of the choice of repeat cover for u). In many cases, we seek to find efficient repeat covers, by which we mean repeat covers \mathcal{R} such that $|\mathcal{R}|$ is small enough for our purposes. In this paper, we only require the bound supplied by the following lemma, which is a slightly weaker version of Lemma 3.8 in [31].

LEMMA 2.5. *Let $F \subset \mathbb{Z}$ be a finite union of intervals of length n , and suppose $u \in \mathcal{A}^F$. Then there exists an n -repeat cover \mathcal{R} for u such that*

$$|\mathcal{R}| \leq 4|F|/n.$$

Additionally, our proofs make use of the following estimate relating the number of unique words of length n in $u \in B_k(X)$ and the cardinality of the repeat area for u .

LEMMA 2.6. *Let $F \subset \mathbb{Z}$ be a finite union of intervals of length n , and let $a = |\{s : s + [0, n) \subset F\}|$. Suppose that $u \in \mathcal{A}^F$ satisfies $|W_n(u)| = j < a$. Then for any repeat cover \mathcal{R} for u ,*

$$|A(\mathcal{R})| \geq a + n - j - 1.$$

PROOF. Let $r = a - j$, which is the number of repeats for u . The lexicographically minimal repeat for u contributes n elements to $A(\mathcal{R})$, and each of the other $r - 1$ repeats must contribute at least one element. Altogether, we must have $|A(\mathcal{R})| \geq n + r - 1 = n + a - j - 1$. \square

In many of the proofs in Section 4, we decompose words into alternating blocks of repeated regions (*i.e.*, regions contained in $A(\mathcal{R})$ for some repeat cover \mathcal{R}) and non-repeated regions. The following definition standardizes some notation that is

useful for such decompositions. We endow \mathbb{Z} with the standard graph structure, in which two nodes $x, y \in \mathbb{Z}$ are connected by an edge whenever $|x - y| = 1$. We then endow all subsets of \mathbb{Z} with the induced subgraph structure, and references to connected components refer to this subgraph structure. Furthermore, we give \mathbb{Z} the standard ordering, and if I and J are disjoint subsets of \mathbb{Z} , then we let $I < J$ whenever $x < y$ for all $x \in I$ and $y \in J$.

DEFINITION 2.7. Let $A \subset [0, k)$ be a union of intervals of length n such that $0 \notin A$. Then the *interval decomposition* of $[0, k)$ induced by A consists of $((I_m)_{m=1}^{N+1}, (J_m)_{m=1}^N)$, where

- each J_m is a non-empty maximal connected component of A , and $\bigcup_m J_m = A$;
- each I_m is a maximal connected component of $[0, k) \setminus A$, and $\bigcup_m I_m = [0, k) \setminus A$;
- only I_{N+1} may be empty;
- for each $m = 1, \dots, N$, we have $I_m < J_m < I_{m+1}$.

Now suppose $b \in \mathcal{A}^k$ and $\mathcal{R} \subset \mathcal{C}_n([0, k)) \times \mathcal{C}_n([0, k))$. Let $A = A(\mathcal{R})$, and let $((I_m)_{m=1}^{N+1}, (J_m)_{m=1}^N)$ be the interval decomposition of $[0, k)$ induced by A . For each m , we let $u_m = b|_{I_m}$ and $v_m = b|_{J_m}$. We refer to $((u_m)_{m=1}^{N+1}, (v_m)_{m=1}^N)$ as the *block decomposition* of b . If \mathcal{R} is a repeat cover for b , then we refer to $((u_m)_{m=1}^{N+1}, (v_m)_{m=1}^N)$ as the *repeat block decomposition* of b . Note that $N \leq |\mathcal{R}|$.

When μ is a Gibbs measure associated to a Hölder continuous potential, the following lemma, which is used several times in Section 4, gives an estimate of the μ -measure of any word b in terms of a block decomposition.

LEMMA 2.8. *Let X be a non-trivial mixing SFT with Hölder continuous potential $f : X \rightarrow \mathbb{R}$ and associated Gibbs measure μ . Let $K > 0$ satisfy the conclusions of Lemma 2.2. Let $b \in B_k(X)$, and suppose that $((u_m)_{m=1}^{N+1}, (v_m)_{m=1}^N)$ is a block decomposition of b . Then*

$$\mu(b) \leq K^{2N} \prod_{m=1}^{N+1} \mu(u_m) \prod_{m=1}^N \mu(v_m).$$

PROOF. Let $((I_m)_{m=1}^{N+1}, (J_m)_{m=1}^N)$ be a block decomposition of b . Let s_m be the minimal element of the corresponding interval I_m , and let t_m be the minimal element of the interval J_m . In order to avoid confusion, in this proof we use proper cylinder set notation: for $u \in B_m(X)$, we let $[u]$ denote the set of points x in X such that $x_0 \dots x_{m-1} = u$. Using conditional probabilities, we have

$$\begin{aligned} \mu([b]) &= \mu([u_1]) \prod_{m=1}^N \mu(\sigma^{-t_m}[v_m] \mid [u_1 v_1 \dots v_{m-1} u_m]) \\ &\quad \times \prod_{m=1}^N \mu(\sigma^{-s_{m+1}}[u_{m+1}] \mid [u_1 v_1 \dots u_m v_m]). \end{aligned}$$

Then by our choice of K , we have

$$\mu([b]) \leq K^{2N} \prod_{m=1}^{N+1} \mu([u_m]) \prod_{m=1}^N \mu([v_m]),$$

as desired. □

3. Pressure of random SFTs

In this section we give a proof of Theorem 1.1. The broad outline of the proof involves finding upper and lower bounds on the pressure in terms of some other random variables, followed by a second moment argument showing that these auxiliary random variables each converge in probability to $P_X(f) + \log \alpha$. For the sake of exposition, we present the argument here and defer the proofs of the many technical lemmas to later sections of the paper. We hope that this presentation helps clarify the main argument and also motivate the technical lemmas. Note that at the beginning of this proof we define some notation and choose some parameters, including the random variables $\phi_{n,k}$ and $\psi_{n,k}$, and we make frequent reference to both the notation and the parameters throughout Sections 4 - 6 in the technical lemmas.

PROOF OF THEOREM 1.1. Let X be a non-trivial mixing SFT. Let $f : X \rightarrow \mathbb{R}$ be a Hölder continuous potential with associated Gibbs measure μ . Choose $\gamma_0 = \gamma_0(X, f)$ as in Definition 2.1, and note that $\gamma_0 < 1$. Let $\alpha \in (\gamma_0, 1]$, and let $\epsilon > 0$. Furthermore, fix K and g_0 as in Lemma 2.2.

We begin by selecting a variety of parameters for our proof. Since $\alpha > \gamma_0$, there exists γ in the interval (γ_0, α) . According to the definition of γ_0 , since $\gamma > \gamma_0$, there exists n_0 such that for all $m \geq n_0$, for all $u \in B_m(X)$, we have $\mu(u) \leq \gamma^m$. We assume throughout that $n \geq n_0$. Choose $\delta > 0$ such that $\delta < \log(\alpha\gamma^{-1})/4$. Fix a sequence $k = k(n)$ such that $n/k \rightarrow 0$ and $k = o(n^2/\log n)$. (For example, one may choose $k = \lceil n^{1+\nu} \rceil$ for any $0 < \nu < 1$.) Now let $\ell = k - n + 1$, which is the number of positions s in $[0, k)$ such that $s + [0, n) \subset [0, k)$.

Having made these parameter choices, we now proceed to define our upper and lower bounds on $P_{Y_n}(f)$. First, for all $m \geq n$, for all $u \in B_m(X)$, recall from Section 2.3 that ξ_u is the random variable that is one if u is allowed (*i.e.* $W_n(u) \cap \mathcal{F}_n = \emptyset$) and zero otherwise. Then define

$$\phi_{n,k} = \sum_{u \in B_k(X)} e^{S_k f(u)} \xi_u.$$

By Lemma 6.1, $\phi_{n,k}$ may be used to provide an upper bound on $P_{Y_n}(f)$:

$$(3.1) \quad P_{Y_n}(f) \leq \frac{1}{k} \log \phi_{n,k}.$$

Now we turn towards the lower bound on $P_{Y_n}(f)$. Recall that we have already defined $\delta > 0$ above. Consider the set of words of length n that are entropy-typical for μ with tolerance δ :

$$E_n = \left\{ u \in B_n(X) : \left| -\frac{1}{n} \log \mu(u) - h(\mu) \right| < \delta \right\}.$$

Then let $G_{n,k}$ be the set of words of length k that begin and end with the same word of length n from E_n :

$$G_{n,k} = \left\{ u \in B_k(X) : u_1^n = u_\ell^k \text{ and } u_1^n \in E_n \right\}.$$

Next we define the random variable

$$\psi_{n,k} = \frac{1}{|E_n|} \sum_{u \in G_{n,k}} e^{S_k f(u)} \xi_u.$$

By Lemma 6.2, we see that $\psi_{n,k}$ may be used to bound $P_{Y_n}(f)$ from below: for all large enough n ,

$$(3.2) \quad P_{Y_n}(f) \geq \frac{1}{k} \log \psi_{n,k} - \epsilon/2.$$

By Lemmas 5.1, 5.2, 5.3, and 5.4, we have the following asymptotic results on the expectation and variance of $\phi_{n,k}$ and $\psi_{n,k}$:

- (I) $\lim_n \frac{1}{k} \log \mathbb{E}[\phi_{n,k}] = P_X(f) + \log(\alpha)$;
- (II) $\lim_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] = P_X(f) + \log(\alpha)$;
- (III) there exists $\rho_1 > 0$ such that for all large enough n ,

$$\frac{\text{Var}[\phi_{n,k}]}{\mathbb{E}[\phi_{n,k}]^2} \leq e^{-\rho_1 n};$$

- (IV) there exists $\rho_2 > 0$ such that for all large enough n ,

$$\frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} \leq e^{-\rho_2 n}.$$

The first two properties indicate that we expect $\phi_{n,k}$ and $\psi_{n,k}$ to be on the correct exponential order of magnitude, while the third and fourth properties show that these random variables are well concentrated around their expected values. Combining these properties with Chebyshev's inequality, we are able to finish the proof as follows.

By the monotonicity of \mathbb{P} under inclusion, the union bound, and the displays (3.1) and (3.2), for all large enough n , we have

$$(3.3) \quad \begin{aligned} & \mathbb{P}\left(|P_{Y_n}(f) - (P_X(f) + \log \alpha)| \geq \epsilon\right) \\ & \leq \mathbb{P}\left(P_{Y_n}(f) \geq P_X(f) + \log \alpha + \epsilon\right) + \mathbb{P}\left(P_{Y_n}(f) \leq P_X(f) + \log \alpha - \epsilon\right) \\ & \leq \mathbb{P}\left(\frac{1}{k} \log \phi_{n,k} \geq P_X(f) + \log \alpha + \epsilon\right) + \mathbb{P}\left(\frac{1}{k} \log \psi_{n,k} - \epsilon/2 \leq P_X(f) + \log \alpha - \epsilon\right) \\ & = \mathbb{P}\left(\phi_{n,k} \geq e^{k(P_X(f) + \log \alpha + \epsilon)}\right) + \mathbb{P}\left(\psi_{n,k} \leq e^{k(P_X(f) + \log \alpha - \epsilon/2)}\right). \end{aligned}$$

We proceed to bound the two terms on the right-hand side separately. For the first, Chebyshev gives

$$\begin{aligned} & \mathbb{P}\left(\phi_{n,k} \geq e^{k(P_X(f) + \log \alpha + \epsilon)}\right) \\ & = \mathbb{P}\left(\phi_{n,k} - \mathbb{E}[\phi_{n,k}] \geq e^{k(P_X(f) + \log \alpha + \epsilon)} - \mathbb{E}[\phi_{n,k}]\right) \\ & = \mathbb{P}\left(\phi_{n,k} - \mathbb{E}[\phi_{n,k}] \geq \text{Var}[\phi_{n,k}]^{1/2} \frac{\mathbb{E}[\phi_{n,k}]}{\text{Var}[\phi_{n,k}]^{1/2}} \left(e^{k(P_X(f) + \log \alpha + \epsilon)} / \mathbb{E}[\phi_{n,k}] - 1\right)\right) \\ & \leq \frac{\text{Var}[\phi_{n,k}]}{\mathbb{E}[\phi_{n,k}]^2} \left(e^{k(P_X(f) + \log \alpha + \epsilon)} / \mathbb{E}[\phi_{n,k}] - 1\right)^{-2}. \end{aligned}$$

Then by properties (I) and (III), there exists $\rho_3 > 0$ such that for all large n ,

$$(3.4) \quad \mathbb{P}\left(\phi_{n,k} \geq e^{k(P_X(f)+\log \alpha+\epsilon)}\right) < e^{-\rho_3 n}.$$

Similarly, for the second term in the last line of (3.3), Chebyshev gives

$$\begin{aligned} & \mathbb{P}\left(\psi_{n,k} \leq e^{k(P_X(f)+\log \alpha-\epsilon/2)}\right) \\ &= \mathbb{P}\left(\psi_{n,k} - \mathbb{E}[\psi_{n,k}] \leq e^{k(P_X(f)+\log \alpha-\epsilon/2)} - \mathbb{E}[\psi_{n,k}]\right) \\ &= \mathbb{P}\left(\psi_{n,k} - \mathbb{E}[\psi_{n,k}] \leq \text{Var}[\psi_{n,k}]^{1/2} \frac{\mathbb{E}[\psi_{n,k}]}{\text{Var}[\psi_{n,k}]^{1/2}} \left(e^{k(P_X(f)+\log \alpha-\epsilon/2)}/\mathbb{E}[\psi_{n,k}] - 1\right)\right) \\ &\leq \frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} \left(e^{k(P_X(f)+\log \alpha-\epsilon/2)}/\mathbb{E}[\psi_{n,k}] - 1\right)^{-2}. \end{aligned}$$

Then by properties (II) and (IV), there exists $\rho_4 > 0$ such that for all large n ,

$$(3.5) \quad \mathbb{P}\left(\psi_{n,k} \leq e^{k(P_X(f)+\log \alpha-\epsilon/2)}\right) < e^{-\rho_4 n}.$$

Combining the inequalities in (3.3), (3.4), and (3.5), we obtain the desired result. \square

4. Repeat probabilities

In this section we bound the μ measure of sets of words that have repeated subwords. By the well-known result of Ornstein and Weiss [32], the first return time of a μ -typical point x to its initial block of length n is approximately $e^{h(\mu)n}$. Then for μ -typical words of polynomial length in n , one would expect to find no repeated words of length n at all. However, to control the expectation and variance of $\phi_{n,k}$ and $\psi_{n,k}$, it is important to give more precise estimates on just how unlikely it is that a word of length k will have exactly j distinct subwords of length n , for each $1 \leq j \leq k - n + 1$.

Throughout this section we use the same environment (notation, parameters, and assumptions) laid out at the beginning of the proof of Theorem 1.1. The results of this section are used in the following section to establish properties (I) - (IV) from the proof of Theorem 1.1.

We begin by considering some sets of words that have exactly j distinct subwords of length n . For $1 \leq j \leq \ell$, we define the following sets:

$$\begin{aligned} B_{n,k}^j &= \{u \in B_k(X) : |W_n(u)| = j\} \\ G_{n,k}^j &= \{u \in G_{n,k} : |W_n(u)| = j\}. \end{aligned}$$

Furthermore, for $1 \leq j \leq 2\ell$, we let

$$\begin{aligned} D_{n,k}^j &= \{(u, v) \in B_k(X) \times B_k(X) : W_n(u) \cap W_n(v) \neq \emptyset, |W_n(u) \cup W_n(v)| = j\} \\ Q_{n,k} &= \{(u, v) \in G_{n,k} \times G_{n,k} : W_n(u) \cap W_n(v) \neq \emptyset\} \\ Q_{n,k}^j &= \{(u, v) \in Q_{n,k} : |W_n(u) \cup W_n(v)| = j\}. \end{aligned}$$

In Lemmas 4.1 and 4.2 we find bounds on the μ measure of the sets $B_{n,k}^j$ and $G_{n,k}$. In subsequent lemmas (Lemmas 4.3 - 4.5), we also find bounds on the $\mu \times \mu$ measure of $D_{n,k}^j$, $Q_{n,k}$, and $Q_{n,k}^j$. These estimates are used in the following section to bound the expectation and variance of $\phi_{n,k}$ and $\psi_{n,k}$.

LEMMA 4.1. *There exists a polynomial $p_1(x)$ such that for all large enough n , for each $1 \leq j \leq \ell$, we have*

$$\mu(B_{n,k}^j) \leq p_1(n)^{k/n} \gamma^{k-j}.$$

PROOF. Consider $n \geq n_0$ and $1 \leq j \leq \ell$. We define a map $\varphi : B_{n,k}^j \rightarrow \{(\mathcal{R}, w) : \mathcal{R} \subset \mathcal{C}_n([0, k]) \times \mathcal{C}_n([0, k]), w \in \mathcal{A}^{[0,k] \setminus A(\mathcal{R})}\}$ as follows. Let b be in $B_{n,k}^j$. Let \mathcal{R} be a repeat cover of b such that $|\mathcal{R}| \leq 4k/n$, which exists by Lemma 2.5. Let $((u_m)_{m=1}^{N+1}, (v_m)_{m=1}^N)$ be a repeat block decomposition of b (as in Definition 2.7). Then set $\varphi(b) = (\mathcal{R}, (v_m)_{m=1}^N)$. Furthermore, note that by Lemma 2.8,

$$\mu(b) \leq K^{2N} \prod_{m=1}^N \mu(v_m) \prod_{m=1}^{N+1} \mu(u_m).$$

Since each block v_m has length at least n (it contains at least one repeated word of length n from b) and $n \geq n_0$, we have that $\mu(v_m) \leq \gamma^{|v_m|}$. Then

$$\mu(b) \leq K^{2N} \gamma^{\sum_m |v_m|} \prod_{m=1}^{N+1} \mu(u_m) = K^{2N} \gamma^{|A(\mathcal{R})|} \prod_{m=1}^{N+1} \mu(u_m)$$

Using that $N \leq |\mathcal{R}| \leq 4k/n$ and $|A(\mathcal{R})| \geq k - j$ (by Lemma 2.6), we see that

$$\begin{aligned} \mu(b) &\leq (K^8)^{k/n} \gamma^{|A(\mathcal{R})|} \prod_{m=1}^{N+1} \mu(u_m) \\ (4.1) \quad &\leq (K^8)^{k/n} \gamma^{k-j} \prod_{m=1}^{N+1} \mu(u_m). \end{aligned}$$

Now define the projection map $\pi : \varphi(B_{n,k}^j) \rightarrow \text{Power}(\mathcal{C}_n(k) \times \mathcal{C}_n(k))$, given by $\pi((\mathcal{R}, w)) = \mathcal{R}$. Let $S = \pi \circ \varphi(B_{n,k}^j)$. Note that $|\mathcal{C}_n([0, k])| \leq k$, and therefore $|\mathcal{C}_n([0, k]) \times \mathcal{C}_n([0, k])| \leq k^2$. Furthermore, since each \mathcal{R} in $\pi \circ \varphi(B_{n,k}^j)$ satisfies $|\mathcal{R}| \leq 4k/n$, we have that $|S| \leq |\mathcal{C}_n([0, k]) \times \mathcal{C}_n([0, k])|^{4k/n} \leq (k^2)^{4k/n} = (k^8)^{k/n}$.

Having established these bounds, we may now estimate the μ measure of $B_{n,k}^j$ as follows. By rearranging the sum, we have

$$\begin{aligned} \mu(B_{n,k}^j) &= \sum_{b \in B_{n,k}^j} \mu(b) \\ &= \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m)) \in \pi^{-1}(\mathcal{R})} \sum_{b \in \varphi^{-1}(\mathcal{R}, (u_m))} \mu(b). \end{aligned}$$

Then by (4.1), we get

$$\begin{aligned} \mu(B_{n,k}^j) &\leq \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m)) \in \pi^{-1}(\mathcal{R})} \sum_{b \in \varphi^{-1}(\mathcal{R}, (u_m))} (K^8)^{k/n} \gamma^{k-j} \prod_{m=1}^{N+1} \mu(u_m) \\ &= (K^8)^{k/n} \gamma^{k-j} \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m)) \in \pi^{-1}(\mathcal{R})} \prod_{m=1}^{N+1} \mu(u_m). \end{aligned}$$

Since μ is a probability measure, the sum of $\mu(u_m)$ over any set of words u_m of the same length is less than or equal to one. Then

$$\begin{aligned} \mu(B_{n,k}^j) &\leq (K^8)^{k/n} \gamma^{k-j} \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m)) \in \pi^{-1}(\mathcal{R})} \prod_{m=1}^{N+1} \mu(u_m) \\ &\leq (K^8)^{k/n} \gamma^{k-j} |S| \\ &\leq (K^8 k^8)^{k/n} \gamma^{k-j}, \end{aligned}$$

where we have used that $|S| \leq (k^8)^{k/n}$ (established in the previous paragraph).

Recall that $k = o(n^2/\log(n))$. Therefore for large enough n , we have $k \leq n^2$. Let $p_1(x) = K^8 x^{16}$. Then for large enough n , for all $1 \leq j \leq \ell$, the previous display yields that $\mu(B_{n,k}^j) \leq p_1(n)^{k/n} \gamma^{k-j}$, as desired. \square

The following lemma gives both upper and lower bounds on the μ measure of $G_{n,k}$.

LEMMA 4.2. *There exists $\rho_0 > 0$ such that for all large enough n ,*

$$K^{-1} e^{-(h(\mu)+\delta)n} (1 - e^{-\rho_0 n}) \leq \mu(G_{n,k}) \leq K e^{-(h(\mu)-\delta)n}.$$

PROOF. For $u \in E_n$, let $G_{n,k}(u) = \{v \in G_{n,k} : v_1^n = u\}$. Then by our choice of K , for large enough n , we have

$$\mu(G_{n,k}(u)) = \sum_{v \in G_{n,k}(u)} \mu(v) \geq K^{-1} \mu(u)^2.$$

Also, note that $G_{n,k} = \bigsqcup_{u \in E_n} G_{n,k}(u)$. Then

$$\mu(G_{n,k}) = \sum_{u \in E_n} \mu(G_{n,k}(u)) \geq K^{-1} \sum_{u \in E_n} \mu(u)^2 \geq K^{-1} e^{-(h(\mu)+\delta)n} \mu(E_n),$$

where the last inequality results from the fact that $\min_{u \in E_n} \mu(u) \geq e^{-(h(\mu)+\delta)n}$. Additionally, using the Gibbs property (2.1) and the large deviations results for Gibbs measures [35], one may check that there exists $\rho_0 > 0$ such that $\mu(E_n) \geq 1 - e^{-\rho_0 n}$ for all large enough n . Combining this fact with the above inequalities yields the desired lower bound.

For the upper bound, for all large enough n and for each $u \in E_n$, we have that

$$\mu(G_{n,k}(u)) = \sum_{v \in G_{n,k}(u)} \mu(v) \leq K \mu(u)^2.$$

Then

$$\mu(G_{n,k}) = \sum_{u \in E_n} \mu(G_{n,k}(u)) \leq K \sum_{u \in E_n} \mu(u)^2 \leq K e^{-(h(\mu)-\delta)n},$$

where we have used that $\max_{u \in E_n} \mu(u) \leq e^{-(h(\mu)-\delta)n}$. \square

In the following three lemmas, we estimate the $\mu \times \mu$ measure of sets of pairs of words with various repeat properties. The outline of these proofs is similar to the outline of the proof of Lemma 4.1, but each proof requires some arguments that are specific to the particular repeat structure of interest.

LEMMA 4.3. *There exists a polynomial $p_2(x)$ such that for all large enough n , for all $1 \leq j \leq 2\ell - 1$,*

$$(\mu \times \mu)(D_{n,k}^j) \leq p_2(n)^{k/n} \gamma^{2\ell-j+n}.$$

PROOF. Consider $n \geq n_0$, and let $1 \leq j \leq 2\ell - 1$. First define the set $F = [0, k] \sqcup [k + 1, 2k]$. Then define a map $\varphi : D_{n,k}^j \rightarrow \{(\mathcal{R}, w) : \mathcal{R} \subset \mathcal{C}_n(F) \times \mathcal{C}_n(F), w \in \mathcal{A}^F \setminus A(\mathcal{R})\}$ as follows. Let $(a, b) \in D_{n,k}^j$. We use the notation $a \sqcup b$ to denote the element in \mathcal{A}^F such that $(a \sqcup b)|_{[0,k]} = a$ and $(a \sqcup b)|_{[k+1,2k]} = b$. Let \mathcal{R} be a repeat cover of $a \sqcup b$ such that $|\mathcal{R}| \leq 4|F|/n = 8k/n$, which exists by Lemma 2.5. Then let $((u_m)_{m=1}^{N_1+1}, (v_m)_{m=1}^{N_1})$ be the repeat block decomposition of a induced by the set $A(\mathcal{R}) \cap [0, k]$, and let $((y_m)_{m=1}^{N_2+1}, (z_m)_{m=1}^{N_2})$ be the repeat block decomposition of b induced by the set $A(\mathcal{R}) \cap [k + 1, 2k]$. Finally, we define $\varphi(a, b) = (\mathcal{R}, (u_m)_{m=1}^{N_1+1}, (y_m)_{m=1}^{N_2+1})$.

By Lemma 2.8, note that

$$\mu(a) \leq K^{2N_1} \prod_{m=1}^{N_1} \mu(v_m) \prod_{m=1}^{N_1+1} \mu(u_m),$$

and

$$\mu(b) \leq K^{2N_2} \prod_{m=1}^{N_2} \mu(z_m) \prod_{m=1}^{N_2+1} \mu(y_m).$$

Furthermore, since each of the blocks v_m and z_m has length at least n , for all large enough n , we have that $\mu(v_m) \leq \gamma^{|v_m|}$ and $\mu(z_m) \leq \gamma^{|z_m|}$. Therefore for all large enough n , we have

$$\mu(a) \leq K^{2N_1} \gamma^{\sum_m |v_m|} \prod_{m=1}^{N_1+1} \mu(u_m),$$

and

$$\mu(b) \leq K^{2N_2} \gamma^{\sum_m |z_m|} \prod_{m=1}^{N_2+1} \mu(y_m).$$

Using that $N_1 + N_2 \leq |\mathcal{R}| \leq 8k/n$ and $\sum_m |v_m| + \sum_m |z_m| = |A(\mathcal{R})| \geq 2k - j - n$ (by Lemma 2.6), we obtain

$$(4.2) \quad \mu(a)\mu(b) \leq (K^{16})^{k/n} \gamma^{2k-j-n} \prod_{m=1}^{N_1+1} \mu(u_m) \prod_{m=1}^{N_2+1} \mu(y_m).$$

Now define the projection $\pi : \varphi(D_{n,k}^j) \rightarrow \text{Power}(\mathcal{C}_n(F) \times \mathcal{C}_n(F))$, given by $\pi(\mathcal{R}, (u_m), (y_m)) = \mathcal{R}$. Let $S = \pi \circ \varphi(D_{n,k}^j)$. Note that $|\mathcal{C}_n(F)| \leq 2k$, and so $|\mathcal{C}_n(F) \times \mathcal{C}_n(F)| \leq (2k)^2$. Since each $\mathcal{R} \in \pi \circ \varphi(D_{n,k}^j)$ satisfies $|\mathcal{R}| \leq 8k/n$, we then have that $|S| = |\pi \circ \varphi(D_{n,k}^j)| \leq |\mathcal{C}_n(F) \times \mathcal{C}_n(F)|^{8k/n} \leq (2k)^{16k/n}$.

Let us now estimate $(\mu \times \mu)(D_{n,k}^j)$. By rearranging the sum, we get

$$\begin{aligned} (\mu \times \mu)(D_{n,k}^j) &= \sum_{(a,b) \in D_{n,k}^j} \mu(a)\mu(b) \\ &= \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m), (y_m)) \in \pi^{-1}(\mathcal{R})} \sum_{(a,b) \in \varphi^{-1}(\mathcal{R}, (u_m), (y_m))} \mu(a)\mu(b). \end{aligned}$$

Applying (4.2) to each term in the sum, we get

$$\begin{aligned}
 & (\mu \times \mu)(D_{n,k}^j) \\
 & \leq (K^{16})^{k/n} \gamma^{2k-j-n} \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m), (y_m)) \in \pi^{-1}(\mathcal{R})} \prod_{m=1}^{N_1+1} \mu(u_m) \prod_{m=1}^{N_2+1} \mu(y_m).
 \end{aligned}$$

Then since the sum of $\mu(u_m)$ over any set of words u_m of the same length is less than or equal to one, we see that

$$(\mu \times \mu)(D_{n,k}^j) \leq (K^{16})^{k/n} \gamma^{2k-j-n} |S|.$$

Combining this estimate with the bound on $|S|$ established in the previous paragraph, we obtain

$$(\mu \times \mu)(D_{n,k}^j) \leq (2^{16} K^{16} k^{16})^{k/n} \gamma^{2k-j-n}.$$

Recall that $k = o(n^2 / \log(n))$. Then for all large enough n , we have $k \leq n^2$. Let $p_2(x) = (2K)^{16} x^{32}$. Then by the previous display, for all large enough n , we obtain that $(\mu \times \mu)(D_{n,k}^j) \leq p_2(n)^{k/n} \gamma^{2k-j-n}$. \square

LEMMA 4.4. *There exists a polynomial $p_3(x)$ such that for all large enough n ,*

$$(\mu \times \mu)(Q_{n,k}) \leq p_3(n) e^{-2n(h(\mu)-\delta)} \gamma^n.$$

PROOF. Consider $n \geq n_0$. First define the set $F = [0, k] \sqcup [k+1, 2k]$. Then we define a map $\varphi : Q_{n,k} \rightarrow \{(\mathcal{R}, w) : \mathcal{J} \subset \mathcal{C}_n(S) \times \mathcal{C}_n(S), w \in \mathcal{A}^{S \setminus A(\mathcal{R})}\}$ as follows. Let $(a, b) \in Q_{n,k}$. We let $a \sqcup b$ denote the element of \mathcal{A}^F such that $(a \sqcup b)|_{[0,k]} = a$ and $(a \sqcup b)|_{[k+1,2k]} = b$. Since $W_n(a) \cap W_n(b) \neq \emptyset$, there exists $I \in \mathcal{C}_n([0, k])$ and $J \in \mathcal{C}_n([k+1, 2k])$ such that $(a \sqcup b)|_I = (a \sqcup b)|_J$, and we assume that J is (lexicographically) minimal among all such intervals. We partition $Q_{n,k}$ into $\tilde{Q}_{n,k}$ and $\hat{Q}_{n,k}$, where $\tilde{Q}_{n,k}$ consists of all pairs (a, b) such that $J \cap [2\ell - n, 2k] = \emptyset$, and $\hat{Q}_{n,k}$ contains the remaining pairs. Our definition of $\varphi(a, b)$ will depend on whether (a, b) is in $\tilde{Q}_{n,k}$ or $\hat{Q}_{n,k}$.

First suppose that $(a, b) \in \tilde{Q}_{n,k}$. Let $\mathcal{R} \subset \mathcal{C}_n(F) \times \mathcal{C}_n(F)$ be the following set containing three pairs of intervals: $\{([0, n], [\ell-1, k]), (I, J), ([k+1, k+n], [2k-n+1, 2k])\}$. Note that the block decomposition of a induced by $A(\mathcal{R}) \cap [0, k]$ has the form (u_1, v_1) , where $v_1 = a|_{[\ell-1, k]}$. Similarly, the block decomposition of b induced by $A(\mathcal{R}) \cap [k+1, 2k]$ has the form $((y_m)_{m=1}^2, (z_m)_{m=1}^2)$, where $z_1 = (a \sqcup b)|_J$ and $z_2 = (a \sqcup b)|_{[2k-n+1, 2k]}$. Finally, we define $\varphi(a, b) = (\mathcal{R}, u_1, (y_m)_{m=1}^2)$. Furthermore, we note that

$$\mu(a) = \mu(u_1 v_1) \leq K \mu(u_1) \mu(v_1),$$

and

$$\mu(b) = \mu(y_1 z_1 y_2 z_2) \leq K^3 \mu(y_1) \mu(y_2) \mu(z_1) \mu(z_2).$$

Since $a, b \in G_{n,k}$, we must have that $v_1, z_2 \in E_n$, and therefore $\mu(v_1) \leq e^{-(h(\mu)-\delta)n}$ and $\mu(z_2) \leq e^{-(h(\mu)-\delta)n}$. Also, since z_1 has length $n \geq n_0$, we have that $\mu(z_1) \leq \gamma^n$. Putting these estimates together, we obtain

$$(4.3) \quad \mu(a) \mu(b) \leq K^4 e^{-2n(h(\mu)-\delta)} \gamma^n \mu(u_1) \mu(y_1) \mu(y_2).$$

Now suppose that $(a, b) \in \hat{Q}_{n,k}$. In this case we let $\mathcal{R} \subset \mathcal{C}_n(F) \times \mathcal{C}_n(F)$ be a different set of three pairs of intervals: $\mathcal{R} = \{([0, n], [\ell-1, k]), (I, J), ([2k -$

$n + 1, 2k], [k + 1, k + n]\}$. Note that the third pair of intervals listed is not in lexicographical order. Let (u_1, v_1) be the repeat block decomposition of a induced by $A(\mathcal{R}) \cap [0, k)$, and let $((z_m)_{m=1}^2, (y_m)_{m=1}^2)$ be the block decomposition induced by $A(\mathcal{R})|_{[k+1, 2k]}$, by which we mean that $(a \sqcup b)|_{[k+1, 2k]} = z_1 y_1 z_2 y_2$, where z_1 and z_2 have length n . In this case, we define $\varphi(a, b) = (\mathcal{R}, u_1, (y_m)_{m=1}^2)$. Note that

$$\mu(a) = \mu(u_1 v_1) \leq K\mu(u_1)\mu(v_1),$$

and

$$\mu(b) = \mu(z_1 y_1 z_2 y_2) \leq K^3 \mu(y_1)\mu(y_2)\mu(z_1)\mu(z_2).$$

Since $a, b \in G_{n,k}$, we have that $v_1, z_1 \in E_n$, and thus $\mu(v_1) \leq e^{-(h(\mu)-\delta)n}$ and $\mu(z_1) \leq e^{-(h(\mu)-\delta)n}$. Also, since z_2 has length $n \geq n_0$, we have that $\mu(z_2) \leq \gamma^n$. Combining these estimates, we see that

$$(4.4) \quad \mu(a)\mu(b) \leq K^4 e^{-2n(h(\mu)-\delta)} \gamma^n \mu(u_1)\mu(y_1)\mu(y_2).$$

Now define the projection map $\pi : \varphi(Q_{n,k}) \rightarrow \text{Power}(\mathcal{C}_n(F) \times \mathcal{C}_n(F))$, given by $\pi(\mathcal{R}, w) = \mathcal{R}$. Let $S = \pi \circ \varphi(Q_{n,k})$. Since $|\mathcal{C}_n(F)| \leq (2k)$, we get $|\mathcal{C}_n(F) \times \mathcal{C}_n(F)| \leq (2k)^2$. Moreover, since each \mathcal{R} in $\pi \circ \varphi(Q_{n,k})$ satisfies $|\mathcal{R}| = 3$, we get $|S| = |\pi \circ \varphi(Q_{n,k})| \leq |\mathcal{C}_n(F) \times \mathcal{C}_n(F)|^3 \leq (2k)^6$.

By rearranging the sum, we find

$$\begin{aligned} (\mu \times \mu)(Q_{n,k}) &= \sum_{(a,b) \in Q_{n,k}} \mu(a)\mu(b) \\ &= \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, u_1, (y_m)_{m=1}^2) \in \pi^{-1}(\mathcal{R})} \sum_{(a,b) \in \varphi^{-1}(\mathcal{R}, u_1, (y_m)_{m=1}^2)} \mu(a)\mu(b). \end{aligned}$$

Then by applying the estimates (4.3) and (4.4) to each term, we get

$$(\mu \times \mu)(Q_{n,k}) \leq \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, u_1, (y_m)_{m=1}^2) \in \pi^{-1}(\mathcal{R})} K^4 e^{-2n(h(\mu)-\delta)} \gamma^n \mu(u_1)\mu(y_1)\mu(y_2).$$

Summing over all u_1, y_1 , and y_2 , we obtain

$$(\mu \times \mu)(Q_{n,k}) \leq K^4 e^{-2n(h(\mu)-\delta)} \gamma^n |S| \leq K^4 (2k)^6 e^{-2n(h(\mu)-\delta)} \gamma^n,$$

where the second inequality uses the bound on $|S|$ established in the previous paragraph

Recall that $k = o(n^2 / \log(n))$, and hence for all large enough n , we have $k \leq n^2$. Let $p_3(x) = 2^6 K^4 x^{12}$. Then by the previous inequality, for all large enough n , we see that $(\mu \times \mu)(Q_{n,k}) \leq p_3(n) e^{-2n(h(\mu)-\delta)} \gamma^n$. \square

LEMMA 4.5. *There exists a polynomial $p_4(x)$ such that for all large enough n and $1 \leq j \leq 2\ell$,*

$$(\mu \times \mu)(Q_{n,k}^j) \leq p_4(k)^{k/n} e^{-2n(h(\mu)-\delta)} \gamma^{2\ell-j}.$$

PROOF. Consider $n \geq n_0$ and $1 \leq j \leq 2\ell$. Let $F = [0, k) \cup [k + 1, 2k]$. We begin by defining a map $\varphi : Q_{n,k}^j \rightarrow \{(\mathcal{R}, w) : \mathcal{R} \subset \mathcal{C}_n(S) \times \mathcal{C}_n(S), w \in \mathcal{A}^{F \setminus A(\mathcal{R})}\}$ as follows. Let $(a, b) \in Q_{n,k}^j$. We let $(a \sqcup b)$ denote the element of \mathcal{A}^F such that $(a \sqcup b)|_{[0,k)} = a$ and $(a \sqcup b)|_{[k+1, 2k]} = b$. Let \mathcal{R} be a repeat cover for $a \sqcup b$ such that $|\mathcal{R}| \leq 4|F|/n = 8k/n$, which exists by Lemma 2.5. Then let $((u_m)_{m=1}^{N_1+1}, (v_m)_{m=1}^{N_1})$ be the repeat block decomposition of a induced by the set $A(\mathcal{R}) \cap [0, k - n)$, and let $((y_m)_{m=1}^{N_2+1}, (z_m)_{m=1}^{N_2})$ be the repeat block decomposition for b induced by the set

$A(\mathcal{R}) \cap [k+1, 2k-n]$. Additionally, let $v_{N_1+1} = a_{[k-n, k]}$ and $z_{N_2+1} = b_{[2k-n+1, 2k]}$. Set $\varphi(a, b) = (\mathcal{R}, (u_m)_{m=1}^{N_1+1}, (y_m)_{m=1}^{N_2+1})$. Furthermore, by Lemma 2.8, we have that

$$\mu(a) \leq K^{2N_1+1} \prod_{m=1}^{N_1+1} \mu(v_m) \prod_{m=1}^{N_1+1} \mu(u_m),$$

and

$$\mu(b) \leq K^{2N_2+1} \prod_{m=1}^{N_2+1} \mu(z_m) \prod_{m=1}^{N_2+1} \mu(y_m).$$

Since $v_{N_1+1}, z_{N_2+1} \in E_n$, we have $\mu(v_{N_1+1}) \leq e^{-n(h(\mu)-\delta)}$ and $\mu(z_{N_2+1}) \leq e^{-n(h(\mu)-\delta)}$. Also, for $m = 1, \dots, N_1-1$, the length of v_m is at least n , and we get $\mu(v_m) \leq \gamma^{|v_m|}$. For v_{N_1} , we always have $\mu(v_{N_1}) \leq \gamma^{|v_{N_1}| - n_0}$. Similarly, for $m = 1, \dots, N_2-1$, the length of z_m is at least n , and we get $\mu(z_m) \leq \gamma^{|z_m|}$. As for z_{N_2} , we always have $\mu(z_{N_2}) \leq \gamma^{|z_{N_2}| - n_0}$. Then for large enough n , we have

$$\mu(a) \leq K^{2N_1+1} \gamma^{\sum_m |v_m| - n_0} e^{-n(h(\mu)-\delta)} \prod_{m=1}^{N_1+1} \mu(u_m),$$

and

$$\mu(b) \leq K^{2N_2+1} \gamma^{\sum_m |z_m| - n_0} e^{-n(h(\mu)-\delta)} \prod_{m=1}^{N_2+1} \mu(y_m).$$

Using that $N_1 + N_2 \leq |\mathcal{R}| \leq 8k/n$ and $\sum_m |v_m| + \sum_m |z_m| = |A(\mathcal{R})| \geq 2\ell - j - 2$ (by Lemma 2.6), we obtain

$$(4.5) \quad \mu(a)\mu(b) \leq (K^{32})^{k/n} \gamma^{2\ell-j-2-2n_0} e^{-2n(h(\mu)-\delta)} \prod_{m=1}^{N_1+1} \mu(u_m) \prod_{m=1}^{N_2+1} \mu(y_m).$$

Now we define the projection map $\pi : \varphi(Q_{n,k}^j) \rightarrow \text{Power}(\mathcal{C}_n(F) \times \mathcal{C}_n(F))$, given by $\pi(\mathcal{R}, (u_m), (y_m)) = \mathcal{R}$. Let $S = \pi \circ \varphi(Q_{n,k}^j)$. Since $|\mathcal{C}_n(F)| \leq 2k$, we see that $|\mathcal{C}_n(F) \times \mathcal{C}_n(F)| \leq (2k)^2$. Moreover, since each \mathcal{R} in S satisfies $|\mathcal{R}| \leq 8k/n$, we estimate $|S| = |\pi \circ \varphi(Q_{n,k}^j)| \leq |\mathcal{C}_n(F) \times \mathcal{C}_n(F)|^{8k/n} \leq (2^{16} k^{16})^{k/n}$.

Let us now estimate $(\mu \times \mu)(Q_{n,k}^j)$. By rearranging the sum, we get

$$\begin{aligned} (\mu \times \mu)(Q_{n,k}^j) &= \sum_{(a,b) \in Q_{n,k}^j} \mu(a)\mu(b) \\ &= \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m), (y_m)) \in \pi^{-1}(\mathcal{R})} \sum_{(a,b) \in \varphi^{-1}(\mathcal{R}, (u_m), (y_m))} \mu(a)\mu(b). \end{aligned}$$

By applying (4.5) to each term in the sum, we see that

$$\begin{aligned} &(\mu \times \mu)(Q_{n,k}^j) \\ &\leq \sum_{\mathcal{R} \in S} \sum_{(\mathcal{R}, (u_m), (y_m)) \in \pi^{-1}(\mathcal{R})} (K^{32} \gamma^{-2n_0-2})^{k/n} \gamma^{2\ell-j} e^{-n(h(\mu)-\delta)} \prod_{m=1}^{N_1+1} \mu(u_m) \prod_{m=1}^{N_2+1} \mu(y_m), \end{aligned}$$

and then summing over all u_m and y_m gives

$$(4.6) \quad (\mu \times \mu)(Q_{n,k}^j) \leq (K^{32} \gamma^{-2n_0-2})^{k/n} e^{-n(h(\mu)-\delta)} \gamma^{2\ell-j} |S|.$$

Combining this estimate with the bound on $|S|$ from the previous paragraph, we obtain

$$(\mu \times \mu)(Q_{n,k}^j) \leq (2^{32} K^{32} \gamma^{-2n_0-2} k^{16})^{k/n} e^{-n(h(\mu)-\delta)\gamma^{2\ell-j}}.$$

Recall that $k = o(n^2/\log(n))$, and hence for all large enough n , we have $k \leq n^2$. Let $p_4(x) = (2K)^{32}\gamma^{-2n_0-2}x^{32}$. Then by the previous displayed inequality, for all large enough n and all $1 \leq j \leq 2\ell$, we have $(\mu \times \mu)(Q_{n,k}^j) \leq p_4(n)^{k/n} e^{-n(h(\mu)-\delta)\gamma^{2\ell-j}}$. \square

5. Moment bounds

In this section we prove properties (I)-(IV) concerning the expectation and variance of $\phi_{n,k}$ and $\psi_{n,k}$, which are used in the proof of Theorem 1.1. Throughout this section, we use the same environment (notation, parameters, and assumptions) as in the proof of Theorem 1.1.

LEMMA 5.1. *For all $n \geq 1$, the expectation of $\phi_{n,k}$ satisfies*

$$\mathbb{E}[\phi_{n,k}] \geq K^{-1}\alpha^\ell e^{Pk}.$$

Furthermore,

$$\lim_n \frac{1}{k} \log \mathbb{E}[\phi_{n,k}] = P + \log(\alpha).$$

PROOF. Let $n \geq 1$. Then by (2.2) and our choice of K , we have

$$\begin{aligned} \mathbb{E}[\phi_{n,k}] &= \sum_{u \in B_k(X)} e^{S_k f(u)} \alpha^{|W_n(u)|} \\ &= \sum_{j=1}^{\ell} \alpha^j \sum_{u \in B_{n,k}^j} e^{S_k f(u)} \\ &\geq K^{-1} e^{Pk} \sum_{j=1}^{\ell} \alpha^j \mu(B_{n,k}^j). \end{aligned}$$

Using that $\alpha^j \geq \alpha^\ell$ for all $j \leq \ell$ and $\sum_j \mu(B_{n,k}^j) = \mu(B_n(X)) = 1$, we get

$$\mathbb{E}[\phi_{n,k}] \geq K^{-1}\alpha^\ell e^{Pk},$$

which establishes the first conclusion of the lemma.

Now we consider letting n tend to infinity. By the first conclusion of the lemma, we have that

$$\liminf_n \frac{1}{k} \log \mathbb{E}[\phi_{n,k}] \geq P + \log \alpha.$$

Also, for any n , our choice of K yields

$$\mathbb{E}[\phi_{n,k}] = \sum_{j=1}^{k-n+1} \alpha^j \sum_{u \in B_{n,k}^j} e^{S_k f(u)} \leq K e^{Pk} \sum_{j=1}^{k-n+1} \alpha^j \mu(B_{n,k}^j).$$

Recall that $\alpha > \gamma$, and therefore $\alpha\gamma^{-1} > 1$. By Lemma 4.1, there exists a polynomial $p_1(x)$ such that for large enough n , we have

$$\begin{aligned} \mathbb{E}[\phi_{n,k}] &\leq Ke^{Pk} \sum_{j=1}^{\ell} \alpha^j \mu(B_{n,k}^j) \\ &\leq Ke^{Pk} \sum_{j=1}^{\ell} \alpha^j p_1(n)^{k/n} \gamma^{k-j} \\ &\leq Ke^{Pk} p_1(n)^{k/n} \gamma^k \sum_{j=1}^{\ell} (\alpha\gamma^{-1})^j \\ &\leq Ke^{Pk} p_1(n)^{k/n} \gamma^k (\alpha\gamma^{-1})^{\ell} \frac{1}{1 - (\alpha^{-1}\gamma)} \\ &= Ke^{Pk} p_1(n)^{k/n} \alpha^{\ell} \gamma^n \frac{1}{1 - (\alpha^{-1}\gamma)}. \end{aligned}$$

Since $n/k \rightarrow 0$ and $n^{-1} \log p_1(n) \rightarrow 0$, we obtain that

$$\limsup_n \frac{1}{k} \log \mathbb{E}[\phi_{n,k}] \leq P + \log \alpha,$$

which finishes the proof. □

LEMMA 5.2. *For all $n \geq 1$, the expectation of $\psi_{n,k}$ satisfies*

$$\mathbb{E}[\psi_{n,k}] \geq |E_n|^{-1} K^{-1} \alpha^{\ell} e^{Pk} \mu(G_{n,k}).$$

Furthermore,

$$\lim_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] = P + \log(\alpha).$$

PROOF. Let $n \geq 1$. Then by (2.2) and our choice of K , we have

$$\begin{aligned} \mathbb{E}[\psi_{n,k}] &= \frac{1}{|E_n|} \sum_{u \in G_{n,k}} e^{S_k f(u)} \alpha^{|W_n(u)|} \\ &= \frac{1}{|E_n|} \sum_{j=1}^{\ell} \alpha^j \sum_{u \in G_{n,k}^j} e^{S_k f(u)} \\ &\geq |E_n|^{-1} K^{-1} e^{Pk} \sum_{j=1}^{\ell} \alpha^j \mu(G_{n,k}^j). \end{aligned}$$

Since $\alpha^j \geq \alpha^{\ell}$ for all $j \leq \ell$ and $\sum_j \mu(G_{n,k}^j) = \mu(G_{n,k})$, we get

$$\mathbb{E}[\psi_{n,k}] \geq |E_n|^{-1} K^{-1} \alpha^{\ell} e^{Pk} \mu(G_{n,k}),$$

which establishes the first conclusion of the lemma.

Now we consider letting n tend to infinity. By the first conclusion of the lemma, we have that

$$\liminf_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] \geq P + \log \alpha + \liminf_n \frac{1}{k} \log |E_n|^{-1} + \frac{1}{k} \log \mu(G_{n,k}).$$

Note that $|E_n| \leq |\mathcal{A}|^n$, and by Lemma 4.2, for large enough n , we have $\mu(G_{n,k}) \geq 2^{-1}K^{-1}e^{-n(h(\mu)+\delta)}$. Therefore

$$\liminf_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] \geq P + \log \alpha + \liminf_n \left(\frac{n}{k} \log |\mathcal{A}|^{-1} + \frac{n}{k} \log(e^{-(h(\mu)+\delta)}) \right).$$

Finally, using that $n/k \rightarrow 0$, we obtain

$$\liminf_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] \geq P + \log \alpha.$$

Also, by Lemma 5.1 and the fact that $\psi_{n,k} \leq \phi_{n,k}$, we have

$$\limsup_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] \leq \limsup_n \frac{1}{k} \log \mathbb{E}[\phi_{n,k}] \leq P + \log \alpha.$$

Taken together, the previous two inequalities yield that

$$\lim_n \frac{1}{k} \log \mathbb{E}[\psi_{n,k}] = P + \log \alpha,$$

as desired. □

LEMMA 5.3. *There exists $\rho_1 > 0$ such that for all large enough n ,*

$$\frac{\text{Var}[\phi_{n,k}]}{\mathbb{E}[\phi_{n,k}]^2} \leq e^{-\rho_1 n};$$

PROOF. Using the fact that the variance of a sum is the sum of the covariances, (2.3), and our choice of K , we have

$$\begin{aligned} \text{Var}[\phi_{n,k}] &= \sum_{u,v \in B_k(X)} \alpha^{|W_n(u) \cup W_n(v)|} (1 - \alpha^{|W_n(u) \cap W_n(v)|}) e^{S_k f(u) + S_k f(v)} \\ &\leq \sum_{j=1}^{2\ell-1} \alpha^j \sum_{(u,v) \in D_{n,k}^j} e^{S_k f(u) + S_k f(v)} \\ &\leq K^2 e^{2Pk} \sum_{j=1}^{2\ell-1} \alpha^j (\mu \times \mu)(D_{n,k}^j). \end{aligned}$$

Let $C = 1/(1 - (\alpha^{-1}\gamma))$. Then by the lower bound on $\mathbb{E}[\phi_{n,k}]$ from Lemma 5.1 and the upper bound on $(\mu \times \mu)(D_{n,k}^j)$ from Lemma 4.3, there exists a polynomial $p_2(x)$ such that for large enough n , we have

$$\begin{aligned} \frac{\text{Var}[\phi_{n,k}]}{\mathbb{E}[\phi_{n,k}]^2} &\leq \frac{K^2 e^{2Pk} \sum_{j=1}^{2\ell-1} \alpha^j (\mu \times \mu)(D_{n,k}^j)}{K^{-2} \alpha^{2\ell} e^{2Pk}} \\ &= K^4 \alpha^{-2\ell} \sum_{j=1}^{2\ell-1} \alpha^j (\mu \times \mu)(D_{n,k}^j) \\ &\leq K^4 \alpha^{-2\ell} \sum_{j=1}^{2\ell-1} \alpha^j p_2(n)^{k/n} \gamma^{2\ell+n-j} \\ &\leq K^4 p_2(n)^{k/n} \alpha^{-2\ell} \gamma^{2\ell+n} \sum_{j=1}^{2\ell-1} (\alpha\gamma^{-1})^j \\ &\leq K^4 p_2(n)^{k/n} \alpha^{-2\ell} \gamma^{2\ell+n} (\alpha\gamma^{-1})^{2\ell} C. \end{aligned}$$

Rewriting this estimate, we find

$$\begin{aligned} \frac{\text{Var}[\phi_{n,k}]}{\mathbb{E}[\phi_{n,k}]^2} &= \exp\left(4\log K + \log C + \frac{k}{n}\log p_2(n) + n\log\gamma\right) \\ &\leq \exp\left(n\left(q\frac{k\log n}{n^2} + \log\gamma + \frac{4}{n}\log K + \frac{1}{n}\log C\right)\right), \end{aligned}$$

where $p_2(x) \leq x^q$ for all large enough x . Since $k = o(n^2/\log n)$ and $\log\gamma < 0$, we obtain the desired bound. \square

LEMMA 5.4. *There exists $\rho_2 > 0$ such that for all large enough n ,*

$$\frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} \leq e^{-\rho_2 n}.$$

PROOF. Let $b = 2\ell - n$. Using the fact that the variance of a sum is the sum of the covariances, (2.3), and our choice of K , we have

$$\begin{aligned} \text{Var}[\psi_{n,k}] &= \frac{1}{|E_n|^2} \sum_{u,v \in G_{n,k}} \alpha^{|W_n(u) \cup W_n(v)|} \left(1 - \alpha^{|W_n(u) \cap W_n(v)|}\right) e^{S_k f(u) + S_k f(v)} \\ &\leq \frac{1}{|E_n|^2} \sum_{j=1}^{2\ell} \alpha^j \sum_{(u,v) \in Q_{n,k}^j} e^{S_k f(u) + S_k f(v)} \\ &\leq K^2 \frac{e^{2Pk}}{|E_n|^2} \sum_{j=1}^{2\ell} \alpha^j (\mu \times \mu)(Q_{n,k}^j). \end{aligned}$$

Dividing by $\mathbb{E}[\psi_{n,k}]$ and using the lower bound on $\mathbb{E}[\psi_{n,k}]$ in Lemma 5.2 and the lower bound on $\mu(G_{n,k})$ in Lemma 4.2, we see that

$$\begin{aligned} \frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} &\leq \frac{K^2 e^{2Pk} \sum_{j=1}^{2\ell-1} \alpha^j (\mu \times \mu)(Q_{n,k}^j)}{K^{-2} \alpha^{2\ell} e^{2Pk} \mu(G_{n,k})^2} \\ &\leq \frac{K^4 \alpha^{-2\ell} \sum_{j=1}^{2\ell-1} \alpha^j (\mu \times \mu)(Q_{n,k}^j)}{(2K)^{-2} e^{-2n(h(\mu)+\delta)}} \\ &= (2K)^6 \alpha^{-2\ell} e^{2n(h(\mu)+\delta)} \left(\sum_{j=1}^{b-1} \alpha^j (\mu \times \mu)(Q_{n,k}^j) + \sum_{j=b}^{2\ell-1} \alpha^j (\mu \times \mu)(Q_{n,k}^j) \right). \end{aligned}$$

Let $C = 1/(1 - (\alpha^{-1}\gamma))$. Note that $\alpha^j \leq \alpha^b$ for $j \geq b$. Applying this fact and the upper bounds on $(\mu \times \mu)(Q_{n,k})$ and $(\mu \times \mu)(Q_{n,k}^j)$ from Lemmas 4.4 and 4.5, respectively, we get that there are polynomials $p_3(x)$ and $p_4(x)$ such that for all

large enough n ,

$$\begin{aligned} \frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} &\leq (2K)^6 \alpha^{-2\ell} e^{2n(h(\mu)+\delta)} \\ &\quad \times \left(\sum_{j=1}^{b-1} \alpha^j p_4(n)^{k/n} e^{-2n(h(\mu)-\delta)} \gamma^{2\ell-j} + \alpha^b p_3(n) e^{-2n(h(\mu)-\delta)} \gamma^n \right) \\ &\leq (2K)^6 \alpha^{-2\ell} e^{4n\delta} \left(p_4(n)^{k/n} \gamma^{2\ell} \sum_{j=1}^{b-1} (\alpha \gamma^{-1})^j + \alpha^b \gamma^n p_3(n) \right) \\ &\leq (2K)^6 \alpha^{-2\ell} e^{4n\delta} \left(p_4(n)^{k/n} \gamma^{2\ell} C(\alpha \gamma^{-1})^b + \alpha^b \gamma^n p_3(n) \right) \\ &= (2K)^6 e^{4n\delta} \gamma^n \alpha^{-n} \left(C p_4(n)^{k/n} + p_3(n) \right). \end{aligned}$$

Rewriting this estimate, we have

$$\frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} \leq \exp \left(n \left(\log(\gamma \alpha^{-1}) + 4\delta + \frac{6}{n} \log(2K) + \frac{1}{n} \log(C p_4(n)^{k/n} + p_3(n)) \right) \right).$$

Let $q > 1$ be such that for all large enough x , we have $C p_4(x)^{k/n} + p_3(x) \leq x^{qk/n}$. Then for all large enough n , we get

$$\frac{\text{Var}[\psi_{n,k}]}{\mathbb{E}[\psi_{n,k}]^2} \leq \exp \left(n \left(\log(\gamma \alpha^{-1}) + 4\delta + \frac{6}{n} \log(2K) + \frac{qk}{n^2} \log n \right) \right).$$

Since $k = o(n^2 / \log(n))$, and $\log(\gamma \alpha^{-1}) + 4\delta < 0$ (by our choice of δ in the proof of Theorem 1.1), we obtain the desired bound. \square

6. Bounds on pressure

In this section we work with the same notation, parameters, and assumptions as in the proof of Theorem 1.1.

LEMMA 6.1. *For each n ,*

$$P_{Y_n}(f) \leq \frac{1}{k} \log \phi_{n,k}.$$

PROOF. By subadditivity in the definition of pressure, we have that for all $m \geq 1$,

$$P_{Y_n}(f) \leq \frac{1}{m} \log \sum_{u \in B_m(Y_n)} e^{S_m f(u)}.$$

We apply this inequality with $m = k$. Also, since $B_k(Y_n) \subset \{u \in B_k(X) : \xi_u = 1\}$, we have

$$\begin{aligned} \frac{1}{k} \log \sum_{u \in B_k(Y_n)} e^{S_m f(u)} &\leq \frac{1}{k} \log \sum_{u \in B_k(X)} e^{S_k f(u)} \xi_u \\ &= \frac{1}{k} \log \phi_{n,k}. \end{aligned}$$

Combining the two previous inequalities yields the desired conclusion. \square

LEMMA 6.2. For any $\epsilon > 0$, for all large enough n ,

$$\frac{1}{k} \log \psi_{n,k} - \epsilon/2 \leq P_{Y_n}(f).$$

PROOF. Let $\mathcal{F} = \mathcal{F}_n$ and $Y = Y_n$. For $v \in E_n$, and $m \geq n$, we let

$$Z_m(v) = \left\{ u \in B_m(X) : W_n(u) \cap \mathcal{F} = \emptyset, \text{ and } \forall q \in \{0, \dots, \lfloor m/\ell - 1 \rfloor\}, u_{q\ell+1}^{q\ell+n} = v \right\}.$$

Note that $\psi_{n,k}$ may be viewed as an average over the set E_n :

$$\psi_{n,k} = \frac{1}{|E_n|} \sum_{v \in E_n} \sum_{u \in Z_k(v)} e^{S_k f(u)}.$$

Since the average over a finite set is always less than or equal to the maximum, there exists $v \in E_n$ such that

$$\psi_{n,k} \leq \sum_{u \in Z_k(v)} e^{S_k f(u)}.$$

For the sake of this proof, if $u \in B_m(X)$, then we let $\underline{S}_m f(u) = \inf_{x \in [u]} \sum_{j=0}^{m-1} f \circ \sigma^j(x)$. Observe that elements of $Z_\ell(v)$ can be arbitrarily concatenated to form words in Y . Hence, for any $q \in \mathbb{N}$, we note that $Z_{q\ell}(v) \subset B_{q\ell}(Y)$, and then we have

$$\begin{aligned} \sum_{u \in B_{q\ell}(Y)} e^{S_{q\ell} f(u)} &\geq \sum_{u \in Z_{q\ell}(v)} e^{S_{q\ell} f(u)} \\ &\geq \sum_{u \in Z_{q\ell}(v)} e^{\underline{S}_{q\ell} f(u)}. \end{aligned}$$

Then by our choice of K , we get

$$\begin{aligned} \sum_{u \in B_{q\ell}(Y)} e^{S_{q\ell} f(u)} &\geq K^{-q} \sum_{u_0 \dots u_{q-1} \in Z_{q\ell}(v)} e^{\sum_{i=0}^{q-1} \underline{S}_\ell f(u_i)} \\ &= K^{-q} \sum_{u_0 \in Z_\ell(v)} \dots \sum_{u_{q-1} \in Z_\ell(v)} e^{\sum_{i=0}^{q-1} \underline{S}_\ell f(u_i)} \\ &= K^{-q} \left(\sum_{u \in Z_\ell(v)} e^{\underline{S}_\ell(u)} \right)^q \\ &\geq K^{-2q} \left(\sum_{u \in Z_\ell(v)} e^{S_\ell(u)} \right)^q \\ &\geq K^{-2q} \left(\sum_{u \in Z_k(v)} e^{S_k(u)} \right)^q e^{-\|f\|_\infty nq}, \end{aligned}$$

where $\|f\|_\infty = \sup_{x \in X} |f(x)|$. Now take logarithm, divide by $q\ell$, and let q tend to infinity:

$$P_Y(f) \geq \frac{1}{\ell} \log \sum_{u \in Z_k(v)} e^{S_k f(u)} - \frac{2q \log K}{\ell} - \|f\|_\infty \frac{n}{\ell}.$$

Then

$$P_Y(f) \geq \frac{1}{k} \log \psi_{n,k} - \frac{2q \log K}{\ell} - \|f\|_\infty \frac{n}{\ell}.$$

Finally, since $n/\ell \rightarrow 0$, we may choose n large enough that

$$\frac{2q \log K}{\ell} + \|f\|_\infty \frac{n}{\ell} < \epsilon/2,$$

which finishes the proof of the lemma. □

7. Connection between pressure and escape rate

Here we relate the notions of pressure and escape rate. For a hole H in an SFT X , we define the *survivor set* to be the set of points that never fall into the hole (in either forward or backward time):

$$Y = X \setminus \left(\bigcup_{m \in \mathbb{Z}} \sigma^{-m}(H) \right).$$

For an SFT (X, σ) , a hole H consisting of a finite union of cylinder sets, and an equilibrium state μ associated to a Hölder continuous potential function f , the following proposition relates the escape rate of μ through the hole H to the pressure of f on the survivor set Y . Although various versions of this result appear to be well-known (see, e.g., [7, 13]), we could not find an explicit reference for it, and we include a proof for completeness. For analogous results in various smooth settings, see the discussion of the *escape rate formula* in [4] and references therein.

PROPOSITION 7.1. *Let X be a non-trivial mixing SFT, $f : X \rightarrow \mathbb{R}$ a Hölder continuous potential, and μ the Gibbs measure associated to f . Further, let H be a finite union of cylinder sets in X , and let Y be the survivor set of the open system (X, σ, H) . Then*

$$-\varrho(\mu : H) = P_X(f) - P_Y(f).$$

PROOF. Let K satisfy the conclusions of Lemma 2.2 for X , f , and μ . We suppose without loss of generality that H is the union of cylinder sets corresponding to words of length n . For $k \geq n$, let $B_k(X, H)$ denote the set of $w \in B_k(X)$ such that w contains no subword in H , and let $P = P_X(f)$. Let $\ell = \ell(k) = k - n + 1$. Recall that

$$M_\ell = \left\{ x \in X : \forall j \in \{0, \dots, \ell - 1\}, \sigma^j(x) \notin H \right\},$$

so that we have $\mu(M_\ell) = \mu(B_k(X, H))$. Note that since n is fixed in this context, we have $\lim_{k \rightarrow \infty} \ell/k = 1$.

By our choice of K and the fact that $B_k(Y) \subset B_k(X, H)$, we have that

$$\begin{aligned} \mu(B_k(X, H)) &= \sum_{w \in B_k(X, H)} \mu(w) \\ &\geq \sum_{w \in B_k(X, H)} K^{-1} e^{-Pk + S_k f(w)} \\ &= K^{-1} e^{-Pk} \sum_{w \in B_k(X, H)} e^{S_k f(w)} \\ &\geq K^{-1} e^{-Pk} \sum_{w \in B_k(Y)} e^{S_k f(w)}. \end{aligned}$$

It follows that

$$\frac{1}{k} \log \mu(B_k(X, H)) \geq -P + \frac{1}{k} \log \Lambda_k(Y) - \frac{1}{k} \log K,$$

and letting k tend to infinity, we see that

$$(7.1) \quad \liminf_k \frac{1}{k} \log \mu(M_\ell) \geq -P + \liminf_{k \rightarrow \infty} \frac{1}{k} \log \Lambda_k(Y) = -P + P_Y(f).$$

Similarly, we have the following upper bound:

$$(7.2) \quad \limsup_k \frac{1}{k} \log \mu(M_\ell) \leq -P + \limsup_{k \rightarrow \infty} \frac{1}{k} \log \left(\sum_{w \in B_k(X, H)} e^{S_k f(w)} \right).$$

Comparing the bounds in (7.1) and (7.2), we see that in order to finish the proof, it suffices to show that

$$(7.3) \quad \limsup_{k \rightarrow \infty} \frac{1}{k} \log \left(\sum_{w \in B_k(X, H)} e^{S_n f(w)} \right) \leq P_Y(f).$$

To get this inequality, we use ideas from [26] to find an invariant measure ν supported on Y such that $h(\nu) + \int f d\nu$. The measure ν is obtained as follows.

For $k \geq 1$, suppose $B_k(X, H) = \{w_1^k, \dots, w_{m_k}^k\}$. Let x_i^k be in $[w_i^k]$ such that $S_k f(w_i^k) = S_m f(x_i^k)$ (which exists by compactness and continuity). Then let

$$\begin{aligned} \mu_k &= \frac{\sum_{j=1}^m e^{S_k f(x_j^k)} \delta_{x_j^k}}{\sum_{j=1}^m e^{S_k f(x_j^k)}} \\ \nu_k &= \frac{1}{k} \sum_{j=0}^{k-1} S^j \mu_k. \end{aligned}$$

Since the space of Borel probability measures on X is weak* compact, there exists a Borel probability measure ν on X and a subsequence (ν_{k_j}) of (ν_k) such that $\nu_{k_j} \rightarrow \nu$ and along which the lim sup in (7.3) is obtained. Note that ν is in $M(X, S)$. Furthermore, we have that

$$\begin{aligned} \log \left(\sum_{w \in B_k(X, H)} e^{S_k f(w)} \right) &= \log \left(\sum_{j=1}^m e^{S_k f(x_j^k)} \right) \\ &= H_{\mu_k}(\xi^k) + \int f d\nu_k, \end{aligned}$$

where ξ is the natural partition of X according to the symbol in the zero coordinate and $\xi^k = \bigvee_{j=0}^{k-1} \sigma^{-j} \xi$. Arguing as in Proposition 3.6 of [26], we obtain that

$$(7.4) \quad \limsup_{k \rightarrow \infty} \frac{1}{k} \log \left(\sum_{w \in B_k(X, H)} e^{S_k f(w)} \right) \leq h(\nu) + \int f d\nu.$$

Now we claim that ν is supported on Y . Let $[w]$ be a cylinder set in X such that $Y \cap [w] = \emptyset$ and w has length N . We show that $\nu([w]) = 0$. Since $Y \cap [w] = \emptyset$ and since X is compact, there must exist k_0 such that for all $k \geq k_0$ and for all u in $B_k(X, H)$, it holds that w is not a subword of u . Then for u in $B_k(X, H)$, x in $[u]$, and $j = 0, \dots, k - N$, we have that $S^j(x) \notin [w]$. Hence $S^j \mu_k(w) = 0$ for $j = 0, \dots, k - N$, and therefore

$$\nu_k(w) = \frac{1}{k} \sum_{j=0}^{k-1} S^j \mu_k(w) = \frac{1}{k} \sum_{j=k-N+1}^{k-1} S^j \mu_k(w) \leq \frac{N}{k}.$$

Letting k tend to infinity along the subsequence (k_j) , we obtain that $\nu([w]) = 0$, as desired. Hence ν is supported on Y .

Then by (7.4) and the variational principle for $P_Y(f)$, we have that

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \log \left(\sum_{w \in B_k(X, H)} e^{S_k f(w)} \right) \leq h(\nu) + \int f d\nu \leq P_Y(f),$$

which establishes (7.3) and finishes the proof. \square

7.1. Proof of Theorem 1.4. Having established Proposition 7.1, we are now in a position to prove Theorem 1.4. The proof simply uses Proposition 7.1 to reduce Theorem 1.4 to Theorem 1.1.

PROOF OF THEOREM 1.4. Let X be a non-trivial mixing SFT, $f : X \rightarrow \mathbb{R}$ a Hölder continuous potential with associated Gibbs measure μ , and $\gamma_0 = \gamma_0(X, f)$ as in Theorem 1.1. Let $\alpha \in (\gamma_0, 1]$. Let $\epsilon > 0$. By Theorem 1.1, there exists $\rho > 0$ such that for all large enough n ,

$$\mathbb{P} \left(|P_{Y_n}(f) - (P_X(f) + \log(\alpha))| \geq \epsilon \right) < e^{-\rho n}.$$

Observe that Y_n is the survivor set of the open system (X, σ, H_n) . Then by Proposition 7.1, we have

$$-\varrho(\mu : H_n) = P_X(f) - P_{Y_n}(f).$$

Then for all large enough n , we see that

$$\mathbb{P} \left(|\varrho(\mu : H_n) - \log(\alpha)| \geq \epsilon \right) = \mathbb{P} \left(|P_{Y_n}(f) - (P_X(f) + \log(\alpha))| \geq \epsilon \right) < e^{-\rho n},$$

as was to be shown. \square

Acknowledgments

The author would like to thank the anonymous referee for helpful comments and suggestions. This work was supported by the National Science Foundation through the grant DMS 1613261.

References

- [1] Miguel Abadi, *Sharp error terms and necessary conditions for exponential hitting times in mixing processes*, Ann. Probab. **32** (2004), no. 1A, 243–264, DOI 10.1214/aop/1078415835. MR2040782
- [2] Rufus Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer-Verlag, Berlin-New York, 1975. MR0442989
- [3] Ryan Broderick, *Finite orbits in random subshifts of finite type*, Qual. Theory Dyn. Syst. **16** (2017), no. 3, 531–545, DOI 10.1007/s12346-017-0224-5. MR3703513
- [4] Henk Bruin, Mark Demers, and Ian Melbourne, *Existence and convergence properties of physical measures for certain dynamical systems with holes*, Ergodic Theory Dynam. Systems **30** (2010), no. 3, 687–728, DOI 10.1017/S0143385709000200. MR2643708
- [5] Leonid A. Bunimovich and Alex Yurchenko, *Where to place a hole to achieve a maximal escape rate*, Israel J. Math. **182** (2011), 229–252, DOI 10.1007/s11856-011-0030-8. MR2783972
- [6] N. Chernov and R. Markarian, *Anosov maps with rectangular holes. Nonergodic cases*, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 2, 315–342, DOI 10.1007/BF01233396. MR1479506
- [7] N. Chernov and R. Markarian, *Ergodic properties of Anosov maps with rectangular holes*, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 2, 271–314, DOI 10.1007/BF01233395. MR1479505

- [8] N. Chernov, R. Markarian, and S. Troubetzkoy, *Conditionally invariant measures for Anosov maps with small holes*, Ergodic Theory Dynam. Systems **18** (1998), no. 5, 1049–1073, DOI 10.1017/S0143385798117492. MR1653291
- [9] N. Chernov, R. Markarian, and S. Troubetzkoy, *Invariant measures for Anosov maps with small holes*, Ergodic Theory Dynam. Systems **20** (2000), no. 4, 1007–1044, DOI 10.1017/S0143385700000560. MR1779391
- [10] H. van den Bedem and N. Chernov, *Expanding maps of an interval with holes*, Ergodic Theory Dynam. Systems **22** (2002), no. 3, 637–654, DOI 10.1017/S0143385702000329. MR1908547
- [11] Pierre Collet, Servet Martínez, and Bernard Schmitt, *The Yorke-Pianigiani measure and the asymptotic law on the limit Cantor set of expanding systems*, Nonlinearity **7** (1994), no. 5, 1437–1443. MR1294552
- [12] Pierre Collet, Servet Martínez, and Bernard Schmitt, *Quasi-stationary distribution and Gibbs measure of expanding systems*, Instabilities and nonequilibrium structures, V (Santiago, 1993), Nonlinear Phenom. Complex Systems, vol. 1, Kluwer Acad. Publ., Dordrecht, 1996, pp. 205–219, DOI 10.1007/978-94-009-0239-8_19. MR1406590
- [13] Pierre Collet, Servet Martínez, and Bernard Schmitt, *The Pianigiani-Yorke measure for topological Markov chains*, Israel J. Math. **97** (1997), 61–70, DOI 10.1007/BF02774026. MR1441238
- [14] Mark F. Demers and Lai-Sang Young, *Escape rates and conditionally invariant measures*, Nonlinearity **19** (2006), no. 2, 377–397, DOI 10.1088/0951-7715/19/2/008. MR2199394
- [15] Mark F. Demers and Bastien Fernandez, *Escape rates and singular limiting distributions for intermittent maps with holes*, Trans. Amer. Math. Soc. **368** (2016), no. 7, 4907–4932, DOI 10.1090/tran/6481. MR3456165
- [16] Mark Demers, Paul Wright, and Lai-Sang Young, *Escape rates and physically relevant measures for billiards with small holes*, Comm. Math. Phys. **294** (2010), no. 2, 353–388, DOI 10.1007/s00220-009-0941-y. MR2579459
- [17] Mark F. Demers, *Markov extensions and conditionally invariant measures for certain logistic maps with small holes*, Ergodic Theory Dynam. Systems **25** (2005), no. 4, 1139–1171, DOI 10.1017/S0143385704000963. MR2158400
- [18] Mark F. Demers, *Markov extensions for dynamical systems with holes: an application to expanding maps of the interval*, Israel J. Math. **146** (2005), 189–221, DOI 10.1007/BF02773533. MR2151600
- [19] Mark F. Demers, *Dispersing billiards with small holes*, Ergodic theory, open dynamics, and coherent structures, Springer Proc. Math. Stat., vol. 70, Springer, New York, 2014, pp. 137–170, DOI 10.1007/978-1-4939-0419-8_8. MR3213499
- [20] Mark F. Demers, Christopher J. Ianzano, Philip Mayer, Peter Morfe, and Elizabeth C. Yoo, *Limiting distributions for countable state topological Markov chains with holes*, Discrete Contin. Dyn. Syst. **37** (2017), no. 1, 105–130, DOI 10.3934/dcds.2017005. MR3583472
- [21] Mark F. Demers and Paul Wright, *Behaviour of the escape rate function in hyperbolic dynamical systems*, Nonlinearity **25** (2012), no. 7, 2133–2150, DOI 10.1088/0951-7715/25/7/2133. MR2947939
- [22] Mark F. Demers, Paul Wright, and Lai-Sang Young, *Entropy, Lyapunov exponents and escape rates in open systems*, Ergodic Theory Dynam. Systems **32** (2012), no. 4, 1270–1301, DOI 10.1017/S0143385711000344. MR2955314
- [23] Andrew Ferguson and Mark Pollicott, *Escape rates for Gibbs measures*, Ergodic Theory Dynam. Systems **32** (2012), no. 3, 961–988, DOI 10.1017/S0143385711000058. MR2995652
- [24] Gary Froyland and Ognjen Stancevic, *Escape rates and Perron-Frobenius operators: Open and closed dynamical systems*, Discrete Contin. Dyn. Syst. Ser. B **14** (2010), no. 2, 457–472, DOI 10.3934/dcdsb.2010.14.457. MR2660868
- [25] Gerhard Keller, *Rare events, exponential hitting times and extremal indices via spectral perturbation*, Dyn. Syst. **27** (2012), no. 1, 11–27, DOI 10.1080/14689367.2011.653329. MR2903242
- [26] François Ledrappier and Peter Walters, *A relativised variational principle for continuous transformations*, J. London Math. Soc. (2) **16** (1977), no. 3, 568–576, DOI 10.1112/jlms/s2-16.3.568. MR0476995
- [27] Carlangelo Liverani and Véronique Maume-Deschamps, *Lasota-Yorke maps with holes: conditionally invariant probability measures and invariant probability measures on the survivor*

- set* (English, with English and French summaries), *Ann. Inst. H. Poincaré Probab. Statist.* **39** (2003), no. 3, 385–412, DOI 10.1016/S0246-0203(02)00005-5. MR1978986
- [28] Artur Lopes and Roberto Markarian, *Open billiards: invariant and conditionally invariant probabilities on Cantor sets*, *SIAM J. Appl. Math.* **56** (1996), no. 2, 651–680, DOI 10.1137/S0036139995279433. MR1381665
- [29] Kevin McGoff, *Random subshifts of finite type*, *Ann. Probab.* **40** (2012), no. 2, 648–694, DOI 10.1214/10-AOP636. MR2952087
- [30] Kevin McGoff and Ronnie Pavlov, *Factor maps and embeddings for random \mathbb{Z}^d shifts of finite type*, *Israel J. Math.*, to appear.
- [31] Kevin McGoff and Ronnie Pavlov, *Random \mathbb{Z}^d -shifts of finite type*, *J. Mod. Dyn.* **10** (2016), 287–330, DOI 10.3934/jmd.2016.10.287. MR3538865
- [32] Donald Samuel Ornstein and Benjamin Weiss, *Entropy and data compression schemes*, *IEEE Trans. Inform. Theory* **39** (1993), no. 1, 78–83, DOI 10.1109/18.179344. MR1211492
- [33] Giulio Pianigiani and James A. Yorke, *Expanding maps on sets which are almost invariant. Decay and chaos*, *Trans. Amer. Math. Soc.* **252** (1979), 351–366, DOI 10.2307/1998093. MR534126
- [34] Peter Walters, *An introduction to ergodic theory*, *Graduate Texts in Mathematics*, vol. 79, Springer-Verlag, New York-Berlin, 1982. MR648108
- [35] Lai-Sang Young, *Large deviations in dynamical systems*, *Trans. Amer. Math. Soc.* **318** (1990), no. 2, 525–543, DOI 10.2307/2001318. MR975689

9201 UNIVERSITY CITY BLVD., CHARLOTTE, NC 28223
Email address: kmcgoff1@uncg.edu