Recurrence rates for shifts of finite type

Demi Allen[1] Simon Baker[2,*] Balázs Bárány[3,4,†]

July 11, 2024

1) Department of Mathematics and Statistics, University of Exeter, Harrison Building, North Park Road, Exeter, EX4 4QF, UK.
Email address: d.d.allen@exeter.ac.uk

2)Department of Mathematical Sciences, Loughborough University, Loughborough, LE11 3TU, UK
Email address: simonbaker412@gmail.com

3) Department of Stochastics, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary.
4) MTA-BME Stochastics Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary
Email address: balubsheep@gmail.com

Abstract

Let Σ_A be a topologically mixing shift of finite type, let $\sigma : \Sigma_A \to \Sigma_A$ be the usual left-shift, and let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. In this paper we study recurrence rates for the dynamical system (Σ_A, σ) that hold μ-almost surely. In particular, given a function $\psi : \mathbb{N} \to \mathbb{N}$ we are interested in the following set

$$R_\psi = \{\mathtt{i} \in \Sigma_A : i_{n+1} \dots i_{n+\psi(n)+1} = i_1 \dots i_{\psi(n)} \text{ for infinitely many } n \in \mathbb{N}\}.$$

We provide sufficient conditions for $\mu(R_\psi) = 1$ and sufficient conditions for $\mu(R_\psi) = 0$. As a corollary of these results, we discover a new critical threshold where the measure of R_ψ transitions from zero to one. This threshold was previously unknown even in the special case of a non-uniform Bernoulli measure defined on the full shift. The proofs of our results combine ideas from Probability Theory and Thermodynamic Formalism. In our final section we apply our results to the study of dynamics on self-similar sets.

1 Introduction

The notion of recurrence is fundamental in the study of Dynamical Systems and Ergodic Theory. The famous Poincaré Recurrence Theorem (see, e.g., [13, Theorem 2.11] or [32, Theorem 1.4]) states that if $T : X \to X$ is a measure-preserving transformation on a probability space $(X, \mathcal{B}, \mu)$, then for any set $A \in \mathcal{B}$ satisfying $\mu(A) > 0$, we have that μ-almost every $x \in A$ satisfies $T^n(x) \in A$

*SB acknowledges support from EPSRC grant EP/W003880/1.

†BB acknowledges support from the grants NKFI K123782, FK134251, K142169, and the grant NKFI KKP144059 "Fractal geometry and applications".

for infinitely many $n \in \mathbb{N}$. Under some modest assumptions this measure theoretic result can be upgraded to a metric one. Indeed if X is equipped with a metric d so that (X, d) is separable and $\mathcal{B}$ is the Borel σ-algebra, then Poincaré's theorem implies that for μ-almost every $x \in X$ we have

$$\liminf_{n\to\infty} d(T^n(x), x) = 0.$$

When d is a metric so that (X, d) is separable and $\mathcal{B}$ is the Borel σ-algebra, we refer to $(X, \mathcal{B}, \mu, T, d)$ as a metric measure-preserving system or a m.m.p.s. It is natural to wonder whether the conclusion $\liminf_{n\to\infty} d(T^n(x), x) = 0$ for μ-almost every x can be strengthened into something more quantitative. In particular, given a function $\psi : \mathbb{N} \to [0, \infty)$ is it true that μ-almost every x satisfies $d(T^n(x), x) \leq \psi(n)$ for infinitely many n. This problem is known as the quantitative recurrence problem. The first result obtained in this direction is the following extremely general quantitative recurrence result obtained by Boshernitzan in [7, Theorem 1.2].

Theorem 1.1 (Boshernitzan, [7]). *Let $(X, \mathcal{B}, \mu, T, d)$ be a m.m.p.s. Assume that for some $\alpha > 0$ the α-dimensional Hausdorff measure $\mathcal{H}^\alpha$ is σ-finite on (X, d). Then for μ-almost every $x \in X$ we have*

$$\liminf_{n\to\infty} n^{1/\alpha} d(T^n(x), x) < \infty.$$

Moreover, if $\mathcal{H}^\alpha(X) = 0$, then for μ-almost every $x \in X$ we have

$$\liminf_{n\to\infty} n^{1/\alpha} d(T^n(x), x) = 0.$$

Throughout this paper we denote by $\mathcal{H}^\alpha(X)$ the *α-dimensional Hausdorff measure* of a set X, and we write $\dim_{\mathrm{H}} X$ to denote the *Hausdorff dimension* of X. We refer the reader to [14] for definitions and further information regarding Hausdorff measures and dimension.

The recurrence rates Boshernitzan's theorem provides do not exhibit a dependence on the measure μ, which is contrary to what one would expect. It could well be the case that Theorem 1.1 does not allow us to conclude an optimal recurrence rate that holds for μ-almost every x. This issue was partly addressed in a paper by Barreira and Saussol [3]. In particular they proved the following statement which appears as [3, Theorem 3].

Theorem 1.2 (Barreira and Saussol, [3]). *If $T : X \to X$ is a Borel measurable map on $X \subset \mathbb{R}^d$, and μ is a T-invariant probability measure on X, then for μ-almost every $x \in X$, we have*

$$\liminf_{n\to\infty} n^{1/\alpha} d(T^n(x), x) = 0 \text{ for any } \alpha > \liminf_{r\to 0} \frac{\log \mu(B(x, r))}{\log r}.$$

Following on from these two important results, two separate research streams have arisen. The first of these streams takes a dynamical system and a recurrence rate, and tries to determine the Hausdorff dimension of the set of points that satisfy this recurrence rate. This line of research was pursued by Tan and Wang in [31] where the underlying dynamical system was the β-transformation. Seuret and Wang obtained similar results in [30] for dynamical systems arising from the study of self-conformal sets. The second stream is more measure theoretic in nature. Given a metric measure-preserving system $(X, \mathcal{B}, \mu, T, d)$ and a function $\psi : \mathbb{N} \to [0, \infty)$, this stream seeks to find simple criteria which determine the measure of the set

$$R(\psi) := \{x \in X : d(T^n(x), x) \leq \psi(n) \text{ for infinitely many } n \in \mathbb{N}\}.$$

This type of problem has been studied in great depth in the context of so called *shrinking target problems*. In the shrinking target framework, instead of studying those x satisfying $d(T^n(x), x) \leq$

$\psi(n)$ for infinitely many n, we fix a $y \in X$ and study those x satisfying $d(T^n(x), y) \leq \psi(n)$ for infinitely many n. For more on shrinking target problems we refer the reader to [1, 11, 16, 17, 22, 25] and the references therein. Part of the motivation behind the recent activity in this area is to bring the quantitative recurrence theory in line with the theory of shrinking targets. Drawing an analogy with shrinking target problems, it is reasonable to expect that $\mu(R(\psi))$ should be determined by some naturally occurring volume sum. In particular, it is reasonable to expect that there exists a monotone function $f : [0, \infty) \to [0, \infty)$ for which the following holds

$$\mu(R(\psi)) = \begin{cases} 1 & \text{if } \sum_{n=1}^{\infty} f(\psi(n)) = \infty, \\ 0 & \text{if } \sum_{n=1}^{\infty} f(\psi(n)) < \infty. \end{cases} \tag{1.1}$$

Typically we might expect f to be of the form $f(x) = x^\gamma$ where $\gamma > 0$ is the Hausdorff dimension of the measure. This principle has been verified in various contexts by several authors. It was shown to be the case for certain natural maps defined on attractors of iterated function systems by Chang et al. in [9], and by the second author and Farmer in [2]. Hussain et al. in [18] gave general conditions for a dynamical system to guarantee that (1.1) holds. This result applies to many well known dynamical systems including the β-transformation and the Gauss map. Further refinements were obtained by Kirsebom et al. in [23]. They obtained positive results for a more general class of dynamical system than was previously considered in [2, 9, 18]. They also considered a more general notion of recurrence where the function ψ was also allowed to depend upon the point x. Recently Kleinbock and Zheng proved a general result that implies a zero-one law for both the shrinking target problem and the quantitative recurrence problem [24]. This result applies to a general family of expanding dynamical systems satisfying suitable bounded distortion assumptions and for which the underlying measure satisfies the following property: there exist constants $c_1, c_2, \alpha > 0$ such that for any $x \in X$ and $r \in (0, Diam(X))$ we have

$$c_1 r^\alpha \leq \mu(B(x, r)) \leq c_2 r^\alpha.$$

Measures satisfying this property are called *Ahlfors regular measures*. We conclude this overview of related works by mentioning a paper by Chazottes and Ugalde [10]. Just as we do in this paper, the authors of [10] study Gibbs measures for shifts of finite type. In their paper they obtained almost sure bounds for the first time a sequence returns under the left-shift to the cylinder determined by its first n entries. For certain functions ψ these bounds can be used to obtain measure statements for $R(\psi)$. However, the class of ψ for which this can be done is quite restrictive and does not contain many natural choices of ψ one would be interested in. This is not surprising given that the bounds stated in [10] apply to all first returns, whereas $R(\psi)$ is defined in terms of a subsequence of returns. In this paper we obtain detailed information on the measure of the set $R(\psi)$ that goes beyond that which can be derived from [10].

In this paper we study recurrence rates for shifts of finite type that hold μ-almost surely with respect to a Gibbs measure μ. Of special interest will be the case where the defining potential of μ has positive variance with respect to μ. For short, we say that μ *has positive variance* in this case. This seemingly innocuous assumption leads to a much richer theory and a wider range of behaviour than was previously observed in [2, 9, 18, 23, 24]. The main results of these papers each assume that the underlying measure μ satisfies some additional uniformity assumption. Often this uniformity assumption is that the measure is Ahlfors regular. This assumption is very useful for technical reasons, and often means that one can study the measure of $R(\psi)$ using a similar toolkit to that which one would use to study shrinking target problems. However, this uniformity assumption rules out measures μ with positive variance. These measures are often

highly non-uniform and exhibit strong multifractal behaviour. Importantly, it is precisely this non-uniformity that leads to the richer theory and the wider range of behaviour mentioned above. We will properly formalise our results in the next section. For now we state the following result which illustrates the phenomenon described in this paragraph. This result is a special case of Theorem 2.4.

Theorem 1.3. *Let $(p_i)_{i=1}^K$ be a non-uniform*[1] *probability vector and let μ be the corresponding Bernoulli measure defined on the full shift $\Sigma = \{1, \dots, K\}^{\mathbb{N}}$. Equip Σ with the metric d given by $d(\mathtt{i}, \mathtt{j}) = K^{-|\,\mathtt{i} \wedge \mathtt{j}\,|}$ where $|\,\mathtt{i} \wedge \mathtt{j}\,| := \inf\{n \geq 0 : i_{n+1} \neq j_{n+1}\}$. Then there exist constants $h, \rho > 0$ such that the following holds: for $\varepsilon > 0$ let $\psi_\varepsilon^+ : \mathbb{N} \to \mathbb{N}$ and $\psi_\varepsilon^- : \mathbb{N} \to \mathbb{N}$ be given by:*

$$\psi_\varepsilon^+(n) = \left\lfloor \frac{\log n}{h} + \frac{(1+\varepsilon)}{h^{3/2}} \sqrt{2\rho \log n \log\log\log n} \right\rfloor$$

and

$$\psi_\varepsilon^-(n) = \left\lfloor \frac{\log n}{h} + \frac{(1-\varepsilon)}{h^{3/2}} \sqrt{2\rho \log n \log\log\log n} \right\rfloor.$$

We then let $\Psi_\varepsilon^+ : \mathbb{N} \to [0, \infty)$ and $\Psi_\varepsilon^- : \mathbb{N} \to [0, \infty)$ be given by $\Psi_\varepsilon^+(n) = K^{-\psi_\varepsilon^+(n)}$ and $\Psi_\varepsilon^-(n) = K^{-\psi_\varepsilon^-(n)}$. Then for any $\varepsilon > 0$ we have $\mu(R(\Psi_\varepsilon^+)) = 0$ and $\mu(R(\Psi_\varepsilon^-)) = 1$.

Theorem 1.3 demonstrates a new critical threshold which to the best of our knowledge has not previously been observed in the study of quantitative recurrence or shrinking target problems. We also emphasise that unlike in [2, 9, 18, 23, 24], the presence of the $\sqrt{2\rho \log n \log\log\log n}$ term appearing in Theorem 1.3 means that for any $\gamma > 0$ the series $\sum_{n=1}^\infty (\Psi_\varepsilon^+)^\gamma$ and $\sum_{n=1}^\infty (\Psi_\varepsilon^-)^\gamma$ both simultaneously converge or diverge. This is a simple check when $\gamma \neq \frac{h}{\log K}$ and when $\gamma = \frac{h}{\log K}$ the terms in both of the series can be bounded above by a constant multiple of $\frac{1}{n(\log n)^{1+\delta}}$ for any $\delta > 0$ and are therefore convergent. Hence there is no $\gamma > 0$ such that for every $\psi : \mathbb{N} \to [0, \infty)$ we have (1.1) for some f of the form $f(x) = x^\gamma$. This observation shows how the quantitative recurrence problem is sensitive to the homogeneity and the fluctuations of the measure. This observation also shows that in the non-Ahflors regular case we obtain faster rates of recurrence. In particular, when $(p_i)_{i=1}^K$ is a non-uniform probability vector, if we take

$$\psi(n) = \left\lfloor \frac{\log n}{-\sum_{i=1}^K p_i \log p_i} + \sqrt{\log n} \right\rfloor$$

and $\Psi(n) = K^{-\psi(n)}$ then $\mu(R(\Psi)) = 1$ by the above, whereas if $(p_i)_{i=1}^K$ were a uniform probability vector then for the analogous Ψ it can be shown that $\mu(R(\Psi)) = 0$ using the Borel-Cantelli lemma.

Theorem 1.3 also has another interesting consequence. Let μ, Σ and d be as in the statement of this theorem. A natural guess for a sum that might determine the measure of a set $R(\Psi)$ is

$$\sum_{n=1}^\infty \mu(\{\mathtt{i} : d(\sigma^n(\mathtt{i}), \mathtt{i}) \leq \Psi(n)\}). \tag{1.2}$$

It is reasonable to expect that if this sum diverges then $\mu(R(\Psi)) = 1$ and if it converges then $\mu(R(\Psi)) = 0$. By the Borel-Cantelli lemma it can be shown that the convergence part of this prediction is true. However, the divergence part of this prediction is false. If we let $\Psi_\alpha(n) = K^{-\alpha \log n}$ where

$$\frac{1}{h} < \alpha < \frac{-1}{\log \sum_{i=1}^K p_i^2} \tag{1.3}$$

[1] $p_i \neq 1/K$ for some i. This condition ensures positive variance.

then the sum (1.2) would diverge. The existence of such α is a consequence of the convexity of $-\log x$. However, since $R(\Psi_\alpha) \subset R(\Psi_\epsilon^+)$ for any $\alpha > 1/h$, it follows from Theorem 1.3 that $\mu(R(\Psi_\alpha)) = 0$. Thus the prediction that divergence of the series (1.2) guarantees full measure is incorrect. Another natural series to consider is

$$\sum_{n=1}^{\infty} \mu(\{\mathtt{i} : \text{ there exists } p \in \Psi^{-1}([K^{-n}, K^{-n+1})) \text{ such that } d(\sigma^p(\mathtt{i}), \mathtt{i}) \leq K^{-n}\}). \tag{1.4}$$

Its convergence by Borel-Cantelli lemma again implies that $\mu(R(\Psi)) = 0$, however, we could not verify nor show a counterexample that the divergence of (1.4) implies $\mu(R(\Psi)) = 1$.

Notation. We end this introductory section by formalising some notation that we will use throughout. Given two positive real valued functions $f, g : S \to (0, \infty)$ defined on some set S, we write $f \ll g$ if there exists a constant $C > 0$ such that $f(x) \leq Cg(x)$ for all $x \in S$. For each $k \in \mathbb{N}$ we write $\log^{(k)}$ to denote the function that is the k-fold composition of log with itself, e.g. $\log^{(3)}(x) = \log\log\log x$. To avoid additional notational complexities, we will adopt the convention throughout that $\log x = 0$ whenever $x \leq 1$.

2 Background and statement of results

2.1 Background on Thermodynamic Formalism

Let $K \geq 2$ be an integer and let

$$\Sigma = \{1, \dots, K\}^{\mathbb{N}}$$

be the set of infinite sequences formed by elements of the set $\{1, \dots, K\}$. For integers $n \geq 0$, let us denote by

$$\Sigma_n = \{1, \dots, K\}^n$$

the set of words of length n, by convention $\Sigma_0 = \emptyset$. We also let

$$\Sigma_* = \bigcup_{n=0}^{\infty} \Sigma_n$$

be the set of all finite words. For $\mathtt{i} \in \Sigma_*$, denote by $|\,\mathtt{i}\,|$ the length of $\mathtt{i}$. For $\mathtt{i}, \mathtt{j} \in \Sigma \cup \Sigma_*$, denote by $\mathtt{i} \wedge \mathtt{j}$ the common prefix of $\mathtt{i}$ and $\mathtt{j}$; that is, let

$$|\,\mathtt{i} \wedge \mathtt{j}\,| := \inf\{n \geq 0 : i_{n+1} \neq j_{n+1}\} \qquad \text{and} \qquad \mathtt{i} \wedge \mathtt{j} = (i_1, \dots, i_{|\,\mathtt{i} \wedge \mathtt{j}\,|}).$$

Note that if $|\,\mathtt{i} \wedge \mathtt{j}\,| = 0$ then we define $\mathtt{i} \wedge \mathtt{j}$ as the empty word, and if $|\,\mathtt{i} \wedge \mathtt{j}\,| = \infty$ or $|\,\mathtt{i}\,| = |\,\mathtt{j}\,| = |\,\mathtt{i} \wedge \mathtt{j}\,|$ then $\mathtt{i} \wedge \mathtt{j} = \mathtt{i} = \mathtt{j}$. We equip Σ with the metric given by

$$d(\mathtt{i}, \mathtt{j}) = K^{-|\,\mathtt{i} \wedge \mathtt{j}\,|}.$$

For a finite word $\mathtt{i} \in \Sigma_*$ and for a sequence $\mathtt{j} \in \Sigma \cup \Sigma_*$, we denote by $\mathtt{i}\,\mathtt{j}$ the concatenation of $\mathtt{i}$ and $\mathtt{j}$. Moreover, for $\mathtt{i} \in \Sigma_*$ and $B \subseteq \Sigma \cup \Sigma_*$, let

$$\mathtt{i}\,B = \{\mathtt{i}\,\mathtt{j} : \mathtt{j} \in B\}.$$

For a word $\mathtt{i} \in \Sigma \cup \Sigma_*$ and $1 \leq n \leq m \leq |\,\mathtt{i}\,|$, we let

$$\mathtt{i}\,|_n^m = (i_n, \dots, i_m).$$

Let $\sigma\colon \Sigma \mapsto \Sigma$ be the left-shift map; that is,

$$\sigma(i_1, i_2, \ldots) = (i_2, i_3, \ldots).$$

Given a $K \times K$ matrix A consisting of zeros and ones we can define the corresponding *shift of finite type* as follows:

$$\Sigma_A := \left\{ \mathtt{i} \in \Sigma : A_{i_l, i_{l+1}} = 1 \text{ for all } l \in \mathbb{N} \right\}.$$

Notice that σ is well defined as a map from Σ_A to Σ_A. Also, note that the full shift corresponds to the case when the matrix A consists entirely of ones. We will always assume that there exists $M \in \mathbb{N}$ such that each entry of A^M is strictly positive. This assumption means that the map $\sigma : \Sigma_A \to \Sigma_A$ is topologically mixing (for the definition of topologically mixing see [21, Definition 1.8.2]). For each $n \in \mathbb{N}$ we let

$$\Sigma_{A,n} := \left\{ \mathtt{i} \in \Sigma_n : A_{i_l, i_{l+1}} = 1 \text{ for all } 1 \leq l \leq n-1 \right\}$$

and write

$$\Sigma_{A,*} = \bigcup_{n=1}^{\infty} \Sigma_{A,n}.$$

For an $\mathtt{i} \in \Sigma_{A,*}$, let $[\mathtt{i}]$ be the corresponding cylinder set of sequences in Σ_A that begin with $\mathtt{i}$, i.e.

$$[\mathtt{i}] := \{ \mathtt{j} \in \Sigma_A : \mathtt{i} = j_1 \ldots j_{|\,\mathtt{i}\,|} \}.$$

For a subset $B \subseteq \Sigma_{A,*}$ we also let $[B] := \bigcup_{\mathtt{i} \in B} [\mathtt{i}]$.

We call a map $f\colon \Sigma_A \mapsto \mathbb{R}$ *Hölder-continuous* if there exist constants $b > 0$ and $0 < \alpha < 1$ such that

$$|f(\mathtt{i}) - f(\mathtt{j})| \leq b\alpha^{|\,\mathtt{i} \wedge \mathtt{j}\,|}$$

for all $\mathtt{i}, \mathtt{j} \in \Sigma_A$. Let us define the *pressure* of such an f in the usual way by

$$P(f) = \lim_{n \to \infty} \frac{1}{n} \log \left(\sum_{\mathtt{i} \in \Sigma_{A,n}} \sup_{\mathtt{j} \in [\mathtt{i}]} \exp \left(\sum_{k=0}^{n-1} f(\sigma^k\, \mathtt{j}) \right) \right).$$

By [8, Theorem 1.4, Proposition 1.13, and Theorem 1.16] for each such f there exists a unique σ-invariant ergodic measure μ on Σ_A for which there exists a constant $C > 1$ such that

$$C^{-1} \leq \frac{\mu([i_1, \ldots, i_n])}{\exp\left(-nP(f) + \sum_{k=0}^{n-1} f(\sigma^k\, \mathtt{i}) \right)} \leq C \tag{2.1}$$

for every $\mathtt{i} \in \Sigma_A$ and $n \in \mathbb{N}$. We call μ the *Gibbs measure* of the potential f. For the purpose of exposition, we suppress the dependence of μ on f throughout. It can be seen from the definition of Gibbs measure that there exists a constant $C > 1$ such that for every $\mathtt{i}, \mathtt{j} \in \Sigma_{A,*}$ satisfying $\mathtt{i}\,\mathtt{j} \in \Sigma_{A,*}$, we have

$$C^{-1} \leq \frac{\mu([\mathtt{i}\,\mathtt{j}])}{\mu([\mathtt{i}])\mu([\mathtt{j}])} \leq C. \tag{2.2}$$

Without loss of generality, we may assume that the constants in (2.1) and (2.2) are equal.

We define the *entropy* of a Gibbs measure μ by

$$h_\mu := \lim_{n \to \infty} \frac{-1}{n} \sum_{\mathtt{i} \in \Sigma_{A,n}} \mu([\mathtt{i}]) \log \mu([\mathtt{i}]).$$

It can be shown, see [8, Theorem 1.22] for example, that

$$h_\mu = P(f) - \int f(\mathtt{i}) d\mu(\mathtt{i}). \tag{2.3}$$

We define the *variance* of a Gibbs measure μ of some potential f to be

$$\rho_\mu := \lim_{n\to\infty} \frac{1}{n} \int \left(\sum_{k=0}^{n-1} f(\sigma^k \mathtt{i}) - n \int f(\mathtt{j})\, d\mu(\mathtt{j}) \right)^2 d\mu(\mathtt{i}).$$

We emphasise that this limit always exists (see [27]). Clearly $\rho_\mu \geq 0$. It is well known that we have equality here if and only if f is cohomologous to a constant, i.e. there exists a continuous function $g : \Sigma_A \to \mathbb{R}$ and $c \in \mathbb{R}$ such that $f = g \circ \sigma - g + c$. For a proof see [27]. We will often assume that f is not cohomologous to a constant and therefore that $\rho_\mu > 0$.

If $(p_i)_{i=1}^K$ is a probability vector and μ is the corresponding Bernoulli measure on the full shift Σ, then μ is the Gibbs measure for the potential $f : \Sigma \to \mathbb{R}$ given by $f(\mathtt{i}) = \log p_{i_1}$. We recall here that the Bernoulli measure corresponding to the probability vector $(p_i)_{i=1}^K$ is defined by assigning cylinders measure

$$\mu([\mathtt{i}]) = p_{i_1} p_{i_2} \dots p_{i_{|\mathtt{i}|}}$$

for each $\mathtt{i} \in \Sigma_*$. An important consequence of this definition is that for any two words $\mathtt{i}, \mathtt{j} \in \Sigma_*$, we have

$$\mu([\mathtt{i}\,\mathtt{j}]) = \mu([\mathtt{i}])\mu([\mathtt{j}]). \tag{2.4}$$

For further details regarding the definition of Bernoulli measures and some of their properties we refer the reader to [21, Chapter 4.2] and [32]. We emphasise that when μ is a Bernoulli measure with corresponding probability vector $(p_i)_{i=1}^K$, then the entropy and variance of μ take the following simpler form:

$$h_\mu = -\sum_{i=1}^K p_i \log p_i$$

and

$$\rho_\mu = \sum_{i=1}^K p_i (\log p_i)^2 - \left(\sum_{i=1}^K p_i \log p_i \right)^2.$$

It can be seen that the potential f corresponding to a Bernoulli measure μ is cohomologous to a constant if and only if f is a constant function. Therefore $\rho_\mu > 0$ if and only if $(p_i)_{i=1}^K$ is not the uniform probability vector.

2.2 Statement of main results and some proofs

Let $\psi\colon \mathbb{N} \mapsto \mathbb{N}$ and let R_ψ be the set of infinite sequences that return infinitely often to the neighbourhood determined by ψ when we apply the left shift. That is, let

$$R_\psi := \{\mathtt{i} \in \Sigma_A : d(\sigma^n \mathtt{i}, \mathtt{i}) \leq K^{-\psi(n)} \text{ for infinitely many } n \in \mathbb{N}\}. \tag{2.5}$$

Note that this definition is slightly different to the definition of $R(\psi)$ given earlier. We emphasise that if we are given $\psi\colon \mathbb{N} \to \mathbb{N}$ and define $\tilde{\psi}\colon \mathbb{N} \to (0, \infty)$ via the equation $\tilde{\psi}(n) = K^{-\psi(n)}$, then $R(\tilde{\psi}) = R_\psi$. The latter definition helps to simplify much of our analysis. In our subsequent

proofs we will often make the assumption that ψ satisfies $\psi(n) \leq n$ for all n sufficiently large. Under this assumption R_ψ has the following simple formulation

$$R_\psi = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \bigcup_{\substack{\mathtt{i}\in\Sigma_{A,\psi(m)} \\ \mathtt{j}\in\Sigma_{A,m-\psi(m)} \\ \mathtt{i}\,\mathtt{j}\,\mathtt{i}\in\Sigma_{A,m+\psi(m)}}} [\mathtt{i}\,\mathtt{j}\,\mathtt{i}]. \tag{2.6}$$

The assumption that ψ satisfies $\psi(n) \leq n$ for all n sufficiently large is not especially restrictive. Indeed for any Gibbs measure μ it can be shown using the Borel–Cantelli Lemma (see Lemma 3.5) and the Shannon–McMillan–Breiman Theorem [28, Chapter 6, Theorem 2.3] that

$$\mu\left(\{\mathtt{i} \in \Sigma_A : d(\sigma^k \mathtt{i}, \mathtt{i}) \leq K^{-k} \text{ for infinitely many } k \in \mathbb{N}\}\right) = 0.$$

As such, for an arbitrary ψ if we let $\tilde{\psi} : \mathbb{N} \to \mathbb{N}$ be given by $\tilde{\psi}(n) = \min\{\psi(n), n\}$, so in particular $\tilde{\psi}(n) \leq n$ for all n, then it follows from the above that $\mu(R_\psi) = \mu(R_{\tilde{\psi}})$. Thus when determining $\mu(R_\psi)$ we can often assume without loss of generality that ψ satisfies $\psi(n) \leq n$ for all n sufficiently large.

The goal of this paper is to find "close to optimal" conditions for establishing a "zero-one law" for the measure of the set R_ψ. The following theorem provides sufficient conditions for R_ψ to have measure zero.

Theorem 2.1. *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. Let $\psi\colon \mathbb{N} \mapsto \mathbb{N}$ be such that there exists $\varepsilon > 0$ for which*

$$\sum_{n=1}^{\infty} e^{-h_\mu \psi(n) + (1+\varepsilon)\sqrt{2\rho_\mu \psi(n) \log\log \psi(n)}} < \infty.$$

Then $\mu(R_\psi) = 0$.

Given $\psi : \mathbb{N} \to \mathbb{N}$ and $n \in \mathbb{N}$, let $\psi^{-1}(n) = \{m \in \mathbb{N} : \psi(m) = n\}$. For any set $A \subseteq \mathbb{N}$ we denote the cardinality of A by $\#A$. The following theorem is the main result of this paper. It gives sufficient conditions for $\mu(R_\psi) = 1$.

Theorem 2.2. *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. Let $\psi\colon \mathbb{N} \mapsto \mathbb{N}$ be such that $\psi(n) \leq n$ for all n sufficiently large and $\lim_{n\to\infty} \psi(n) = \infty$. Suppose that there exists an increasing sequence $(n_k)_{k=1}^{\infty}$ and a function $g\colon \mathbb{N} \mapsto [1, \infty)$ satisfying $\lim_{n\to\infty} \frac{g(n)}{g(n+1)} = 1$, such that*

$$\mu\left(\left\{\mathtt{i} \in \Sigma_A : \limsup_{k\to\infty} \frac{\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k} g(n_k)} > 1\right\}\right) = 1 \tag{2.7}$$

and

$$\lim_{k\to\infty} \#\psi^{-1}(n_k) e^{-h_\mu n_k + \sqrt{2\rho_\mu n_k} g(n_k)} = \infty. \tag{2.8}$$

Then

$$\mu\left(\left\{\mathtt{i} \in \Sigma_A : d(\sigma^p \mathtt{i}, \mathtt{i}) \leq K^{-n_k} \text{ for some } p \in \psi^{-1}(n_k) \text{ for infinitely many } k \in \mathbb{N}\right\}\right) = 1.$$

Moreover, $\mu(R_\psi) = 1$.

We emphasise that Theorem 2.2 also holds without the assumption $\lim_{n\to\infty}\psi(n)=\infty$. It can be shown using the arguments given in this paper that whenever $\liminf_{n\to\infty}\psi(n)<\infty$ we have $\mu(R_\psi)=1$. Theorem 2.2 implies the following corollary.

Corollary 2.3. *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. Let $\psi\colon \mathbb{N}\mapsto\mathbb{N}$ be such that $\psi(n)\leq n$ for all n sufficiently large and $\lim_{n\to\infty}\psi(n)=\infty$. Suppose that there exists $\varepsilon>0$ for which*

$$\left\lceil e^{h_\mu n-(1-\varepsilon)\sqrt{2\rho_\mu n\log\log n}}\right\rceil \ll \#\psi^{-1}(n) \tag{2.9}$$

for all n sufficiently large. Then $\mu(R_\psi)=1$.

Proof. We take $(n_k)_{k=1}^\infty=(k)_{k=1}^\infty$ and $g(n)=(1-\varepsilon/2)\sqrt{\log\log n}$. Then, by (2.9), we have that

$$\#\psi^{-1}(n_k)e^{-h_\mu n_k+\sqrt{2\rho_\mu n_k}g(n_k)}\gg\exp\left(\frac{\varepsilon}{2}\sqrt{2\rho_\mu k\log\log k}\right)$$

for all k sufficiently large and so (2.8) follows readily. The fact assumption (2.7) holds is a consequence of the Law of the Iterated Logarithm for Gibbs measures, see Theorem 3.2 below. By Theorem 2.2 we now have $\mu(R_\psi)=1$. □

One case our analysis does not cover is when ψ only satisfies the weaker assumption that

$$\left\lceil e^{h_\mu n-\sqrt{2\rho_\mu n\log\log n}}\right\rceil \ll \#\psi^{-1}(n)$$

for all n sufficiently large. It seems to be a challenging problem to determine $\mu(R_\psi)$ under this assumption.

Combining Theorem 2.1 and Corollary 2.3 we obtain the following result. It identifies a critical threshold for the measure of R_ψ. This result implies Theorem 1.3 stated in the introduction.

Theorem 2.4. *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. For $\varepsilon>0$ let $\psi_\varepsilon^+:\mathbb{N}\to\mathbb{N}$ and $\psi_\varepsilon^-:\mathbb{N}\to\mathbb{N}$ be given by:*

$$\psi_\varepsilon^+(n)=\left\lfloor\frac{\log n}{h_\mu}+\frac{(1+\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu\log n\log^{(3)}n}\right\rfloor$$

and

$$\psi_\varepsilon^-(n)=\left\lfloor\frac{\log n}{h_\mu}+\frac{(1-\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu\log n\log^{(3)}n}\right\rfloor.$$

Then for any $\varepsilon>0$ we have $\mu(R_{\psi_\varepsilon^+})=0$ and $\mu(R_{\psi_\varepsilon^-})=1$.

Proof. This result is a consequence of Theorem 2.1 and Corollary 2.3. We split our proof into two parts.

Proof that $\mu(R_{\psi_\varepsilon^+})=0$. Fix $\varepsilon>0$. We begin by noting that for all n sufficiently large we have

$$\psi_\varepsilon^+(n)=\left\lfloor\frac{\log n}{h_\mu}+\frac{(1+\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu\log n\log^{(3)}n}\right\rfloor\leq\left(1+\frac{\varepsilon}{2}\right)^{2/3}\frac{\log n}{h_\mu} \tag{2.10}$$

and

$$\log\log\left(\left(1+\frac{\varepsilon}{2}\right)^{2/3}\frac{\log n}{h_\mu}\right)\leq\left(1+\frac{\varepsilon}{2}\right)^{2/3}\log^{(3)}n. \tag{2.11}$$

Writing $e(x) = e^x$, it follows from the definition of ψ_ε^+, (2.10) and (2.11) that

$$\begin{aligned}
&\sum_{n=1}^{\infty} e^{-h_\mu \psi_\varepsilon^+(n) + (1+\varepsilon/2)^{1/3}\sqrt{2\rho_\mu \psi_\varepsilon^+(n) \log\log \psi_\varepsilon^+(n)}} \\
\ll &\sum_{n=1}^{\infty} \frac{1}{n} e\left(-\frac{(1+\varepsilon)}{h_\mu^{1/2}} \sqrt{2\rho_\mu \log n \log^{(3)} n} + \left(1+\frac{\varepsilon}{2}\right)^{1/3} \sqrt{2\rho_\mu \psi_\varepsilon^+(n) \log^{(2)} \psi_\varepsilon^+(n)} \right) \\
\overset{(2.10)}{\ll} &\sum_{n=1}^{\infty} \frac{1}{n} e\left(-\frac{(1+\varepsilon)}{h_\mu^{1/2}} \sqrt{2\rho_\mu \log n \log^{(3)} n} + \left(1+\frac{\varepsilon}{2}\right)^{1/3} \sqrt{2\rho_\mu \left(1+\frac{\varepsilon}{2}\right)^{2/3} \frac{\log n}{h_\mu} \log^{(2)} \left(\left(1+\frac{\varepsilon}{2}\right)^{2/3} \frac{\log n}{h_\mu} \right)} \right) \\
\overset{(2.11)}{\ll} &\sum_{n=1}^{\infty} \frac{1}{n} e\left(-\frac{(1+\varepsilon)}{h_\mu^{1/2}} \sqrt{2\rho_\mu \log n \log^{(3)} n} + \frac{(1+\varepsilon/2)}{h_\mu^{1/2}} \sqrt{2\rho_\mu \log n \log^{(3)} n} \right) \\
= &\sum_{n=1}^{\infty} \frac{1}{n} e\left(-\frac{\varepsilon/2}{h_\mu^{1/2}} \sqrt{2\rho_\mu \log n \log^{(3)} n} \right) \\
< &\infty.
\end{aligned}$$

The final line holds because the exponential term decays to zero faster than any negative power of the logarithm. Now applying Theorem 2.1 we have that $\mu(R_{\psi_\varepsilon^+}) = 0$.

Proof that $\mu(R_{\psi_\varepsilon^-}) = 1$. Fix $0 < \varepsilon < 1$. Define $\tilde{\psi}_\varepsilon^- : [0,\infty) \to [0,\infty)$ by

$$\tilde{\psi}_\varepsilon^-(x) = \frac{\log x}{h_\mu} + \frac{(1-\varepsilon)}{h_\mu^{3/2}} \sqrt{2\rho_\mu \log x \log^{(3)} x}.$$

Note that $\lfloor \tilde{\psi}_\varepsilon^-(n) \rfloor = \psi_\varepsilon^-(n)$ for all $n \in \mathbb{N}$. It is a straightforward albeit tedious calculation to verify that there exist constants $C_1, C_2 > 0$ such that

$$\frac{C_1}{x} \leq (\tilde{\psi}_\varepsilon^-)'(x) \leq \frac{C_2}{x} \tag{2.12}$$

for all x sufficiently large. We omit the details. It is also straightforward to show that $(\tilde{\psi}_\varepsilon^-)' : [0,\infty) \to [0,\infty)$ is decreasing. We define $m(n) \in [0,\infty)$ implicitly as the unique solution to the equation $\tilde{\psi}_\varepsilon^-(m(n)) = n$. Using the Mean Value Theorem together with (2.12) and the fact that $(\tilde{\psi}_\varepsilon^-)'$ is decreasing, we may deduce the following

$$\begin{aligned}
\#(\psi_\varepsilon^-)^{-1}(n) = \#\{m \in \mathbb{N} : \tilde{\psi}_\varepsilon^-(m) \in [n, n+1)\} &\gg |m(n+1) - m(n)| \\
&\geq ((\tilde{\psi}_\varepsilon^-)'(m(n)))^{-1} \\
&\gg \frac{m(n)}{C_2}.
\end{aligned}$$

It follows from the above that to obtain the cardinality bounds required by Corollary 2.3, it is sufficient to obtain a good lower bound for $m(n)$. By the definition of $m(n)$ and using the fact that $\tilde{\psi}_\varepsilon^-$ is increasing, it follows that to obtain the desired bound it is sufficient to show that

$$\tilde{\psi}_\varepsilon^-(e^{h_\mu n - (1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n \log^{(2)} n}}) < n$$

for all n sufficiently large. This we do below.

The following holds for all n sufficiently large:

$$\begin{aligned}
&\tilde{\psi}_\varepsilon^-(e^{h_\mu n-(1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n\log^{(2)}n}})\\
=&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}\\
&+\frac{(1-\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu\log\left(e^{h_\mu n-(1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n\log^{(2)}n}}\right)\log^{(3)}\left(e^{h_\mu n-(1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n\log^{(2)}n}}\right)}\\
=&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}\\
&+\frac{(1-\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu\left(h_\mu n-\left(1-\frac{\varepsilon}{2}\right)\sqrt{2\rho_\mu n\log^{(2)}n}\right)\log^{(3)}\left(e^{h_\mu n-(1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n\log^{(2)}n}}\right)}\\
\leq&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}+\frac{(1-\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu h_\mu n\log^{(3)}\left(e^{h_\mu n-(1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n\log^{(2)}n}}\right)}\\
=&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}+\frac{(1-\varepsilon)}{h_\mu}\sqrt{2\rho_\mu n\log^{(3)}\left(e^{h_\mu n-(1-\frac{\varepsilon}{2})\sqrt{2\rho_\mu n\log^{(2)}n}}\right)}\\
\leq&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}+\frac{(1-\varepsilon)}{h_\mu}\sqrt{2\rho_\mu n\log^{(3)}\left(e^{h_\mu n}\right)}\\
=&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}+\frac{(1-\varepsilon)}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}h_\mu n}\\
\leq&n-\frac{(1-\frac{\varepsilon}{2})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}+\frac{(1-\frac{3\varepsilon}{4})}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}\\
=&n-\frac{\frac{\varepsilon}{4}}{h_\mu}\sqrt{2\rho_\mu n\log^{(2)}n}\\
<&n.
\end{aligned}$$

It follows from the above that

$$m(n)\geq e^{h_\mu n-(1-\varepsilon/2)\sqrt{2\rho_\mu n\log\log n}}$$

for all n sufficiently large. This in turn implies that

$$\#(\psi_\varepsilon^-)^{-1}(n)\gg e^{h_\mu n-(1-\varepsilon/2)\sqrt{2\rho_\mu n\log\log n}}$$

for all n sufficiently large. Therefore the function ψ_ε^- satisfies the bound (2.9) and by Corollary 2.3 we may conclude that $\mu(R_{\psi_\varepsilon^-})=1$.

□

The following result gives a divergence condition in the form of (1.1) which guarantees that R_ψ has full measure.

Corollary 2.5. *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. Let $\psi\colon \mathbb{N}\mapsto\mathbb{N}$ be such that $\psi(n)\leq n$ for all $n\in\mathbb{N}$ and $\lim_{n\to\infty}\psi(n)=\infty$. Suppose that there exists a constant $K>0$ for which*

$$\sum_{n=1}^{\infty}e^{-h_\mu\psi(n)+K\sqrt{2\rho_\mu\psi(n)}}=\infty.$$

Then $\mu(R_\psi)=1$.

Proof. By our divergence assumption we have

$$\sum_{n=1}^{\infty} \#\psi^{-1}(n) e^{-h_\mu n + K\sqrt{2\rho_\mu n}} = \sum_{n=1}^{\infty} e^{-h_\mu \psi(n) + K\sqrt{2\rho_\mu \psi(n)}} = \infty.$$

Hence, there exists a sequence $(n_k)_{k=1}^{\infty}$ such that

$$\#\psi^{-1}(n_k) e^{-h_\mu n_k + K\sqrt{2\rho_\mu n_k}} > n_k^{-2} \tag{2.13}$$

holds for all $k \in \mathbb{N}$. Consider the set

$$B := \left\{ \mathtt{i} \in \Sigma_A : \limsup_{k\to\infty} \frac{\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k}} = \infty \right\}.$$

Appealing to the fact that μ satisfies (2.2), it can be shown that being an element of B is independent of any initial string of digits. Therefore B is a *tail event*, i.e. B belongs to the tail σ-algebra given by $\bigcap_{n=0}^{\infty} \sigma^{-n}\mathcal{B}^+$ where $\mathcal{B}^+$ is the Borel σ-algebra on $\Sigma_{A,*}$. It is well known that Gibbs measures for shifts of finite type are exact, i.e. the tail σ-algebra is the trivial σ-algebra (see [27, Remark 3, Page 29]). It follows from exactness that either $\mu(B) = 0$ or $\mu(B) = 1$. Let us suppose that $\mu(B) = 0$. Then there exists $M \in \mathbb{R}$ sufficiently large for which

$$B_M = \left\{ \mathtt{i} \in \Sigma_A : \limsup_{k\to\infty} \frac{\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k}} \leq M \right\}$$

has positive measure. Importantly B_M is also a tail event. This can also be shown using (2.2). Therefore by exactness we must have $\mu(B_M) = 1$. However this contradicts the Central Limit Theorem for Gibbs measures, see Theorem 3.1. Therefore we must have $\mu(B) = 1$.

Now fix an arbitrary $K' \in \mathbb{N}$ with $K' > K$ and define $g : \mathbb{N} \to \mathbb{N}$ according to the rule $g(n) = K'$. Using the fact that $\mu(B) = 1$ we may conclude that (2.7) from Theorem 2.2 is satisfied for this choice of g. Using (2.13) we have that

$$\#\psi^{-1}(n_k) e^{-h_\mu n_k + K'\sqrt{2\rho_\mu n_k}} > \frac{e^{(K'-K)\sqrt{2\rho_\mu n_k}}}{n_k^2} \to \infty \text{ as } k \to \infty.$$

Therefore (2.8) from Theorem 2.2 is satisfied. Our result now follows from Theorem 2.2. □

Observe that the negation of (2.8) from Theorem 2.2 is

$$\limsup_{n\to\infty} \#\psi^{-1}(n) e^{-h_\mu n + \sqrt{2\rho_\mu n} g(n)} < \infty.$$

Unfortunately, this condition does not correspond to the convergent series condition given in Theorem 2.1. Therefore we cannot use this condition to obtain a simple criterion for determining the measure of R_ψ.

After examining Theorem 2.1 and recalling the analogy made earlier with works on shrinking targets where formulas of the form (1.1) play an important role, a natural candidate for a divergence criterion for ensuring R_ψ will have full measure is

$$\sum_{n=1}^{\infty} e^{-h_\mu \psi(n) + \sqrt{2\rho_\mu \psi(n) \log\log \psi(n)}} = \infty.$$

The next theorem demonstrates that this intuitive guess is incorrect and that this divergence criterion alone is insufficient for determining the measure of R_ψ.

Theorem 2.6. *Let μ be a non-uniform Bernoulli measure and let $g\colon \mathbb{N} \mapsto [1,\infty)$ be a function satisfying $\limsup_{n\to\infty} \frac{g(n)}{\sqrt{\log\log(n)}} \leq 1$ and $\lim_{n\to\infty} g(n) = \infty$. Then there exists a monotone increasing function $\psi\colon \mathbb{N} \mapsto \mathbb{N}$ such that $\psi(n) \leq n$ for all $n \in \mathbb{N}$ and*

$$\sum_{n=1}^{\infty} e^{-h_\mu \psi(n) + \sqrt{2\rho_\mu \psi(n)} g(\psi(n))} = \infty$$

but $\mu(R_\psi) = 0$.

If we take $g(n) = \sqrt{\log\log n}$ in Theorem 2.6, we see that it is in conflict with the prediction made just before the statement of this theorem and thus this prediction cannot be correct.

What remains of the paper is organised as follows. In Section 3 we recall some useful results from Probability Theory. In Sections 4 and 5 we prove Theorems 2.1 and 2.2 respectively. In Section 6 we prove Theorem 2.6. Last of all, in Section 7 we apply our results to the study of dynamics on self-similar sets.

3 Probabilistic tools

In this section we recall some probabilistic results relating to Gibbs measures. The following two statements are the Central Limit Theorem and the Law of the Iterated Logarithm for Gibbs measures. These statements follow from [12, Corollary 1 and Corollary 2] and (2.1). We also refer the reader to [20, Proposition 2.1 and Proposition 2.2] where these results are stated in our language.

Theorem 3.1 (Central Limit Theorem)**.** *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. Then we have*

$$\lim_{n\to\infty} \mu\left(\left\{\mathtt{i} \in \Sigma_A : \frac{\log\mu([\mathtt{i}\,|_1^n]) + h_\mu n}{\sqrt{n}} \leq t\right\}\right) = \frac{1}{\sqrt{2\pi\rho_\mu^2}} \int_{-\infty}^{t} e^{-x^2/2\rho_\mu^2}\, dx$$

Theorem 3.2 (Law of the Iterated Logarithm)**.** *Let μ be the Gibbs measure for a Hölder continuous potential that is not cohomologous to a constant. Then we have*

$$\limsup_{n\to\infty} \frac{\log\mu([\mathtt{i}\,|_1^n]) + h_\mu n}{\sqrt{2\rho_\mu n \log\log n}} = 1 \quad \text{for } \mu\text{-almost every } \mathtt{i} \in \Sigma_A.$$

We will also use the following well known property of Gibbs measures for Hölder continuous potentials which states that they have exponential decay of correlations. A proof of the following statement is contained within the proof of [8, Proposition 1.14].

Theorem 3.3 (Exponential Decay of Correlations)**.** *Let μ be the Gibbs measure for a Hölder continuous potential. Then there exist constants $D > 0$ and $0 < \gamma < 1$ such that for every $\mathtt{i}, \mathtt{j} \in \Sigma_{A,*}$ and $n \geq |\,\mathtt{i}\,|$ we have*

$$\left|\mu([\mathtt{i}] \cap \sigma^{-n}[\mathtt{j}]) - \mu([\mathtt{i}])\mu([\mathtt{j}])\right| \leq D\gamma^{n-|\,\mathtt{i}\,|}\mu([\mathtt{i}])\mu([\mathtt{j}]).$$

In the special case of Bernoulli measures, we will make use of the following more sophisticated estimate in the Central Limit Theorem. A proof of this statement can be found in [6, Theorem 9.4].

Theorem 3.4. *Let μ be a non-uniform Bernoulli measure and let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers such that $\lim_{n\to\infty} a_n = \infty$ and $\lim_{n\to\infty} \frac{a_n}{\sqrt{n}} = 0$. Then*

$$\mu\left(\left\{\mathtt{i} \in \Sigma : \frac{\log \mu([\mathtt{i}\,|_1^n]) + h_\mu n}{\sqrt{\rho_\mu}} \geq a_n\sqrt{n}\right\}\right) = e^{-a_n^2(1+\zeta_n)/2},$$

for some sequence of real numbers $(\zeta_n)_{n=1}^{\infty}$ satisfying $\lim_{n\to\infty} \zeta_n = 0$.

Finally, from time-to-time, we will also call upon the first Borel–Cantelli Lemma, see for example [15, Lemma 1.2].

Lemma 3.5. *Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $(A_j)_{j=1}^{\infty}$ be a collection of measurable subsets of X. If*

$$\sum_{j=1}^{\infty} \mu(A_j) < \infty,$$

then

$$\mu\left(\{x \in X : x \in A_j \text{ for infinitely many } j \in \mathbb{N}\}\right) = 0.$$

4 Proof of Theorem 2.1

In this section we prove Theorem 2.1. Our proof relies upon careful use of Theorem 3.2 and Lemma 3.5. To successfully apply Theorem 3.2 we exhaust Σ_A up to a set of measure zero by sets wherein the conclusion of Theorem 3.2 can be made uniform.

Proof. Let $\psi\colon \mathbb{N} \mapsto \mathbb{N}$ be a function satisfying

$$\sum_{n=1}^{\infty} e^{-h_\mu\psi(n)+(1+\varepsilon)\sqrt{2\rho_\mu\psi(n)\log\log\psi(n)}} < \infty \tag{4.1}$$

for some $\varepsilon > 0$. Define $\tilde{\psi} : \mathbb{N} \to \mathbb{N}$ according to the rule

$$\tilde{\psi}(n) = \begin{cases} \psi(n) & \text{if } \psi(n) \leq n, \\ n & \text{if } \psi(n) > n. \end{cases}$$

Notice that $\tilde{\psi}(n) \leq \psi(n)$ for all $n \in \mathbb{N}$. This implies $R_\psi \subset R_{\tilde{\psi}}$. We also have

$$\begin{aligned} &\sum_{n=1}^{\infty} e^{-h_\mu\tilde{\psi}(n)+(1+\varepsilon)\sqrt{2\rho_\mu\tilde{\psi}(n)\log\log\tilde{\psi}(n)}} \\ \leq &\sum_{n=1}^{\infty} e^{-h_\mu\psi(n)+(1+\varepsilon)\sqrt{2\rho_\mu\psi(n)\log\log\psi(n)}} + \sum_{n=1}^{\infty} e^{-h_\mu n+(1+\varepsilon)\sqrt{2\rho_\mu n\log\log n}} \\ <&\infty. \end{aligned}$$

Therefore $\tilde{\psi}$ also satisfies (4.1). This means that without loss of generality we may assume that our original ψ satisfies $\psi(n) \leq n$ for all $n \in \mathbb{N}$. The rest of our proof is given under this assumption.

By Theorem 3.2, $\mu\left(\bigcup_{N=1}^{\infty} E_N\right) = 1$, where

$$E_N = \left\{ \mathtt{i} \in \Sigma_A : \mu([\mathtt{i}\,|_1^n]) \leq e^{-h_\mu n + (1+\varepsilon)\sqrt{2\rho_\mu n \log\log n}} \text{ for all } n \geq N \right\}.$$

Thus, to prove $\mu(R_\psi) = 0$ it is enough to show that $\mu(E_N \cap R_\psi) = 0$ for every $N \in \mathbb{N}$. This we do below. In what follows $N \in \mathbb{N}$ is fixed.

For each $n \in \mathbb{N}$ let

$$F_{A,n} = \{\mathtt{i} \in \Sigma_{A,n} : \mu([\mathtt{i}]) \leq e^{-h_\mu n + (1+\varepsilon)\sqrt{2\rho_\mu n \log\log n}}\}.$$

We also let

$$M(N) := \max\{k \geq 1 : \psi(k) < N\} + 1.$$

By the expression for R_ψ given in (2.6), we know that

$$E_N \cap R_\psi = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \bigcup_{\substack{\mathtt{i}\in\Sigma_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j\,i}\in\Sigma_{A,k+\psi(k)}}} E_N \cap [\mathtt{i\,j\,i}].$$

Therefore to prove $\mu(R_\psi) = 0$, by the first Borel–Cantelli Lemma (Lemma 3.5) it is sufficient to show that

$$\sum_{k=1}^{\infty} \sum_{\substack{\mathtt{i}\in\Sigma_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j\,i}\in\Sigma_{A,k+\psi(k)}}} \mu(E_N \cap [\mathtt{i\,j\,i}]) < \infty. \tag{4.2}$$

This we do presently. We begin by observing that

$$\begin{aligned}
\sum_{k=1}^{\infty} \sum_{\substack{\mathtt{i}\in\Sigma_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j\,i}\in\Sigma_{A,k+\psi(k)}}} \mu(E_N \cap [\mathtt{i\,j\,i}]) &\leq M(N) + \sum_{k=M(N)+1}^{\infty} \sum_{\substack{\mathtt{i}\in\Sigma_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j\,i}\in\Sigma_{A,k+\psi(k)}}} \mu(E_N \cap [\mathtt{i\,j\,i}]) \\
&\leq M(N) + \sum_{k=M(N)+1}^{\infty} \sum_{\substack{\mathtt{i}\in F_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j\,i}\in\Sigma_{A,k+\psi(k)}}} \mu([\mathtt{i\,j\,i}]).
\end{aligned}$$

Using (2.2) and the definition of $F_{A,\psi(k)}$, it follows that

$$\begin{aligned}
&\sum_{k=1}^{\infty} \sum_{\substack{\mathtt{i}\in\Sigma_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j\,i}\in\Sigma_{A,k+\psi(k)}}} \mu(E_N \cap [\mathtt{i\,j\,i}]) \\
\leq & M(N) + C \sum_{k=M(N)+1}^{\infty} \sum_{\substack{\mathtt{i}\in F_{A,\psi(k)} \\ \mathtt{j}\in\Sigma_{A,k-\psi(k)} \\ \mathtt{i\,j}\in\Sigma_{A,k}}} \mu([\mathtt{i\,j}]) e^{-h_\mu \psi(k) + (1+\varepsilon)\sqrt{2\rho_\mu \psi(k) \log\log \psi(k)}} \\
\leq & M(N) + C \sum_{k=M(N)+1}^{\infty} e^{-h_\mu \psi(k) + (1+\varepsilon)\sqrt{2\rho_\mu \psi(k) \log\log \psi(k)}}
\end{aligned}$$

$<\infty$.

In the final line we have used (4.1). Thus we have shown that (4.2) holds and this completes our proof. □

5 Proof of Theorem 2.2

In this section we prove Theorem 2.2. We now fix μ, ψ, $(n_k)_{k=1}^{\infty}$, and g so that the assumptions of this theorem are satisfied. Let

$$S_{\psi,(n_k)} := \left\{ \mathtt{i} \in \Sigma_A : d(\sigma^p \mathtt{i}, \mathtt{i}) \leq K^{-n_k} \text{ for some } p \in \psi^{-1}(n_k) \text{ for infinitely many } k \in \mathbb{N} \right\}.$$

By changing the values that ψ takes for at most finitely many natural numbers we may ensure that $\psi(n) \leq n$ for all $n \in \mathbb{N}$. This assumption does not change the set $S_{\psi,(n_k)}$ or any of our underlying assumptions. As such, without loss of generality we can give our proof under this assumption.

To prove that $\mu(S_{\psi,(n_k)}) = 1$, we will show that there exists a constant $c > 0$ such that for any $\mathtt{i} \in \Sigma_{A,*}$ we have

$$\mu([\mathtt{i}] \cap S_{\psi,(n_k)}) \geq c\mu([\mathtt{i}]). \tag{5.1}$$

It then follows from [4, Lemma 6] that $\mu(S_{\psi,(n_k)}) = 1$. We can think of [4, Lemma 6] as essentially providing us with an analogue of the classical Lebesgue Density Theorem (see, for example, [26, Corollary 2.14]) that holds for more general measures. Since $S_{\psi,(n_k)} \subset R_\psi$ we also immediately have that $\mu(R_\psi) = 1$. Therefore to complete the proof of Theorem 2.2 it remains to show that (5.1) holds. To prove that (5.1) holds we construct a suitable subset of $[\mathtt{i}] \cap S_{\psi,(n_k)}$ as the infinite intersection of sets $D_k(\mathtt{i})$ defined inductively. Each $D_k(\mathtt{i})$ will be a finite union of cylinders. Given one of these cylinders we construct a family of sub-cylinders with the important property that the recurrence inequality will be satisfied at times where the corresponding prefix we return to has large measure. Here large measure means that the inequality appearing in (2.7) in Theorem 2.2 is satisfied. We define $D_{k+1}(\mathtt{i})$ to be the resulting collection of cylinders. Crucially, we can choose the times appearing in this construction to be a suitably sparse set so that we can apply Theorem 3.3 in a meaningful way. In particular, we can conclude that $D_{k+1}(\mathtt{i})$ has the same measure as $D_k(\mathtt{i})$ up to an error that converges to zero sufficiently fast as $k \to \infty$. It is at this point in the argument where we use our assumption (2.8).

We split our proof of (5.1) into three parts below.

Part 1. Construction of the auxiliary sets.

It follows from (2.2) that the set

$$\left\{ \mathtt{i} \in \Sigma_A : \limsup_{k\to\infty} \frac{\log \mu([\mathtt{i}|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k} g(n_k)} > 1 + \varepsilon \right\}$$

is σ-invariant for every $\varepsilon > 0$. Using the fact that μ is ergodic together with (2.7), we may conclude that there exists $\varepsilon > 0$ such that

$$\mu\left(\left\{ \mathtt{i} \in \Sigma_A : \limsup_{k\to\infty} \frac{\log \mu([\mathtt{i}|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k} g(n_k)} > 1 + \varepsilon \right\} \right) = 1. \tag{5.2}$$

For the rest of our proof $\varepsilon > 0$ is fixed so that (5.2) is satisfied.

Without loss of generality, we may assume that our sequence $(n_k)_{k=1}^{\infty}$ is such that

$$\frac{e^{\varepsilon\sqrt{2n_k\rho_\mu}g(n_k)/4}}{1+n_k+\lceil\frac{-h_\mu}{\log\gamma}\rceil\cdot n_k}\geq D+1 \tag{5.3}$$

for every $k\in\mathbb{N}$. Here and throughout our proof, $\gamma\in(0,1)$ and $D>0$ are as in Theorem 3.3.

Given $\delta>0$ we let $a_1(\delta)\in\mathbb{N}$ be such that for every $n_k\geq a_1(\delta)$ we have

$$\exp\left(-\#\psi^{-1}(n_k)e^{-h_\mu n_k+\sqrt{2\rho_\mu n_k}g(n_k)}\right)\leq\delta. \tag{5.4}$$

The fact $a_1(\delta)$ always exists is a consequence of (2.8).

For simplicity, we let

$$F:=\min_{\mathtt{i}\in\Sigma_A}f(\mathtt{i})-P(f).$$

For $\delta>0$ and $m\in\mathbb{N}$, let $a(m,\delta)>\max\{m,a_1(\delta)\}$ be such that for every $n\geq a(m,\delta)$ we have

$$\left|\frac{-2\log C+m(F+h_\mu)}{\sqrt{2\rho_\mu n}g(n)}\right|<\varepsilon/4\quad\text{and}\quad\frac{\sqrt{n-m}g(n-m)}{\sqrt{n}g(n)}\geq\frac{1+\varepsilon/2}{1+3\varepsilon/4}. \tag{5.5}$$

The latter inequality is possible because of our assumption that $\lim_{n\to\infty}\frac{g(n)}{g(n+1)}=1$. Here C is the constant appearing in (2.1) and (2.2). The equations (2.1), (2.2) and (5.5) imply the following estimate.

Lemma 5.1. *For every $\mathtt{i}\in\Sigma_{A,*}$, if $\mathtt{j}\in\Sigma_{A,*}$ satisfies $|\,\mathtt{j}\,|>a(|\,\mathtt{i}\,|,\delta)$, $\mathtt{i}\,\mathtt{j}\in\Sigma_{A,*}$, and*

$$\frac{\log\mu([\mathtt{j}])+h_\mu|\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{j}\,|}g(|\,\mathtt{j}\,|)}\geq1+\frac{3\varepsilon}{4},$$

then

$$\frac{\log\mu([\mathtt{i}\,\mathtt{j}])+h_\mu|\,\mathtt{i}\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}\geq1+\frac{\varepsilon}{4}.$$

Proof. First, we note that it follows from (2.2) that

$$\frac{\log\mu([\mathtt{i}\,\mathtt{j}])+h_\mu|\,\mathtt{i}\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}\geq\frac{-\log C+\log\mu([\mathtt{i}])+h_\mu|\,\mathtt{i}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}+\frac{\log\mu([\mathtt{j}])+h_\mu|\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}.$$

It then follows from (2.1) that

$$\begin{aligned}\frac{\log\mu([\mathtt{i}\,\mathtt{j}])+h_\mu|\,\mathtt{i}\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}&\geq\frac{-2\log C+\sum_{k=0}^{|\,\mathtt{i}\,|-1}f(\sigma^k\,\mathtt{i})-|\,\mathtt{i}\,|P(f)+h_\mu|\,\mathtt{i}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}+\frac{\log\mu([\mathtt{j}])+h_\mu|\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}\\&\geq\frac{-2\log C+|\,\mathtt{i}\,|(F+h_\mu)}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}+\frac{\log\mu([\mathtt{j}])+h_\mu|\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}.\end{aligned}$$

Next we rewrite the fraction on the far right-hand side and then apply the second inequality of (5.5) and the assumption of the lemma to obtain

$$\begin{aligned}\frac{\log\mu([\mathtt{i}\,\mathtt{j}])+h_\mu|\,\mathtt{i}\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}&\geq\frac{-2\log C+|\,\mathtt{i}\,|(F+h_\mu)}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}+\left(\frac{\log\mu([\mathtt{j}])+h_\mu|\,\mathtt{j}\,|}{\sqrt{2\rho_\mu|\,\mathtt{j}\,|}g(|\,\mathtt{j}\,|)}\right)\left(\frac{\sqrt{2\rho_\mu|\,\mathtt{j}\,|}g(|\,\mathtt{j}\,|)}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}\right)\\&\geq\frac{-2\log C+|\,\mathtt{i}\,|(F+h_\mu)}{\sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)}+\left(1+\frac{\varepsilon}{2}\right).\end{aligned}$$

Finally, the claimed bound is now obtained by applying the first inequality of (5.5). □

For every $\delta > 0$, and $m \in \mathbb{N}$, we define $b(m,\delta)$ to be a sufficiently large natural number for which the set

$$G_{h,m,\delta} := \Bigg\{ \mathtt{j} \in [h] : \text{there exists } n_k \in [a(m,\delta), b(m,\delta)] \text{ such that } \frac{\log \mu([\mathtt{j}\,|_1^{n_k-m}]) + h_\mu(n_k-m)}{\sqrt{2\rho_\mu(n_k-m)}g(n_k-m)} \geq 1 + \frac{3\varepsilon}{4} \Bigg\}$$

satisfies

$$\mu\left(G_{h,m,\delta}\right) \geq (1-\delta)\mu([h]) \tag{5.6}$$

for all $h \in \{1,\dots,K\}$. The existence of $b(m,\delta)$ follows from (5.2) and an application of (2.1), (2.2), and our assumptions on g. By definition we have

$$G_{h,m,\delta} \cap G_{h',m,\delta} = \emptyset \text{ for } h \neq h'. \tag{5.7}$$

We can partition each $G_{h,m,\delta}$ into a disjoint collection of cylinders as follows. For each $\ell \in \mathbb{N}$ such that $n_\ell \in [a(m,\delta), b(m,\delta)]$ let

$$\begin{aligned} G^{(\ell)}_{h,m,\delta} = \Bigg\{ \mathtt{j} \in \Sigma_{A,n_\ell-m} : j_1 = h, \quad & \frac{\log \mu([\mathtt{j}\,|_1^{n_\ell-m}]) + h_\mu(n_\ell-m)}{\sqrt{2\rho_\mu(n_\ell-m)}g(n_\ell-m)} \geq 1 + \frac{3\varepsilon}{4} \\ & \text{but } \frac{\log \mu([\mathtt{j}\,|_1^{n_k-m}]) + h_\mu(n_k-m)}{\sqrt{2\rho_\mu(n_k-m)}g(n_k-m)} < 1 + \frac{3\varepsilon}{4} \text{ for all } n_k \in [a(m,\delta), n_\ell - 1] \Bigg\}. \end{aligned} \tag{5.8}$$

Then

$$G_{h,m,\delta} = \bigcup_{\ell : n_\ell \in [a(m,\delta), b(m,\delta)]} [G^{(\ell)}_{h,m,\delta}] \quad \text{and} \quad [G^{(\ell_1)}_{h,m,\delta}] \cap [G^{(\ell_2)}_{h,m,\delta}] = \emptyset$$

for $\ell_1 \neq \ell_2$.

Now we define two sets of words whose corresponding elements of Σ_A either satisfy or fail a recurrence property defined in terms of a cylinder of length n_ℓ. Given $m \in \mathbb{N}$ and $\delta > 0$ let

$$c(m,\delta) := \max \left\{ \bigcup_{k=1}^{b(m,\delta)} \psi^{-1}(k) \right\} + b(m,\delta).$$

This significance of the quantity $c(m,\delta)$ is seen as follows. Given $\mathtt{i} \in \Sigma_{A,*}$ and $\delta > 0$, using the fact that $\psi(n) \leq n$ for all $n \in N$ and the definition of $c(|\,\mathtt{i}\,|,\delta)$, it follows that for any $n_\ell \in [a(|\,\mathtt{i}\,|,\delta), b(|\,\mathtt{i}\,|,\delta)]$ we have

$$\psi^{-1}(n_\ell) \subset [n_\ell, c(|\,\mathtt{i}\,|,\delta) - n_\ell]. \tag{5.9}$$

This means that if we are interested in sequences satisfying $d(\sigma^p \mathtt{j}, \mathtt{j}) \leq K^{-n_\ell}$ for some $n_\ell \in [a(|\,\mathtt{i}\,|,\delta), b(|\,\mathtt{i}\,|,\delta)]$ and $p \in \psi^{-1}(n_\ell)$, then it is sufficient to know only the first $c(|\,\mathtt{i}\,|,\delta)$ entries in $\mathtt{j}$. This fact underpins the definitions of $B^{(\ell)}_{\mathtt{i},\delta}$ and $C^{(\ell)}_{\mathtt{i},\delta}$ below.

For a word $\mathtt{i} \in \Sigma_{A,*}$, $\delta > 0$ and $l \in \mathbb{N}$ such that $n_\ell \in [a(|\,\mathtt{i}\,|,\delta), b(|\,\mathtt{i}\,|,\delta)]$, let

$$B^{(\ell)}_{\mathtt{i},\delta} := \Bigg\{ \mathtt{j} \in \Sigma_{A,c(|\,\mathtt{i}\,|,\delta)-|\,\mathtt{i}\,|} : \mathtt{j}\,|_1^{n_\ell-|\,\mathtt{i}\,|} \in \bigcup_{h:\,\mathtt{i}\,h \in \Sigma_{A,*}} G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta} \text{ and } \mathtt{j}\,|_{p-|\,\mathtt{i}\,|+1}^{p-|\,\mathtt{i}\,|+n_\ell} = (\mathtt{i}\,\mathtt{j})|_1^{n_\ell} \text{ for some } p \in \psi^{-1}(n_\ell) \Bigg\}$$

and

$$C^{(\ell)}_{\mathtt{i},\delta} := \Big\{ \mathtt{j} \in \Sigma_{A,c(|\mathtt{i}|,\delta)-|\mathtt{i}|} : \mathtt{j}\,|_1^{n_\ell-|\mathtt{i}|} \in \bigcup_{h:\,\mathtt{i}\,h\in\Sigma_{A,*}} G^{(\ell)}_{h,|\mathtt{i}|,\delta} \text{ and } \mathtt{j}\,|_{p-|\mathtt{i}|+1}^{p-|\mathtt{i}|+n_\ell} \neq (\mathtt{i}\,\mathtt{j})|_1^{n_\ell} \text{ for every } p \in \psi^{-1}(n_\ell) \Big\}.$$

We also let

$$B_{\mathtt{i},\delta} := \bigcup_{\ell: n_\ell \in [a(|\mathtt{i}|,\delta), b(|\mathtt{i}|,\delta)]} B^{(\ell)}_{\mathtt{i},\delta}.$$

An important consequence of the definition of $B^{(\ell)}_{\mathtt{i},\delta}$ is that if $\mathtt{j}' \in \mathtt{i}\,B^{(\ell)}_{\mathtt{i},\delta}$ then

$$d(\sigma^p\,\mathtt{j}', \mathtt{j}') \le K^{-n_\ell} \tag{5.10}$$

for some $p \in \psi^{-1}(n_\ell)$.

Part 2. Measure properties of our auxiliary sets.

Lemma 5.2. *For every $\mathtt{i} \in \Sigma_{A,*}$, $\delta > 0$, and $\ell \in \mathbb{N}$ such that $n_\ell \in [a(|\mathtt{i}|,\delta), b(|\mathtt{i}|,\delta)]$, we have*

$$\mu([\mathtt{i}\,C^{(\ell)}_{\mathtt{i},\delta}]) \le C\delta \sum_{h:\,\mathtt{i}\,h\in\Sigma_{A,*}} \mu([\mathtt{i}\,G^{(\ell)}_{h,|\mathtt{i}|,\delta}]).$$

Proof. Let $r_\ell = \lceil \frac{-h_\mu}{\log\gamma} \rceil \cdot n_\ell$. Let S denote the $(n_\ell + r_\ell)$-separated subset of $\psi^{-1}(n_\ell)$ defined inductively as follows. Let $S_1 = \{\min \psi^{-1}(n_\ell)\}$ and for $k \ge 2$ let

$$S_k = S_{k-1} \cup \left\{ \min \psi^{-1}(n_\ell) \cap [\max S_{k-1} + n_\ell + r_\ell + 1, \infty) \right\}.$$

We then let $S = \bigcup_k S_k$. Clearly,

$$\frac{\#\psi^{-1}(n_\ell)}{n_\ell + r_\ell + 1} \le \#S \le \#\psi^{-1}(n_\ell). \tag{5.11}$$

By (5.9) we know that $S \subset [n_\ell, c(\mathtt{i},\delta) - n_\ell]$. It follows now from the definition of $C^{(\ell)}_{\mathtt{i},\delta}$ and (5.7) that

$$\begin{aligned}
&\mu([\mathtt{i}\,C^{(\ell)}_{\mathtt{i},\delta}]) \\
\le&\mu\left(\left\{ \mathtt{j} \in \Sigma_A : \mathtt{j}\,|_1^{|\mathtt{i}|} = \mathtt{i},\ \mathtt{j}\,|_{|\mathtt{i}|+1}^{n_\ell} \in \bigcup_{h:\,\mathtt{i}\,h\in\Sigma_{A,*}} G^{(\ell)}_{h,|\mathtt{i}|,\delta} \text{ and } \mathtt{j}\,|_{p+1}^{p+n_\ell} \neq \mathtt{j}\,|_1^{n_\ell} \text{ for every } p \in S \right\} \right) \\
=& \sum_{h:\,\mathtt{i}\,h\in\Sigma_{A,*}} \sum_{\mathtt{j}\in G^{(\ell)}_{h,|\mathtt{i}|,\delta}} \mu\left([\mathtt{i}\,\mathtt{j}] \cap \left(\bigcap_{p\in S} \sigma^{-p}[\mathtt{i}\,\mathtt{j}]^c \right) \right).
\end{aligned}$$

Combining the above with Theorem 3.3 and (2.2), we have

$$\mu([\mathtt{i}\,C^{(\ell)}_{\mathtt{i},\delta}]) \le C \sum_{h:\,\mathtt{i}\,h\in\Sigma_{A,*}} \sum_{\mathtt{j}\in G^{(\ell)}_{h,|\mathtt{i}|,\delta}} \mu([\mathtt{i}\,\mathtt{j}])\,(1 + D\gamma^{r_\ell})^{\#S}\,(1 - \mu([\mathtt{i}\,\mathtt{j}]))^{\#S}.$$

Now using the fact that $1+x \le e^x$ for all $x \in \mathbb{R}$ together with (5.11), it follows that

$$\begin{aligned}\mu([\mathtt{i}\, C^{(\ell)}_{\mathtt{i},\delta}]) &\le C \sum_{h\colon \mathtt{i}\, h \in \Sigma_{A,*}} \sum_{\mathtt{j} \in G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}} \mu([\mathtt{i}\,\mathtt{j}]) \exp\left(\#SD\gamma^{r_\ell}\right) \exp\left(-\#S\mu([\mathtt{i}\,\mathtt{j}])\right) \\ &\le C \sum_{h\colon \mathtt{i}\, h \in \Sigma_{A,*}} \sum_{\mathtt{j} \in G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}} \mu([\mathtt{i}\,\mathtt{j}]) \exp\left(D\gamma^{r_\ell}\#\psi^{-1}(n_\ell) - \mu([\mathtt{i}\,\mathtt{j}])\frac{\#\psi^{-1}(n_\ell)}{n_\ell + r_\ell + 1}\right).\end{aligned}$$

By definition, see (5.8), any word $\mathtt{j} \in G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}$ has $|\,\mathtt{j}\,| = n_\ell - |\,\mathtt{i}\,|$. Furthermore, it also follows directly from the definition of $G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}$ that the hypothesis of Lemma 5.1 holds, and so we have

$$\mu([\mathtt{i}\,\mathtt{j}]) \ge \exp\left(-h_\mu|\,\mathtt{i}\,\mathtt{j}\,| + \sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)(1+\varepsilon/4)\right).$$

Putting this all together, we conclude that

$$\begin{aligned}&\mu([\mathtt{i}\, C^{(\ell)}_{\mathtt{i},\delta}]) \\ \le& C \sum_{h\colon \mathtt{i}\, h \in \Sigma_{A,*}} \sum_{\mathtt{j} \in G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}} \mu([\mathtt{i}\,\mathtt{j}]) \exp\left(D\gamma^{r_\ell}\#\psi^{-1}(n_\ell) - \frac{\#\psi^{-1}(n_\ell)}{n_\ell + r_\ell + 1} e^{-h_\mu|\,\mathtt{i}\,\mathtt{j}\,| + \sqrt{2\rho_\mu|\,\mathtt{i}\,\mathtt{j}\,|}g(|\,\mathtt{i}\,\mathtt{j}\,|)(1+\varepsilon/4)}\right) \\ \le& C \sum_{h\colon \mathtt{i}\, h \in \Sigma_{A,*}} \mu([\mathtt{i}\, G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}]) \exp\left(D\gamma^{r_\ell}\#\psi^{-1}(n_\ell) - \frac{\#\psi^{-1}(n_\ell)}{n_\ell + r_\ell + 1} e^{-h_\mu n_\ell + \sqrt{2\rho_\mu n_\ell}g(n_\ell)(1+\varepsilon/4)}\right).\end{aligned}$$

By equation (5.3) and the definition of r_ℓ we have

$$e^{-h_\mu n_\ell + \sqrt{2\rho_\mu n_\ell}g(n_\ell)}\left(\frac{e^{\varepsilon\sqrt{2\rho_\mu n_\ell}g(n_\ell)/4}}{n_\ell + r_\ell + 1} - 1\right) \ge D\gamma^{r_\ell}.$$

Thus, it follows that

$$\mu([\mathtt{i}\, C^{(\ell)}_{\mathtt{i},\delta}]) \le C \sum_{h\colon \mathtt{i}\, h \in \Sigma_{A,*}} \mu([\mathtt{i}\, G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}]) \exp\left(-\#\psi^{-1}(n_\ell) e^{-h_\mu n_\ell + \sqrt{2\rho_\mu n_\ell}g(n_\ell)}\right).$$

Finally, from the definition of $a(|\,\mathtt{i}\,|,\delta)$ and (5.4), we see that

$$\mu([\mathtt{i}\, C^{(\ell)}_{\mathtt{i},\delta}]) \le C\delta \sum_{h\colon \mathtt{i}\, h \in \Sigma_{A,*}} \mu([\mathtt{i}\, G^{(\ell)}_{h,|\,\mathtt{i}\,|,\delta}]),$$

which completes our proof. □

Part 3. Proof of (5.1)**.**

Let us fix now an arbitrary $\mathtt{i} \in \Sigma_{A,*}$ and set about proving (5.1). Recall that we defined

$$S_{\psi,(n_k)} := \left\{\mathtt{i} \in \Sigma_A : d(\sigma^p\,\mathtt{i}, \mathtt{i}) \le K^{-n_k} \text{ for some } p \in \psi^{-1}(n_k) \text{ for infinitely many } k \in \mathbb{N}\right\}.$$

We now define a subset of $S_{\psi,(n_k)}$ by induction as follows. First, let k_1 be such that $D\gamma^{k_1} < 1$ and let

$$D_1(\mathtt{i}) := \bigcup_{\substack{\mathtt{k} \in \Sigma_{A,1+k_1} \\ \mathtt{i}\,\mathtt{k} \in \Sigma_{A,*}}} [\mathtt{i}\,\mathtt{k}\, B_{\mathtt{i}\,\mathtt{k},2^{-1}/C}].$$

By the definition of $D_1(\mathtt{i})$, and noting that each $B_{\mathtt{i}\,\mathtt{k},2^{-1}/C}$ is a finite union of cylinders, we may choose a finite set of words $W_1(\mathtt{i}) \subset \Sigma_{A,*}$ such that

$$D_1(\mathtt{i}) = \bigcup_{\mathtt{j}\in W_1(\mathtt{i})} [\mathtt{j}] \text{ and } [\mathtt{j}_1] \cap [\mathtt{j}_2] = \emptyset \text{ for } \mathtt{j}_1 \neq \mathtt{j}_2\,.$$

Now suppose $D_k(\mathtt{i}) \subset \Sigma_A$ and $W_k(\mathtt{i}) \subset \Sigma_{A,*}$ are defined for some $k \geq 1$. We then define

$$D_{k+1}(\mathtt{i}) = \bigcup_{\mathtt{j}\in W_k(\mathtt{i})} \bigcup_{\substack{\mathtt{k}\in\Sigma_{A,k+1+k_1}\\ \mathtt{j}\,\mathtt{k}\in\Sigma_{A,*}}} [\mathtt{j}\,\mathtt{k}\,B_{\mathtt{j}\,\mathtt{k},2^{-k-1}/C}].$$

We also let $W_{k+1}(\mathtt{i}) \subset \Sigma_{A,*}$ be a finite set of words such that

$$D_{k+1}(\mathtt{i}) = \bigcup_{\mathtt{j}\in W_{k+1}(\mathtt{i})} [\mathtt{j}] \text{ and } [\mathtt{j}_1] \cap [\mathtt{j}_2] = \emptyset \text{ for } \mathtt{j}_1 \neq \mathtt{j}_2\,.$$

Proceeding inductively we define $D_k(\mathtt{i})$ and $W_k(\mathtt{i})$ for all $k \in \mathbb{N}$. We observe that

$$D_{k+1}(\mathtt{i}) \subset D_k(\mathtt{i}) \tag{5.12}$$

for all $k \geq 1$, and

$$\bigcap_{k=1}^{\infty} D_k(\mathtt{i}) \subset [\mathtt{i}] \cap S_{\psi,(n_k)}. \tag{5.13}$$

This last inclusion follows from (5.10). Therefore to prove (5.1) it is sufficient to obtain lower bounds for the measure of $\cap_{k=1}^{\infty} D_k(\mathtt{i})$. The following lemma allows us to do this and bounds how much measure is lost as we pass from $D_k(\mathtt{i})$ to $D_{k+1}(\mathtt{i})$.

Lemma 5.3. *For every $k \geq 2$ we have*

$$\mu(D_k(\mathtt{i})) \geq \mu(D_{k-1}(\mathtt{i}))(1 - D\gamma^{k+k_1})(1 - 2^{-k})(1 - 2^{-k}/C),$$

and in the case when $k = 1$ we have

$$\mu(D_1(\mathtt{i})) \geq \mu([\mathtt{i}])(1 - D\gamma^{k_1})(1 - 2^{-1})(1 - 2^{-1}/C).$$

Proof. Let $k \geq 2$. By the definition of $D_k(\mathtt{i})$, and recalling the definitions of $B_{\mathtt{i},\delta}$, $B^{(\ell)}_{\mathtt{i},\delta}$, and $C^{(\ell)}_{\mathtt{i},\delta}$ as well as noting that for different values of ℓ the sets $G^{(\ell)}_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C}$ are disjoint by construction, we have

$$\begin{aligned}
&\mu(D_k(\mathtt{i}))\\
&\quad= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\substack{\mathtt{k}\in\Sigma_{A,k+k_1}\\ \mathtt{j}\,\mathtt{k}\in\Sigma_{A,*}}} \mu([\mathtt{j}\,\mathtt{k}\,B_{\mathtt{j}\,\mathtt{k},2^{-k}/C}])\\
&\quad= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\substack{\mathtt{k}\in\Sigma_{A,k+k_1}\\ \mathtt{j}\,\mathtt{k}\in\Sigma_{A,*}}} \sum_{\ell: n_\ell \in [a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C), b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \mu([\mathtt{j}\,\mathtt{k}\,B^{(\ell)}_{\mathtt{j}\,\mathtt{k},2^{-k}/C}])\\
&\quad= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\substack{\mathtt{k}\in\Sigma_{A,k+k_1}\\ \mathtt{j}\,\mathtt{k}\in\Sigma_{A,*}}} \sum_{\ell: n_\ell \in [a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C), b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \sum_{h:\,\mathtt{j}\,\mathtt{k}\,h\in\Sigma_{A,*}} \mu([\mathtt{j}\,\mathtt{k}\,G^{(\ell)}_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C}])\\
&\qquad\qquad - \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\substack{\mathtt{k}\in\Sigma_{A,k+k_1}\\ \mathtt{j}\,\mathtt{k}\in\Sigma_{A,*}}} \sum_{\ell: n_\ell \in [a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C), b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \mu([\mathtt{j}\,\mathtt{k}\,C^{(\ell)}_{\mathtt{j}\,\mathtt{k},2^{-k}/C}]).
\end{aligned}$$

Now using Lemma 5.2 in the above we see that

$$\begin{aligned}
&\frac{\mu(D_k(\mathtt{i}))}{1-2^{-k}}\\
&\geq \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\substack{\mathtt{k}\in\Sigma_{A,k+k_1}\\ \mathtt{j}\,\mathtt{k}\in\Sigma_{A,*}}} \sum_{\ell:n_\ell\in[a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C),b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \sum_{h:\,\mathtt{j}\,\mathtt{k}\,h\in\Sigma_{A,*}} \mu([\mathtt{j}\,\mathtt{k}\,G^{(\ell)}_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C}])\\
&= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\ell:n_\ell\in[a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C),b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \sum_{h\in\{1,\dots,K\}} \sum_{\substack{\mathtt{k}\in\Sigma_{A,k+k_1}\\ \mathtt{j}\,\mathtt{k}\,h\in\Sigma_{A,*}}} \mu([\mathtt{j}\,\mathtt{k}\,G^{(\ell)}_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C}])\\
&= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\ell:n_\ell\in[a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C),b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \sum_{h\in\{1,\dots,K\}} \mu([\mathtt{j}]\cap\sigma^{-|\,\mathtt{j}\,|-k-k_1}[G^{(\ell)}_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C}]).
\end{aligned}$$

Next, by Theorem 3.3 we have

$$\begin{aligned}
&\frac{\mu(D_k(\mathtt{i}))}{1-2^{-k}}\\
&\geq \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{\ell:n_\ell\in[a(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C),b(|\,\mathtt{j}\,|+k+k_1,2^{-k}/C)]} \sum_{h\in\{1,\dots,K\}} \mu([\mathtt{j}])\mu([G^{(\ell)}_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C}])(1-D\gamma^{k+k_1})\\
&= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{h\in\{1,\dots,K\}} \mu([\mathtt{j}])\mu(G_{h,|\,\mathtt{j}\,|+k+k_1,2^{-k}/C})(1-D\gamma^{k+k_1}).
\end{aligned}$$

Now using (5.6) in the above we see that

$$\begin{aligned}
\mu(D_k(\mathtt{i})) &\geq \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \sum_{h\in\{1,\dots,K\}} \mu([\mathtt{j}])\mu([h])(1-D\gamma^{k+k_1})(1-2^{-k})(1-2^{-k}/C)\\
&= \sum_{\mathtt{j}\in W_{k-1}(\mathtt{i})} \mu([\mathtt{j}])(1-D\gamma^{k+k_1})(1-2^{-k})(1-2^{-k}/C)\\
&= \mu(D_{k-1}(\mathtt{i}))(1-D\gamma^{k+k_1})(1-2^{-k})(1-2^{-k}/C).
\end{aligned}$$

The proof of the second claim is similar. □

It now follows from (5.12), (5.13), and Lemma 5.3 that

$$\begin{aligned}
\mu([i]\cap S_{\psi,(n_k)}) &\geq \mu\left(\bigcap_{k=1}^{\infty} D_k(\mathtt{i})\right)\\
&= \lim_{k\to\infty}\mu(D_k(\mathtt{i}))\\
&\geq \mu([\mathtt{i}])\prod_{k=k_1}^{\infty}(1-D\gamma^k)\prod_{k=1}^{\infty}(1-2^{-k})\prod_{k=1}^{\infty}(1-2^{-k}/C).
\end{aligned}$$

Since the series $\sum_{k=k_1}^{\infty} D\gamma^k$ and $\sum_{k=1}^{\infty} 2^{-k}$ are convergent we can take

$$c=\prod_{k=k_1}^{\infty}(1-D\gamma^k)\prod_{k=1}^{\infty}(1-2^{-k})\prod_{k=1}^{\infty}(1-2^{-k}/C)$$

and (5.1) holds (recall that $0<\gamma<1$ and that $k_1\in\mathbb{N}$ was chosen such that $D\gamma^{k_1}<1$). This completes our proof.

6 Proof of Theorem 2.6

We begin our proof of Theorem 2.6 by showing that for every function $g\colon \mathbb{N} \mapsto [1,\infty)$ satisfying $\limsup_{n\to\infty} \frac{g(n)}{\sqrt{\log\log n}} \leq 1$ and $\lim_{n\to\infty} g(n) = \infty$, there exists a sequence $(n_k)_{k=1}^{\infty}$ such that for a typical $\mathtt{i} \in \Sigma$ the quantity $\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k$ eventually satisfies a useful upper bound formulated in terms of our function g.

Lemma 6.1. *Let μ be a non-uniform Bernoulli measure and let $g\colon \mathbb{N} \mapsto [1,\infty)$ be a function such that $\limsup_{n\to\infty} \frac{g(n)}{\sqrt{\log\log n}} \leq 1$ and $\lim_{n\to\infty} g(n) = \infty$. Then there exists a sequence of positive integers $(n_k)_{k=1}^{\infty}$ such that*

$$\mu\left(\left\{\mathtt{i} \in \Sigma : \limsup_{k\to\infty} \frac{\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k} g(n_k)} \leq \frac{1}{2}\right\}\right) = 1.$$

Proof. We define the sequence (n_k) via the equation $n_k = \max\{n \geq 1 : g(n)^2 < k\} + 1$. Let the sequence $(\zeta_n)_{n=1}^{\infty}$ be as in Theorem 3.4 with $(a_n)_{n=1}^{\infty} = (\frac{\sqrt{2}g(n)}{2})_{n=1}^{\infty}$. Now applying Theorem 3.4 and using the fact that $\lim_{n\to\infty} \zeta_n = 0$ we have

$$\begin{aligned}\sum_{k=1}^{\infty} \mu\left(\left\{\mathtt{i} \in \Sigma : \frac{\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k} g(n_k)} \geq \frac{1}{2}\right\}\right) &= \sum_{k=1}^{\infty} e^{\frac{-g(n_k)^2(1+\zeta_{n_k})}{4}} \\ &\ll \sum_{k=1}^{\infty} e^{-k/5} \\ &< \infty.\end{aligned}$$

Our claim now follows by the Borel–Cantelli Lemma (Lemma 3.5). □

Equipped with Lemma 6.1 we are now in a position to prove Theorem 2.6. Our proof relies upon first of all constructing a function ψ using the sequence (n_k) provided by this lemma. We then apply Theorem 3.2, Lemma 3.5, and a uniformisation argument in an analogous way to the proof of Theorem 2.1.

Proof of Theorem 2.6. Let $\mathcal{G} = \{n_k : k \geq 1\} \subseteq \mathbb{N}$, where $(n_k)_{k=1}^{\infty}$ is the sequence defined in Lemma 6.1. Let $(a_n)_{n=1}^{\infty}$ be the sequence of natural numbers defined as follows

$$a_n := \begin{cases} \left\lceil e^{h_\mu n - \sqrt{2\rho_\mu n} g(n)} \right\rceil & \text{if } n \in \mathcal{G}, \\ \left\lfloor n^{-2} e^{h_\mu n - \frac{5}{4}\sqrt{2\rho_\mu n \log\log n}} \right\rfloor & \text{if } n \notin \mathcal{G}. \end{cases}$$

Then let $\psi(n) := \max\{k \geq 1 : \sum_{m=1}^{k-1} a_m \leq n\}$. Notice that $\#\psi^{-1}(n) = a_n$ for all $n \in \mathbb{N}$.

Thus,

$$\begin{aligned}\sum_{n=1}^{\infty} e^{-h_\mu \psi(n) + \sqrt{2\rho_\mu \psi(n)} g(\psi(n))} &= \sum_{n=1}^{\infty} \#\psi^{-1}(n) e^{-h_\mu n + \sqrt{2\rho_\mu n} g(n)} \\ &\geq \sum_{k=1}^{\infty} \#\psi^{-1}(n_k) e^{-h_\mu n_k + \sqrt{2\rho_\mu n_k} g(n_k)}\end{aligned}$$

$$\geq \sum_{k=1}^{\infty} 1 = \infty.$$

For each $N \in \mathbb{N}$ we let

$$E_N^{(1)} = \left\{ \mathtt{i} \in \Sigma : \frac{\log \mu([\mathtt{i}\,|_1^{n_k}]) + h_\mu n_k}{\sqrt{2\rho_\mu n_k} g(n_k)} < \frac{3}{4} \text{ for every } n_k \geq N \right\}$$

and

$$E_N^{(2)} = \left\{ \mathtt{i} \in \Sigma : \frac{\log \mu([\mathtt{i}\,|_1^{n}]) + h_\mu n}{\sqrt{2\rho_\mu n \log\log n}} < \frac{5}{4} \text{ for every } n \geq N \right\}.$$

By Theorem 3.2 and Lemma 6.1 we have $\mu(\bigcup_{N=1}^{\infty} E_N^{(1)} \cap E_N^{(2)}) = 1$.

Let us also define

$$Q(n) = \begin{cases} e^{-h_\mu n + \frac{3}{4}\sqrt{2\rho_\mu n} g(n)} & \text{if } n \in \mathcal{G}, \\ e^{-h_\mu n + \frac{5}{4}\sqrt{2\rho_\mu n \log\log n}} & \text{if } n \notin \mathcal{G}. \end{cases}$$

It follows from the above that if $\mathtt{j} \in \cup_{n=N}^{\infty} \Sigma_n$ satisfies $[\mathtt{j}] \cap E_N^{(1)} \cap E_N^{(2)} \neq \emptyset$ then we have

$$\mu([\mathtt{j}]) \leq Q(|\,\mathtt{j}\,|). \tag{6.1}$$

We will now show that $\mu\left(E_N^{(1)} \cap E_N^{(2)} \cap R_\psi\right) = 0$ for any $N \geq 1$. Since $\mu(\bigcup_{N=1}^{\infty} E_N^{(1)} \cap E_N^{(2)}) = 1$, this will complete our proof.

Let us now fix $N \geq 1$. By the first Borel–Cantelli Lemma (Lemma 3.5), to show that $\mu\left(E_N^{(1)} \cap E_N^{(2)} \cap R_\psi\right) = 0$, it suffices to show that

$$\sum_{n=1}^{\infty} \mu\left(\left\{ \mathtt{i} \in E_N^{(1)} \cap E_N^{(2)} : \sigma^n \mathtt{i} \in [\mathtt{i}\,|_1^{\psi(n)}] \right\}\right) < \infty. \tag{6.2}$$

Let us start by defining $M_N \in \mathbb{N}$ to be sufficiently large such that for all $n \geq M_N$ we have $\psi(n) \geq N$. We now show that (6.2) holds. We first note that it follows from (2.6) and the fact that μ is a Bernoulli measure, so (2.4) holds, that

$$\begin{aligned} \sum_{n=1}^{\infty} \mu\left(\left\{ \mathtt{i} \in E_N^{(1)} \cap E_N^{(2)} : \sigma^n \mathtt{i} \in [\mathtt{i}\,|_1^{\psi(n)}] \right\}\right) &\leq M_N + \sum_{n=M_N}^{\infty} \sum_{\substack{\mathtt{j} \in \Sigma_{\psi(n)} \\ \mathtt{i} \in \Sigma_{n-\psi(n)} \\ [\mathtt{j}] \cap E_N^{(1)} \cap E_N^{(2)} \neq \emptyset}} \mu([\mathtt{j}\,\mathtt{i}\,\mathtt{j}]) \\ &= M_N + \sum_{n=M_N}^{\infty} \sum_{\substack{\mathtt{j} \in \Sigma_{\psi(n)} \\ \mathtt{i} \in \Sigma_{n-\psi(n)} \\ [\mathtt{j}] \cap E_N^{(1)} \cap E_N^{(2)} \neq \emptyset}} \mu([\mathtt{j}])\mu([\mathtt{i}])\mu([\mathtt{j}]) \\ &\leq M_N + \sum_{n=M_N}^{\infty} \sum_{\substack{\mathtt{j} \in \Sigma_{\psi(n)} \\ [\mathtt{j}] \cap E_N^{(1)} \cap E_N^{(2)} \neq \emptyset}} \mu([\mathtt{j}])^2. \end{aligned}$$

Next, from (6.1) and the fact that $\psi(n) \geq N$ for all $n \geq M_N$, we see that

$$\sum_{n=1}^{\infty} \mu\left(\left\{\mathtt{i} \in E_N^{(1)} \cap E_N^{(2)} : \sigma^n \mathtt{i} \in [\mathtt{i}\,|_1^{\psi(n)}]\right\}\right) \leq M_N + C \sum_{n=M_N}^{\infty} \sum_{\substack{\mathtt{j} \in \Sigma_{\psi(n)} \\ [\mathtt{j}] \cap E_N^{(1)} \cap E_N^{(2)} \neq \emptyset}} Q(\psi(n))\mu([\mathtt{j}]).$$

Finally, recalling the definitions of $Q(n)$, a_n, and the fact that $\#\psi^{-1}(n) = a_n$ for all $n \in \mathbb{N}$, we observe that

$$\begin{aligned}
&\sum_{n=M_N}^{\infty} \sum_{\substack{\mathtt{j} \in \Sigma_{\psi(n)} \\ [\mathtt{j}] \cap E_N^{(1)} \cap E_N^{(2)} \neq \emptyset}} Q(\psi(n))\mu([\mathtt{j}]) \\
&\leq \sum_{n=1}^{\infty} \#\psi^{-1}(n) Q(n) \\
&= \sum_{n \in \mathcal{G}} \left\lceil e^{h_\mu n - \sqrt{2\rho_\mu n} g(n)} \right\rceil e^{-h_\mu n + \frac{3}{4}\sqrt{2\rho_\mu n} g(n)} + \sum_{n \notin \mathcal{G}} \left\lfloor n^{-2} e^{h_\mu n - \frac{5}{4}\sqrt{2\rho_\mu n \log^{(2)} n}} \right\rfloor e^{-h_\mu n + \frac{5}{4}\sqrt{2\rho_\mu n \log^{(2)} n}} \\
&\ll \sum_{n=1}^{\infty} e^{-\frac{1}{4}\sqrt{2\rho_\mu n} g(n)} + \sum_{n=1}^{\infty} n^{-2} < \infty.
\end{aligned}$$

Thus it follows that (6.2) holds and our proof is complete. □

7 Applications to dynamics on self-similar sets

In this section we apply our results to the study of dynamics on self-similar sets. Before that it is necessary to define some preliminary notions. We call a map $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ a contracting similarity if there exists $r \in (0,1)$ such that

$$\|\varphi(x) - \varphi(y)\| = r\|x - y\| \quad \text{for all } x, y \in \mathbb{R}^d.$$

In this section, we define an *iterated function system, or IFS* for short, to be a finite set of contracting similarities $\Phi := \{\varphi_i\}_{i=1}^K$. When each similarity in an IFS has the same contraction ratio the IFS is said to be *homogeneous*. A well known result due to Hutchinson [19] states that for any IFS there exists a unique, non-empty, compact set $X \subset \mathbb{R}^d$ satisfying

$$\bigcup_{i=1}^{K} \varphi_i(X) = X.$$

The set X is called the *self-similar set* or *attractor* of Φ. Self-similar sets are important and well studied objects in the field of Fractal Geometry. For more on these sets we refer the reader to Falconer's book [14]. Self-similar sets can be viewed as the image of $\Sigma = \{1, \ldots, K\}^{\mathbb{N}}$ under an appropriate map. Let $\pi : \Sigma \to X$ be given by

$$\pi(\mathtt{i}) = \lim_{n \to \infty} \left(\varphi_{i_1} \circ \cdots \circ \varphi_{i_n}\right)(0).$$

The map π is continuous and surjective. Importantly π allows us to take measures defined on Σ, most notably Gibbs measures, and to project them forward on to X. The measures defined on X as pushforwards under π are an important and well studied class. The most well studied

measures amongst this class are the self-similar measures which are the pushforwards of Bernoulli measures.

If Φ satisfies the additional assumption that

$$\varphi_i(X) \cap \varphi_j(X) = \emptyset \quad \text{for all } i \neq j$$

then Φ is said to satisfy the *strong separation condition*. This condition is equivalent to the map π being a bijection. If an IFS satisfies the strong separation condition we can define a map $T : X \to X$ according to the rule

$$T(x) = \varphi_i^{-1}(x) \quad \text{if} \quad x \in \varphi_i(X).$$

Because $X = \bigcup_{i=1}^{K} \varphi_i(X)$ and this union is disjoint, the map T is well defined.

Given a homogeneous IFS with common contraction ratio r which satisfies the strong separation condition and a function $\psi : \mathbb{N} \to [0, \infty)$, we can define a recurrence set as follows:

$$\tilde{R}_\psi := \left\{ x \in X : \|T^n(x) - x\| \leq r^{\psi(n)} \text{ for infinitely many } n \in \mathbb{N} \right\}.$$

Here $\|\cdot\|$ denotes the Euclidean norm. This family of recurrence sets was studied previously in [2] and [9]. Combining these papers with the mass transference principle of Beresnevich and Velani [5], it is possible to obtain a detailed description of the metric properties of $\tilde{R}_\psi$ in terms of Hausdorff measure. As we will see, the results of the current paper allow us to prove new statements on the $\pi^*\mu$ measure of $\tilde{R}_\psi$. Here $\pi^*\mu$ denotes the pushforward of some Gibbs measure μ under the map π; that is, $\pi^*\mu = \mu \circ \pi^{-1}$. The key proposition that allows us to translate our previous results for R_ψ (recall the definition in (2.5)) into statements for $\tilde{R}_\psi$ is the following:

Proposition 7.1. *Let Φ be a homogeneous IFS satisfying the strong separation condition with associated self-similar set X and let $\psi : \mathbb{N} \to [0, \infty)$. There exists $N \in \mathbb{N}$ depending only upon Φ such that we have the following inclusions:*

$$\pi(R_{\lfloor \psi \rfloor + N}) \subseteq \tilde{R}_\psi \subseteq \pi(R_{\lfloor \psi \rfloor - N}).$$

Proof. This follows from the observation that $\pi(\sigma\, \mathtt{i}) = T(\pi\, \mathtt{i})$ together with the following fact from Fractal Geometry.

Fact: Let $x, y \in X$ and $\mathtt{i}, \mathtt{j} \in \Sigma$ be such that $\pi(\mathtt{i}) = x$ and $\pi(\mathtt{j}) = y$. Then there exists $N \in \mathbb{N}$ such that:

$$\text{if } \|x - y\| \leq r^{\psi(n)} \text{ then } |\, \mathtt{i} \wedge \mathtt{j} \,| \geq \lfloor \psi(n) \rfloor - N \tag{7.1}$$

and

$$\text{if } |\, \mathtt{i} \wedge \mathtt{j} \,| \geq \lfloor \psi(n) \rfloor + N \text{ then } \|x - y\| \leq r^{\psi(n)}. \tag{7.2}$$

We include a proof of this fact for completion.

Let $N \in \mathbb{N}$ be sufficiently large that

$$\inf_{i,j:i \neq j} d(\varphi_i(X), \varphi_j(X)) \geq r^N \text{ and } r^{N-1} \cdot Diam(X) < 1.$$

Here d is the Euclidean metric. Note that $\inf_{i,j:i\neq j} d(\varphi_i(X), \varphi_j(X)) > 0$ because of the strong separation condition. Therefore N is well defined. We now prove that this N satisfies the desired

properties.

Let $x, y \in X$ be such that $\|x-y\| \leq r^{\psi(n)}$. By considering inverses and the maximal common prefix $\mathtt{i} \wedge \mathtt{j}$, we see that

$$r^{-|\,\mathtt{i}\wedge\mathtt{j}\,|}\|x-y\| = \|(\varphi_{i_1} \circ \cdots \circ \varphi_{i_{|\,\mathtt{i}\wedge\mathtt{j}\,|}})^{-1}(x) - (\varphi_{j_1} \circ \cdots \circ \varphi_{j_{|\,\mathtt{i}\wedge\mathtt{j}\,|}})^{-1}(y)\| \geq r^N.$$

Therefore $\|x-y\| \geq r^{N+|\,\mathtt{i}\wedge\mathtt{j}\,|}$ and we have $\psi(n) \leq N + |\,\mathtt{i}\wedge\mathtt{j}\,|$. Taking integer parts we see that $|\,\mathtt{i}\wedge\mathtt{j}\,| \geq \lfloor \psi(n) \rfloor - N$ and (7.1) holds.

Now suppose $|\,\mathtt{i}\wedge\mathtt{j}\,| \geq \lfloor \psi(n) \rfloor + N$. Then

$$\begin{aligned}
\|x-y\| &\leq Diam((\varphi_{i_1} \circ \cdots \circ \varphi_{i_{|\,\mathtt{i}\wedge\mathtt{j}\,|}})(X)) \\
&\leq r^{|\,\mathtt{i}\wedge\mathtt{j}\,|} \cdot Diam(X) \\
&\leq r^{\lfloor \psi(n) \rfloor + N} \cdot Diam(X) \\
&\leq r^{\psi(n)-1+N} \cdot Diam(X) \\
&\leq r^{\psi(n)}.
\end{aligned}$$

Therefore (7.2) holds and our proof is complete. □

Equipped with Proposition 7.1, it is possible to translate Theorems 2.1, 2.2, 2.4, and 2.6 into the setting of dynamics on self-similar sets and pushforwards of Gibbs measures. The key point that allows us to establish these analogues is that the parameter N, whose existence is asserted by Proposition 7.1, only depends upon the underlying IFS. As such, if $\psi : \mathbb{N} \to [0,\infty)$ satisfies some appropriate hypothesis analogous to that given in one of the theorems listed above, then the functions $\lfloor \psi(n) \rfloor - N$ and $\lfloor \psi(n) \rfloor + N$ will also satisfy the hypothesis formulated in the original theorem. Using this observation together with the inclusions given in Proposition 7.1, we may prove appropriate analogues of Theorems 2.1, 2.2, 2.4, and 2.6 in this setting. For the sake of brevity we do not give the statement of each of these analogues here. We instead content ourselves with the following analogue of Theorem 2.4.

Theorem 7.2. *Let Φ be a homogeneous IFS satisfying the strong separation condition and let μ be the Gibbs measure of a Hölder continuous potential that is not cohomologous to a constant. For $\varepsilon > 0$ let $\psi_\varepsilon^+ : \mathbb{N} \to [0,\infty)$ and $\psi_\varepsilon^- : \mathbb{N} \to [0,\infty)$ be given by*

$$\psi_\varepsilon^+(n) = \frac{\log n}{h_\mu} + \frac{(1+\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu \log n \log^{(3)} n}$$

and

$$\psi_\varepsilon^-(n) = \frac{\log n}{h_\mu} + \frac{(1-\varepsilon)}{h_\mu^{3/2}}\sqrt{2\rho_\mu \log n \log^{(3)} n}.$$

Then for any $\varepsilon > 0$ we have $\pi^*\mu(\tilde{R}_{\psi_\varepsilon^+}) = 0$ and $\pi^*\mu(\tilde{R}_{\psi_\varepsilon^-}) = 1$.

Theorem 7.2 demonstrates that the critical threshold observed for Gibbs measures and shifts of finite type persists in the setting of dynamics on self-similar sets. Theorem 7.2 also yields some interesting metric properties for the sets $\tilde{R}_{\psi_\varepsilon^-}$ that do not follow from [2] or [9]. It is well known[2] that for any Gibbs measure μ for which the defining potential is not cohomologous to a constant

[2]This fact follows from [27, Proposition 3.6] and the well known fact that the uniform measure is the unique measure of maximal entropy for the full shift (see [8]).

we have $h_\mu < \log K$, and so by a covering argument, $\mathcal{H}^{\dim_{\mathrm{H}} X}(\tilde{R}_{\psi_\varepsilon^-}) = 0$ for any $\varepsilon \in (0,1)$. Here $\dim_{\mathrm{H}} X$ is the Hausdorff dimension of the self-similar set X. Moreover, using the mass transference principle of Beresnevich and Velani [5], it is possible to show that $\dim_{\mathrm{H}}(\tilde{R}_{\psi_\varepsilon^-}) = \dim_{\mathrm{H}} X$ for any $\varepsilon > 0$. Therefore, despite being null in terms of the $\dim_{\mathrm{H}}(\tilde{R}_{\psi_\varepsilon^-})$-dimensional Hausdorff measure, Theorem 7.2 tells us that $\tilde{R}_{\psi_\varepsilon^-}$ is large in terms of μ for any $\varepsilon > 0$.

Acknowledgements. The authors would like to thank the referees for their valuable feedback.

References

[1] J. Athreya. Logarithm laws and shrinking target properties. *Proc. Indian Acad. Sci. Math. Sci.* **119** (2009), no. 4, 541–557.

[2] S. Baker and M. Farmer. Quantitative recurrence properties for self-conformal sets. *Proc. Amer. Math. Soc.* **149**(3) (2021), 1127–1138.

[3] L. Barreira and B. Saussol. Hausdorff dimension of measures via Poincaré recurrence. *Comm. Math. Phys.* **219**(2) (2001), 443–463.

[4] V. Beresnevich, D. Dickinson and S. Velani. Measure theoretic laws for lim sup sets. *Mem. Amer. Math. Soc.* **179** (2006), no. 846, x+91 pp.

[5] V. Beresnevich and S. Velani. A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures. *Ann. of Math.* (2) **164** (2006), no. 3, 971–992.

[6] P. Billingsley. *Probability and measure.* Third edition. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1995.

[7] M. Boshernitzan. Quantitative recurrence results. *Invent. Math.* **113** (3) (1993), 617–631.

[8] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Second revised edition. Lecture Notes in Mathematics, 470. Springer-Verlag, Berlin, 2008.

[9] Y. Chang, M. Wu and W. Wu. Quantitative recurrence properties and homogeneous self-similar sets. *Proc. Amer. Math. Soc.* **147**(4) (2019), 1453–1465.

[10] J.-R. Chazottes and E. Ugalde. Entropy estimation and fluctuations of hitting and recurrence times for Gibbsian sources. *Discrete Contin. Dyn. Syst. Ser. B* **5**(3) (2005), 565–586.

[11] N. Chernov and D. Kleinbock. Dynamical Borel-Cantelli lemmas for Gibbs measures. *Israel J. Math.* **122** (2001), 1–27.

[12] M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory & Dynam. Systems* **4**(4) (1984), 541–552.

[13] M. Einsiedler, T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, 259. Springer-Verlag London, Ltd., London, 2011. xviii+481 pp. ISBN: 978-0-85729-020-5.

[14] K. Falconer. *Fractal geometry. Mathematical foundations and applications*, Third edition. John Wiley & Sons, Ltd., Chichester, 2014. xxx+368 pp. ISBN: 978-1-119-94239-9.

[15] G. Harman, *Metric number theory.* London Mathematical Society Monographs. New Series, 18. The Clarendon Press, Oxford University Press, New York, 1998. xviii+297 pp.

[16] R. Hill and S. Velani, The ergodic theory of shrinking targets. *Invent. Math.* **119** (1995), no. 1, 175–198.

[17] R. Hill and S. Velani, The shrinking target problem for matrix transformations of tori. *J. London Math. Soc.* (2) **60** (1999), no. 2, 381–398.

[18] M. Hussain, B. Li, D. Simmons, and B. Wang. Dynamical Borel–Cantelli lemma for recurrence theory. *Ergodic Theory & Dynam. Systems*, **42** (2022), no. 6, 1994–2008.

[19] J. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747.

[20] T. Jordan and M. Pollicott. Multifractal analysis and the variance of Gibbs measures. *J. Lond. Math. Soc.,* **76**(1) (2007), 57–72.

[21] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Cambridge University Press, Cambridge, 1995. xviii+802 pp. ISBN: 0-521-34187-6

[22] D. Kelmer and S. Yu. Shrinking target problems for flows on homogeneous spaces. *Trans. Amer. Math. Soc.* **372** (2019), no. 9, 6283–6314.

[23] M. Kirsebom, P. Kunde, and T. Persson. On shrinking targets and self-returning points. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* 2023: Vol. XXIV, issue 3

[24] D. Kleinbock and J. Zheng. Dynamical Borel-Cantelli Lemma for Recurrence under Lipschitz Twists. *Nonlinearity*, Vol 36, 2.

[25] B. Li, L. Liao, S. Velani, E. Zorin, B. Wang. The Shrinking Target Problem for Matrix Transformations of Tori: revisiting the standard problem. *Adv. Math.* **421** (2023), 108994.

[26] P. Mattila. *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability.* Cambridge Studies in Advanced Mathematics, 44. Cambridge University Press, Cambridge, 1995.

[27] W. Parry and M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics.* Astérisque **187-188**, (1990) 1–268.

[28] K. Petersen, *Ergodic theory.* Cambridge Studies in Advanced Mathematics, 2. Cambridge University Press, Cambridge, 1983. xii+329 pp. ISBN: 0-521-23632-0

[29] W. Philipp. Some metrical theorems in number theory. *Pacific J. Math.* **20** (1967), 109–127.

[30] S. Seuret and B. Wang. Quantitative recurrence properties in conformal iterated function systems. *Adv. Math.* **280** (2015), 472–505.

[31] B. Tan and B. Wang. Quantitative recurrence properties of beta dynamical systems. *Adv. Math.* **228** (2011) 2071–2097.

[32] P. Walters. *An introduction to ergodic theory.* Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.