COMPLEXITY OF DYNAMICAL SYSTEMS ARISING FROM RANDOM SUBSTITUTIONS IN ONE DIMENSION

by

ANDREW MITCHELL

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
The University of Birmingham
June 2023

ACKNOWLEDGEMENTS

This journey would not have been possible without the support and encouragement of my supervisors, Tony Samuel and Dan Rust, whose guidance has been invaluable in my development as an independent researcher. Further, I thank Rob Leek for his support throughout my PhD, especially in the early stages. I also thank my further collaborators, Philipp Gohlke and Alex Rutar, for sharing many illuminating ideas with me and for the time and effort they devoted to writing our joint papers.

I was fortunate enough to spend a one-month research visit at Bielefeld University, Germany during the summer of 2022. I am grateful to Michael Baake for this invitation, and to the SFB 1283 for funding this visit. Further, I thank Michael, Philipp and the other members of the Aperiodic Order research group for making me feel so welcome during my visit. I will remember fondly our daily lunchtime discussions and our trips to Siegfriedplatz.

My PhD journey would not have been nearly as enjoyable without the support of my friends and colleagues in the School of Mathematics at Birmingham. This margin is too small to thank all of you individually, but you know who you are.

I thank the EPSRC and University of Birmingham for their financial support.

Last but not least, I thank my family and Daniel, for their constant love and support, without which this journey would not have been possible.

Abstract

This thesis is based on three papers the author wrote while a PhD student [37, 54, 55], which concern different notions of complexity for dynamical systems arising from random substitutions.

Before presenting our main results, we first provide an introduction to random substitutions. In Chapter 2, we give the main definitions that we work with throughout, and prove several basic properties of random substitutions and their associated subshifts. We define the *frequency measure* corresponding to a random substitution, and prove a key result concerning such measures which will be of fundamental importance in our work.

Chapter 3 is based on the solo-author paper [54] and concerns word complexity and topological entropy of random substitution subshifts. In contrast to previous work, we do not assume that the underlying random substitution is compatible. In our main results, we show that the subshift of a primitive random substitution has zero topological entropy if and only if it can be obtained as the subshift of a deterministic substitution – answering in the affirmative an open question of Rust and Spindeler [70] – and provide a systematic approach to calculating the topological entropy for subshifts of constant length random substitutions. We also consider word complexity for constant length random substitutions and show that, without primitivity, the complexity function can exhibit features not possible in the deterministic or primitive random settings.

Chapters 4 and 5 are based on the paper [37], which is joint work with P. Gohlke, D. Rust and T. Samuel. These chapters focus on measure theoretic entropy and its relationship to topological entropy. In Chapter 4, we introduce a new measure of complexity for primitive random substitutions called *measure theoretic inflation word entropy* and show that this coincides with the measure theoretic entropy of the subshift with respect to the corresponding frequency measure. This allows the measure theoretic entropy to be explicitly calculated in many cases. In Chapter 5, we provide sufficient conditions under which a random substitution subshift supports a frequency measure of maximal entropy and, under more restrictive conditions, show that this measure is the unique measure of maximal entropy. Notably, we show that random substitutions can give rise to intrinsically ergodic subshifts that do not satisfy Bowen's specification property [10] or the weaker specification property of Climenhaga and Thompson [13], thus providing an

interesting new class of intrinsically ergodic subshifts. We conclude this chapter by showing that the random period doubling substitution is intrinsically ergodic.

Finally, Chapter 6 is based on the paper [55], which is joint work with A. Rutar. Here, we consider multifractal properties of frequency measures. Specifically, we study the multifractal spectrum and L^q-spectrum of frequency measures corresponding to primitive and compatible random substitutions. We introduce a new notion called the *inflation word* L^q*-spectrum* of a random substitution and show that this coincides with the L^q-spectrum of the corresponding frequency measure for all $q \geq 0$. Under an additional assumption (*recognisability*) we show that the two notions coincide for all $q \in \mathbb{R}$. Further, under these assumptions, we show that the multifractal formalism holds. The techniques we develop allow the L^q-spectrum and multifractal spectrum to be obtained for many frequency measures.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Background and context

The discovery of quasicrystals – naturally occurring structures which exhibit long-range order but lack translation symmetry – came as a surprise to physicists, chemists and materials scientists, and was honoured with the 1999 Wolf Prize in Physics and the 2011 Nobel Prize in Chemistry. In 1982, Shechtman [71] discovered an Aluminium–Manganese alloy that exhibits a diffraction pattern with five-fold rotational symmetry (see Figure 1.1, right). This not possible for crystals, which always have two-, three-, four- or six-fold rotation symmetry, by the crystallographic restriction theorem [8]. Since then, many examples of naturally occurring quasicrystals have been observed, including examples with other symmetries not possible for crystals, such as eight-, 12- and 20-fold rotational symmetry. This has stimulated a wealth of research in the field of *aperiodic order*: the mathematical study of quasicrystals.

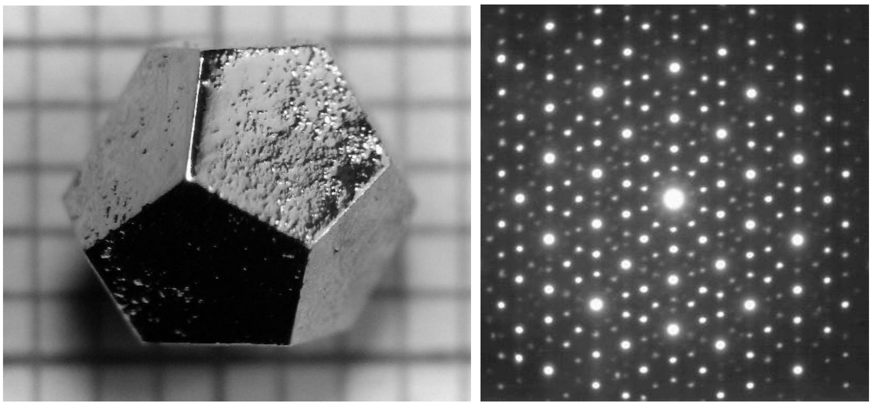


Figure 1.1: A Ho-Mg-Zn quasicrystal (left – image credit [32]) and the diffraction spectrum of an Al-Mn alloy observed by Shechtman (right – image credit [71]).

While a complete classification of the possible configurations of crystals is known, a classification of aperiodically ordered structures remains distant [5, 6].

> *It is fair to say that a classification of a hierarchy of (aperiodic) order has not only not been achieved yet, but is actually not even in sight.* — Baake and Grimm [5, Page 9].

One of the primary objectives in aperiodic order research is to establish such a classification.

Central to understanding properties of quasicrystals from a mathematical viewpoint is the careful analysis of dynamical systems associated with aperiodic sequences. One of the primary methods for generating such sequences is via *substitutions*. A deterministic substitution is a rule that replaces each symbol from a finite set with a concatenation of symbols from the same set. For example, the *Fibonacci substitution* $\theta\colon a \mapsto ab$, $b \mapsto a$ is the rule that replaces every letter a in a given string with the word ab and every b with an a. To a given substitution, a dynamical system (subshift) can be associated in a natural way. Subshifts of deterministic substitutions are well-studied dynamical systems that possess a high degree of long-range order, have low complexity, and provide theoretical models for physical quasicrystals. Many of the key topological and dynamical properties of substitution subshifts are well understood. For instance, subshifts of deterministic substitutions always have zero topological entropy and, under a mild assumption (primitivity), support a unique ergodic measure [67].

Random substitutions are a generalisation of deterministic substitutions where the substituted image of a letter is chosen from a fixed finite set according to a probability distribution. Similarly to deterministic substitutions, a subshift can be associated to a given random substitution in a canonical manner. However, in contrast to their deterministic counterparts, subshifts of random substitutions typically have positive topological entropy [35] and often support uncountably many ergodic measures [39]. Nonetheless, random substitution subshifts maintain many of the features of long-range order witnessed for subshifts of deterministic substitutions. While they have positive entropy, indicating disorder, random substitutions often admit long-range correlations presenting as a non-trivial pure-point component in their diffraction spectrum [7, 34, 56]. This competition between order and disorder, and between long- and short-range correlations suggests an intricate combinatorial structure that warrants careful study. Further, physical quasicrystals are unlikely to possess perfect order, but instead exhibit local defects. Thus, a good theoretical model should exhibit features of both long-range order but also local disorder, features often possessed by subshifts of random substitutions.

The non-trivial topological entropy of random substitution subshifts provides a new invariant in their study not available for their deterministic counterparts. Further, to each random substitution an ergodic measure which captures the underlying probability distribution can be associated in a natural way [39, 70], the properties of which can provide a means of classifying random substitutions. In this thesis, we develop techniques to quantify topological entropy and three measures of complexity for measures arising from random substitutions; namely, measure theoretic entropy, the multifractal spectrum and the L^q-spectrum. With this work, we make the first steps towards providing a hierarchical classification of random substitutions in terms of their complexity.

1.2 Symbolic dynamics

Throughout this thesis, we will be concerned with symbolic dynamical systems. The following symbolic notation will be standard throughout, and is generally in line with the notation used in [49, 75].

An *alphabet* $\mathcal{A} = \{a_1, \ldots, a_d\}$ is a finite collection of symbols, which we call *letters*. We call a finite concatenation of letters a *word*, and let $\mathcal{A}^+$ denote the set of all non-empty finite words with letters from $\mathcal{A}$. For a given word $u = u_1 \cdots u_n$, where $n \in \mathbb{N}$ and $u_j \in \mathcal{A}$ for all $j \in \{1, \ldots, n\}$,we write $|u| = n$ for the length of u and, for each $a_i \in \mathcal{A}$, let $|u|_{a_i}$ denote the number of occurrences of a_i in u. To avoid conflicting notation, we denote the cardinality of a given set B by $\#B$ throughout. Given $u \in \mathcal{A}^+$ and $i, j \in \{1, \ldots, |u|\}$, we write $u_{[i,j]} = u_i \cdots u_j$. The *abelianisation* of a word $u \in \mathcal{A}^+$ is the vector $\Phi(u) \in \mathbb{Z}^{\#\mathcal{A}}$ defined by $\Phi(u)_i = |u|_{a_i}$ for all $i \in \{1, \ldots, d\}$. For two words $u, v \in \mathcal{A}^+$, with $|v| \leq |u|$, we write $|u|_v$ for the number of distinct occurrences v as a subword of u, namely, $|u|_v = \#\{j \in \{1, \ldots, |u| - |v| + 1\} \colon u_{[j,j+|v|-1]} = v\}$.

For a given alphabet $\mathcal{A}$, we let $\mathcal{A}^{\mathbb{Z}}$ denote the set of all bi-infinite sequences of letters in $\mathcal{A}$, and endow $\mathcal{A}^{\mathbb{Z}}$ with the discrete product topology. We let $S \colon \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ denote the (left) shift map, defined by $S(x)_j = x_{j+1}$ for all $j \in \mathbb{Z}$. Note that S is continuous with respect to the topology on $\mathcal{A}^{\mathbb{Z}}$. Since S is invertible on $\mathcal{A}^{\mathbb{Z}}$ and has continuous inverse, S defines a homeomorphism. We call the dynamical system $(\mathcal{A}^{\mathbb{Z}}, S)$ the *full shift* on the alphabet $\mathcal{A}$. A *subshift* is a closed and S-invariant subspace X of the full shift $\mathcal{A}^{\mathbb{Z}}$, that is, a subspace for which $S^{-1}X = X$. We endow X with the subspace topology inherited from $\mathcal{A}^{\mathbb{Z}}$. We highlight that we work with *two-sided* shifts throughout. It is also possible to define the one-sided shift $\mathcal{A}^{\mathbb{N}}$ of right-infinite sequences of letters in $\mathcal{A}$; however, we do not concern ourselves with one-sided shifts in this thesis.

If $i, j \in \mathbb{Z}$ with $i \leq j$, and $x = \cdots x_{-1}x_0x_1\cdots$ is an element of a subshift X, then we write $x_{[i,j]} = x_i x_{i+1} \cdots x_j$. We let $\mathcal{L}(X)$ denote the *language* of the subshift X: namely, the set of all finite words $u \in \mathcal{A}^{\mathbb{Z}}$ such that there exist $x \in X$ and $j \in \mathbb{Z}$ for which $x_{[j,j+|u|-1]} = u$. Observe that if X and Y are subshifts, then $X = Y$ if and only if $\mathcal{L}(X) = \mathcal{L}(Y)$. For each $n \in \mathbb{N}$, we write $\mathcal{L}^n(X) = \{u \in \mathcal{L}(X) \colon |u| = n\}$ for the subset of $\mathcal{L}(X)$ consisting of words of length n.

An example of a subshift is given by the set of all bi-infinite sequences over the alphabet $\mathcal{A} = \{a, b\}$ that do not admit the word bb as a subword. This subshift is known as the *golden mean shift*. The golden mean shift is an example of a *shift of finite type*; namely, a subshift whose elements can be defined via a finite set of forbidden words. Shifts of finite type (in one dimension) are a well-studied class of dynamical systems whose topological and ergodic properties are largely well understood.

For a given subshift X, $u \in \mathcal{L}(X)$ and $m \in \mathbb{Z}$, the cylinder set of u in position m is the subset of X defined by

$$[u]_m = \left\{x \in X \colon x_{[m,m+|u|-1]} = u\right\}.$$

In the case $m = 0$, we omit the dependence on m and write $[u]$ for the cylinder set of u positioned at the origin. Let $\xi(X)$ denote the collection of all cylinder sets that specify the origin; namely,

$$\xi(X) = \{[u]_m \colon u \in \mathcal{L}(X),\, 1 - |u| \leq m \leq 0\},$$

together with the empty set $\varnothing$. The collection $\xi(X)$ forms a generating algebra for the Borel sigma-algebra on X. Thus, any pre-measure defined on cylinder sets extends uniquely to a measure on the Borel sigma-algebra on X by the Hahn–Kolmogorov extension theorem.

The *Bernoulli measures* are a well-studied family of measures on the full shift. On the full shift on two symbols $X = \{a, b\}^{\mathbb{Z}}$, they are defined as follows. Given $p \in (0, 1)$, let $\mu_p \colon \xi(X) \to [0, 1]$ be defined by $\mu_p([u]_m) = p^{|u|_a}(1 - p)^{|u|_b}$ for all $u \in \mathcal{L}(X)$ and $m \in \{1 - |u|, 2 - |u|, \ldots, 0\}$. The set function μ_p defines a pre-measure on the algebra $\xi(X)$, which extends uniquely to a measure on $\mathcal{B}(X)$ by the Hahn–Kolmogorov extension theorem. The measure μ_p is called the $(p, 1-p)$-*Bernoulli measure* on $\{a, b\}^{\mathbb{Z}}$. Bernoulli measures can also be defined on the full shift over larger alphabets in an analogous manner.

1.3 Topological entropy and word complexity

One of the most well-studied measures of complexity for subshifts is *topological entropy*, which quantifies the asymptotic growth rate of words admitted by a given subshift. For a subshift X, the topological entropy $h_{\mathrm{top}}(X, S)$ of X is the quantity defined by[1]

$$h_{\mathrm{top}}(X, S) = \lim_{n\to\infty} \frac{1}{n} \log \#\mathcal{L}^n(X);$$

that this limit exists is a routine consequence of Fekete's lemma [27]. For ease of notation, we often omit the explicit dependence on S in the notation and write $h_{\mathrm{top}}(X)$ for the topological entropy of a subshift X.

For a subshift X, the function $p_X\colon \mathbb{N} \to \mathbb{N}$ defined by $p_X(n) = \#\mathcal{L}^n(X)$ for all $n \in \mathbb{N}$ is called the *complexity function* of X. If X is a subshift with positive topological entropy, then the complexity function $p_X(n)$ grows exponentially in n. We note that the topological entropy of a subshift on a finite alphabet $\mathcal{A}$ cannot exceed $\log(\#\mathcal{A})$, so the complexity function cannot grow super-exponentially [75]. On the other hand, if X has zero topological entropy, then the complexity function grows sub-exponentially. In many cases, this growth is polynomial, but there also exist zero topological entropy subshifts with complexity function that grows faster than any polynomial [11]. In this latter case, we say that the complexity function has *intermediate growth*.

Classifying the functions that can be obtained as the complexity function of a subshift is a central problem in symbolic dynamics [9, 30]. One of the most famous results in this direction is due to Morse and Hedlund [57], which states that the complexity function of a subshift that is the orbit closure of a bi-infinite sequence either grows at least linearly or is bounded above by a constant. Ehrenfeucht and Rozenberg [20] later showed that the same holds for every subshift over a finite alphabet.

Proposition 1.3.1 ([20])**.** Let X be a subshift over a finite alphabet. Then, the complexity function p_X is either bounded above by a constant or $p_X(n) \geq n + 1$ for all $n \in \mathbb{N}$.

The dichotomy that the complexity function always grows at least linearly or is bounded above by a constant is sometimes referred to as the *Morse–Hedlund complexity gap* [12].

[1] Topological entropy can be defined more generally for an arbitrary topological dynamical system. We provide this more general definition in Appendix A.1.

1.4 Invariant measures and measure theoretic entropy

The classification of measure preserving dynamical systems is an important problem in ergodic theory. This problem dates back to von Neumann, who asked whether the full shift on three symbols equipped with the $(1/3, 1/3, 1/3)$-Bernoulli measure is (measure theoretically) isomorphic to the full shift on two symbols equipped with the $(1/2, 1/2)$-Bernoulli measure. In 1959, this question was answered in the negative by Kolmogorov and Sinai [44, 45, 73], who showed that these systems have different measure theoretic entropy and that this is an invariant of dynamical systems. Later, in 1970, it shown by Ornstein [61] that two Bernoulli shifts are isomorphic if and only if they have the same measure theoretic entropy.[2] Thus, entropy provides a powerful tool in the classification of dynamical systems.

Before we define the measure theoretic entropy of a subshift, we first recall some definitions from ergodic theory. A measure μ, supported on a subshift X, is called shift-invariant (or S-invariant) if $\mu(S^{-1}B) = \mu(B)$ for every Borel-measurable set B. We let $\mathcal{M}(X, S)$ denote the set of all shift-invariant measures supported on X. A shift-invariant measure μ is called *ergodic* if whenever $S^{-1}B = B$, we have $\mu(B) = 0$ or $\mu(B) = 1$. It is well known that every subshift supports at least one shift-invariant (respectively, ergodic) measure [75]. If a subshift X supports a unique shift-invariant measure, then we say that X is *uniquely ergodic*. In this case, the unique shift-invariant measure is ergodic [75].

For a subshift X and shift-invariant measure μ supported on X, the measure theoretic entropy $h_\mu(X, S)$ of X with respect to μ is the quantity defined by[3]

$$h_\mu(X, S) = \lim_{n\to\infty} \frac{1}{n} \sum_{u\in\mathcal{L}^n(X)} -\mu([u]) \log \mu([u]),$$

where, for each $n \in \mathbb{N}$, $\mathcal{L}^n(X)$ denotes the set of all words of length n admitted by the subshift. Again, this limit always exists by Fekete's lemma [75]. Similarly to topological entropy, we often suppress the dependence on S in the notation and write $h_\mu(X)$ for the measure theoretic entropy of X.

Measure theoretic and topological entropy are related by the *variational principle*, which

[2]We note that Ornstein's theorem only holds in the case of two-sided shifts. The isomorphism problem for one-sided shifts is more subtle – see [41] for more detials.

[3]We give a more general definition for an arbitrary measure-preserving dynamical system in Appendix A.1.

states that

$$h_{\text{top}}(X) = \sup_{\mu \in \mathcal{M}(X,S)} h_\mu(X).$$

In particular, $h_\mu(X) \leq h_{\text{top}}(X)$ for every $\mu \in \mathcal{M}(X,S)$. If μ is a measure such that $h_\mu(X) = h_{\text{top}}(X)$, then we say that μ is a *measure of maximal entropy*. For subshifts, there always exists at least one measure of maximal entropy.[4] If a subshift X has a unique measure of maximal entropy, then we say that X is *intrinsically ergodic*.

Measure theoretic entropy can often be difficult to calculate. However, it is well understood for many classes of subshifts, such as shifts of finite type. For example, if X is the full shift on two symbols, $p \in (0,1)$ and μ_p is the $(p, 1-p)$-Bernoulli measure, then

$$h_\mu(X) = -(p \log p + (1-p)\log(1-p)).$$

In the case $p = 1/2$, we have $h_\mu(X) = \log 2 = h_{\text{top}}(X)$, so the $(1/2, 1/2)$-Bernoulli measure is a measure of maximal entropy. Moreover, it can be shown that this is the unique measure of maximal entropy, so the subshift X is intrinsically ergodic.

Determining conditions under which a subshift is intrinsically ergodic is an important question in symbolic dynamics and ergodic theory. It was shown by Parry [63] that for all topologically transitive shifts of finite type, the *Parry measure* is a measure of maximal entropy. Adler and Weiss [2] later showed that this is the unique measure of maximal entropy. Hence, all transitive shifts of finite type are intrinsically ergodic. The approach developed by Parry and Adler and Weiss was extended by Bowen [10], who showed that all subshifts satisfying the *specification property* are intrinsically ergodic. Verifying the specification property has become the prototypical method for proving intrinsic ergodicity of subshifts, and several relaxations of this property have been provided in recent years. For a survey of recent progress, we refer the reader to [15].

1.5 Multifractal analysis

The local scaling properties of measures can be studied using tools from *multifractal analysis*. For a subshift X, the local scaling behaviour of a measure μ, supported on X, is quantified by the *local dimensions* and *multifractal spectrum*. For each $x \in X$, the local dimension of μ at x is

[4] We note that this does not hold for dynamical systems in general.

defined by

$$\dim_{\mathrm{loc}}(\mu, x) = \lim_{n\to\infty} \frac{-\log \mu([x_{[-n,n]}])}{2n+1},$$

provided this limit exists. The multifractal spectrum of μ quantifies the size of the set of points in X that have a given local dimension. Specifically, for each $\alpha \in [0,\infty)$, let

$$F_\mu(\alpha) = \{x \in X\colon \dim_{\mathrm{loc}}(\mu, x) = \alpha\}\,. \tag{1.1}$$

The size of the set $F_\mu(\alpha)$ is quantified via its *Hausdorff dimension*, with respect to a metric defined as follows. If $x, y \in \mathcal{A}^{\mathbb{Z}}$ are such that $x \neq y$ but $x_0 = y_0$, then we let $n(x,y)$ denote the largest integer such that $x_j = y_j$ for all $|j| \leq n$. The map $d\colon \mathcal{A}^{\mathbb{Z}} \times \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ given by

$$d(x,y) = \begin{cases} 1 & \text{if } x_0 \neq y_0 \\ \mathrm{e}^{-(2n(x,y)+1)} & \text{if } x \neq y \text{ but } x_0 = y_0 \\ 0 & \text{if } x = y \end{cases}$$

for all $x, y \in \mathcal{A}^{\mathbb{Z}}$ defines a metric on $\mathcal{A}^{\mathbb{Z}}$, which generates the discrete product topology – see [49] for more details. The *multifractal spectrum* of μ is the function $f_\mu\colon [0,\infty) \to [0,\infty)$ given by

$$f_\mu(\alpha) = \dim_{\mathrm{H}} F_\mu(\alpha)$$

for all $\alpha \in [0,\infty)$.

Computing the local dimensions and multifractal spectrum of a measure is often difficult. However, there is a related notion, called the L^q-spectrum, which is typically easier to establish for a given measure. The L^q-spectrum is the function $\tau_\mu\colon \mathbb{R} \to \mathbb{R}$ defined by

$$\tau_\mu(q) = \lim_{n\to\infty} -\frac{1}{n} \log \left(\sum_{u \in \mathcal{L}^n(X)} \mu([u])^q \right)$$

for all $q \in \mathbb{R}$, provided this limit exists. In some cases, the multifractal spectrum coincides with the concave conjugate of the L^q-spectrum. If this is the case, then we say that the *multifractal formalism* holds. It is an important question in multifractal analysis to determine settings in which the multifractal formalism holds, and to find qualitative conditions describing its failure.

1.6 Deterministic substitutions

Sequences associated with deterministic substitutions are the prototypical examples of mathematical quasicrystals. A deterministic substitution is a rule which replaces each symbol in a finite or infinite string over an alphabet $\mathcal{A}$ with a finite word over the same alphabet. Typically, deterministic substitutions are defined via their action on letters, which then extends to finite and infinite strings by concatenation. Two of the most well-studied deterministic substitutions are the *Fibonacci* and *period doubling* substitutions, which are defined over the two-letter alphabet $\mathcal{A} = \{a, b\}$ by

$$\theta_{\text{Fib}}\colon \begin{cases} a \mapsto ab, \\ b \mapsto a, \end{cases} \qquad \text{and} \quad \theta_{\text{PD}}\colon \begin{cases} a \mapsto ab, \\ b \mapsto aa, \end{cases}$$

respectively. Observe that powers of a deterministic substitution can be defined by iteration. Namely, if θ is a deterministic substitution and $k \in \mathbb{N}$, the k^{th} power of θ is defined by $\theta^k = \theta \circ \theta^{k-1}$. For example, for the Fibonacci substitution, we have $\theta_{\text{Fib}}^2(a) = \theta_{\text{Fib}}(\theta_{\text{Fib}}(a)) = \theta_{\text{Fib}}(ab) = aba$.

A subshift can be associated to a given deterministic substitution θ in a natural way. We say that a word $u \in \mathcal{A}^+$ is θ-legal if there exists a letter $a \in \mathcal{A}$ and a positive integer $k \in \mathbb{N}$ such that u is a subword of $\theta^k(a)$. Then, we let X_θ be the subspace of $\mathcal{A}^{\mathbb{Z}}$ defined by

$$X_\theta = \{x \in \mathcal{A}^{\mathbb{Z}} \colon \text{ every subword of } x \text{ is } \theta\text{-legal}\}.$$

Since the set of θ-legal words defines a language, X_θ is an S-invariant subspace of $\mathcal{A}^{\mathbb{Z}}$. Thus, X_θ defines a subshift. We call X_θ the *subshift associated with* θ.

A standard assumption in the study of deterministic substitutions is *primitivity*. We say that a deterministic substitution θ over a finite alphabet $\mathcal{A}$ is primitive if there exists a $k \in \mathbb{N}$ such that, for all $a \in \mathcal{A}$, every letter in $\mathcal{A}$ appears in $\theta^k(a)$. For example, the Fibonacci substitution is primitive since $\theta_{\text{Fib}}^2(a) = aba$ and $\theta_{\text{Fib}}^2(b) = ab$, and both of these words contain an occurrence of an a and a b. Similarly, the period doubling substitution is primitive. Many topological and dynamical properties of subshifts associated with primitive deterministic substitutions are well understood. For instance, such subshifts are always non-empty, minimal and either finite or homeomorphic to a Cantor set [5, 67]. Further, subshifts of primitive deterministic substitutions are always uniquely ergodic [52, 67].

One of the most striking properties of deterministic substitution subshifts is that they always have zero topological entropy [67], which is an indication of low complexity. Moreover, it was shown by Pansiot [62] that the complexity function of a deterministic substitution never grows faster than quadratic. In fact, Pansiot provided a complete classification of the possible asymptotic growth rates the complexity function of a deterministic substitution subshift can exhibit. Thus, for subshifts of deterministic substitutions, the class of possible complexity functions is significantly more restricted than is permitted for subshifts in general. In the following, and throughout, we say that a function $g\colon \mathbb{N} \to \mathbb{N}$ is $\Theta(f)$, for some function $f\colon \mathbb{N} \to \mathbb{N}$, if there exist constants $C_1, C_2 > 0$ such that $C_1 < g(n)/f(n) < C_2$ for all sufficiently large $n \in \mathbb{N}$. We also write g is $O(f)$ if the ratio $g(n)/f(n)$ is uniformly bounded from above, but not necessarily from below, and g is $o(f)$ if $g(n)/f(n)$ converges to 0 as $n \to \infty$.

Proposition 1.6.1 ([62])**.** Let θ be a deterministic substitution. Then, the complexity function of the subshift X_θ is either $\Theta(1)$, $\Theta(n)$, $\Theta(n \log\log n)$, $\Theta(n \log n)$ or $\Theta(n^2)$. Moreover, if θ is primitive, then the complexity function is $\Theta(1)$ or $\Theta(n)$.

1.7 Random substitutions

Random substitutions are a generalisation of deterministic substitutions where the substituted image of a letter is chosen from a fixed finite set according to a probability distribution. For example, given $p \in (0,1)$, we define the *random Fibonacci* and *random period doubling* substitutions by

$$\vartheta_{\mathrm{RF}}\colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto a \quad \text{with probability } 1, \end{cases}$$

and

$$\vartheta_{\mathrm{RPD}}\colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1, \end{cases}$$

respectively. The action of a random substitution can be extended to finite words by applying the random substitution independently to each letter. In a similar manner to deterministic substitutions, a subshift can be associated to a random substitution in a natural way (we highlight that the subshift is independent of the choice of non-degenerate probabilities assigned to the random substitution). However, in contrast to subshifts of deterministic substitutions, random substitution subshifts often have positive topological entropy and support uncountably many ergodic measures [70].

Random substitutions were first studied in the physics community, in the context of percolation theory. A seminal paper of Mandelbrot [50] on turbulence in a fluid initiated the study of fractal percolation, a phenomenon that random substitutions have proved useful in modelling [16, 17, 18]. The systematic study of random substitutions from a mathematical viewpoint was initiated by Godrèche and Luck [34] in 1989, who focused on a single example, the random Fibonacci substitution. Notably, it was there that positive topological entropy for random substitution subshifts was first identified. Following this discovery, the topological entropy was calculated for several families of random substitution subshifts, for example by Koslicki [46], Nilsson [58, 59], Spindeler [74] and Wing [76].

The mathematical theory of random substitutions and their associated subshifts has developed rapidly in recent years. In 2018, Baake, Spindeler and Strungaru [7] computed the diffraction measure for a family of random substitutions and Rust and Spindeler [70] established some key topological and dynamical properties of random substitution subshifts. For example, they showed that, under mild conditions, a random substitution subshift is topologically transitive and has uncountably many minimal components. Further, they provided a weak condition under which a random substitution subshift has positive topological entropy. Following these works, Rust [69] provided sufficient conditions under which a random substitution subshift does not admit periodic points and Miro et al [53] provided sufficient conditions for a random substitution subshift to be topologically mixing. In addition, it was shown by Gohlke, Rust and Spindeler [38] that every topologically transitive shift of finite type can be obtained as the subshift of a random substitution. However, despite this recent progress, the systematic study of random substitutions is still in its infancy, and many properties remain poorly understood.

The notion of *primitivity* for deterministic substitutions extends naturally to the random setting (we give the definition in Chapter 2). As in the deterministic setting, primitivity is a standard assumption in the study of random substitutions and is assumed in the majority of the aforementioned works. To each primitive random substitution, a measure can be associated in a canonical manner [70]. This measure, called the *frequency measure* corresponding to the random substitution, reflects the underlying probability distribution. It was shown by Gohlke and Spindeler [39] that for every primitive random substitution, the corresponding frequency measure is ergodic with respect to the shift action [39].

A systematic approach to calculating topological entropy of random substitution subshifts was provided by Gohlke [35] in 2020. There, the notion of *inflation word entropy* for a random substitution was introduced and shown to coincide with the topological entropy of the associated subshift, under the assumption that the underlying random substitution is primitive and *compatible* (we give the definition in Section 2.2). This allows the topological entropy to be calculated or accurately estimated for a broad class of random substitution subshifts. However, a limitation of Gohlke's result is that it requires the somewhat restrictive condition of compatibility. In Chapter 3 of this thesis, we continue the development of the theory of topological entropy for random substitution subshifts. In contrast to [35], we do not require the underlying assumption of compatibility in our work.

A limitation of topological entropy as a measure of complexity is that it is blind to the choice of probabilities attached to the random substitution. This is not the case for many important properties of random substitutions, such as word frequencies, which can be viewed as almost sure properties with respect to the corresponding frequency measure [70]. In Chapters 4–6, we view complexity from the perspective of this measure. In particular, we develop a theory of measure theoretic entropy for random substitution subshifts and study regularity properties of frequency measures from the perspective of multifractal analysis. Motivated by Gohlke's inflation word approach to topological entropy, we introduce new notions called the *measure theoretic inflation word entropy* and the *inflation word L^q-spectrum* of a primitive random substitution. We show that, in many cases, these notions coincide with the measure theoretic entropy and L^q-spectrum, respectively, where the measure in question is the frequency measure. Thus, our approach allows us to develop a robust theory of measure theoretic entropy and multifractal analysis for frequency measures arising from random substitutions.

1.8 Outline of thesis

This thesis is based on three papers the author wrote while a Ph.D. student [37, 54, 55]. We introduce the key definitions that we work with throughout in Chapter 2 and prove several basic properties of random substitutions. The majority of the results stated here are either proved in the paper [37] or available in the literature on random substitutions. However, at the end of this chapter, we provide a new proof that every Bernoulli measure on the full shift can be obtained as the frequency measure corresponding to a primitive random substitution.

Chapter 3 contains the main results from the solo-authored paper [54], on topological entropy and word complexity of random substitution subshifts. The main results proved in this chapter are the following.

- A primitive random substitution subshift has zero topological entropy if and only if it is the subshift of a deterministic substitution. This answers in the positive an open question of Rust and Spindeler [70].

- For all primitive constant length random substitutions, the topological entropy of the associated subshift coincides with the notion of inflation word entropy introduced by Gohlke in [35]. Together with the main theorem in [35], this shows that the inflation word entropy of a primitive random substitution coincides with the topological entropy of the associated subshift for all random substitutions for which it is well-defined.

- Without primitivity, a wide range of complexity behaviour can occur that is not possible in the deterministic or primitive random settings. For example, there exist non-primitive random substitution subshifts with intermediate growth complexity function. A partial classification of complexity functions for subshifts of constant length random substitutions is provided.

Chapters 4 and 5 are based on the paper [37], which was written in collaboration with P. Gohlke, D. Rust and T. Samuel. Chapter 4 focuses on measure theoretic entropy; then, in Chapter 5,we consider conditions under which a primitive random substitution gives rise to a frequency measure of maximal entropy. We give the definition of *measure theoretic inflation word entropy* in Section 4.1. The main results from these chapters are the following.

- For all primitive random substitutions, the notion of measure theoretic inflation word entropy coincides with the measure theoretic entropy of the associated subshift with respect to the corresponding frequency measure.
- In many cases, a closed-form formula for the measure theoretic entropy can be obtained.
- Under mild assumptions, a primitive random substitution subshift supports a frequency measure of maximal entropy.
- Under more restrictive assumptions, this measure is shown to be the unique measure of maximal entropy.
- The random period doubling substitution subshift is intrinsically ergodic.

Chapter 6 concerns multifractal properties of frequency measures. This chapter is based on the paper [55], which is joint work with A. Rutar. We introduce the notion of the *inflation word L^q-spectrum* of a random substitution in Section 6.2. In the main results of this chapter, we show the following.

- For primitive and compatible random substitutions, the inflation word L^q-spectrum coincides with the L^q-spectrum of the corresponding frequency measure for all $q > 0$.
- Under an additional assumption (recognisability), the two notions coincide for all $q \in \mathbb{R}$.
- In this latter setting, the multifractal formalism holds.

CHAPTER 2

RANDOM SUBSTITUTIONS

In this chapter we provide a systematic introduction to random substitutions and their associated subshifts. In Section 2.1, we give the definition of a random substitution and its associated subshift and state some basic properties of random substitution subshifts. Then, in Section 2.2, we introduce various conditions that it is natural to impose on random substitutions and discuss their consequences. Section 2.3 concerns frequency measures, the class of measures that arise naturally from random substitutions, which are our main objects of study in Chapters 4–6. After giving their definition and key properties, we provide the statement of a result which is fundamental to the proofs of our main results in Chapters 4–6.

Since the content of this chapter is largely foundational, many of the results stated here are not originally due to the author. We have taken care to provide citations to the original work where this is the case.

2.1 Random substitutions and their subshifts

2.1.1 Random substitutions

In a similar manner to [35, 37], we define a random substitution by the data required to determine its action on letters. We then extend this to a random map on words. In the following, and throughout, we let $\mathcal{F}(\cdot)$ denote the collection of all finite subsets of a given set.

Definition 2.1.1. Let $\mathcal{A} = \{a_1, \ldots, a_d\}$ be a finite alphabet. A *random substitution* $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is a *set-valued substitution* $\vartheta \colon \mathcal{A} \to \mathcal{F}(\mathcal{A}^+)$ together with a set of non-degenerate probability vectors

$$\mathbf{P} = \left\{ \mathbf{p}_i = (p_{i,1}, \ldots, p_{i,r_i}) : r_i = \#\vartheta(a_i);\ \mathbf{p}_i \in (0,1]^{r_i};\ \sum_{j=1}^{r_i} p_{i,j} = 1 \text{ for all } i \in \{1, \ldots, d\} \right\},$$

such that

$$\vartheta_{\mathbf{P}} \colon a_i \mapsto \begin{cases} s^{(i,1)} & \text{with probability } p_{i,1}, \\ \vdots & \vdots \\ s^{(i,r_i)} & \text{with probability } p_{i,r_i}, \end{cases}$$

for every $1 \leq i \leq d$, where $\vartheta(a_i) = \{s^{(i,j)}\}_{1 \leq j \leq r_i}$.

We call each $s^{(i,j)}$ a *realisation* of $\vartheta_{\mathbf{P}}(a_i)$. If there exists an $\ell \in \mathbb{N}$ such that $|s^{(i,j)}| = \ell$ for all $i \in \{1, \ldots, d\}$ and $j \in \{1, \ldots, r_i\}$, then we say that $\vartheta_{\mathbf{P}}$ is of *constant length* ℓ.

If θ is a deterministic substitution, defined over the alphabet $\mathcal{A}$, such that $\theta(a)$ is a realisation of $\vartheta_{\mathbf{P}}(a)$ for all $a \in \mathcal{A}$, then we say that θ is a *marginal* of $\vartheta_{\mathbf{P}}$.

Similarly to deterministic substitutions, the action of a random substitution can be extended to finite and bi-infinite words. In the following, we describe how a random substitution $\vartheta_{\mathbf{P}}$ determines a Markov process with state space $\mathcal{A}^+$ and transition matrix Q, indexed by $\mathcal{A}^+ \times \mathcal{A}^+$. We interpret the entry $Q_{u,v}$ as the probability of mapping a word u to a word v under the random substitution. Formally, $Q_{a_i, s^{(i,j)}} = p_{i,j}$ for all $j \in \{1, \ldots, r_i\}$ and $Q_{a_i,v} = 0$ if $v \notin \vartheta(a_i)$. We extend the action of $\vartheta_{\mathbf{P}}$ to finite words by mapping each letter *independently* to one of its realisations. More precisely, given $n \in \mathbb{N}$, $u = a_{i_1} \cdots a_{i_n} \in \mathcal{A}^n$ and $v \in \mathcal{A}^+$ with $|v| \geq n$, we let

$$\mathcal{D}_n(v) = \{(v^{(1)}, \ldots, v^{(n)}) \in (\mathcal{A}^+)^n : v^{(1)} \cdots v^{(n)} = v\}$$

denote the set of all decompositions of v into n individual words and set

$$Q_{u,v} = \sum_{(v^{(1)}, \ldots, v^{(n)}) \in \mathcal{D}_n(v)} \prod_{j=1}^{n} Q_{a_{i_j}, v^{(j)}}.$$

In other words, $\vartheta_{\mathbf{P}}(u) = v$ with probability $Q_{u,v}$.

For $u \in \mathcal{A}^+$, let $(\vartheta_{\mathbf{P}}^n(u))_{n \in \mathbb{N}}$ be a stationary Markov chain on some probability space

$(\Omega_u, \mathcal{F}_u, \mathbb{P}_u)$, with transition matrix given by Q; that is,

$$\mathbb{P}_u[\vartheta_{\mathbf{P}}^{n+1}(u) = w \mid \vartheta_{\mathbf{P}}^{n}(u) = v] = \mathbb{P}_v[\vartheta_{\mathbf{P}}(v) = w] = Q_{v,w}$$

for all $v, w \in \mathcal{A}^+$, and $n \in \mathbb{N}$. In particular,

$$\mathbb{P}_u[\vartheta_{\mathbf{P}}^{n}(u) = v] = (Q^n)_{u,v}$$

for all $u, v \in \mathcal{A}^+$, and $n \in \mathbb{N}$. We often write $\mathbb{P}$ for $\mathbb{P}_u$ if the initial word is understood. In this case, we also write $\mathbb{E}$ for the expectation with respect to $\mathbb{P}$. As before, we call v a *realisation* of $\vartheta_{\mathbf{P}}^{n}(u)$ if $(Q^n)_{u,v} > 0$ and set

$$\vartheta^n(u) = \{v \in \mathcal{A}^+ : (Q^n)_{u,v} > 0\}$$

to be the set of all realisations of $\vartheta_{\mathbf{P}}^{n}(u)$. Conversely, we may regard $\vartheta_{\mathbf{P}}^{n}(u)$ as the set $\vartheta^n(u)$ endowed with the additional structure of a probability vector. If $u = a \in \mathcal{A}$ is a letter, we call a word $v \in \vartheta^n(a)$ a *level-n inflation word*, or *exact inflation word*.

We now give the definitions of two of the most well-studied random substitutions: the random Fibonacci substitution and the random period doubling substitution. We will often refer back to these two guiding examples throughout this thesis.

Example 2.1.2 (Random Fibonacci)**.** Let $\mathcal{A} = \{a, b\}$ and let $p \in (0, 1)$. The *random Fibonacci substitution* $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is the random substitution given by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto a \quad \text{with probability } 1, \end{cases}$$

with defining data $r_a = 2$, $r_b = 1$, $s^{(a,1)} = ab$, $s^{(a,2)} = ba$, $s^{(b,1)} = a$, $\mathbf{P} = \{\mathbf{p}_a = (p, 1-p), \mathbf{p}_b = (1)\}$ and corresponding set-valued substitution $\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{a\}$.

Example 2.1.3 (Random period doubling)**.** Let $\mathcal{A} = \{a, b\}$ and let $p \in (0, 1)$. The *random* *period doubling substitution* $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1 \end{cases}$$

with defining data $r_a = 2$, $r_b = 1$, $s^{(a,1)} = ab$, $s^{(a,2)} = ba$, $s^{(b,1)} = aa$, $\mathbf{P} = \{\mathbf{p}_a = (p, 1-p), \mathbf{p}_b = (1)\}$ and corresponding set-valued substitution $\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{aa\}$. Observe that $\vartheta_{\mathbf{P}}$ is a constant length random substitution of length 2.

For both the random Fibonacci and random period doubling substitutions, for all $a \in \mathcal{A}$ and $k \in \mathbb{N}$, every realisation of $\vartheta_{\mathbf{P}}^k(a)$ has the same length. Where this is the case, we write $|\vartheta^k(a)|$ for the common length. This is not always the case for random substitutions and, in general, the length $|\vartheta_{\mathbf{P}}(a)|$ is a random variable. We also write $|\vartheta| = \max_{a \in \mathcal{A}} \max_{v \in \vartheta(a)} |v|$ for the maximal length of an exact inflation word.

2.1.2 The subshift associated to a random substitution

To a given random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$, one can associate a subshift in a similar manner to the deterministic setting, via a language generated by the random substitution.

Definition 2.1.4. Let $\vartheta = (\vartheta, \mathbf{P})$ be a random substitution over a finite alphabet $\mathcal{A}$. We say that a word $u \in \mathcal{A}^+$ is *(ϑ-)legal* if there exists an $a_i \in \mathcal{A}$ and $k \in \mathbb{N}$ such that u appears as a subword of some word in $\vartheta^k(a_i)$. The *language of* ϑ is the subset of $\mathcal{A}^+$ defined by $\mathcal{L}_\vartheta = \{u \in \mathcal{A}^+ : u \text{ is } \vartheta\text{-legal}\}$.

Definition 2.1.5. The *random substitution subshift* of a random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is the system (X_ϑ, S), where $X_\vartheta = \{x \in \mathcal{A}^{\mathbb{Z}} : \text{ every subword of } x \text{ is } \vartheta\text{-legal}\}$ and S denotes the usual (left) shift map.

Under very mild assumptions, the space X_ϑ is non-empty (we give sufficient conditions in Section 2.1.3). We endow X_ϑ with the subspace topology inherited from $\mathcal{A}^{\mathbb{Z}}$. Since X_ϑ is defined in terms of a language, it is a compact S-invariant subspace of $\mathcal{A}^{\mathbb{Z}}$; hence, X_ϑ is a subshift. For $n \in \mathbb{N}$, we write $\mathcal{L}_\vartheta^n = \{u \in \mathcal{L}_\vartheta \colon |u| = n\}$ for the subset of $\mathcal{L}_\vartheta$ consisting of words of length n.

The notation X_ϑ reflects the fact that the random substitution subshift does not depend on the choice of (non-degenerate) probabilities $\mathbf{P}$. In fact, this is the case for many structural properties of $\vartheta_{\mathbf{P}}$. In these cases, one sometimes refers to ϑ instead of $\vartheta_{\mathbf{P}}$ as a random substitution, see for instance [35, 38, 69, 70]. On the other hand, for some applications, one needs additional structure on the probability space. For example, the measure theoretic properties we consider in Chapters 4–6 are in general dependent on the explicit choice of probabilities.

The set-valued function ϑ extends naturally to X_ϑ, where for $w = \cdots w_{-2}w_{-1} \cdot w_0 w_1 \cdots \in X_\vartheta$, we let $\vartheta(w)$ denote the (possibly infinite) set of sequences of the form $x = \cdots x_{-2}x_{-1} \cdot x_0 x_1 \cdots$ with $x_j \in \vartheta(w_j)$ for all $j \in \mathbb{Z}$. It follows routinely from the definition of X_ϑ that $\vartheta(X_\vartheta) \subseteq X_\vartheta$. Some properties of ϑ are reminiscent of continuous functions, although ϑ itself is *not* a function. In the following, we recall that we equip $\mathcal{A}^{\mathbb{Z}}$ with the discrete product topology and a given subshift of $\mathcal{A}^{\mathbb{Z}}$ with the subspace topology inherited from $\mathcal{A}^{\mathbb{Z}}$.

Lemma 2.1.6. If $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is a random substitution and $X \subset \mathcal{A}^{\mathbb{Z}}$ is compact, then $\vartheta(X)$ is compact.

Proof. It suffices to show that $\vartheta(X)$ is closed. Let $(y^{(n)})_{n\in\mathbb{N}}$ be a sequence in $\vartheta(X)$ and assume that this sequence converges to some $y \in \mathcal{A}^{\mathbb{Z}}$. We show that $y \in \vartheta(X)$. To this end, let $(x^{(n)})_{n\in\mathbb{N}}$ be a sequence in X with $y^{(n)} \in \vartheta(x^{(n)})$ for all $n \in \mathbb{N}$. By the compactness of X, this sequence has an accumulation point $x = \cdots x_{-1}x_0x_1 \cdots \in X$. By restricting to an appropriate subsequence, we may assume that

$$x^{(m)}_{[-n,n]} = x_{[-n,n]}$$

for all $m, n \in \mathbb{N}$ with $m \geq n$. In which case,

$$y^{(n)}_{[-n,n]} = w^{(n)}_{-n} \cdots w^{(n)}_{-1} \cdot w^{(n)}_0 \cdots w^{(n)}_n$$

with $w^{(n)}_j \in \vartheta(x_j)$ for all $j \in \{-n, \ldots, n\}$. As $(y^{(m)})_{m\in\mathbb{N}}$ converges to y, we may assume, for $n \in \mathbb{N}$,

$$y_{[-n,n]} = w^{(n)}_{-n} \cdots w^{(n)}_{-1} \cdot w^{(n)}_0 \cdots w^{(n)}_n,$$

again by possibly restricting to an appropriate subsequence. By a standard diagonal argument utilising the pigeonhole principle, we can choose $w_j \in \vartheta(x_j)$ for all $j \in \mathbb{Z}$ such that $y = \cdots w_{-2}w_{-1}.w_0w_1w_2\cdots$. Namely, we have that $y \in \vartheta(x)$. □

2.1.3 Primitive random substitutions

Similarly to deterministic substitutions, the action of a random substitution on the abelianisation of a word can be encoded by a matrix. This allows a natural analogue of primitivity for random substitutions to be defined. In general, the matrix of a random substitution depends on the choice of probabilities.

Definition 2.1.7. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a random substitution over a finite alphabet $\mathcal{A} = \{a_1, \ldots, a_d\}$. The *matrix of* $\vartheta_{\mathbf{P}}$ is the $d \times d$ matrix $M = M_{\vartheta_{\mathbf{P}}}$ defined by

$$M_{i,j} = \mathbb{E}[|\vartheta_{\mathbf{P}}(a_j)|_{a_i}] = \sum_{k=1}^{r_j} p_{j,k} |s^{(j,k)}|_{a_i}.$$

Definition 2.1.8. We say that $\vartheta_{\mathbf{P}}$ is *primitive* if the matrix of $\vartheta_{\mathbf{P}}$ is a primitive matrix.

In contrast to the deterministic setting, primitivity is not sufficient to guarantee that the subshift associated to a given random substitution is non-empty. For example, the random substitution $\vartheta \colon a, b \mapsto \{a, b\}$ is primitive, but $X_\vartheta = \varnothing$. However, it was shown by Rust and Spindeler [70, Prop. 9] that a primitive random substitution gives rise to an empty subshift if and only if, for every $a \in \mathcal{A}$, $\vartheta(a)$ consists only of realisations of length 1. To keep the statements of our main results simple, we exclude these pathological cases and from now on restrict the definition of primitivity to those random substitutions that give rise to a non-empty subshift.

By the Perron–Frobenius theorem, if $\vartheta_{\mathbf{P}}$ is a primitive random substitution, then there exists a real number $\lambda \geq 1$ such that λ is an eigenvalue of the matrix of $\vartheta_{\mathbf{P}}$, and every other eigenvalue is strictly smaller than λ in modulus. Moreover, λ is a simple eigenvalue, so has a one-dimensional eigenspace. Further, the left and right eigenvectors corresponding to the eigenvalue λ consist of positive real entries. We normalise $\mathbf{R}$ such that its entries sum to 1 and normalise $\mathbf{L}$ such that $\mathbf{L}^\top \cdot \mathbf{R} = 1$. For simplicity, we call λ the *Perron–Frobenius eigenvalue* of $\vartheta_{\mathbf{P}}$ (as opposed to the Perron–Frobenius eigenvalue of the substitution matrix of $\vartheta_{\mathbf{P}}$). Similarly, we call $\mathbf{L}$ and $\mathbf{R}$ the *left* and *right eigenvectors* of $\vartheta_{\mathbf{P}}$. Together, we call $(\lambda, \mathbf{L}, \mathbf{R})$ the *Perron–Frobenius data* of $\vartheta_{\mathbf{P}}$.

The Perron–Frobenius data of a random substitution encodes information on expected inflation lengths and letter frequencies. This is summarised in the following.

Proposition 2.1.9 ([39, Lemma 3.5])**.** Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with Perron–Frobenius data $(\lambda, \mathbf{L}, \mathbf{R})$. Then, for all $a \in \mathcal{A}$, we have

$$\frac{\mathbb{E}[|\vartheta_{\mathbf{P}}^m(a)|]}{\mathbb{E}[|\vartheta_{\mathbf{P}}^{m-1}(a)|]} \to \lambda \quad \text{and} \quad \frac{\mathbb{E}[|\vartheta_{\mathbf{P}}^m(a)|]}{\lambda^m} \to L_a$$

as $m \to \infty$. Moreover, for all $b \in \mathcal{A}$, we have

$$\frac{\mathbb{E}[|\vartheta_{\mathbf{P}}^m(a)|_b]}{\mathbb{E}[|\vartheta_{\mathbf{P}}^m(a)|]} \to R_b$$

as $m \to \infty$.

If $\vartheta_{\mathbf{P}}$ is a random substitution and M is the matrix of $\vartheta_{\mathbf{P}}$, then for all $k \in \mathbb{N}$, M^k is the matrix of the random substitution $\vartheta_{\mathbf{P}}^k$. Thus, if $\vartheta_{\mathbf{P}}$ is primitive, then $\vartheta_{\mathbf{P}}^k$ is primitive for all $k \in \mathbb{N}$. Moreover, if $(\lambda, \mathbf{L}, \mathbf{R})$ is the Perron–Frobenius data for $\vartheta_{\mathbf{P}}$, then $(\lambda^k, \mathbf{L}, \mathbf{R})$ is the Perron–Frobenius data for $\vartheta_{\mathbf{P}}^k$.

While the matrix of a random substitution is dependent on the choice of probabilities, primitivity itself is independent of the choice of (non-degenerate) probabilities. In fact, primitivity can be characterised entirely in terms of the underlying set-valued substitution.

Proposition 2.1.10 ([70, Def. 4])**.** Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be random substitution over a finite alphabet $\mathcal{A}$. Then $\vartheta_{\mathbf{P}}$ is primitive if and only if there exists a positive integer K such that, for all $a, b \in \mathcal{A}$, the letter a appears as a subword of some realisation of $\vartheta^K(b)$.

We emphasise that while primitivity is independent of the choice of probabilities, the Perron–Frobenius data corresponding to the random substitution does, in general, depend on the probabilities. This is highlighted by the following example.

Example 2.1.11. Let $p \in (0,1)$, and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \to \begin{cases} aa & \text{with probability } p, \\ ab & \text{with probability } 1-p, \end{cases} \\ b \to a \quad \text{with probability } 1. \end{cases}$$

We have that $\vartheta_{\mathbf{P}}(a) = aa$ with probability p and $\vartheta_{\mathbf{P}}(a) = ab$ with probability $1-p$, so $\mathbb{E}[|\vartheta_{\mathbf{P}}(a)|_a] = 2p + (1-p) = 1+p$. Similarly, $\mathbb{E}[|\vartheta_{\mathbf{P}}(a)|_b] = 1-p$, $\mathbb{E}[|\vartheta_{\mathbf{P}}(b)|_a] = 1$ and $\mathbb{E}[|\vartheta_{\mathbf{P}}(b)|_a] = 0$, so the matrix of $\vartheta_{\mathbf{P}}$ is given by

$$M = \begin{pmatrix} \mathbb{E}[|\vartheta_{\mathbf{P}}(a)|_a] & \mathbb{E}[|\vartheta_{\mathbf{P}}(b)|_a] \\ \mathbb{E}[|\vartheta_{\mathbf{P}}(a)|_b] & \mathbb{E}[|\vartheta_{\mathbf{P}}(b)|_b] \end{pmatrix} = \begin{pmatrix} 1+p & 1 \\ 1-p & 0 \end{pmatrix}.$$

The Perron–Frobenius eigenvalue of M is $\lambda = (p+1+\sqrt{p^2-2p+5})/2$ and the corresponding (normalised) right eigenvector is $\mathbf{R} = (1-f(p), f(p))$, where

$$f(p) = \frac{2(p-1)}{3p-1-\sqrt{p^2-2p+5}}.$$

The above example also illustrates that primitivity is not preserved under passing to marginals. The marginal at $p=1$ is the deterministic substitution $\theta\colon a \mapsto aa$, $b \mapsto a$, which is not primitive since b does not appear as a subword of $\theta^k(a)$ or $\theta^k(b)$ for any $k \in \mathbb{N}$.

2.2 Special classes of random substitutions

Since primitive random substitutions give rise to a wide variety of subshifts, including all topologically transitive shifts of finite type [38] and all primitive deterministic substitution subshifts, it is reasonable to impose additional conditions in their study. In this section, we introduce some common conditions imposed on random substitutions.

2.2.1 Compatible random substitutions

One of the most common assumptions in the study of random substitutions is *compatibility*, which states that the inflated image of every word has a well-defined abelianisation. Compatibility is a fundamental assumption in Gohlke's work on topological entropy [35], which we summarise in Chapter 3, and is also assumed in the main results of [7, 33, 53, 69]. In the following, recall that for a given $u \in \mathcal{A}^{\mathbb{Z}}$, we let $\Phi(u) = (|u|_a)_{a \in \mathcal{A}}$ denote the abelianisation of u.

Definition 2.2.1. We say that a random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is *compatible* if for all $a \in \mathcal{A}$ and $u, v \in \vartheta(a)$, we have $\Phi(u) = \Phi(v)$.

Compatibility guarantees that the length $|\vartheta^k(a)|$ is well-defined for all $a \in \mathcal{A}$ and $k \in \mathbb{N}$. Further, if $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is compatible, then for all $u \in \mathcal{A}^+$, $a \in \mathcal{A}$ and $s, t \in \vartheta(u)$, we have that $|s|_a = |t|_a$. We write $|\vartheta(u)|_a$ to denote this common value. It is straightforward to verify that if $\vartheta_{\mathbf{P}}$ is compatible, then $\vartheta_{\mathbf{P}}^k$ is compatible for all $k \in \mathbb{N}$.

Note that the matrix of a compatible random substitution is independent of the choice of probabilities $\mathbf{P}$. Therefore, if a random substitution is both primitive and compatible, the Perron–Frobenius eigenvalue and corresponding right eigenvector do not depend on $\mathbf{P}$. In fact, there is a *uniform* inflation rate and *uniform* letter frequencies, which are encoded by the Perron–Frobenius data. Specifically, we have the following.

Proposition 2.2.2 ([35, Prop. 13])**.** If $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is a primitive and compatible random substitution with Perron–Frobenius data $(\lambda_1, \mathbf{L}, \mathbf{R})$, then for all $\varepsilon > 0$ there is an integer N such that every word $v \in \mathcal{L}_\vartheta$ of length at least N satisfies

$$|v|(R_a - \varepsilon) < |v|_a < |v|(R_a + \varepsilon)$$

for all $a \in \mathcal{A}$, where R_a is the entry of $\mathbf{R}$ corresponding to a. Consequently, for all $a, b \in \mathcal{A}$,

$$\frac{|\vartheta^k(b)|_a}{|\vartheta^k(b)|} \to R_a$$

as $k \to \infty$. Hence, for all $x \in X_\vartheta$,

$$\frac{|x_{[-n,n]}|_a}{2n+1} \to R_a$$

as $n \to \infty$.

The following is proved in [67, Prop. 5.8] for primitive deterministic substitutions. The proof presented there extends to primitive and compatible random substitutions.

Proposition 2.2.3. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive and compatible random substitution, with Perron–Frobenius data $(\lambda_1, \mathbf{L}, \mathbf{R})$. Further, let λ_2 denote the second largest (in absolute value) eigenvalue of the substitution matrix. Then, there exists a constant $c > 0$ such that for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$,

$$\lambda_1^m L_a - c|\lambda_2|^m \leq |\vartheta^m(a)| \leq \lambda_1^m L_a + c|\lambda_2|^m,$$

where $|\lambda_2|$ denotes the modulus of λ_2. In particular,

$$\frac{|\vartheta^m(a)|}{\lambda_1^m} \to L_a.$$

as $m \to \infty$ for all $a \in \mathcal{A}$.

The random Fibonacci substitution defined in Example 2.1.2 is compatible, since $\Phi(ab) = \Phi(ba) = (1,1)$. For any choice of probabilities, the right Perron–Frobenius eigenvector is given by $(\tau^{-1}, \tau^{-2})^\top$, where τ denotes the golden ratio. Thus, in every element of the associated subshift, the letter a occurs with frequency τ^{-1} and the letter b occurs with frequency τ^{-2}. Similarly, the random period doubling substitution defined in Example 2.1.3 is compatible, with right Perron–Frobenius eigenvector $(2/3, 1/3)^\top$.

2.2.2 Generalisations of compatibility

For compatible random substitutions, the Perron–Frobenius data $(\lambda, \mathbf{L}, \mathbf{R})$ is independent of the choice of probabilities. A natural generalisation of compatibility is to allow $\mathbf{R}$ to depend on the probabilities but still insist that λ and $\mathbf{L}$ are independent of the probabilities. Such random substitutions are called *geometrically compatible*, since this is the natural setting in which a random substitution can be viewed as a random inflation rule on an associated tiling dynamical system. For more details, we refer the reader to [36].

Definition 2.2.4. We say that a primitive random substitution $\vartheta_{\mathbf{P}}$ is *geometrically compatible* if there is a real number $\lambda > 1$ and a vector $\mathbf{L}$ with strictly positive entries, such that $\mathbf{L}$ is a left eigenvector with eigenvalue λ for all marginals of $\vartheta_{\mathbf{P}}$.

As well as encompassing all compatible primitive random substitutions, the class of geometrically compatible random substitutions also includes all primitive random substitutions of constant length.

Lemma 2.2.5. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution over a finite alphabet $\mathcal{A}$. If $\vartheta_{\mathbf{P}}$ is compatible or constant length, then $\vartheta_{\mathbf{P}}$ is geometrically compatible.

Proof. For compatible random substitutions, the result follows from the fact that the substitution matrix is independent of the choice of probabilities and is preserved under passing to marginals, so it remains to consider the constant length case. To this end, let $\vartheta_{\mathbf{P}}$ be a primitive random substitution of constant length ℓ and let θ be a marginal of $\vartheta_{\mathbf{P}}$. The constant length property is preserved under passing to marginals, so θ has Perron–Frobenius eigenvalue $\lambda = \ell$. Let $\mathbf{L} = (1, \ldots, 1)^{\top} \in \mathbb{R}^{\#\mathcal{A}}$. Since θ is of constant length ℓ, we have that $\sum_{b \in \mathcal{A}} \mathbb{E}[|\vartheta(a)|_b] = \ell$ for all $a \in \mathcal{A}$. In particular, the column sums of the substitution matrix M are all equal to ℓ. Hence, $\ell \, \mathbf{L}^{\top} = M^{\top} \mathbf{L}^{\top}$, so $\mathbf{L}$ is a left eigenvector for the eigenvalue ℓ and we conclude that $\vartheta_{\mathbf{P}}$ is geometrically compatible. $\square$

The reverse inclusion in Lemma 2.2.5 does not hold. Namely, there exist geometrically compatible random substitutions that are neither compatible nor constant length. This is illustrated by the following example.

Example 2.2.6. Let $\vartheta_{\mathbf{P}}$ be the primitive random substitution defined over the alphabet $\mathcal{A} = \{a, b\}$ by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto & abb \quad \text{with probability } 1, \\ b \mapsto & \begin{cases} a & \text{with probability } p, \\ bb & \text{with probability } 1 - p. \end{cases} \end{cases}$$

The random substitution $\vartheta_{\mathbf{P}}$ is geometrically compatible with $\mathbf{L} = (2, 1)^{\top}$ and $\lambda = 2$. We highlight that $\vartheta_{\mathbf{P}}$ is neither constant length nor compatible.

Beyond geometric compatibility, there is the class of random substitutions with *unique realisation paths*. Recall that for $v = v_1 \cdots v_n$, the random word $\vartheta_{\mathbf{P}}(v) = \vartheta_{\mathbf{P}}(v_1) \cdots \vartheta_{\mathbf{P}}(v_n)$ can be written as a concatenation of the random variables $\vartheta_{\mathbf{P}}(v_1), \ldots, \vartheta_{\mathbf{P}}(v_n)$. In general, there might be several realisations of $(\vartheta_{\mathbf{P}}(v_1), \ldots, \vartheta_{\mathbf{P}}(v_n))$ that concatenate to the same realisation of $\vartheta_{\mathbf{P}}(v)$. For random substitutions with unique realisation paths, this phenomenon can be excluded.

Definition 2.2.7. We say that $\vartheta_{\mathbf{P}}$ has *unique realisation paths* if for every $v \in \mathcal{L}_{\vartheta}$ and $k \in \mathbb{N}$, the vector $(\vartheta_{\mathbf{P}}^k(v_1), \ldots, \vartheta_{\mathbf{P}}^k(v_{|v|}))$ is completely determined by $\vartheta_{\mathbf{P}}^k(v)$.

While the definition above is most adequate for our purposes, we note that the property of having unique realisation paths does not depend on the choice of $\mathbf{P}$. Indeed, it is straightforward to verify that a random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ has unique realisation paths if and only if for all $v \in \mathcal{L}_\vartheta$ and $k \in \mathbb{N}$ the concatenation map

$$\vartheta^k(v_1) \times \cdots \times \vartheta^k(v_{|v|}) \to \mathcal{L}_\vartheta, \quad (w^1, \ldots, w^{|v|}) \mapsto w^1 \cdots w^{|v|}$$

is injective.

Lemma 2.2.8. Every primitive, geometrically compatible random substitution has unique realisation paths.

Proof. Let $\vartheta_{\mathbf{P}}$ be primitive and geometrically compatible. Since the same holds for $\vartheta_{\mathbf{P}}^k$, we may restrict to the case $k = 1$ in the following. Let $v \in \mathcal{L}_\vartheta$ and let u be a realisation of the random word

$$\vartheta_{\mathbf{P}}(v) = \vartheta_{\mathbf{P}}(v_1) \cdots \vartheta_{\mathbf{P}}(v_{|v|})$$

and $(u^1, \ldots, u^{|v|})$ a corresponding realisation of $(\vartheta_{\mathbf{P}}(v_1), \ldots, \vartheta_{\mathbf{P}}(v_{|v|}))$ satisfying

$$u = u^1 \cdots u^{|v|}.$$

Let M_1 be the substitution matrix of a marginal of $\vartheta_{\mathbf{P}}$ with $v_1 \mapsto u_1$. Since $\mathbf{L}$ has strictly positive entries, there is a unique $1 \leq m \leq |u|$ such that

$$\mathbf{L}\Phi(u_{[1,m]}) = \mathbf{L}\Phi(u^1) = \mathbf{L}M_1\Phi(v_1) = \lambda \mathbf{L}_{v_1}.$$

This determines $u^1 = u_{[1,m]}$ unambiguously. Inductively, we find that u^j is uniquely determined by u for all $1 \leq j \leq |v|$. □

For the reader's convenience, we summarise the relation between different characterisations of primitive random substitutions in Figure 2.1.

2.2.3 Separation conditions

In this section, we introduce additional conditions that either (1) impose a certain separation on inflation words, or (2) impose a certain uniformity on the inflation and the probabilities.

$$\begin{tikzcd}[column sep=small]
\text{Constant Length} \arrow[dr, Rightarrow] & & \text{Compatibility} \arrow[dl, Rightarrow] \\
& \text{Geometric Compatibility} \arrow[d, Rightarrow] & \\
& \text{Unique Realisation Paths} &
\end{tikzcd}$$

Figure 2.1: Implication diagram for some conditions on primitive random substitutions.

Definition 2.2.9. A random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ satisfies the *disjoint set condition* if

$$u, v \in \vartheta(a) \text{ with } u \neq v \implies \vartheta^k(u) \cap \vartheta^k(v) = \varnothing$$

for all $a \in \mathcal{A}$ and $k \in \mathbb{N}$. It satisfies the *identical set condition* if

$$u, v \in \vartheta(a) \implies \vartheta^k(u) = \vartheta^k(v)$$

for all $a \in \mathcal{A}$ and $k \in \mathbb{N}$. Moreover, we say that $\vartheta_{\mathbf{P}}$ has *identical production probabilities* if for all $a \in \mathcal{A}$, $k \in \mathbb{N}$ and $v \in \vartheta^k(a)$,

$$\mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(u_1) = v] = \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(u_2) = v]$$

for all $u_1, u_2 \in \vartheta(a)$.

Observe that the identical and disjoint set condition depend only on the underlying set-valued substitution. However, the property of having identical production probabilities is dependent on the probability distribution.

Remark 2.2.1. The conditions introduced in Definition 2.2.9 can also be rephrased in probabilistic terms. Specifically, $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition if and only if $\vartheta_{\mathbf{P}}(a)$ is determined by $\vartheta_{\mathbf{P}}^n(a)$ for all $n \in \mathbb{N}$ and $a \in \mathcal{A}$. The identical set condition with identical production probabilities holds for $\vartheta_{\mathbf{P}}$ if and only if the random words $\vartheta_{\mathbf{P}}(a)$ and $\vartheta_{\mathbf{P}}^n(a)$ are independent for all $n \geq 2$ and $a \in \mathcal{A}$. This formulation will be useful when we consider measure theoretic entropy in Chapter 4.

In general, it is not straightforward to verify whether a given random substitution satisfies the identical or disjoint set condition. However, for some families of random substitutions there are sufficient conditions that are easy to check.

Proposition 2.2.10. If $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is a random substitution such that $\vartheta(a) = \vartheta(b)$ for all $a, b \in \mathcal{A}$, then $\vartheta_{\mathbf{P}}$ satisfies the identical set condition.

Proof. Since $\vartheta(a) = \vartheta(b)$ for all $a, b \in \mathcal{A}$, it follows inductively that $\vartheta^k(a) = \vartheta^k(b)$ for all $k \in \mathbb{N}$. Hence, $\vartheta_{\mathbf{P}}$ satisfies the identical set condition. □

The sufficient condition for the identical set condition given by Proposition 2.2.10 is not a necessary condition. For example, any random substitution defined over the set-valued substitution $\vartheta \colon a, b \mapsto \{abc, bac\}$, $c \mapsto \{acb, bca\}$ satisfies the identical set condition, but $\vartheta(a) \neq \vartheta(c)$.

Proposition 2.2.11. If $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is a constant length random substitution such that $\vartheta(a) \cap \vartheta(b) = \varnothing$ for all $a, b \in \mathcal{A}$ with $a \neq b$, then $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition.

Proof. We show that, for each $k \in \mathbb{N}$ and each level-$(k+1)$ inflation word w, there is a *unique* level-k inflation word v such that $w \in \vartheta(v)$. By the constant length property, w can be decomposed into ℓ^k level-1 inflation words $w = w^1 \cdots w^{\ell^k}$ so that $w^i \in \vartheta(v_i)$ for all $i \in \{1, \ldots, \ell^k\}$. Moreover, this is the unique such decomposition of w into level-1 inflation words. By assumption, v_i is the unique letter such that $w^i \in \vartheta(v_i)$, so it follows that v is the unique level-k inflation word such that $w \in \vartheta(v)$. It then follows inductively that there is a unique level-1 inflation word u such that $w \in \vartheta^k(u)$. Since this holds for all $k \in \mathbb{N}$ and all level-$(k+1)$ inflation words, we conclude that the disjoint set condition is satisfied. □

Example 2.2.12. Given $p \in (0,1)$, let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random period doubling substitution

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1. \end{cases}$$

Since ϑ is constant length and $\vartheta(a) \cap \vartheta(b) = \varnothing$, Proposition 2.2.11 gives that $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition.

Example 2.2.13. Given $p \in (0,1)$, let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random Fibonacci substitution

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto a \quad \text{with probability } 1. \end{cases}$$

Observe that $ab, ba \in \vartheta(a)$ and $aba \in \vartheta(ab) \cap \vartheta(ba)$, so $\vartheta_{\mathbf{P}}$ does not satisfy the disjoint set condition. On the other hand, $aab \in \vartheta(ba) \setminus \vartheta(ab)$, so $\vartheta(ab) \neq \vartheta(ba)$ and thus $\vartheta_{\mathbf{P}}$ does not satisfy the identical set condition either.

The above example illustrates that the constant length assumption in Proposition 2.2.11 cannot be dropped. The issue in the non-constant length setting is that an inflation word can have multiple different decompositions into lower level inflation words. For the random Fibonacci substitution, this happens for the level-2 inflation word aba, which can be decomposed into level-1 inflation words as (ab, a) or (a, ba).

2.2.4 Recognisable random substitutions

A consequence of the disjoint set condition is that for every $a \in \mathcal{A}$, $k \in \mathbb{N}$ and $w \in \vartheta^k(a)$, there is a unique $v \in \vartheta^{k-1}(a)$ such that $w \in \vartheta(v)$. In other words, every exact inflation word can be uniquely "de-substituted" to another exact inflation word. The following definition extends this idea from inflation words to all elements in the subshift.

Definition 2.2.14. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ denote a random substitution over a finite alphabet $\mathcal{A}$, and suppose that $|\vartheta(a)|$ is well-defined for all $a \in \mathcal{A}$. We call $\vartheta_{\mathbf{P}}$ *recognisable* if, for all $x \in X_\vartheta$, there exist a unique $y = \cdots y_{-1} y_0 y_1 \cdots \in X_\vartheta$ and a unique integer $k \in \{0, \ldots, |\vartheta(y_0)| - 1\}$ with $S^{-k}(x) \in \vartheta(y)$.

In other words, for all $i \in \mathbb{Z}$, there exist words $w^i \in \vartheta(y_i)$ such that x can be uniquely decomposed into inflation words as $(\ldots, w^{-1}, w^0, w^1, \ldots)$. We call each w^i an *inflation tile* or *supertile* of x and call each index j_i such that $x_{j_i} = w^i_0$ a *cutting point* of inflation tiles.

A common strategy for verifying that a given random substitution is recognisable is to first show that, in every element of the subshift, the cutting points of inflation tiles are uniquely determined, and then show that each inflation tile has a unique preimage. This is especially useful for constant length random substitutions, since once one cutting point is determined, this determines all cutting points of inflation tiles in a given inflation word decomposition.

Example 2.2.15. Let $p_1, p_2 \in (0,1)$ and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} abbaa & \text{with probability } p_1, \\ aabba & \text{with probability } 1 - p_1, \end{cases} \\ b \mapsto \begin{cases} babaa & \text{with probability } p_2, \\ baaba & \text{with probability } 1 - p_2. \end{cases} \end{cases}$$

Observe that ab appears as a subword in every element of X_ϑ. Since, for every $x \in X_\vartheta$, there exists a $y \in X_\vartheta$ and $j \in \{0, \ldots, |\vartheta(y_1)|\}$ such that $x \in S^j(\vartheta(y))$, it follows that every $x \in X_\vartheta$ contains the word $v = bbuaba$ as a subword, for some (possibly empty) legal word u that does not contain bb or aba as a subword. Since u does not contain aba as a subword, we deduce that the occurrence of aba in v must lie on the overlap of an a and a b supertile. Since $\vartheta_{\mathbf{P}}$ is constant length, this determines the positions of all cutting points of inflation tiles and, as all exact inflation words are distinct, the inflation word decomposition of x is uniquely determined. Hence, we conclude that $\vartheta_{\mathbf{P}}$ is recognisable.

In the following, we present an example of a non-constant length random substitution that is recognisable.

Example 2.2.16. Let $p \in (0,1)$ and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} abb & \text{with probability } p, \\ bab & \text{with probability } 1 - p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1. \end{cases}$$

We show that $\vartheta_{\mathbf{P}}$ is recognisable. Observe that, for each element in X_ϑ, if one cutting point between inflation tiles is known, then all can be determined by reading from left to right, or right to left, since the first two and last two letters of every realisation are distinct from each other. Further, note that no letter can appear more than three times consecutively. We deduce recognisability by showing that the cutting points of inflation tiles are determined by the repeated occurrences of letters. First, observe that any occurrence of *bbb* must lie on the overlap of two a supertiles, and that *abba* can only occur on the overlap of an a and a b supertile, or two a supertiles, and reading the letter immediately to the right of the occurrence of *abba* determines which case we are in. Thus, any word that contains multiple occurrences of b has a unique decomposition into inflation words. Since every element of the subshift X_ϑ admits an occurrence of bb, we thus deduce that every element has a unique decomposition into inflation words. Hence, $\vartheta_{\mathbf{P}}$ is recognisable.

Like many of the properties we have encountered before, recognisability does not depend on the choice of probabilities, and could equivalently be defined in terms of the underlying set-valued substitution. An alternative characterisation of recognisability is the following *local* version. Intuitively, local recognisability means that applying a *finite window* to a sequence is enough to determine the position and the type of the inflation word in the middle of that window.

Lemma 2.2.17. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution over an alphabet $\mathcal{A}$, and suppose that $|\vartheta(a)|$ is well-defined for all $a \in \mathcal{A}$. If $\vartheta_{\mathbf{P}}$ is recognisable, then there exists a smallest natural number $\kappa(\vartheta)$, called the *recognisability radius of* ϑ, with the following property. If $x \in \vartheta([a])$ for some $a \in \mathcal{A}$ and $x_{[-\kappa(\vartheta),\kappa(\vartheta)]} = y_{[-\kappa(\vartheta),\kappa(\vartheta)]}$ for some $y \in X_\vartheta$, then $y \in \vartheta([a])$.

Proof. By way of contradiction, suppose there is no radius of recognisability. In which case, there exists a sequence of tuples $((x^{(k)}, y^{(k)}))_{k\in\mathbb{N}}$ with $(x^{(k)}, y^{(k)}) \in \vartheta([a]) \times (X_\vartheta \setminus \vartheta([a]))$ and $x^{(k)}_{[-k,k]} = y^{(k)}_{[-k,k]}$ for all $k \in \mathbb{N}$. Let $(x, y) \in X_\vartheta \times X_\vartheta$ be an accumulation point of this sequence. By recognisability,

$$X_\vartheta = \bigsqcup_{b\in\mathcal{A}} \bigsqcup_{k=0}^{|\vartheta(b)|-1} S^k(\vartheta([b])),$$

and by construction, $x = y$. Due to Lemma 2.1.6, and since S is continuous, we have that $S^k(\vartheta([b]))$ is compact for all $b \in \mathcal{A}$ and $k \in \mathbb{Z}$. Hence, both $\vartheta([a])$ and $X_\vartheta \setminus \vartheta([a])$ are compact. It therefore follows that $x \in \vartheta([a])$ and $x = y \in X_\vartheta \setminus \vartheta([a])$, leading to a contradiction. $\square$

As a consequence of the local characterisation of recognisability, for every legal word u with length greater than twice the radius of recognisability there exists an inflation word w, appearing as a subword of u, that has a unique decomposition into exact inflation words. We call the largest such w the *recognisable core* of u.

Every recognisable random substitution satisfies the disjoint set condition. However, in contrast to the disjoint set condition, recognisability is preserved under taking powers. Moreover, the recognisability radius of $\vartheta_{\mathbf{P}}^m$ grows (asymptotically) at most with the inflation factor as m increases.

Lemma 2.2.18. If a random substitution $\vartheta_{\mathbf{P}}$ is recognisable, then it satisfies the disjoint set condition.

Proof. By way of contradiction, suppose that $\vartheta_{\mathbf{P}}$ does not satisfy the disjoint set condition. In which case, there exist $a \in \mathcal{A}$, and s and $t \in \vartheta(a)$ with $s \neq t$ and $\vartheta(s) \cap \vartheta(t) \neq \varnothing$. For $x \in [a]$, observe that there exist y and $z \in \vartheta(x)$ such that $y_{[0,|\vartheta(a)|-1]} = s$, $z_{[0,|\vartheta(a)|-1]} = t$, and y coincides with z at all other positions. Hence, there exists a $w \in \vartheta(y) \cap \vartheta(z)$ that can be constructed by mapping s and t to the same word $v \in \vartheta(s) \cap \vartheta(t)$. This is a contradiction to recognisability, so we conclude that $\vartheta_{\mathbf{P}}$ must satisfy the disjoint set condition. $\square$

Lemma 2.2.19. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution for which $|\vartheta(a)|$ is well-defined for all $a \in \mathcal{A}$. If $\vartheta_{\mathbf{P}}$ is recognisable, then $\vartheta_{\mathbf{P}}^k$ is recognisable for all $k \in \mathbb{N}$. Moreover, if $\vartheta_{\mathbf{P}}$ is of constant length ℓ, then for all $m \in \mathbb{N}$ we have that

$$\kappa(\vartheta^m) \leq \frac{\ell^m - 1}{\ell - 1} \kappa(\vartheta).$$

Proof. Let $x \in X_\vartheta$ and assume that there exist $y, y' \in X_\vartheta$ and $i \in \{0, \ldots, |\vartheta^k(y_0)| - 1\}$, $j \in \{0, \ldots, |\vartheta^k(y_0')| - 1\}$ such that $x \in S^i(\vartheta^k(y))$ and $x \in S^j(\vartheta^k(y'))$. We show that $y = y'$ and $i = j$. By definition, there exist elements $w_0 = x, w_1, \ldots, w_{k-1}, w_k = y$, $z_0 = x, z_1, \ldots z_{k-1}, z_k = y'$ and integers $i_r \in \{0, \ldots, |\vartheta(w_0^r)| - 1\}$, $j_r \in \{0, \ldots, |\vartheta(z_0^r)| - 1\}$ such that $w^r \in S^{i_r}(\vartheta(w^{r-1}))$ and $w^r \in S^{j_r}(\vartheta(z^{r-1}))$ for all $r \in \{1, \ldots, k\}$. Since $\vartheta_{\mathbf{P}}$ is recognisable, we must have that $i_1 = j_1$ and $w_1 = z_1$. Then, it follows inductively that $i_r = j_r$ and $w_r = z_r$ for all $r \in \{1, \ldots, k\}$. In particular, we must have that $y = y'$ and $i = j$. Since $x \in X_\vartheta$ was arbitrary, we thus conclude that $\vartheta_{\mathbf{P}}^k$ is recognisable.

Now assume that $\vartheta_{\mathbf{P}}$ is of constant length ℓ. We prove the bounds on the recognisability radius by induction. The claim is immediate for $m = 1$. Assume it holds for some $m \in \mathbb{N}$, and note, by primitivity, that $X_\vartheta = X_{\vartheta^m}$. Let $a \in \mathcal{A}$, $x \in \vartheta^{m+1}([a])$ and $y \in X_\vartheta$ with $x_{[-k,k]} = y_{[-k,k]}$ for $k = \ell\kappa(\vartheta^m) + \kappa(\vartheta)$; in particular, $y \in \vartheta(X_\vartheta)$. Let $v \in \vartheta^m([a])$ be such $x \in \vartheta(v)$, and let

$w \in X_\vartheta$ such that $y \in \vartheta(w)$. Applying the local characterisation of recognisability given by Lemma 2.2.17 to the pair $(S^{j\ell}x, S^{j\ell}y)$ for each $j \in \{-\kappa(\vartheta^m), \dots, \kappa(\vartheta^m)\}$, in combination with Lemma 2.2.18, we obtain that $v_{[-\kappa(\vartheta^m),\kappa(\vartheta^m)]} = w_{[-\kappa(\vartheta^m),\kappa(\vartheta^m)]}$. By the definition of $\kappa(\vartheta^m)$, this implies $w \in \vartheta^m([a])$ and so $y \in \vartheta(w) \subseteq \vartheta^{m+1}([a])$, yielding

$$\kappa(\vartheta^{m+1}) \leq \ell\kappa(\vartheta^m) + \kappa(\vartheta) = \kappa(\vartheta)\sum_{j=0}^{m} \ell^j = \frac{\ell^m - 1}{\ell - 1}\kappa(\vartheta),$$

where the second to last equality follows from the inductive hypothesis. □

While every recognisable substitution satisfies the disjoint set condition, the reverse implication does not hold. A counterexample is given by the random period doubling substitution, which satisfies the disjoint set condition but is not recognisable.

Example 2.2.20. Let $p \in (0,1)$ and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random period doubling substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1. \end{cases}$$

We show that the subshift X_ϑ admits elements that do not have a unique inflation word decomposition. For example, let $x = (aab)^{\mathbb{Z}}$. Since, for all $k \in \mathbb{N}$, we have $b(aab)^{(2^{2k}-1)/3} \in \vartheta^{2k}(b)$, it follows that $x \in X_\vartheta$. We note that $(2^{2k}-1)/3$ is always an integer since $2^{2k}-1 = (2^k-1)(2^k+1)$, and one of these factors must be divisible by 3. Observe that both

$$\cdots \underbrace{aa}_{b}\underbrace{ba}_{a}\underbrace{ab}_{a} \cdot \underbrace{aa}_{b}\underbrace{ba}_{a}\underbrace{ab}_{a} \cdots \quad \text{and} \quad \cdots a\underbrace{ab}_{a}\underbrace{aa}_{b}\underbrace{b \cdot a}_{a}\underbrace{ab}_{a}\underbrace{aa}_{b}b\cdots,$$

are valid inflation word decompositions of x. Hence, x does not have a unique inflation word decomposition and so $\vartheta_{\mathbf{P}}$ is not recognisable.

2.3 Frequency measures

If θ is a primitive deterministic substitution, then the subshift X_θ supports a unique shift-invariant measure μ. For each $v \in \mathcal{L}_\theta$, μ assigns to the cylinder set $[v]$ mass equal to the word frequency of

v under repeated iteration of the substitution; namely,

$$\mu([v]) = \lim_{k\to\infty} \frac{|\theta^k(a)|_v}{|\theta^k(a)|}.$$

By primitivity, this limit always exists and is independent of the choice of a. For random substitutions, the associated subshift is typically not uniquely ergodic. However, a shift-invariant measure can be constructed in a similar manner. In particular, for a given random substitution $\vartheta_{\mathbf{P}}$, there exists a shift-invariant measure $\mu_{\mathbf{P}}$ that is compatible with the word frequencies that arise almost surely in the limit of large inflation words. Moreover, it was shown by Gohlke and Spindeler [39] that this measure is ergodic with respect to the shift action.

2.3.1 Definition and basic properties

Recall that for every subshift X, the algebra $\xi(X)$ of cylinder sets that specify the origin, together with the empty set, generates the Borel sigma-algebra $\mathcal{B}(X)$. Frequency measures assign to a given cylinder the expected frequency of the word under repeated substitution.

Definition 2.3.1. Given a primitive random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$, we define the *expected frequency* of a word $v \in \mathcal{L}_\vartheta$ by

$$\mathrm{freq}(v) = \lim_{k\to\infty} \frac{\mathbb{E}[|\vartheta_{\mathbf{P}}^k(a)|_v]}{\mathbb{E}[|\vartheta_{\mathbf{P}}^k(a)|]}. \tag{2.1}$$

We note that, by primitivity, the limit in (2.1) always exists and is independent of the choice of a. In fact, we have the stronger property that the word frequencies exist $\mathbb{P}$-almost surely in the limit of large inflation words and are given by $\mathrm{freq}(v)$ for all $v \in \mathcal{L}_\vartheta$ (see [39] for further details).

Definition 2.3.2. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution and define $\mu_{\mathbf{P}} \colon \xi(X_\vartheta) \to [0,1]$ by $\mu_{\mathbf{P}}(\varnothing) = 0$, $\mu_{\mathbf{P}}(X_\vartheta) = 1$ and $\mu_{\mathbf{P}}([v]_m) = \mathrm{freq}(v)$ for all $v \in \mathcal{L}_\vartheta$ and $m \in \{1-|v|, 2-|v|, \dots, 0\}$. The set function $\mu_{\mathbf{P}}$ defines a pre-measure on the algebra $\xi(X_\vartheta)$, which extends uniquely to a measure on $\mathcal{B}(X_\vartheta)$. We call the measure $\mu_{\mathbf{P}}$ the *frequency measure* corresponding to the random substitution $\vartheta_{\mathbf{P}}$.

Proposition 2.3.3 ([39, Proposition 5.3 and Theorem 5.9])**.** Let $\vartheta_{\mathbf{P}}$ be a primitive random substitution. Then, the corresponding frequency measure is a shift-invariant and ergodic probability measure on the subshift X_ϑ.

Since $\mu_{\mathbf{P}}$ is ergodic, it follows by Birkhoff's ergodic theorem that for every $u \in \mathcal{L}_\vartheta$, we have

$$\frac{1}{2n+1} \# \left\{ i \in \{-n, \ldots, n\} \colon x_{[i,i+|u|-1]} = u \right\} \xrightarrow{n \to \infty} \mu_{\mathbf{P}}([u])$$

for $\mu_{\mathbf{P}}$-almost all $x \in X_\vartheta$. This motivates the term *frequency measure*.

For all $a \in \mathcal{A}$, we have that $\mu_{\mathbf{P}}([a]) = R_a$, the entry of the right Perron–Frobenius eigenvector corresponding to a. It is also possible to define the measures of longer words in a similar manner, as the entries of the right Perron–Frobenius eigenvector of an associated induced random substitution. However, Definition 2.3.2 is sufficient for our purposes so we do not include this alternative characterisation here. For more details, we refer the reader to [39, 70].

Note that frequency measures are dependent on the probabilities of the random substitution. As such, primitive random substitution subshifts often support uncountably many ergodic measures. This is in contrast to subshifts of primitive deterministic substitutions, which are always uniquely ergodic [52, 67].

2.3.2 Renormalisation lemma

The presence of an inherent hierarchical structure allows for the application of renormalisation techniques in the study of random substitutions. Since the frequency measure corresponding to a primitive random substitution reflects the underlying Markov process, it is natural to expect this measure will exhibit some kind of dynamical self-similarity. This is given by the following self-consistency relation, which was shown as the first step in the proof of [39, Prop. 5.8].

Lemma 2.3.4. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution, with corresponding frequency measure $\mu_{\mathbf{P}}$. Then, for all $n \in \mathbb{N}$ and $u \in \mathcal{L}_\vartheta^n$, we have

$$\mu_{\mathbf{P}}([u]) = \frac{1}{\lambda} \sum_{v \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([v]) \sum_{m=1}^{|\vartheta|} \sum_{j=1}^{m} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u \text{ and } |\vartheta_{\mathbf{P}}(v_1)| = m].$$

Lemma 2.3.4 relates the measures of cylinder sets via the action of the random substitution. This is a very powerful tool for studying frequency measures, and is fundamental to the proofs of our main results on measure theoretic entropy and multifractal properties of frequency measures in Chapters 4–6.

If additional assumptions are imposed on the random substitution, then simpler formulations of

Lemma 2.3.4 can be obtained. For example, if the inflated image of every letter has a well-defined length, then we have the following.

Lemma 2.3.5. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Additionally assume that, for all $a \in \mathcal{A}$, the length $|\vartheta(a)|$ is well-defined. Fix $n \in \mathbb{N}$ and let k be an integer such that every $v \in \mathcal{L}_{\vartheta}^{k}$ has $|\vartheta(v)| \geq n + |\vartheta(v_1)|$. Then, for every $u \in \mathcal{L}_{\vartheta}^{n}$,

$$\mu_{\mathbf{P}}([u]) = \frac{1}{\lambda} \sum_{v \in \mathcal{L}_{\vartheta}^{k}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{|\vartheta(v_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+m-1]} = u].$$

Proof. For all $v \in \mathcal{L}_{\vartheta}^{k}$, the length $|\vartheta(v_1)|$ is well-defined, so for all $u \in \mathcal{L}_{\vartheta}^{n}$ and $j \in \{1, \ldots, |\vartheta(v_1)|\}$, we have

$$\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u \text{ and } |\vartheta_{\mathbf{P}}(v_1)| = m] = \begin{cases} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u] & \text{if } m = |\vartheta(v_1)|, \\ 0 \quad \text{if } m \neq |\vartheta(v_1)|. \end{cases}$$

Hence, it follows by Lemma 2.3.4 that

$$\mu_{\mathbf{P}}([u]) = \frac{1}{\lambda} \sum_{v \in \mathcal{L}_{\vartheta}^{k}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{|\vartheta(v_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+m-1]} = u],$$

which completes the proof. □

We note that if $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is compatible or constant length, then the condition that $|\vartheta(a)|$ is well-defined for all $a \in \mathcal{A}$ is satisfied.

2.3.3 Frequency measures on the full shift

We conclude this chapter by providing conditions under which a class of frequency measures on the full shift are Bernoulli measures. We highlight the following is a new result which has not been presented in any of the papers on which this thesis is based.

Proposition 2.3.6. Given $p_1, p_2, p_3, p_4 \in (0, 1)$ such that $\sum_{i=1}^{4} p_i = 1$, let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon a, b \mapsto \begin{cases} aa & \text{with probability } p_1, \\ ab & \text{with probability } p_2, \\ ba & \text{with probability } p_3, \\ bb & \text{with probability } p_4, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Then, the associated subshift X_ϑ is the full shift $\{a,b\}^{\mathbb{Z}}$. Further, the measure $\mu_{\mathbf{P}}$ is Bernoulli if and only if there exists a $p \in (0,1)$ such that $p_1 = p^2$, $p_2 = p_3 = p(1-p)$ and $p_4 = (1-p)^2$. In this case, $\mu_{\mathbf{p}}$ is the $(p, 1-p)$-Bernoulli measure on $\{a,b\}^{\mathbb{Z}}$.

Proof. For each $k \in \mathbb{N}$, we have that $\vartheta^k(a) = \vartheta^k(b)$ is the set of all words of length 2^k over the alphabet $\{a,b\}$, so it follows by the definition of the subshift associated to a random substitution that X_ϑ is the full shift $\{a,b\}^{\mathbb{Z}}$.

To prove the necessary and sufficient conditions under which the frequency measure $\mu_{\mathbf{P}}$ is a Bernoulli measure, we use the renormalisation lemma stated in Section 2.3.2. In particular, since $\vartheta_{\mathbf{P}}$ is constant length, we can apply the version stated in Lemma 2.3.5. For each $n \in \mathbb{N}$ and $u \in \mathcal{L}_\vartheta^n$, Lemma 2.3.5 gives that

$$\mu_{\mathbf{P}}([u]) = \frac{1}{2} \sum_{v \in \mathcal{L}_\vartheta^{\lfloor n/2 \rfloor + 1}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{2} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j, |u|+j-1]} = u].$$

For all $v^1, v^2 \in \mathcal{L}_\vartheta^{\lfloor n/2 \rfloor + 1}$ and $j \in \{1,2\}$ we have

$$\mathbb{P}[\vartheta_{\mathbf{P}}(v^1)_{[j,|u|+j-1]} = u] = \mathbb{P}[\vartheta_{\mathbf{P}}(v^2)_{[j,|u|+j-1]} = u],$$

so the above expression simplifies to

$$\mu_{\mathbf{P}}([u]) = \frac{1}{2} \left(\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[1,|u|]} = u] + \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[2,|u|+1]} = u] \right) \tag{2.2}$$

for any choice of $v \in \mathcal{L}_\vartheta^{\lfloor n/2 \rfloor + 1}$, noting that $\sum_{v \in \mathcal{L}_\vartheta^{\lfloor n/2 \rfloor + 1}} \mu_{\mathbf{P}}([v]) = 1$.

We first show that when $p_1 = p^2$, $p_2 = p_3 = p(1-p)$ and $p_4 = (1-p)^2$, the frequency measure

$\mu_{\mathbf{P}}$ is the $(p, 1-p)$-Bernoulli measure. Observe that for every $w \in \vartheta(a) = \vartheta(b)$, we have

$$\mathbb{P}[\vartheta_{\mathbf{P}}(a) = w] = \mathbb{P}[\vartheta_{\mathbf{P}}(b) = w] = p^{|w|_a}(1-p)^{|w|_b}$$

and, if $j \in \{1, 2\}$ and $c \in \{a, b\}$, then

$$\mathbb{P}[\vartheta_{\mathbf{P}}(c)_j = a] = p \quad \text{and} \quad \mathbb{P}[\vartheta_{\mathbf{P}}(c)_j = b] = 1 - p.$$

Hence, splitting each of the probabilities in (2.2) into inflation tiles, we obtain

$$\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[1,|u|]} = u] = \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[2,|u|]+1} = u] = p^{|u|_a}(1-p)^{|u|_b},$$

so it follows by (2.2) that $\mu_{\mathbf{P}}([u]) = p^{|u|_a}(1-p)^{|u|_b}$. Since this is true for all $u \in \mathcal{L}_\vartheta = \{a, b\}^*$, we conclude that $\mu_{\mathbf{P}}$ is the $(p, 1-p)$-Bernoulli measure.

Now, we show that if $\mu_{\mathbf{P}}$ is a Bernoulli measure, then there exists a $p \in (0, 1)$ such that $p_1 = p^2$, $p_2 = p_3 = p(1-p)$ and $p_4 = (1-p)^2$. If $\mu_{\mathbf{P}}$ is a Bernoulli measure, then $\mu_{\mathbf{P}}([u]) = \prod_{i=1}^{|u|} \mu_{\mathbf{P}}([u_i])$ for all $u \in \mathcal{L}_\vartheta$. Thus, we require $\mu_{\mathbf{P}}([ab]) = \mu_{\mathbf{P}}([ba])$. By Equation (2.2), we have

$$\begin{aligned}
\mu_{\mathbf{P}}([ab]) &= \frac{1}{2}\left(p_2 + (p_1 + p_3)(p_3 + p_4)\right) \\
\mu_{\mathbf{P}}([ba]) &= \frac{1}{2}\left(p_3 + (p_1 + p_2)(p_2 + p_4)\right),
\end{aligned}$$

and these two quantities coincide if and only if $p_2 = p_3$. Next, note that the right Perron–Frobenius eigenvector of $\vartheta_{\mathbf{P}}$ is given by

$$\mathbf{R} = \left(\frac{1 + p_1 - p_4}{2}, \frac{1 - p_1 + p_4}{2}\right)^\top,$$

so

$$\mu_{\mathbf{P}}([a]) = \frac{1 + p_1 - p_4}{2} = p_1 + p_2,$$

where in the second equality we have used that $p_4 = 1 - p_1 - p_2 - p_3$ and $p_2 = p_3$. Thus, by (2.2), if $\mu_{\mathbf{P}}$ is a Bernoulli measure, we require

$$(p_1 + p_2)^2 = \mu_{\mathbf{P}}([a])\mu_{\mathbf{P}}([a]) = \mu_{\mathbf{P}}([aa]) = \frac{1}{2}(p_1 + (p_1 + p_2)^2).$$

In other words, we require $(p_1 + p_2)^2 = p_1$. If $p \in (0, 1)$ is the real number such that $p_1 = p^2$, then this identity holds if and only if $p_2 = p(1 - p)$. Since $p_3 = p_2$, we also require $p_3 = p(1 - p)$. Finally, since $\sum_{i=1}^{4} p_i = 1$, we have $p_4 = 1 - p^2 - 2p(1 - p) = (1 - p)^2$. This completes the proof. $\square$

CHAPTER 3

TOPOLOGICAL ENTROPY AND WORD COMPLEXITY

In contrast to subshifts of deterministic substitutions, random substitution subshifts often have positive topological entropy. As such, topological entropy provides a new invariant in the study of random substitution subshifts not available for their deterministic counterparts.

Positive topological entropy was first identified for subshifts of random substitutions in the pioneering work of Godrèche and Luck [34], who showed that the random Fibonacci substitution gives rise to a subshift with positive topological entropy. Following this, the topological entropy was calculated for several families of random substitution subshifts, for example, see [46, 56, 58, 59, 74, 76]. In each of these references, the topological entropy was calculated by first quantifying the asymptotic growth rate of inflation words and then showing the topological entropy coincides with this quantity. This approach was unified by Gohlke [35], who showed that for every primitive and compatible random substitution, the topological entropy of the associated subshift coincides with the notion of *inflation word entropy*, which is characterised entirely in terms of the substitution branching process, as opposed to the subshift. Further, Gohlke showed that if the underlying random substitution satisfies either the identical or disjoint set condition, then a closed-form formula for the topological entropy can be obtained.

In this chapter, we continue to develop the theory of topological entropy for random substitution subshifts. In Section 3.1, we provide necessary and sufficient conditions under which a primitive random substitution gives rise to a subshift of positive topological entropy. In particular, we show that a primitive random substitution subshift has zero topological entropy if and only if it can be obtained as the subshift of a deterministic substitution, answering in the positive an open question of Rust and Spindeler [70]. Then, in Section 3.2, we develop techniques for calculating topological entropy for subshifts of constant length random substitutions, without the requirement of compatibility. This allows the topological entropy to be calculated for a broad class of non-compatible random substitution subshifts. Further, we obtain general bounds on the topological entropy of constant length random substitutions that hold even without primitivity. In contrast to the primitive setting, non-primitive random substitutions can give rise to subshifts that cannot be obtained as the subshift of a deterministic substitution. In Section 3.3, we show that these subshifts exhibit a rich variety of complexity behaviour not witnessed in the deterministic or primitive random settings.

The results in this chapter are largely based on the solo-authored paper [54].

Notation. Throughout this chapter, we write ϑ for a random substitution instead of $\vartheta_{\mathbf{P}}$, to reflect the fact that the topological entropy of the associated subshift does not depend on the choice of probabilities. We note that all the results in this chapter can be framed in terms of set-valued substitutions. However, we still refer to ϑ as a random substitution to keep in line with the rest of the thesis.

3.1 Positivity of topological entropy for random substitution subshifts

3.1.1 Statement of main results

While random substitution subshifts typically have positive topological entropy, every deterministic substitution subshift is itself the subshift of a random substitution, so it is not true that all random substitution subshifts have positive topological entropy. Rust and Spindeler [70] conjectured that, under primitivity, the zero entropy random substitution subshifts are precisely those that can be obtained as the subshift of a deterministic substitution. Here, we provide a proof of this conjecture.

Theorem 3.1.1. Let X be the subshift of a primitive random substitution. Then $h_{\mathrm{top}}(X) = 0$ if and only if X is a deterministic substitution subshift.

We give the proof of Theorem 3.1.1 in Section 3.1.2. By combining Theorem 3.1.1 with the classification of complexity functions for subshifts of primitive deterministic substitutions given by Pansiot (Proposition 1.6.1), we obtain the following. In particular, we highlight that there do not exist primitive random substitution subshifts with intermediate growth complexity function.

Corollary 3.1.2. If ϑ is a primitive random substitution such that $h_{\mathrm{top}}(X_\vartheta) = 0$, then the complexity function of X_ϑ is either $\Theta(1)$ or $\Theta(n)$.

Corollary 3.1.2 illustrates that there exists a *complexity gap* for subshifts of primitive random substitutions. In particular, any function that grows faster than linearly but sub-exponentially cannot be obtained as the complexity function of a primitive random substitution subshift. On the other hand, there is no exponential complexity gap, as Gohlke, Rust and Spindeler [38, Theorem 42] showed that the set $\{h_{\mathrm{top}}(X_\vartheta) \colon \vartheta \text{ is a primitive random substitution}\}$ is dense in $[0, \infty)$.

Without primitivity, the conclusion of Theorem 3.1.1 does not hold. In Section 3.3, we show that non-primitive random substitutions give rise to subshifts whose complexity functions exhibit a rich variety of behaviour not witnessed for subshifts of deterministic or primitive random substitutions. In particular, we show that there exist zero entropy subshifts of non-primitive random substitutions that have intermediate growth complexity function, as well as ones with polynomial growth not possible for deterministic substitutions. A key feature of these examples is the existence of letters that occur with frequency zero, which is not possible under primitivity.

3.1.2 Proof of Theorem 3.1.1

We now turn towards the proof of Theorem 3.1.1. Our proof uses the notion of a *splitting pair* for a random substitution introduced by Rust and Spindeler [70]. There, it was shown that this notion provides a sufficient condition for positive entropy of the associated subshift.

Definition 3.1.3. Given $u, v \in \mathcal{A}^+$, we say that u is an *affix* of v if u is either a prefix or a suffix of v, namely, if $u = v_{[1,|u|]}$ or $u = v_{[|v|-|u|+1,|v|]}$. If u is both a prefix and a suffix of v, then we call u a *strong affix* of v.

Definition 3.1.4. If ϑ is a random substitution and $a \in \mathcal{A}$ is such that there exist realisations $u, v \in \vartheta(a)$ with $|u| \leq |v|$ for which u is not a strong affix of v, then we say that *a admits a splitting pair for ϑ*.

Proposition 3.1.5 ([70, Corollary 34])**.** If ϑ is a primitive random substitution for which there exists a letter that admits a splitting pair, then $h_{\mathrm{top}}(X_\vartheta) > 0$.

If ϑ is a primitive random substitution, then for all $m \in \mathbb{N}$ the random substitution ϑ^m gives rise to the same subshift as ϑ. Hence, to ascertain that the subshift X_ϑ has positive topological entropy, it is sufficient to verify that there exists a positive integer m such that ϑ^m admits a splitting pair. Thus, to obtain Theorem 3.1.1, it suffices to show that every primitive random substitution for which no power admits a splitting pair can be obtained as the subshift of a deterministic substitution. In particular, we show that any such random substitution gives rise to the same subshift as one of its marginals.

Proof of Theorem 3.1.1. Let ϑ be a primitive random substitution with $h_{\mathrm{top}}(X_\vartheta) = 0$. It follows by Proposition 3.1.5 that no letter in $\mathcal{A}$ admits a splitting pair for any power of ϑ. Thus, for every $m \in \mathbb{N}$ and $a \in \mathcal{A}$, there exists a unique realisation of $\vartheta^m(a)$ of greatest length, since if $u, v \in \vartheta^m(a)$ have $|u| = |v|$ and u is a strong affix of v, then we require $u = v$. Let θ be the marginal of ϑ that maps each letter $a \in \mathcal{A}$ to the realisation of $\vartheta(a)$ of greatest length. By construction, for each $a \in \mathcal{A}$ and $m \in \mathbb{N}$, $\theta^m(a)$ is the realisation of $\vartheta^m(a)$ of greatest length. Since, for every $a \in \mathcal{A}$, a does not admit a splitting pair for ϑ, every realisation of $\vartheta(a)$ appears as a subword of $\theta(a)$, so primitivity transfers to θ. Now, let $u \in \mathcal{L}(X_\vartheta)$. Since ϑ is primitive, we have $\mathcal{L}(X_\vartheta) = \mathcal{L}_\vartheta$, so there exist $m \in \mathbb{N}$, $a \in \mathcal{A}$ and $v \in \vartheta^m(a)$ such that u appears as a subword of v. But every realisation of $\vartheta^m(a)$ appears as a subword of $\theta^m(a)$, so $u \in \mathcal{L}_\theta$. Primitivity of θ gives that $\mathcal{L}(X_\theta) = \mathcal{L}_\theta$, so we have that $u \in \mathcal{L}(X_\vartheta)$. Hence, $\mathcal{L}(X_\vartheta) = \mathcal{L}(X_\theta)$ and so we conclude that $X_\vartheta = X_\theta$. □

While Theorem 3.1.1 guarantees that every zero entropy primitive random substitution subshift can be obtained as the subshift of a deterministic substitution, it does not guarantee that *every* random substitution that gives rise to that subshift must be deterministic. For example, the primitive random substitution $\vartheta\colon a \mapsto \{a, aba\}$, $b \mapsto \{bab\}$ gives rise to the finite subshift $X_\vartheta = \{(ab)^{\mathbb{Z}}, (ba)^{\mathbb{Z}}\}$. However, if a primitive random substitution is additionally assumed to have unique realisation paths (recall Definition 2.2.7), then its subshift has zero topological entropy if and only if it is itself deterministic.

Proposition 3.1.6. If ϑ is a primitive random substitution with unique realisation paths and there exists a letter $b \in \mathcal{A}$ such that $\#\vartheta(b) \geq 2$, then $h_{\mathrm{top}}(X_\vartheta) > 0$.

Proposition 3.1.6 is a consequence of a lower bound on measure theoretic entropy that we prove in Chapter 4. In particular, there exists a frequency measure μ, supported on the subshift

X_ϑ, such that X_ϑ has positive measure theoretic entropy with respect to the measure μ. Positivity of topological entropy then follows from the fact that measure theoretic entropy provides a lower bound for topological entropy. We refer the reader to Proposition 4.1.5 for the exact details.

3.2 Inflation word entropy

A systematic approach to calculating topological entropy for subshifts of primitive and compatible random substitutions was provided by Gohlke [35]. There, it was shown that topological entropy coincides with the notion of inflation word entropy, which is defined in terms of the underlying random substitution.

Definition 3.2.1. Let ϑ be a random substitution such that the length $|\vartheta^m(a)|$ is well-defined for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$. The lower and upper inflation word entropy of type a, respectively, are defined by

$$\underline{h}_a = \liminf_{m\to\infty} \frac{1}{|\vartheta^m(a)|} \log(\#\vartheta^m(a))$$
$$\overline{h}_a = \limsup_{m\to\infty} \frac{1}{|\vartheta^m(a)|} \log(\#\vartheta^m(a)).$$

When these limits coincide, we denote their common value by h_a, which we call the *inflation word entropy* of type a.

Note that the length $|\vartheta^m(a)|$ is well-defined for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$ if and only if ϑ is compatible or of constant length.

3.2.1 Compatible random substitutions

For primitive and compatible random substitutions, Gohlke [35] showed that the inflation word entropy h_a exists and coincides with the topological entropy for all $a \in \mathcal{A}$. In this section, we summarise the main results from [35] and show how these can be applied to obtain the topological entropy for some examples.

Proposition 3.2.2 ([35, Theorem 17])**.** Let ϑ be a primitive and compatible random substitution with Perron–Frobenius eigenvalue λ and right eigenvector $\mathbf{R}$. Then, for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$, the following bounds hold:

$$\frac{1}{\lambda^m}\sum_{a\in\mathcal{A}} R_a \log(\#\vartheta^m(a)) \leq \underline{h}_a \leq \overline{h}_a \leq h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\lambda^m - 1}\sum_{a\in\mathcal{A}} R_a \log(\#\vartheta^m(a)).$$

Moreover, the lower bound is non-decreasing in m. In particular, the inflation word entropy h_a is well-defined and independent of $a \in \mathcal{A}$, and the topological entropy can be calculated via

$$h_{\mathrm{top}}(X_\vartheta) = h_a = \lim_{m\to\infty}\frac{1}{\lambda^m}\sum_{a\in\mathcal{A}} R_a \log(\#\vartheta^m(a)) = \sup_{m\in\mathbb{N}}\frac{1}{\lambda^m}\sum_{a\in\mathcal{A}} R_a \log(\#\vartheta^m(a)).$$

If a primitive and compatible random substitution ϑ satisfies the identical or disjoint set condition, then a closed-form formula for the topological entropy of the associated subshift can be obtained.

Proposition 3.2.3 ([35, Corollary 18])**.** Let ϑ be a primitive and compatible random substitution. If ϑ satisfies the identical set condition, then

$$h_{\mathrm{top}}(X_\vartheta) = \frac{1}{\lambda}\sum_{a\in\mathcal{A}} R_a \log(\#\vartheta(a)).$$

If ϑ satisfies the disjoint set condition, then

$$h_{\mathrm{top}}(X_\vartheta) = \frac{1}{\lambda - 1}\sum_{a\in\mathcal{A}} R_a \log(\#\vartheta(a)).$$

Proposition 3.2.3 allows the exact value of the topological entropy to be obtained for a broad class of random substitution subshifts. We now present several examples where this is possible. Our first example satisfies the identical set condition and our second satisfies the disjoint set condition.

Example 3.2.4. Let ϑ be the random substitution defined by

$$\vartheta\colon a, b \mapsto \begin{cases} ab, \\ ba. \end{cases}$$

Since ϑ satisfies the identical set condition and has Perron–Frobenius eigenvalue $\lambda = 2$ and right eigenvector $\mathbf{R} = (1/2, 1/2)^\top$, by Proposition 3.2.3 we have

$$h_{\text{top}}(X_\vartheta) = \frac{1}{\lambda}\left(R_a \log(\#\vartheta(a)) + R_b \log(\#\vartheta(b))\right) = \frac{1}{2}\log 2.$$

Example 3.2.5. Let ϑ be the random period doubling substitution

$$\vartheta\colon \begin{cases} a \mapsto \begin{cases} ab, \\ ba, \end{cases} \\ b \mapsto aa. \end{cases}$$

We have previously shown that ϑ satisfies the disjoint set condition and has Perron–Frobenius eigenvalues $\lambda = 2$ and right eigenvector $\mathbf{R} = (2/3, 1/3)^\top$. Thus, it follows by Proposition 3.2.3 that

$$h_{\text{top}}(X_\vartheta) = \frac{1}{\lambda - 1}\left(R_a \log(\#\vartheta(a)) + R_b \log(\#\vartheta(b))\right) = \frac{2}{3}\log 2.$$

The following example does not satisfy the identical set condition or disjoint set condition, so Proposition 3.2.3 cannot be applied. However, a precise estimate for the topological entropy can be obtained using the bounds given by Proposition 3.2.2.

Example 3.2.6. Let ϑ be the random Thue-Morse substitution defined by

$$\vartheta\colon \begin{cases} a \mapsto \begin{cases} ab, \\ ba, \end{cases} \\ b \mapsto ba. \end{cases}$$

Noting that ϑ has Perron–Frobenius eigenvalue $\lambda = 2$ and right eigenvector $\mathbf{R} = (1/2, 1/2)^\top$, Proposition 3.2.2 gives that

$$\frac{1}{2^{m+1}}\left(\log(\#\vartheta^m(a)) + \log(\#\vartheta^m(b))\right) \leq h_{\text{top}}(X_\vartheta) \leq \frac{1}{2^{m+1} - 2}\left(\log(\#\vartheta^m(a)) + \log(\#\vartheta^m(b))\right).$$

Using these bounds, a computer-assisted calculation of the cardinalities $\#\vartheta^m(a)$ and $\#\vartheta^m(b)$ gives that, to four decimal places, the topological entropy of the subshift X_ϑ is 0.2539.

For some random substitutions, an exact value can be obtained for the topological entropy of the associated subshift even if neither the identical set condition nor the disjoint set condition are satisfied. This is the case for the random Fibonacci substitution, where a closed-form recursive relation for the cardinalities of inflation sets allows the topological entropy of the associated subshift to be expressed as an infinite sum.

Example 3.2.7. Let ϑ be the random Fibonacci substitution

$$\vartheta\colon \begin{cases} a \mapsto \begin{cases} ab, \\ ba, \end{cases} \\ b \mapsto a. \end{cases}$$

It was shown in [34, 58] that

$$h_{\text{top}}(X_\vartheta) = \sum_{n=2}^{\infty} \frac{\log n}{\tau^{n+2}} \approx 0.444399,$$

where τ denotes the golden ratio.

3.2.2 Constant length random substitutions

Proposition 3.2.2 and Proposition 3.2.3 provide a means of calculating the topological entropy of a broad class of random substitution subshifts. However, a limitation is the somewhat restrictive assumption of compatibility. The random substitutions that give rise to many well-known subshifts, such as shifts of finite type [38], are typically not compatible. As such, it is desirable to better understand topological entropy for non-compatible random substitution subshifts. However, there are two main barriers to extending the inflation word entropy approach to the non-compatible setting. Firstly, the lower and upper inflation word entropy are only well-defined for compatible and constant length random substitutions. Secondly, subshifts of non-compatible random substitutions do not have uniform letter frequencies, a property fundamental to the proof of Proposition 3.2.2 in [35]. Here, we develop techniques to circumvent the latter of these issues in the constant length setting. We show that for primitive random substitutions of constant length, the topological entropy of the associated subshift can be calculated via inflation word entropy. This allows the topological entropy to be calculated for a broad class of non-compatible random substitution subshifts, which previously has only been possible for isolated examples.

Theorem 3.2.8. Let ϑ be a primitive random substitution of constant length. Then, for all $a \in \mathcal{A}$, the inflation word entropy h_a exists and coincides with $h_{\mathrm{top}}(X_\vartheta)$.

We present the proof of Theorem 3.2.8 in Section 3.2.3. In general, the conclusion of Theorem 3.2.8 does not hold without primitivity. For example, consider the non-primitive random substitution defined by $\vartheta\colon a \mapsto \{ab, ba\}$, $b \mapsto \{aa\}$, $c \mapsto \{cc\}$. The associated subshift X_ϑ contains the subshift of the random period doubling substitution as a subsystem. In Example 3.2.5, we showed that this subshift has topological entropy $\log(4)/3$, so $h_{\mathrm{top}}(X_\vartheta) \geq \log(4)/3 > 0$. However, since $\#\vartheta^m(c) = 1$ for all $m \in \mathbb{N}$, we have $h_c = 0$.

We now present two examples of non-compatible random substitutions where Theorem 3.2.8 can be applied to obtain the exact value of the topological entropy.

Example 3.2.9. Let ϑ be the primitive random substitution defined by $\vartheta\colon a \mapsto \{aa, bb\}$, $b \mapsto \{aa\}$, and let X_ϑ denote the corresponding subshift. Using Theorem 3.2.8, we show that

$$h_{\mathrm{top}}(X_\vartheta) = \frac{1}{2} \log 2.$$

Observe that $\vartheta^m(b) \subseteq \vartheta^m(a)$ for all $m \in \mathbb{N}$. Hence, by the constant length property, we have

$$\vartheta^m(a) = \vartheta^{m-1}(a)\vartheta^{m-1}(a) \cup \vartheta^{m-1}(b)\vartheta^{m-1}(b) = \vartheta^{m-1}(a)\vartheta^{m-1}(a)$$

for all $m \in \mathbb{N}$, so $\#\vartheta^m(a) = (\#\vartheta^{m-1}(a))^2$. It follows inductively that

$$\frac{1}{2^m} \log(\#\vartheta^m(a)) = \frac{1}{2} \log(\#\vartheta(a)) = \frac{1}{2} \log 2$$

for all m, so we conclude by Theorem 3.2.8 that $h_{\mathrm{top}}(X_\vartheta) = h_a = \log(2)/2$.

Example 3.2.10. Let ϑ be the primitive random substitution defined by $\vartheta\colon a \mapsto \{aa, ab\}$, $b \mapsto \{ba\}$, and let X_ϑ denote the corresponding subshift. The topological entropy of X_ϑ is

$$h_{\mathrm{top}}(X_\vartheta) = \sum_{n=1}^{\infty} \frac{1}{2^n} \log n \approx 0.507834.$$

To see this, note that since ϑ is primitive and constant length, it follows by Theorem 3.2.8 that $h_{\mathrm{top}}(X_\vartheta) = h_a = h_b$, so we can calculate $h_{\mathrm{top}}(X_\vartheta)$ by computing h_a or h_b. Since for all distinct realisations $u, v \in \vartheta(a) \cup \vartheta(b)$ we have $u \neq v$ and ϑ is constant length, we have

$$\vartheta^{m+1}(a) = \vartheta^m(aa) \cup \vartheta^m(ab) \text{ and } \vartheta^{m+1}(b) = \vartheta^m(ba)$$

for all $m \in \mathbb{N}$, where the union is disjoint. It follows that

$$\#\vartheta^{m+1}(a) = \#\vartheta^m(a)(\#\vartheta^m(a) + \#\vartheta^m(b)) \text{ and } \#\vartheta^{m+1}(b) = \#\vartheta^m(a)\#\vartheta^m(b),$$

noting that $\vartheta^{m+1}(u) = \vartheta^{m+1}(u_1) \cdots \vartheta^{m+1}(u_{|u|})$ for all $u \in \mathcal{L}_\vartheta$. We next show that $\#\vartheta^m(a) = (m+1)\#\vartheta^m(b)$ for all $m \in \mathbb{N}$. For $m = 1$, the identity clearly holds since $\#\vartheta(a) = 2$ and $\#\vartheta(b) = 1$. For $m \geq 2$, we have

$$\frac{\#\vartheta^m(a)}{\#\vartheta^m(b)} = \frac{\#\vartheta^{m-1}(a)(\#\vartheta^{m-1}(a) + \#\vartheta^{m-1}(b))}{\#\vartheta^{m-1}(a)\#\vartheta^{m-1}(b)} = \frac{\#\vartheta^{m-1}(a)}{\#\vartheta^{m-1}(b)} + 1;$$

thus, it follows by induction that $\#\vartheta^m(a)/\#\vartheta^m(b) = m+1$ for all $m \in \mathbb{N}$. Specifically, $\#\vartheta^m(a) = (m+1)\#\vartheta^m(b)$. Hence,

$$\log(\#\vartheta^m(b)) = \log(\#\vartheta^{m-1}(a)\#\vartheta^{m-1}(b)) = \log m + 2\log(\#\vartheta^{m-1}(b)),$$

and it follows inductively that

$$\frac{1}{2^m}\log(\#\vartheta^m(b)) = \sum_{n=1}^{m} \frac{1}{2^n}\log n.$$

Letting $m \to \infty$, we obtain

$$h_{\text{top}}(X_\vartheta) = h_b = \sum_{n=1}^{\infty} \frac{1}{2^n}\log n.$$

We note that the topological entropy of the subshift in Example 3.2.10 can also be written as $\log\sigma$, where σ is Somos's quadratic recurrence constant:

$$\sigma = \sqrt{1\sqrt{2\sqrt{3\sqrt{4\cdots}}}} = \prod_{k=1}^{\infty}\left(1 + \frac{1}{k}\right)^{2^{-k}}.$$

It is an open question as to whether σ is algebraic or transcendental [31, p. 446]. Similarly, it is not known whether the topological entropy of the random Fibonacci substitution (Example 3.2.7) is the logarithm of an algebraic number. Gohlke, Rust and Spindeler asked whether the topological entropy of a primitive random substitution is always the logarithm of an algebraic number [38, Question 44]. This question was recently answered in the negative by Escolano, Manibo and Miro [22], who presented an example of a primitive random substitution for which the topological entropy of the associated subshift is the logarithm of a transcendental number.

In the compatible setting, the exact value of the topological entropy can be obtained if either the identical or disjoint set condition are satisfied (recall Proposition 3.2.3). For constant length random substitutions, this is also possible, provided the random substitution has uniform inflation set cardinalities.

Proposition 3.2.11. Let ϑ be a primitive random substitution of constant length $\ell \geq 2$ for which there exists an $N \in \mathbb{N}$ such that $\#\vartheta(a) = N$ for all $a \in \mathcal{A}$.
If ϑ satisfies the identical set condition, then

$$h_{\mathrm{top}}(X_\vartheta) = \frac{1}{\ell} \log N.$$

If ϑ satisfies the disjoint set condition, then

$$h_{\mathrm{top}}(X_\vartheta) = \frac{1}{\ell - 1} \log N.$$

Proof. First, assume ϑ satisfies the identical set condition. We prove, by induction, that $\#\vartheta^k(a) = N^{\ell^{k-1}}$ for all $k \in \mathbb{N}$ and $a \in \mathcal{A}$. This is immediate in the case $k = 1$, so assume it holds for all $k \leq m \in \mathbb{N}$. The identical set condition gives that, for each $a \in \mathcal{A}$, we have $\vartheta^m(u) = \vartheta^m(v)$ for all $u, v \in \vartheta(a)$. Hence, for any fixed $u \in \vartheta(a)$, we have

$$\#\vartheta^{m+1}(a) = \#\vartheta^m(u) = \prod_{i=1}^{\ell} (\#\vartheta^m(u_j)) = (N^{\ell^{m-1}})^\ell = N^{\ell^m},$$

where the second equality follows by the constant length property and the third follows by the inductive hypothesis. Thus, for all $k \in \mathbb{N}$ and $a \in \mathcal{A}$, we have

$$\frac{1}{\ell^k} \log(\#\vartheta^k(a)) = \frac{1}{\ell} \log(\#\vartheta(a)) = \frac{1}{\ell} \log N.$$

Letting $k \to \infty$, it follows by Theorem 3.2.8 that $h_{\mathrm{top}}(X_\vartheta) = \ell^{-1} \log N$.

Now assume ϑ satisfies the disjoint set condition. We show that $\#\vartheta^k(a) = N^{\sum_{j=0}^{k-1} \ell^j}$ for all $k \in \mathbb{N}$ and $a \in \mathcal{A}$. Again, this is immediate for $k = 1$, so assume it holds for all $k \leq m \in \mathbb{N}$. By the disjoint set condition, for all $u, v \in \vartheta(a)$ with $u \neq v$, we have $\vartheta^m(u) \cap \vartheta^m(v) = \varnothing$, so it follows that

$$\#\vartheta^{m+1}(a) = \sum_{u \in \vartheta(a)} \#\vartheta^m(u) = \sum_{u \in \vartheta(a)} \prod_{j=1}^{\ell} (\#\vartheta^m(u_j)) = (\#\vartheta(a)) \left(N^{\sum_{j=0}^{m-1} \ell^j} \right)^{\ell} = N^{\sum_{j=0}^{m} \ell^j},$$

where the second inequality follows by the constant length property, the third follows by the inductive hypothesis and the fourth by the fact that $\#\vartheta(a) = N$. Hence, for all $k \in \mathbb{N}$ and $a \in \mathcal{A}$, we have

$$\frac{1}{\ell^k} \log(\#\vartheta^k(a)) = \left(\frac{1}{\ell^k} \sum_{j=0}^{k-1} \ell^k \right) \log N \to \frac{1}{\ell - 1} \log N$$

as $k \to \infty$. Thus, it follows by Theorem 3.2.8 that

$$h_{\mathrm{top}}(X_\vartheta) = \frac{1}{\ell - 1} \log N,$$

and this completes the proof. $\square$

3.2.3 Proof of Theorem 3.2.8

If ϑ is a primitive random substitution of constant length ℓ, we have $\vartheta^m(a) \subseteq \mathcal{L}_\vartheta^{\ell^m} = \mathcal{L}^{\ell^m}(X_\vartheta)$ for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$, so $\underline{h}_a \leq \overline{h}_a \leq h_{\mathrm{top}}(X_\vartheta)$. Thus, to prove Theorem 3.2.8 it suffices to show that $h_{\mathrm{top}}(X_\vartheta) \leq \underline{h}_a$ for all $a \in \mathcal{A}$. The following upper bound for $h_{\mathrm{top}}(X_\vartheta)$ is central to our proof of this inequality. We highlight that the following does not require primitivity.

Proposition 3.2.12. Let ϑ be a random substitution of constant length ℓ and let $m \in \mathbb{N}$. For each k, let u^k denote the level-k inflation word for which $\#\vartheta^m(u^k)$ is maximised. Then, for all $k \in \mathbb{N}$, the following inequality holds:

$$h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\ell^m - 1} \sum_{a \in \mathcal{A}} \frac{|u^k|_a}{\ell^k} \log(\#\vartheta^m(a)).$$

Proof. Fix $m, k \in \mathbb{N}$ and let $n \in \mathbb{N}$. By definition, every finite word admitted by the subshift is a subword of a realisation of $\vartheta^m(v)$ for some $v \in \mathcal{L}(X_\vartheta)$. Moreover, since ϑ is constant length, for every legal word of length $n\ell^m$, there exists such a v with length $n+1$, so

$$\mathcal{L}^{n\ell^m}(X_\vartheta) \subseteq \bigcup_{v\in\mathcal{L}^{n+1}(X_\vartheta)} \bigcup_{j=1}^{\ell^m} \vartheta^m(v)_{[j,j+n\ell^m-1]}. \tag{3.1}$$

By the definition of the subshift X_ϑ, for every $v \in \mathcal{L}^{n+1}(X_\vartheta)$ there exists a legal word u and $w \in \vartheta^k(u)$ such that v is a subword of w. Moreover, by the constant length property, there exists such a u with length $k_n = \lceil \ell^{-k}(n+1)\rceil + 2$, so we have that

$$\#\vartheta^m(v) \leq \#\vartheta^m(w) \leq (\#\vartheta^m(u^k))^{k_n} = \prod_{a\in\mathcal{A}} \left((\#\vartheta^m(a))^{|u^k|_a}\right)^{k_n},$$

where the second inequality follows by breaking w into level-k inflation words and the fact that u^k is the level-k inflation word for which $\#\vartheta^m(u^k)$ is maximised. Since this bound is independent of the choice of $v \in \mathcal{L}_\vartheta^{n+1}$, it follows by (3.1) that

$$\frac{1}{n\ell^m}\log(\#\mathcal{L}^{n\ell^m}(X_\vartheta)) \leq \frac{1}{n\ell^m}\log \ell^m + \frac{1}{n\ell^m}\log(\#\mathcal{L}^{n+1}(X_\vartheta)) + \frac{k_n}{n\ell^m}\sum_{a\in\mathcal{A}}|u^k|_a \log(\#\vartheta^m(a)).$$

Noting that $k_n/n \to \ell^{-k}$ as $n \to \infty$, we deduce that

$$\left(1 - \frac{1}{\ell^m}\right) h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\ell^m}\sum_{a\in\mathcal{A}} \frac{|u^k|_a}{\ell^k}\log(\#\vartheta^m(a)).$$

Dividing by $1 - \ell^{-m}$ completes the proof. □

We now give the proof of Theorem 3.2.8.

Proof of Theorem 3.2.8. It suffices to show that $\underline{h}_a \geq h_{\mathrm{top}}(X_\vartheta)$ for all $a \in \mathcal{A}$. To this end, let $n \in \mathbb{N}$ and, for each $k \in \mathbb{N}$, let u^k be the level-k inflation word for which $\#\vartheta^n(u^k)$ is maximised. By Proposition 3.2.12, we have

$$h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\ell^n - 1}\sum_{a\in\mathcal{A}} \frac{|u^k|_a}{\ell^k}\log(\#\vartheta^n(a))$$

for all $k \in \mathbb{N}$, and so

$$h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\ell^n - 1}\liminf_{k\to\infty}\sum_{a\in\mathcal{A}} \frac{|u^k|_a}{\ell^k}\log(\#\vartheta^n(a)). \tag{3.2}$$

We show that, for every $b \in \mathcal{A}$, the right hand side is bounded above by $\ell^n(\ell^n - 1)^{-1}\underline{h}_b$. For each $b \in \mathcal{A}$ and $k \in \mathbb{N}$, let v_b^k denote the realisation of $\vartheta^k(b)$ for which $\#\vartheta^n(v_b^k)$ is maximised. By definition, for every $k \in \mathbb{N}$ there exists a $b \in \mathcal{A}$ such that $v_b^k = u^k$. For each $k \in \mathbb{N}$, let $b(k) \in \mathcal{A}$ be a letter such that $\#\vartheta^n(v_{b(k)}^k) \leq \#\vartheta^n(v_a^k)$ for all $a \in \mathcal{A}$. By primitivity, there is an integer K such that, for all $a \in \mathcal{A}$, there is a realisation of $\vartheta^K(a)$ in which every letter appears at least once. By construction, there exists an inflation word $w \in \vartheta^K(b)$ such that $v_b^k \in \vartheta^{k-K}(w)$. Moreover, since v_b^k was chosen so that $\#\vartheta^n(v_b^k)$ is maximised, for all $w \in \vartheta^K(b)$ we have

$$\#\vartheta^n(v_b^k) \geq \prod_{j=1}^{\ell^K} \#\vartheta^n(v_{w_j}^{k-K}) = \prod_{a \in \mathcal{A}} (\#\vartheta^n(v_a^{k-K}))^{|w|_a}.$$

Since, by primitivity, there exists a realisation of $\vartheta^K(b)$ in which every letter appears at least once and, for all k, there exists a $c \in \mathcal{A}$ for which $v_c^{k-K} = u^{k-K}$, it follows that

$$\log(\#\vartheta^n(v_{b(k)}^k)) \geq \log(\#\vartheta^n(u^{k-K})) + (\ell^K - 1)\log(\#\vartheta^n(v_{b(k-K)}^{k-K})).$$

Iterating the above, we obtain that for all $m \in \mathbb{N}$ and $k \geq Km$, we have

$$\log(\#\vartheta^n(v_{b(k)}^k)) \geq \sum_{j=1}^{m} (\ell^K - 1)^{j-1} \log(\#\vartheta^n(u^{k-jK})) + (\ell^K - 1)^m \log(\#\vartheta^n(v_{b(k-mK)}^{k-mK})).$$

Noting that all terms in the above are non-negative, it follows that for all $b \in \mathcal{A}$ and $m \in \mathbb{N}$, we have

$$\begin{aligned} \underline{h}_b = \liminf_{k\to\infty} \frac{1}{\ell^{n+k}} \log(\#\vartheta^{n+k}(b)) &\geq \frac{1}{\ell^n} \liminf_{k\to\infty} \frac{1}{\ell^k} \log(\#\vartheta^n(v_b^k)) \\ &\geq \frac{1}{\ell^{n+K}} \sum_{j=1}^{m} \left(\frac{\ell^K - 1}{\ell^K}\right)^{j-1} \liminf_{k\to\infty} \frac{1}{\ell^{k-jK}} \log(\#\vartheta^n(u^{k-jK})). \end{aligned}$$

By the constant length property and (3.2), we have

$$\liminf_{k\to\infty} \frac{1}{\ell^{k-jK}} \log(\#\vartheta^n(u^{k-jK})) = \liminf_{k\to\infty} \sum_{a \in \mathcal{A}} \frac{|u^{k-jK}|_a}{\ell^{k-jK}} \log(\#\vartheta^n(a)) \geq (\ell^n - 1) h_{\mathrm{top}}(X_\vartheta).$$

for all $j \in \{1, \ldots, m\}$. Hence, we obtain

$$\underline{h}_b \geq \frac{\ell^n - 1}{\ell^{n+K}} h_{\mathrm{top}}(X_\vartheta) \sum_{j=1}^{m} \left(\frac{\ell^K - 1}{\ell^K} \right)^{j-1} \xrightarrow{m \to \infty} \frac{\ell^n - 1}{\ell^n} h_{\mathrm{top}}(X_\vartheta),$$

noting that $\sum_{j=1}^{\infty} ((\ell^K - 1)/\ell^K)^{j-1} = \ell^K$. Finally, since this bound holds for all $n \in \mathbb{N}$, we conclude that $\underline{h}_b \geq h_{\mathrm{top}}(X_\vartheta)$, which completes the proof. □

3.2.4 General bounds for constant length random substitutions

The coincidence of topological entropy and inflation word entropy given by Theorem 3.2.8 provides a mechanism for calculating the topological entropy for a broad class of non-compatible random substitution subshifts. However, a limitation of Theorem 3.2.8 is that it does not give any information about the rate of convergence. Moreover, it does not provide a means of calculating topological entropy for non-primitive random substitutions. In this section, we prove general bounds on the topological entropy for constant length random substitutions, which provide a means of obtaining good estimates, even in cases where a closed-form expression cannot be obtained via Theorem 3.2.8. Central to our approach is the following definition.

Definition 3.2.13. Let ϑ be a random substitution over some alphabet $\mathcal{A}$. We say that a vector $\nu = (\nu_a)_{a \in \mathcal{A}} \in [0,1]^{\# A}$ is a *permissible letter frequency vector* for ϑ if there exists a letter $b \in \mathcal{A} \cap \mathcal{L}(X_\vartheta)$, a sequence of integers $(n_k)_k$ and sequence of exact inflation words $v^k \in \vartheta^{n_k}(b)$ such that $|v^k|_a / |v^k| \to \nu_a$ for all $a \in \mathcal{A}$.

The assumption that $b \in \mathcal{L}(X_\vartheta)$ guarantees that the sequence v^k is admitted by the subshift X_ϑ. The inclusion $\mathcal{A} \subseteq \mathcal{L}(X_\vartheta)$ always holds under primitivity, however, there are non-primitive random substitutions for which there exists a letter not admitted by the subshift. For example, for the random substitution $\vartheta \colon a \mapsto \{ab, ba\}$, $b \mapsto \{aa\}$, $c \mapsto \{bb\}$ we have that $c \notin \mathcal{L}(X_\vartheta)$.

Recall from Proposition 3.2.2 that if ϑ is a primitive and compatible random substitution, with Perron–Frobenius eigenvalue $\lambda > 1$ and corresponding right eigenvector $\mathbf{R} = (R_a)_{a \in \mathcal{A}}$, then the following bounds hold for all $m \in \mathbb{N}$:

$$\frac{1}{\lambda^m} \sum_{a \in \mathcal{A}} R_a \log(\# \vartheta^m(a)) \leq h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\lambda^m - 1} \sum_{a \in \mathcal{A}} R_a \log(\# \vartheta^m(a)).$$

Under compatibility, the only permissible letter frequency vector is $\mathbf{R}$. However, without compatibility there often exists a continuum of permissible letter frequency vectors. We prove an analogue of the above bounds for constant length random substitutions that holds without compatibility, using the notion of a permissible letter frequency vector. In the lower bound, we can replace $\mathbf{R}$ with any permissible letter frequency vector; however, for the upper bound we require a particular choice.

Proposition 3.2.14. Let ϑ be a random substitution of constant length $\ell \geq 2$ and let $\nu = (\nu_a)_{a\in\mathcal{A}}$ be a permissible letter frequency vector. Then,

$$h_{\mathrm{top}}(X_\vartheta) \geq \frac{1}{\ell} \sum_{a\in\mathcal{A}} \nu_a \log(\#\vartheta(a)).$$

Moreover, there exists a permissible letter frequency vector $\eta = (\eta_a)_{a\in\mathcal{A}}$ such that

$$h_{\mathrm{top}}(X_\vartheta) \leq \frac{1}{\ell - 1} \sum_{a\in\mathcal{A}} \eta_a \log(\#\vartheta(a)).$$

In particular, η can be taken to be any permissible letter frequency vector that maximises the quantity $\sum_{a\in\mathcal{A}} \eta_a \log(\#\vartheta(a))$.

Proof. We first prove the lower bound. For any permissible letter frequency vector ν, there exists a letter $b \in \mathcal{A} \cap \mathcal{L}(X_\vartheta)$, a sequence of positive integers $(n_k)_k$ with $n_k \to \infty$ and realisations $v^k \in \vartheta^{n_k}(b)$ such that $|v^k|_a/|v^k| \to \nu_a$ as $k \to \infty$, for all $a \in \mathcal{A}$. Since $\vartheta(X_\vartheta) \subseteq X_\vartheta$, every realisation of $\vartheta(v^k)$ is admitted by the subshift X_ϑ. As every realisation of $\vartheta(v^k)$ has length ℓ^{n_k+1}, it follows that $\mathcal{L}^{\ell^{n_k+1}}(X_\vartheta) \geq \#\vartheta(v^k)$. Moreover, the constant length property gives that $\#\vartheta(v^k) = \prod_{a\in\mathcal{A}} (\#\vartheta(a))^{|v^k|_a}$, so we have

$$\frac{1}{\ell^{n_k+1}} \log(\#\mathcal{L}^{\ell^{n_k+1}}(X_\vartheta)) \geq \frac{1}{\ell} \sum_{a\in\mathcal{A}} \frac{|v^k|_a}{\ell^{n_k}} \log(\#\vartheta(a)).$$

Letting $k \to \infty$, we obtain

$$h_{\mathrm{top}}(X_\vartheta) \geq \frac{1}{\ell} \sum_{a\in\mathcal{A}} \nu_a \log(\#\vartheta(a)).$$

The upper bound is largely a consequence of Proposition 3.2.12. Letting $(u^k)_k$ denote the sequence of level-k inflation words for which $\#\vartheta(u^k)$ is maximised, Proposition 3.2.12 gives

$$h_{\text{top}}(X_\vartheta) \leq \frac{1}{\ell - 1} \sum_{a \in \mathcal{A}} \frac{|u^k|_a}{\ell^k} \log(\#\vartheta^m(a)). \tag{3.3}$$

Since $\mathcal{A}$ is finite, there exists a letter $b \in \mathcal{A}$ such that $u^k \in \vartheta^k(b)$ for infinitely many k. Thus, by the compactness of $\mathbb{R}^{\#\mathcal{A}}$, there exists a sequence of positive integers $(k_n)_n$ such that $u^{k_n} \in \vartheta^{k_n}(b)$ for all $n \in \mathbb{N}$ and $|u^{k_n}|_a/\ell^{k_n}$ converges for all $a \in \mathcal{A}$. By definition, this limit is a permissible letter frequency vector η. Passing to limits along the subsequence $(k_n)_n$ in (3.3), we obtain that

$$h_{\text{top}}(X_\vartheta) \leq \frac{1}{\ell - 1} \sum_{a \in \mathcal{A}} \eta_a \log(\#\vartheta(a)).$$

In particular, the above holds for any permissible letter frequency vector that maximises the quantity $\sum_{a \in \mathcal{A}} \eta_a \log(\#\vartheta(a))$. This completes the proof. □

As a consequence of Proposition 3.2.14, we obtain the following characterisation of positive topological entropy for constant length random substitutions. We emphasise that, in contrast to the results in Section 3.1, we do not assume primitivity in the following.

Corollary 3.2.15. If ϑ is a constant length random substitution, then $h_{\text{top}}(X_\vartheta) > 0$ if and only if there exists a permissible letter frequency vector ν and $a \in \mathcal{A}$ such that $\nu_a > 0$ and $\#\vartheta(a) \geq 2$.

We conclude this section by presenting an example of a primitive random substitution where Theorem 3.2.8 is insufficient to obtain an exact formula for the topological entropy but where the bounds given by Proposition 3.2.14 allow us to obtain a good estimate.

Example 3.2.16. Let ϑ be the primitive random substitution defined by $\vartheta\colon a \mapsto \{aa, bb\}$, $b \mapsto \{ab\}$, and let X_ϑ denote the associated subshift. Since for all $m \in \mathbb{N}$, there exists a realisation of $\vartheta^m(a)$ in which a is the only letter that appears, the vector $(1,0)^\top$ is a permissible letter frequency vector for every power of ϑ. For all $m \in \mathbb{N}$, $\#\vartheta^m(a) \geq \#\vartheta^m(b)$, so for every permissible letter frequency vector $\nu = (\nu_a, \nu_b)$, we have $\nu_a \log(\#\vartheta^m(a)) + \nu_b \log(\#\vartheta^m(b)) \leq \log(\#\vartheta^m(a))$, noting that $\nu_a + \nu_b = 1$. Hence, it follows by Proposition 3.2.14 (applied to ϑ^m) that

$$\frac{1}{2^m} \log(\#\vartheta^m(a)) \leq h_{\text{top}}(X_\vartheta) \leq \frac{1}{2^m - 1} \log(\#\vartheta^m(a)).$$

For all $m \in \mathbb{N}$, the inflation sets $\vartheta^m(a)$ satisfy the following relations:

$$\vartheta^m(a) = \vartheta^{m-1}(aa) \cup \vartheta^{m-1}(bb) \quad \text{and} \quad \vartheta^m(b) = \vartheta^{m-1}(ab).$$

Since ϑ is constant length and $\vartheta(a) \cap \vartheta(b) = \varnothing$, we have that $\vartheta^m(a) \cap \vartheta^m(b) = \varnothing$ for all $m \in \mathbb{N}$. Hence, it follows from the above relations that

$$\#\vartheta^m(a) = \left(\#\vartheta^{m-1}(a)\right)^2 + \left(\#\vartheta^{m-1}(b)\right)^2 = \left(\#\vartheta^{m-1}(a)\right)^2 + \prod_{i=1}^{m-2} \left(\#\vartheta^i(a)\right)^2.$$

A computer-assisted calculation for the case $m = 14$ gives that, to four decimal places, the topological entropy of the subshift X_ϑ is 0.4115.

3.3 Non-primitive random substitutions

Recall from Corollary 3.1.2 that if a primitive random substitution subshift has zero topological entropy, then the complexity function is either $\Theta(1)$ or $\Theta(n)$. Without primitivity, the picture is very different. In this section, we show that, without primitivity, constant length random substitutions give rise to a broad variety of subshifts with complexity functions not witnessed in the deterministic or primitive random settings.

3.3.1 Polynomial growth

There is no polynomial complexity gap for subshifts of constant length random substitutions, beyond that given by the Morse–Hedlund theorem [20, 57]. In particular, we have the following.

Proposition 3.3.1. The set of α for which there exists a constant length random substitution subshift with $\Theta(n^\alpha)$ complexity function is dense in $[1, \infty)$.

Proof. We show that there exists a set A that is dense in $[1, \infty)$ such that, for every $\alpha \in A$, there exists a constant length random substitution whose subshift has $\Theta(n^\alpha)$ complexity function. To this end, let $\ell \geq 3$, let $\mathcal{A} = \{a_1, \ldots, a_{\ell+2}\}$ be an alphabet of $\ell + 2$ letters, and let

$$\mathcal{R}_\ell = \left\{ u \in \mathcal{A}^\ell \colon \text{ for all } i \in \{1, \ldots, \ell\},\, a_i \text{ appears in } u \text{ precisely once} \right\}.$$

Note that every $u \in \mathcal{R}_\ell$ has length $|u| = \ell$. Given a non-empty subset $\mathcal{S}$ of $\mathcal{R}_\ell$, let $\vartheta_{\ell,\mathcal{S}}$ be the random substitution of constant length ℓ defined over the alphabet $\mathcal{A}$ by

$$\vartheta_{\ell,\mathcal{S}}\colon \begin{cases} a_i \mapsto a_i \cdots a_i \quad \text{for all } i \in \{1, \dots, \ell\}, \\ a_{\ell+1} \mapsto \mathcal{S}, \\ a_{\ell+2} \mapsto a_{\ell+1} a_{\ell+2} a_1 \cdots a_1, \end{cases}$$

and let $X_{\ell,\mathcal{S}}$ denote the corresponding subshift. For notational convenience, we write $p_{\ell,\mathcal{S}}$ for the complexity function of $X_{\ell,\mathcal{S}}$. We show that $p_{\ell,\mathcal{S}}$ is $\Theta(n^{1+\log_\ell(\#S)})$. We first prove this in the case $n = \ell^k$ for some $k \in \mathbb{N}$ and then extend to all $n \in \mathbb{N}$ in the second step. To this end, we first compute the cardinalities $\#\vartheta^m(a_i)$ for each $m \in \mathbb{N}$ and $i \in \{1, \dots, \ell + 2\}$. Observe that $\#\vartheta^m(a_i) = 1$ for all $m \in \mathbb{N}$ and $i \in \{1, \dots, \ell\}$. Further, since every word in the set $\mathcal{S}$ contains only letters from the set $\{a_1, \dots, a_\ell\}$, we have that $\#\vartheta^m(a_{\ell+1}) = \#\mathcal{S}$ for all $m \in \mathbb{N}$. Finally, we have

$$\#\vartheta^m(a_{\ell+2}) = (\#\vartheta^{m-1}(a_{\ell+1}))(\#\vartheta^{m-1}(a_{\ell+2})) = (\#\mathcal{S})(\#\vartheta^{m-1}(a_{\ell+2})) = (\#\mathcal{S})^{m-1}$$

for all $m \in \mathbb{N}$. Every letter in the alphabet $\mathcal{A}$ is admitted by the subshift, so $\mathcal{L}(X_{\ell,\mathcal{S}}) = \mathcal{L}_{\vartheta_{\ell,\mathcal{S}}}$ and thus $p_{\ell,\mathcal{S}}(n) = \#\mathcal{L}_\vartheta^n$ for all $n \in \mathbb{N}$. By the constant length property, if $k \in \mathbb{N}$, $m \in \{1, \dots, k\}$ and u is a legal word of length ℓ^k, then there exists a legal word v of length $\ell^{k-m} + 1$ and an integer $j \in \{1, \dots, \ell^{k-m}\}$ such that $u \in \vartheta^m(v)_{[j,j+\ell^k]}$; hence,

$$\mathcal{L}_{\vartheta_{\ell,\mathcal{S}}}^{\ell^k} = \bigcup_{v \in \mathcal{L}_{\vartheta_{\ell,\mathcal{S}}}^{\ell^{k-m}+1}} \bigcup_{j=1}^{\ell^m} \vartheta^m(v)_{[j,j+\ell^k-1]}. \tag{3.4}$$

For all $k \geq 2$, we have that $\#\vartheta^k(a_{\ell+2}) \geq \#\vartheta^k(a_{\ell+1}) \geq \#\vartheta^k(a_i)$ for all $i \in \{1, \dots, \ell\}$. Since $a_{\ell+2}a_{\ell+2}$ is not a legal word, it follows that $\#\vartheta^k(v) \leq \#\vartheta^k(a_{\ell+1})\#\vartheta^k(a_{\ell+2}) = (\#\mathcal{S})^k$ for all $v \in \mathcal{L}_\vartheta^2$. Hence, it follows by (3.4) in the case $m = k$ that

$$p_{\ell,\mathcal{S}}(\ell^k) \leq p_{\ell,\mathcal{S}}(2)\ell^k(\#\mathcal{S})^k = p_{\ell,\mathcal{S}}(2)\,(\ell^k)^{1+\log_\ell(\#S)}. \tag{3.5}$$

For the lower bound, observe that the $\ell + 1$ letter word $w = a_{\ell+1}a_{\ell+2}a_1 \cdots a_1$ is legal, so it follows by (3.4) in the case $m = k - 1$ that

$$\mathcal{L}_{\vartheta_{\ell,\mathcal{S}}}^{\ell^k} \supseteq \bigcup_{j=1}^{\ell^{k-1}} \vartheta^{k-1}(w)_{[j,j+\ell^k-1]}. \tag{3.6}$$

Every realisation of $\vartheta^{k-1}(w)$ contains precisely one occurrence of the letter $a_{\ell+2}$. Further, there is a positive integer $\ell^{k-1} < n < 2\ell^{k-1}$ such that for every realisation of $\vartheta^{k-1}(w)$, the letter $a_{\ell+2}$ appears in position n. Thus, the above union is disjoint. Moreover, for all $j \in \{1, \ldots, \ell^{k-1}\}$, the inflated image of the second letter of w is contained in the corresponding realisation of $\vartheta^{k-1}(w)_{[j,j+\ell^k-1]}$, so $\#\vartheta^{k-1}(w)_{[j,j+\ell^k-1]} \geq \#\vartheta^{k-1}(a_{\ell+2}) = (\#\mathcal{S})^{k-2}$. Hence, it follows by (3.6) that

$$p_{\ell,\mathcal{S}}(\ell^k) \geq \ell^{k-1}(\#\mathcal{S})^{k-2} = \ell^{-1}(\#\mathcal{S})^{-2}(\ell^k)^{1+\log_\ell(\#\mathcal{S})}. \tag{3.7}$$

Now, let $n \in \mathbb{N}$ and let k be the unique integer such that $\ell^k \leq n < \ell^{k+1}$. By the monotonicity of the complexity function, we have

$$p_{\ell,\mathcal{S}}(\ell^k) \leq p_{\ell,\mathcal{S}}(n) \leq p_{\ell,\mathcal{S}}(\ell^{k+1})$$

and so it follows by (3.5) and (3.7) that

$$\ell^{-2}(\#\mathcal{S})^{-3} n^{1+\log_\ell(\#\mathcal{S})} \leq p_{\ell,\mathcal{S}}(n) \leq \ell\, p_{\ell,\mathcal{S}}(2)\, n^{1+\log_\ell(\#\mathcal{S})}.$$

Hence, we conclude that the complexity function of $X_{\ell,\mathcal{S}}$ is $\Theta(n^{1+\log_\ell(\#\mathcal{S})})$. Since the set

$$A = \{1 + \log_\ell k \colon \ell \in \{3, 4, 5, \ldots\},\, k \in \{1, \ldots, \ell!\}\}$$

is dense in $[1, \infty)$ and, for every $\alpha \in A$, there exists an $\ell \in \{3, 4, 5, \ldots\}$ and a subset $\mathcal{S}$ of $\mathcal{R}_\ell$ such that the subshift $X_{\ell,\mathcal{S}}$ has $\Theta(n^\alpha)$ complexity function, the result follows. $\square$

By combining Proposition 3.3.1 with Pansiot's classification of complexity functions for deterministic substitutions (Proposition 1.6.1), we obtain the following.

Corollary 3.3.2. There exist subshifts of constant length random substitutions with zero topological entropy that cannot be obtained as the subshift of a deterministic substitution.

Example 3.3.3. Let ϑ be the random substitution defined by $\vartheta\colon a \mapsto \{aaa\}$, $b \mapsto \{bbb\}$, $c \mapsto \{ccc\}$, $d \mapsto \{abc,\, acb\}$, $e \mapsto \{dea\}$. The associated subshift has $\Theta(n^{1+\log_3 2})$ complexity function. By Proposition 1.6.1, there does not exist a deterministic substitution subshift with this complexity function, so the subshift X_ϑ cannot be obtained as the subshift of a deterministic substitution.

3.3.2 Intermediate growth

It is also possible for the subshift of a non-primitive constant length random substitution to have an intermediate growth complexity function. In the following, we give sufficient conditions for this phenomenon to occur.

Proposition 3.3.4. Let ϑ be a constant length random substitution for which the following hold:

- for all $a \in \mathcal{A}$ with $\#\vartheta(a) \geq 2$ and all permissible letter frequency vectors ν, we have $\nu_a = 0$;
- there exists a letter $b \in \mathcal{A}$ for which $\#\vartheta(b) \geq 2$ and a realisation $v \in \vartheta(b)$ with $|v|_b \geq 2$.

Then, the associated subshift X_ϑ has intermediate growth complexity function.

Proof. The first condition guarantees that $h_{\text{top}}(X_\vartheta) = 0$ by Corollary 3.2.15. Meanwhile, it follows inductively from the second condition that for all $m \in \mathbb{N}$, there exists a realisation $v^m \in \vartheta^m(b)$ with $|v^m|_b \geq 2^m$. This gives that the letter b is admitted by the subshift, so for all $m \in \mathbb{N}$ we have $\vartheta^m(b) \subseteq \mathcal{L}^{\ell^m}(X_\vartheta)$; hence,

$$p_{X_\vartheta}(\ell^m) \geq \#\vartheta^m(b) \geq \#\vartheta(v^{m-1}) \geq 2^{|v^m|_b} \geq 2^{2^m} = 2^{(\ell^m)^{\log_\ell 2}}.$$

Thus, if $n \in \mathbb{N}$ and m is the integer such that $\ell^m \leq n < \ell^{m+1}$, then by monotonicity of the complexity function we have

$$p_{X_\vartheta}(n) \geq p_{X_\vartheta}(\ell^m) \geq 2^{(\ell^m)^{\log_\ell 2}} \geq 2^{2^{-1} n^{\log_\ell 2}},$$

which grows faster than any polynomial. Hence, we conclude that the complexity function of X_ϑ has intermediate growth. $\square$

Example 3.3.5. Let ϑ be the random substitution defined by $\vartheta\colon a \mapsto \{aaa\}$, $b \mapsto \{abb, bba\}$ and let X_ϑ denote the corresponding subshift. We have that $|\vartheta^m(a)|_b = 0$ and $|\vartheta^m(b)|_b = 2^m$ for all $m \in \mathbb{N}$, so for every realisation $v \in \vartheta^m(a) \cup \vartheta^m(b)$ we have $|v|_b/3^m \leq 2^m/3^m \to 0$. Hence, the only permissible letter frequency vector is $(1,0)^\top$; since $\#\vartheta(a) = 1$, the first condition of Proposition 3.3.4 is satisfied. Moreover, since $abb \in \vartheta(b)$ and $\#\vartheta(b) \geq 2$, the second condition is also satisfied. Therefore, it follows by Proposition 3.3.4 that X_ϑ has intermediate growth complexity function. Specifically, one can show that $p_{X_\vartheta}(n) \geq 2^{2^{-1}n^{\log_3 2}}$ for all $n \in \mathbb{N}$.

Propositions 3.3.1 and 3.3.4, together with our results on topological entropy, illustrate that a broad variety of functions can be obtained as the complexity function of the subshift of a non-primitive random substitution. However, these results are far from a complete classification, and further work is required to better understand the full range of behaviour that can be exhibited, particularly the different possible intermediate growth rates.

CHAPTER 4

MEASURE THEORETIC ENTROPY

Topological entropy is, almost by definition, blind to the generating probabilities assigned to a random substitution. This is not the case for aspects such as word frequencies and diffraction spectra, which are almost-sure properties in the limit of an appropriate substitution Markov process [65]. Alternatively, these properties can be associated with the corresponding frequency measure. It is therefore reasonable to treat entropy on the same footing: interpreting it as a quantity that is generic with respect to the frequency measure. Moreover, this perspective more closely reflects the original context considered by Godrèche and Luck [34], who were interested in random substitutions providing models for physical quasicrystals, the entropy of which depend on the underlying Markov process.

In this chapter we provide a systematic treatment of measure theoretic entropy for random substitution subshifts, where the measure in question is a frequency measure. We introduce a new notion called *measure theoretic inflation word entropy* which, similarly to the notion of inflation word entropy studied in Chapter 3 for topological entropy, is defined in terms of the underlying random substitution, as opposed to the subshift. We show that for all primitive random substitutions, this notion coincides with the measure theoretic entropy of the associated subshift with respect to the frequency measure. This allows the measure theoretic entropy to be calculated for many random substitution subshifts and establishes a natural analogue to the results on topological entropy presented in Chapter 3. However, we emphasise that our present setting is more general, as we do not assume that the underlying random substitution is compatible or constant length.

We give the definition of measure theoretic inflation word entropy in Section 4.1, followed by the statements of our main results. We then present some consequences of these results and give examples of random substitution subshifts where our inflation word analogue allows the measure theoretic entropy to be calculated or accurately estimated. Following this, in Section 4.2, we present the proofs of our main results.

The results presented in this chapter are based on [37, Section 3], which is joint work with P. Gohlke, D. Rust and T. Samuel and has been published in *Annales Henri Poincaré*.

4.1 Measure theoretic inflation word entropy

4.1.1 Definitions and main results

For a primitive random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$, $m \in \mathbb{N}$ and $a \in \mathcal{A}$, we let

$$H_{m,a} = \sum_{v \in \vartheta^m(a)} -\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = v]$$

and let $\mathbf{H}_m$ be the vector given by $\mathbf{H}_m = (H_{m,a})_a$. To reflect the notation for entropy of a partition, we often denote $H_{m,a}$ by $H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^m(a))$. Our main result of this section is that the measure theoretic entropy $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ of the subshift X_ϑ, with respect to the corresponding frequency measure $\mu_{\mathbf{P}}$, coincides with the quantity

$$\lim_{m \to \infty} \frac{1}{\lambda^m} \mathbf{H}_m \cdot \mathbf{R},$$

where $\mathbf{R}$ is the right Perron–Frobenius eigenvector of the substitution. Before giving the formal statements of our main results, we first introduce some notation. Given $p \in (0,1)$, we write $H(p)$ for the entropy of the vector $(p, 1-p)$. That is,

$$H(p) = -p \log p - (1-p) \log(1-p).$$

Our most general result on the relation between the entropy of $\mu_{\mathbf{P}}$ and the sequence of entropies assigned to the Markov processes $(\vartheta_{\mathbf{P}}^m(a))_m$, with $a \in \mathcal{A}$, takes the following form.

Theorem 4.1.1. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with Perron–Frobenius eigenvalue λ and right eigenvector $\mathbf{R}$, and let $\mu_{\mathbf{P}}$ be the frequency measure corresponding to $\vartheta_{\mathbf{P}}$.

Then, for all $m \in \mathbb{N}$, the following bounds hold:

$$\frac{1}{\lambda^m}\mathbf{H}_m \cdot \mathbf{R} - H(\lambda^{-m}) \leq h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq \frac{1}{\lambda^m - 1}\mathbf{H}_m \cdot \mathbf{R}.$$

In particular,

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \lim_{m\to\infty} \frac{1}{\lambda^m}\mathbf{H}_m \cdot \mathbf{R}.$$

If the random substitution $\vartheta_{\mathbf{P}}$ is additionally assumed to have unique realisation paths (recall Definition 2.2.7), then the $H(\lambda^{-m})$ counter-term in the lower bound can be removed. Further, under this additional assumption, we obtain necessary and sufficient conditions for when the lower and upper bounds are achieved.

Theorem 4.1.2. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with unique realisation paths, with Perron–Frobenius eigenvalue λ and right eigenvector $\mathbf{R}$, and let $\mu_{\mathbf{P}}$ be the frequency measure corresponding to $\vartheta_{\mathbf{P}}$. Then, for all $m \in \mathbb{N}$, the following bounds hold:

$$\frac{1}{\lambda^m}\mathbf{H}_m \cdot \mathbf{R} \leq h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq \frac{1}{\lambda^m - 1}\mathbf{H}_m \cdot \mathbf{R}.$$

The upper bound is an equality if and only if $\vartheta_{\mathbf{P}}^m$ satisfies the disjoint set condition; the lower bound is an equality if and only if $\vartheta_{\mathbf{P}}^m$ satisfies the identical set condition with identical production probabilities. Further, the sequence of lower bounds $(\lambda^{-m}\mathbf{H}_m \cdot \mathbf{R})_m$ is non-decreasing in m.

We present the proofs of Theorems 4.1.1 and 4.1.2 in Section 4.2.

It is natural to enquire whether the counter-term $H(\lambda^{-m})$ can also be dropped in the more general case. However, this is not possible, as can be seen from the following example.

Example 4.1.3. Given $p \in (0, 1)$, let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} a & \text{with probability } p, \\ aba & \text{with probability } 1-p, \end{cases} \\ b \mapsto \quad bab \quad \text{with probability } 1, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. The random substitution $\vartheta_{\mathbf{P}}$ gives rise to the finite subshift $X_\vartheta = \{(ab)^{\mathbb{Z}}, (ba)^{\mathbb{Z}}\}$; thus, $h_{\mu_{\mathbf{P}}}(X_\vartheta) = 0$. On the other hand, $\vartheta_{\mathbf{P}}$ is primitive and $H_{\mathbb{P}}(\vartheta_{\mathbf{P}}(a)) = -(p \log p + (1-p)\log(1-p)) > 0$.

4.1.2 Applications

In general, for a random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ with corresponding frequency measure $\mu_{\mathbf{P}}$, the measure theoretic entropy $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ depends on the choice of $\mathbf{P}$. This is in contrast to topological entropy, which is blind to the choice of probabilities. As a consequence of Theorem 4.1.1 we obtain that the dependence on the probability parameters is continuous. In the following, we regard $\mathbf{P}$ as a vector in $\mathbb{R}^r$ equipped with the Euclidean topology, where $r = \sum_{i=1}^{d} r_i = \sum_{i=1}^{d} \#\vartheta(a_i)$ and d is the cardinality of the alphabet. We emphasise that we assume that $\mathbf{P}$ is non-degenerate in the sense that all probabilities are assumed to be strictly positive.

Corollary 4.1.4. Assume the setting of Theorem 4.1.1. The map $\mathbf{P} \mapsto h_{\mu_{\mathbf{P}}}(X_\vartheta)$ is continuous.

Proof. For $0 < \varepsilon < 1$, let D_ε be the domain of those $\mathbf{P}$ such that all entries of $\mathbf{P}$ are greater than ε. Since the complete domain of $\mathbf{P}$ can be obtained as a (nested) union over all D_ε, it is enough to show that the map $\mathbf{P} \mapsto h_{\mu_{\mathbf{P}}}(X_\vartheta)$ is continuous on D_ε for arbitrary ε. The general strategy of the proof is to represent $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ as a uniform limit of continuous functions on D_ε via Theorem 4.1.1.

Recall that $\lambda, \mathbf{H}_m$ and $\mathbf{R}$ all depend implicitly on $\mathbf{P}$. By primitivity, the Perron–Frobenius eigenvalue $\lambda > 1$ is a *simple* eigenvalue for all $\mathbf{P}$. Since the substitution matrix depends analytically on the probability parameters, we can resort to standard results in perturbation theory; compare for example [42]. In particular, λ depends analytically on $\mathbf{P} \in D_\varepsilon$ and since λ is simple, so does $\mathbf{R}$. The entries of $\mathbf{H}_m$ inherit continuity from the fact that the maps $\mathbf{P} \mapsto \mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = u]$ are continuous for all $a \in \mathcal{A}$ and $u \in \mathcal{A}^+$. Hence, the function

$$s_m \colon \mathbf{P} \mapsto \frac{1}{\lambda^m} \mathbf{H}_m \cdot \mathbf{R},$$

is continuous in $\mathbf{P}$ for all $m \in \mathbb{N}$. With this notation, Theorem 4.1.1 can be rephrased as

$$\frac{\lambda^m - 1}{\lambda^m} h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq s_m(\mathbf{P}) \leq h_{\mu_{\mathbf{P}}}(X_\vartheta) + H(\lambda^{-m}), \tag{4.1}$$

for all $m \in \mathbb{N}$. Note that $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ is uniformly bounded from above by the topological entropy of X_ϑ and λ is bounded from below by its minimal value $\lambda_\varepsilon > 1$ on the compact set D_ε. Therefore, the convergence

$$\lim_{m \to \infty} s_m(\mathbf{P}) = h_{\mu_{\mathbf{P}}}(X_\vartheta)$$

is uniform on D_ε, and the result follows. $\square$

Recall that measure theoretic entropy is always bounded above by topological entropy. Thus, if $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ is a primitive random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$, then $h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq h_{\mathrm{top}}(X_\vartheta)$. As a consequence of Theorem 4.1.2, we obtain the following sufficient conditions for positive topological entropy for subshifts of primitive random substitutions with unique realisation paths.

Proposition 4.1.5. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with unique realisation paths for which there exists a letter $b \in \mathcal{A}$ such that $\#\vartheta(b) \geq 2$ and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Then, $0 < h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq h_{\mathrm{top}}(X_\vartheta)$.

Proof. Since $\#\vartheta(b) \geq 2$, there exists a realisation $v \in \vartheta(b)$ such that $0 < \mathbb{P}[\vartheta_{\mathbf{P}}(b) = v] < 1$. By primitivity, the right Perron–Frobenius eigenvector $\mathbf{R}$ of $\vartheta_{\mathbf{P}}$ has strictly positive entries, so it follows that $\lambda^{-1}\mathbf{H}_1 \cdot \mathbf{R} > 0$. Hence, we conclude by Theorem 4.1.2 that $0 < h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq h_{\mathrm{top}}(X_\vartheta)$. $\square$

4.1.3 Examples

We now present several examples where Theorems 4.1.1 and 4.1.2 allow the measure theoretic entropy to be calculated or accurately estimated. We first give some examples for which Theorem 4.1.2 can be applied to obtain a closed-form formula for the measure theoretic entropy.

Example 4.1.6 (Random period doubling)**.** Given $p \in (0,1)$, let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random period doubling substitution

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Recall that $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition and has Perron–Frobenius data $\lambda = 2$, $\mathbf{R} = (2/3, 1/3)^\top$, so Theorem 4.1.2 gives that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \frac{1}{\lambda - 1}\mathbf{H}_1 \cdot \mathbf{R} = \frac{2}{3}(p \log p + (1-p)\log(1-p)).$$

A plot of $h_{\mu_{\mathbf{P}}}(X_\vartheta)$, for $p \in (0,1)$, is given in Figure 4.1 (left).

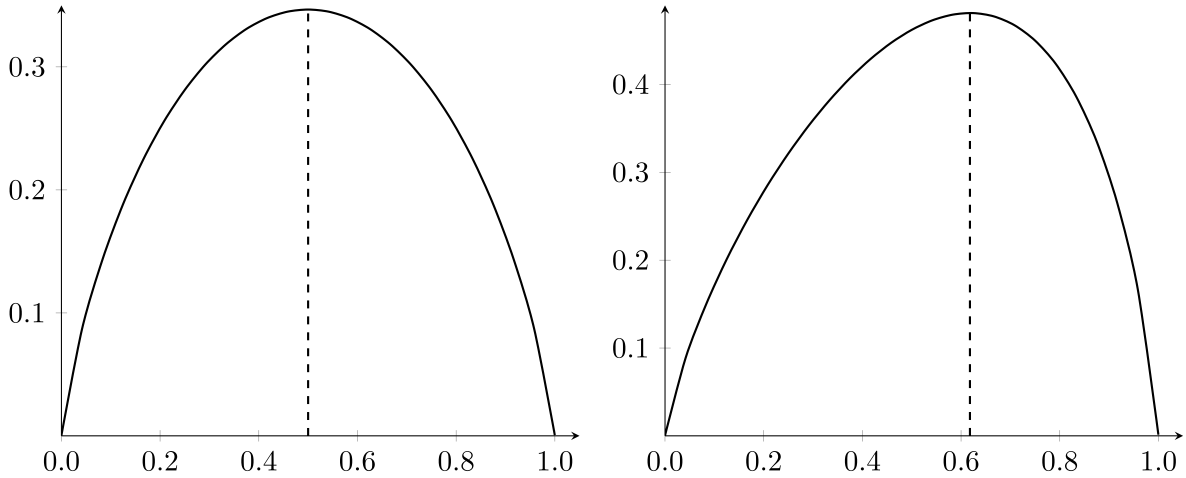


Figure 4.1: Plots of measure theoretic entropy for Example 4.1.6 (left) and Example 4.1.7 (right), for $p \in (0, 1)$.

Example 4.1.7. Let $p \in (0, 1)$ and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} aa & \text{with probability } p, \\ ab & \text{with probability } 1 - p, \end{cases} \\ b \mapsto ba \quad \text{with probability } 1, \end{cases}$$

with corresponding frequency measure $\mu_{\mathbf{P}}$. Since $\vartheta_{\mathbf{P}}$ is constant length, it has unique realisation paths, so Theorem 4.1.2 can be applied. Moreover, since $\vartheta_{\mathbf{P}}$ satisfies the conditions of Proposition 2.2.11, the disjoint set condition is satisfied. The random substitution $\vartheta_{\mathbf{P}}$ has Perron–Frobenius data $\lambda = 2$ and

$$\mathbf{R} = \left(\frac{1}{2 - p}, \frac{1 - p}{2 - p} \right);$$

thus, it follows by Theorem 4.1.2 that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \frac{1}{\lambda - 1} \mathbf{H}_1 \cdot \mathbf{R} = -\frac{1}{2 - p} (p \log p + (1 - p) \log(1 - p)).$$

A plot of the measure theoretic entropy, for $p \in (0, 1)$, is given in Figure 4.1 (right).

Example 4.1.8. Let $p \in (0,1)$ and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon a, b \mapsto \begin{cases} aa & \text{with probability } p, \\ ab & \text{with probability } 1-p, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. The random substitution $\vartheta_{\mathbf{P}}$ satisfies the identical set condition with identical production probabilities and has unique realisation paths, so Theorem 4.1.2 can be applied to obtain an exact formula for the measure theoretic entropy. Since $\vartheta_{\mathbf{P}}$ has Perron–Frobenius data $\lambda = 2$ and $\mathbf{R} = (1/2, 1/2)^{\top}$, Theorem 4.1.2 gives that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \frac{1}{\lambda} \mathbf{H}_1 \cdot \mathbf{R} = -\frac{1}{2}(p \log p + (1-p)\log(1-p)).$$

Recall that the random Fibonacci substitution does not satisfy the identical or disjoint set condition. However, we can still obtain bounds on the measure theoretic entropy using Theorem 4.1.2.

Example 4.1.9 (Random Fibonacci)**.** Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random Fibonacci substitution with uniform probabilities, namely

$$\vartheta_{\mathrm{RF}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } 1/2, \\ ba & \text{with probability } 1/2, \end{cases} \\ b \mapsto a \quad \text{with probability } 1, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Since $\vartheta_{\mathbf{P}}$ is primitive and compatible, thus has unique realisation paths, Theorem 4.1.2 gives that

$$\frac{1}{\lambda^k} \mathbf{H}_k \cdot \mathbf{R} \leq h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq \frac{1}{\lambda^k - 1} \mathbf{H}_k \cdot \mathbf{R}$$

for all $k \in \mathbb{N}$. Obtaining the exact value of $\mathbf{H}_k \cdot \mathbf{R}$ for arbitrary k is difficult. However, a computer-assisted calculation for the case $k = 7$ yields the following bounds:

$$0.4164 < h_{\mu_{\mathbf{P}}}(X_\vartheta) < 0.4313.$$

4.2 Proof of main results

The majority of the work in proving Theorems 4.1.1 and 4.1.2 lies in proving the sequence of lower and upper bounds on $h_{\mu_{\mathbf{P}}}(X_\vartheta)$. Central to proving these bounds is the renormalisation lemma (Lemma 2.3.4). Before we give the proof of Theorems 4.1.1 and 4.1.2, we first provide a reformulation of Lemma 2.3.4 in terms of random variables that will be appropriate for our purposes. We also summarise some of the standard properties of entropy and conditional entropy that we require in our proofs. For ease of notation, we write B^C for the complement of a given set B in what follows.

4.2.1 Renormalisation

The renormalisation lemma (Lemma 2.3.4) is a fundamental ingredient in our proof of Theorems 4.1.1 and 4.1.2. We first prove a reformulation of Lemma 2.3.4 that will be more appropriate for our purposes, by interpreting the expression appearing in Lemma 2.3.4 via the distribution of an appropriate random variable that mirrors the action of $\vartheta_{\mathbf{P}}$ on the initial distribution $\mu_{\mathbf{P}}$, together with the choice of the origin in the inflation word decomposition. For each $n \in \mathbb{N}$ and $w \in \mathcal{L}_\vartheta^n$, we write

$$\mu^{(n)}(w) = \mu_{\mathbf{P}}([w]).$$

Lemma 4.2.1. For each $n \in \mathbb{N}$, $\mu^{(n)}$ is the distribution of a random word $\mathcal{W}_n$ on a finite probability space (Ω_n, P_n), defined as follows. The space

$$\Omega_n = \{(v, u_1, \cdots, u_n, j) : v \in \mathcal{L}_\vartheta^n, u_i \in \vartheta(v_i), 1 \leq j \leq |u_1|\}$$

is equipped with the probability vector

$$P_n \colon (v, u_1, \cdots, u_n, j) \mapsto \frac{1}{\lambda} \mu_{\mathbf{P}}([v]) \prod_{i=1}^{n} \mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = u_i].$$

The random word $\mathcal{W}_n$ is defined via

$$\mathcal{W}_n \colon (v, u_1, \cdots, u_n, j) \mapsto (u_1 \cdots u_n)_{[j, j+n-1]}.$$

Proof. Let $w \in \mathcal{L}_\vartheta^n$. We note that $\mathcal{W}_n^{-1}(\{w\})$ comprises all those elements in Ω_n such that the property $(u_1 \cdots u_n)_{[j,j+n-1]} = w$ holds. That is,

$$P_n(\mathcal{W}_n = w) = \sum_{v \in \mathcal{L}_\vartheta^n} \sum_{u_1,\ldots,u_n} \sum_{j=1}^{|u_1|} \frac{1}{\lambda} \mu_{\mathbf{P}}([v]) \prod_{i=1}^{n} \mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = u_i] \, \delta_{w,(u_1 \cdots u_n)_{[j,j+n-1]}}.$$

Comparing with the expression in Lemma 2.3.4, we further note that

$$\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = w \wedge |\vartheta_{\mathbf{P}}(v_1)| = m] = \sum_{u_1,\ldots,u_n} \prod_{i=1}^{n} \mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = u_i] \, \delta_{m,|u_1|} \, \delta_{w,(u_1 \cdots u_n)_{[j,j+n-1]}}.$$

From this, we obtain that $P_n(\mathcal{W}_n = w) = \mu_{\mathbf{P}}([w])$ and the claim follows. □

Remark 4.2.1. We may interpret the factors occurring in the definition of P_n in terms of the renormalisation step. The term λ^{-1} corresponds to a change of scale due to the expansion of the length of words, $\mu_{\mathbf{P}}([v])$ reflects the choice of a word before the inflation step, and each of $\mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = u_i]$ gives the probability of mapping v_i to the particular word u_i as we apply the random substitution. Marginalised to (prefixes of) v, the distribution induced by P_n and $\mu_{\mathbf{P}}$ are closely related but different in general. To be more precise, we will be interested in the random variable

$$\mathcal{V}_{[1,m]} \colon (v, u_1 \cdots u_n, j) \mapsto v_{[1,m]}$$

for some $m \leqslant n$. Integrating out the dependencies on $u_2, \ldots, u_n$ and j in the first step, we obtain

$$P_n(\mathcal{V}_{[1,m]} = v') = \frac{1}{\lambda} \sum_{v, v_{[1,m]} = v'} \mu_{\mathbf{P}}([v]) \sum_{u_1} |u_1| \mathbb{P}[\vartheta_{\mathbf{P}}(v_1) = u_1] = \frac{1}{\lambda} \mu_{\mathbf{P}}([v']) \mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|].$$

The additional factor $\lambda^{-1}\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]$ accounts for the fact that starting the inflation word decomposition of a word within some $u_1 \in \vartheta(v_1)$ is more probable if $\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]$ is large.

Lemma 4.2.1 provides us with an alternative way to calculate the measure theoretic entropy that will be instrumental for the proof of our main theorems.

Lemma 4.2.2. The measure theoretic entropy $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ satisfies

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \lim_{n \to \infty} \frac{1}{n} H_{P_n}(\mathcal{W}_n).$$

Proof. Let $I_n\colon v \mapsto v$ be the identity map on $\mathcal{L}_\vartheta^n$. By the definition of measure theoretic entropy,

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \lim_{n\to\infty} \frac{1}{n} H_{\mu^{(n)}}(I_n).$$

Since $\mu^{(n)} = P_n \circ W_n^{-1}$ by Lemma 4.2.1, it follows that $H_{\mu^{(n)}}(I_n) = H_{P_n}(\mathcal{W}_n)$. $\square$

4.2.2 Properties of entropy and conditional entropy

In the proof of Theorems 4.1.1 and 4.1.2, we use properties of conditional entropy. For the reader's convenience, we state here the key definitions and properties that we require.

Definition 4.2.3. Let $(X, \mathcal{B}, \mu)$ be a probability space and let η and ξ be measurable partitions of X. We write $H_\mu(\xi)$ for the entropy of the partition ξ. For each set $A \in \mathcal{B}$, let $\mu_A\colon B \mapsto \mu(A \cap B)/\mu(A)$ denote the normalised restriction of μ to the set A. The entropy of ξ given η is the quantity defined by

$$H_\mu(\xi \,|\, \eta) = -\sum_{A\in\eta} \mu(A) H_{\mu_A}(\xi).$$

We will mostly be concerned with partitions that are generated by some random map $\mathcal{U}$, that is, a measurable function on a probability space $(\Omega, \mathcal{F}, \mu)$. More precisely, if $\mathcal{U}$ has a finite image $\mathrm{Im}(\mathcal{U})$, then it generates the partition

$$\xi(\mathcal{U}) = \{\mathcal{U}^{-1}(u) : u \in \mathrm{Im}(\mathcal{U})\}.$$

To avoid heavy notation, in such situations we set

$$H_\mu(\mathcal{U}) := H_\mu(\xi(\mathcal{U})).$$

If we are dealing with two such random maps $\mathcal{U}$ and V, we set

$$H_\mu(\mathcal{U}, \mathcal{V}) := H_\mu(\xi((\mathcal{U}, \mathcal{V}))),$$

where

$$\xi((\mathcal{U}, \mathcal{V})) = \xi(\mathcal{U}) \vee \xi(\mathcal{V}) = \{A \cap B : A \in \xi(\mathcal{U}), B \in \xi(\mathcal{V})\}$$

is the common refinement of the partitions generated by $\mathcal{U}$ and $\mathcal{V}$. Conditional entropies are defined accordingly. Namely, $H_\mu(\mathcal{U} \,|\, \mathcal{V}) = H_\mu(\xi(\mathcal{U}) \,|\, \xi(\mathcal{V}))$, $H_\mu(\mathcal{U}, \mathcal{V} \,|\, \mathcal{W}) = H_\mu(\xi(\mathcal{U}) \vee \xi(\mathcal{V}) \,|\, \xi(\mathcal{W}))$ and $H_\mu(\mathcal{U} \,|\, \mathcal{V}, \mathcal{W}) = H_\mu(\xi(\mathcal{U}) \,|\, \xi(\mathcal{V}) \vee \xi(\mathcal{W}))$, where $\mathcal{U}, \mathcal{V}$ and $\mathcal{W}$ are random maps on $(\Omega, \mathcal{F}, \mu)$.

In the proof of Theorems 4.1.1 and 4.1.2, we utilise several standard properties of entropy and conditional entropy, which we summarise in the following. For more further details, we refer the reader to [75, Ch. 4].

Lemma 4.2.4. Let $\mathcal{U}, \mathcal{V}$ and $\mathcal{W}$ be (measurable) random maps with finite image as above. Then, the following hold:

(1) $H_\mu(\mathcal{U}) \leq \log(\#\mathrm{Im}(\mathcal{U}))$, with equality precisely if $\mu \circ \mathcal{U}^{-1}$ is equidistributed;

(2) $H_\mu(\mathcal{U}) \leq H_\mu(\mathcal{U}, \mathcal{V})$, with equality if and only if $\xi(\mathcal{U})$ is a refinement, up to nullsets, of $\xi(\mathcal{V})$;

(3) $H_\mu(\mathcal{U}, \mathcal{V}) = H_\mu(\mathcal{V}) + H_\mu(\mathcal{U} \,|\, \mathcal{V})$;

(4) $H_\mu(\mathcal{U} \,|\, \mathcal{V}) \leq H_\mu(\mathcal{W} \,|\, \mathcal{V}) + H(\mathcal{U} \,|\, \mathcal{W})$;

(5) $H_\mu(\mathcal{U} \,|\, \mathcal{V}) \leq H_\mu(\mathcal{U})$, with equality if and only if $\mathcal{U}$ and $\mathcal{V}$ are independent;

(6) $H_\mu(\mathcal{U} \,|\, \mathcal{V}, \mathcal{W}) \leq H_\mu(\mathcal{U} \,|\, \mathcal{V})$;

(7) $H_\mu(\mathcal{U}, \mathcal{V} \,|\, \mathcal{W}) = H_\mu(\mathcal{U} \,|\, \mathcal{W}) + H_\mu(\mathcal{V} \,|\, \mathcal{U}, \mathcal{W})$.

4.2.3 Control over large deviations

Recall that the Perron–Frobenius eigenvalue λ of a primitive random substitution $\vartheta_{\mathbf{P}}$ can be regarded as an inflation factor. In the case that $\vartheta_{\mathbf{P}}$ is of constant length ℓ, this interpretation is exact in the sense that, for all $v \in \mathcal{A}^+$, every realisation of $\vartheta_{\mathbf{P}}(v)$ has length $\ell|v|$. If $\vartheta_{\mathbf{P}}$ is compatible, then it still holds that every realisation of $\vartheta_{\mathbf{P}}(v)$ has the same length, but this length might deviate slightly from $\lambda|v|$. However, we still obtain that λ is arbitrarily close to the ratio $|\vartheta(v)|/|v|$ for sufficiently long legal words v. In general, such a strong statement does not hold if we drop the assumption of compatibility. However, the probability that $|\vartheta_{\mathbf{P}}(v)|$ deviates by a positive fraction from $\lambda|v|$ decays quickly with the length of v for typical choices of v. We make this more precise in the following.

Lemma 4.2.5. Let $\lambda_- < \lambda < \lambda_+$ and let $(m_n)_n$ be a sequence of integers such that $m < n$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} m_n = \infty$. Further, let

$$A_n = \{(v, u_1, \ldots, u_n, j) \in \Omega_n : \lambda_- m_n \leq |u_2 \cdots u_{m_n}| \leq \lambda_+ m_n\}$$

for all $n \in \mathbb{N}$. Then, $\lim_{n\to\infty} P_n(A_n) = 1$.

Proof. Let $A_n^u := \{(u_2, \ldots, u_{m_n}) \colon \lambda_- m_n \leq |u_2 \cdots u_{m_n}| \leq \lambda_+ m_n\}$ be the set of $(u_2, \ldots, u_{m_n})$-tuples that extend to elements in A_n. By definition of P_n and A_n,

$$\begin{aligned} P_n(A_n) &= \frac{1}{\lambda} \sum_{v_{[1,m_n]}} \mu_{\mathbf{P}}([v_{[1,m_n]}]) \sum_{u_1} |u_1| \mathbb{P}[\vartheta_{\mathbf{P}}(v_1) = u_1] \sum_{(u_2,\ldots,u_{m_n}) \in A_n^u} \prod_{i=1}^{m_n} \mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = u_i] \\ &= \frac{1}{\lambda} \sum_{v_{[1,m_n]}} \mu_{\mathbf{P}}([v_{[1,m_n]}]) \mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|] \, \mathbb{P}[\lambda_- m_n \leq |\vartheta_{\mathbf{P}}(v_2 \cdots v_{m_n})| \leq \lambda_+ m_n]. \end{aligned}$$

We show that for $\mu_{\mathbf{P}}$-almost every $x \in X_\vartheta$, we have

$$\lim_{n\to\infty} \mathbb{P}[\lambda_- m_n \leq |\vartheta_{\mathbf{P}}(x_2 \cdots x_{m_n})| \leq \lambda_+ m_n] = 1. \tag{4.2}$$

Since $\mu_{\mathbf{P}}$ is ergodic, Birkhoff's ergodic theorem (Theorem A.1.5) gives that for $\mu_{\mathbf{P}}$-almost every $x \in X_\vartheta$ and every $\delta > 0$, it holds that

$$m_n(R_a - \delta) \leq |x_{[2,m_n]}|_a \leq m_n(R_a + \delta),$$

for each $a \in \mathcal{A}$ and large enough $n \in \mathbb{N}$. In this case, it follows by Cramér's theorem (Theorem A.2.1) that for all $\delta' > 0$,

$$\sum_{i, v_i = a} |\vartheta_{\mathbf{P}}(v_i)| \leq (1 + \delta') m_n (R_a + \delta) \mathbb{E}[|\vartheta_{\mathbf{P}}(a)|], \tag{4.3}$$

up to a set $E = E(n, v, \delta, \delta')$ whose probability decays exponentially with n. By the definition of the substitution matrix M, we have

$$\mathbb{E}[|\vartheta_{\mathbf{P}}(a)|] = \sum_{b \in \mathcal{A}} \mathbb{E}[|\vartheta_{\mathbf{P}}(a)|_b] = \sum_{b \in \mathcal{A}} M_{ba}.$$

Summing over $a \in \mathcal{A}$ in (4.3), we obtain that

$$|\vartheta_{\mathbf{P}}(v_2 \cdots v_{m_n})| \leq m_n(1+\delta')\bigg(\sum_{a,b\in\mathcal{A}} M_{ba}R_a + \delta|\vartheta|\bigg) = m_n(1+\delta')(\lambda + \delta|\vartheta|),$$

up to an exponentially decaying probability. Choosing δ, δ' small enough, we get

$$|\vartheta_{\mathbf{P}}(v_2 \cdots v_{m_n})| \leq \lambda_+ m_n$$

in these cases. The estimate for the lower bound follows by analogous arguments. Hence, there exist $c = c(v) > 0$ and $n_0 = n_0(v)$ such that

$$\mathbb{P}[\lambda_- m_n \leq |\vartheta_{\mathbf{P}}(v_2 \cdots v_{m_n})| \leq \lambda_+ m_n] \geq 1 - \mathrm{e}^{-m_n c},$$

for all $n \geq n_0$. In particular, (4.2) holds $\mu_{\mathbf{P}}$-almost surely and it follows by the dominated convergence theorem that

$$\lim_{n\to\infty} P_n(A_n) = \frac{1}{\lambda}\int_{X_\vartheta} \mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]\,\mathrm{d}\mu_{\mathbf{P}}(v) = 1. \qquad \square$$

4.2.4 The upper bound

As a first step towards the proof of our main results, we establish the sequence of upper bounds for the measure theoretic entropy stated in Theorem 4.1.1 and Theorem 4.1.2. For ease of notation, we let $\varphi\colon [0,\infty) \to \mathbb{R}$ denote the function defined by $\varphi(0) = 0$ and $\varphi(x) = -x\log x$ for all $x \in (0,\infty)$.

Proposition 4.2.6. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then,

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq \frac{1}{\lambda^k - 1}\mathbf{H}_k \cdot \mathbf{R}$$

for all $k \in \mathbb{N}$.

Proof. It suffices to show the relation for $k = 1$, since for all $k \in \mathbb{N}$, $\mu_{\mathbf{P}}$ is the frequency measure corresponding to $\vartheta_{\mathbf{P}}^k$ and λ^k is the Perron–Frobenius eigenvalue of $\vartheta_{\mathbf{P}}^k$. By Lemma 4.2.2, it is possible to control $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ via the entropy of $\mathcal{W}_n$. We wish to refer to data in Ω_n via a set of appropriate random variables. To this end we define (or recall in the case of $\mathcal{V}_{[1,m]}$)

- $\mathcal{V}_{[1,m]}\colon (v, u_1 \cdots u_n, j) \mapsto v_{[1,m]}$ for all $1 \leq m \leq n$,
- $\mathcal{J}\colon (v, u_1, \ldots, u_n, j) \mapsto j$,
- $\mathcal{U}_k\colon (v, u_1 \cdots u_n, j) \mapsto u_k$ for all $1 \leq k \leq n$,
- $\mathcal{U}_{[k,\ell]} = (\mathcal{U}_k, \ldots, \mathcal{U}_\ell)$ for $1 \leq k \leq \ell \leq n$.

Also, recall that $\mathcal{W}_n$ is given by $(u_1 \cdots u_n)_{[j,j+n-1]}$. On average, the words u_k have length λ, and therefore, in typical situations, $\mathcal{W}_n$ in fact only depends on u_k with $1 \leq k \leq m(n)$, for some integer $m(n)$ which deviates from n/λ by at most a constant. This motivates the following notation. Fix $\varepsilon > 0$ and let $\lambda_- = \lambda - \varepsilon$. Further, let $n \in \mathbb{N}$ and

$$m = m_+(n) = \left\lceil \frac{n}{\lambda_-} \right\rceil.$$

As a first step, we bound the entropy by

$$H_{P_n}(\mathcal{W}_n) \leq H_{P_n}(\mathcal{U}_{[1,m]}, \mathcal{J}) + H_{P_n}(\mathcal{W}_n \,|\, \mathcal{U}_{[1,m]}, \mathcal{J}),$$

using properties (2) and (3) of Lemma 4.2.4. Setting

$$A_n = \{(v, u_1, \ldots, u_n, j) \in \Omega_n \mid |u_2 \cdots u_m| \geq n\},$$

we note that on A_n, $\mathcal{W}_n$ is given by $(u_1 \cdots u_m)_{[j,j+n-1]}$ and hence is completely determined by $\mathcal{U}_{[1,m]}$ and $\mathcal{J}$. On A_n^C, we can bound the (conditioned) entropy of $\mathcal{W}_n$ by

$$\log(\#\mathcal{L}_\vartheta^n) \leq n \log(\#\mathcal{A}),$$

using property (1) of Lemma 4.2.4. Combining these two observations, we obtain

$$H_{P_n}(\mathcal{W}_n \,|\, \mathcal{U}_{[1,m]}, \mathcal{J}) \leq P_n(A_n^C)\, n \log(\#\mathcal{A}).$$

By Lemma 4.2.5, the term $P_n(A_n^C)$ converges to 0 as $n \to \infty$, so we have

$$H_{P_n}(\mathcal{W}_n) \leq H_{P_n}(\mathcal{U}_{[1,m]}, \mathcal{J}) + o(n).$$

On the other hand, since both $\mathcal{J}$ and $\mathcal{U}_1$ have a bounded number of realisations,

$$H_{P_n}(\mathcal{U}_{[1,m]}, \mathcal{J}) = H_{P_n}(\mathcal{U}_{[2,m]}) + O(1).$$

Conditioning on $\mathcal{V}_{[1,m]}$, we therefore obtain

$$H_{P_n}(\mathcal{W}_n) \leq H_{P_n}(\mathcal{U}_{[2,m]}) + o(n) \leq H_{P_n}(\mathcal{V}_{[1,m]}) + H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]}) + o(n). \tag{4.4}$$

For the calculation of the entropy $H_{P_n}(\mathcal{V}_{[1,m]})$, recall from Remark 4.2.1 that

$$P_n(\mathcal{V}_{[1,m]} = v_{[1,m]}) = \frac{1}{\lambda}\mu_{\mathbf{P}}([v_{[1,m]}])\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]. \tag{4.5}$$

In the following, we show that the modification by the factor $\lambda^{-1}\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]$ is inessential for our purposes. To this end, we make use of the general observation that $\varphi(pq) = p\varphi(q) + q\varphi(p)$. For an arbitrary probability vector $(p_i)_{i\in I}$ and a finite sequence of real numbers $q = (q_i)_{i\in I}$, this implies that

$$\sum_{i\in I} \varphi(p_i q_i) \leqslant \max_{i\in I} \varphi(q_i) + \sum_{i\in I} q_i \varphi(p_i).$$

Using this for $I = \mathcal{L}_\vartheta^m$, and the probability vector with entries $\mu_{\mathbf{P}}([v_{[1,m]}])$, we obtain via (4.5) that

$$\begin{aligned} H_{P_n}(\mathcal{V}_{[1,m]}) &= \sum_{v_{[1,m]}\in\mathcal{L}_\vartheta^m} \varphi\left(\frac{1}{\lambda}\mu_{\mathbf{P}}([v_{[1,m]}])\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]\right) \\ &= \sum_{v_{[1,m]}\in\mathcal{L}_\vartheta^m} \frac{\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]}{\lambda}\varphi(\mu_{\mathbf{P}}([v_{[1,m]}])) + O(1). \end{aligned}$$

Recall that $m = m(n)$ implicitly depends on n and note that we can rewrite

$$\frac{1}{n}\sum_{v_{[1,m]}\in\mathcal{L}_\vartheta^m} \frac{\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]}{\lambda}\varphi(\mu_{\mathbf{P}}([v_{[1,m]}])) = \frac{m}{n}\int_{X_\vartheta} \frac{-\log(\mu_{\mathbf{P}}([v_{[1,m]}]))}{m}\,\frac{\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]}{\lambda}\,\mathrm{d}\mu_{\mathbf{P}}(v).$$

Due to the ergodicity of $\mu_{\mathbf{P}}$ and the Shannon–McMillan–Breiman theorem (Theorem A.1.9), we have that $-\log(\mu_{\mathbf{P}}([v_{[1,m]}]))/m$ converges to $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ in $L^1(X_\vartheta, \mu_{\mathbf{P}})$. It follows that the product with any uniformly bounded function g also converges in $L^1(X_\vartheta, \mu_{\mathbf{P}})$. Applying this to $g\colon v \mapsto \lambda^{-1}\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]$ yields

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} H_{P_n}(\mathcal{V}_{[1,m]}) &= \frac{1}{\lambda_-} h_{\mu_{\mathbf{P}}}(X_\vartheta) \sum_{v_1\in\mathcal{A}} \mu_{\mathbf{P}}([v_1]) \frac{\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]}{\lambda} \\ &= \frac{1}{\lambda_-} h_{\mu_{\mathbf{P}}}(X_\vartheta) \frac{1}{\lambda} \sum_{a,b\in\mathcal{A}} M_{ba} R_a = \frac{1}{\lambda_-} h_{\mu_{\mathbf{P}}}(X_\vartheta).\end{aligned} \tag{4.6}$$

We next turn to the calculation of the conditional entropy $H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]})$. Denoting by $P_{n,v_{[1,m]}}$ the normalised restriction of P_n to $\{\mathcal{V}_{[1,m]} = v_{[1,m]}\}$, we have

$$P_{n,v_{[1,m]}}[\mathcal{U}_{[2,m]} = (u_2, \cdots, u_m)] = \prod_{i=2}^{m} \mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = u_i],$$

and thereby

$$H_{P_{n,v_{[1,m]}}}(\mathcal{U}_{[2,m]}) = \sum_{i=2}^{m} H_{\mathbb{P}}(\vartheta_{\mathbf{P}}(v_i)) = \mathbf{H}_1 \cdot \Phi(v_{[2,m]}).$$

Using (4.5), this yields

$$H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]}) = \frac{1}{\lambda} \sum_{v_{[1,m]}\in\mathcal{L}_\vartheta^m} \mu_{\mathbf{P}}([v_{[1,m]}]) \mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|] \, \mathbf{H}_1 \cdot \Phi(v_{[2,m]}).$$

For the corresponding asymptotic behaviour we note that, by Birkhoff's ergodic theorem (Theorem A.1.5), $\Phi(v_{[2,m]})/m$ converges to $\mathbf{R}$ for $\mu_{\mathbf{P}}$-almost every v, so

$$\lim_{n\to\infty} \frac{1}{n} H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]}) = \frac{1}{\lambda_-} \mathbf{H}_1 \cdot \mathbf{R} \sum_{v_1\in\mathcal{A}} \mu_{\mathbf{P}}([v_1]) \frac{\mathbb{E}[|\vartheta_{\mathbf{P}}(v_1)|]}{\lambda} = \frac{1}{\lambda_-} \mathbf{H}_1 \cdot \mathbf{R}. \tag{4.7}$$

Hence, combining the contributions from (4.6) and (4.7), we obtain by (4.4) that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \lim_{n\to\infty} \frac{1}{n} H_{P_n}(\mathcal{W}_n) \le \frac{1}{\lambda_-} \big(h_{\mu_{\mathbf{P}}}(X_\vartheta) + \mathbf{H}_1 \cdot \mathbf{R}\big).$$

As $\varepsilon \to 0$, we have $\lambda_- \to \lambda$, so it follows that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) \le \frac{1}{\lambda - 1} \mathbf{H}_1 \cdot \mathbf{R},$$

which completes the proof. $\square$

The sequence of vectors $(\mathbf{H}_n)_{n\in\mathbb{N}}$ can be bounded via a matrix-recursion that involves the substitution matrix.

Proposition 4.2.7. Let $\vartheta_{\mathbf{P}}$ be a primitive random substitution. Then, for each $a \in \mathcal{A}$ and all $n, k \in \mathbb{N}$, we have that

$$H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k}(a)) \leq \mathbf{H}_n \cdot (M^k \mathbf{e}_a) + H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^k(a)),$$

where M is the matrix of $\vartheta_{\mathbf{P}}$ and $\mathbf{e}_a$ is the unit vector corresponding to the letter a. In particular,

$$\mathbf{H}_{n+k} \cdot \mathbf{R} \leq \lambda^k \mathbf{H}_n \cdot \mathbf{R} + \mathbf{H}_k \cdot \mathbf{R}.$$

If $\vartheta_{\mathbf{P}}$ has unique realisation paths, equality occurs precisely if $\vartheta_{\mathbf{P}}^n(a)$ is completely determined by $\vartheta_{\mathbf{P}}^{n+k}(a)$.

Proof. First, let $v \in \mathcal{L}_{\vartheta}^m$, for some $m \in \mathbb{N}$, and note that the random variable $\vartheta_{\mathbf{P}}^n(v)$ can be written as a function of $(\vartheta_{\mathbf{P}}^n(v_1), \cdots, \vartheta_{\mathbf{P}}^n(v_m))$. Due to the independence of the random variables in the last tuple, we obtain that

$$H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^n(v)) \leq H_{\mathbb{P}}\big(\vartheta_{\mathbf{P}}^n(v_1), \ldots, \vartheta_{\mathbf{P}}^n(v_m)\big) = \sum_{i=1}^{m} H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^n(v_i)) = \mathbf{H}_n \cdot \Phi(v).$$

If $\vartheta_{\mathbf{P}}$ has unique realisation paths, then we have equality. Using the Markov property of the substitution process in the first step, for each $a \in \mathcal{A}$, we have

$$\begin{aligned} H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k}(a) | \vartheta_{\mathbf{P}}^k(a)) &= \sum_{v \in \vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v] H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^n(v)) \leq \mathbf{H}_n \cdot \left(\sum_{v \in \vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v] \Phi(v) \right) \\ &= \mathbf{H}_n \cdot \mathbb{E}[\Phi(\vartheta_{\mathbf{P}}^k(a))] = \mathbf{H}_n \cdot (M^k \mathbf{e}_a), \end{aligned}$$

with equality if $\vartheta_{\mathbf{P}}$ has unique realisation paths. Therefore, for all $a \in \mathcal{A}$,

$$H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k}(a)) \leq H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k}(a) | \vartheta_{\mathbf{P}}^k(a)) + H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^k(a)) \leq \mathbf{H}_n^{\top} M^k \mathbf{e}_a + H_{k,a}.$$

The result now follows by summing over all $a \in \mathcal{A}$, weighted with respect to the entries of the right Perron–Frobenius eigenvector. The first inequality in the above is an equality precisely if $\vartheta_{\mathbf{P}}^{k}(a)$ is completely determined by $\vartheta_{\mathbf{P}}^{n+k}(a)$ and the second inequality is an equality provided that $\vartheta_{\mathbf{P}}$ has unique realisation paths. $\square$

Corollary 4.2.8. Let $\vartheta_{\mathbf{P}}$ be a primitive random substitution. Then, for all $n \in \mathbb{N}$,

$$\frac{1}{\lambda^n - 1}\mathbf{H}_n \cdot \mathbf{R} \leq \frac{1}{\lambda - 1}\mathbf{H}_1 \cdot \mathbf{R}.$$

If $\vartheta_{\mathbf{P}}$ has unique realisation paths, we have equality for all $n \in \mathbb{N}$ if and only if $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition.

Proof. Given $n \geq 2$, iterating the relation $\mathbf{H}_n \cdot \mathbf{R} \leq \lambda^{n-1}\mathbf{H}_1 \cdot \mathbf{R} + \mathbf{H}_{n-1} \cdot \mathbf{R}$ yields

$$\mathbf{H}_n \cdot \mathbf{R} \leq \mathbf{H}_1 \cdot \mathbf{R} \sum_{k=0}^{n-1} \lambda^k = \frac{\lambda^n - 1}{\lambda - 1}\mathbf{H}_1 \cdot \mathbf{R},$$

immediately giving the required inequality. Under the assumption of unique realisation paths, equality holds if and only if $\vartheta_{\mathbf{P}}^{n}(a)$ completely determines $\vartheta_{\mathbf{P}}(a)$ for all $a \in \mathcal{A}$ and $n \in \mathbb{N}$. This is just a reformulation of the disjoint set condition (compare Remark 2.2.1). $\square$

4.2.5 The lower bound

We now establish the lower bounds for the measure theoretic entropy in Theorem 4.1.1 and Theorem 4.1.2. Again, our proof relies heavily on the self-consistency relation for $\mu_{\mathbf{P}}$ presented in Section 4.2.1.

Proposition 4.2.9. Let $\vartheta_{\mathbf{P}}$ be a primitive random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then,

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) \geq \frac{1}{\lambda^k}\mathbf{H}_k \cdot \mathbf{R} - H(\lambda^{-k}),$$

for all $k \in \mathbb{N}$. If $\vartheta_{\mathbf{P}}$ has unique realisation paths, then

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) \geq \frac{1}{\lambda^k}\mathbf{H}_k \cdot \mathbf{R},$$

for all $k \in \mathbb{N}$.

Proof. Again, it suffices to consider the case $k = 1$. We use the same notation as the proof of Proposition 4.2.6, with one modification. For $\varepsilon > 0$, we now consider $\lambda_+ = \lambda + \varepsilon$ and set

$$m = m_-(n) = \left\lceil \frac{n}{\lambda_+} \right\rceil.$$

This is to ensure that $\mathcal{W}_n$ and $\mathcal{J}$ determine $\mathcal{U}_2 \cdots \mathcal{U}_m$ on a set of large probability, given by

$$B_n = \{(v, u_1, \ldots, u_n, j) : |u_2 \cdots u_m| \leq n - |\vartheta|\},$$

where we recall that $|\vartheta| = \max_{a \in \mathcal{A}} \max_{v \in \vartheta(a)} |v|$. By properties (4) and (5) of Lemma 4.2.4, we have

$$H_{P_n}(\mathcal{W}_n) \geq H_{P_n}(\mathcal{W}_n \,|\, \mathcal{V}_{[1,m]}) \geq H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]}) - H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n). \tag{4.8}$$

and it follows by analogous arguments to those in the proof of Proposition 4.2.6 that

$$\lim_{n \to \infty} \frac{1}{n} H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]}) = \frac{1}{\lambda_+} \mathbf{H}_1 \cdot \mathbf{R}.$$

It remains to find an adequate upper bound for $H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n)$. To that end, we introduce an additional random variable on Ω_n, namely,

$$\ell_m \colon (v, u_1, \ldots, u_n, j) \mapsto |u_2 \cdots u_m|.$$

It then follows by properties (6) and (7) of Lemma 4.2.4 that

$$\begin{aligned} H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n) &\leq H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n, \mathcal{J}, \ell_m) + H_{P_n}(\mathcal{J}, \ell_m \,|\mathcal{W}_n) \\ &= H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n, \mathcal{J}, \ell_m) + O(\log(m)). \end{aligned} \tag{4.9}$$

The last step follows because the number of distinct realisations of $(\mathcal{J}, \ell_m)$ can be bounded from above by $|\vartheta|^2 m$. Conditioned on $\mathcal{W}_n, \mathcal{J}, \ell_m$, and provided $\ell_m \leq n - |\vartheta|$, knowledge of $\mathcal{U}_{[2,m]}$ is equivalent to knowledge of

$$|\mathcal{U}|_{[2,m]} \colon (v, u_1, \ldots, u_n, j) \mapsto (|u_2|, \ldots, |u_m|).$$

Indeed, on the set B_n (that is, if $\ell_m \leq n - |\vartheta|$) we observe that $\mathcal{W}_n, \mathcal{J}, \ell_m$ determines the word

$u_2 \cdots u_m$, such that knowing the lengths of the individual words allows us to infer $(u_2, \ldots, u_m)$. By conditioning,

$$H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n, \mathcal{J}, \ell_m) \leq H_{P_n}(|\mathcal{U}|_{[2,m]} \,|\, \mathcal{W}_n, \mathcal{J}, \ell_m) + H_{P_n}(\mathcal{U}_{[2,m]} \,|\, |\mathcal{U}|_{[2,m]}, \mathcal{W}_n, \mathcal{J}, \ell_m).$$

Let $M = \max_{a \in \mathcal{A}} \#\vartheta(a)$, implying $\#\mathrm{Im}(\mathcal{U}_{[2,m]}) \leq M^m$. By the observations above, we can bound

$$H_{P_n}(\mathcal{U}_{[2,m]} \,|\, |\mathcal{U}|_{[2,m]}, \mathcal{W}_n, \mathcal{J}, \ell_m) \leq P_n(B_n^C)\, m \log(M).$$

Since $P(B_n^C) \to 0$ as $n \to \infty$ by Lemma 4.2.5, it follows that

$$H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n, \mathcal{J}, \ell_m) \leq H_{P_n}(|\mathcal{U}|_{[2,m]} \,|\, \ell_m) + o(n). \tag{4.10}$$

If $\vartheta_{\mathbf{P}}$ has unique realisation paths, then $\mathcal{W}_n, \mathcal{J}, \ell_m$ determines $\mathcal{U}_{[2,m]}$ completely on B_n, yielding

$$H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n, \mathcal{J}, \ell_m) = o(n),$$

by an analogous argument. Given $\ell_m = \ell$, the number of possible values of $|U|_{[2,m]}$ is bounded above by the number of choices to decompose a block of length ℓ into $m-1$ smaller blocks, that is, by the binomial coefficient $\binom{\ell-1}{m-2}$. Using this bound on B_n and the fixed bound M^m on B_n^C, we obtain

$$\begin{aligned} H_{P_n}(|\mathcal{U}|_{[2,m]} \,|\, \ell_m) &\leq \sum_{\ell=m-1}^{n-|\vartheta|} P_n[\ell_m = \ell] \log \binom{\ell-1}{m-2} + P_n(B_n^C)\, m \log(M) \\ &\leq \log \binom{n}{m-2} + o(n) \leq n\, H((m-2)/n) + o(n). \end{aligned}$$

Since we have seen in (4.9) and (4.10) that $H_{P_n}(|\mathcal{U}|_{[2,m]} \,|\, \ell_m)$ bounds $H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n)$ up to a term of order $o(n)$, we obtain from (4.8) that

$$\begin{aligned} h_{\mu_{\mathbf{P}}}(X_\vartheta) = \lim_{n\to\infty} \frac{1}{n} H_{P_n}(\mathcal{W}_n) &\geq \lim_{n\to\infty} \frac{1}{n} H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{V}_{[1,m]}) - \limsup_{n\to\infty} H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n) \\ &\geq \frac{1}{\lambda_+} \mathbf{H}_1 \cdot \mathbf{R} - H(\lambda_+^{-1}) \xrightarrow{\varepsilon\to 0} \frac{1}{\lambda} \mathbf{H}_1 \cdot \mathbf{R} - H(\lambda^{-1}). \end{aligned}$$

If $\vartheta_{\mathbf{P}}$ has unique realisation paths, then $H_{P_n}(\mathcal{U}_{[2,m]} \,|\, \mathcal{W}_n) = o(n)$, which gives the stronger bound

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) \geq \frac{1}{\lambda}\mathbf{H}_1 \cdot \mathbf{R}.$$

This completes the proof. □

For the remainder of this section, we restrict to the case of unique realisation paths.

Proposition 4.2.10. Let $\vartheta_{\mathbf{P}}$ be a primitive random substitution with unique realisation paths. Then, for each $a \in \mathcal{A}$, we have

$$H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k}(a)) \geq \mathbf{H}_n \cdot (M^k \mathbf{e}_a)$$

for all $n, k \in \mathbb{N}$, where $\mathbf{e}_a$ is the unit vector corresponding to a. Equality holds if and only if $\vartheta_{\mathbf{P}}^{n+k}(a)$ is independent of $\vartheta_{\mathbf{P}}^{n}(a)$ for all $a \in \mathcal{A}$.

Proof. In the proof of Proposition 4.2.7, we showed that $H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k} \,|\, \vartheta_{\mathbf{P}}^{n}(a)) = \mathbf{H}_n \cdot (M^k \mathbf{e}_a)$ for each $a \in \mathcal{A}$. Thus, we have

$$H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k}(a)) \geq H_{\mathbb{P}}(\vartheta_{\mathbf{P}}^{n+k} \,|\, \vartheta_{\mathbf{P}}^{n}(a)) = \mathbf{H}_n \cdot (M^k \mathbf{e}_a),$$

for each $a \in \mathcal{A}$, and equality holds if and only if $\vartheta_{\mathbf{P}}^{n+k}(a)$ and $\vartheta_{\mathbf{P}}^{n}(a)$ are independent random variables. □

Corollary 4.2.11. Let $\vartheta_{\mathbf{P}}$ be a primitive random substitution with unique realisation paths. Then, for all $m \leq n$,

$$\frac{1}{\lambda^m}\mathbf{H}_m \cdot \mathbf{R} \leq \frac{1}{\lambda^n}\mathbf{H}_n \cdot \mathbf{R}.$$

Equality holds for all $m \leq n$ if and only if $\vartheta_{\mathbf{P}}$ satisfies the identical set condition with identical production probabilities.

Proof. It follows by Proposition 4.2.10 that

$$\frac{1}{\lambda^n}\mathbf{H}_n \cdot \mathbf{R} \geq \frac{1}{\lambda^n}\mathbf{H}_m \cdot (M^{n-m}\mathbf{R}) = \frac{1}{\lambda^m}\mathbf{H}_m \cdot \mathbf{R}.$$

Equality for all $m \leq n$ holds precisely if

$$\frac{1}{\lambda^n}\mathbf{H}_n \cdot \mathbf{R} = \frac{1}{\lambda}\mathbf{H}_1 \cdot \mathbf{R},$$

for all $n \in \mathbb{N}$. This is the case if and only if for all $a \in \mathcal{A}$, $\vartheta_{\mathbf{P}}(a)$ is independent from $\vartheta_{\mathbf{P}}^n(a)$ for all $n \in \mathbb{N}$, which means that $\vartheta_{\mathbf{P}}^{n-1}(v)$ has the same distribution for all possible realisations v of $\vartheta_{\mathbf{P}}(a)$. This is precisely the identical set condition with identical production probabilities (compare Remark 2.2.1). $\square$

4.2.6 Proof of main results

Our main results now follow in a straightforward manner from the results we have already established.

Proof of Theorem 4.1.1. The fact that $\lambda^{-m}\mathbf{H}_m \cdot \mathbf{R} - H(\lambda^{-m}) \leq h_{\mu_{\mathbf{P}}}(X_\vartheta) \leq (\lambda^m - 1)^{-1}\mathbf{H}_m \cdot \mathbf{R}$ for all $m \in \mathbb{N}$ follows directly by combining Proposition 4.2.6 and Proposition 4.2.9. The convergence of $\lambda^{-m}\mathbf{H}_m \cdot \mathbf{R}$ as $m \to \infty$ can be seen from the reformulation of this relation in (4.1). $\square$

Proof of Theorem 4.1.2. The upper and lower bounds for $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ were established in Proposition 4.2.6 and Proposition 4.2.9. The statements on the equivalent conditions for equality with the lower or upper bound are given in Corollary 4.2.8 and Corollary 4.2.11. The fact that the sequence of lower bounds is non-decreasing is also contained in Corollary 4.2.11. $\square$

CHAPTER 5

MEASURES OF MAXIMAL ENTROPY

Recall that a measure μ supported on a subshift X is called a *measure of maximal entropy* if $h_\mu(X) = h_{\text{top}}(X)$ and that for every subshift there always exists at least one measure of maximal entropy. In this chapter, we combine the theory developed in Chapters 3 and 4 on topological and measure theoretic entropy to ascertain conditions under which a primitive random substitution gives rise to a frequency measure of maximal entropy. This allows us to show that there exists a frequency measure of maximal entropy for a broad class of random substitution subshifts. Further, in a more general setting, we show that a measure of maximal entropy can be constructed as the weak*-limit of a sequence of frequency measures.

Recall that we say a subshift is *intrinsically ergodic* if it has a unique measure of maximal entropy. We prove intrinsic ergodicity for several families of random substitution subshifts and show that the unique measure of maximal entropy is a frequency measure. These subshifts often do not satisfy the specification property, the prototypical method for verifying intrinsic ergodicity, thus provide an interesting new class of intrinsically ergodic subshifts.

The results in this chapter are largely based on [37, Section 4]. However, we highlight that some of the results in Section 5.1 on the existence of frequency measures of maximal entropy are slightly more general than those presented in [37], since we apply results from Chapter 3 on topological entropy which were proved after the paper [37] was published. Nonetheless, the proofs are very similar to those in [37]. Also, our proof that the subshift associated with the random period doubling substitution is intrinsically ergodic is new.

5.1 Frequency measures of maximal entropy

5.1.1 Existence of frequency measures of maximal entropy

Using our results on topological and measure theoretic entropy, we can show that for many random substitution subshifts, there exists a frequency measure of maximal entropy. As a guiding example, let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random period doubling substitution

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1, \end{cases}$$

for some $p \in (0,1)$, and let $\mu_{\mathbf{P}}$ be the corresponding frequency measure. Since $\vartheta_{\mathbf{P}}$ is primitive and compatible, both topological entropy and measure theoretic entropy of the associated subshift coincide with the inflation word analogues introduced in Chapters 3 and 4. Moreover, since $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition, closed-form formulae can be obtained for each. Specifically, in Example 3.2.5, we showed that

$$h_{\mathrm{top}}(X_\vartheta) = \frac{2}{3} \log 2$$

and, in Example 4.1.6, we showed that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = -\frac{2}{3} \left(p \log p + (1-p) \log(1-p) \right).$$

Observe that when $p = 1/2$, $h_{\mu_{\mathbf{P}}}(X_\vartheta)$ and $h_{\mathrm{top}}(X_\vartheta)$ coincide. Hence, the frequency measure corresponding to $p = 1/2$ is a measure of maximal entropy. More generally, by combining the theory established so far for measure theoretic entropy and topological entropy, we can show that there exists a frequency measure of maximal entropy for many random substitution subshifts. In the following, we provide sufficient conditions under which this occurs.

Theorem 5.1.1. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution satisfying either the identical set condition or disjoint set condition, for which $\mathbb{P}[\vartheta_{\mathbf{P}}(a) = v] = 1/(\#\vartheta(a))$ for all $a \in \mathcal{A}$ and $v \in \vartheta(a)$. Further, assume that at least one of the following holds:

- $\vartheta_{\mathbf{P}}$ is compatible;

- $\vartheta_{\mathbf{P}}$ is constant length and there exists an $N \in \mathbb{N}$ such that $\#\vartheta(a) = N$ for all $a \in \mathcal{A}$.

Then, the corresponding frequency measure $\mu_{\mathbf{P}}$ is a measure of maximal entropy for the subshift X_ϑ.

Proof. For all $a \in \mathcal{A}$ and $v \in \vartheta(a)$, we have that $\mathbb{P}[\vartheta_{\mathbf{P}}(a) = v] = 1/(\#\vartheta(a))$; hence,

$$\mathbf{H}_1 \cdot \mathbf{R} = \sum_{a \in \mathcal{A}} R_a \log(\#\vartheta(a)).$$

If $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition, then by Theorem 4.1.2 we have

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \frac{1}{\lambda - 1} \sum_{a \in \mathcal{A}} R_a \log(\#\vartheta(a)).$$

Thus, if $\vartheta_{\mathbf{P}}$ is compatible, it follows by Proposition 3.2.3 that $h_{\mu_{\mathbf{P}}}(X_\vartheta) = h_{\text{top}}(X_\vartheta)$, and so $\mu_{\mathbf{P}}$ is a measure of maximal entropy. On the other hand, if $\vartheta_{\mathbf{P}}$ is constant length and $\#\vartheta(a) = N$ for all $a \in \mathcal{A}$, then we have $h_{\mu_{\mathbf{P}}}(X_\vartheta) = (\lambda - 1)^{-1} \log N$, which is equal to $h_{\text{top}}(X_\vartheta)$ by Proposition 3.2.11.

Now assume that $\vartheta_{\mathbf{P}}$ satisfies the identical set condition. Before we can apply Theorem 4.1.2, we first need to verify that $\vartheta_{\mathbf{P}}$ has identical production probabilities. To this end, let $a \in \mathcal{A}$, and $u, v \in \vartheta(a)$. Since $\vartheta_{\mathbf{P}}$ is compatible, $|u|_b = |v|_b$ for all $b \in \mathcal{A}$. Hence, if $t \in \vartheta^2(a)$, it follows that

$$\mathbb{P}[\vartheta_{\mathbf{P}}(u) = t] = \prod_{b \in \mathcal{A}} (\#\vartheta(b))^{-|u|_b} = \prod_{b \in \mathcal{A}} (\#\vartheta(b))^{-|v|_b} = \mathbb{P}[\vartheta_{\mathbf{P}}(v) = t].$$

By way of induction, let $n \in \mathbb{N}$ and assume that $\mathbb{P}[\vartheta_{\mathbf{P}}^{n-1}(u) = w] = [\vartheta_{\mathbf{P}}^{n-1}(v) = w]$ for all $w \in \vartheta^n(a)$. Since $\vartheta_{\mathbf{P}}$ satisfies the identical set condition, for all $t \in \vartheta^{n+1}(a)$ we have $t \in \vartheta^n(u) \cap \vartheta^n(v)$, so

$$\begin{aligned}
\mathbb{P}[\vartheta_{\mathbf{P}}^n(u) = t] &= \sum_{w \in \vartheta^{n-1}(u)} \mathbb{P}[\vartheta_{\mathbf{P}}^{n-1}(u) = w] \, \mathbb{P}[\vartheta_{\mathbf{P}}(w) = t] \\
&= \sum_{w \in \vartheta^{n-1}(v)} \mathbb{P}[\vartheta_{\mathbf{P}}^{n-1}(v) = w] \, \mathbb{P}[\vartheta_{\mathbf{P}}(w) = t] = \mathbb{P}[\vartheta_{\mathbf{P}}^n(v) = t].
\end{aligned}$$

Therefore, by induction, $\vartheta_{\mathbf{P}}$ has identical production probabilities, and thus by Theorem 4.1.2 we have

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = \frac{1}{\lambda} \sum_{a \in \mathcal{A}} R_a \log(\#\vartheta(a)).$$

Comparing this with the expressions for topological entropy given by Propositions 3.2.3 and 3.2.11, we conclude that $h_{\mu_{\mathbf{P}}}(X_\vartheta) = h_{\text{top}}(X_\vartheta)$. Namely, $\mu_{\mathbf{P}}$ is a measure of maximal entropy. $\square$

For the random Fibonacci substitution, the conclusion of Theorem 5.1.1 does *not* hold. Thus, the assumption that either the identical set condition or disjoint set condition is satisfied cannot be dropped in general.

Example 5.1.2. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random Fibonacci substitution with uniform probabilities,

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } 1/2, \\ ba & \text{with probability } 1/2, \end{cases} \\ b \mapsto a \quad \text{with probability } 1, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. In Example 4.1.9, we showed that the measure theoretic entropy of the subshift X_ϑ with respect to the measure $\mu_{\mathbf{P}}$ satisfies the following bounds:

$$0.4164 < h_{\mu_{\mathbf{P}}}(X_\vartheta) < 0.4314.$$

However, $h_{\text{top}}(X_\vartheta) = \sum_{m=2}^{\infty} \log(m)/\tau^{m+2} \approx 0.4444$, so $\mu_{\mathbf{P}}$ is not a measure of maximal entropy.

The following example shows that the assumption that the cardinalities of inflation sets coincide cannot be dropped for constant length random substitutions. Further, it illustrates that the frequency measure of greatest entropy may not correspond to uniform probabilities in the non-compatible setting.

Example 5.1.3. Let $p \in (0,1)$ and let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} aa & \text{with probability } p, \\ ab & \text{with probability } 1-p, \end{cases} \\ b \mapsto ba \quad \text{with probability } 1, \end{cases}$$

with corresponding frequency measure $\mu_{\mathbf{P}}$ and subshift X_ϑ. Recall that $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition. In Example 3.2.10 we showed that the topological entropy of the subshift X_ϑ is

$$h_{\text{top}}(X_\vartheta) = \sum_{n=1}^{\infty} \frac{1}{2^n} \log n \approx 0.507834.$$

Further, in Example 4.1.7, we showed that

$$h_{\mu_{\mathbf{P}}}(X_\vartheta) = -\frac{1}{2-p}(p \log p + (1-p)\log(1-p)),$$

which is maximised when $p = \tau^{-1}$, where τ denotes the golden ratio. Letting $\mu_{\max}$ denote the measure corresponding to this value of p, we have

$$h_{\mu_{\max}}(X_\vartheta) = \log \tau \approx 0.481212.$$

Hence, $\mu_{\max}$ is not a measure of maximal entropy for X_ϑ.

5.1.2 Weak*-limits of sequences of frequency measures

Examples 5.1.2 and 5.1.3 demonstrate that neither constant length nor compatibility are sufficient to guarantee the frequency measure corresponding to uniform probabilities is a measure of maximal entropy. However, in both of these settings, a measure of maximal entropy can be obtained as the weak*-limit of a sequence of frequency measures. In particular, the measures in this sequence can be taken to be frequency measures corresponding to powers of the random substitution with uniform probabilities.

Theorem 5.1.4. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution. Further, assume that $\vartheta_{\mathbf{P}}$ is compatible or constant length. Then, there exists a sequence of frequency measures $(\mu_n)_n$ such that μ_n converges weak* to a measure of maximal entropy for the system subshift X_ϑ.

Proof. For each $n \in \mathbb{N}$, let $\mathbf{P}_n$ denote the family of probability vectors corresponding to uniform probabilities on ϑ^n. Since the subshift of a random substitution is independent of the choice of probabilities, the random substitution $(\vartheta^n, \mathbf{P}_n)$ gives rise to the subshift X_ϑ. Let μ_n denote the frequency measure corresponding to the random substitution $(\vartheta^n, \mathbf{P}_n)$. Since the space of shift-invariant probability measures on X_ϑ is weak*-compact, there exists a shift-invariant probability measure μ and a sequence $(n_k)_k$ of natural numbers such that $(\mu_{n_k})_{k\in\mathbb{N}}$ converges weak* to μ. As we have assumed that $\vartheta_{\mathbf{P}}$ is compatible or constant length, $\vartheta_{\mathbf{P}}$ has unique realisation paths. Thus, by Theorem 4.1.2 we have

$$h_{\mu_{n_k}}(X_\vartheta) \geq \frac{1}{\lambda^{n_k}} \sum_{a \in \mathcal{A}} R_a \log(\# \vartheta^{n_k}(a)) \tag{5.1}$$

for all $k \in \mathbb{N}$. If $\vartheta_{\mathbf{P}}$ is compatible, then the right hand side converges to $h_{\text{top}}(X_\vartheta)$ by Proposition 3.2.2. On the other hand, if $\vartheta_{\mathbf{P}}$ is of constant length ℓ, then

$$\frac{1}{\lambda^{n_k}} \sum_{a \in \mathcal{A}} R_a \log(\# \vartheta^{n_k}(a)) = \sum_{a \in \mathcal{A}} R_a \frac{\log(\# \vartheta^{n_k}(a))}{\ell^{n_k}}$$

for all k, which converges to $h_{\text{top}}(X_\vartheta)$ by Theorem 3.2.8. Hence, in either case, we have

$$\lim_{k \to \infty} h_{\mu_{n_k}}(X_\vartheta) \geq h_{\text{top}}(X_\vartheta),$$

and so it follows by the upper semi-continuity of measure theoretic entropy that $h_\mu(X_\vartheta) = h_{\text{top}}(X_\vartheta)$. □

Notice that in the proof of Theorem 5.1.4, the only place we have used compatibility or the constant length property is in establishing the coincidence of topological entropy with the quantity in (5.1). As such, we expect the full strength of compatibility or the constant length property are not required for the conclusion of Theorem 5.1.4 to hold.

5.2 Intrinsic ergodicity of random substitution subshifts

Theorem 5.1.1 gives that for a broad class of random substitution subshifts, there exists a frequency measure of maximal entropy. It is natural to enquire whether this measure is unique as a measure of maximal entropy. In this section, we show that this is indeed the case for many random substitutions satisfying the conditions of Theorem 5.1.1.

5.2.1 Techniques for proving intrinsic ergodicity

A common technique for proving a subshift is intrinsically ergodic is to show that there exists a measure of maximal entropy μ for which there are constants $A, B > 0$ such that

$$A\mathrm{e}^{-|u|h} \leq \mu([u]) \leq B\mathrm{e}^{-|u|h} \tag{5.2}$$

for every legal word u, where h denotes the topological entropy of the subshift. Such a uniform restriction on the scaling of cylinder sets is often called a *Gibbs property.* This technique was first used by Adler and Weiss [2] to show that the Parry measure is the unique measure of maximal entropy for every irreducible shift of finite type (although we note that Parry [63] previously showed that this measure is a measure of maximal entropy). Generally, verifying the property (5.2) holds can be difficult. However, Bowen [10] introduced a sufficient condition, called the *specification property*, under which this property holds, which is typically easier to verify than proving the bounds in (5.2) directly. Verifying the specification property has become a standard technique for proving intrinsic ergodicity of subshifts and, in recent years, several weaker versions of this property have been introduced, which are still sufficient to establish intrinsic ergodicity. For example, Climenhaga and Thompson [13, 14] introduced a weak specification property that holds for all subshift factors of β-shifts and S-gap shifts, for which Bowen's specification property is not satisfied, thus establishing intrinsic ergodicity for a broad class of subshifts. However, their approach still relies on the existence of a measure of maximal entropy that satisfies the property (5.2).

5.2.2 A class of random substitutions satisfying the identical set condition

For a class of subshifts arising from constant length random substitutions satisfying the identical set condition, we can show that the measure of maximal entropy given by Theorem 5.1.1 satisfies the Gibbs property in (5.2), thus is the unique measure of maximal entropy. These subshifts were shown to be intrinsically ergodic by Gohlke and Spindeler [39]. In their proof, they showed that these subshifts are coded shifts, which are well known to satisfy Bowen's specification property. Here, we present an alternative proof, which does not rely on specification, but instead uses the renormalisation lemma (Lemma 2.3.5) to directly verify the Gibbs property (5.2).

Proposition 5.2.1. Let $\mathcal{A}$ be a finite alphabet, $\ell \geq 2$ and $s^1, \ldots, s^k \in \mathcal{A}^+$ be distinct words of length ℓ, such that for every letter $a \in \mathcal{A}$ there exists an $i \in \{1, \ldots, k\}$ such that a appears in s^i. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon a \mapsto \begin{cases} s^1 & \text{with probability } 1/k, \\ \vdots & \\ s^k & \text{with probability } 1/k, \end{cases}$$

for all $a \in \mathcal{A}$ and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Then, the measure $\mu_{\mathbf{P}}$ satisfies the Gibbs property (5.2). Thus, the subshift X_ϑ is intrinsically ergodic and $\mu_{\mathbf{P}}$ is the unique measure of maximal entropy.

Proof. First observe that, since for any $a \in \mathcal{A}$, every letter in $\mathcal{A}$ appears in a realisation of $\vartheta(a)$, the random substitution $\vartheta_{\mathbf{P}}$ is primitive; hence, $\mathcal{L}(X_\vartheta) = \mathcal{L}_\vartheta$. Let $u \in \mathcal{L}_\vartheta$. We prove the Gibbs property (5.2) by application of the renormalisation lemma. Since $\vartheta_{\mathbf{P}}$ is constant length, Lemma 2.3.5 gives that

$$\mu_{\mathbf{P}}([u]) = \frac{1}{\ell} \sum_{v \in \mathcal{L}_\vartheta^{\lceil |u|/\ell \rceil + 1}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{\ell} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j, j+|u|+1]} = u]. \tag{5.3}$$

We have that $\vartheta(a) = \vartheta(b)$ for all $a, b \in \mathcal{A}$ and $\mathbb{P}[\vartheta_{\mathbf{P}}(a) = w] = 1/k$ for all $w \in \vartheta(a)$. Hence, it follows by the Markov property of $\vartheta_{\mathbf{P}}$ that, for all $v \in \mathcal{L}_\vartheta^{\lceil |u|/\ell \rceil + 1}$ and $j \in \{1, \ldots, \ell\}$, if $u \in \vartheta(v)_{[j, j+|u|-1]}$, then

$$\left(\frac{1}{k} \right)^{\lceil |u|/\ell \rceil + 1} \leq \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j, j+|u|-1]} = u] \leq \left(\frac{1}{k} \right)^{\lceil |u|/\ell \rceil - 1}.$$

Otherwise, if $u \notin \vartheta(v)_{[j, j+|u|-1]}$, then $\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j, j+|u|-1]} = u] = 0$. Since u is legal, there exists at least one $j \in \{1, \ldots, \ell\}$ such that $u \in \vartheta(v)_{[j, j+|u|-1]}$, so it follows by (5.3) that

$$\frac{1}{\ell k} \mathrm{e}^{-|u| h_{\mathrm{top}}(X_\vartheta)} = \frac{1}{\ell k} \mathrm{e}^{-|u| \ell^{-1} \log k} \leq \mu_{\mathbf{P}}([v]) \leq k \mathrm{e}^{-|u| \ell^{-1} \log k} = k \mathrm{e}^{-|u| h_{\mathrm{top}}(X_\vartheta)},$$

noting that $\sum_{v \in \mathcal{L}_\vartheta^{\lceil |u|/\ell \rceil + 1}} \mu_{\mathbf{P}}([v]) = 1$ and $h_{\mathrm{top}}(X_\vartheta) = \ell^{-1} \log k$ by Proposition 3.2.11. Thus, we conclude that the Gibbs property (5.2) holds, so the subshift X_ϑ is intrinsically ergodic. $\square$

Example 5.2.2. Let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon a, b \mapsto \begin{cases} ab & \text{with probability } 1/2, \\ ba & \text{with probability } 1/2, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Since $\vartheta_{\mathbf{P}}$ satisfies the conditions of Proposition 5.2.1, the associated subshift is intrinsically ergodic, and $\mu_{\mathbf{P}}$ is the unique measure of maximal entropy.

5.2.3 Recognisable random substitutions

While the class of intrinsically ergodic random substitution subshifts given by Proposition 5.2.1 satisfy the Gibbs property (5.2), this is not the case in general for random substitution subshifts. For a broad class of recognisable random substitutions, the frequency measure of maximal entropy given by Theorem 5.1.1 does *not* satisfy this Gibbs property. Nonetheless, we can prove that this frequency measure is the unique measure of maximal entropy by other means.

Theorem 5.2.3. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a recognisable primitive random substitution of constant length ℓ and assume that at least one of the following holds:

(i) there exists an $N \in \mathbb{N}$ such that $\#\vartheta(a) = N$ for all $a \in \mathcal{A}$;

(ii) $\vartheta_{\mathbf{P}}$ is compatible and ℓ is the only non-zero eigenvalue of the substitution matrix.

Under these hypotheses, the system (X_ϑ, S) is intrinsically ergodic. Moreover, the unique measure of maximal entropy is the frequency measure corresponding to uniform probabilities.

We present the proof of Theorem 5.2.3 in Section 5.2.5. Our technique is to show that for the frequency measure of maximal entropy given by Theorem 5.1.1, a weaker Gibbs property holds for cylinder sets of exact inflation words. This, together with the underlying assumptions of recognisability and constant length, is sufficient to obtain the conclusion of Theorem 5.2.3.

Example 5.2.4. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} abbaa & \text{with probability } 1/2, \\ aabba & \text{with probability } 1/2, \end{cases} \\ b \mapsto \begin{cases} babaa & \text{with probability } 1/2, \\ baaba & \text{with probability } 1/2, \end{cases} \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. In Example 2.2.15, we showed that $\vartheta_{\mathbf{P}}$ is recognisable. Thus, since $\#\vartheta(a) = \#\vartheta(b)$, the conditions of Theorem 5.2.3 (specifically (i)) are satisfied. Hence, the subshift X_ϑ is intrinsically ergodic and $\mu_{\mathbf{P}}$ is the unique measure of maximal entropy.

5.2.4 Gibbs properties of frequency measures

In this section we prove the (weak) Gibbs property satisfied by the measure of maximal entropy for subshifts of random substitutions satisfying the conditions of Theorem 5.2.3. This is the content of Lemma 5.2.8, which utilises the following auxiliary results.

Lemma 5.2.5. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Further assume that $\vartheta_{\mathbf{P}}$ is compatible or constant length. Then, for all $v \in \mathcal{L}_\vartheta$ and $w \in \vartheta(v)$,

$$\mu_{\mathbf{P}}([w]) \geq \frac{1}{\lambda} \mu_{\mathbf{P}}([v]) \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w].$$

In the case that $\vartheta_{\mathbf{P}}$ is constant length and recognisable and $|\vartheta(v)| > 2\kappa(\vartheta)$, we have

$$\mu_{\mathbf{P}}([w]) = \frac{1}{\lambda} \sum_{u \in \mathcal{L}_\vartheta^{|v|}} \mu_{\mathbf{P}}([u]) \mathbb{P}[\vartheta_{\mathbf{P}}(u) = w].$$

Proof. Let $v \in \mathcal{L}_\vartheta$ and let $w \in \vartheta(v)$ be fixed. Let $n = |w|$ and $\mathcal{J}_n(v) = \{u \in \mathcal{L}_\vartheta^n : u_{[1,|v|]} = v\}$. Since $\vartheta_{\mathbf{P}}$ is compatible or constant length, Lemma 2.3.5 gives that

$$\mu_{\mathbf{P}}([w]) = \frac{1}{\lambda} \sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u]) \sum_{j=1}^{|\vartheta(u_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(u)_{[j,j+|w|-1]} = w]$$

Noting that $[v]$ is the union of all $[u]$ with $u \in \mathcal{J}_n(v)$, we thus obtain

$$\begin{aligned}\mu_{\mathbf{P}}([w]) &\geq \frac{1}{\lambda} \sum_{u \in \mathcal{J}_n(v)} \mu_{\mathbf{P}}([u]) \mathbb{P}[\vartheta_{\mathbf{P}}(u)_{[1,|w|]} = w] = \frac{1}{\lambda} \sum_{u \in \mathcal{J}_n(v)} \mu_{\mathbf{P}}([u]) \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w] \\ &= \frac{1}{\lambda} \mu_{\mathbf{P}}([v]) \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w].\end{aligned}$$

If $\vartheta_{\mathbf{P}}$ is constant length and recognisable, and $|\vartheta(v)| > 2\kappa(\vartheta)$, then there is a unique way to decompose w into inflation words. However, there might still be several words $u \in \mathcal{L}_\vartheta$ with $|u| = |v|$ such that $w \in \vartheta(u)$. Hence, it follows by Lemma 2.3.5 that

$$\mu_{\mathbf{P}}([w]) = \frac{1}{\lambda} \sum_{u \in \mathcal{L}_\vartheta^{|v|}} \mu_{\mathbf{P}}([u]) \mathbb{P}[\vartheta_{\mathbf{P}}(u) = w]. \qquad \square$$

Lemma 5.2.6. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive random substitution satisfying the disjoint set condition. Assume that $\mathbb{P}[\vartheta_{\mathbf{P}}(a) = s] = 1/\#\vartheta(a)$ for all $a \in \mathcal{A}$ and $s \in \vartheta(a)$ and that at least one of the following conditions is satisfied:

(i) $\vartheta_{\mathbf{P}}$ is of constant length ℓ and there is an $N \in \mathbb{N}$ such that $\#\vartheta(a) = N$ for all $a \in \mathcal{A}$;

(ii) $\vartheta_{\mathbf{P}}$ is compatible and the second largest eigenvalue τ of the substitution matrix satisfies $|\tau| < 1$.

Under these hypotheses, there exists a constant $c > 0$ such that $\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w] \geq c\mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)}$ for all $m \in \mathbb{N}$, $a \in \mathcal{A}$ and $w \in \vartheta^m(a)$. In particular, when $\vartheta_{\mathbf{P}}$ is of constant length, we have that $\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w] = \mathrm{e}^{h_{\mathrm{top}}(X_\vartheta)} \mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)}$.

Proof. As $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition, it follows by the Markov property of $\vartheta_{\mathbf{P}}$ and induction that, for $a \in \mathcal{A}$, $m \in \mathbb{N}$ and $w \in \vartheta^m(a)$, we have

$$\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w] = \frac{1}{\#\vartheta^m(a)}. \tag{5.4}$$

Let us first consider case (i). It follows by the constant length property and disjoint set condition that $\#\vartheta^m(a) = N^{1+\ell+\cdots+\ell^{m-1}}$ for all $m \in \mathbb{N}$. By Proposition 3.2.11, we have that $h_{\mathrm{top}}(X_\vartheta) =$

$\log N/(\ell-1)$. Noting that $\sum_{j=0}^{m-1}\ell^j = (\ell^m-1)/(\ell-1)$, it follows that

$$\log(\#\vartheta^m(a)) = \log N \sum_{j=0}^{m-1} \ell^j = (\ell^m - 1) h_{\mathrm{top}}(X_\vartheta) = |w| h_{\mathrm{top}}(X_\vartheta) - h_{\mathrm{top}}(X_\vartheta)$$

for all $m \in \mathbb{N}$, $a \in \mathcal{A}$ and $w \in \vartheta^m(a)$. Taking the exponential of both sides, we conclude from (5.4) that

$$\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w] = \mathrm{e}^{h_{\mathrm{top}}(X_\vartheta)} \mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)}.$$

Let us now consider case (ii). Since the Perron–Frobenius eigenvalue λ of $\vartheta_{\mathbf{P}}$ is simple, we can split the substitution matrix M as $M = \lambda \mathbf{R}\mathbf{L}^\top + N$, where $\mathbf{R}$ and $\mathbf{L}$ are respectively the right and left Perron–Frobenius eigenvectors of $\vartheta_{\mathbf{P}}$ and where $N\mathbf{R}\mathbf{L}^\top = 0 = \mathbf{R}\mathbf{L}^\top N$. For each $m \in \mathbb{N}$, we let $\mathbf{q}_m$ denote the vector indexed by $\mathcal{A}$ defined by $q_{m,a} = \log(\#\vartheta^m(a))$ for all $a \in \mathcal{A}$. It was shown in [35, Lemma 10] that for all primitive and compatible random substitutions satisfying the disjoint set condition, $\mathbf{q}_m^\top = \mathbf{q}_1^\top \sum_{k=0}^{m-1} M^k$ for all $m \in \mathbb{N}$. Hence,

$$\begin{aligned}\mathbf{q}_m^\top = \mathbf{q}_1^\top \sum_{k=0}^{m-1} M^k &= \mathbf{q}_1^\top \sum_{k=0}^{m-1} \lambda^k \mathbf{R}\mathbf{L}^\top + \mathbf{q}_1^\top \sum_{k=0}^{m-1} N^k \\ &= \frac{\lambda^m - 1}{\lambda - 1} \mathbf{q}_1^\top \mathbf{R}\mathbf{L}^\top + \mathbf{q}_1^\top \sum_{k=0}^{m-1} N^k = (\lambda^m - 1) h_{\mathrm{top}}(X_\vartheta) \mathbf{L}^\top + \mathbf{q}_1^\top \sum_{k=0}^{m-1} N^k.\end{aligned}$$

By construction, τ is the dominant eigenvalue of N, and so there exists a $c > 0$ and $n \in \mathbb{N}$ such that $\|N^k\|_\infty < c k^n |\tau|^k$ for all $k \in \mathbb{N}$. Hence, there is $r \in \mathbb{R}$ with $|\tau| < r < 1$ such that $\|N^k\|_\infty < c r^k$. We therefore obtain

$$\begin{aligned}\log(\#\vartheta^m(a)) &\leq (\lambda^m - 1) L_a h_{\mathrm{top}}(X_\vartheta) + \|\mathbf{q}_1\|_\infty \sum_{k=0}^{m-1} \|N^k\|_\infty \\ &\leq (\lambda^m - 1) L_a h_{\mathrm{top}}(X_\vartheta) + \frac{c}{1-r} \|\mathbf{q}_1\|_\infty.\end{aligned}$$

On the other hand, by Proposition 2.2.3, we have that

$$|\vartheta^m(a)| \geq L_a \lambda^m - D|\tau|^m \geq L_a \lambda^m - D,$$

for some $D > 0$. Hence, there exists a constant $C > 0$ such that $\log(\#\vartheta^m(a)) \leq |\vartheta^m(a)| h_{\mathrm{top}}(X_\vartheta) + C$. Taking the exponential of both sides, we conclude from (5.4) that $\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w] \geq$

$\mathrm{e}^{-|w|h_{\mathrm{top}}(X_\vartheta)}\mathrm{e}^{-C}$. Setting $c = e^{-C}$ completes the proof. If $\vartheta_{\mathbf{P}}$ is additionally assumed to be of constant length, then $\tau = 0$ since the eigenvalues of the matrix associated to a constant length substitution are integers. In this case, the matrix M satisfies $M = \lambda \mathbf{R}\mathbf{L}^\top$, where $\mathbf{L} = (1, \ldots, 1)$ by the constant length property. Thus, it follows by the same arguments as above that $\log(\#\vartheta^m(a)) = (\lambda^m - 1)h_{\mathrm{top}}(X_\vartheta)$. Taking the exponential of both sides, it follows from (5.4) that $\mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w] = \mathrm{e}^{h_{\mathrm{top}}(X_\vartheta)}\mathrm{e}^{-|w|h_{\mathrm{top}}(X_\vartheta)}$. $\square$

Lemma 5.2.7. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a random substitution satisfying either of the conditions of Lemma 5.2.6 and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Then, there exists a constant $c > 0$ such that

$$\mu_{\mathbf{P}}([w]) \geq \mu_{\mathbf{P}}([v]) \frac{c^{|v|}}{|w|\mathrm{e}^{|w|h_{\mathrm{top}}(X_\vartheta)}}$$

for all $v \in \mathcal{L}_\vartheta$, $m \in \mathbb{N}$ and $w \in \vartheta^m(v)$. If, in addition, $\vartheta_{\mathbf{P}}$ is constant length and recognisable and $|v| > 2\kappa(\vartheta)$, then

$$\mu_{\mathbf{P}}([w]) \leq \frac{|v|\mathrm{e}^{|v|h_{\mathrm{top}}(X_\vartheta)}}{|w|\mathrm{e}^{|w|h_{\mathrm{top}}(X_\vartheta)}}.$$

Proof. Let $v \in \mathcal{L}_\vartheta$, $m \in \mathbb{N}$ and $w \in \vartheta^m(v)$ be fixed. Applying Lemma 5.2.5 to $\vartheta_{\mathbf{P}}^m$ yields

$$\mu_{\mathbf{P}}([w]) \geq \frac{1}{\lambda^m} \mu_{\mathbf{P}}([v]) \mathbb{P}[\vartheta_{\mathbf{P}}^m(v) = w]. \tag{5.5}$$

Since $\vartheta_{\mathbf{P}}$ is compatible or constant length, we can decompose w into subwords $w = w^1 \cdots w^{|v|}$ such that $w^j \in \vartheta^m(v_j)$ for all $j \in \{1, \ldots, |v|\}$. Hence, it follows by Lemma 5.2.6 that there is a constant $c > 0$ such that

$$\mathbb{P}[\vartheta_{\mathbf{P}}^m(v) = w] = \prod_{j=1}^{|v|} \mathbb{P}[\vartheta_{\mathbf{P}}^m(v_j) = w^j] \geq \prod_{j=1}^{|v|} c\, \mathrm{e}^{-|w^j|h_{\mathrm{top}}(X_\vartheta)} = c^{|v|}\mathrm{e}^{-|w|h_{\mathrm{top}}(X_\vartheta)}. \tag{5.6}$$

By Proposition 2.2.3, there is a universal constant $D > 0$ such that $\lambda^m \leq D|\vartheta^m(a)|$ for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$. Combining this with (5.5) and (5.6) yields the desired result.

Now, assume additionally that $\vartheta_{\mathbf{P}}$ is recognisable and of constant length ℓ. Then by Lemma 5.2.6 we have that $\mathbb{P}[\vartheta_{\mathbf{P}}^m(u) = w] = \mathrm{e}^{|u|h_{\mathrm{top}}(X_\vartheta)}\mathrm{e}^{-|w|h_{\mathrm{top}}(X_\vartheta)}$ for every $u \in \mathcal{L}_\vartheta^{|v|}$ with $w \in \vartheta(u)$. Thus, the lower bound follows by identical arguments to the above, taking $c = e^{h_{\mathrm{top}}(X_\vartheta)}$.

For the upper bound, observe that if $|v| > 2\kappa(\vartheta)$, we also have $|\vartheta^m(v)| = \ell^m |v| > 2\kappa(\vartheta^m)$, for all $m \in \mathbb{N}$, by Lemma 2.2.19. Hence, noting that $|u| = |v|$ and $\ell^{-m} = |v|/|w|$, Lemma 5.2.5 gives that

$$\mu_{\mathbf{P}}([w]) = \frac{1}{\ell^m} \sum_{u \in \mathcal{L}_\vartheta^{|v|}} \mu_{\mathbf{P}}([u]) \mathbb{P}[\vartheta_{\mathbf{P}}^m(u) = w] \leq \frac{|v| \mathrm{e}^{|v| h_{\mathrm{top}}(X_\vartheta)}}{|w| \mathrm{e}^{|w| h_{\mathrm{top}}(X_\vartheta)}}. \qquad \square$$

In the proof of Theorem 5.2.3 we only require the lower bound of Lemma 5.2.7. However, the upper bound allows us to show that the subshifts we consider in Theorem 5.2.3 do not satisfy the Gibbs property (5.2), therefore do not satisfy Bowen's specification property or the weaker specification property of Climenhaga and Thompson. This is illustrated by the following.

Lemma 5.2.8. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a random substitution satisfying the conditions of Theorem 5.2.3, and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Then there exist constants $c_1, c_2 > 0$ such that for all $a \in \mathcal{A}$, $m \in \mathbb{N}$ and $w \in \vartheta^m(a)$,

$$\frac{c_1}{|w|} \mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)} \leq \mu_{\mathbf{P}}([w]) \leq \frac{c_2}{|w|} \mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)}.$$

Proof. The lower bound follows immediately from Lemma 5.2.7, taking $c_1 = \min_{a \in \mathcal{A}} c \mu_{\mathbf{P}}([a])$ where c is the constant given by Lemma 5.2.7. For the upper bound, let M be the least integer such that $\ell^M > 2\kappa(\vartheta)$ and set $c_2 = \max_{u \in \mathcal{L}_\vartheta, |u| \leq \ell^M} |u| \mathrm{e}^{|u| h_{\mathrm{top}}(X_\vartheta)}$. Since $\mu_{\mathbf{P}}$ is a probability measure, $\mu_{\mathbf{P}}([w]) \leq c_2 \mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)}/|w|$ if $|w| \leq \ell^M$. On the other hand, if $m > M$ and $w \in \vartheta^m(a)$ then it follows by Lemma 5.2.7 that there is a $v \in \vartheta^M(a)$ such that

$$\mu_{\mathbf{P}}([w]) \leq \frac{|v| \mathrm{e}^{|v| h_{\mathrm{top}}(X_\vartheta)}}{|w| \mathrm{e}^{|w| h_{\mathrm{top}}(X_\vartheta)}} \leq \frac{c_2}{|w|} \mathrm{e}^{-|w| h_{\mathrm{top}}(X_\vartheta)}.$$

This completes the proof. $\square$

The upper bound on $\mu_{\mathbf{P}}$ in Lemma 5.2.8 is irreconcilable with the bound for the unique measure of maximal entropy on subshifts with a weak specification property established in [14, Lemma 5.12]. In particular, the subshifts for which Theorem 5.2.3 establishes intrinsic ergodicity do not satisfy the weak specification property in [14].

5.2.5 Proof of Theorem 5.2.3

We now present the proof of Theorem 5.2.3. In addition to the Gibbs property proved in the previous section, we also utilise the following result, which is proved in [21]. For ease of notation, we let $\triangle$ denote the symmetric difference of two sets. That is, given two sets A and B, we write $A \triangle B = (A \cup B) \setminus (A \cap B)$.

Lemma 5.2.9 ([21, Lemma 8.8])**.** Let (X, d) be a compact metric space and let ϱ be a Borel probability measure on X. If $B \subset X$ is measurable and $(\xi_n)_{n \in \mathbb{N}}$ is a sequence of finite measurable partitions of X for which $\lim_{n\to\infty} \max_{P \in \xi_n} \operatorname{diam}(P) = 0$, then there exists a sequence of sets $(A_n)_{n\in\mathbb{N}}$ with $A_n \in \sigma(\xi_n)$ and $\lim_{n\to\infty} \varrho(A_n \triangle B) = 0$. Here, $\sigma(\xi_n)$ denotes the sigma algebra generated by the partition ξ_n.

Proof of Theorem 5.2.3. Let μ denote the frequency measure of maximal entropy given by Theorem 5.1.1 and let $m \in \mathbb{N}$. For each $k \in \{0, \ldots, \ell^m - 1\}$, let $X_{m,k}$ denote the subset of X_ϑ defined by $X_{m,k} = S^k(\vartheta^m(X_\vartheta))$. It follows by recognisability that these subsets are pairwise disjoint for different choices of k. Note, by Lemma 2.1.6, that the subsets $X_{m,k}$ are closed, and since by the constant length property we have

$$S^{\ell^m}(X_{m,k}) = S^{\ell^m}(S^k(\vartheta^m(X_\vartheta))) = S^k(\vartheta^m(SX_\vartheta)) = S^k(\vartheta^m(X_\vartheta)) = X_{m,k},$$

it follows that $X_{m,k}$ is a S^{ℓ^m}-invariant subspace of X_ϑ. Since every $x \in X_\vartheta$ can be split into level-m inflation words, we have

$$X_\vartheta = \bigsqcup_{k=0}^{\ell^m - 1} X_{m,k},$$

where the union is disjoint due to recognisability. Lemma 2.2.19 implies that $r = \lceil \kappa(\vartheta)/(\ell - 1) \rceil + 1$ satisfies

$$\ell^m r > \frac{\ell^m - 1}{\ell - 1} \kappa(\vartheta) + \ell^m \geq \kappa(\vartheta^m) + \ell^m.$$

By the constant length property, this ensures that every word of length at least $2r\ell^m$ has a unique decomposition into inflation words. This together with Lemma 2.2.17 implies, for all $u \in \mathcal{L}_\vartheta^{2r}$ and $w \in \vartheta^m(u)$, that $|w| = 2r\ell^m$ and $S^{r\ell^m}([w]) \subset \vartheta^m(X_\vartheta)$. Let us consider the following partition of

$X_{m,k}$:

$$\xi_{m,k} = S^{r\ell^m}\left(\left\{S^k([w]) : w \in \vartheta^m(u) \text{ and } u \in \mathcal{L}_\vartheta^{2r}\right\}\right).$$

This in turn yields a partition of X_ϑ, namely

$$\xi_m = \bigcup_{k=0}^{\ell^m-1} \xi_{m,k}.$$

By way of a contradiction, assume that $\nu \neq \mu$ is another ergodic measure of maximal entropy. Since distinct ergodic measures are mutually singular, there exists an S-invariant set B with $\mu(B) = 0$ and $\nu(B) = 1$. Note that the diameter of the atoms of ξ_m tends uniformly to zero as m tends to infinity, so $(\xi_m)_{m\in\mathbb{N}}$ meets the requirements of Lemma 5.2.9. Applying it to the measure $\varrho' = (\mu+\nu)/2$ we obtain that, given $\varepsilon > 0$, there exist $m \in \mathbb{N}$ and $A_m \in \sigma(\xi_m)$ such that

$$(\mu+\nu)(A_m \triangle B) < \varepsilon. \tag{5.7}$$

For $k \in \{0,\dots,\ell^m-1\}$, let $A_{m,k} = A_m \cap X_{m,k}$ and $B_{m,k} = B \cap X_{m,k}$, and define the conditional probability measures $\mu_{m,k}$ and $\nu_{m,k}$ by

$$\mu_{m,k} = \frac{1}{\mu(X_{m,k})}\,\mu|_{X_{m,k}} \quad \text{and} \quad \nu_{m,k} = \frac{1}{\nu(X_{m,k})}\,\nu|_{X_{m,k}}.$$

For all $j \in \{0,\dots,\ell^m-1\}$, we have $S^{k-j}(X_{m,j}) = X_{m,k}$, and since μ and ν are S-invariant and the sets $X_{m,k}$ are disjoint, it follows that

$$\mu(X_{m,k}) = \mu(X_{m,j}) = \frac{1}{\ell^m} \quad \text{and} \quad \nu(B\cap X_{m,k}) = \nu(B \cap X_{m,j}) = \frac{1}{\ell^m}.$$

Consequently, $\nu_{m,k}(B_{m,k}) = \ell^m\,\nu(B\cap X_{m,k}) = 1$. On the other hand, $\mu_{m,k}(B_{m,k}) = \ell^m\mu(B \cap X_{m,k}) = 0$. Since $\{X_{m,k} : k \in \{0,\dots,\ell^m-1\}\}$ forms a partition of X_ϑ, we can rewrite (5.7) as

$$\begin{aligned}\sum_{k=0}^{\ell^m-1}(\mu_{m,k}+\nu_{m,k})(A_{m,k}\triangle B_{m,k}) &= \ell^m \sum_{k=0}^{\ell^m-1}(\mu+\nu)((A_m\triangle B)\cap X_{m,k})\\ &= \ell^m(\mu+\nu)(A_m\triangle B) < \ell^m\varepsilon.\end{aligned}$$

Hence, there exists a k' such that

$$(\mu_{m,k'} + \nu_{m,k'})(A_{m,k'} \triangle B_{m,k'}) < \varepsilon. \tag{5.8}$$

Here, we observe that $A_{m,k'} \in \sigma(\xi_{m,k'})$, and recall that if $|v| \geq 2\ell^m r$, then the word v has a unique inflation word decomposition under ϑ^m. Therefore, there exists a unique $j \in \{0, \ldots, \ell^m - 1\}$ such that $[v] \subset X_{m,j}$.

Note that the system $(X_{m,j}, S^{\ell^m})$ equipped with the measure $\nu_{m,j}$ is an induced subshift obtained from (X_ϑ, S) equipped with the measure ν by inducing on $X_{m,j}$. Hence, by Abramov's formula (Lemma A.1.10), we have

$$h_\nu(X_\vartheta, S) = \frac{1}{\ell^m} h_{\nu_{m,j}}(X_{m,j}, S^{\ell^m}).$$

The remainder of the proof follows a similar line of arguments to Adler and Weiss' [2] proof of intrinsic ergodicity for topologically transitive shifts of finite type, applied to the system $(X_{m,k'}, S^{\ell^m})$ and the S^{ℓ^m}-invariant measures $\mu_{m,k'}$ and $\nu_{m,k'}$. For ease of notation, in the following we write $k = k'$ and $T = S^{\ell^m}$. Note that

$$\alpha_{m,k} = \{S^k([w]) : w \in \vartheta^m(a), a \in \mathcal{A}\}$$

forms a generating partition of $X_{m,k}$, and by the fact that $\vartheta_{\mathbf{P}}$ is of constant length and recognisable,

$$\xi_{m,k} = \bigvee_{j=-r}^{r-1} T^{-j}(\alpha_{m,k}).$$

Let $\eta_m = \{A_{m,k}, X_{m,k} \setminus A_{m,k}\}$ and for a given set $A \subseteq X_{m,k}$ denote by $t_m(A)$ the number of atoms in $\xi_{m,k}$ that intersect A. Then, we have

$$\begin{aligned}
2r\ell^m h_\nu(X_\vartheta, S) &= 2r h_{\nu_{m,k}}(X_{m,k}, S^{\ell^m}) \leq H_{\nu_{m,k}}(\xi_{m,k}) \\
&\leq H_{\nu_{m,k}}(\eta_m) + H_{\nu_{m,k}}(\xi_{m,k} \,|\, \eta_m) \\
&\leq \log 2 + \nu_{m,k}(A_{m,k}) \log(t_m(A_{m,k})) + \nu_{m,k}(X_{m,k} \setminus A_{m,k}) \log(t_m(X_{m,k} \setminus A_{m,k})).
\end{aligned}$$

Let $S^{r\ell^m+k}[w] \in \xi_{m,k}$, with $w \in \vartheta^m(v)$ for some $v \in \mathcal{L}_\vartheta^{2r}$. By Lemma 5.2.7, we have that

$$\mu_{m,k}(S^{r\ell^m+k}([w])) = \ell^m \mu([w]) \geq \mu([v]) \frac{c^{2r}}{2r\mathrm{e}^{2\ell^m r h_{\mathrm{top}}(X_\vartheta)}} \geq C\mathrm{e}^{-2r\ell^m h_{\mathrm{top}}(X_\vartheta,S)},$$

taking $C = c^{2r}(\min_{v\in\mathcal{L}_\vartheta^{2r}} \mu([v]))/2r$. We have that $C > 0$ since $\mu([v]) > 0$ for all $v \in \mathcal{L}_\vartheta^r$. Hence,

$$t_m(A_{m,k}) \leq \frac{1}{C}\,\mu(A_{m,k})\mathrm{e}^{2\ell^m r h_{\mathrm{top}}(X_\vartheta,S)}$$

and

$$t_m(X_{m,k} \setminus A_{m,k}) \leq \frac{1}{C}\,\mu(X_{m,k} \setminus A_{m,k})\mathrm{e}^{2\ell^m r h_{\mathrm{top}}(X_\vartheta,S)}.$$

This yields that $0 \leq \log(2) - \log(C) + \nu_{m,k}(A_{m,k})\log(\mu_{m,k}(A_{m,k}))$. By (5.8), we have that $\mu_{m,k}(A_{m,k}) < \varepsilon$ and $\nu_{m,k}(A_{m,k}) > 1 - \varepsilon$. This implies the following contradiction:

$$0 \leq \lim_{\varepsilon\to 0}\left(\log 2 - \log(C) + (1-\varepsilon)\log\varepsilon\right) = -\infty. \qquad \square$$

From Lemma 5.2.7, we have used only the lower bound in the proof of Theorem 5.2.3. Since this inequality holds under less restrictive conditions, it seems natural to enquire whether Theorem 5.2.3 can be sharpened accordingly by replacing the constant length assumption with a weaker condition. However, a closer inspection reveals that the last part of the proof relies on the detailed control that the constant length assumption provides. A definite answer therefore remains an open problem.

5.2.6 Random period doubling

We have shown that for the random period doubling substitution

$$\vartheta_{\mathbf{P}}\colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1, \end{cases}$$

the frequency measure μ corresponding to $p = 1/2$ is a measure of maximal entropy for the associated subshift X_ϑ. We now show that this is the unique measure of maximal entropy.

Theorem 5.2.10. The subshift associated to the random period doubling substitution is intrinsically ergodic. Moreover, the unique measure of maximal entropy is the frequency measure corresponding to $p = 1/2$.

Our strategy of proof is to construct a random substitution $\varphi_{\mathbf{P}} = (\varphi, \mathbf{P})$ satisfying the conditions of Theorem 5.2.3 that gives rise to a subshift with the same topological entropy as X_ϑ, for which there exists a factor map $\pi \colon X_\varphi \to X_\vartheta$. That is, a continuous surjective map for which $\pi \circ S = S \circ \pi$. Uniqueness of the measure of maximal entropy then follows by the intrinsic ergodicity of the subshift X_φ and the fact that every measure of maximal entropy on X_ϑ is the push-forward of a measure of maximal entropy on X_φ. A proof of this latter fact is given in [51].

Lemma 5.2.11 ([51, Theorem 3.3])**.** Let X and Y be subshifts with $h_{\mathrm{top}}(X) = h_{\mathrm{top}}(Y)$ and let $\pi \colon X \to Y$ be a factor map. Then, every measure of maximal entropy on Y is the push-forward of a measure of maximal entropy on X.

We now give the proof of Theorem 5.2.10.

Proof of Theorem 5.2.10. Let $\varphi_{\mathbf{P}} = (\varphi, \mathbf{P})$ be the random substitution defined over the alphabet $\mathcal{A} = \{a_0, a_1, b_0, b_1\}$ by

$$\varphi_{\mathbf{P}} \colon \begin{cases} a_0 \mapsto \begin{cases} a_0 b_0 & \text{with probability } 1/2, \\ b_0 a_1 & \text{with probability } 1/2, \end{cases} \\ a_1 \mapsto \begin{cases} a_0 b_1 & \text{with probability } 1/2, \\ b_1 a_1 & \text{with probability } 1/2, \end{cases} \\ b_0 \mapsto a_0 a_0 \quad \text{with probability } 1, \\ b_1 \mapsto a_1 a_1 \quad \text{with probability } 1, \end{cases}$$

and let ν denote the corresponding frequency measure. We claim that $\varphi_{\mathbf{P}}^2$ satisfies the conditions of Theorem 5.2.3. Since $\varphi_{\mathbf{P}}^2$ gives rise to the same frequency measure as $\varphi_{\mathbf{P}}$, this gives that the frequency measure ν is the unique measure of maximal entropy for the subshift X_φ. It is clear that $\varphi_{\mathbf{P}}$ is a primitive random substitution of constant length, therefore so is $\varphi_{\mathbf{P}}^2$. Also, observe that $\#\varphi^2(a_0) = \#\varphi^2(a_1) = \#\varphi^2(b_0) = \#\varphi^2(b_1) = 4$, so the only condition of Theorem 5.2.3 that it remains to verify is that $\varphi_{\mathbf{P}}^2$ is recognisable. By Lemma 2.2.19, it suffices to show that $\varphi_{\mathbf{P}}$ is recognisable. To show this, we prove that for all $x \in X$, there is a unique $j \in \{0,1\}$ and $y \in X$ such that $x \in S^{-k}(\varphi(y))$. Since all exact inflation words are distinct, for each $j \in \{0,1\}$, if $x \in X$ then there is a unique $y \in X$ such that $x \in S^{-j}(\varphi(y))$. Hence, it suffices to show that there is a unique $j \in \{0,1\}$ such that $x \in S^{-j}(\varphi(X))$. Observe that every $x \in X$ contains an occurrence of b_0 or b_1. If an occurrence of either of these letters is followed by a_0, b_0 or b_1, then this uniquely determines the choice of j, since these two letter words cannot appear on the overlap of two exact inflation words. On the other hand, if the following letter is an a_1, then j is uniquely determined by the number of occurrences of a_1 before another letter is observed. In particular, if there are an even number of occurrences of a_1, then $b_j a_1$ lies on the overlap of two inflation tiles and if there are an odd number, $b_j a_1$ is contained within a single inflation tile. Thus, $\varphi_{\mathbf{P}}$ is recognisable, therefore satisfies the conditions of Theorem 5.2.3. Hence, the subshift X_φ is intrinsically ergodic.

Observe that a factor map $\pi\colon X_\varphi \to X_\vartheta$ can be defined onto the random period doubling subshift X_ϑ by $\pi(a_0) = \pi(a_1) = a$, $\pi(b_0) = \pi(b_1) = b$. Thus, by Lemma 5.2.11, every measure of maximal entropy on X_ϑ is the push-forward of a measure of maximal entropy on X_φ. Since the subshift X_φ is intrinsically ergodic, it follows that the only measure of maximal entropy on X_ϑ is the push-forward of the unique measure of maximal entropy on X_φ. In particular, the subshift X_ϑ is intrinsically ergodic. By Theorem 5.1.1, the unique measure of maximal entropy is the frequency measure corresponding to $p = 1/2$. □

5.2.7 The golden mean shift

It was shown in [38] that every topologically transitive shift of finite type can be obtained as the subshift of a primitive random substitution. For the golden mean shift, we show that the unique measure of maximal entropy, the Parry measure, can be obtained as the weak*-limit of a sequence of frequency measures corresponding to primitive random substitutions.

Example 5.2.12 (The golden mean shift)**.** The golden mean shift is the shift of finite type over the alphabet $\{a,b\}$ defined by the forbidden word set $\mathcal{F} = \{bb\}$. The subshift X can be obtained as the subshift of the random substitution $\vartheta_{\mathbf{P}}$ defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} aa & \text{with probability } \tau^{-1}, \\ aba & \text{with probability } \tau^{-2}, \end{cases} \\ b \mapsto b \quad \text{with probability } 1. \end{cases}$$

However, this random substitution is not primitive, so we cannot directly apply our results. To circumvent this issue, let $\varepsilon \in (0,1)$ and let ϑ_ε be the random substitution defined by

$$\vartheta_\varepsilon \colon \begin{cases} a \mapsto \begin{cases} aa & \text{with probability } \tau^{-1}, \\ aba & \text{with probability } \tau^{-2}, \end{cases} \\ b \mapsto \begin{cases} b & \text{with probability } 1-\varepsilon, \\ abb & \text{with probability } \varepsilon, \end{cases} \end{cases}$$

and let μ_ε denote the corresponding frequency measure. For all $\varepsilon \in (0,1)$, ϑ_ε is a primitive random substitution with unique realisation paths, satisfying the disjoint set condition. Since the space of shift-invariant measures on X is weak*-compact, there exists a shift-invariant measure μ such that μ_ε converges weak* to μ as $\varepsilon \to 0$. One can show that $R_{a,\varepsilon}/(\lambda_\varepsilon - 1) \to \tau^2/(\tau^2+1)$ as $\varepsilon \to 0$, where λ_ε and $R_{a,\varepsilon}$ are the Perron–Frobenius eigenvalue and the entry of the right Perron–Frobenius eigenvector corresponding to the letter a, respectively, and τ denotes the golden ratio. Thus, it follows by the upper semi-continuity of entropy and Theorem 4.1.2 that

$$\begin{aligned} h_\mu(X) &\geq \limsup_{\varepsilon\to 0} h_{\mu_\varepsilon}(X) = \limsup_{\varepsilon\to 0} \frac{1}{\lambda_\varepsilon - 1} \mathbf{H}_1^\top \mathbf{R} \geq \limsup_{\varepsilon\to 0} \frac{-1}{\lambda_\varepsilon - 1} R_{a,\varepsilon}(\tau^{-2}\log\tau^{-2} + \tau^{-1}\log\tau^{-1}) \\ &= \frac{\tau^2}{\tau^2+1}(2\tau^{-2}+\tau^{-1})\log\tau = \log\tau, \end{aligned}$$

where the last equality is a consequence of the characteristic equation $\tau^2 = \tau + 1$. Since $h_{\text{top}}(X) = \log\tau$ and the Parry measure is the unique measure of maximal entropy [2, 63] on X, we conclude that μ must be the Parry measure.

In [38], it was shown that for any topologically transitive shift of finite type, there is an algorithm which constructs a primitive random substitution that gives rise to that subshift. Thus, this algorithm yields a primitive random substitution that gives rise to the golden mean shift. However, a closer inspection reveals that if the corresponding frequency measure is the Parry measure then we require two of the realisations to occur with probability zero and the resulting random substitution is the random substitution $\vartheta_{\mathbf{P}}$ defined in Example 5.2.12, which is not primitive. As to whether there exists a primitive random substitution for which the Parry measure is the corresponding frequency measure remains open.

5.2.8 An example with multiple measures of maximal entropy

So far, all of the examples we have considered in this section have been intrinsically ergodic. However, there exist subshifts of primitive random substitutions that support multiple measures of maximal entropy. This is illustrated by the following example, the *Dyck shift*, which supports two distinct ergodic measures of maximal entropy.

Example 5.2.13 (The Dyck shift)**.** For each $i \in \{1, 2, 3, 4\}$, let $\mathbf{p}_i = (p_{i,1}, p_{i,2}, p_{i,3})$ be a probability vector and let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined over the alphabet $\mathcal{A} = \{(,), [,]\}$ by

$$\vartheta_{\mathbf{P}} : \begin{cases} (\;\mapsto \begin{cases} (& \text{with probability } p_{1,1}, \\ (\,() & \text{with probability } p_{1,2}, \\ (\,[\,] & \text{with probability } p_{1,3}, \end{cases} &) \;\mapsto \begin{cases}) & \text{with probability } p_{2,1}, \\ ()\,) & \text{with probability } p_{2,2}, \\ [\,]\,) & \text{with probability } p_{2,3}, \end{cases} \\ [\;\mapsto \begin{cases} [& \text{with probability } p_{3,1}, \\ [\,() & \text{with probability } p_{3,2}, \\ [\,[\,] & \text{with probability } p_{3,3}, \end{cases} &] \;\mapsto \begin{cases}] & \text{with probability } p_{4,1}, \\ ()\,] & \text{with probability } p_{4,2}, \\ [\,]\,] & \text{with probability } p_{4,3}. \end{cases} \end{cases}$$

Gohlke and Spindeler [39] showed that the corresponding subshift is the Dyck shift, which supports two distinct ergodic measures of maximal entropy [47]. The random substitution $\vartheta_{\mathbf{P}}$ does not have unique realisation paths since, for example, the word (()) can be obtained as two different realisations of () under $\vartheta_{\mathbf{P}}$.

5.2.9 Outlook

We have established intrinsic ergodicity for several families of random substitution subshifts. However, we are still far from a classification of intrinsic ergodicity for random substitution subshifts. In particular, we highlight that there is a large gap between the theory of measure theoretic entropy developed in Chapter 4, which applies to all primitive random substitutions, and the more restrictive conditions that we require to establish intrinsic ergodicity. We conclude this chapter by presenting several open questions and possible directions for future work.

The intrinsic ergodicity of random period doubling substitution subshift suggests a relaxation of the recognisability assumption in Theorem 5.2.3 to the disjoint set condition may be possible. In this more general setting, the conditions of Theorem 5.1.1 are still satisfied, so the frequency measure corresponding to uniform probabilities is a measure of maximal entropy. Moreover, this measure satisfies the lower bound of the Gibbs property used in the proof of Theorem 5.2.3. However, the proof of Theorem 5.2.3 relies on the existence of a finite recognisability radius, which we no longer have without the assumption of recognisability. Therefore, a different approach will be required. One possible approach could be to extend the techniques we used to show intrinsic ergodicity for the subshift of the random period doubling substitution. Namely, by constructing a surjective measure preserving factor map from the subshift of a recognisable random substitution satisfying the conditions of Theorem 5.2.3.

It remains open as to whether the subshift associated with the random Fibonacci substitution is intrinsically ergodic. In Example 5.1.2, we showed that the frequency measure corresponding to uniform probabilities is not a measure of maximal entropy. However, by Theorem 5.1.4, a measure of maximal entropy can be obtained as the weak*-limit of a sequence of frequency measures. A careful analysis of the properties of this measure could provide an indication as to whether or not it is unique as a measure of maximal entropy.

The Dyck shift illustrates that, in general, primitive random substitutions need not be intrinsically ergodic. However, the random substitution that gives rise to the Dyck shift is a somewhat pathological example as it does not have unique realisation paths. We are not currently aware of any random substitutions with unique realisation paths for which it is known that the corresponding subshift supports multiple measures of maximal entropy.

CHAPTER 6

MULTIFRACTAL PROPERTIES OF FREQUENCY MEASURES

So far, we have developed techniques to study the measure theoretic entropy of the subshift X_ϑ of a primitive random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$, with respect to the corresponding frequency measure $\mu_{\mathbf{P}}$. It follows by the Shannon–McMillan–Breiman theorem (Theorem A.1.9) that

$$\lim_{n\to\infty} \frac{\log \mu_{\mathbf{P}}([x_{[-n,n]}])}{2n+1} = h_{\mu_{\mathbf{P}}}(X_\vartheta)$$

for $\mu_{\mathbf{P}}$-almost all $x \in X_\vartheta$. However, entropy does not give any information about the scaling behaviour of $\mu_{\mathbf{P}}([x_{[-n,n]}])$ for non-typical x. In this section, we develop techniques to study this behaviour. Specifically, we study the fine scaling properties of frequency measures from the perspective of *multifractal analysis*. This perspective is relevant in a wide variety of contexts, such as the geometry of fractal sets and measures and in dynamical systems, with typical applications to geometric measure theory and number theory. In our setting, our primary objects of study are the *L^q-spectrum*, a parametrised family of quantities that capture the inhomogeneous scaling properties of a measure, and the *local dimensions*, which capture the exponential growth rate of a measure around a point. The L^q-spectrum and local dimensions are related through a heuristic relationship known as the *multifractal formalism*, first introduced and studied in a physical context in [40]. It is an important and well-studied question to determine settings in which the multifractal formalism holds, and to determine qualitative conditions describing its failure.

Much of the work on multifractal analysis has been done in the setting of local dimensions of self-similar measures (see [3, 28, 29, 48, 72] for some examples) and Birkhoff sums of potentials in dynamical systems with a finite type property (see, for example, [25, 66] and the references therein). As a notable recent example, in Shmerkin's proof of Furstenberg's intersection conjecture [72], the L^q-spectrum is computed for a large class of dynamically self-similar measures and related to the multifractal analysis of slices of sets. However, we note that alternative proofs of Furstenberg's intersection conjecture have been provided by Austin [4] and Wu [77], which do not rely on techniques from multifractal analysis. For more details on the geometry of measures and multifractal analysis, we refer the reader to the foundational work of Olsen [60] and the classic texts of Falconer [23, 24] and Pesin [64].

We have already seen that random substitution subshifts have characteristic features of (dynamical) self-similarity, but in many cases are far from being shifts of finite type. This is highlighted by the fact that Theorem 5.2.3 provides a broad class of random substitutions for which the associated subshift supports a frequency measure that is the unique measure of maximal entropy but does not satisfy the Gibbs property. Thus, frequency measures typically do not fall into any of the classes for which the main multifractal properties are already well understood.

Here, we derive symbolic expressions for the L^q-spectrum of frequency measures corresponding to primitive and compatible random substitutions. We introduce an inflation word analogue of the L^q-spectrum that reflects the underlying Markov process, in a similar vein to the analogue of measure theoretic entropy introduced in Chapter 4. For $q \geq 0$, we show that this coincides with the L^q-spectrum and often yields a closed-form formula. For $q < 0$, it provides a lower bound for the L^q-spectrum but the two notions need not coincide in general. Our results on the L^q-spectrum are presented in Section 6.2. Under the additional assumption of recognisability, we show that the two notions coincide for all $q \in \mathbb{R}$. Moreover, we establish that the multifractal formalism holds, which allows the multifractal spectrum to be obtained. We present our results on the multifractal spectrum and multifractal formalism in Section 6.3.

The results in this chapter are based on the paper [55], which is joint work with A. Rutar.

6.1 Multifractal analysis and the L^q-spectrum

6.1.1 Local dimensions

One of the primary methods of studying local properties of measures in multifractal analysis is via local dimensions and the multifractal spectrum. Local dimensions quantify the scaling behaviour of a measure at a given point.

Definition 6.1.1. Let μ be a Borel measure supported on a compact metric space X. For each $x \in X$, we define the *lower and upper local dimension of μ at x*, respectively, by

$$\underline{\dim}_{\mathrm{loc}}(\mu, x) = \liminf_{r \to 0} \frac{\log \mu\big(B(x,r)\big)}{\log r}$$
$$\overline{\dim}_{\mathrm{loc}}(\mu, x) = \limsup_{r \to 0} \frac{\log \mu\big(B(x,r)\big)}{\log r}.$$

When these quantities coincide, we refer to their common value as the *local dimension of μ at x*, which we denote by $\dim_{\mathrm{loc}}(\mu, x)$.

Observe that the definition of local dimensions requires a metric structure. For subshifts, a natural metric can be defined as follows. With this metric, the lower and upper local dimensions can be expressed in terms of cylinder sets.

Definition 6.1.2. Let $\mathcal{A}$ be a finite alphabet. If $x, y \in \mathcal{A}^{\mathbb{Z}}$ are such that $x \neq y$ but $x_0 = y_0$, then we let $n(x,y)$ denote the largest integer such that $x_j = y_j$ for all $|j| \leq n$. A metric $d \colon \mathcal{A}^{\mathbb{Z}} \times \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ can be defined by

$$d(x,y) = \begin{cases} 1 & \text{if } x_0 \neq y_0 \\ \mathrm{e}^{-(2n(x,y)+1)} & \text{if } x \neq y \text{ but } x_0 = y_0 \\ 0 & \text{if } x = y \end{cases}$$

for all $x, y \in \mathcal{A}^{\mathbb{Z}}$.

We note that the metric d generates the (discrete product) topology on $\mathcal{A}^{\mathbb{Z}}$ – see [49] for more details. Throughout this chapter, we assume that the subshifts we consider are equipped with this metric. If $x \in \mathcal{A}^{\mathbb{Z}}$, and $n \in \mathbb{N}$, then $B(x, \mathrm{e}^{-(2n+1)})$ is precisely the set of elements in X that agree with x on positions $-n$ up to n; that is, the cylinder set $[x_{[-n,n]}]$. Moreover, for all $r > 0$, we have that $B(x,r) = [x_{[-n(r),n(r)]}]$, where $n(r)$ is the integer such that $\mathrm{e}^{-(2n(r)+1)} \leq r < \mathrm{e}^{-(2n(r)-1)}$. Hence, the local dimensions of a subshift can be characterised as follows.

Proposition 6.1.3. Let μ be a Borel probability measure supported on a subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$. For all $x \in X$, the lower and upper local dimensions of μ at x are given by

$$\begin{aligned}
\underline{\dim}_{\mathrm{loc}}(\mu, x) &= \liminf_{n\to\infty} \frac{\log \mu\big([x_{[-n,n]}]\big)}{2n+1} \\
\overline{\dim}_{\mathrm{loc}}(\mu, x) &= \limsup_{n\to\infty} \frac{\log \mu\big([x_{[-n,n]}]\big)}{2n+1}.
\end{aligned}$$

The *multifractal spectrum* quantifies the size of the set of points that have a given local dimension. For $\alpha \geq 0$, we define

$$F_\mu(\alpha) = \{x \in X \colon \dim_{\mathrm{loc}}(\mu, x) = \alpha\}. \tag{6.1}$$

We quantify the size of each of the sets $F_\mu(\alpha)$ via their *Hausdorff dimension*, which is defined as follows.

Definition 6.1.4. Let E be a subset of a compact metric space X and let $s \geq 0$. For each $\delta > 0$, let

$$\mathcal{H}^s_\delta(E) = \inf \left\{ \sum_{i=1}^{\infty} (\mathrm{diam}(U_i))^s \colon \{U_i\} \text{ is a countable cover of } E \text{ by sets of diameter at most } \delta \right\}.$$

The *s-dimensional Hausdorff measure* of X is defined by

$$\mathcal{H}^s(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E).$$

The *Hausdorff dimension* of E is then defined by

$$\dim_{\mathrm{H}} E = \inf\{s \geq 0 \colon \mathcal{H}^s(E) = 0\}.$$

We use the convention that $\dim_{\mathrm{H}} \varnothing = -\infty$, which is standard in multifractal analysis.

Definition 6.1.5. Let μ be a Borel measure supported on a compact metric space X. The *multifractal spectrum* of μ is the function $f_\mu \colon [0, \infty) \to [0, \infty)$ defined by

$$f_\mu(\alpha) = \dim_{\mathrm{H}} F_\mu(\alpha)$$

for all $\alpha \geq 0$.

Related to the local dimensions and multifractal spectrum is the notion of Hausdorff dimension of a measure. While the multifractal spectrum provides information about the *local* scaling behaviour of a measure, the Hausdorff dimension is a *global* property.

Definition 6.1.6. Let μ be a Borel probability measure supported on a compact metric space X. The *Hausdorff dimension of* μ is defined by

$$\dim_{\mathrm{H}} \mu = \inf\{\dim_{\mathrm{H}} E : \mu(E) > 0\},$$

where the infimum is taken over all Borel-measurable sets E.

Proposition 6.1.7 ([23, Prop. 10.1])**.** Let μ be a Borel probability measure supported on a compact metric space X. Then,

$$\dim_{\mathrm{H}} \mu = \sup\{s : \underline{\dim}_{\mathrm{loc}}(\mu, x) \geq s \text{ for } \mu\text{-a.e. } x\}.$$

6.1.2 The L^q-spectrum

The *L^q-spectrum* is a well-studied quantity in multifractal analysis which encodes scaling properties of a measure, in a weak sub-exponential sense. The L^q-spectrum and multifractal spectrum are closely related, however the L^q-spectrum is generally easier to compute. In the following, we give the definition for the L^q-spectrum of a shift-invariant measure supported on a subshift. We note that it is possible to define the L^q-spectrum more generally for measures on arbitrary metric spaces, however, the following is sufficient for our purposes.

Definition 6.1.8. Let X be a subshift and μ be a shift-invariant Borel probability measure supported on X. The *L^q-spectrum of* μ is defined by

$$\tau_\mu(q) = \liminf_{n\to\infty} -\frac{1}{n} \log \left(\sum_{u \in \mathcal{L}^n(X)} \mu([u])^q \right).$$

We also define the upper variant $\overline{\tau}_\mu$ by taking limit superior in place of the limit inferior.

We first list some basic properties of the L^q-spectrum of the measure μ. Here, (a) is well-known and is a routine consequence of Hölder's inequality and (b) is proved in [72, Lemma 1.4].

Lemma 6.1.9. Let μ be a shift-invariant measure on X.

(a) The L^q-spectrum τ_μ is continuous, increasing and concave on $\mathbb{R}$.

(b) Let $\alpha_{\min} = \lim_{q\to\infty} \tau_\mu(q)/q$ and $\alpha_{\max} = \lim_{q\to-\infty} \tau_\mu(q)/q$. Then, for every $s < \alpha_{\min} \leq \alpha_{\max} < t$, all n sufficiently large and $u \in \mathcal{L}^n$, $e^{-tn} \leq \mu([u]) \leq e^{-sn}$. In particular, the local dimensions satisfy

$$\alpha_{\min} \leq \inf_{x\in X} \underline{\dim}_{\mathrm{loc}}(\mu, x) \leq \sup_{x\in X} \overline{\dim}_{\mathrm{loc}}(\mu, x) \leq \alpha_{\max}.$$

We also note that the L^q-spectrum encodes both the measure theoretic and topological entropy. In the following, (a) follows immediately from the definition and (b) was proved in [26, Theorem 1.4].

Proposition 6.1.10. Let μ be a fully-supported shift-invariant measure supported on a subshift X. Then,

(a) $\tau_\mu(0) = h_{\mathrm{top}}(X)$, and

(b) if τ_μ is differentiable at $q = 1$, then $\tau_\mu'(1) = h_\mu(X)$.

We conclude this section by presenting a technical result concerning the L^q-spectrum that will be useful in later proofs.

Lemma 6.1.11. Let μ be a shift-invariant probability measure supported on a subshift X and let $\zeta > 1$. Then

$$\underline{\tau}_\mu(q) = \frac{1}{\zeta} \liminf_{n\to\infty} -\frac{1}{n} \log \left(\sum_{u\in\mathcal{L}^{\lfloor \zeta n\rfloor}(X)} \mu([u])^q \right) \tag{6.2}$$

and

$$\overline{\tau}_\mu(q) = \frac{1}{\zeta} \limsup_{n\to\infty} -\frac{1}{n} \log \left(\sum_{u\in\mathcal{L}^{\lfloor \zeta n\rfloor}(X)} \mu([u])^q \right). \tag{6.3}$$

Proof. It follows by standard properties of limits inferior and superior that

$$\underline{\tau}_\mu(q) \leq \frac{1}{\zeta} \liminf_{n\to\infty} -\frac{1}{n} \log \left(\sum_{u\in\mathcal{L}^{\lfloor \zeta n\rfloor}(X)} \mu([u])^q \right)$$
$$\overline{\tau}_\mu(q) \geq \frac{1}{\zeta} \limsup_{n\to\infty} -\frac{1}{n} \log \left(\sum_{u\in\mathcal{L}^{\lfloor \zeta n\rfloor}(X)} \mu([u])^q \right),$$

so it suffices to show the opposite inequalities. For ease of notation, we write $S_{k,\mu}(q) = \sum_{u\in\mathcal{L}^k(X)} \mu([u])^q$ for each $k \in \mathbb{N}$ and $q \in \mathbb{R}$. First, let $q < 0$ and let $n \in \mathbb{N}$ be arbitrary. Let k_n be minimal so that $\lfloor \zeta k_n \rfloor \geq n$. Observe that there is some $M \in \mathbb{N}$ (independent of n) so that $\lfloor \zeta k_n \rfloor \leq n + M$: it follows that $\lim_{n\to\infty} n/k_n = \zeta$. Then if $v \in \mathcal{L}^{\lfloor \zeta k_n \rfloor}(X)$ is arbitrary, $[v] \subset [u]$ for some $u \in \mathcal{L}^n(X)$ and $\mu([v])^q \geq \mu([u])^q$. Thus

$$S_{\lfloor \zeta k_n \rfloor,\mu}(q) \geq S_{n,\mu}(q).$$

which gives (6.2) for $q < 0$ since $\lim_{n\to\infty} n/k_n = \zeta$.

Similarly, for $q \geq 0$, since there are at most $(\#\mathcal{A})^M$ words $v \in \mathcal{L}^{\lfloor \zeta k_n \rfloor}(X)$ with $[v] \subseteq [u]$, for each $u \in \mathcal{L}^n(X)$ there is some $v(u) \in \mathcal{L}^{\lfloor \zeta k_n \rfloor}(X)$ such that $\mu([v(u)])^q \geq (\#\mathcal{A})^{-qM}\mu([u])^q$. Hence,

$$S_{\lfloor \zeta k_n \rfloor,\mu}(q) \geq (\#\mathcal{A})^{-qM} S_{n,\mu}(q).$$

This gives (6.2) for $q \geq 0$. The arguments for (6.3) follow similarly, by choosing k_n maximal so that $\lfloor \zeta k_n \rfloor \leq n$. $\square$

6.1.3 Relation between the L^q-spectrum and multifractal spectrum

In general, it is difficult to compute the multifractal spectrum of a measure. However, in some cases, the multifractal coincides with the *concave conjugate* of the L^q-spectrum.

Definition 6.1.12. Let $g \colon \mathbb{R} \to \mathbb{R}$ be a concave function. The concave conjugate of g is the function g^* defined by

$$g^*(\alpha) = \inf_{q\in\mathbb{R}} \{q\alpha - g(q)\}.$$

The function g^* is itself concave since it is the infimum of a family of affine functions. For more detail concerning the theory of concave functions, we refer the reader to [68].

In general, the concave conjugate of the L^q-spectrum provides an upper bound for the multifractal spectrum (this is the content of Proposition 6.1.13). Further, in some cases, it coincides exactly with the multifractal spectrum. When this is the case, we say that the *multifractal formalism* holds. The multifractal formalism is a powerful tool for computing the multifractal spectrum of a measure. As such, determining conditions under which it holds is a central question in multifractal analysis.

Proposition 6.1.13. Let μ be a shift-invariant measure on a subshift X. Then $f_\mu(\alpha) \leq \tau_\mu^*(\alpha)$ for all $\alpha \in \mathbb{R}$.

Proof. Let $\alpha \in \mathbb{R}$, $n \in \mathbb{N}$ and $\varepsilon > 0$, and let

$$\mathcal{M}_{n,\epsilon}(\alpha) = \Big\{u \in \mathcal{L}^n(X) \colon \mathrm{e}^{-(2n+1)(\alpha+\epsilon)} \leq \mu([u]) \leq \mathrm{e}^{-(2n+1)(\alpha-\epsilon)}\Big\}.$$

In other words, $\mathcal{M}_{n,\epsilon}(\alpha)$ is an ϵ-approximation of $F_\mu(\alpha)$ at level n. Our strategy is to control the size of the sets $\mathcal{M}_{n,\epsilon}(\alpha)$ in terms of the L^q-spectrum of μ, and then use these sets to build a good cover of $F_\mu(\alpha)$. Since τ_μ^* is a concave function, the left and right derivatives exist at every point. Let q^- and q^+ denote the left and right derivative of τ_μ^* at α, respectively. By concavity, $q^- \geq q^+$. Let $q \in [q^+, q^-]$. For the remainder of the proof, we assume that $q \geq 0$; the case $q < 0$ is analogous. Observe that

$$\sum_{u \in \mathcal{L}^n(X)} \mu([u])^q \geq \sum_{u \in \mathcal{M}_{n,\varepsilon}(\alpha)} \mu([u])^q \geq \mathrm{e}^{-(2n+1)(\alpha+\varepsilon)q} \# \mathcal{M}_{n,\varepsilon}(\alpha). \tag{6.4}$$

Since, by Lemma 6.1.11, $\tau_\mu(q) = \liminf_{n\to\infty} (\sum_{u \in \mathcal{L}^{2n+1}(X)} \mu([u])^q)/(2n+1)$, there is a positive integer N_ε such that for all $n \geq N$, we have $\sum_{u \in \mathcal{L}^n(X)} \mu([u])^q < \mathrm{e}^{-(2n+1)(\tau_\mu(q)-\varepsilon)}$. Further, since $q \in [q^+, q^-]$, we have that $\alpha q - \tau_\mu(q) = \tau_\mu^*(\alpha)$, so it follows by (6.4) that,

$$\# \mathcal{M}_{n,\epsilon}(\alpha) \leq e^{-(2n+1)(\tau(q)-\epsilon)} \cdot e^{(2n+1)(\alpha+\epsilon)q} = e^{(2n+1)(\tau^*(\alpha)+(q+1)\epsilon)} \tag{6.5}$$

for all $n \geq N_\epsilon$.

Now for each $x \in F_\mu(\alpha)$, we can find some $n_x \in \mathbb{N}$ so that for all $n \geq n_x$, $\mu([x_{[-n,n]}]) \geq e^{-(2n+1)(\alpha+\epsilon)}$. In particular,

$$\mathcal{G}_\epsilon \coloneqq \bigcup_{n=N_\epsilon}^{\infty} \mathcal{M}_{n,\epsilon}(\alpha)$$

is a Vitali cover for $F_\mu(\alpha)$.

Now suppose $\{I_j\}_{j=1}^\infty$ is any disjoint sub-collection of $\mathcal{G}_\epsilon$. Then, setting $s = \tau^*(\alpha) + 2\epsilon(1+q)$, we have

$$\begin{aligned}
\sum_{j=1}^{\infty}(\operatorname{diam} I_j)^s &\leq \sum_{n=N_\epsilon}^{\infty} \sum_{u\in\mathcal{M}_{n,\epsilon}(\alpha)} (\operatorname{diam}[u])^s \leq \sum_{n=N_\epsilon}^{\infty} e^{-(2n+1)s} \#\mathcal{M}_{n,\epsilon}(\alpha) \\
&\leq \sum_{n=N_\epsilon}^{\infty} e^{-(2n+1)s} e^{(2n+1)(\tau^*(\alpha)+(q+1)\epsilon)} \\
&= \sum_{n=N_\epsilon}^{\infty} (e^{-(1+q)\epsilon})^{2n+1} < \infty,
\end{aligned}$$

where the third inequality follows by (6.5). Thus, by the Vitali covering theorem, there is a cover $\{E_i\}_{i=1}^{\infty}$ for $F_\mu(\alpha)$ such that

$$\mathcal{H}^s(F_\mu(\alpha)) \leq \sum_{i=1}^{\infty}(\operatorname{diam} E_i)^s < \infty$$

and so $\dim_{\mathrm{H}} F_\mu(\alpha) \leq \tau^*(\alpha) + 2\epsilon(1+q)$. Since $\epsilon > 0$ was arbitrary, the desired result follows. □

6.2 L^q-spectra of frequency measures

6.2.1 Inflation word L^q-spectra

Given a primitive random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$, we can define an analogue of the L^q-spectrum in terms of its production probabilities, in a similar manner to the inflation word analogue of measure theoretic entropy introduced in [37]. We will see that, in many cases, this notion coincides with the L^q-spectrum of the frequency measure associated to $\vartheta_{\mathbf{P}}$. For each $k \in \mathbb{N}$ and $q \in \mathbb{R}$, we define

$$\varphi_k(q) = -\sum_{a\in\mathcal{A}} R_a \log\left(\sum_{s\in\vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = s]^q\right),$$

where $\mathbf{R} = (R_a)_{a\in\mathcal{A}}$ is the right Perron–Frobenius eigenvector of $\vartheta_{\mathbf{P}}$. We define the *inflation word L^q-spectrum of $\vartheta_{\mathbf{P}}$* by

$$T_{\vartheta,\mathbf{P}}(q) = \liminf_{k\to\infty} \frac{\varphi_k(q)}{\lambda^k}.$$

We similarly define the upper variant $\overline{T}_{\vartheta,\mathbf{P}}$ by taking a limit supremum in place of the limit infimum.

We first state some key properties of $T_{\vartheta,\mathbf{P}}(q)$ that follow easily from the definition. Firstly, if $\vartheta_{\mathbf{P}}$ is a primitive and compatible random substitution that satisfies either the disjoint set condition or identical set condition with identical production probabilities, then the limit defining $T_{\vartheta,\mathbf{P}}(q)$ exists for all $q \in \mathbb{R}$ and is given by a closed-form expression. We will see later that, for $q \geq 0$, these expressions transfer to the L^q-spectrum.

Proposition 6.2.1. Let $\vartheta_{\mathbf{P}}$ be a primitive and compatible random substitution and $q \in \mathbb{R}$. If $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition, then the limit defining $T_{\vartheta,\mathbf{P}}(q)$ exists and

$$T_{\vartheta,\mathbf{P}}(q) = \frac{1}{\lambda - 1}\varphi_1(q).$$

If $\vartheta_{\mathbf{P}}$ satisfies the identical set condition and has identical production probabilities, then the limit defining $T_{\vartheta,\mathbf{P}}(q)$ exists and

$$T_{\vartheta,\mathbf{P}}(q) = \frac{1}{\lambda}\varphi_1(q).$$

Proof. Fix $q \in \mathbb{R}$. By the Markov property of $\vartheta_{\mathbf{P}}$, for all $a \in \mathcal{A}$, $k \in \mathbb{N}$ and $v \in \vartheta^k(a)$,

$$\mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v] = \sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]\, \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = v]. \tag{6.6}$$

If $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition, then for every $v \in \vartheta^k(a)$ there is a unique $s(v) \in \vartheta(a)$ such that $v \in \vartheta^{k-1}(s(v))$. Thus, for all $s \in \vartheta(a)$ such that $s \neq s(v)$, we have $\mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = v] = 0$, and so it follows by (6.6) that

$$\begin{aligned}
\sum_{v \in \vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v]^q &= \sum_{v \in \vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s(v)]^q\, \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s(v)) = v]^q \\
&= \sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \sum_{u \in \vartheta^{k-1}(s)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = u]^q \\
&= \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right) \cdot \prod_{b \in \mathcal{A}} \left(\sum_{u \in \vartheta^{k-1}(b)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(b) = u]^q \right)^{|\vartheta(a)|_b},
\end{aligned}$$

where in the final equality we use compatibility to split the second sum into inflation tiles. Thus,

$$\begin{aligned}\varphi_k(q) = &-\sum_{a\in\mathcal{A}} R_a \sum_{b\in\mathcal{A}} |\vartheta(a)|_b \log\left(\sum_{u\in\vartheta^{k-1}(b)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(b) = u]^q\right) \\ &-\sum_{a\in\mathcal{A}} R_a \log\left(\sum_{s\in\vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q\right) \\ = &\ \lambda\varphi_{k-1}(q) + \varphi_1(q),\end{aligned}$$

noting that $\sum_{a\in\mathcal{A}} R_a|\vartheta(a)|_b = \lambda R_b$. It follows inductively that

$$\frac{1}{\lambda^k}\varphi_k(q) = \sum_{j=1}^{k} \frac{1}{\lambda^j}\varphi_1(q) \xrightarrow{k\to\infty} \frac{1}{\lambda - 1}\varphi_1(q),$$

so the limit defining $T_{\vartheta,\mathbf{P}}(q)$ exists and is equal to $(\lambda-1)^{-1}\varphi_1(q)$.

On the other hand, if $\vartheta_{\mathbf{P}}$ satisfies the identical set condition and has identical production probabilities, then $\mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s^1) = u] = \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s^2) = u]$ for all $s^1, s^2 \in \vartheta(a)$. Hence, it follows by (6.6) that

$$\sum_{v\in\vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v]^q = \sum_{v\in\vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = v]^q$$

for any choice of $s \in \vartheta(a)$. By compatibility and the Markov property of $\vartheta_{\mathbf{P}}$, we have

$$\sum_{v\in\vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v]^q = \prod_{b\in\mathcal{A}} \left(\sum_{u\in\vartheta^{k-1}(b)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(b) = u]^q\right)^{|\vartheta(a)|_b};$$

thus,

$$\varphi_k(q) = \sum_{b\in\mathcal{A}}\sum_{a\in\mathcal{A}} R_a|\vartheta(a)|_b \log\left(\sum_{v\in\vartheta^{k-1}(b)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(b) = v]^q\right) = \lambda\varphi_{k-1}(q),$$

noting that $\sum_{a\in\mathcal{A}} R_a|\vartheta(a)|_b = R_b$. It follows by induction that $\varphi_k(q)/\lambda^k = \varphi_1(q)/\lambda$ for all $k \in \mathbb{N}$, so we conclude that $T_{\vartheta,\mathbf{P}}(q)$ exists and equals $\lambda^{-1}\varphi_1(q)$. □

Proposition 6.2.2. Let $\vartheta_{\mathbf{P}}$ be a primitive and compatible random substitution. For all $q > 1$ and $q < 0$, the sequence $(\lambda^{-k}\varphi_k(q))_k$ is non-decreasing; for all $0 < q < 1$, the sequence is non-increasing.

Proof. This is largely a consequence of Jensen's inequality. Note that on the interval $(0, 1]$, the function $x \mapsto x^q$ is convex if $q > 1$ or $q < 0$, and concave if $0 < q < 1$. We prove the claim for the case $q > 1$ or $q < 0$; the other case follows similarly. For all $a \in \mathcal{A}$, $k \in \mathbb{N}$ with $k \geq 2$ and $v \in \vartheta^k(a)$, it follows by the Markov property of $\vartheta_{\mathbf{P}}$ that

$$\begin{aligned}
\sum_{v \in \vartheta^k(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v]^q &= \sum_{v \in \vartheta^k(a)} \left(\sum_{s \in \vartheta(a)\colon v \in \vartheta^{k-1}(s)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s] \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = v] \right)^q \\
&\leq \sum_{v \in \vartheta^k(a)} \left(\frac{\sum_{s \in \vartheta(a)\colon v \in \vartheta^{k-1}(s)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s] \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = v]^q}{\sum_{s \in \vartheta(a)\colon v \in \vartheta^{k-1}(s)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]} \right) \\
&= \prod_{b \in \mathcal{A}} \left(\sum_{w \in \vartheta^{k-1}(b)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(b) = w]^q \right)^{|\vartheta(a)|_b} .
\end{aligned}$$

In the second line we have applied Jensen's inequality and in the third we have used compatibility to decompose each probability $\mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(s) = w]$ into inflation tiles. It follows that

$$\frac{1}{\lambda^k} \varphi_k(q) \geq -\frac{1}{\lambda^k} \sum_{b \in \mathcal{A}} R_b \sum_{a \in \mathcal{A}} R_a |\vartheta(a)|_b \log \left(\sum_{w \in \vartheta^{k-1}(b)} \mathbb{P}[\vartheta_{\mathbf{P}}^{k-1}(b) = w]^q \right) = \frac{1}{\lambda^{k-1}} \varphi_{k-1}(q),$$

noting that $\sum_{a \in \mathcal{A}} R_a |\vartheta(a)|_b = \lambda$.

The $0 < q < 1$ case follows by similar arguments, with Jensen's inequality giving the opposite inequality since $x \mapsto x^q$ is concave. $\square$

6.2.2 L^q-spectra for non-negative q

For $q \geq 0$, we show that for every primitive and compatible random substitution, the L^q-spectrum of the corresponding frequency measure coincides with the inflation word analogue introduced in Section 6.2.1. In particular, we have the following.

Theorem 6.2.3. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive and compatible random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then, the limits defining $\tau_{\mu_{\mathbf{P}}}(q)$ and $T_{\vartheta,\mathbf{P}}(q)$ exist and coincide for all $q \geq 0$. Moreover, the following bounds hold.

(1) For all $0 \leq q \leq 1$,

$$\frac{1}{\lambda^k - 1} \varphi_k(q) \leq \tau_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda^k} \varphi_k(q) \tag{6.7}$$

and $(\lambda^{-k} \varphi_k(q))_{k=1}^{\infty}$ converges monotonically to $\tau_{\mu_{\mathbf{P}}}(q)$ from above.

(2) For all $q \geq 1$,

$$\frac{1}{\lambda^k}\varphi_k(q) \leq \tau_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda^k - 1}\varphi_k(q) \tag{6.8}$$

and $(\lambda^{-k}\varphi_k(q))_{k=1}^{\infty}$ converges monotonically to $\tau_{\mu_{\mathbf{P}}}(q)$ from below.

We give the proof of Theorem 6.2.3 at the end of this subsection. If $\vartheta_{\mathbf{P}}$ satisfies the disjoint or identical set condition, then Proposition 6.2.1 and Theorem 6.2.3 together provide a closed-form formula for the L^q-spectrum of the corresponding frequency measure.

Corollary 6.2.4. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive and compatible random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$ and let $q \geq 0$.

(1) If $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition, then

$$\tau_{\mu_{\mathbf{P}}}(q) = \frac{1}{\lambda - 1}\varphi_1(q).$$

(2) If $\vartheta_{\mathbf{P}}$ satisfies the identical set condition with identical production probabilities, then

$$\tau_{\mu_{\mathbf{P}}}(q) = \frac{1}{\lambda}\varphi_1(q).$$

In particular, if $\vartheta_{\mathbf{P}}$ satisfies the disjoint set condition or identical set condition with identical production probabilities, then $\tau_{\mu_{\mathbf{P}}}$ is analytic on $(0, \infty)$.

We now apply Theorem 6.2.3 to calculate the L^q-spectrum for some familiar examples. First, we obtain a closed-form expression for the L^q-spectrum on $[0, \infty)$ for the frequency measure corresponding to the random period doubling substitution.

Example 6.2.5. Given $p \in (0, 1)$, let $\vartheta_{\mathbf{P}}$ be the random period doubling substitution

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. The random substitution $\vartheta_{\mathbf{P}}$ satisfies

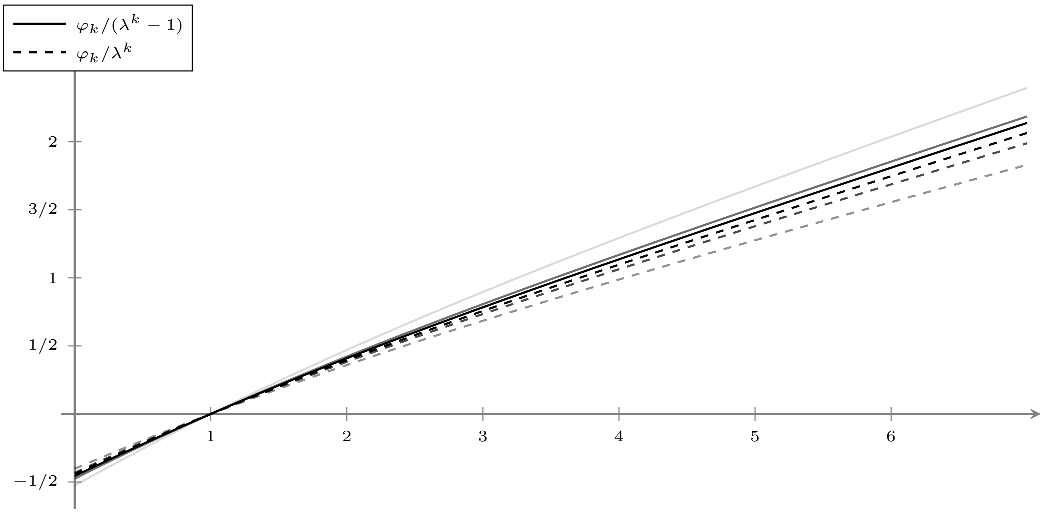


Figure 6.1: Lower and upper bounds on the L^q-spectrum of the frequency measure corresponding to the random Fibonacci substitution with $p = 1/2$, for $k = 3$ (faint lines), $k = 5$ (grey lines) and $k = 7$ (black lines).

the disjoint set condition, so Corollary 6.2.4 gives that

$$\tau_{\mu_p}(q) = \frac{1}{\lambda - 1}\varphi_1(q) = -\frac{2}{3}\log(p^q + (1-p)^q)$$

for all $q \geq 0$.

As we have already seen, the random Fibonacci substitution does not satisfy the identical set condition or disjoint set condition. However, in a similar vein to measure theoretic entropy, we can obtain a sequence of converging lower and upper bounds for the L^q-spectrum of the corresponding frequency measure from Theorem 6.2.3.

Example 6.2.6. Given $p \in (0,1)$, let $\vartheta_{\mathbf{P}}$ be the random Fibonacci substitution

$$\vartheta_p\colon \begin{cases} a \mapsto \begin{cases} ab & \text{with probability } p, \\ ba & \text{with probability } 1-p, \end{cases} \\ b \mapsto a \quad \text{with probability } 1, \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. Theorem 6.2.3 provides lower and upper bounds for $\tau_{\mu_{\mathbf{P}}}$ which hold for all $q \geq 0$. See Figure 6.1 for a plot of these bounds for $k \in \{3, 5, 7\}$, in the case that $p = 1/2$. The bounds displayed were obtained computationally.

The majority of the work in proving Theorem 6.2.3 lies in proving the bounds in (6.7) and

(6.8). To achieve this, we will make use of the following result, which follows routinely from standard properties of compatible random substitutions.

Lemma 6.2.7. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive and compatible random substitution and let $q \geq 1$. For all $\varepsilon > 0$, there is an $M \in \mathbb{N}$ such that for every $m \geq M$ and $v \in \mathcal{L}_{\vartheta}^{m}$,

$$\prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{m(R_a + \varepsilon)} \leq \sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q \leq \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{m(R_a - \varepsilon)} .$$

For $q \leq 1$, the same result holds with reversed inequalities.

Proof. Since $\vartheta_{\mathbf{P}}$ is compatible, the cutting points of inflation tiles are well-defined, so breaking the sum into inflation tiles we obtain

$$\begin{aligned} \sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q &= \sum_{w^1 \in \vartheta(v_1)} \mathbb{P}[\vartheta_{\mathbf{P}}(v_1) = w^1]^q \sum_{w^2 \in \vartheta(v_2)} \cdots \sum_{w^m \in \vartheta(v_m)} \mathbb{P}[\vartheta_{\mathbf{P}}(v_m) = w^m]^q \\ &= \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{|v|_a} . \end{aligned}$$

The result then follows by applying Proposition 2.2.2 to bound $|v|_a$, noting that for all $a \in \mathcal{A}$ we have $\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \leq 1$ if $q \geq 1$ and $\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \geq 1$ if $q \leq 1$. $\square$

Proposition 6.2.8. Let $\vartheta_{\mathbf{P}}$ be a primitive and compatible random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then, for all $q > 1$, we have

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda - 1} \varphi_1(q).$$

Proof. Fix $q > 1$. Let $\varepsilon > 0$ and, for each $n \in \mathbb{N}$, let $m(n)$ be the integer defined by

$$m(n) = \left\lceil \frac{n}{\lambda - \varepsilon} \right\rceil .$$

Then the integers n and $m(n)$ satisfy the conditions of Lemma 2.3.5, so it follows that

$$\sum_{u \in \mathcal{L}_{\vartheta}^{n}} \mu_{\mathbf{P}}([u])^q = \sum_{u \in \mathcal{L}_{\vartheta}^{n}} \left(\frac{1}{\lambda} \sum_{v \in \mathcal{L}_{\vartheta}^{m(n)}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{|\vartheta(v_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j, j+n-1]} = u] \right)^q .$$

Since $q > 1$, the function $x \mapsto x^q$ is super-additive on the interval $[0,1]$, so

$$\begin{aligned}\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q &\geq \sum_{u \in \mathcal{L}_\vartheta^n} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \left(\frac{1}{\lambda} \sum_{j=1}^{|\vartheta(v_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u] \right)^q \\ &\geq \frac{1}{\lambda^q} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \sum_{j=1}^{|\vartheta(v_1)|} \sum_{u \in \mathcal{L}_\vartheta^n} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q.\end{aligned}$$

We now bound the probability on the right of this expression by the production probability of an inflation word. If $w(u) \in \vartheta(v)$ contains u as a subword in position j, then $\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u] \geq \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w(u)]$. Hence,

$$\sum_{u \in \mathcal{L}_\vartheta^n} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q \geq \sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q$$

for all $j \in \{1, \ldots, |\vartheta(v_1)|\}$.

Since $\vartheta_{\mathbf{P}}$ is compatible, by Lemma 6.2.7 there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and all $v \in \mathcal{L}_\vartheta^{m(n)}$

$$\sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q \geq \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{m(n)(R_a + \varepsilon)}.$$

Hence,

$$\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \geq \frac{1}{\lambda^q} \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{m(n)(R_a + \varepsilon)} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q.$$

Taking logarithms, rearranging and dividing by n gives

$$\begin{aligned}-\frac{1}{n} \log \left(\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \right) \leq\ &-\frac{1}{n} \log \left(\sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \right) + \frac{1}{n} \log \lambda^q \\ &- \frac{m(n)}{n} \sum_{a \in \mathcal{A}} (R_a + \varepsilon) \log \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right).\end{aligned}$$

Noting that $m(n)/n \to (\lambda - \varepsilon)^{-1}$ as $n \to \infty$, it follows by Lemma 6.1.11 that

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda - \varepsilon}\overline{\tau}_{\mu_{\mathbf{P}}}(q) + \frac{1}{\lambda - \varepsilon}\sum_{a \in \mathcal{A}} \log\left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q\right) + c\varepsilon$$

where $c := (\#\mathcal{A}) \max_{a \in \mathcal{A}} \log(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q)$. But $\varepsilon > 0$ was arbitrary; letting $\varepsilon \to 0$ and rearranging we obtain

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda - 1}\varphi_1(q),$$

which completes the proof. □

We now prove the corresponding lower bound.

Proposition 6.2.9. For all $q > 1$,

$$\tau_{\mu_{\mathbf{P}}}(q) \geq \frac{1}{\lambda}\varphi_1(q).$$

Proof. Let $\varepsilon > 0$ and, for each $n \in \mathbb{N}$, let $m(n)$ be the integer defined by

$$m(n) = \left\lceil \frac{n}{\lambda - \varepsilon} \right\rceil.$$

Since $q > 1$, the function $x \mapsto x^q$ is convex on the interval $[0, 1]$. Hence, it follows by Lemma 2.3.5 and two applications of Jensen's inequality that

$$\begin{aligned}
\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q &= \sum_{u \in \mathcal{L}_\vartheta^n} \left(\frac{1}{\lambda} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{|\vartheta(v_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]\right)^q \\
&\leq \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v]) \sum_{u \in \mathcal{L}_\vartheta^n} \left(\frac{1}{\lambda} \sum_{j=1}^{|\vartheta(v_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]\right)^q \\
&\leq \frac{|\vartheta|^{q-1}}{\lambda^q} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v]) \sum_{j=1}^{|\vartheta(v_1)|} \sum_{u \in \mathcal{L}_\vartheta^n} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q.
\end{aligned}$$

We bound above the probability on the right of this expression by the production probability of a sufficiently large inflation word contained in u. By compatibility, there is an integer $k(n)$ such that $j + n \leq |\vartheta(v_{[1,m(n)-k(n)]})|$ for all $n \in \mathbb{N}$ and $v \in \mathcal{L}_\vartheta^{m(n)}$, where $\lim k(n)/n = 0$. In particular, for every $v \in \mathcal{L}_\vartheta^n$, a realisation of $\vartheta(v_{[2,m(n)-k(n)]})$ is contained in u as an inflation word, so

$$\sum_{u \in \mathcal{L}_\vartheta^n} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q \leq \sum_{w \in \vartheta(v_2 \cdots v_{m(n)-k(n)})} \mathbb{P}[\vartheta_{\mathbf{P}}(v_2 \cdots v_{m(n)-k(n)}) = w]^q.$$

We now bound this quantity uniformly for all $v \in \mathcal{L}_\vartheta^{m(n)}$. By Lemma 6.2.7 and the above, there is an $N \in \mathbb{N}$ such that for all $n \geq N$

$$\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \leq \frac{|\vartheta|^{q-1}}{\lambda^q} \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{(m(n)-k(n)-1)(R_a - \varepsilon)}.$$

Taking logarithms, rearranging and dividing by n gives

$$\begin{aligned} -\frac{1}{n} \log \left(\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \right) &\geq \frac{m(n) - k(n) - 1}{n} \sum_{a \in \mathcal{A}} (R_a - \varepsilon) \log \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right) \\ &\quad - \frac{\log(|\vartheta|^{q-1}/\lambda^q)}{n} \\ &\xrightarrow{n \to \infty} \frac{1}{\lambda - \varepsilon} \sum_{a \in \mathcal{A}} (R_a - \varepsilon) \log \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right), \end{aligned}$$

But $\varepsilon > 0$ was arbitrarily, so

$$\tau_{\mu_{\mathbf{P}}}(q) \geq \frac{1}{\lambda} \varphi_1(q),$$

which completes the proof. □

We now state the bounds for the $q \in (0,1)$ case. We do not give a proof here since the arguments mirror the proofs of Propositions 6.2.8 and 6.2.9, except with reversed inequalities: since $x \mapsto x^q$ is concave rather than convex and subadditive as opposed to superadditive.

Proposition 6.2.10. If $q \in (0,1)$, then

$$\frac{1}{\lambda - 1} \varphi_1(q) \leq \tau_{\mu_{\mathbf{P}}}(q) \leq \overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda} \varphi_1(q).$$

We are now in a position to prove Theorem 6.2.3.

Proof of Theorem 6.2.3. By primitivity, for each $k \in \mathbb{N}$, the random substitution $\vartheta_{\mathbf{P}}^k$ gives rise to the same frequency measure as $\vartheta_{\mathbf{P}}$. Applying Propositions 6.2.8, 6.2.9 and 6.2.10 to $\vartheta_{\mathbf{P}}^k$, we obtain that

$$\frac{1}{\lambda^k}\varphi_k(q) \le \tau_{\mu_{\mathbf{P}}}(q) \le \overline{\tau}_{\mu_{\mathbf{P}}}(q) \le \frac{1}{\lambda^k - 1}\varphi_k(q)$$

for all $q > 1$ and

$$\frac{1}{\lambda^k - 1}\varphi_k(q) \le \tau_{\mu_{\mathbf{P}}}(q) \le \overline{\tau}_{\mu_{\mathbf{P}}}(q) \le \frac{1}{\lambda^k}\varphi_k(q)$$

for all $0 < q < 1$. Letting $k \to \infty$ gives

$$\tau_{\mu_{\mathbf{P}}}(q) = \overline{\tau}_{\mu_{\mathbf{P}}}(q) = T_{\vartheta,\mathbf{P}}(q) = \overline{T}_{\vartheta,\mathbf{P}}(q)$$

for all $q \in (0,1) \cup (1,\infty)$, so the limits defining $\tau_{\mu_{\mathbf{P}}}(q)$ and $T_{\vartheta,\mathbf{P}}(q)$ both exist and coincide. The same holds for $q = 0$ and $q = 1$ by continuity. The monotonicity of the bounds $\lambda^{-k}\varphi_k(q)$ follows by Proposition 6.2.2. □

6.2.3 L^q-spectra for negative q

For $q < 0$, the inflation word L^q-spectrum provides a general lower bound for the L^q-spectrum.

Proposition 6.2.11. Let $\vartheta_{\mathbf{P}}$ be a primitive and compatible random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then, for all $k \in \mathbb{N}$ and $q < 0$, we have

$$\tau_{\mu_{\mathbf{P}}}(q) \ge \frac{1}{\lambda^k - 1}\varphi_k(q). \tag{6.9}$$

In particular,

$$\tau_{\mu_{\mathbf{P}}}(q) \ge \overline{T}_{\vartheta,\mathbf{P}}(q) \ge T_{\vartheta,\mathbf{P}}(q).$$

Proof. We prove (6.9) for $k = 1$. The bounds for $k > 1$ then follow by considering higher powers of $\vartheta_{\mathbf{P}}$. Let $\varepsilon > 0$ be sufficiently small and for n sufficiently large, let $m(n)$ be the integer defined by

$$m(n) = \left\lceil \frac{n}{\lambda - \varepsilon} \right\rceil.$$

To avoid division by zero, we rewrite Lemma 2.3.5 in a form where we do not sum over elements equal to zero. Here, we write $u \blacktriangleleft \vartheta(v)$ to mean there exists $w \in \vartheta(v)$ for which u appears as a subword of w. For each $v \in \mathcal{L}_{\vartheta}^{m(n)}$ and $u \in \mathcal{L}_{\vartheta}^{n}$, let $\mathcal{J}(v,u) = \{j \in \{1,\ldots,|\vartheta(v_1)|\} : u \in \vartheta(v)_{[j,j+n-1]}\}$. If $j \notin \mathcal{J}(u,v)$, then $\mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u] = 0$, and if u does not appear as a subword of any realisations of $\vartheta(v)$, then $\mathcal{J}(u,v)=\varnothing$. Therefore, we can rewrite Lemma 2.3.5 as

$$\mu_{\mathbf{P}}([u]) = \frac{1}{\lambda} \sum_{\substack{v \in \mathcal{L}_\vartheta^{m(n)} \\ u \blacktriangleleft \vartheta(v)}} \mu_{\mathbf{P}}([v]) \sum_{j \in \mathcal{J}(v,u)} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u].$$

Hence, by subadditivity of the function $x \mapsto x^q$ on the domain $(0,1]$,

$$\begin{aligned}
\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q &= \sum_{u \in \mathcal{L}_\vartheta^n} \left(\frac{1}{\lambda} \sum_{\substack{v \in \mathcal{L}_\vartheta^{m(n)} \\ u \blacktriangleleft \vartheta(v)}} \mu_{\mathbf{P}}([v]) \sum_{j \in \mathcal{J}(v,u)} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u] \right)^q \\
&\leq \frac{1}{\lambda^q} \sum_{u \in \mathcal{L}_\vartheta^n} \sum_{\substack{v \in \mathcal{L}_\vartheta^{m(n)} \\ u \blacktriangleleft \vartheta(v)}} \mu_{\mathbf{P}}([v])^q \sum_{j \in \mathcal{J}(v,u)} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q \\
&= \frac{1}{\lambda^q} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \sum_{\substack{u \in \mathcal{L}_\vartheta^n \\ u \blacktriangleleft \vartheta(v)}} \sum_{j \in \mathcal{J}(v,u)} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q.
\end{aligned}$$

For each $j \in \mathcal{J}(v,u)$, let $w_j(u) \in \vartheta(v)$ be a word such that $w_j(u)_{[j,j+n-1]} = u$. Note that there are at most $K \coloneqq 2|\vartheta|(\#\mathcal{A})^{|\vartheta|}$ different $u \in \mathcal{L}_\vartheta^n$ such that $w_j(u)_{[j,j+n-1]} = u$. Hence,

$$\begin{aligned}
\sum_{\substack{u \in \mathcal{L}_\vartheta^n \\ u \blacktriangleleft \vartheta(v)}} \sum_{j \in \mathcal{J}(v,u)} \mathbb{P}[\vartheta_{\mathbf{P}}(v)_{[j,j+n-1]} = u]^q &\leq \sum_{\substack{u \in \mathcal{L}_\vartheta^n \\ u \blacktriangleleft \vartheta(v)}} \sum_{j \in \mathcal{J}(v,u)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w_j(u)]^q \\
&\leq K \sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q
\end{aligned}$$

and it follows that

$$\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \leq \lambda^{-q} K \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q.$$

Thus, by Lemma 6.2.7, for all $\varepsilon > 0$ there is an integer N such that for all $n \geq N$, we have

$$\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \leq \lambda^{-q} K \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{m(n)(R_a + \varepsilon)} \left(\sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \right).$$

Taking logarithms, rearranging and dividing by n gives

$$\begin{aligned}-\frac{1}{n}\log\left(\sum_{u\in\mathcal{L}_\vartheta^n}\mu_{\mathbf{P}}([u])^q\right) \geq &-\frac{1}{n}\log\left(\sum_{v\in\mathcal{L}_\vartheta^{m(n)}}\mu_{\mathbf{P}}([v])^q\right)+\frac{1}{n}\log(\lambda^{-q}K)\\ &-\frac{m(n)}{n}\sum_{a\in\mathcal{A}}(R_a+\varepsilon)\log\left(\sum_{s\in\vartheta(a)}\mathbb{P}[\vartheta_{\mathbf{P}}(a)=s]^q\right).\end{aligned}$$

Noting that $m(n)/n \to (\lambda-\varepsilon)^{-1}$ as $n\to\infty$, it follows by Lemma 6.1.11 that

$$\tau_{\mu_{\mathbf{P}}}(q) \geq \frac{1}{\lambda-\varepsilon}\tau_{\mu_{\mathbf{P}}}(q)+\frac{1}{\lambda-\varepsilon}\sum_{a\in\mathcal{A}}\log\left(\sum_{s\in\vartheta(a)}\mathbb{P}[\vartheta_{\mathbf{P}}(a)=s]^q\right)+c\varepsilon$$

where $c := (\#\mathcal{A})\max_{a\in\mathcal{A}}\log(\sum_{s\in\vartheta(a)}\mathbb{P}[\vartheta_{\mathbf{P}}(a)=s]^q)$. Letting $\varepsilon\to 0$ and rearranging, we obtain

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \geq \frac{1}{\lambda-1}\varphi_1(q),$$

and by considering higher powers of $\vartheta_{\mathbf{P}}^k$ we obtain (6.9) for all $k\in\mathbb{N}$. Finally, letting $k\to\infty$ in (6.9) gives that $\tau_{\mu_{\mathbf{P}}}(q) \geq \overline{T}_{\vartheta,\mathbf{P}}(q) \geq T_{\vartheta,\mathbf{P}}(q)$, which completes the proof. □

In general, the corresponding upper bound does *not* hold for $q<0$, even under compatibility. In particular, the quantities $\tau_{\mu_{\mathbf{P}}}(q)$ and $T_{\vartheta,\mathbf{P}}(q)$ need not coincide. This is illustrated by the following two examples. In the first, we show that this can occur for frequency measures on the full shift. Then, in the second we provide a compatible example, which is a slight modification of the first.

Example 6.2.12. Let $p_1 < p_2 \in (0,1)$ such that $p_1+3p_2=1$ and let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}}\colon a,b\mapsto\begin{cases} ab & \text{with probability } p_1,\\ ba & \text{with probability } p_2,\\ aa & \text{with probability } p_2,\\ bb & \text{with probability } p_2.\end{cases}$$

We show for all sufficiently small $q<0$ that $\tau_{\mu_{\mathbf{P}}}(q) > T_{\vartheta,\mathbf{P}}(q)$. Observe that, for each $k\in\mathbb{N}$, the word $v^k = (ab)^{2^k} \in \vartheta^{k+1}(a) \cap \vartheta^{k+1}(b)$ occurs with probability

$$\mathbb{P}[\vartheta_{\mathbf{P}}^{k+1}(a) = v^k] = \mathbb{P}[\vartheta_{\mathbf{P}}^{k+1}(b) = v^k] = p_1^{2^k}.$$

Since $p_1 < p_2$, this is the minimal possible probability with which a level-k inflation word can occur, so it follows that

$$\lim_{q \to -\infty} \frac{T_{\vartheta,\mathbf{P}}(q)}{q} = -\frac{1}{2} \log p_1.$$

Now, let $u \in \mathcal{L}_{\vartheta}^{2^{k+1}}$ be arbitrary. We show that $\mu_{\mathbf{P}}([u]) \geq p_1^{2^{k-1}} p_2^{2^{k-1}}/2$. Since $\vartheta(a) = \vartheta(b)$ with identical production probabilities, it follows by Lemma 2.3.5 that, for any choice of $w \in \mathcal{L}_{\vartheta}^{2^k+1}$,

$$\mu_{\mathbf{P}}([u]) = \frac{1}{2}\left(\mathbb{P}[\vartheta_{\mathbf{P}}(w)_{[1,2^{k+1}]} = u] + [\vartheta_{\mathbf{P}}(w)_{[2,2^{k+1}+1]} = u]\right).$$

If $\mathbb{P}[\vartheta_{\mathbf{P}}(w)_{[1,2^{k+1}]} = u] \geq p_1^{2^{k-1}} p_2^{2^{k-1}}$, then we are done. Otherwise, at least half of the letters in v must be mapped to ab. But then for u to appear from the second letter, at least half of the letters in v must be mapped to ba or bb, so $\mathbb{P}[\vartheta_{\mathbf{P}}(w)_{[2,2^{k+1}+1]} = u] \geq p_1^{2^{k-1}} p_2^{2^{k-1}}$. Hence, $\mu_{\mathbf{P}}([u]) \geq p_1^{2^{k-1}} p_2^{2^{k-1}}/2$ so, in particular,

$$\min_{u \in \mathcal{L}_{\vartheta}^{2^{k+1}}} \mu_{\mathbf{P}}([u]) \geq \frac{1}{2} p_1^{2^{k-1}} p_2^{2^{k-1}}.$$

It follows that

$$\lim_{q \to -\infty} \frac{\tau_{\mu_{\mathbf{P}}}(q)}{q} \leq -\frac{1}{4}(\log p_1 + \log p_2) < -\frac{1}{2} \log p_1 = \lim_{q \to -\infty} \frac{T_{\vartheta,\mathbf{P}}(q)}{q}.$$

Example 6.2.13. Let $p_1 < p_2 \in (0,1)$ such that $p_1 + 3p_2 = 1$ and let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon a, b \mapsto \begin{cases} ab\,ba & \text{with probability } p_1, \\ ba\,ab & \text{with probability } p_2, \\ ab\,ab & \text{with probability } p_2, \\ ba\,ba & \text{with probability } p_2. \end{cases}$$

By similar arguments to the previous example, it can be shown that

$$\lim_{q\to-\infty}\frac{\tau_{\mu_{\mathbf{P}}}(q)}{q}\leq-\frac{1}{8}(\log p_1+\log p_2)<-\frac{1}{4}\log p_1=\lim_{q\to-\infty}\frac{T_{\vartheta,\mathbf{P}}(q)}{q}.$$

In particular, there exists a $q_0<0$ such that $\tau_{\mu_{\mathbf{P}}}(q)<T_{\vartheta,\mathbf{P}}(q)$ for all $q<q_0$.

6.2.4 Recognisable random substitutions

A key feature of the random substitutions in Examples 6.2.12 and 6.2.13 is the existence of legal words u and v for which there are distinct $j_1,j_2\in\{1,\dots,|\vartheta(v_1)|\}$ such that $u\in\vartheta(v)_{[j_1,j_1+|v|-1]}\cap\vartheta(v)_{[j_2,j_2+|v|-1]}$. The inflation word L^q-spectrum does not capture the averaging procedure across these realisations, thus it is possible for $T_{\vartheta,\mathbf{P}}(q)$ and $\tau_{\mu_{\mathbf{P}}}(q)$ to be disparate for $q<0$. If we additionally assume recognisability, then we can exclude this possibility.

Theorem 6.2.14. Let $\vartheta_{\mathbf{P}}$ be a primitive, compatible, and recognisable random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then for all $q\in\mathbb{R}$,

$$\tau_{\mu_{\mathbf{P}}}(q)=T_{\vartheta,\mathbf{P}}(q)=\frac{1}{\lambda-1}\varphi_1(q).$$

It follows from Proposition 6.2.11 that $(\lambda-1)^{-1}\varphi_1(q)$ is a lower bound for the L^q-spectrum, so it only remains to show the upper bound. Central to the proof of this bound is the following version of the renormalisation lemma. In the proof of Theorem 6.2.14, we only require the upper bound in the following. However, we will use the lower bound when we consider the multifractal formalism in Section 6.3. Recall that, for a recognisable random substitution, the recognisable core of a legal word u with length at least twice the recognisability radius, is the largest inflation word, appearing as a subword of u, that has a unique decomposition into exact inflation words.

Lemma 6.2.15. Let $\vartheta_{\mathbf{P}}=(\vartheta,\mathbf{P})$ be a primitive and compatible random substitution, with corresponding frequency measure $\mu_{\mathbf{P}}$ and let $u\in\mathcal{L}_\vartheta$. If $v\in\mathcal{L}_\vartheta$ and $w\in\vartheta(v)$ contains u as a subword, then

$$\mu_{\mathbf{P}}([u])\geq\frac{1}{\lambda}\mu_{\mathbf{P}}([v])\mathbb{P}[\vartheta_{\mathbf{P}}(v)=w].$$

If, additionally, $\vartheta_{\mathbf{P}}$ is recognisable, $|u|>2\kappa(\vartheta)$ and w' is the recognisable core of u with $v'\in\mathcal{L}_\vartheta$

the unique legal word such that $w' \in \vartheta(v')$, then

$$\mu_{\mathbf{P}}([u]) \leq \frac{\kappa(\vartheta)}{\lambda} \mu_{\mathbf{P}}([v']) \mathbb{P}[\vartheta_{\mathbf{P}}(v') = w'].$$

Proof. If u is a subword of $w \in \vartheta(v)$, then $\mu_{\mathbf{P}}([u]) \geq \mu_{\mathbf{P}}([w])$. Thus by Lemma 2.3.5 applied to $\mu_{\mathbf{P}}([w])$,

$$\mu_{\mathbf{P}}([u]) \geq \frac{1}{\lambda} \mu_{\mathbf{P}}([v]) \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w].$$

Now, assume that $\vartheta_{\mathbf{P}}$ is recognisable, $|u| > 2\kappa(\vartheta)$ and $w' \in \vartheta(v')$ is the recognisable core of u. Let k be an integer such that every $t \in \mathcal{L}_{\vartheta}^k$ has $|\vartheta(t)| \geq k + |\vartheta(v_1)|$. Since there are at most $\kappa(\vartheta)$ letters of u preceding the recognisable core, if $t \in \mathcal{L}_{\vartheta}^k$ is a word for which $u \in \vartheta(t)_{[j,j+|u|-1]}$ for some $j \in \{1, \ldots, |\vartheta(t_1)|\}$, then $t_i \cdots t_{i+|v|-1} = v'$ for some $i \in \{1, \ldots, \kappa(\vartheta)\}$. Moreover, since there is a unique way to decompose w' into exact inflation words, for each $t \in \mathcal{L}_{\vartheta}^k$ there can be at most one $j \in \{1, \ldots, \vartheta(t_1)\}$ such that $u \in \vartheta(t)_{[j,j+|u|-1]}$. Hence, it follows by Lemma 2.3.5 that

$$\begin{aligned}
\mu_{\mathbf{P}}([u]) &= \frac{1}{\lambda} \sum_{t \in \mathcal{L}^k} \mu_{\mathbf{P}}([t]) \sum_{j=1}^{|\vartheta(t_1)|} \mathbb{P}[\vartheta_{\mathbf{P}}(t)_{[j,j+|u|-1]} = u] \\
&\leq \frac{1}{\lambda} \sum_{i=1}^{\kappa(\vartheta)} \sum_{\substack{t \in \mathcal{L}_{\vartheta}^k \\ t_i \cdots t_{i+|v|-1} = v'}} \mu_{\mathbf{P}}([t]) \mathbb{P}[\vartheta_{\mathbf{P}}(v') = w'] \\
&= \frac{\kappa(\vartheta)}{\lambda} \mu_{\mathbf{P}}([v']) \mathbb{P}[\vartheta_{\mathbf{P}}(v') = w'],
\end{aligned}$$

which completes the proof. □

Using Lemma 6.2.15, we prove the following upper bound on the L^q-spectrum for $q < 0$.

Proposition 6.2.16. If $\vartheta_{\mathbf{P}}$ is a primitive, compatible and recognisable random substitution, then, for all $k \in \mathbb{N}$ and $q < 0$, we have

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda - 1} \varphi_1(q).$$

Proof. Again, it suffices to verify the bounds in the case $k = 1$. To this end, let $\varepsilon > 0$ be sufficiently small and, for each $n \in \mathbb{N}$ sufficiently large, let $m(n)$ be the integer defined by

$$m(n) = \left\lfloor \frac{n}{\lambda - \varepsilon} \right\rfloor .$$

For each $u \in \mathcal{L}_\vartheta^{n+2\kappa(\vartheta)}$, let $w(u)$ denote the recognisable core of u. Further, let $v(u)$ denote the unique legal word such that $w(u) \in \vartheta(v(u))$. Then, by Lemma 6.2.15,

$$\mu_{\mathbf{P}}([u]) \leq \frac{\kappa(\vartheta)}{\lambda} \mu_{\mathbf{P}}([v(u)]) \mathbb{P}[\vartheta_{\mathbf{P}}(v(u)) = w(u)]. \tag{6.10}$$

For all $u \in \mathcal{L}_\vartheta^{n+2\kappa(\vartheta)}$, the recognisable core $w(u)$ has length at least n so, by compatibility, there is an integer N such that if $n \geq N$, then $|v(u)| \geq m(n)$ for all $u \in \mathcal{L}_\vartheta^{n+2\kappa(\vartheta)}$. In particular, for every u there exists a $v \in \mathcal{L}_\vartheta^{m(n)}$ such that $\mu_{\mathbf{P}}([v(u)]) \leq \mu_{\mathbf{P}}([v])$ and a $w \in \vartheta(v)$ such that $\mathbb{P}[\vartheta_{\mathbf{P}}(v(u)) = w(u)] \leq \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]$. Hence, it follows by (6.10) and Lemma 6.2.7 that

$$\begin{aligned}\sum_{u \in \mathcal{L}_\vartheta^{n+2\kappa(\vartheta)}} \mu_{\mathbf{P}}([u])^q &\geq \frac{1}{\lambda^q} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \sum_{w \in \vartheta(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = w]^q \\ &\geq \frac{1}{\lambda^q} \prod_{a \in \mathcal{A}} \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right)^{m(R_a - \varepsilon)} \sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q,\end{aligned}$$

noting that since $q < 0$, the function $x \mapsto x^q$ is decreasing on $(0, 1]$. Taking logarithms, rearranging and dividing by n gives

$$\begin{aligned}-\frac{1}{n} \log \left(\sum_{u \in \mathcal{L}_\vartheta^n} \mu_{\mathbf{P}}([u])^q \right) \leq &-\frac{1}{n} \log \left(\sum_{v \in \mathcal{L}_\vartheta^{m(n)}} \mu_{\mathbf{P}}([v])^q \right) + \frac{1}{n} \log \lambda^q \\ &- \frac{m(n)}{n} \sum_{a \in \mathcal{A}} (R_a - \varepsilon) \log \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right).\end{aligned}$$

Noting that $m(n)/n \to (\lambda - \varepsilon)^{-1}$ as $n \to \infty$, it follows by Lemma 6.1.11 that

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda - \varepsilon} \overline{\tau}_{\mu_{\mathbf{P}}}(q) + \frac{1}{\lambda - \varepsilon} \sum_{a \in \mathcal{A}} \log \left(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q \right) + c\varepsilon$$

where $c := (\#\mathcal{A}) \max_{a \in \mathcal{A}} \log(\sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s]^q)$. Letting $\varepsilon \to 0$ and rearranging, we obtain

$$\overline{\tau}_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda - 1} \varphi_1(q),$$

which completes the proof. □

We now give the proof of Theorem 6.2.14.

Proof of Theorem 6.2.14. In the case $q \geq 0$, the result follows from Corollary 6.2.4, noting that every recognisable random substitution satisfies the disjoint set condition. For $q < 0$, the conclusion follows by combining the upper bound on $\tau_{\mu_{\mathbf{P}}}(q)$ given by Proposition 6.2.16 with the lower bound given by Proposition 6.2.11. □

6.2.5 Recovering entropy from the L^q-spectrum

Since the L^q-spectrum encodes both topological and measure theoretic entropy, Theorem 6.2.3 provides an alternative means of proving the coincidence of these quantities with the inflation word analogues introduced in Chapters 3 and 4 for all random substitutions that satisfy the conditions of Theorem 6.2.3. In particular, we obtain the conclusion of Proposition 3.2.2 in full generality and the conclusion of Theorem 4.1.2 under the additional assumption of compatibility.

Theorem 6.2.17. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive and compatible random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$.

(1) The limit

$$\lim_{k \to \infty} \frac{1}{\lambda^k} \sum_{a \in \mathcal{A}} R_a \log(\#\vartheta^k(a))$$

exists and is equal to $h_{\mathrm{top}}(X_\vartheta)$.

(2) The L^q-spectrum of $\mu_{\mathbf{P}}$ is differentiable at 1. Moreover, the limit

$$\lim_{k \to \infty} \frac{1}{\lambda^k} \sum_{a \in \mathcal{A}} R_a \sum_{v \in \vartheta^k(a)} -\mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v] \log(\mathbb{P}[\vartheta_{\mathbf{P}}^k(a) = v])$$

exists and is equal to $\tau'_{\mu_{\mathbf{P}}}(1) = h_{\mu_{\mathbf{P}}}(X_\vartheta) = \dim_{\mathrm{H}} \mu_{\mathbf{P}}$.

Proof of Theorem 6.2.17. We first establish the result for topological entropy. By Theorem 6.2.3, the limit defining $T_{\vartheta,\mathbf{P}}(0)$ exists; in particular,

$$\lim_{m\to\infty}\frac{1}{\lambda^m}\sum_{a\in\mathcal{A}}R_a\log(\#\vartheta^m(a))$$

exists. Since $h_{\mathrm{top}}(X_\vartheta)=-\tau_{\mu_\mathbf{P}}(0)=-T_{\vartheta,\mathbf{P}}(0)$, we conclude that

$$h_{\mathrm{top}}(X_\vartheta)=-\lim_{m\to\infty}\frac{1}{\lambda^m}\sum_{a\in\mathcal{A}}R_a\log(\#\vartheta^m(a))$$

as claimed.

Now we consider measure theoretic entropy. For notational simplicity, we set

$$\rho_k=-\sum_{a\in\mathcal{A}}R_a\sum_{s\in\vartheta^k(a)}\mathbb{P}[\vartheta^k_\mathbf{P}(a)=s]\log(\mathbb{P}[\vartheta^k_\mathbf{P}(a)=s]).$$

for each $k\in\mathbb{N}$. We first make the following elementary observation: if f and g are concave functions with $f(1)=g(1)$ and $f(x)\le g(x)$ for all $x\ge 1$, then $f^+(1)\le g^+(1)$. Indeed, for all $\epsilon>0$,

$$\frac{f(1+\epsilon)-f(1)}{\epsilon}\le\frac{g(1+\epsilon)-g(1)}{\epsilon},$$

and taking the limit as $\epsilon\to 0$ (which always exists by concavity) yields the desired inequality.

Recall that $\tau_{\mu_\mathbf{P}}$ and $\lambda^{-k}\varphi_k$ are concave functions with $\tau_{\mu_\mathbf{P}}(1)=\varphi_k(1)=0$ for all $k\in\mathbb{N}$. Moreover, φ_k is differentiable for all $k\in\mathbb{N}$ with $\varphi_k'(1)=\rho_k$ and, by Proposition 6.2.2, $\left(\lambda^{-k}\varphi_k\right)_k$ converges monotonically to $\tau_{\mu_\mathbf{P}}$ from below. Thus, it follows by concavity, and the fact that $\tau_{\mu_\mathbf{P}}(1)=\lambda^{-k}\varphi_k(1)=0$ for all k, that ρ_k/λ^k is a monotonically increasing sequence bounded above by $\tau^+_{\mu_\mathbf{P}}(1)$, so the limit exists. Therefore, by these observations,

$$\tau^+_{\mu_\mathbf{P}}(1)=\lim_{k\to\infty}\frac{\rho_k}{\lambda^k},$$

since $\varphi_k(q)/(\lambda^k-1)\ge\tau_{\mu_\mathbf{P}}(q)$ for all $q\in(0,\infty)$.

The result for $\tau^-_{\mu_\mathbf{P}}(1)$ follows by an identical argument, instead using monotonicity and the corresponding bounds for $q\in(0,1)$. Thus $\tau'_{\mu_\mathbf{P}}(1)=\lim_{k\to\infty}\rho_k/\lambda^k$, so the desired result follows by Lemma 6.1.9(c). $\square$

6.3 A multifractal formalism for frequency measures

For primitive, compatible and recognisable random substitutions, Theorem 6.2.14 provides a closed-form formula for the L^q-spectrum for all $q \in \mathbb{R}$. In this section, we prove that the multifractal formalism holds for all such random substitutions. This allows the multifractal spectrum to be obtained from the L^q-spectrum.

Theorem 6.3.1. Let $\vartheta_{\mathbf{P}}$ be a primitive, compatible and recognisable random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. Then, the multifractal formalism holds for $\mu_{\mathbf{P}}$. Moreover, $f_{\mu_{\mathbf{P}}} = \tau^*_{\mu_{\mathbf{P}}}$ is an analytic and concave function.

Our strategy of proof is to establish a variational principle by considering typical local dimensions of one frequency measure $\mu_{\mathbf{P}}$ relative to another frequency measure $\mu_{\mathbf{Q}}$. The multifractal formalism then follows from this dimensional result combined with the formula for the L^q-spectrum proved in Theorem 6.2.14 – we give the proof in Section 6.3.2.

6.3.1 Non-typical local dimensions

To prove the multifractal formalism holds for a given frequency measure $\mu_{\mathbf{P}}$, we show that for every $\alpha \in [\alpha_{\min}, \alpha_{\max}]$, there exists another frequency measure $\mu_{\mathbf{Q}}$ such that $\dim_{\mathrm{H}} \mu_{\mathbf{Q}} \geq \tau^*_{\mu_{\mathbf{P}}}(\alpha)$ and $\dim_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) = \alpha$ for $\mu_{\mathbf{Q}}$-almost every $x \in X_\vartheta$. Given a primitive set-valued substitution ϑ, permissible probabilities $\mathbf{P}$ and $\mathbf{Q}$, $m \in \mathbb{N}$ and $a \in \mathcal{A}$, we define the quantity $H^{m,a}_{\mathbf{P},\mathbf{Q}}$ by

$$H^{m,a}_{\mathbf{P},\mathbf{Q}} = \sum_{v \in \vartheta^m(a)} -\mathbb{P}[\vartheta^m_{\mathbf{Q}}(a) = v] \log \mathbb{P}[\vartheta^m_{\mathbf{P}}(a) = v].$$

Further, let $\mathbf{H}^m_{\mathbf{P},\mathbf{Q}}$ denote the vector $(H^{m,a}_{\mathbf{P},\mathbf{Q}})_{a \in \mathcal{A}}$. We first prove some properties of the quantity $\mathbf{H}^m_{\mathbf{P},\mathbf{Q}}$ that we will use in the proof of Proposition 6.3.6.

Lemma 6.3.2. If ϑ is a primitive and compatible set-valued substitution and $\mathbf{P}$ and $\mathbf{Q}$ are permissible probabilities, then for all $m \in \mathbb{N}$, $a \in \mathcal{A}$ and $s \in \vartheta(a)$,

$$\sum_{v \in \vartheta^m(s)} \mathbb{P}[\vartheta^m_{\mathbf{Q}}(s) = v] \log \mathbb{P}[\vartheta^m_{\mathbf{P}}(s) = v] = \sum_{b \in \mathcal{A}} |\vartheta(a)|_b \, H^{m,b}_{\mathbf{P},\mathbf{Q}}.$$

Proof. Since ϑ is compatible, we can decompose each $v \in \vartheta^m(s)$ into inflation words $v =$

$v^1 \cdots v^{|\vartheta(a)|}$. By the Markov property of $\vartheta_{\mathbf{P}}$ (respectively $\vartheta_{\mathbf{Q}}$), we have

$$\mathbb{P}[\vartheta_{\mathbf{P}}^m(s) = v] = \mathbb{P}[\vartheta_{\mathbf{P}}^m(s_1) = v^1] \cdots \mathbb{P}[\vartheta_{\mathbf{P}}^m(s_{|\vartheta(a)|}) = v^{|\vartheta(a)|}].$$

and it follows that

$$\begin{aligned}
\sum_{v \in \vartheta^m(s)} \mathbb{P}[\vartheta_{\mathbf{Q}}^m(s) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}^m(s) = v] &= \sum_{b \in \mathcal{A}} |\vartheta(a)|_b \sum_{w \in \vartheta^m(b)} \mathbb{P}[\vartheta_{\mathbf{Q}}^m(b) = w] \log \mathbb{P}[\vartheta_{\mathbf{P}}^m(b) = w] \\
&= \sum_{b \in \mathcal{A}} |\vartheta(a)|_b \, H_{\mathbf{P},\mathbf{Q}}^{m,b},
\end{aligned}$$

which completes the proof. □

Lemma 6.3.3. If ϑ is a primitive and compatible set-valued substitution satisfying the disjoint set condition, with right Perron–Frobenius eigenvector $\mathbf{R}$, and $\mathbf{P}$ and $\mathbf{Q}$ are permissible probabilities, then

$$\frac{1}{\lambda^m} \mathbf{H}_{\mathbf{P},\mathbf{Q}}^m \cdot \mathbf{R} \to \frac{1}{\lambda - 1} \mathbf{H}_{\mathbf{P},\mathbf{Q}}^1 \cdot \mathbf{R}$$

as $m \to \infty$.

Proof. Since ϑ satisfies the disjoint set condition, for all $m \in \mathbb{N}$ and $a \in \mathcal{A}$,

$$\begin{aligned}
\mathbf{H}_{\mathbf{P},\mathbf{Q}}^{m+1} \cdot \mathbf{R} &= \sum_{a \in \mathcal{A}} R_a \sum_{v \in \vartheta^{m+1}(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}^{m+1}(a) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}^{m+1}(a) = v] \\
&= \sum_{a \in \mathcal{A}} R_a \sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = s] \log \mathbb{P}[\vartheta_{\mathbf{P}}(a) = s] \\
&\quad + \sum_{a \in \mathcal{A}} R_a \sum_{s \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = s] \sum_{v \in \vartheta^m(s)} \mathbb{P}[\vartheta_{\mathbf{Q}}^m(s) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}^m(s) = v] \\
&= \mathbf{H}_{\mathbf{P},\mathbf{Q}}^1 \cdot \mathbf{R} + \sum_{b \in \mathcal{A}} H_{\mathbf{P},\mathbf{Q}}^{m,b} \sum_{a \in \mathcal{A}} |\vartheta(a)|_b R_a \\
&= \mathbf{H}_{\mathbf{P},\mathbf{Q}}^1 \cdot \mathbf{R} + \lambda \sum_{b \in \mathcal{A}} R_b H_{\mathbf{P},\mathbf{Q}}^{m,b} \\
&= \mathbf{H}_{\mathbf{P},\mathbf{Q}}^1 \cdot \mathbf{R} + \lambda \mathbf{H}_{\mathbf{P},\mathbf{Q}}^m \cdot \mathbf{R}.
\end{aligned}$$

In the second equality we have used the Markov property of $\vartheta_{\mathbf{P}}$ and $\vartheta_{\mathbf{Q}}$, laws of logarithms, and that $\sum_{v \in \vartheta^m(s)} \mathbb{P}[\vartheta_{\mathbf{Q}}^m(s) = v] = 1$ for all $s \in \vartheta(a)$; in the third we have applied Lemma 6.3.2 and the fourth follows from the fact that λ is an eigenvalue of the substitution matrix. Applying the above inductively,

$$\frac{1}{\lambda^m}\mathbf{H}^m_{\mathbf{P},\mathbf{Q}} \cdot \mathbf{R} = \sum_{j=1}^{m} \frac{1}{\lambda^j}\mathbf{H}^1_{\mathbf{P},\mathbf{Q}} \cdot \mathbf{R} \xrightarrow{m\to\infty} \frac{1}{\lambda - 1}\mathbf{H}^1_{\mathbf{P},\mathbf{Q}} \cdot \mathbf{R},$$

which completes the proof. □

Given a recognisable random substitution $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$, any bi-infinite sequence $x \in X_\vartheta$ can be written uniquely as a bi-infinite concatenation of exact inflation words: $x = \cdots w^{-1}w^0w^1\cdots$. Moreover, there is a unique $y \in X_\vartheta$ and $j \in \{0, \ldots, |\vartheta(y_0)| - 1\}$ such that $x \in S^j(\vartheta(y))$. For each $a \in \mathcal{A}$ and $w \in \vartheta(a)$, we define the *inflation word frequency* of (a, w) in $x \in X_\vartheta$ by

$$\begin{aligned} f_x(a,w) &= \lim_{n\to\infty} f_x^n(a,w) \\ f_x^n(a,w) &= \frac{1}{2n+1}\#\{m\colon a_m = a, w^m = w, w^m \text{ is contained in } x_{[-n,n]}\}, \end{aligned}$$

provided the limit exists. For a given frequency measure $\mu_{\mathbf{P}}$, the inflation word frequency of a $\mu_{\mathbf{P}}$-typical word is determined by the production probabilities. Specifically, we have the following.

Lemma 6.3.4. Let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be a primitive, compatible and recognisable random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$. For $\mu_{\mathbf{P}}$-almost every $x \in X_\vartheta$, the inflation word frequency exists and is given by

$$f_x(a,w) = \frac{1}{\lambda} R_a \mathbb{P}[\vartheta_{\mathbf{P}}(a) = w],$$

for all $a \in \mathcal{A}$ and $w \in \vartheta(a)$.

Proof. Let $A_{a,w}$ be the set of points $x \in X_\vartheta$ such that the above does *not* hold. We show that $A_{a,w}$ is a null set. Taking the complement and then the intersection over all a, w gives a full-measure set with the required property. Given $\varepsilon > 0$, let $E(n, \varepsilon)$ be the set of $x \in X_\vartheta$ such that

$$\left|f_x^n(a,w) - \frac{1}{\lambda} R_a \mathbb{P}[\vartheta_{\mathbf{P}}(a) = w]\right| > \varepsilon.$$

By the Borel–Cantelli lemma, it suffices to show that

$$\sum_{n\in\mathbb{N}} \mu_{\mathbf{P}}(E(n,\varepsilon)) < \infty$$

for all $\varepsilon > 0$ in order to conclude that $A_{a,w}$ is a nullset. To this end, we show that $\mu_{\mathbf{P}}(E(n,\varepsilon))$ decays exponentially with n. Given u with $|u| = 2n+1 > 2\kappa(\vartheta)$, let u^R denote the recognisable core of u, which has length at least $|u| - 2\kappa(\vartheta)$. Lemma 6.2.15 gives that

$$\mu_{\mathbf{P}}([u]) \leq \frac{\kappa(\vartheta)}{\lambda} \mu_{\mathbf{P}}([v]) \mathbb{P}[\vartheta_{\mathbf{P}}(v) = u^R] = \frac{\kappa(\vartheta)}{\lambda} \mu_{\mathbf{P}}([v]) \prod_{i=1}^{|v|} \mathbb{P}[\vartheta_{\mathbf{P}}(v_i) = w^{i,v_i}]$$

where each w^{i,v_i} is the inflated image of v_i in u^R. By compatibility, we can choose an integer N such that every v of length at least N satisfies $|v|(R_a - \varepsilon/3) \leq |v|_a \leq |v|(R_a + \varepsilon/3)$ for all $a \in \mathcal{A}$. For each v and $a \in \mathcal{A}$, let $A_a(v)$ denote the set of $u' \in \vartheta(v)$ such that the frequency of indices $i \in \{j : a_j = a\}$ with $w^{i,a} = w$ deviates from $\mathbb{P}[\vartheta_{\mathbf{P}}(a) = w]$ by more than $\varepsilon/3$. Since $\vartheta_{\mathbf{P}}$ acts independently on letters, it follows by Cramér's theorem (Theorem A.2.1) that the sum $\sum_{u' \in A(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = u']$ decays exponentially with $|v|_a$ (and hence with $|v|$). In particular, there is a constant $C > 0$, independent of the choice of v, such that

$$\sum_{u' \in A(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = u'] \leq e^{-Cn}. \tag{6.11}$$

Note that if u is a sufficiently long legal word and has $[u] \cap E(n,\varepsilon) = \varnothing$, then we require that $u^R \in A(v)$. Indeed, if $u' \notin A(v)$ and $|v| \geq N$, then the relative inflation word frequency of w is bounded above by

$$\begin{aligned} \frac{\{j : a_j = a\}}{|v|} \frac{|v|}{|u|} \left(\mathbb{P}[\vartheta_{\mathbf{P}}(a) = w] + \frac{\varepsilon}{3} \right) &\leq \frac{1}{\lambda} \left(R_a + \frac{\varepsilon}{3} \right) \left(\mathbb{P}[\vartheta_{\mathbf{P}}(a) = w] + \frac{\varepsilon}{3} \right) \\ &\leq \frac{1}{\lambda} R_a \mathbb{P}[\vartheta_{\mathbf{P}}(a) = w] + \varepsilon \end{aligned}$$

and, similarly, bounded below by $R_a \mathbb{P}[\vartheta_{\mathbf{P}}(a) = w]/\lambda - \varepsilon$; hence, $[u^R] \cap E(n,\varepsilon) = \varnothing$. Let $\mathcal{V}_n$ denote set of all words that appear as the (unique) preimage of the recognisable core of a word of length n. It then follows by Lemma 6.2.15 that

$$\mu_{\mathbf{P}}(E(n,\varepsilon)) \leqslant \sum_{\substack{u \in \mathcal{L}_\vartheta^n \\ [u] \cap E(n,\varepsilon) \neq \varnothing}} \mu_{\mathbf{P}}([u]) \leq \frac{\kappa(\vartheta)}{\lambda} \sum_{v \in \mathcal{V}_n} \mu_{\mathbf{P}}([v]) \sum_{u' \in A(v)} \mathbb{P}[\vartheta_{\mathbf{P}}(v) = u'] \leqslant e^{-Cn},$$

where in the final inequality we have used (6.11) and that

$$\sum_{v \in \mathcal{V}_n} \mu_{\mathbf{P}}([v]) \leq \sum_{j=1}^{n} \sum_{v \in \mathcal{L}_\vartheta^j} \mu_{\mathbf{P}}([v]) \leq n,$$

absorbing this contribution and the $\kappa(\vartheta)/\lambda$ factor into the constant C. It follows that

$$\sum_{n=1}^{\infty} \mu_{\mathbf{P}}(E(n,\varepsilon)) \leq \sum_{n=1}^{\infty} e^{-Cn} < \infty,$$

and the result then follows by the Borel–Cantelli lemma. □

Finally, we require the following bounds on the exponential scaling rate of measures of cylinders, which is a consequence of Theorem 6.2.3 and standard properties of the L^q-spectrum and multifractal spectrum. In particular, these give bounds on the possible local dimensions of the measure.

Proposition 6.3.5. If $\vartheta_{\mathbf{P}}$ is a primitive and compatible random substitution with corresponding frequency measure $\mu_{\mathbf{P}}$, then there are values $0 < s_1 < s_2 < \infty$ and $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$ and $v \in \mathcal{L}^n(X_\vartheta) = \mathcal{L}_\vartheta^n$, we have

$$s_1 n + c_1 \leq \log \mu_{\mathbf{P}}([v]) \leq s_2 n + c_2$$

Proof. By Theorem 6.2.3, for all $k \in \mathbb{N}$ and $q > 1$,

$$\tau_{\mu_{\mathbf{P}}}(q) \leq \frac{1}{\lambda^k - 1} \varphi_k(q);$$

and for $q < 0$,

$$\frac{1}{\lambda^k - 1} \varphi_k(q) \leq \tau_{\mu_{\mathbf{P}}}(q),$$

Moreover, for each k, with

$$\beta_{k,\min} := \lim_{q\to\infty} \frac{\varphi_k(q)}{q(\lambda^k - 1)} = -\frac{1}{\lambda^k - 1} \sum_{a\in\mathcal{A}} R_a \log\left(\min_{v\in\vartheta^k(a)} \mathbb{P}[\vartheta^k_{\mathbf{P}}(a) = v] \right)$$

$$\beta_{k,\max} := \lim_{q\to-\infty} \frac{\varphi_k(q)}{q(\lambda^k - 1)} = -\frac{1}{\lambda^k - 1} \sum_{a\in\mathcal{A}} R_a \log\left(\max_{v\in\vartheta^k(a)} \mathbb{P}[\vartheta^k_{\mathbf{P}}(a) = v] \right),$$

it follows that $[\beta_{k,\min}, \beta_{k,\max}] \subset (0, \infty)$ is a decreasing nested sequence of intervals, so with $\beta_{\min} = \lim_{k\to\infty} \beta_{k,\min}$ and $\beta_{\max} = \lim_{k\to\infty} \beta_{k,\max}$,

$$0 < \beta_{\min} \le \lim_{q\to\infty} \tau_{\mu_{\mathbf{P}}}(q) \le \lim_{q\to-\infty} \tau_{\mu_{\mathbf{P}}}(q) \le \beta_{\max} < \infty.$$

Applying Lemma 6.1.9(b) gives the result. □

Finally, we obtain our main conclusion concerning relative local dimensions.

Proposition 6.3.6. Let ϑ be a primitive, compatible and recognisable set-valued substitution, let $\mathbf{P}$ and $\mathbf{Q}$ be permissible probabilities, and let $\mu_{\mathbf{P}}$ and $\mu_{\mathbf{Q}}$ denote the respective frequency measures. Then, for $\mu_{\mathbf{Q}}$-almost all $x \in X_\vartheta$,

$$\dim_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) = \frac{1}{\lambda - 1} \sum_{a\in\mathcal{A}} R_a \sum_{v\in\vartheta(a)} -\mathbb{P}[\vartheta^m_{\mathbf{Q}}(a) = v] \log \mathbb{P}[\vartheta^m_{\mathbf{P}}(a) = v]. \tag{6.12}$$

Proof. Fix $m \in \mathbb{N}$. It follows by Lemma 2.2.19 that since $\vartheta_{\mathbf{P}}$ is recognisable, so is $\vartheta^m_{\mathbf{P}}$. For each $x \in X_\vartheta$ and $n \in \mathbb{N}$ with $n > \kappa(\vartheta^m)$, let $u^n_-(x)$ denote the recognisable core of $x_{[-n,n]}$ and let $u^n_+(x)$ denote an inflation word of minimal length that contains $x_{[-n,n]}$. By compatibility, $|u^n_-(x)|/(2n+1) \to \lambda^{-m}$ and $|u^n_+(x)|/(2n+1) \to \lambda^{-m}$ as $n \to \infty$. Further, let $v^n_-(x)$ be the legal word such that $u^n_-(x) \in \vartheta^m(v^n_-(x))$ and $v^n_+(x)$ be the legal word such that $u^n_+(x) \in \vartheta^m(v^n_+(x))$. Then, it follows by Lemma 6.2.15 and the definition of local dimension that

$$\begin{aligned}
&\liminf_{n\to\infty} \left(-\frac{1}{2n+1} \log \mu_{\mathbf{P}}([u^n_-(x)]) - \frac{1}{2n+1} \log \mathbb{P}[\vartheta_{\mathbf{P}}(v^n_-(x)) = u^n_-(x)] \right) \\
&\quad \le \underline{\dim}_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) \le \overline{\dim}_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) \\
&\quad \le \limsup_{n\to\infty} \left(-\frac{1}{2n+1} \log \mu_{\mathbf{P}}([u^n_+(x)]) - \frac{1}{2n+1} \log \mathbb{P}[\vartheta_{\mathbf{P}}(v^n_+(x)) = u^n_+(x)] \right).
\end{aligned}$$

By Proposition 6.3.5, there exists a constant $C \geq 0$ such that for all $x \in X_\vartheta$,

$$0 \leq \liminf_{n\to\infty} -\frac{1}{2n+1} \log \mu_{\mathbf{P}}([u_-^n(x)]) \leq \limsup_{n\to\infty} -\frac{1}{2n+1} \log \mu_{\mathbf{P}}([u_+^n(x)]) \leq C.$$

Hence, it follows from the above that

$$\begin{aligned}
&\liminf_{n\to\infty} -\frac{1}{2n+1} \log \mathbb{P}[\vartheta_{\mathbf{P}}(v_-^n(x)) = u_-^n(x)] \\
&\qquad \leq \underline{\dim}_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) \leq \overline{\dim}_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) \\
&\qquad \leq \limsup_{n\to\infty} -\frac{1}{2n+1} \log \mathbb{P}[\vartheta_{\mathbf{P}}(v_+^n(x)) = u_+^n(x)] + \frac{C}{\lambda^m}.
\end{aligned} \tag{6.13}$$

We now show that for $\mu_{\mathbf{Q}}$-almost all $x \in X_\vartheta$,

$$\begin{aligned}
\liminf_{n\to\infty} -\frac{1}{n} \log \mathbb{P}[\vartheta_{\mathbf{P}}(v_-^n(x)) = u_-^n(x)] &= \limsup_{n\to\infty} -\frac{1}{n} \log \mathbb{P}[\vartheta_{\mathbf{P}}(v_+^n(x)) = u_+^n(x)] \\
&= \frac{1}{\lambda^m} \mathbf{H}_{\mathbf{P},\mathbf{Q}}^m \cdot \mathbf{R}.
\end{aligned}$$

By compatibility, we can decompose the production probabilities into inflation tiles as

$$\mathbb{P}[\vartheta_{\mathbf{P}}^m(v_-^n(x)) = u_-^n(x)] = \prod_{a\in\mathcal{A}} \prod_{w\in\vartheta^m(a)} \mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = w]^{N_{a,w}(x,n)},$$

where, for each $a \in \mathcal{A}$ and $w \in \vartheta^m(a)$, $N_{a,w}(x,n)$ denotes the number of a's in $v_-^n(x)$ that map to w. It follows by Lemma 6.3.4, applied to $\vartheta_{\mathbf{Q}}^m$, that for $\mu_{\mathbf{Q}}$-almost all $x \in X_\vartheta$,

$$\frac{1}{2n+1} N_{a,w}(x,n) \to \frac{1}{\lambda^m} R_a \mathbb{P}[\vartheta_{\mathbf{Q}}^m(a) = w]$$

as $n \to \infty$ for all $a \in \mathcal{A}$ and $w \in \vartheta^m(a)$. Hence, it follows that

$$\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{2n+1} \log \mathbb{P}[\vartheta_{\mathbf{P}}^m(v_-^n(x)) = u_-^n(x)] \\
&\qquad = \frac{1}{\lambda^m} \sum_{a\in\mathcal{A}} R_a \sum_{v\in\vartheta^m(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}^m(a) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}^m(a) = v] \\
&\qquad = \frac{1}{\lambda^m} \mathbf{H}_{\mathbf{P},\mathbf{Q}}^m \cdot \mathbf{R},
\end{aligned}$$

with the same convergence holding for $u_+^n(x)$ by identical arguments. Thus, it follows from (6.13) that

$$\frac{1}{\lambda^m}\mathbf{H}^m_{\mathbf{P},\mathbf{Q}}\cdot\mathbf{R}\le\underline{\dim}_{\mathrm{loc}}(\mu_{\mathbf{P}},x)\le\overline{\dim}_{\mathrm{loc}}(\mu_{\mathbf{P}},x)\le\frac{1}{\lambda^m}\mathbf{H}^m_{\mathbf{P},\mathbf{Q}}\cdot\mathbf{R}+\frac{C}{\lambda^m}.$$

Since the above holds for all $m\in\mathbb{N}$, by letting $m\to\infty$ it follows by Lemma 6.3.3 that $\dim_{\mathrm{loc}}(\mu_{\mathbf{P}},x)$ exists and

$$\dim_{\mathrm{loc}}(\mu_{\mathbf{P}},x)=\frac{1}{\lambda-1}\mathbf{H}^1_{\mathbf{P},\mathbf{Q}}\cdot\mathbf{R},$$

which completes the proof. □

6.3.2 Proof of the multifractal formalism

In this section, we apply the results obtained in the previous subsection, along with results on the L^q-spectrum under recognisability, to prove Theorem 6.3.1.

Proof of Theorem 6.3.1. In light of Proposition 6.1.13, it remains to show that $f_{\mu_{\mathbf{P}}}(\alpha)\ge\tau^*_{\mu_{\mathbf{P}}}(\alpha)$ for each $\alpha\in\mathbb{R}$. By Theorem 6.2.14, for all $q\in\mathbb{R}$, we have

$$\tau_{\mu_{\mathbf{P}}}(q)=\frac{1}{\lambda-1}\varphi_1(q)=\frac{1}{\lambda-1}\sum_{a\in\mathcal{A}}R_aT_a(q)$$

where, for each $a\in\mathcal{A}$,

$$T_a(q)=-\log\left(\sum_{s\in\vartheta(a)}\mathbb{P}[\vartheta_{\mathbf{P}}(a)=s]^q\right).$$

First, fix $\alpha\in(\alpha_{\min},\alpha_{\max})$ and let $q\in\mathbb{R}$ be chosen so that $\tau'_{\mu_{\mathbf{P}}}(q)=\alpha$. Observe that $q\alpha-\tau_{\mu_{\mathbf{P}}}(q)=\tau^*_{\mu_{\mathbf{P}}}(\alpha)$. Then define $\mathbf{Q}$ by the rule

$$\mathbb{P}[\vartheta_{\mathbf{Q}}(a)=s]=\mathbb{P}[\vartheta_{\mathbf{P}}(a)=s]^qe^{T_a(q)}$$

for all $a \in \mathcal{A}$ and $s \in \vartheta(a)$. Then by Theorem 6.2.17,

$$\begin{aligned}
\dim_{\mathrm{H}} \mu_{\mathbf{Q}} &= \frac{1}{\lambda - 1} \sum_{a \in \mathcal{A}} R_a \Big(- \sum_{v \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = v] \log \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = v] \Big) \\
&= q \cdot \frac{1}{\lambda - 1} \sum_{a \in \mathcal{A}} R_a \Big(- \sum_{v \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}(a) = v] \Big) \\
&\quad - \frac{1}{\lambda - 1} \sum_{a \in \mathcal{A}} R_a T_a(q) \sum_{v \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = v] \\
&= q\alpha - \tau_{\mu_{\mathbf{P}}}(q) = \tau_{\mu_{\mathbf{P}}}^*(\alpha)
\end{aligned}$$

since

$$\begin{aligned}
\tau'_{\mu_{\mathbf{P}}}(q) &= \frac{1}{\lambda - 1} \sum_{a \in \mathcal{A}} R_a \frac{- \sum_{v \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{P}}(a) = v]^q \log \mathbb{P}[\vartheta_{\mathbf{P}}(a) = v]}{e^{-T_a(q)}} \\
&= \frac{1}{\lambda - 1} \sum_{a \in \mathcal{A}} R_a \Big(- \sum_{v \in \vartheta(a)} \mathbb{P}[\vartheta_{\mathbf{Q}}(a) = v] \log \mathbb{P}[\vartheta_{\mathbf{P}}(a) = v] \Big).
\end{aligned}$$

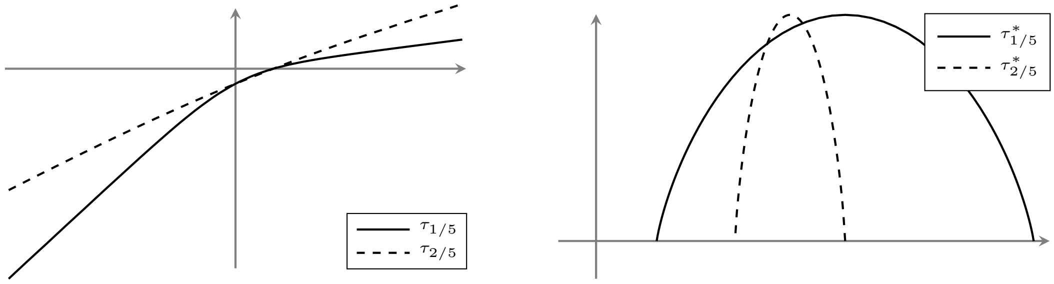
In fact, this shows that $\dim_{\mathrm{loc}}(\mu_{\mathbf{P}}, x) = \alpha$ for $\mu_{\mathbf{Q}}$-almost all $x \in X_\vartheta$ by Proposition 6.3.6. Thus $f_{\mu_{\mathbf{P}}}(\alpha) \geq \dim_{\mathrm{H}} \mu_{\mathbf{Q}} = \tau_{\mu_{\mathbf{P}}}^*(\alpha)$, as required.

The result for $\alpha = \alpha_{\min}$ (resp. $\alpha = \alpha_{\max}$) follows similarly by taking a degenerate probability vector $\mathbf{Q}$ assigning equal value to the realisations of $\vartheta(a)$ with maximal (resp. minimal) probabilities given by $\mathbf{P}$, and zero otherwise. The corresponding non-degenerate sub-substitution is also compatible and recognisable, so the same arguments yield the corresponding bounds. □

6.3.3 Examples

Example 6.3.7. Let $p > 0$ and let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} : \begin{cases} a \mapsto \begin{cases} abb & \text{with probability } p, \\ bab & \text{with probability } 1 - p, \end{cases} \\ b \mapsto aa \quad \text{with probability } 1, \end{cases}$$



(a) L^q-spectrum (b) Multifractal spectrum

Figure 6.2: L^q-spectrum and multifractal spectrum corresponding to the frequency measure in Example 6.3.7 for $p \in \{1/5, 2/5\}$.

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. The random substitution $\vartheta_{\mathbf{P}}$ is compatible, with corresponding primitive substitution matrix

$$M = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix},$$

Perron–Frobenius eigenvalue $\lambda = (1+\sqrt{17})/2$, and (normalised) right Perron–Frobenius eigenvector

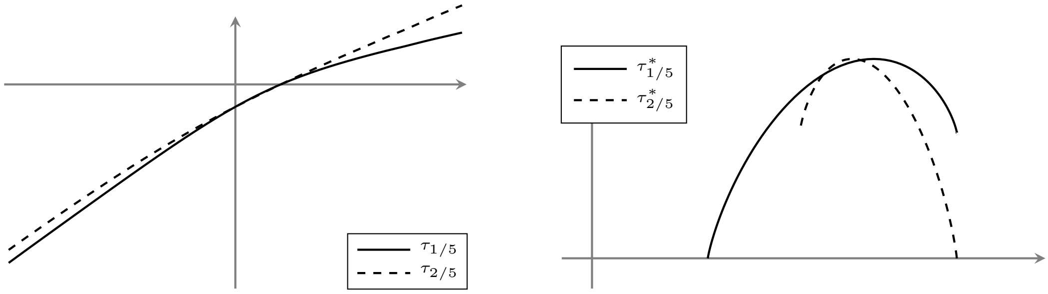
$$\left(\frac{-3+\sqrt{17}}{2}, \frac{5-\sqrt{17}}{2} \right).$$

We showed in Example 2.2.15 that the random substitution $\vartheta_{\mathbf{P}}$ is recognisable. Hence, Theorem 6.2.14 gives that for all $q \in \mathbb{R}$, we have

$$\tau_{\mu_p}(q) = T_{\vartheta,\mathbf{P}}(q) = \frac{1}{\lambda - 1}\varphi_1(q) = -\frac{7-\sqrt{17}}{8}\log(p^q + (1-p)^q).$$

Further, it follows by Theorem 6.3.1 that $\mu_{\mathbf{P}}$ satisfies the multifractal formalism. A plot of the L^q-spectrum and multifractal spectrum for two choices of p is given in Figure 6.2. We note that when $p = 1/2$, the L^q-spectrum of the measure $\mu_{\mathbf{P}}$ is a straight line and the multifractal spectrum is equal to $h_{\mathrm{top}}(X_\vartheta)$ at $h_{\mathrm{top}}(X_\vartheta)$, and $-\infty$ otherwise.

In the following example, we highlight that the multifractal spectrum need not have value 0 at the endpoints.



(a) L^q-spectrum (b) Multifractal spectrum

Figure 6.3: L^q-spectrum and multifractal spectrum corresponding to the frequency measure in Example 6.3.8 for $p \in \{1/5, 2/5\}$.

Example 6.3.8. Let $\vartheta_{\mathbf{P}}$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} abb & \text{with probability } p, \\ bab & \text{with probability } p, \\ bba & \text{with probability } 1-2p, \end{cases} \\ b \mapsto aaa \quad \text{with probability } 1. \end{cases}$$

Similarly to Example 6.3.7, $\vartheta_{\mathbf{P}}$ is primitive, compatible and recognisable, so Theorem 6.3.1 gives that the multifractal formalism holds. By Theorem 6.2.14, we have

$$\tau_{\mu_{\mathbf{P}}}(q) = -\frac{3}{10} \log(2p^q + (1-2p)^q),$$

with the multifractal spectrum given by the concave conjugate. For $p = 1/5$ and $p = 2/5$, the L^q-spectrum and multifractal spectrum of $\mu_{\mathbf{P}}$ are plotted in Figure 6.3.

Example 6.3.9. Given $p \in (0,1)$, let $\vartheta_{\mathbf{P}} = (\vartheta, \mathbf{P})$ be the random substitution defined by

$$\vartheta_{\mathbf{P}} \colon \begin{cases} a \mapsto \begin{cases} abbaa & \text{with probability } p, \\ aabba & \text{with probability } 1-p, \end{cases} \\ b \mapsto \begin{cases} babaa & \text{with probability } p, \\ baaba & \text{with probability } 1-p, \end{cases} \end{cases}$$

and let $\mu_{\mathbf{P}}$ denote the corresponding frequency measure. We have previously shown that $\vartheta_{\mathbf{P}}$

is primitive, compatible and recognisable, so by Theorem 6.3.1, $\mu_{\mathbf{P}}$ satisfies the multifractal formalism. In Example 5.2.4, we showed that the subshift X_ϑ is intrinsically ergodic, with the frequency measure corresponding to $p = 1/2$ being the unique measure of maximal entropy. Further, we showed that this measure does *not* satisfy the Gibbs property (5.2).

6.4 Outlook

We conclude this chapter with a list of open questions which we feel could be potentially interesting directions for future work.

(1) What is the L^q-spectrum of the frequency measure corresponding to a primitive and compatible random substitution when $q < 0$? Example 6.2.12 demonstrates that, without recognisability, the L^q-spectrum and inflation word L^q-spectrum need not coincide for $q < 0$. However, even for this example, we do not know an exact formula for the L^q-spectrum when $q < 0$. Obtaining precise results for $q < 0$ is substantially more challenging than for $q \geq 0$, since the sum in the definition of the L^q-spectrum depends on the measure of cylinders with very small (but non-zero) measure. For example, in the self-similar case, without the presence of strong separation assumptions, little is known. This is in stark contrast to the $q \geq 0$ case, which is generally well understood.

(2) Without the disjoint set condition and the identical set condition, what can be said about differentiability of the L^q-spectrum? For $q \geq 0$, we give the L^q-spectrum as a uniform limit of analytic functions: however, aside from the exceptional point $q = 1$ where we can say more, this is not enough to give information about differentiability.

(3) Does the frequency measure corresponding to the random period doubling substitution satisfy the multifractal formalism? More generally, can the assumption of recognisability in Theorem 6.3.1 be relaxed to a weaker condition, such as the disjoint set condition?

APPENDIX A

MATHEMATICAL BACKGROUND

A.1 Ergodic theory and dynamical systems

Much of our work concerns properties of topological and measure theoretic dynamical systems. Here, we summarise some of the key concepts from ergodic theory that we work with throughout, and give the definitions and key properties of measure theoretic and topological entropy in the more general setting of topological dynamical systems. For a detailed introduction to ergodic theory, we refer the reader to Walters' book [75].

Throughout this section, we assume that a given topological space X is equipped with the Borel sigma-algebra $\mathcal{B}(X)$.

A.1.1 Invariant and ergodic measures

Definition A.1.1. Let $T\colon X \to X$ be a transformation of a topological space X. We say that a probability measure $\mu\colon \mathcal{B}(X) \to \mathbb{R}$ is T-*invariant* if $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}(X)$. We call the triple (X, μ, T) a *measure-preserving transformation*. Further, we let $\mathcal{M}(X, T)$ denote the set of all T-invariant Borel probability measures on X.

Definition A.1.2. We say that a sequence of measures $(\mu_n)_n \subseteq \mathcal{M}(X, T)$ converges weak* to $\mu \in \mathcal{M}(X, T)$ if for every continuous $f\colon X \to \mathbb{R}$, we have

$$\int f\, d\mu_n \to \int f\, d\mu$$

as $n \to \infty$.

Under the assumption that the topological space X is compact and metrisable, and the transformation $T\colon X \to X$ is continuous, the set of all T-invariant probability measures on X is weak*-compact.

Proposition A.1.3 ([75, Theorem 6.10])**.** Let $T\colon X \to X$ be a continuous mapping of a compact metric space. Then, the space $\mathcal{M}(X,T)$ is weak*-compact.

Definition A.1.4. Let $T\colon X \to X$ be a continuous transformation of a compact metric space and let $\mu \in \mathcal{M}(X,T)$. We say that μ is an *ergodic* measure for T if for every $B \in \mathcal{B}(X)$ with $T^{-1}B = B$, we have $\mu(B) = 0$ or $\mu(B) = 1$.

A notable consequence of ergodicity is the following.

Theorem A.1.5 (Birkhoff's ergodic theorem)**.** Let (X,μ,T) be an ergodic measure-preserving transformation and $f\colon X \to X$ be an L^1-function. Then,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) = \int f \,\mathrm{d}\,\mu$$

for μ-almost every $x \in X$.

A.1.2 Measure theoretic entropy

Here, we give the general definition of measure theoretic entropy for an invertible measure-preserving transformation. We first define the entropy of a measurable partition.

Definition A.1.6. Let $(X,\mathcal{B},\mu)$ be a probability space and let η be a measurable partition of X. The entropy $H_\mu(\eta)$ of η with respect to μ is the quantity defined by

$$H_\mu(\eta) = \sum_{A\in\eta} -\mu(A)\log\mu(A).$$

Definition A.1.7. Let (X,μ,T) be a measure-preserving transformation. Further, let ξ be a finite measurable partition of X with $\vee_{i\in\mathbb{N}} T^{-i}(\xi) = \mathcal{B}(X)$, up to null sets. For each $n \in \mathbb{N}$, let $\xi_n = \vee_{i=0}^{n-1} T^{-i}(\xi)$. The *measure theoretic entropy* of the system (X,μ,T) is the quantity defined by

$$h_\mu(X,T) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_n).$$

That the limit exists follows from the fact that the sequence $(H_\mu(\xi_n))_n$ is sub-additive and Fekete's lemma [27].

If the space X is compact and metrisible and the transformation $T\colon X \to X$ is continuous, then the entropy map $\mu \to h_\mu(X,T)$ is upper semi-continuous.

Proposition A.1.8 ([75, Theorem 8.2])**.** Let $T\colon X \to X$ be a continuous transformation of a compact topological space X and let $(\mu_n)_n$ be a sequence of T-invariant measures which converge weak* to a T-invariant measure μ. Then,

$$h_\mu(X,T) \geq \limsup_{n\to\infty} h_{\mu_n}(X,T).$$

If μ is additionally assumed to be ergodic, then the measure theoretic entropy $h_\mu(X,T)$ can be obtained from the individual points almost surely. This result is commonly known as the Shannon–McMillan–Breiman theorem. We give the statement here for the special case of subshifts. For a more general result, we refer the reader to Keller's book [43].

Theorem A.1.9 (Shannon–McMillan–Breiman)**.** Let X be a subshift over a finite alphabet, equipped with an ergodic probability measure μ. Then,

$$-\frac{1}{n}\log\mu([x_{[1,n]}]) \to h_\mu(X)$$

as $n \to \infty$, both almost-surely and in L^1.

If μ is an S-invariant probability measure on a shift space (X,S), then the entropy of the systems (X,S^m), $m \in \mathbb{N}$, are related to the entropy of the system (X,S) by *Abramov's formula* [1]. We note that Abramov's result applies more generally to flows on Lebesgue spaces. However, the following formulation is sufficient for our purposes.

Lemma A.1.10. Let X be a subshift equipped with an S-invariant probability measure μ. For all $m \in \mathbb{N}$, the following identity holds:

$$h_\mu(X,S^m) = m h_\mu(X,S).$$

A.1.3 Topological entropy

Topological entropy can similarly be defined more generally for topological dynamical systems (X, T), where X is a compact space and $T\colon X \to X$ is a continuous transformation of X.

Definition A.1.11. Let X be a compact topological space and let α be an open cover of X. We let $N(\alpha)$ denote the smallest cardinality of a sub-cover of α. By compactness, $N(\alpha) < \infty$. We define the entropy of α to be

$$H(\alpha) = \log N(\alpha).$$

Definition A.1.12. Let X be a compact topological space, let α, β be open covers of X and let $T\colon X \to X$ a continuous transformation of X. We let $\alpha \vee \beta$ denote the open cover of X by sets of the form $A \cap B$, where $A \in \alpha$, $B \in \beta$ and let $T^{-1}\alpha$ denote the open cover by sets of the form $T^{-1}A$, $A \in \alpha$.

Definition A.1.13. Let $T\colon X \to X$ be a continuous transformation of a compact topological space X and let α be an open cover of X. The *topological entropy of* (X, T) *relative to* α is defined by

$$h_{\text{top}}(X, T, \alpha) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{j=0}^{n-1} T^{-j}\alpha\right). \tag{A.1}$$

The *topological entropy of* (X, T) is then defined by

$$h_{\text{top}}(X, T) = \sup\{h_{\text{top}}(X, T, \alpha)\colon \alpha \text{ is an open cover of } X\}. \tag{A.2}$$

The limit in (A.1) always exists due to the sub-additivity of the sequence $(\vee_{j=0}^{n-1} T^{-j}\alpha)_n$ and Fekete's lemma [27]. See [75] for the precise details.

A.2 Large deviations theory

Some of our proofs utilise results from large deviations theory. Here, we provide a brief overview of the theory of large deviations and give the statement of Cramér's theorem, which we apply in the proofs of Theorems 4.1.1 and 6.3.1. For a more detailed introduction to large deviations, we refer the reader to the excellent book by Den Hollander [19].

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables on a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$, with mean $\zeta \in \mathbb{R}$ and standard deviation σ. For each $n \in \mathbb{N}$, let $S_n = X_1 + \cdots + X_n$ denote the n^{th} partial sum. By the strong law of large numbers,

$$\frac{1}{n} S_n \to \zeta$$

$\mathbb{P}$-almost surely. However, the strong law of large numbers does not provide any quantitative information about the rate of convergence. Observe that the central limit theorem gives that

$$\frac{1}{\sigma\sqrt{n}}(S_n - \zeta n) \to Z$$

in distribution, with respect to $\mathbb{P}$, where Z is the standard normal distribution. In particular, the central limit theorem quantifies the probability that S_n differs from ζn by an amount of order $\sqrt{n}$. Such deviations are referred to as *normal deviations*. The theory of *large deviations* is concerned with quantifying the size of the set of points for which S_n differs from ζn by an amount of order n. One of the most famous results in this direction is Cramér's theorem, which states the following.

Theorem A.2.1. Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of independent, identically distributed random variables and assume that

$$\varphi(t) = \mathbb{E}[\mathrm{e}^{tX_1}] < \infty$$

for all $t \in \mathbb{R}$. For each $n \in \mathbb{N}$, let S_n denote the n^{th} partial sum $S_n = \sum_{i=1}^{n} X_i$. Then, for all $\alpha > \mathbb{E}[X_1]$,

$$\frac{1}{n} \log \mathbb{P}[S_n \geq \alpha n] \to -I(\alpha)$$

as $n \to \infty$, where

$$I(x) = \sup_{t \in \mathbb{R}} \{xt - \log \phi(t)\}.$$

In particular, there exists a constant $C > 0$ such that

$$\mathbb{P}[S_n \geq \alpha n] \leq \mathrm{e}^{-Cn}$$

for all $n \in \mathbb{N}$.

BIBLIOGRAPHY

[1] L. M. Abramov. On the entropy of a flow. *Dokl. Akad. Nauk SSSR*, 128:873–875, 1959.

[2] R. L. Adler and B. Weiss. Similarity of automorphisms of the torus. *Mem. Amer. Math. Soc.*, 98:ii+43, 1970.

[3] M. Arbeiter and N. Patzschke. Random Self-Similar Multifractals. *Math. Nachr.*, 181:5–42, 1996.

[4] T. Austin. A new proof of the Shmerkin–Wu theorem. *J. Modern Dynam.*, 18:1–11, 2022.

[5] M. Baake and U. Grimm. *Aperiodic Order. Vol. 1.* Number 149 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge ; New York, 2013.

[6] M. Baake and U. Grimm, editors. *Aperiodic Order. Vol. 2.* Number 166 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge ; New York, 2017.

[7] M. Baake, T. Spindeler, and N. Strungaru. Diffraction of compatible random substitutions in one dimension. *Indag. Math. (N.S.)*, 29:1031–1071, 2018.

[8] J. Bamberg, G. Cairns, and D. Kilminster. The crystallogrphic restriction, permutations, and goldbach's conjecture. *Amer. Math. Monthly*, 110:202–209, 2003.

[9] V. Berthé and M. Rigo, editors. *Combinatorics, Automata and Number Theory.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010.

[10] R. Bowen. Some systems with unique equilibrium states. *Math. Systems Theory*, 8:193–202, 1974.

[11] J. Cassaigne. Constructing infinite words of intermediate complexity. In *Developments in Language Theory*, pages 173–184. Springer Berlin Heidelberg, 2003.

[12] J. Cassaigne and F. Nicholas. On the Morse–Hedlund complexity gap. Available at arXiv:0903.1627.

[13] V. Climenhaga and D. Thompson. Intrinsic ergodicity beyond specification: β-shifts, S-gap shifts, and their factors. *Israel J. Math.*, 192:785–817, 2012.

[14] V. Climenhaga and D. Thompson. Intrinsic ergodicity via obstruction entropies. *Ergod. Dyn. Sys.*, 34:1816–1831, 2014.

[15] V. Climenhaga and D. Thompson. Beyond Bowen's Specification Property. In M. Pollicott and S. Vaienti, editors, *Thermodynamic Formalism*, volume 2290, pages 3–82. Springer International Publishing, Cham, 2021.

[16] F. M. Dekking and G. R. Grimmett. Superbranching processes and projections of random Cantor sets. *Probab. Th. Relat. Fields*, 78:335–355, 1988.

[17] F. M. Dekking and R. W. J. Meester. On the structure of Mandelbrot's percolation process and other random Cantor sets. *J. Stat. Phys.*, 58:1109–1126, 1990.

[18] F. M. Dekking and P. vd Wal. Fractal percolation and branching cellular automata. *Probab. Th. Relat. Fields*, 120:277–308, 2001.

[19] F. Den Hollander. *Large deviations.* Fields Institute Monographs. American Mathematical Society, 2000.

[20] A. Ehrenfeucht and G. Rozenberg. On subword complexities of homomorphic images. *Informatique Théoretique et Applications*, 16(4):303–316, 1982.

[21] M. Einsiedler, E. Lindenstrauss, and T. Ward. Entropy in ergodic theory and homogeneous dynamics. Available at tbward0. wixsite.com/books/entropy.

[22] G. Escolano, N. Manibo, and E. Miro. Mixing properties and entropy bounds of a family of pisot random substitutions. *Indag. Math.*, 33:965–991, 2022.

[23] K. J. Falconer. *Techniques in Fractal Geometry.* Wiley, Chichester ; New York, 1997.

[24] K. J. Falconer. *Fractal Geometry – Mathematical Foundations and Applications.* Wiley, 3rd edition, 2014.

[25] A.-H. Fan, D.-J. Feng, and J. Wu. Recurrence, dimension, and entropy. *J. Lond. Math. Soc.*, 64:229–244, 2001.

[26] A.-H. Fan, K.-S. Lau, and H. Rao. Relationships between Different Dimensions of a Measure. *Monatsh. Math.*, 135:191–201, 2002.

[27] M. Fekete. Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten. *Math. Zeitschrift*, 17(1):228–249, 1923.

[28] D.-J. Feng. The limited Rademacher functions and Bernoulli convolutions associated with Pisot numbers. *Adv. Math.*, 195:24–101, 2005.

[29] D.-J. Feng and K.-S. Lau. Multifractal formalism for self-similar measures with weak separation condition. *J. Math. Pures Appl.*, 92:407–428, 2009.

[30] S. Ferenczi. Complexity of sequences and dynamical systems. *Discrete Math.*, 206:145–154, 1999.

[31] S. Finch. *Mathematical Constants.* Cambridge University Press, 2003.

[32] I. Fisher, K. Cheon, A. Panchula, P. Canfield, M. Chernikov, H. Ott, and K. Dennis. Magnetic and transport properties of single grain R-Mg-Zn icosahedral quasicrystals. *Phys. Rev. B*, 59:308–321, 1999.

[33] R. Fokkink, D. Rust, and V. Salo. Automorphism groups of random substitution subshifts. Pre-print: arXiv:2203.13545.

[34] C. Godrèche and J. Luck. Quasiperiodicity and randomness in tilings of the plane. *J. Stat. Phys.*, 55:1–28, 1989.

[35] P. Gohlke. Inflation word entropy for semi-compatible random substitutions. *Monatsh. Math.*, 192:93–110, 2020.

[36] P. Gohlke. *Aperiodic order and singular spectra.* PhD thesis, Bielefeld University, 2021.

[37] P. Gohlke, A. Mitchell, D. Rust, and T. Samuel. Measure Theoretic Entropy of Random Substitution Subshifts. *Ann. Henri Poincaré*, 24:277–323, 2023.

[38] P. Gohlke, D. Rust, and T. Spindeler. Shifts of finite type and random substitutions. *Discrete Contin. Dyn. Syst.*, 39:5085–5103, 2019.

[39] P. Gohlke and T. Spindeler. Ergodic frequency measures for random substitutions. *Studia Math.*, 255:265–301, 2020.

[40] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman. Fractal measures and their singularities: The characterization of strange sets. *Phys. Rev. A*, 33:1141–1151, 1986.

[41] C. Hoffman. Isomorphism Theory in Ergodic Theory. In R. Meyers, editor, *Mathematics of Complexity and Dynamical Systems*, pages 783–795. Springer, New York, 2012.

[42] T. Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.

[43] G. Keller. *Equilibrium States in Ergodic Theory.* Cambridge University Press, Cambridge, 1998.

[44] A. N. Kolmogorov. A new metric invariant of transitive dynamic systems and automorphisms in lebesgue spaces. *Dokl. Akad. Nauk SSSR*, 119:861–864, 1958.

[45] A. N. Kolmogorov. On the entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk SSSR*, 124:754–755, 1959.

[46] D. Koslicki. *Substitution Markov chains with applications to molecular evolution.* PhD thesis, Pennsylvania State Univ., 2012.

[47] W. Kriger. On the uniqueness of the equilibrium state. *Math. Systems Theory*, 8:97–104, 1974.

[48] K.-S. Lau and S.-M. Ngai. Multifractal Measures and a Weak Separation Condition. *Adv. Math.*, 141:45–96, 1999.

[49] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding.* Cambridge University Press, first edition, 1995.

[50] B. Mandelbrot. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, 62:331–358, 1974.

[51] R. Meester and J. Steif. Higher-dimensional shifts of finite type, factor maps and measures of maximal entropy. *Pacific J. Math.*, 200:497–510, 2001.

[52] P. Michel. Stricte ergodicité d'ensembles minimaux de substitutions. *C. R. Acad. Sc. Paris*, 278:811–813, 1974.

[53] E. D. P. Miro, D. Rust, L. Sadun, and G. S. Tadeo. Topologiocal mixing of random substitutions. *Israel J. Math.*, 2022. To appear.

[54] A. Mitchell. On word complexity and topological entropy of random substitution subshifts. Pre-print: arXiv:2305.04817.

[55] A. Mitchell and A. Rutar. Multifractal analysis of measures arising from random substitutions. Pre-print: arXiv:2301.04958.

[56] M. Moll. Diffraction of random noble means words. *J. Stat. Phys.*, 156:1221–1236, 2014.

[57] M. Morse and G. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60:815–866, 1938.

[58] J. Nilsson. On the entropy of a family of random substitutions. *Monatsh. Math.*, 168:563–577, 2012.

[59] J. Nilsson. On the entropy of a two step random Fibonacci substitution. *Entropy*, 15:3312–3324, 2013.

[60] L. Olsen. A Multifractal Formalism. *Adv. Math.*, 116:82–196, 1995.

[61] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Adv. Math.*, 4:337–352, 1970.

[62] J.-J. Pansiot. Complexité des facteurs des mots infinis engendrés par morphismes itérés. In *Automata, Languages and Programming*, pages 380–389, Berlin, Heidelberg, 1984. Springer Berlin Heidelberg.

[63] W. Parry. Intrinsic Markov chains. *Trans. Amer. Math. Soc.*, 112:55–66, 1964.

[64] Y. B. Pesin. *Dimension Theory in Dynamical Systems: Contemporary Views and Applications.* Chicago Lectures in Mathematics Series. University of Chicago Press, Chicago, 1997.

[65] J. Peyrière. Substitutions aléatoires itérées. *Séminaire de théorie des nombres de Bordeaux*, 80–81:1701–1709, 1981.

[66] M. Pollicott and H. Weiss. Multifractal Analysis of Lyapunov Exponent for Continued Fraction and Manneville-Pomeau Transformations and Applications to Diophantine Approximation. *Comm. Math. Phys.*, 207:145–171, 1999.

[67] M. Queffélec. *Substitution dynamical systems: Spectral analysis.* Springer, 2nd edition, 2010.

[68] R. T. Rockafellar. *Convex Analysis.* Number 28 in Princeton Mathematical Series. Princeton University Press, Princeton, N.J, 1970.

[69] D. Rust. Periodic points in random substitution subshifts. *Monatsh. Math*, 193:683–704, 2020.

[70] D. Rust and T. Spindeler. Dynamical systems arising from random substitutions. *Indag. Math.*, 29:1131–1155, 2018.

[71] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn. Metallic phase with long-range orientational order and no translational symmetry. *Phys. Rev. Lett.*, 53:1951–1953, 1984.

[72] P. Shmerkin. On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions. *Ann. Math.*, 189:319–391, 2019.

[73] Y. G. Sinai. On the notion of entropy of a dynamical system. *Dokl. Akad. Nauk SSSR*, 124:768–771, 1959.

[74] T. Spindeler. *On Spectral Theory of Compatible Random Inflation Systems.* PhD thesis, Bielefeld University, 2017.

[75] P. Walters. *An introduction to ergodic theory.* Graduate Texts in Mathematics. Springer-Verlag, 1982.

[76] D. Wing. *Notions of complexity in substitution dynamical systems.* PhD thesis, Oregon State University, 2011.

[77] M. Wu. A proof of Furstenberg's conjecture on the intersections of $\times p-$ and $\times q-$ invariant sets. *Ann. Math.*, 189:707–751, 2019.